Viral Membrane Proteins: Structure, Function, and Drug Design

PROTEIN REVIEWS

Recent Volumes in this Series

VIRAL MEMBRANE PROTEINS: STRUCTURE, FUNCTION, AND DRUG DESIGN
Edited by Wolfgang B. Fischer

Viral Membrane Proteins: Structure, Function, and Drug Design

Edited by

WOLFGANG B. FISCHER
Department of Biochemistry, University of Oxford, United Kingdom

Kluwer Academic / Plenum Publishers
New York, Boston, Dordrecht, London, Moscow

Library of Congress Cataloging-in-Publication Data

Viral membrane proteins: structure, function, and drug design / edited by
Wolfgang Fischer.
p. cm. – (Protein reviews)
Includes bibliographical references and index.
ISBN 0-306-48495-1
1. Viral proteins. I. Fischer, Wolfgang (Wolfgang B.) II. Series.

QR460.V55 2005
616.9′101–dc22

2004042175

ISBN 0-306-48495-1 Hardback

http://www.kluweronline.com

10 9 8 7 6 5 4 3 2 1

A C.I.P. record for this book is available from the Library of Congress.

Printed in the United States of America

Contents

Part II. Fusion Proteins

3. HIV gp41: A Viral Membrane Fusion Machine

Sergio G. Peisajovich and Yechiel Shai

4. Diversity of Coronavirus Spikes: Relationship to Pathogen Entry and Dissemination

Edward B. Thorp and Thomas M. Gallagher

5. Aspects of the Fusogenic Activity of Influenza Hemagglutinin Peptides by Molecular Dynamics Simulations

L. Vaccaro, K.J. Cross, S.A. Wharton, J.J. Skehel, and F. Fraternali

Part III. Viral Ion Channels/viroporins

6. Viral Proteins that Enhance Membrane Permeability

María Eugenia González and Luis Carrasco

7. FTIR Studies of Viral Ion Channels

Itamar Kass and Isaiah T. Arkin

15. Virus Ion Channels Formed by Vpu of HIV-1, the 6K Protein of Alphaviruses and NB of Influenza B Virus

Peter W. Gage, Gary Ewart, Julian Melton, and Anita Premkumar

16. The Alphavirus 6K Protein

M.A. Sanz, V. Madan, J.L. Nieva, and Luis Carrasco

Part IV. Membrane-Spanning/Membrane Associated

17. The Structure, Function, and Inhibition of Influenza Virus Neuraminidase

Elspeth Garman and Graeme Laver

18. Interaction of HIV-1 Nef with Human CD4 and Lck

Dieter Willbold

List of Contributors

Michael J. Adams, Plant Pathogen Interactions Division, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK.

John F. Antoniw, Plant Pathogen Interactions Division, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK.

Isaiah T. Arkin, The Alexander Silberman Institute of Life Science, Department of Biological Chemistry, The Hebrew University, Givat-Ram, Jerusalem, 91904, Israel.

Burkhard Bechinger, Université Louis Pasteur, Faculté de chimie, ILB, 4 rue Blaise Pascal, 67070 Strasbourg, France.

S. Bour, Bioinformatics and Cyber Technology Center, Office of Technology and Information Systems, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 4 Center Drive, Bethesda, MD, USA.

Luis Carrasco, Centro de Biología Molecular Severo Ochoa, Facultad de Ciencias, Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain.

K. Cross, National Research for Medical Research, Mill Hill, London NW7 1AA, UK.

Gary Ewart, Division of Molecular Bioscience, John Curtin School of Medical Research, Australian National University, Canberra, Australia.

D. Fischer, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK.

W.B. Fischer, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK, and Bionanotechnology Interdisciplinary Research Consortium, Clarendon Laboratory, Department of Physics, Oxford University, Parks Road, Oxford OX1 3SU, UK.

F. Fraternali, National Research for Medical Research, Mill Hill, London NW7 1AA, UK.

Peter W. Gage, Division of Molecular Bioscience, John Curtin School of Medical Research, Australian National University, Canberra, Australia.

Thomas M. Gallagher, Department of Microbiology and Immunology, Loyola University Medical Center, Maywood, IL, USA.

Elspeth Garman, Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.

María Eugenia González, Unidad de Expresión Viral, Centro Nacional de Microbiologia, Instituto de Salud Carlos III, Carretera de Majadahonda-Pozuelo, Km 2, Majadahonda 28220, Madrid, Spain.

Peter Henklein, Humboldt Universität, Institut für Biochemie, Hessische Str 3-4, 10115 Berlin, Germany.

D. Jones, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.

Itamar Kass, The Alexander Silberman Institute of Life Science, Department of Biological Chemistry, The Hebrew University, Givat-Ram, Jerusalem 91904, Israel.

C.G. Kim, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK.

Jared Knopman, Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA.

Y.H. Lam, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK.

Robert A. Lamb, Department of Biochemistry, Molecular Biology and Cell Biology, and Howard Hughes Medical Institute, Northwestern University, Evanston, IL, USA.
Graeme Laver, Barton Highway, Murrumbateman, NSW 2582, Australia.
S. Lee, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
V. Lemaitre, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK, and Nestec S.A., BioAnalytical Science Department, Vers-Chez-Les-Blanc, CH-1000 Lausanne 26, Switzerland.
Tse-I Lin, Tibotec BVDV, Gen. De. Wittelaan 11B-3, B-2800 Mechelen, Belgium.
V. Madan, Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Facultad de Ciencias, Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain.
F.M. Marassi, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA, USA.
Julian Melton, Division of Molecular Bioscience, John Curtin School of Medical Research, Australian National University, Canberra, Australia.
M. Mesleh, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
M. Montal, Section of Neurobiology Division of Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
Anna Moroni, Dipartimento di Biologia and CNR—IBF, Unità di Milano; Istituto Nazionale di Fisica della Materia, Unità di Milano-Università, Milano, Italy.
A. Mrse, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
A. Nevzorov, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
J.L. Nieva, Unidad de Biofísica (CSIC-UPV/EHU), Departamento de Bioquímica, Universidad del País Vasco, Aptdo. 644, 48080 Bilbao, Spain.
M. Oblatt-Montal, Section of Neurobiology Division of Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
S.J. Opella, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
S. Park, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA.
Sergio G. Peisajovich, Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel.
Lawrence H. Pinto, Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA.
Anita Premkumar, Division of Molecular Bioscience, John Curtin School of Medical Research, Australian National University, Canberra, Australia.
M.A. Sanz, Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Facultad de Ciencias, Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain.
Cornelia Schroeder, Abteilung Virologie, Institut für Mikrobiologie und Hygiene, Universitätskliniken Homburg/Saar, Germany.
Ulrich Schubert, Institute for Clinical and Molecular Virology, University of Erlangen-Nürnberg, Schlossgarten 4, D-91054 Erlangen, Germany.
Yechiel Shai, Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel.
J.J. Skehel, National Research for Medical Research, Mill Hill, London NW7 1AA, UK.
K. Strebel, Bioinformatics Core Facility, Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 4 Center Drive, Bethesda, MD, USA.
Yajun Tang, Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA.

