NMR of TMV

NUCLEAR MAGNETIC RESONANCE of TOBACCO MOSAIC VIRUS

Promotor: dr.T.J.Schaafsma, hoogleraar in de moleculaire fysica

J.L. DE WIT

NMR of TMV

NUCLEAR MAGNETIC RESONANCE of TOBACCO MOSAIC VIRUS

proefschrift

ter verkrijging van de graad
van doctor in de landbouwwetenschappen,
op gezag van de Rector Magnificus,
dr. H.C. van der Plas,
hoogleraar in de organische scheikunde,
in het openbaar te verdedigen
op vrijdag 26 mei 1978
des namiddags te vier uur in de aula
van de Landbouwhogeschool te Wageningen

NN08201, 726

STELLINGEN

1

Een eenduidig model ter verklaring van T_1 waarden van eiwitten kan alleen worden gebaseerd op frekwentie afhankelijke ^{13}C en ^{1}H T_1 metingen.

Dit proefschrift.

2

Het door Butler et al. gegeven model voor de vorming van het initiatiekomplex voor TMV assemblage bevat innerlijke tegenstrijdigheden.

Butler, P.J.G., Bloomer, A.C., Bricogne, G., Champness, J.N., Graham, J., Guilley, H., Klug, A. en Zimmern, D. (1976) Proc. of the third John Innes Symp. 101-110.

3

Het door Lebeurier et al. en Butler et al. voorgestelde elongatiemodel voor TMV assemblage is, gezien de secundaire en tertiaire RNA struktuur, uit thermodynamisch oogpunt niet mogelijk.

Lebeurier, G., Nicolaieff, A. en Richards, K.E. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 149-153.

Butler, P.J.G., Finch, J.T. en Zimmern, D. (1977) Nature 262, 217-219.

4

De juistheid van de gangbare opvatting, dat onverzadigde vetzuren in hun algemeenheid het cholesterol gehalte verlagen, is twijfelachtig en gaat ondermeer voorbij aan het feit, dat *trans* onverzadigde vetzuren cholesterol afbraak tegengaan.

Rivers, J. (1977) Nature 270, 2.

5

De door Packer et al., op grond van ^{13}C shifts van de 2', 6' koolstofatomen van tyrosine geschatte exchange snelheid tussen de ijzer-zwavel centra van ferredoxine, houdt geen rekening met een ferredoxine-methylviologeen evenwicht.

Packer, E.L., Sternlicht, H. en Rabinowitz, J.C. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 3278-3282.

6

De berekeningen van Chothia en Janin verschaffen aanwijzingen, dat de eiwit eenheden in de door deze auteurs onderzochte eiwitkomplexen beweeglijk zijn.

Chothia, C. en Janin, J. (1975) Nature 256, 705-708.

7

De door Sykes et al. gevonden relaxatietijden van verschillende eiwitten kunnen beter verklaard worden door rotatiediffusie om alle koolstof-koolstof bindingsassen over kleine hoeken aan te nemen.

Sykes, B.D., Hull, W.E. en Snyder, G.H. (1978) Biophys. Journal 21, 137-145.

8

Menig schoolboek zou beter tot zijn recht komen, wanneer het in stripvorm zou verschijnen.

Marx? Nooit van gehoord; Rius, Kritakfonds Nederland, Postbus 636, 3800 AP, Amersfoort.

9

Het is betreurenswaardig, dat Bohandy en Kim pas na een reeks publikaties over porfyrines in trifenyleen tot de konklusie komen, dat deze gastheer ongeschikt is.

Bohandy, J. en Kim, B.F. (1976) Spectrochim. Acta 32A, 1083-1088.

10

Het *in vitro* model voor antenne chlorophyll van Katz et al. vertoont meer inconsistenties met *in vivo* gegevens dan het door deze auteurs gepostuleerde en speculatief genoemde model voor het reaktiecentrum chlorophyllkomplex.

Katz, J.J., Norris, J.R. en Shipman, L.L. (1976) Brookhaven Symp. in Biol. 28, 16-55.

11

Het rapport "Het dilemma van de nationale landschapsparken" van de stichting "Natuur en Milieu" rechtvaardigt een gegrond wantrouwen en verzet van agrariërs tegen plannen voor wijziging van de landinrichting.

Proefschrift van J.L. de Wit.
Wageningen, april 1978.

aan mijn ouders
aan Afra

CURRICULUM VITAE

In 1965 heb ik het H.B.S.-B diploma gehaald aan het PIUS X college te Beverwijk. In hetzelfde jaar ben ik m'n studie aan de Landbouwhogeschool begonnen. Het kandidaatsexamen in de Levensmiddelentechnologie afgelegd in 1971. Het ingenieursdiploma is behaald in januari 1973, in de vakken Levensmiddelenchemie, Biochemie en Natuurkunde. In februari 1973 ben ik als wetenschappelijk medewerker in dienst getreden van de Landbouwhogeschool op de afdeling Moleculaire Fysica.

VOORWOORD

Toen ik ongeveer 5 jaar geleden dit onderzoek startte, beperkte het gebruik van kernspin resonantie (NMR) in de biologie zich tot kleine biomolekulen in oplossing, zoals eiwitten, substraten, oligonucleotiden en koolhydraten. De laatste jaren is het aantal toepassingen, door de ontwikkeling van geavanceerde 270 en 360 MHz kernspin resonantie apparatuur, echter sterk toegenomen. De traditionele school (kleine eiwitten, nucleinezuren e.d.) heeft zich verder ontwikkeld, maar daarnaast ontstaan er nieuwe toepassingen aan cellen en zelfs komplete organen. NMR biedt vooral in het laatste geval veel perspectief omdat dan, door het niet destructieve gebruik, de *in vivo* situatie benaderd wordt. Tot op de dag van vandaag worden vrijwel alle NMR studies echter nog steeds gedaan aan de bovengenoemde kleine biomolekulen. Op grond van theoretische argumenten neemt men aan, dat NMR aan grote biologische systemen niets of slechts ongestructureerde brede "bulten" laat zien.

Deze dissertatie is een bewijs van het tegendeel. Het is geboren uit behoefte naar een bruikbare techniek, waarmee men op moleculaire schaal grote biologische systemen *in oplossing* kan bestuderen. Röntgendiffraktie en elektronen mikroskopie hebben in deze hun wel bekende beperkingen; kernspin resonantie, zou hierin een taak kunnen vervullen. Gebruik van stabiele isotopen (^{15}N, ^{2}H, ^{13}C) zou deze taak verder kunnen verlichten. NMR en stabiele isotopen vormen de sleutelwoorden voor dit onderzoek. Tabaks mozaik virus is als model systeem gekozen, omdat het groot, simpel en goed gedefiniëerd is. Na bekende moeilijke tijden van vallen en opstaan kon een deel van het onderzoek afgerond worden. Tabaks mozaiek virus lijkt namelijk minder op een statisch bouwwerk, waarin alles zijn plaats heeft, dan op een levend organisme, flexibel en intern beweeglijk.

Ik wil ieder danken, die aan het eigenlijke onderzoek heeft meegewerkt. De plezierige sfeer in het laboratorium voor Moleculaire Fysica is, ondanks de nijpende behuizing, steeds van positieve invloed geweest op mijn werklust. Tjeerd Schaafsma wil ik danken voor de, ook in slechte tijden, morele en geestelijke steun aan dit onderzoek. Marleen Boerjan bedank ik voor de meer dan normale inzet gedurende de tijd, dat zij als part-time analiste bij dit

onderzoek betrokken was. Wij hebben voor de preparatieve kant van het werk een groot deel van onze tijd in het laboratorium van de Vakgroep Virologie doorgebracht. Mijn dank aan deze Vakgroep, waar ik als gast altijd plezierig heb gewerkt en van alle faciliteiten gebruik kon maken, is groot. De Vakgroep Moleculaire Biologie wil ik danken voor de verleende faciliteiten voor een deel van het onderzoek. Colette Alma en Lambert Dorssers hebben als doktoraal studenten in belangrijke mate bijgedragen tot het welslagen van dit werk. Ik ben het Instituut voor Plantenziektekundig Onderzoek zeer erkentelijk voor advies over en gebruik van de analytische ultracentrifuge. De CCMV, BMV en BBMV preparaten zijn in dank verkregen van Dick Verduin van de Vakgroep Virologie.

Tenslotte wil ik mijn waardering uitspreken voor de wijze, waarop Tine de Vries, Lia van Baren, dhr. A. Hoogeveen van de Vakgroep Moleculaire Biologie en mevr. W.M. Lao van de afdeling Tekstverwerking, zorg hebben gedragen voor het typen en tekenklaar maken van dit proefschrift.