Gerhard Thiel, Institute for Botany, University of Technology, Darmstadt, Germany.
Edward B. Thorp, Department of Microbiology and Immunology, Loyola University Medical Center, Maywood, IL, USA.
L. Vaccaro, National Research for Medical Research, Mill Hill, London NW7 1AA, UK.
James Van Etten, Department of Plant Pathology, Nebraska Center for Virology, University of Nebraska, Lincoln, NE, USA.
Padmavati Venkataraman, Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA.
Gregory A. Voth, Department of Chemistry and Henry Eyring Center for Theoretical Chemistry, University of Utah, 315 S. 1400 E. Rm 2020, Salt Lake City, UT, USA.
A. Watts, Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK.
S.A. Wharton, National Research for Medical Research, Mill Hill, London NW7 1AA, UK.
Dieter Willbold, Institut für Physikalische Biologie, Heinrich-Heine-Universität, 40225 Düsseldorf, Germany und Forschungszentrum Jülich, IBI-2, 52425 Jülich, Germany.
Victor Wray, Department of Structural Biology, German Research Centre for Biotechnology, Mascheroder Weg 1, D-38124 Braunschweig, Germany.
Yujie Wu, Department of Chemistry and Henry Eyring Center for Theoretical Chemistry, University of Utah, 315 S. 1400 E. Rm 2020, Salt Lake City, UT, USA.

Preface

Viruses enter cells and modulate the biosynthetic machinery of the host for the synthesis of their own building blocks. These building blocks assemble in an organized fashion and large numbers of viral copies finally leave the host, ready to enter the next host cell. This life style would be of little concern if it were not that many viruses have a fatal side effect to the host, leading to the death of the host cell and sometimes to the death of the whole organism. Under these circumstances, the first two sentences could be rewritten and the words "enter" and "leave" replaced by more drastic words such as "attack," "invade," or "kill." However, not all the attacks lead to an immediate cell death or to a phenotypic manifestation in the organism.

Viruses are under constant survival pressure and have evolved mechanisms to resist environmental pressure by having for example a high mutation rate. This may lead finally also to an increased spreading to novel hosts, which can have a devastating effect on the invaded organism, including humans, especially if the species barrier is crossed in an unpredicted way. In the modern world, the large density of populations and travel habits can lead to a rapid spread of the virus, with a possible major impact on our social behavior and the economy. The recent appearance of the SARS virus or avian influenza viruses in humans represents such an immediate threat. Once within a host, some viruses, such as HIV-1, replicate but rather than produce faithful copies of the parent virus, constantly mutate making it almost impossible to produce a vaccine and limiting the success of drug therapy. We are also directly or indirectly affected by animal or plant viruses. The last foot and mouth outbreak within the United Kingdom and other European countries resulted in the slaughter of large numbers of farm animals to prevent the rapid spread of the disease. As an indirect effect, the country side, dependent to a very large extent on tourism, had to be closed down to avoid any further spread of the disease. Plant viruses threatening our annual harvest and can through price rises add to inflation. However, not all viruses cause harm to us, some of the plant viruses may even cheer us up such as the tulip mosaic virus that causes the striping pattern of tulip petals.

In this book, we aim to summarize the current knowledge on a special class of viral molecules, the membrane proteins, from the full range of viruses, including plant viruses. Research on these membrane proteins has been limited by a number of technical difficulties, and rate of progress compared with globular proteins has been slow. Membrane proteins are involved at the stage of viral entry into the host cell, in modulating subcellular electrochemical gradients and/or shuffling proteins across cell membranes.

The first section is dedicated to viral membrane proteins from plant viruses with the most recent computational research on the viral genome revealing the first experimental evidence of a K^+ channel encoded by a plant virus. The second section in the book is dedicated to the proteins involved in the early event of the life cycle of the viruses in the host cell, the fusion proteins. The third section summarizes, in several chapters, the current state of the research on ion channels and viroporins, which are known to modulate the electrochemical balance in the virus itself and subcellular compartments in the host cell. The fourth section

describes membrane-bound and membrane-associated viral proteins. All chapters include functional and structural data and address, where possible, the development of antiviral drugs. A large number of techniques are described by the authors, revealing the way in which a wide range of approaches are required to shed light on the molecular life of viruses.

I wish to thank the editorial team at Kluwer Academic Publishers for their enthusiasm and physical support during the generation of this synthesis of our recent advances in viral membrane protein research. Thanks go also to all the authors for their willingness and patience while working on the book. My acknowledgment includes also Judy Armitage (Oxford) and my colleagues in the lab for stimulating and helpful discussions.

Part I

Membrane Proteins from Plant Viruses

1

Membrane Proteins in Plant Viruses

Michael J. Adams and John F. Antoniw

1. Introduction

The plant cell wall is a substantial barrier preventing direct entry of viruses and therefore, unlike many animal viruses, plant viruses cannot initiate infection by any independent ability to cross membranes. A few plant viruses enter through microscopic wounds but most are introduced into their hosts by a vector, most frequently a leaf-feeding insect. Small pores between adjacent cells (plasmodesmata) provide cytoplasmic continuity and thus a channel for transport of nutrients and some larger molecules, and viruses exploit this route for cell-to-cell transport. While some viruses remain restricted to a small area around the initial site of infection and may be limited to certain cell types, many exhibit long distance movement via the plant vascular system.