CONTENTS

1. GENERAL INTRODUCTION

1.1 TOBACCO MOSAIC VIRUS

Tobacco Mosaic Virus (TMV) has a long scientific history. Purified and studied for the first time by Stanley in 1936 (1), it became a favoured model system in the fields of e.g. molecular biology, virology, physics, biochemics, biophysics and genetics. Since the preparation and properties of

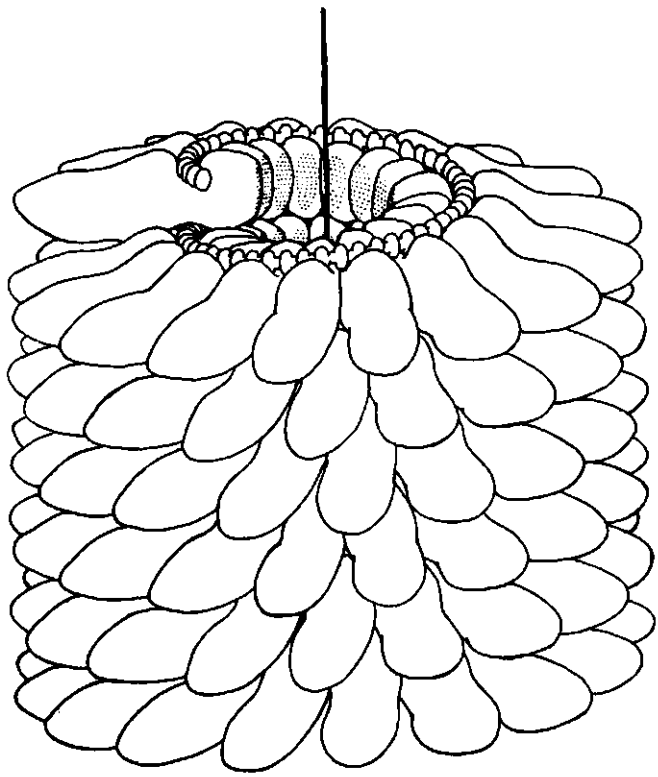


Fig. 1. Schematic model of TMV showing the helical arrangement of protein subunits and RNA (represented as a chain); reprinted from Caspar, D.L.D. (22), by Courtesy of Academic Press. New York.

the virus are well-characterized, it provides an attractive object for physical studies aimed at the elucidation of mechanisms underlying the elementary processes in molecular biology. The book of Kaper (7) gives a survey of the scientific history of TMV. TMV is a rod-shaped plant virus (molecular weight $\sim 42 \times 10^6$) consisting of 2,200 identical protein subunits (molecular weight $\sim$ 17,450) and one RNA chain of $\sim$ 6,600 nucleotides (2) (molecular weight $\sim 2 \times 10^6$). Its structure is schematically shown in Fig. 1. That the rod-like shape is based on a property of the protein subunits, became apparent when their assembly behaviour without RNA was studied. The protein subunits, depending on conditions such as protein concentration, temperature, pH, ionic strength, form a number of oligo- or polymers (3). Fig. 2. visualizes (from left to right) the rod-like protein oligomer which titrates anomalously, the "lock-washer" which is considered to be a hypothetical oligomer, the double disk

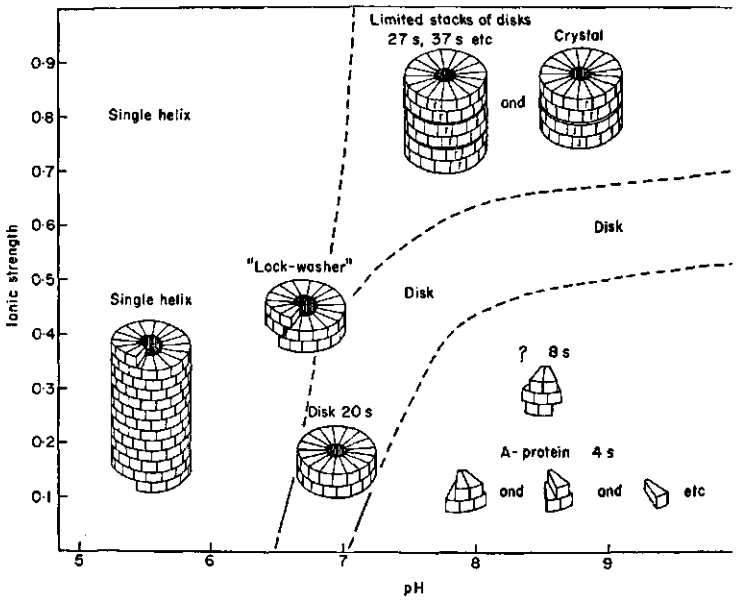


Fig. 2. Phase diagram containing stable protein oligomers as a function of pH and ionic strength. Other conditions: concn: 5 mg/ml; temperature: 20°C; reprinted from Durham, A.C.H. (3), by courtesy of Academic Press.

consisting of 34 protein subunits, stacks of double disks irreversibly formed at high ionic strength and smaller protein oligomers, such as pentamers, trimers and monomers. Purely monomer is only obtainable under extreme conditions (22).

The structure of the stacked disk has been analyzed by X-ray diffraction at 0.5 nm resolution (11). One of these oligomers, the so called double disk, occurs

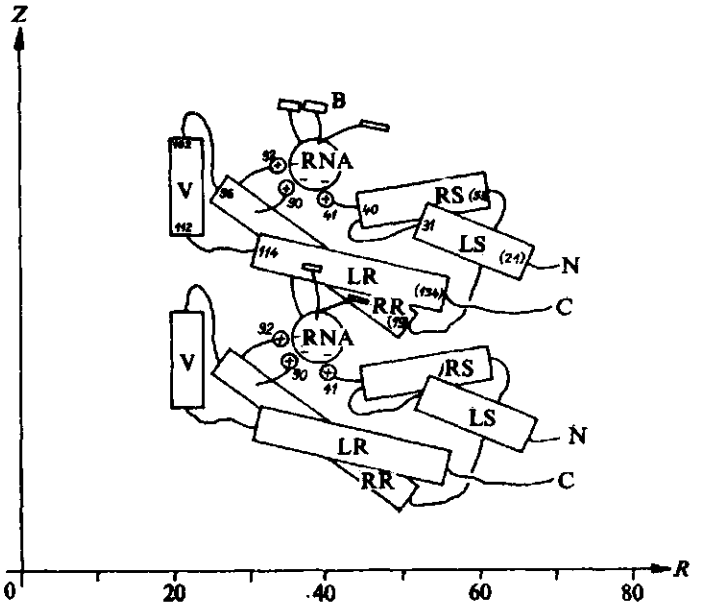
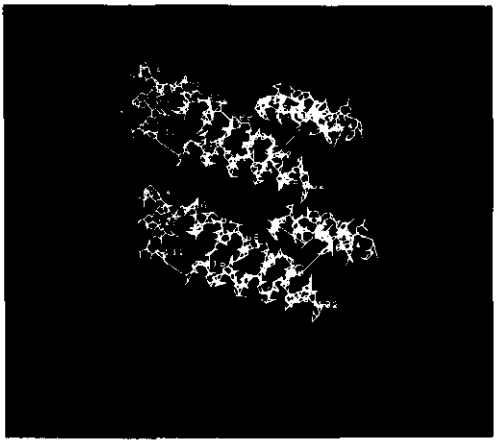


Fig. 3. Schematic drawing of the RNA and protein backbone in TMV (side view), based on X-ray diffraction data. B are RNA bases; LS, RS, LR, RR and V forms the nomenclature for the protein α-helix sections; numbers refer to amino acid sequence coding (Appendix 1); Z and R are the axial and radial axis of TMV respectively, the latter with a radius of ∿ 80 Å; reprinted from Stubbs, G. et al. (10), by courtesy of Macmillan Journals Ltd.

at physiological conditions and recognizes a TMV RNA section at about 1000 nucleotides from the 3' RNA end after which the RNA is encapsidated (4,5). The folding of RNA and protein in the virus is schematically shown in Fig. 3. This figure is obtained from X-ray diffraction at 0.4 nm resolution (10) on oriented gels of TMV, a study which presently can be considered as one of the major achievements of X-ray diffraction in molecular biology. The X-ray study also reveals a remarkable feature in the protein of TMV, the so-called carboxyl cage containing a number of aspartic and glumatic acids and arginines, in close contact with each other. These acid amino acids are thought to be reponsible for the anomalous titration behaviour (22) of the rod-like oligomers with pK's of carboxyl groups shifted to ∿ 7. Fig. 4A and B represent

A



B

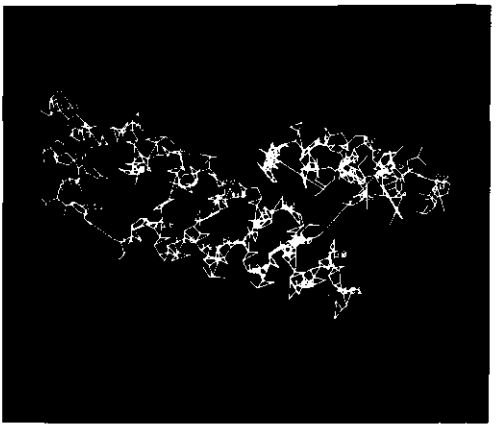


Fig. 4. This figure is a more detailed presentation of Fig. 3, without RNA. A: computer display of two protein subunits in TMV (side view). Numbering is according to. Appendix 1. B: computer display of one protein subunit in TMV at slightly larger magnification. Unknown regions are indicated by bars.

computer simulations of TMV protein subunits based on coordinates kindly made available by Dr. K.C. Holmes, Heidelberg. These pictures do not represent the actual structure in complete and reliable detail since only coordinates of the

α-helix regions have been displayed and, furthermore, since the phase of the α-helices is subject to some uncertainty. In general however these figures give a good overall impression of the geometry and relative orientation of the protein subunits in TMV. These simulations were generated by Ms. T. Trienekens in collaboration with the Computer Center and the Department of Mathematics of this University according to a modified program obtained from Dr.R.J. Feldman, Division of Computer Research and Technology, N.I.H., Bethesda, U.S.A.

From the point of view of a (molecular) biologist and virologist, a number of relevant- and sometimes intriguing- questions about TMV can be put forward.

a) how does TMV penetrate the cell (6)?
b) how does TMV dissociation occur under physiological conditions (7)?
c) how does the RNA coding for protein synthesis proceed (8)?
d) how is a specific section of the RNA chain recognized by the double-disk (4,5)?
e) how proceeds the RNA encapsidation and elongation (9)?

In order to answer these questions it is necessary to have a thorough knowledge about protein-protein and protein-RNA interactions in TMV in solution. Furthermore, from the viewpoint of a molecular physicist, it is fruitful to incorporate in any model answering the above mentioned questions that the system must obey thermodynamical and statistical principles even at the level, where observations suggest a mechanistic model.

The above mentioned interactions have in the past been studied by a large variety of methods such as X-ray diffraction (10,11), electron microscopy (12) and analytical ultracentrifuge experiments (13). Although this resulted in a rather detailed model of TMV and its protein oligomers, knowledge about protein-protein and protein-RNA interactions at the molecular level *in solution* is scanty because at the moment no method is suitable for this purpose, except Nuclear Magnetic Resonance (NMR). NMR can be considered to be one of the most powerful methods for the study of structure and dynamics of biomolecules in solution (14). Contrary to what one would expect following standard NMR theory, stating that NMR is only valuable for studying small biomolecules (15), this Thesis endeavours to demonstrate that it is possible and attractive to apply NMR to TMV and other large biosystems. Nature provided TMV with such properties that the study of TMV by NMR became feasible, yielding at least a partial answer to the questions a through e, mentioned above.

1.2 NUCLEAR MAGNETIC RESONANCE

In this Thesis all NMR experiments were carried out using High Resolution Fourier Transform (HR FT) NMR (16,17). The main differences with previous NMR studies on biomolecules are: the use of the highest available magnetic field stength to obtain the largest possible signal to noise ratio and resolution (17) (8.4 Tesla (T), 360 MHz for ^{1}H, 90.5 MHz for ^{13}C), the size of the biological systems studied (TMV, Cowpea Chlorotic Mottle Virus (CCMV), *E. coli* ribosomes). ^{13}C enrichment of the biological systems, and, finally the way the NMR spectra are interpreted. These interpretations are based on the spin-lattice (T_1) and spin-spin relaxation times (T_2) (17) and to a lesser extent on chemical shifts and spectral intensity measurements.

A simple introduction to the parameters T_1 and T_2 can be found in several textbooks (13,14,15). In large biological systems T_1 and T_2 are of crucial importance, since they determine the spectral intensity (signal to noise ratio) and resolution of NMR spectra. Presently, NMR spectra are mostly recorded in FT mode (16,17) and the resulting signal to noise (S/N) ratio of a number of accumulated spectra is largely determined by T_1. S/N linearly increases with $\sqrt{n}$, where n is the number of accumulations (18); for each spectrum to be properly recorded a practical rule says one should wait a few times T_1 before starting acquisition of the next spectrum (16,17). The T_2 parameter for a Lorentzian NMR line shape (16,17) can be calculated from $(\pi T_2)^{-1}$ which is the line width at half height. The resolution in NMR spectra, consisting of many partly resolved resonances, can be improved by increasing the magnetic field strength, so that these resonance are being increasingly separated. Commercially available superconducting magnets have for NMR purposes an upper limit of 8.4 Tesla at the moment.

T_1 and T_2 for rigid systems can be calculated using the equations for dipolar relaxation (19) (Appendix 3). The correlation time τ_R in these equations describing the rotational reorientation of a *rigid* sytem in solution can be approached by the Stokes-Einstein equation for spheres (20) (Appendix 3). As a rule of thumb, the line width can be obtained by dividing the molecular weight of the system by 10^3, e.g. $\sim$ 64 Hz for hemoglobine with molecular weight $\sim$ 64,000; $\sim$ 42,000 Hz for TMV. The values of T_1 in both cases would be $\sim$ 1 and $\sim$ 2000 s, respectively. Note that the maximum chemical shift range for ^{1}H at 360 MHz is $\sim$ 4000 Hz and that a normal acquisition time is $\sim$1 s. For TMV this implies that most of the spectra presented in this Thesis could only be obtained because TMV is in fact not at all rigid. In other words, the nuclei in

TMV are internally mobile (21,23,24). T_1 and T_2 can now be used to describe this internal mobility (Appendix 4) Molecular information of a system as large as TMV now also can be obtained in solution because of internal mobility.

This Thesis presents NMR spectra and relaxation data of both ^1H and ^{13}C nuclei in TMV. Although ^{13}C is $\sim$ 6000 times more insensitive than ^1H (13-15), the resolution at 8.4 Tesla (90.5 MHz) is so much better than that for protons that for enriched TMV and its protein oligomers more molecular information could be obtained than from ^1H NMR.

In Chapter 2 ^{13}C spectra of TMV and its protein oligomers are presented and it shown that TMV and oligomers are internally mobile. In Chapter 3 the molecular motion within the double disk-like oligomer is described based on T_1 and T_2 data. The short T_1's, observed in a number of other proteins have been explained assuming spin diffusion (25,26). The magnetic dilution of ^{13}C nuclei also in 12% ^{13}C enriched TMV protein warrants the absence of this mechanism. Chapter 4 contains the titration behaviour of spectral intensity of rod-like polymers of TMV protein which is interpreted in the framework of a model for internal mobility. In Chapter 5 internal mobility in sections of the protein chain in TMV, resulted in a more quantitive description of molecular dynamics in TMV and its protein oligomers. In Chapter 6 the enrichment procedures are reported and evidence is presented that TMV protein subunits in solution are structured not with standing the fact that there is considerable internal motion. A model for assembly of TMV from protein double disk and RNA is put forward in Chapter 7, based on internal mobility in the double disk-like oligomer, NMR results from Chapter 5, and thermodynamics. In this context the dissociation of TMV when penetrating the cell is discussed. Chapter 8 is an excursion to other biological systems, such as Cowpea Chlorotic Mottle Virus, *E. coli* ribosomes, aimed at demonstrating that NMR can be used to study other large biological systems. In Appendix 1 the amino acid sequence of TMV protein including sections with α-helix configuration is given; Appendix 2 summarizes chemical shifts of all carbons of the amino acids; Appendix 3 contains dipolar relaxation equations for rigid systems; Appendix 4 gives the dipolar relaxation equations describing internal mobility.

NMR spectra were taken by making use of three different machines: a Perkin-Elmer 23B 60 MHz spectrometer equipped for ^1H, made available by the Dept. of Organic Chemistry; A Varian XL-100 Fourier Transform (FT) spectrometer operating at 100 MHz. and equipped for ^1H and ^{13}C; a Bruker SPX 360 Supercon

FT NMR spectrometer, equipped for ^{1}H and ^{13}C, and made available by the Netherlands Organization for the Advancement of Pure Research.

Using the latter spectrometer the method of Convolution Difference (CD) spectroscopy (27) was applied to TMV ^{13}C spectra. In this method, the free induction signal originating from the ^{13}C nuclei, is multiplied by an exponential (e^{-at}) and stored. Then, the process is repeated with a second exponential (e^{-bt}) and the difference between the two resulting signals is subsequently fourier transformed, resulting in a Convolution Difference spectrum. In this way, broad spectral features are eliminated and the spectral resolution is enhanced.

1.3 REFERENCES

1. Stanley, W.M. (1936) Phytopathology 26, 305-307.
2. Lauffer, M.A. (1975) Entropy-Driven Processes in Biology, Springer-Verlag, Heidelberg.
3. Durham, A.C.H. (1972) J. Mol. Biol. 67, 289-305.
4. Butler, P.J.G., Bloomer, A.C., Bricogne, G., Champness, J.N., Graham, J., Guilley, H., Klug, A. and Zimmern, D. (1976) Proc. of the Third John Innes Symp., 101-110.
5. Guilley, H., Jonard, G., Richards, K.E. and Hirth, L. (1975) Eur. J. Biochem. 54, 145-153.
6. Durham, A.C.H. and Butler, P.J.G. (1975) Eur. J. Biochem. 53, 397-404.
7. The chemical basis of virus structure, dissociation and reassembly (1975) Kaper, J.M., North-Holland Publishing, Company, Amsterdam.
8. Bruening, G., Beachy, R.N., Scalla, R. and Zaitlin, M. (1976) Virology 71, 498-517.
9. Otsuki, Y., Takebe, I., Ohno, T., Fukuda, M. and Okada, Y. (1977) Proc. Natl. Acad, Sci. U.S.A. 74, 1913-1917.
10. Stubbs, G., Warren, S. and Holmes, K.C. (1977) Nature 267, 216-221.
11. Champness, J.N., Bloomer, A.C., Bricogne, G., Butler, R.J.G. and Klug, A. (1976) Nature 259, 20-24.
12. Durham, A.C.H. and Finch, J.T. (1972) J. Mol. Biol. 67, 307-314.
13. Dwek, R.A. (1975) Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems, Clarendon Press, Oxford.
14. NMR in Biology (1977) eds. Dwek, R.A., Campbell, I.D., Richards, R.E. and

Williams, R.J.P., Academic Press, London.

15. Wüthrich, K. (1976) NMR in Biological Research: peptides and proteins North-Holland Publishing Company, Amsterdam.
16. Farrar, T.C. and Becker, E.D. (1971) Pulse and Fourier Transform NMR, Introduction to Theory and Method, Academic Press, New York.
17. Gillies, D.G. and Show, D. (1972) in Ann. Rep. on NMR Spec., 560-673 ed. by Mooney, E., Academic Press, London.
18. Ernst, R.R. in Advances in Magnetic Resonance vol 2, 1-131 (1966) ed. by Waugh, J.S., Academic Press, New York.
19. Lyerla, J.R. and Levy, G.C. in Topics in Carbon-13 NMR Spectroscopy I, 79-143, (1974) ed. by Levy, G.C., Wiley-Interscience, London.
20. Hallenga, K. and Koenig, S.H. (1976) Biochemistry 15, 4255-4264.
21. Levine, Y.K., Partington, P., Roberts, G.C.K. (1973) Mol. Phys. 25, 497-514.
22. Caspar, D.L.S. (1963) Adv. in Protein Chemistry 18, 37-121.
23. Woessner, D.E. (1962) J. Chem. Phys. 36, 1-4.
24. Doddrell, D., Glushko, V. and Allerhand, A. (1972) J. Chem. Phys. 56, 3683-3689.
25. Kalk, A. and Berendsen, H.J.C. (1976) J. Magn. Resonance 24, 343-366.
26. Sykes, B.D., Hull, W.E. and Snyder, G.H. (1978) Biophys. J. 21, 137-146.
27. Campbell, I.D., Dobson, C.M;, Williams, R.J.P. and Xavier, A.V. (1973) J. Magn. Resonance 11, 172-181.