Among the plant viruses, cell-to-cell movement depends on one or more virus-encoded movement proteins (MPs) and many of these are integral membrane proteins that interact with the endoplasmic reticulum (ER). Many RNA viruses multiply within the cell cytoplasm and there is recent evidence that replication proteins of such viruses are also targeted to membranes. In addition, membrane proteins may play a role in plant virus transmission for those viruses that enter the cells of their vectors. In this chapter, we survey the occurrence of membrane proteins among all plant viruses and review the literature on their biological role. We also present and discuss the limited structural information on plant virus integral membrane proteins.

2. Survey of Transmembrane Proteins in Plant Viruses

In preparation for this chapter, we have used a plant virus sequence database that we developed to make a comprehensive survey of all published complete gene sequences of all plant viruses.

2.1. The Database

The database used was developed from files originally prepared for the electronic version of the Association of Applied Biologists (AAB) Descriptions of Plant Viruses

Michael J. Adams and John F. Antoniw • Plant Pathogen Interactions Division, Rothamsted Research, Harpenden, United Kingdom.

Viral Membrane Proteins: Structure, Function, and Drug Design, edited by Wolfgang Fischer.
Kluwer Academic / Plenum Publishers, New York, 2005.

(Adams *et al*., 1998). As part of that project, we provided software (DPVMap) to display selected virus sequences interactively. A separate enhanced feature table (EFT) file was written for each sequence containing the start and end nucleotide positions of the features (e.g., open-reading frames (ORFs), untranslated regions) within the sequence. In DPVMap any of the features of the sequence could be dragged into a sequence editor to display either its nucleotide sequence (as RNA or DNA), or the predicted amino acid sequence of an ORF. Annotations provided for the correct display of reverse complementary sequences and of those incorporating a frameshift or intron. Sequence features were checked for accuracy and, as far as possible, nomenclature for genes and proteins were standardized within genera and families to make it easier to compare features from different viruses. From a modest beginning, the number of sequences provided has been increased and now includes all complete sequences of plant viruses, viroids and satellites, and all sequences that contain at least one complete gene. The information contained in the individual EFT files is valuable because it has been checked for accuracy and is often more detailed than that provided in the original sequence file from EMBL or Genbank. However, the EFT files can only be used with DPVMap and to examine one sequence at a time. We therefore decided to transfer this information together with the sequences themselves into a database table, so that multiple data sets could be selected and extracted easily and then used for further analysis.

The database was prepared in MySQL on a Linux PC and includes up to date taxonomic information and a table of sequence data containing all the information from the individual EFT files. The version used here was based on sequences available from the public databases at the end of November 2002 and includes a total of 4,687 accessions. It therefore records the start and end positions of all important features and genes in every one of the significant plant virus sequences. The database has been placed on a public Internet site (DPVweb at http://www.dpvweb.net) where it may be accessed using client software.

2.2. Software

A web-enabled Windows client application was written in Delphi for IBM-compatible PCs to scan the database tables, translate each complete ORF into its amino acid sequence, and then to predict transmembrane (TM) regions using TMPRED (Hofmann and Stoffel, 1993). A summary of the results was exported to a Microsoft®Excel spreadsheet and examined for consistency within species and genera. The results have been used to inform the presentation and discussion of the different types and function of plant virus membrane proteins (below) and some ambiguous results were checked using the web-based software HMMTOP (Tusnády and Simon, 1998), TMHMM (Sonnhammer *et al*., 1998), and TopPred 2 (von Heijne, 1992).

3. Cell-to-Cell Movement Proteins

Most plant viruses encode one or more specific MPs that are required for the virus to spread between adjacent host cells. Functions assigned to these proteins include nucleic-acid binding (some viruses move as nucleic acid–MP complexes), modification of the size exclusion limit of the plasmodesmata (the connections between adjacent cells), and targeting to the inter- and intracellular membrane system, the ER. A number of groups of MPs have been identified and at least some of these are integral membrane proteins.

3.1. The "30K" Superfamily

A very large number of plant viruses have MPs that share common structural features, which led Mushegian and Koonin (1993) to propose the name "30K" superfamily for them. This grouping has most recently been reviewed and defended by Melcher (2000). It includes a surprisingly diverse range of viruses including those with DNA genomes (the pararetroviruses and the ssDNA viruses in the genus *Begomovirus*) and many different groups of both positive sense (*Bromoviridae, Comoviridae, Capillovirus, Dianthovirus, Furovirus, Idaeovirus, Tobamovirus, Tobravirus, Tombusvirus, Trichovirus, Umbravirus*) and negative sense (*Nucleorhabdovirus, Tospovirus*) RNA viruses. These have been assigned by computer predictions showing the presence of a core domain consisting of two α-helices separated by a series of β-elements.

The best-studied virus from this group is *Tobacco mosaic virus* (TMV, genus *Tobamovirus*). Its MP has been shown to increase the size exclusion limit of plasmodesmata, and specifically at the leading edge of expanding lesions (Oparka *et al*., 1997). It has non-specific RNA-binding activity, forming a viral RNA–MP complex that moves between cells (Citovsky and Zambryski, 1991). It can also bind to the cytoskeleton (Heinlein *et al*., 1998; Reichel and Beachy, 1998; Reichel *et al*., 1999; Boyko *et al*., 2000) but it remains uncertain whether this property is essential for cell-to-cell movement as recent evidence suggests that TMV can replicate and move in the absence of microtubules (Gillespie *et al*., 2002). There remains much to be discovered about the interaction of the MP with host cell components and how this facilitates cell-to-cell movement of the viral RNA, but a combination of CD spectroscopy, trypsin treatment, and mass spectroscopy has helped to develop a topological model (Brill *et al*., 2000). This confirms the role of the two core α-helices as TM domains resistant to trypsin, and indicates that the N- and C-termini would be exposed in the cytoplasm and a short loop in the ER lumen (Figure 1.1).

There is less experimental information for the other MPs in this group but they are likely to have a similar association with membranes. For example, the movement protein (3a) of *Alfalfa mosaic virus* (genus *Alfamovirus*, family *Bromoviridae*), used as a MP–GFP (green fluorescent protein) construct, co-localized with ER in tobacco protoplasts and onion cells and moved between adjacent onion cells. Fractionation and biochemical studies in insect cells demonstrated that the MP–GFP was an integral membrane protein (Mei and Lee, 1999) although no ER targeting signal has been identified. Some other "30K" superfamily MPs that have been shown to interact with membranes are the ORF3 products of *Grapevine virus A* and *Grapevine virus B* (genus *Vitivirus*) (Saldarelli *et al*., 2000), the P22 of *Tomato bushy stunt virus* (genus *Tombusvirus*, family *Tombusviridae*) (Desvoyes *et al*., 2002), and the BC1 protein of *Abutilon mosaic virus* (genus *Begomovirus*, family *Geminiviridae*) (Zhang *et al*., 2001, 2002; Aberle *et al*., 2002).