2. A ^{13}C NMR STUDY OF TOBACCO MOSAIC VIRUS AND ITS PROTEIN

J.L. de Wit, N.C.M. Alma, M.A. Hemminga, T.J. Schaafsma
submitted to Biochemistry.

2.1 INTRODUCTION

Tobacco mosaic virus (TMV) is a well-known plant virus (molecular weight 42 x 10^6) with length and radius of 300 and 9 nm respectively, containing 2,200 identical protein subunits (molecular weight 17,500) complexed in a helix with 6,600 RNA nucleotides (1). Both virus and its coat protein oligomers (2) have been widely studied by various techniques such as electron microscopy (3), circular dichroism (4), sedimentation analysis (5) and X-ray diffraction (6).

Three dimensional electron density maps for crystals of stacks of protein double disks and gels of the oriented virus with 0.5 and 0.4 nm resolution, respectively, have recently been obtained (6,7). Since the virus assembly is based upon protein-protein and protein-RNA interaction in solution, a comparison of the molecular models of the virus and protein (6,7) obtained from X-ray data and the molecular structure of TMV-RNA and protein in solution is necessary. As has been previously shown, Fourier Transform Nuclear Magnetic Resonance (FT NMR) can be used for this purpose (8). X-ray and NMR results on protein and RNA structure in crystal and solution have been compared in a few cases (8,9).

NMR increasingly focuses attention on molecular dynamics manifesting itself as mobility of protein side chain groups modulating protein structure (10). It has usually been assumed that this technique can only be successfully applied to fairly small protein molecules (molecular weight < 50,000), because one expects longer spin-lattice (T_1) and shorter spin-spin (T_2) relaxation times for larger proteins, resulting in a loss of signal to noise ratio (S/N) and spectral resolution, respectively.

We report ^{13}C NMR spectra for 12% enriched TMV (11) (strain Vulgare) and its coat protein of which pH-induced rod-like and temperature-induced double disk-like polymerization has been studied. A general discussion of these spectra in view of molecular dynamics for TMV and its protein oligomers is also presented.

2.2 ^{13}C NMR AT 90.5 MHz

The potential of conventional high resolution ^{13}C FT NMR at 15.1 MHz on small protein molecules (molecular weight < 50,000) has been shown (12). The much larger spectral resolution for ^{13}C in comparison with ^{1}H at the same magnetic field strength is an advantage, but the low sensitivity as a result of the 1.1% natural abundance and low gyromagnetic ratio is a large handicap for ^{13}C NMR. Although for small proteins a number of resonances (carboxyl, carbamide, aromatic, arginine and backbone carbons) with large chemical shift anisotropy broaden with increasing frequency (up to 90.5 MHz), both spectral resolution and S/N in general improve, especially for aliphatic carbons (13). For large rigid protein molecules (molecular weight > 50,000) and spectral repetition times (T) of $\sim$ 1 s, T_1 is found to be $\gtrsim$ T and the S/N is related to T_1 by (14),

$$S/N = QT_2 \ (1/T_1)^{\frac{1}{2}} \qquad [1]$$

where Q contains such factors as sample volume, concentration, probe quality and magnetic field. For these large proteins with rotational correlation times $\tau_R > 10^{-9}$ s, T_1 increases with increasing frequency in the absence of internal mobility and a decrease in S/N is expected (14). In such a case the experiments on enriched TMV, at 90.5 MHz described in this paper would be laborious with a need for large amounts of material. Since TMV-protein solubility is limited ($\lesssim$ 70 mg/ml), probes for sample tubes with diameters > 1 cm would be necessary. These are commercially not available at high magnetic fields (8.5 Tesla). Therefore, it was an unexpected result, that the S/N increased with about a factor 8 for protonated carbons of oligomers of TMV coat protein (molecular weight > 50,000) when comparing ^{13}C spectra at 90.5 and 25.4 MHz. All ^{13}C NMR experiments described in this paper are performed on a SPX Bruker supercon spectrometer in FT mode employing $\sim$ 5 W continuous wave ^{1}H decoupling power with 5 kHz bandwidth, an observe frequency for ^{13}C of 90.5 MHz and quadrature detection. A D_2O lock proved to be unnecessary for our measurements so that isotope effects of D_2O are avoided by using H_2O solutions. This is an additional advantage for ^{13}C NMR with respect fo ^{1}H NMR. The use of ^{13}C enriched TMV is necessary for sufficient S/N. The enrichment is easy and inexpensive with $^{13}CO_2$ as the carbon source (11). The optimum enrichment (15) at which no appreciable carbon-carbon J coupling is present is 10-15%.

2.3 90.5 MHz ^{13}C SPECTRA OF TMV

The ^{13}C spectrum of 12% enriched TMV purified according to Leberman (16) is given in figure 1A. A sedimentation coefficient of 130 S is found and electron micrographs show rod-like particles with length of 300 nm for the virus. No significant saturation of the observed protonated ^{13}C resonances is observed with T = 1 s, since a control experiment with T = 2 s showed that the intensity increase upon changing T from 1 to 2 s was less then the noise level (which was about 30% of the spectral intensity in this particular experiment). With the resonance intensity (I) given by (17),

$$I = I_{\infty} \left[1-\exp\left(-T/T_1\right)\right] \qquad [2]$$

this results in $T_1 \lesssim 1$ s.

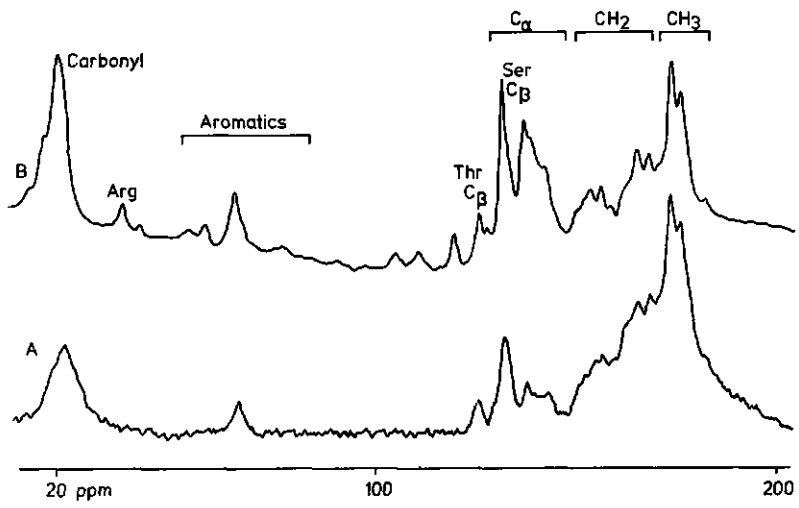


Fig. 1. Broadband ^{1}H decoupled 90.5 MHz ^{13}C NMR spectra of TMV and its RNA and protein constituents taken both at ∿ 30 °C with a concentration of 60 mg/ml (1 ml total volume) and 50,000 accumulations with spectral repetition time of 1 s and sensitivity enhancement of 30 Hz. The ppm scale is referenced to CS_2 assuming 125.8 ppm for the β-Thr carbon position. The vertical scales of A and B are different. Further conditions for: A, 1 mM sodiumphosphate pH 11; B, 1 mM sodiumphosphate pH 7.2.

In figure 1B the ^{13}C NMR spectrum of a clear solution of TMV dissociated in its RNA and native protein oligomers (average molecular weight < 50,000) is presented. The intensity of the native TMV spectrum (fig. 1A) compared to that of the spectrum of its free RNA and protein constituents for the backbone region between 120 and 150 ppm and the aliphatic region between 150 and 200 ppm is 6 and 17% respectively. The RNA resonances have low intensity and are mainly found in the region 40 to 125 ppm. The spectrum in figure 1B contains resonances of all carbons of the protein in the sample tube resulting from the molecular weight (< 50,000) of the protein oligomers under these conditions and a comparison with a ^{13}C spectrum of a known amount of lysozyme (12). On the average about 90% of the protein carbon resonances are not observed in the TMV spectrum (fig. 1A).

The line width parameter $(1/\pi T_2)$ for ^{13}C-H dipole-dipole relaxation in large rigid proteins for $\omega_C^2\tau_R^2 >> 1$, is given by (18),

$$1/T_2 = 0.2(\mu_o/4\pi)^2 N\gamma_C^2\gamma_H^2\hbar^2\tau_R/r_{CH}^6 \quad , \qquad [3]$$

where γ_C and γ_H are the gyromagnetic ratios in rad/s for ^{13}C and ^1H respectively, τ_R is the rotational correlation time, r_{CH} is the C-H distance and ω_C is the Larmor frequency in rad.s^{-1}, N is the number of protons interacting with a ^{13}C nucleus, $(\mu_o/4\pi)^2\gamma_C^2\gamma_H^2\hbar^2/r_{CH}^6 = 2.147 \times 10^{10} s^{-2}$ with $r_{CH} = 0.109$ nm. For spherical proteins τ_R can be calculated from the Stokes-Einstein equation (19),

$$\tau_R = M\bar{v}\eta/RT, \qquad [4]$$

where η is the viscosity, M is the molecular weight and $\bar{v}$ the partial specific volume. From equations [3] and [4], it can be calculated that the lines in the ^{13}C spectrum of the virus should be broadened with a factor of 800 compared to the lines in the spectrum of native protein oligomers (molecular weight < 50,000) and free RNA.