Some of the superfamily member MPs act in a rather different fashion by producing tubules that extend through the plasmodesmata. This has been best studied in *Cowpea mosaic virus* (CPMV, genus *Comovirus*, family *Comoviridae*) (Van Lent *et al*., 1991). In these examples, the virus has been shown to move as intact virions and therefore to require the coat protein (CP), but it appears that some "30K" superfamily MPs have both tubule-forming and RNA-binding activities (Perbal *et al*., 1993; Jansen *et al*., 1998; Canto and Palukaitis, 1999; Nurkiyanova *et al*., 2001). Unlike TMV, the CPMV MP does not localize to either the microtubules or the ER and the mechanism of its delivery to the cell periphery is not known. The tubules themselves are thought to arise from the host protein plasma membrane (Pouwels *et al*., 2002).

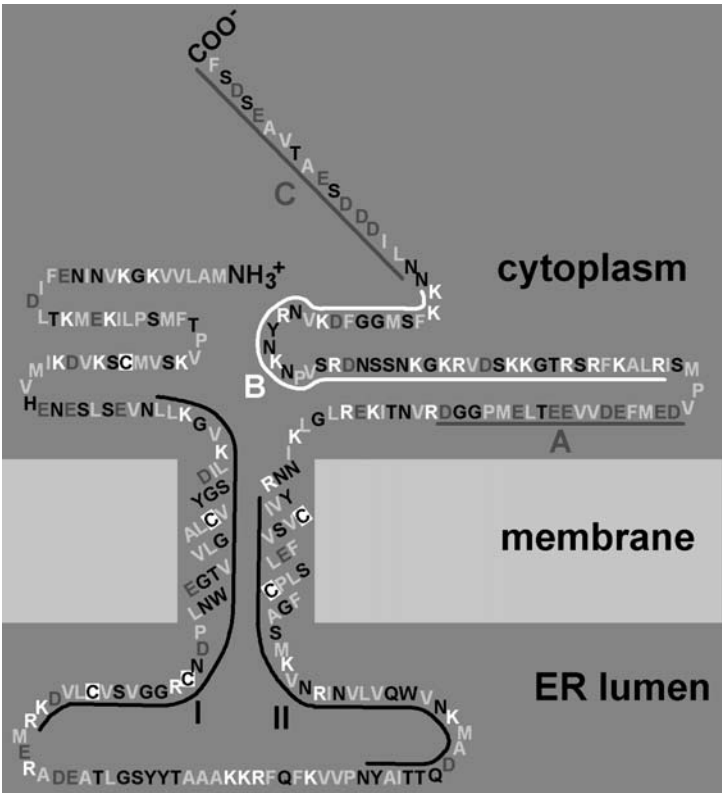

Figure 1.1. Topological model of the *Tobacco mosaic virus* movement protein, re-drawn from Brill *et al*. (2000), and used by kind permission of Prof. R.N. Beachy. Hydrophobic amino acid residues are shown in pale grey. Basic residues (in white) are concentrated in domain B and acidic residues, shown in dark grey are concentrated in domains A and C; Cys residues are shown with white background. Domains I and II are conserved amongst tobamoviruses.

TMPRED correctly identified the position and orientation of the two TM domains of TMV and 7 other tobamoviruses (out of 16 different species sequenced). Among the "30K" superfamily generally, a TM domain was identified in most viruses, but two domains were predicted in only 20 out of more than 60 species.

3.2. Triple Gene Block

Some positive-sense ssRNA filamentous and rod-shaped viruses do not have the single MP exemplified by TMV or CPMV but a group of three, partially overlapping, proteins known as the "triple gene block" (TGB). The structure and function of the TGB has been recently reviewed by Morozov and Solovyev (2003). All three TGB proteins are required for movement and the two smaller proteins, TGBp2 and TGBp3, are TM proteins. These were

not only strongly predicted by computer analyses but also by *in vitro* studies (e.g., Morozov *et al*., 1991), localization to membrane fractions of infected plant tissues (e.g., Gorshkova *et al*., 2003), and by microscopical studies of proteins fused to GFP showing them to be localized to the ER or membrane bodies as well as to plasmodesmata (e.g., Solovyev *et al*., 2000; Cowan *et al*., 2002; Zamyatnin *et al*., 2002; Gorshkova *et al*., 2003). No detailed structural studies have been reported, but all TGBp2 molecules (11–14 kDa) contain two TM segments and it is predicted that the N- and C-termini are in the cytoplasm. TGBp3 molecules are of three different types. Those of the filamentous viruses (genera *Allexivirus, Carlavirus, Foveavirus, Potexvirus*) are 6–13 kDa in size and have a single TM segment, while those of the rod-shaped viruses are larger (15–24 kDa) and have two segments, and those of the genus *Benyvirus*, having a different arrangement to those in the genera *Hordeivirus, Pecluvirus*, and *Pomovirus*. The C-termini of TGBp3 molecules are predicted to be in the cytoplasm.

Transiently expressed TGBp2–GFP fusions localize to the ER, while TGBp3 fusions are found in membrane bodies near the plant cell periphery but in the presence of TGBp3, TGBp2 is re-targeted to the peripheral bodies (Solovyev *et al*., 2000; Tamai and Meshi, 2001; Cowan *et al*., 2002; Zamyatnin *et al*., 2002; Gorshkova *et al*., 2003). TGBp2 and TGBp3 together appear to be responsible for targeting rod-shaped virus TGBp1 to plasmodesmata (Erhardt *et al*., 1999, 2000; Lawrence and Jackson, 2001), but the (smaller) TGBp1 of the filamentous viruses can move independently (e.g., Morozov *et al*., 1999). TGBp2 is also involved in increasing the size exclusion limit of plasmodesmata and it has been suggested that this may occur via the regulation of callose deposition or degradation. Recent evidence that TGBp2 interacts with TIP, a host protein regulator of β-1,3-glucanase (a key enzyme of callose turnover), strengthens this hypothesis (Fridborg *et al*., 2003).