The factor 800 is a lower limit for the increase in line width. Taking into account the rod-like shape of the virus particle, (the known rotational correlation time of rod rotation about the lateral axis of 1.1×10^{-3} s(21) yields $\sim 1 \times 10^{-4}$ s for rod rotation about the axial axis from the moments of inertia) and the deviations from the Stokes-Einstein equation (20) still higher numbers are calculated.

In the absence of internal rotation and taking $\tau_R \sim 1 \times 10^{-4}$ s, T_1 is > 2000 s for the ^{13}C virus resonances as can be calculated from dipole-dipole spin-lattice relaxation for a C-H fragment (18), which is in contrast with the observed $T_1 \lesssim 1$ s.

These calculations demonstrate that then under our NMR conditions with a receiver dead time of ∿ 50 μsec and short spectral repetition times, virus resonances should be broadened and saturated beyond detection (22). The observable resonances of the ^{13}C virus spectrum must therefore arise from ^{13}C nuclei of the TMV coat protein with rotational degrees of freedom within the virus with correlation times $<< 10^{-4}$ s.

From the spectral regions indicated in figure 1 it can be seen that side chain resonances (β-Ser, β-Thr and aliphatic carbons) and a number of backbone resonances (C_α and carbonyl) have internal rotational degrees of freedom.

It is impossible to determine the increase in line width for the non-resolved resonances in the ^{13}C virus spectrum in comparison with the spectrum of its free constituents, because there is a large resonance overlap in the aliphatic region. From the ^{1}H virus spectrum at 100 MHz however, it can be concluded that the observable aliphatic ^{13}C resonances in the ^{13}C virus spectrum in any case have ^{13}C line widths < 300 Hz (23).

2.4 pH INDUCED ROD-LIKE PROTEIN POLYMERIZATION

The ^{13}C FT NMR spectrum of small oligomers of TMV protein purified according to Durham (2) with slight modifications (23) is shown in figure 2A. These protein oligomers have a $E_{282/252} > 2.5$ and a single sedimentation coefficient of 3.8 S in a Spinco analytical ultracentrifuge in 0.12 M Tris HCl pH 8.6 at 5 °C. This spectrum (Fig. 2A) is comparable with that given in Figure 1B, the only difference being the absence of RNA resonances. In figure 2B the ^{13}C NMR spectrum of TMV protein at pH 5.3 is shown. At this pH the protein polymer has a rod-like configuration with the protein subunits arranged in a helix, similar to the structure of the complete virus (24). An average intensity decrease of 90% is found with respect to figure 2A. Although this number is comparable to the intensity decrease observed for the virus, it is remarkable that now no significant line broadening is observed in comparing the high field flanks (180 ppm) of figure 2A and 2B. In analytical ultracentrifuge experiments no components with sedimentation coefficients < 40 S have been found at conditions given in the caption of figure 2B, in agreement with

literature data (25).

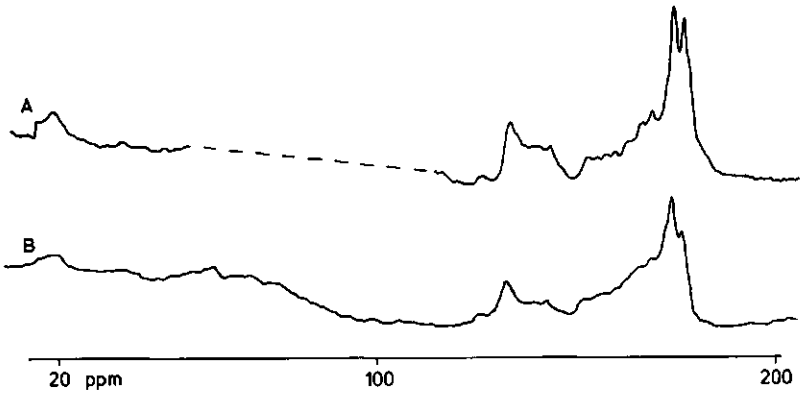


Fig. 2. Broadband ^{1}H decoupled 90.5 MHz ^{13}C NMR spectra of TMV protein oligomers with a concentration of 60 mg/ml (1 ml total volume) in 0.1 M NaCl taken at 7 ^{o}C with a spectral repetition rate of 0.5 s, a sensitivity enhancement of 45 Hz and 17,000 as the number of accumulations. The ppm scale is referenced to CS_2, assuming 125.8 ppm for the β-Thr carbon position. The vertical scales of A and B are different. Further conditions for A, pH 10.0; B, pH 5.3.

From the sedimentation coefficient an increase in molecular weight of at least 40 is calculated when going from 3.8 to > 40 S polymers. T_1 and T_2 should increase and decrease with this factor, respectively, if the coat protein oligomer is considered to be rigid, resulting in saturation and strong broadening of the ^{13}C resonances. Consequently, here again the residual intensity of about 10% must also arise from rotational degrees of freedom within the protein polymers.

2.5 TEMPERATURE-INDUCED DOUBLE DISK-LIKE PROTEIN POLYMERIZATION

Under "physiological" conditions, polymerization of TMV-protein results in a double disk-like configuration which has been shown to be important for virus assembly (25). In figure 3B and 3A ^{13}C spectra at 5 and 30 ^{o}C of oligomers with sedimentation coefficients 4 and 18 S are shown. From this figure and 100 MHz ^{1}H NMR experiments (23) it is concluded that no disappearance of spectral intensity and no line broadening is observed upon double disk-like polymerization. Consequently, the conclusion of rotational degrees of freedom

can be extended to all nuclei of both backbone and side chains within the double disk-like protein oligomer.

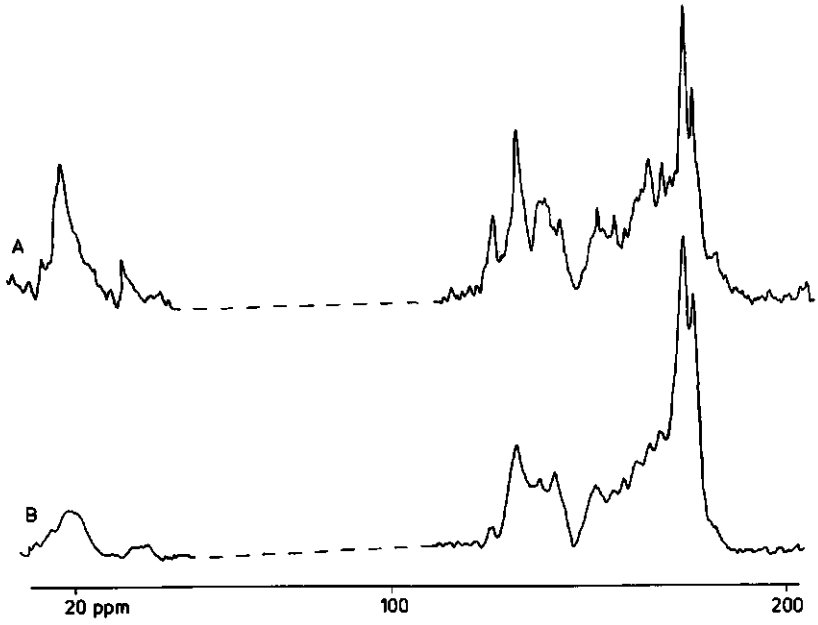


Fig. 3. Broadband ^{1}H decoupled 90.5 MHz ^{13}C NMR spectra of TMV protein oligomers in 0.1 M NaCl at pH 7.3 with 1 ml total volume. The ppm scale is referenced to CS_2, assuming 125.8 ppm for the β-Thr carbon resonance position. The aromatic regions are omitted because of large humps in the spectra. The vertical scales of A and B are different. Further conditions for A, 30 °C, 40 mg/ml, 4000 accumulations with T = 1 s and a sensitivity enhancement of 40 Hz; B, 7 °C, 60 mg/ml, 17,000 accumulations with T = 0.5 s and a sensitivity enhancement of 45 Hz.

In contrast figure 3 clearly shows a decrease in line width with increasing temperature for a number of resolved carbon resonances both in the spectral region for backbone and side chain carbons. At certain spectral positions a sharpening of the resonances is recognized for carbonyl-, δ-Arg-, β-Thr-, β-Ser- and other aliphatic carbons. A comparison of the total integrated aliphatic spectral region of figure 3A and the regions under the resolved resonances with exception of the methyl resonances, yields ∿ 10% for aliphatic resonances of which the line width decrease with an increase in temperature (fig. 3A and 3B); if methyl groups are included this number increases to ∿ 20%. NOE has been ignored in this calculation, although an effect is observed for methyl resonances in this case (unpublished results).

2.6 INTERNAL ROTATIONAL MOTIONS

In the presence of internal rotations the relaxation times T_1 and T_2 differ in their molecular weight dependence.