In TMPRED, the expected TM domains were consistently and strongly detected in all TGBp2 and TGBp3 sequences. TGBp2 proteins were 104–154 amino acids (aa) long, with a loop between the two predicted TM domains of 39–61 aa. The two classes of TGBp3 proteins were correctly identified; in the rod-shaped viruses with two TM domains, the second domain is consistently at the C-terminus (within 2–5 aa).

3.3. *Carmovirus*-Like

Members of the genus *Carmovirus* are among the smallest RNA viruses (genome ~4 kb). They do not have MPs of the "30K" superfamily, nor a TGB, but two small, overlapping, internal ORFs are involved in cell-to-cell movement (Hacker *et al*., 1992). The first, and slightly smaller, of these proteins is a soluble protein with RNA-binding capacity, while the second contains two potential TM domains. In experiments using the type member, *Carnation mottle virus* (CarMV), the two putative TM domains of the p9 protein were inserted into the *Escherichia coli* inner membrane protein Lep and then tested for insertion into dog pancreas microsomes. The experiments demonstrated TM activity and that the N- and C-termini of the protein were located in the cytoplasm. It was proposed that the charged C-terminus of p9 would interact with the C-terminal domain of the smaller p7 protein that had already bound to viral genomic RNA (Vilar *et al*., 2002). Results of a spatio-temporal analysis are consistent with this hypothesis (Garcia-Castillo *et al*., 2003). Our analysis confirmed the consistent presence of two TM domains in most members of the genus, but only indicated one such domain in *Melon necrotic spot virus* and in members of the related genus *Necrovirus*, where the protein seems to be smaller than in CarMV.

3.4. Other Movement Proteins

In *Maize streak virus* (genus *Mastrevirus*, family *Geminiviridae*), the MP is encoded by ORF V2, the smaller of the two ORFs translated in the positive sense and a central α-helical domain has been predicted to have TM properties (Boulton *et al*., 1993). This is supported by studies showing its localization to plasmodesmata (Dickinson *et al*., 1996) and by the occurrence of similar domains in other members of the genus (confirmed by our analyses), but it has not yet been proved experimentally that the hydrophobic domain is required for membrane association (Boulton, 2002).

In *Banana bunchy top virus* (genus *Babuvirus*, family *Nanoviridae*), GFP-tagging showed that the protein encoded by DNA-4, which possesses a hydrophobic N-terminus, was found to localize exclusively to the cell periphery. Deletion of the N-terminal region abolished its ability to localize to the cell periphery (Wanitchakorn *et al*., 2000). Our analyses show similar domains in other viruses of this family.

Within the genus *Tymovirus*, the first ORF, which almost completely overlaps with the large replication protein, has been identified as a MP (Bozarth *et al*., 1992). This protein is much larger than those discussed above (69–85 kDa) and is proline-rich. Its localization within cells has not been reported. There were no strongly hydrophobic regions in the sequences of this gene for any of the members of the genus and the few possible TM regions identified in our analysis were not strongly supported and did not appear at a consistent position within an alignment of the MPs of the genus members.

In *Beet yellows virus* (genus *Closterovirus*, family *Closteroviridae*), the 70K HSP70h (heat shock protein 70 homolog) has been shown to be absolutely required for cell-to-cell movement (Peremyslov *et al*., 1999) and can be localized in plasmodesmatal channels (Medina *et al*., 1999). The protein acts as a molecular chaperone and is incorporated into the tail of the functional virion (Alzhanova *et al*., 2001). This activity appears to be related to its ATPase activity and it is not clear whether any membrane-targeting activity is involved, although our studies show several potential TM domains within the protein, one of which appears to be fairly consistent among all members of the family.

3.5. General Comments

At least for the better studied viral MPs (TMV, TGB proteins, *Carmovirus*), it seems probable that they enter the ER co-translationally and that the hydrophobic regions then migrate into the ER membrane. Movement to the cell periphery probably occurs as complexes with virions (or other nucleic acid–protein associations) in membrane-bound bodies and may use the cytoskeleton-based pathway. The complexes are thus delivered to the neck of the plasmodesmata. None of the plant host proteins that interact with viral MPs have yet been unequivocally identified but it is interesting that the NS_M movement protein of *Tomato spotted wilt virus* (genus *Tospovirus*, family *Bunyaviridae*), which has been classified in the "30K" superfamily, has been shown to interact with the viral CP, to bind viral RNA and, in a yeast two-hybrid screen, to bind to two plant proteins of the DnaJ family, that are in turn known to bind plant HSP70s (Soellick *et al*., 2000). There are at least hints here of common links between what appear to be very dissimilar viral MPs. It is also interesting that there is increasing evidence that some plant proteins ("non-cell-autonomously replicating proteins," NCAPs) have properties similar to viral MPs in their effects upon plasmodesmatal size exclusion limits and in transporting RNA (see, for example, the detailed review by Roberts and

Oparka, 2003). It therefore appears likely that plant virus MPs mimic various aspects of the plant's own machinery for trafficking of large molecules.

4. Replication Proteins

Positive-strand RNA viruses assemble their RNA replication complexes on intracellular membranes and some progress has been made in identifying the proteins and sequences responsible.

In the genus *Tombusvirus* (family *Tombusviridae*), ORF1 encodes a polymerase with a readthrough (RT) domain and the smaller product contains an N-terminal hydrophilic portion followed by two predicted hydrophobic TM segments. In the type member, *Tomato bushy stunt virus*, the protein is localized to membrane fractions of cell extracts (Scholthof *et al*., 1995). Infection of *Nicotiana benthamiana* cells with *Cymbidium ringspot virus* (CymRSV) or *Carnation Italian ringspot virus* (CIRV) results in the formation of conspicuous membranous bodies, which develop from modified peroxisomes or mitochondria, respectively. The ORF1 proteins can be localized in these membranous bodies (Bleve-Zacheo *et al*., 1997) and have been shown to be integral membrane proteins with their N- and C-termini in the cytoplasm (Rubino and Russo, 1998; Rubino *et al*., 2000, 2001; Weber-Lotfi *et al*., 2002). These domains were consistently identified in all sequenced members of the genus by our TMPRED analysis; in the other genera of the *Tombusviridae*, although TM domains were identified they did not appear to be in corresponding positions, or at similar spacing, within the protein.

Members of the family *Bromoviridae* have three RNAs and the major products of both RNA1 and RNA2 (1a and 2a proteins) are required for replication. In both *Brome mosaic virus* (BMV, genus *Bromovirus*) and *Alfalfa mosaic virus* (genus *Alfamovirus*) proteins 1a and 2a co-localize to membranes, but respectively to the ER and tonoplast (Restrepo-Hartwig and Ahlquist, 1999; Heijden *et al*., 2001). In BMV, the 1a protein is primarily responsible for this localization and a region, C-terminal to the core methyltransferase motif, has been identified by membrane floatation gradient analysis as sufficient for high-affinity ER membrane association although other regions are probably also involved (den Boon *et al*., 2001). The 1a protein is fully susceptible to proteolytic digestion in the absence of detergent, suggesting that it does not span the membrane, but has an association with membranes that is stronger (resistant to high salt and high pH conditions) than is usual for a peripheral membrane protein. The 2a protein is then recruited to the membrane through its interaction with 1a and the N-terminal 120 amino-acid segment of 2a is sufficient for this (Chen and Ahlquist, 2000). Neither experimental evidence, nor computer predictions, suggest that a TM domain is involved with this interaction, although TMPRED does identify some (rather weak) regions in most *Bromoviridae* 1a proteins.

Members of the family *Comoviridae* have two RNAs, each of which encodes a polyprotein. Products of RNA1 are involved in replication, which has been associated with ER membranes in CPMV (Carette *et al*., 2000, 2002) and in *Grapevine fanleaf virus* (genus *Nepovirus*) (Ritzenthaler *et al*., 2002). In particular, the nucleoside triphosphate binding protein is believed to act as a membrane anchor for the replication complex and in *Tomato ringspot virus* (genus *Nepovirus*) a region at its C-terminus has been shown to have TM properties (Han and Sanfaçon, 2003). This is strongly confirmed by our TMPRED analyses for viruses in all genera of the family (*Comovirus, Fabavirus*, and *Nepovirus*).

Some progress in identifying the plant proteins with which the viral replication proteins interact has been made in the genus *Tobamovirus*. Western blot studies of membrane-bound *Tomato mosaic virus* (ToMV) replication complexes indicated the presence of a plant protein related to the 54.6-kDa GCD10 protein, the RNA-binding subunit of yeast eIF-3 (Osman and Buck, 1997). More recently, studies of *Arabidopsis* mutants have revealed several genes that are necessary for efficient multiplication of tobamoviruses. In particular TOM1 has been identified as a 7-pass TM protein of 291aa that interacts with the helicase domain of tobamovirus replication proteins and TOM2A, a 4-pass TM protein of 280 aa that interacts with TOM1. GFP-tagging had demonstrated that these proteins co-localize with the replication proteins to vacuolar (tonoplast) membranes in plant cells (Yamanaka *et al*., 2000, 2002; Hagiwara *et al*., 2003; Tsujimoto *et al*., 2003).

There is less detailed evidence for the involvement of membrane targeting in the replication of other plant viruses but the replication proteins of *Peanut clump virus* (genus *Pecluvirus*) have been localized to membranes (Dunoyer *et al*., 2002). In the genus *Potyvirus*, there have been suggestions that the 6K2 product of the polyprotein of *Tobacco etch virus* is involved with replication and that it binds to membranes (Restrepo-Hartwig and Carrington, 1994) and this is supported by recent results showing that the CI-6K2 protein of *Potato virus A* was associated with membrane fractions but that fully processed CI was not (Merits *et al*., 2002). Our analyses show that there is a strongly predicted TM domain in all 6K2 proteins in the family *Potyviridae*.

5. Proteins Involved in Transmission by Vectors

To initiate infection of a host plant, viruses have to be introduced into a cell across the substantial barrier posed by the cell wall. Many plant viruses are dependent upon vectors for this step. Some virus–vector interactions involve adsorption onto, and release from, an external surface and this is typified by the nonpersistent, stylet-borne transmission by aphids of many viruses, for example in the genus *Potyvirus*. In other viruses, there is a more intimate and lasting ("persistent") relationship with the vector, in which the virus enters the host cells of its vector ("circulative") and, in some cases may replicate within it ("propagative") as well as within the plant host. Viral membrane proteins may therefore play an important role in the transmission of some viruses.

5.1. Insect Transmission

5.1.1. Persistent Transmission by Aphids

Persistent (circulative but not propagative) transmission has been best studied in members of the family *Luteoviridae*. Electron microscopy indicates that virus particles cross the gut into the aphid haemocoel in coated vesicles by receptor-mediated endocytosis (Gildow, 1993; Garret *et al*., 1996). While the aphid gut acts as a barrier against the uptake of some morphologically similar viruses, uptake of different luteoviruses is not always related to the efficiency of virus transmission and it therefore appears that endocytosis is only partially selective. It is likely that the CP is primarily involved in interactions with the receptor but evidence for the role of the CP–RT is not entirely consistent. Mutants of *Barley yellow dwarf virus-PAV* lacking the RT were taken up through the aphid gut (although not subsequently transmitted) (Chay *et al*., 1996) but some mutations in the *Beet western yellows virus* RT

domain apparently affect acquisition across the gut membrane (Brault *et al*., 2000). Changes in the CP and/or the RT of *Potato leafroll virus* (PLRV) have also been shown to hinder passage across the gut membrane (Rouze-Jouan *et al*., 2001). Virions taken up into the haemocoel appear to be bound to a protein (symbionin) produced by endosymbiotic bacteria of the genus *Buchnera*. This appears to be important for virus survival within the vector (see review of Reavy and Mayo, 2002). If a virus is to be transmitted, it must then cross a membrane into the accessory salivary gland and this, also, is a specific, receptor-mediated process. The aphid and virus determinants of this process have not been characterized in detail but virus-like particles of PLRV consisting of CP (without the RT) and no genomic RNA could be exported into the salivary duct canal suggesting that the virus determinants are located within the CP alone (Gildow *et al*., 2000). Our analyses do not suggest that there are TM domains in the CP or RT and it is likely, therefore, that their association with membranes is peripheral.