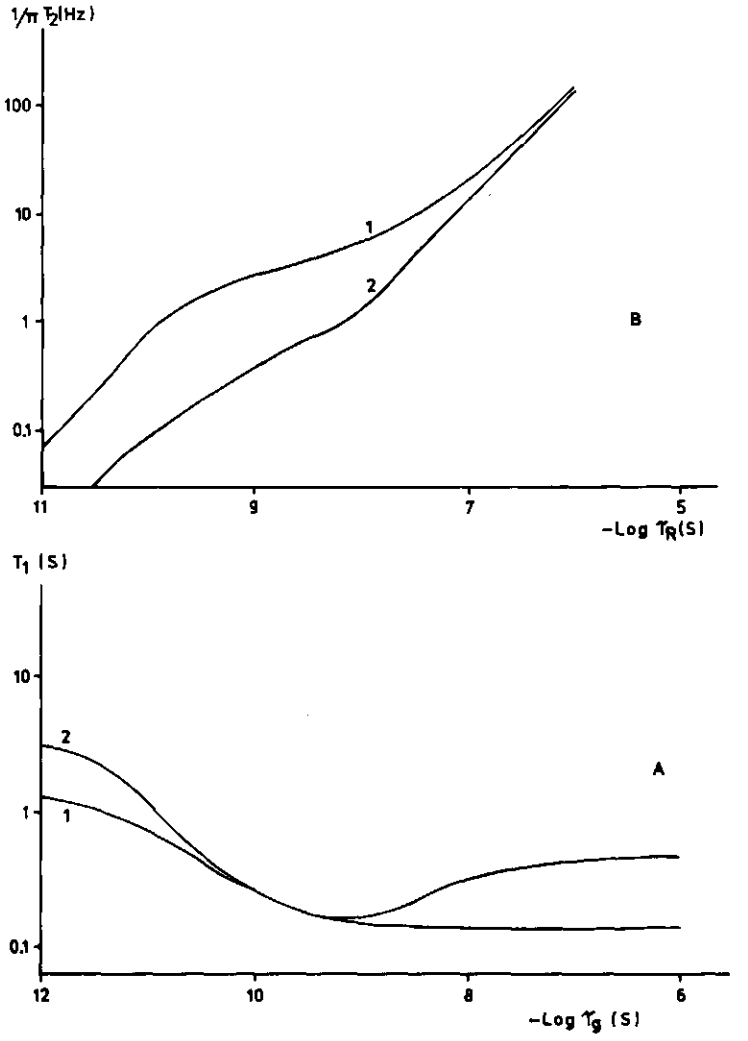


Fig. 4. A, the spin lattice relaxation time T_1 plotted against the internal rotational correlation time τ_g on a double logarithmic scale for overall rotational correlation times τ_R of 10^{-9} s (1) and 10^{-8} s (2); B, the line width parameter ($1/\pi T_2$) plotted against the overall rotational correlation time τ_R on a double logarithmic scale for internal rotational correlation times τ_g of 10^{-9} s (1) and 10^{-12} s (2).

According to Doddrell et al (26), taking a correlation time for one degree of freedom of internal rotation (τ_g) for a tetrahedral C-H fragment about its C-C axis over 360° and τ_R for the isotropic reorientation of the oligomer, the T_1 and T_2 dependence on τ_g and τ_R is graphically represented in figure 4 at an observe frequency of 90.5 MHz. According to Van Putte (28) a similar behaviour of T_1 with τ_g (fig. 4A) is found, when τ_g describes restricted rotations over angles << 360°. Furthermore, it can be shown that in proteins with a secondary and tertiary structure restricted rotations are possible (27).

From figure 4B it can be seen that T_2 depends on τ_R only when $\tau_R > 10^{-7}$. According to Doddrell (26) (with $\tau_g << \tau_R$) this τ_R-dependency of T_2 is given by equation [3] with a reduction factor of 1/9.

2.7 ROTATIONAL DEGREES OF FREEDOM WITHIN TMV AND ITS PROTEIN OLIGOMERS

From X-ray data it can be shown that TMV-protein double disks and virus possess distinct secondary, tertiary and quaternary structure (6, 7). This structure must also be present in solution in view of the following observations: 1) a large number of backbone-nitrogen bound protons is found to be non-exchangeable (29); 2) there are amino acid side chains which are non-reactive towards chemical agents because they are buried in the protein subunit (30) or protein-oligomer; 3) the presence of an RNA combining site; 4) a substantial amount of α-helix detected by ORD/CD (31); 5) a protein subunit structure and intersubunit interactions necessary for the rigidity of the protein oligomers; 6) ^{13}C NMR spectra indicate chemical shift inequivalence for at least a number of resolved resonances caused by the folding of the protein (unpublished results). Therefore, conclusions from our NMR data showing internal motion to a large extent in virus, rod-like protein oligomers and double disk-like protein oligomer must be interpreted keeping this in mind.

The experimental value of T_1 (< 1 s) for the virus can easily be accounted for by small angle rotations (27, 28) in contrast with an explanation for the behaviour of T_2. A τ_R of 1.1×10^{-6} s can be calculated based on Doddrell's model (26), if internal rotation is included using the reduction factor of 1/9 and a ^{13}C line width of 300 Hz (23) for virus carbons in equation [3]. Even under these conditions with a small τ_g value (< 10^{-9} s) this number is still 100 times smaller than the lower limit for the time constant of virus rotation of 1×10^{-4} s. Therefore, T_2 shows that the observable resonances represent regions in protein subunits within the virus with more than one degree of freedom of rotation.

An impression of what type of carbons are involved in these rotational motions can be acquired by comparing the weighted surfaces of different regions in the spectra (Fig. 1A and 1B). The observable 17% of the amino acid side chain groups with spectral positions between 148 and 200 ppm corresponds to a total number of 47 aliphatic carbons and consists of about 18 CH_3 carbons between 172 and 200 ppm and about 29 CH_2 and CH carbons between 148 and 172 ppm. The peak at 125 ppm corresponds to $\sim$ 2 β-Thr carbons. Depending on the number of β-Ser carbons at 132 ppm, a number of $\sim$ 10 α-carbons is calculated to be observable in the region 120 to 150 ppm. This number is comparable with the number of observable carbonyl resonances. No δ-Arg carbons are found and the resonance in the aromatic region at 65 ppm could correspond to 1 Phe side chain group. The number of side chain carbons showing rotational motions exceeds the number of mobile α-carbons with a factor of about 5. The accuracy of the calculations is estimated to be within $\pm$ 20%, the largest inaccuracy contribution arising from the determination of the relative spectral intensities (Fig. 1A and 1B). The observed α-carbons in the virus must belong to very mobile regions of the protein subunits, because backbone carbon rotations with more than one degree of freedom also force connected carbons into rotational motion. Another possibility is that the observed α-carbons among themselves are connected and belong to the same backbone section.

Although rod-like protein oligomers and virus yield comparable residual spectral intensity, the shape of their spectra differ considerably. For rod-like protein polymerization a reduction in intensity for the ^{13}C spectra is observed with only minor changes in spectral position and linewidth. Since RNA is now absent, the protein subunits can only interact with adjacent subunits. Rotational and translational motions of the protein subunit within the oligomer, motions of parts of amino acids within the protein subunit, and exchange of subunits between protein oligomers must be considered.

The ^{13}C spectra of the temperature-induced double disk-like oligomers differ from those of the pH-induced rod-like oligomers in that in the former case no loss in intensity and line broadening is observed upon polymerization leading to the conclusion that all nuclei of both backbone and side chains possess rotanional motions within the double disk-like oligomer. For about 20% of the protein resonances including a number of α-carbons, a molecular weight independent behaviour for the line width is found. This can not be explained using figure 4B, and these carbons therefore must have more than one degree of freedom for rotation. Since the remaining resonances are unresolved and $1/\pi T_2$ is calculated to be $<< \delta$, where δ is the chemical shift range for a number of unresolved

^{13}C resonances, the overall line shape of these resonances is less sensitive to an increase in molecular weight ($\sim$ 10x) upon protein polymerization from trimer to double disk. From an analysis of ^{1}H NMR spectra, however, it can be concluded that all nuclei within the double disk-like oligomer show a molecular weight independent line width behaviour, indicating also the presence of more than one degree of freedom of internal rotation (23) on a time scale $\sim 4 \times 10^{-8}$ s.

Internal rotational motions with correlation times $\lesssim 4 \times 10^{-8}$ may arise from one or more of the following causes: mobility within the protein subunit, mobility of the protein subunit within the oligomers and fast exchange ($> 2.5 \times 10^{7}$ s^{-1}) of protein subunits between oligomers. Exchange does not occur at this time scale, however, because of the large free energy decrease for protein subunits binding in an oligomer (25,23) and the absence of small molecular weight oligomers in analytical ultracentrifuge runs.

2.8 IMPLICATIONS FOR X-RAY DIFFRACTION

The 0.4 nm resolution electron density map for the virus allows for internal mobility of amino acids primarily on the exterior of the virus (6). In a paper by Caspar et al. (32) very small coordinated motions are suggested within the Dahlemense strain of TMV. It must be noted that there are large differences in extension and time scale between our NMR measurements and their suggestion.

Because only low resolution X-ray and NMR data are available, a correct comparison is as yet impossible, so that at the moment the number of nuclei showing τ_R-independent internal rotational motions is not inconsistent with X-ray analysis. For the rod-like protein oligomer no X-ray data have been published.

A comparison of the 0.5 nm resolution electron density map for crystals of stacks of double disks with NMR data for double disk-like oligomers in solution reveals large differences in molecular dynamical behaviour. Our data indicate that all nuclei within the double disk are mobile and even if these nuclei possess only one degree of freedom of internal rotation the electron density of the protein subunit should be completely smeared out. It is interesting to note that the number of nuclei showing line width decrease upon double disk-like polymerization (Fig. 3) roughly corresponds to the number of amino acids with a smeared-out electron density in the crystal (6). X-ray data have been obtained for stacks of double disks crystallized from 0.8 ionic strength solutions while NMR experiments were performed on oligomers in 0.1 to 0.2 ionic strength

solutions containing double disks. It is likely that the protein in the crystal is "frozen" by the constraints of crystal packing and periodicity. It must be realized that the widely different dynamical behaviour of protein subunits in the crystal and in solution must also imply conformational differences (10): a number of conformations must be present in solution to account for the observed internal mobility. Preliminary experiments show that at high ionic strength the spectral intensity starts to disappear for the double disk-like oligomer with a spectral behaviour similar as observed for the rod-like oligomer. It must be realized, however, that amino acids with smeared out electron density in an X-ray diffraction pattern may arise from spatial or time-dependent disorder while NMR measures only time-dependent disorder on a short time scale.