5.1.2. Transmission by Hoppers

Viruses transmitted by leafhoppers, planthoppers, and treehoppers include members of the genera *Mastrevirus, Curtovirus*, and *Topocuvirus* (family *Geminiviridae*) which have circulative, but not propagative, transmission. There is little experimental work to determine how these enter their vector, but chimerical clones based on the whitefly-transmitted *African cassava mosaic virus* (genus *Begomovirus*, family *Geminiviridae*) with the CP of the leafhopper transmitted *Beet curly top virus* (genus *Curtovirus*) could be transmitted by the leafhopper, demonstrating that the CP was the major determinant of vector specificity (Briddon *et al*., 1990). A single TM domain is predicted in the CP of all these viruses by TMPRED (but not in the whitefly-transmitted geminiviruses) but it is not known whether this is related to any role in vector transmission.

Hopper-transmitted viruses that are propagative include members of the genera *Marafivirus* (family *Tymoviridae*) and *Tenuivirus*, some plant rhabdoviruses and all plant-infecting members of the family *Reoviridae*. In *Rice dwarf virus* (genus *Phytoreovirus*), a nontransmissible isolate that could not infect cells of the vector was shown to lack the P2 outer capsid protein, one of the six structural proteins of the virus (Tomaru *et al*., 1997). It was subsequently shown that this protein was required for adsorption to cells of the insect vector (Omura *et al*., 1998). In another reovirus, *Rice ragged stunt virus* (genus *Oryzavirus*), the spike protein encoded by S9 was expressed in bacteria, fed to the vector, and shown to inhibit transmission. Its ability to bind a 32-kDa insect membrane protein indicated that this might be a virus receptor that interacts with the spike protein (Zhou *et al*., 1999). Within the genus *Tenuivirus*, the larger RNA2 product pC2, encoded in a negative sense, has several typical features of viral membrane glycoproteins (Takahashi *et al*., 1993; Miranda *et al*., 1996) and these are strongly detected by TMPRED, but its structure and function have not been studied in detail.

5.1.3. Transmission by Thrips

Viruses in the genus *Tospovirus* (family *Bunyaviridae*) are transmitted by thrips in a propagative manner, and the best studied is the type member, *Tomato spotted wilt virus* (TSWV). Virus enters its vector after ingestion of infected plant material and involves endocytosis by fusion at the apical plasmalemma of midgut epithelial cells. It is believed that one or both of the membrane glycoproteins (GP1 and GP2) serve as virus attachment proteins,

binding to vector receptor proteins. The evidence for this, largely derived from electron microscopy has recently been summarized by Ullman *et al*. (2002). The use of anti-idiotypic antibodies has indicated that GP1 and GP2 bind thrips proteins of about 50 kDa (Bandla *et al*., 1998; Meideros *et al*., 2000) but the receptors have not been characterized in detail. Experiments in mammalian cells show that transporting and targeting of TSWV glycoproteins is probably very similar to that in animal-infecting bunyaviruses (e.g., *Uukeniemi virus* and *Bunyamwera virus*). The glycoprotein precursor was efficiently cleaved and the resulting GP1 and GP2 glycoproteins were transported from the ER to the Golgi complex, where they were retained. GP2 alone was retained in the Golgi complex, while GP1 alone was retained in the ER, irrespective of whether it contained the precursor's signal sequence or its own N-terminal hydrophobic sequence (Kikkert *et al*., 2001). TMPRED predicts 5–10 TM segments in the precursor glycoprotein of different tospoviruses.

5.1.4. Persistent Transmission by Whiteflies

Viruses in the genus *Begomovirus* (family *Geminiviridae*) are transmitted by whiteflies in a circulative, but not propagative, manner. The route of transmission is similar to that described above for aphids (Section 5.1.1.) and it is therefore likely that receptor-mediated endocytosis is involved, both in crossing the gut into the haemocoel and then in viral transmission through the salivary glands. Several experiments indicate that the specificity for this resides in the CP. For example, *Abutilon mosaic virus* has lost its ability to be transmitted by whiteflies (probably because it has been maintained in plants by cuttings) and does not move into the haemocoel (Morin *et al*., 2000). However, this ability can be restored by substitution of the CP by that of *Sida golden mosaic virus* (Hofer *et al*., 1997) or by mutation at 2 or 3 positions (aa 124, 149, 174) in the CP (Hohnle *et al*., 2001). Conversely, replacement of two amino acids (129 Q to P, 134 Q to H) in the CP of *Tomato yellow leaf curl Sardinia virus* was sufficient to abolish transmission (Norris *et al*., 1998). There is not yet any detailed information on the interaction between the CP and putative whitefly receptors but our TMPRED results show that this is unlikely to involve a TM protein.

5.2. Fungus Transmission

A range of single-stranded RNA viruses are transmitted by plasmodiophorid "fungi," obligate intracellular parasites that are confined to plant roots. Although traditionally regarded as fungi by plant pathologists, these organisms are of uncertain taxonomic affinity but appear to be more closely related to protists than to the true fungi. In some of these, the viruses are carried within the vector and both acquisition and transmission involves transport across the membrane that separates the cytoplasm of the vector from that of its host (Adams, 2002; Kanyuka *et al*., 2003). For rod-shaped viruses of the genera *Benyvirus, Furovirus*, and *Pomovirus*, deletions in the CP–RT domain abolish transmissibility (Tamada and Kusume, 1991; Schmitt *et al*., 1992; Reavy *et al*., 1998), while for filamentous viruses of the genus *Bymovirus* (family *Potyviridae*), deletions in the P2 domain have a similar effect (Adams *et al*., 1988; Jacobi *et al*., 1995; Peerenboom *et al*., 1996). In *Beet necrotic yellow vein virus* (BNYVV, genus *Benyvirus*), substitution of two amino acids (KTER to ATAR at 553–556) in the CP–RT prevented transmission by the vector, *Polymyxa betae* (Tamada *et al*., 1996). Computer predictions by TMPRED and other software suggest that all the CP–RTs and P2 proteins have two hydrophobic regions. Directional alignment of these two helices also shows

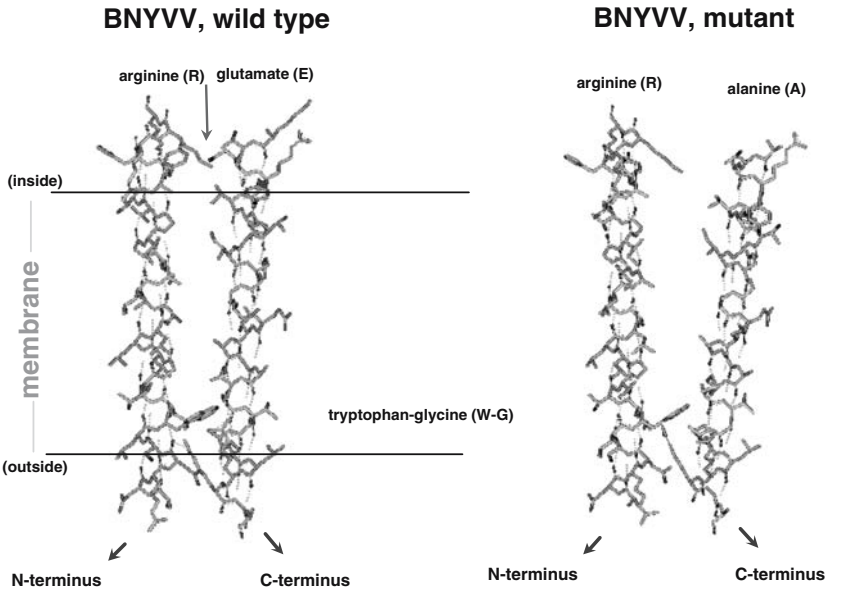

Figure 1.2. Models of the predicted helices and interfacial regions of the TM domains in the CP–RT of *Beet necrotic yellow vein virus* (BNYVV), showing the effects of the KTER>ATAR substitution that abolishes transmission by the plasmodiophorid vector *Polymyxa betae*, modeled using MOLMOL (Ver. 2.6) and displayed using the WebLab Viewer (from Adams *et al*., 2001). Electrostatic interactions are shown dotted.

evidence of compatibility between their amino acids, with groupings of amino acids that are either identical or in the same hydrophobicity group and evidence of possible fits between the small residues on one helix and the larger aromatic ones on the other. From these patterns, and from calculation of relative helix tilts, structural arrangements consistent with tight packing of TM helices were detected. These included ridge/groove arrangements between the two helices and strong electrostatic associations at the interfacial regions of the membrane. This suggests that the two TM domains could be paired within a membrane and with their C- and N-termini on the outside of the membrane. Nontransmissible deletion mutants lack the second of these putative TM regions and modeling of the BNYVV substitution suggests that it would disrupt the alignment of the polypeptide at a critical position adjacent to the second TM domain (Adams *et al*., 2001) (Figure 1.2). As there are few other similarities between the genomes of some of these viruses, it seems probable that the TM regions are instrumental in assisting virus particles to move across the vector membrane.

6. Other Membrane Proteins

Studies with *Southern cowpea mosaic virus* (genus *Sobemovirus*) have investigated the interaction of the CP with artificial membranes using a liposome dye-release assay and circular dichroism. The native CP and the R domain (which binds RNA and is usually on the inside of the spherical particle, but which is externalized under certain pH and salt conditions)

were shown to interact with liposomes *in vitro*. Studies of mutants, mapped the region responsible to residues 1–30 and analysis of this region by circular dichroism indicated that it assumes an alpha-helical structure when exposed to liposomes composed of anionic lipids (Lee *et al*., 2001). It has not yet been shown if this occurs *in vivo* but, if it does, the authors suggest that it could be related to cell-to-cell movement, to replication, to beetle transmission, or alteration of ion flux into or out of the cytoplasm. In our analyses, the region is not predicted to have TM properties.

Our TMPRED analyses indicate a few other plant virus proteins with strong TM properties, for which functions have not been assigned. These include the nonstructural protein P9-2 in *Rice black streaked dwarf virus* (RBSDV, genus *Fijivirus*, family *Reoviridae*) and its homologs in other members of the genus (P9-2 of *Nilaparvarta lugens reovirus*, P9-2 of *Fiji disease virus*, P10-2 of *Oat sterile dwarf virus*, and P8-2 of *Maize rough dwarf virus*). The TM domains occur in similar positions in the middle of the protein with the N- and C-termini exposed to the outside and a loop of 20–25 aa between them. No protein with similar properties can be identified in other plant-infecting reoviruses and it has not been detected within infected plants (Isogai *et al*., 1998). A further example is the small P6 protein encoded by Barley yellow dwarf viruses of the genus *Luteovirus* (family *Luteoviridae*), which all contain a single, strongly predicted TM region. Viruses assigned to other genera in the family do not appear to have this ORF and its function is not known.

7. Conclusions

It is clear that MPs play an essential role in the pathogenesis and movement within the plant of many plant viruses. However, studies of the structure and function of such proteins are still in their infancy. Substantial progress may be expected in the next few years, particularly in the area of cell-to-cell movement where viruses are proving useful tools to study the basic processes of macromolecular trafficking between adjacent plant cells.

Acknowledgments

We thank Prof. R.N. Beachy for permission to reproduce Figure 1.1 and Drs A. Tymon and K. Hammond-Kosack for helpful comments on the manuscript. Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

References

Aberle, H.J., Rutz, M.L., Karayavuz, M., Frischmuth, S., Wege, C., Hulser, D. *et al*. (2002). Localizing the movement proteins of Abutilon mosaic geminivirus in yeast by subcellular fractionation and freeze-fracture immuno-labelling. *Arch. Virol*. **147**, 103–117.

Adams, M.J. (2002). Fungi. In R.T. Plumb (ed.), *Plant Virus Vector Interactions* (*Adv. Bot. Res*. 36). Academic Press, San Diego, CA, pp. 47–64.

Adams, M.J., Swaby, A.G., and Jones, P. (1988). Confirmation of the transmission of barley yellow mosaic virus (BaYMV) by the fungus *Polymyxa graminis. Ann. Appl. Biol*. **112**, 133–141.

Adams, M.J., Antoniw, J.F., and Mullins, J.G.L. (2001). Plant virus transmission by plasmodiophorid fungi is associated with distinctive transmembrane regions of virus-encoded proteins. *Arch. Virol*. **146**, 1139–1153.