2.9 IMPLICATIONS FOR NMR

Because of large internal rotational motions TMV, and its oligomers with different types of assembly (rod-like and disk-like oligomers) can be studied in solution by NMR. The molecular weight does not appear to be the limiting factor for application of ^{1}H and ^{13}C NMR. From preliminary experiments it is also concluded that TMV is no special case, since high internal mobility is found in a number of viruses, phages and ribosomes. Conventional FT NMR is therefore not restricted to low molecular weight biomolecules and opens possibilities to approach the in vivo situation, by obtaining information on molecular level of high molecular weight biosystems in solution.

2.10 ACKNOWLEDGEMENTS

The technical assistance of Miss M. Boerjan with the preparation of TMV protein is gratefully acknowledged. We are indebted to the Department of Virology for providing laboratory facilities. The NMR investigations at high field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

2.11 REFERENCES

1. Finch, J.T., J. Mol. Biol. 66, 291-294 (1972).
2. Durham, A.C.H., J. Mol. Biol. 67, 289-305 (1972).

3. Durham, A.C.H. and Finch, J.T., J. Mol. Biol. 67, 307-314 (1972).
4. Schubert, D. and Krafczyk, B., Biochim. Biophys. Acta 188, 155-157 (1969).
5. Paglini, S. and Lauffer, M.A., Biochemistry 7, 1827-1835 (1968).
6. Stubbs, G., Warren, S. and Holmes, K., Nature 267, 216-221 (1977).
7. Champness, J.N., Bloomer, A.C., Bricogne, G., Butler, P.J.G. and Klug, A., Nature 259, 20-24 (1976).
8. NMR in Biology (ed. Dwek, R.A., Cambell, I.D., Richards and Williams, R.J.P., Academic Press, London 1977).
9. Robillard, G.T., Tarr, C.E., Vosman, F. and Berendsen, H.J.C., Nature 262, 363-369 (1976).
10. Gelin, B.R. and Karplus, M., Proc. Nat. Acad. Sci. USA 72, 2002-2006 (1975).
11. De Wit, J.L., Zeestraten-Alma, N.C.M. and Schaafsma, T.J., to be published, Chapter 6.
12. Oldfield, E., Norton, R.S. and Allerhand, A., J. Biol. Chem. 250, 6381-6402 (1975).
13. Topics in carbon-13 NMR spectroscopy (ed. Levy, G.C., J. Wiley, Interscience, New York, 1974).
14. Egan, W., Shindo, H. and Cohen, J.S., Ann. Rev. Biophys. Bioeng 6, 383-417 (1977).
15. Wüthrich, K. NMR in Biological Research: peptides and proteins. (North Holland Publishing Company, Amsterdam, 1976).
16. Leberman, R., Virology 30, 341-347 (1966).
17. Freeman, R. and Hill, H.D.W., J. Chem. Phys. 54, 3367-3377 (1971).
18. Oldfield, E., Norton, R.S. and Allerhand, A., J. Biol. Chem. 250, 6368-6380 (1975).
19. Dwek, R.A., Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems (Clarendon, Oxford, 1973).
20. Hallenga, K. and Koenig, S.H., Biochemistry 15, 4255-4264 (1976).
21. Tanford, C., Physical Chemistry of Macromolecules (Wiley, New York, 1961).
22. Torchia, D.A. and VanderHart, D.L., J. Mol. Biol. 104, 315-321 (1976).
23. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J., J. Magn. Resonance, in press,Chapter 3.
24. Caspar, D.L.D., Adv. in Protein Chemistry 18, 37-121 (1963).
25. Lauffer, M.A., Entropy-driven processes in Biology (Springer, New York, 1975).
26. Doddrell, D., Glushke, V. and Allerhand, A., J. Chem. Phys. 56, 3683-3689 (1972).

27. Mc Cammon, J.A., Gelin, D.R., Karplus, M. and Wolynes, P.G., Nature 262, 325-326 (1976).
28. Van Putte, K., J. Magn. Resonance 2, 23-34 (1970).
29. Budzynski, A.Z. and Fraenkel-Conrat, H., Biochemistry 9, 3301-3309 (1970).
30. Durham, A.C.H. and Butler, P.J.G., Eur. J. Biochem. 53, 397-404 (1975).
31. Simmons, N.S. and Blout, E.R., Photochem. Photobiol. 8, 81-92 (1968).
32. Caspar, D.L.D. and Holmes, K.C., J. Mol. Biol. 46, 99-133 (1969).

3. A ^{13}C AND ^{1}H NMR STUDY OF THE DYNAMIC BEHAVIOUR OF TOBACCO MOSAIC VIRUS PROTEIN

J.L. de Wit, M.A. Hemminga and T.J. Schaafsma
J. Magn. Resonance, in press.

3.1 INTRODUCTION

Tobacco Mosaic Virus (TMV) is a rod-shaped plant virus consisting of RNA and 2,130 identical protein subunits with a molecular weight of 17,500. The total molecular weight of the virus is 42 x 10^6 (1). A protein/RNA model of the virus has recently been published based on a 4Å X-ray diffraction map (2). The coat protein of TMV has been well-characterized by various techniques such as sedimentation analysis (3), electron microscopy (4) and X-ray diffraction (5).

There is an increasing amount of NMR literature on proteins. Because one expects long-spin lattice relaxation times (T_1) and severe line broadening ($1/\pi T_2$) for large rigid biomolecules, conventional ^{1}H and ^{13}C FT NMR have been restricted to fairly small systems with molecular weight < 50,000 (6).

Recently we introduced ^{13}C FT NMR as a promising technique for the study of the dynamic behaviour of TMV and its coat protein (7). Although one normally assumes a rigid structure for TMV and its protein oligomers, as indicated by electron microscopy (4) and X-ray diffraction (5), about 10% of the carbon atoms of the protein in the virus shows a dynamic behaviour on a time scale upon which conventional FT NMR is sensitive (7). In addition we found increased mobility in the oligomers of TMV coat protein without RNA (7).

In this paper we present ^{1}H and ^{13}C frequency dependent T_1 and T_2 data for different types of protein oligomers, without discussing single resonance behaviour. The results are interpreted in terms of a high degree of internal motion within the protein oligomers and thermodynamic parameters defining their stability.

3.2 EXPERIMENTAL

3.2.1 TMV-protein preparation

For ^{13}C NMR experiments 10-15% randomly ^{13}C-enriched TMV was used. The

preparation of the enriched virus will be described elsewhere (Chapter 6). Normal and ^{13}C-enriched TMV strain Vulgare was purified according to Lebermann (8). TMV-protein was prepared according to Durham (3). This method was slightly modified in that 0.1 mg/ml Macaloid was added to all solutions and the virus was dissociated in protein and RNA in 0.01 M-NaOH until the solution became clear. The time for clarifying the solution was always less than 30 minutes. Protein stock solutions were stored in water at 5 ^{o}C at pH $\approx$ 5 and were characterized by sedimentation analysis, gel electrophoresis and optical density measurements. For sedimentation analysis a Spinco analytical ultracentrifuge was used, equipped with an automatic speed control. The protein always showed a single boundary with 3-4 S in 0.12 M-Tris HCl, pH 8.6 at 10 ^{o}C and a concentration of 30 mg/ml. Also a single protein band was found on 5% gel electrophoresis in 0.12 M-Tris HCl, pH 8.6 at 5 ^{o}C, using recrystallized acrylamide (9). In all cases the optical density ratio $E_{282/252}$ was > 2.5. The residual nucleotide content of the protein samples was < 10 µg/g protein as determined by phosphate analysis (10). Protein concentrations were determined spectrophotometrically with $E_{280}^{0.1\%} = 1.27$. For all ^{1}H NMR measurements the protein was extensively dialyzed against 99.8% D_2O (Biorad) at pD $\approx$ 6. NaOH and NaOD were used to adjust the pH and pD. For the pD measurements, the pH meter reading was used (11).

3.2.2 NMR measurements

100 MHz ^{1}H and 25.2 MHz ^{13}C NMR measurements were carried out on a Varian XL-100 spectrometer in FT mode, using 10 Watt ^{1}H noise decoupling with 1.5 kHz bandwidth for ^{13}C. 360 MHz ^{1}H and 90.5 MHz ^{13}C NMR were performed on a Bruker SPX-360 supercon spectrometer in FT mode with quadrature detection using 4 Watt ^{1}H noise decoupling with 5 kHz bandwidth for ^{13}C. A Hitachi Perkin-Elmer R 24B was used for the 60 MHz ^{1}H NMR measurements. Spin lattice relaxation times T_1 were determined both by progressive saturation and inversion recovery (12). On the Bruker SPX-360 and Varian XL-100 spectrometers temperature was kept constant within 5 ^{o}C over a period of 8 hours by flowing cold evaporated nitrogen from a liquid nitrogen container directly through the probe, permitting long time averaging experiments.

3.3 RESULTS

3.3.1 T_1 data for ^1H at 100 MHz and 360 MHz

As shown in Fig. 1, the resonance overlap in protein oligomers is very severe. Relaxation can only be measured for various spectral positions, each consisting of a number of overlapping resonances, as indicated in Fig. 1. It is found that this relaxation is approximately exponential with T_1 given in Table 1. It is clear that such a T_1 parameter is a measure of the time decay of the exponentials of different overlapping resonances. Therefore the real differences between the T_1 values of resolved resonances can be much larger as indicated in Table 1.

The molecular weight of the protein oligomers ranges from 0.5 to 1 x 10^6, as calculated from sedimentation analysis. This is in agreement with results obtained by Paglini and Lauffer (13) and molecular weights predicted from the linear condensation theory (14).
The oligomers are mainly in the double disk configuration, consisting of 34 protein subunits (15).

For dipole-dipole interaction between protons, T_1 is given by (16)

$$1/T_1 = 0.6\ (\mu_o/4\pi)^2\ N\ \gamma_H^4\ \hbar^2\ r_{HH}^{-6}\ \omega_H^{-2}\ \tau_c^{-1}, \qquad [1]$$

assuming $\omega_H^2\ \tau_c^2 >> 1$. N is the number of protons interacting with the observed proton, γ_H is the ^1H gyromagnetic ratio, r_{HH} is the proton-proton distance, ω_H is the Larmor frequency in rad. sec^{-1} and τ_c is the dipolar correlation time. The term $(\mu_o/4\pi)^2\ \gamma_H^4\ \hbar^2\ r_{HH}^{-6} = 1.791 \times 10^{10}\ sec^{-2}$ with $r_{HH} = 0.178$ nm for the H-H distance in a CH_3-group.

Because of the overlap in the ^1H NMR spectrum, it is not possible to study the relaxation behaviour of single resonances. Therefore Eq. [1] will be used to determine upper limits for τ_c in accordance with the largest experimental values of T_1.

The dipolar correlation time τ_c can be written as

$$\tau_c^{-1} = \tau_R^{-1} + \tau_m^{-1}, \qquad [2]$$

where τ_R is the rotational correlation time of rigid protein oligomers and τ_m the residence lifetime of protein subunits in an oligomer. τ_m is given by (17)

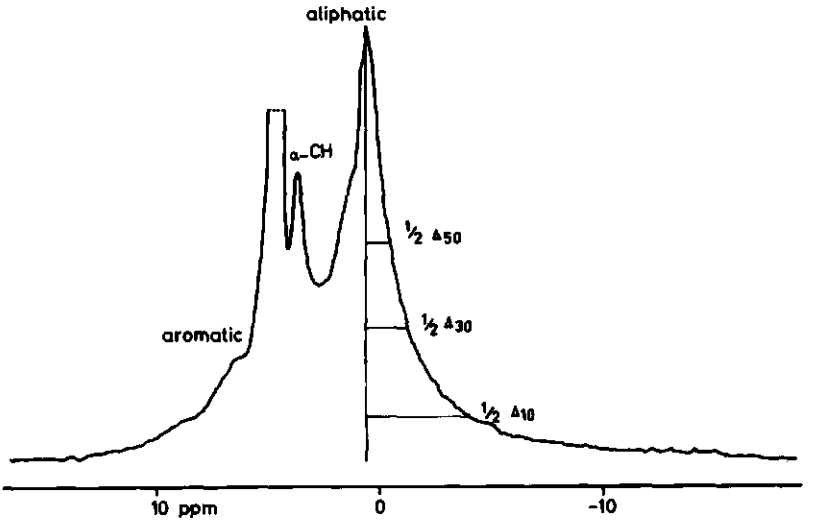


Fig. 1. 360 MHz ^{1}H NMR spectrum of oligomers of TMV coat protein in D_2O. The molecular weight is between 0.5 and 1 x 10^6. Conditions: conc. 40 mg/ml, temp. 31 °C, pD 7.6, 0.1 M-NaCl. No reference for the ppm scale was used. Arrows indicate positions where T_1 is determined. The regions in the spectrum at about 4.0 ppm and 2.5 ppm are omitted because of the strong resonance of HDO and a strong spike arising from the quadrature detection, respectively. Acquisition time plus pulse delay: 1 sec, number of accumulations: 447, sensitivity enhancement: 1.25 Hz.

Table 1. Spin lattice relaxation times T_1 for ^{1}H at 100 and 360 MHz for various spectral positions[a]

Position[b]	1	2	3	4	5	6	7
T_1(s) at 100 MHz[c]	-	0.32	0.32	0.32	0.32	-	-
T_1(s) at 360 MHz[d]	0.36	0.70	0.97	0.97	0.73	1.1	1.1

[a] See caption of Fig. 1 for experimental conditions. T_1 values are uncorrected for the number of protons interacting with the observed protons. The accuracy of the T_1 values is ± 10%.

[b] Numbers correspond to arrows in Fig.1.

[c] T_1 determined by inversion recovery.

[d] T_1 determined by progressive saturation.

$$\tau_m^{-1} = (kT/h) \exp(-\Delta G^a/RT), \quad [3]$$

where k is the Boltzmann constant, T the absolute temperature and ΔG^a is the free energy of activation. The free energy of binding (ΔG) of protein subunits in the smallest oligomer, the trimer, is -50 kJ mole^{-1} (18). Because $\Delta G^a > \Delta G$, a lower limit for τ_m of 8.7×10^{-5} sec is calculated.

The rotational correlation times for protein oligomers, in the absence of motions within the oligomers, can be calculated from their molecular weights using the Stokes-Einstein equation (17)

$$\tau_R = M \bar{v} \eta / RT, \quad [4]$$

assuming a spherical geometry and neglecting hydration effects. M is the molecular weight, $\bar{v}$ the partial specific volume (0.71×10^{-3} m^3 kg^{-1}) and η is the viscosity of the solvent. Hallenga and Koenig (19) have shown that Eq. [4] can be used for a large range of molecular weights including non-spherical macromolecules, to calculate approximate values for τ_R. The range of molecular weights of 0.5 to 1×10^6 for the protein oligomers corresponds to τ_R values varying between 1.4 and 2.8×10^{-7} sec, respectively. Since $\tau_m \gg \tau_R$, Eq. [2] shows that $\tau_c = \tau_R$. Therefore the contribution of protein exchange is ignored in this paper.

Inserting in Eq. [1] the smallest value 1.4×10^{-7} sec for τ_R, it is found that T_1 = 67 sec at 360 MHz and T_1 = 5.1 sec at 100 MHz for N = 1. There is a large deviation between the calculated and experimental T_1 values given in Table 1. Part of this deviation can be attributed to paramagnetic effects of oxygen (20). The effect of dissolved oxygen could be important for the region $T_1 \gtrsim 2$ sec (20). The longest relaxation time in Table 1 is about 1 sec for aromatic protons and a maximum oxygen effect of 30% may be expected for these protons, which is too small to explain the discrepancy between theoretical and experimental T_1 values. Rather, we are led to conclude that there must be internal motions within the protein oligomers. These internal motions may arise from mobile protein subunits within the oligomers and/or mobile protons within the subunit.

If we now calculate T_1 again, assuming the extreme case that the protein subunits with molecular weight of 17,500 are isotropically reorienting in the oligomers with a τ_R of 5×10^{-9} sec, Eq. [1] yields T_1 = 2.4 sec at 360 MHz. η is assumed to be the viscosity of water. This T_1 value is still larger than

that given in Table 1, demonstrating that the existence of subunit motion alone is not sufficient to explain the T_1 data. Therefore it is concluded that there are motions of protons within the subunit as well, their correlation times being smaller than 5×10^{-9} sec. At 100 MHz a value of 0.18 sec is calculated for T_1. Although this value is lower than the T_1 values in Table 1, it does not affect the above mentioned conclusion, but only indicates that the assumption of isotropic rotational motion of the subunit with τ_R of 5×10^{-9} sec is not correct.

From Table 1 it can be seen that at higher frequency the T_1 values start to deviate from each other. In the theory of spin diffusion (21), where the diffusion process tends to equalize the T_1 values, smaller T_1 differences are predicted at higher frequency. This contradicts with the data in Table 1 and therefore spin diffusion is excluded.

3.3.2 T_1 data for ^{13}C at 25.2 and 90.5 MHz

Fig. 2 gives the 90.5 MHz ^{13}C NMR spectrum of about 12% ^{13}C-enriched protein oligomers with molecular weights 0.5 to 1 x 10^6. The spectral resolution is much better than the resolution of the 360 MHz spectrum of Fig. 1, because of the larger chemical shift range for ^{13}C and the absence of ^{13}C-^{13}C J-coupling. This allows a determination of T_1 for groups of carbons in the protein, such as CH_3, CH_2, C_α, C_β of serine, C_β of threonine, C_β of arginine and carbonyls. T_1 data of the various spectral positions of Fig. 2 obtained at 25.2 and 90.5 MHz are collected in Table 2. In all cases the relaxation of the spectral regions turned out to be approximately exponential.

The ^{13}C dipolar spin lattice relaxation time T_1 of aliphatic carbons is given by (22)

$$1/T_1 = (\mu_o/4\pi)^2 \frac{N\gamma_C^2 \gamma_H^2 \hbar^2}{10 r_{CH}^6}\left[\frac{\tau_R}{1+(\omega_H-\omega_C)^2\tau_R^2} + \frac{3\tau_R}{1+\omega_C^2\tau_R^2} + \frac{6\tau_R}{1+(\omega_H+\omega_C)^2\tau_R^2}\right], \qquad [5]$$

where γ_C is the ^{13}C gyromagnetic ratio, r_{CH} the proton-carbon distance and ω_C the ^{13}C Larmor frequency. The term $(\mu_o/4\pi)^2\gamma_C^2\gamma_H^2\hbar^2 r_{CH}^{-6} = 2.147 \times 10^{10}$ sec^{-2} with $r_{CH} = 0.109$ nm for a CH, CH_2, or CH_3 group. Taking the smaller value of τ_R, (1.4×10^{-7} sec), for the double disks, Eq. [5] yields $T_1 = 7.0$ sec at 90.5 MHz and $T_1 = 0.55$ sec at 25.4 MHz for N = 1. The values of T_1 obtained at 90.5 and 25.2 MHz are larger than those found experimentally, showing again the presence of motions within the protein oligomers, in accordance with the ^1H results presented before.

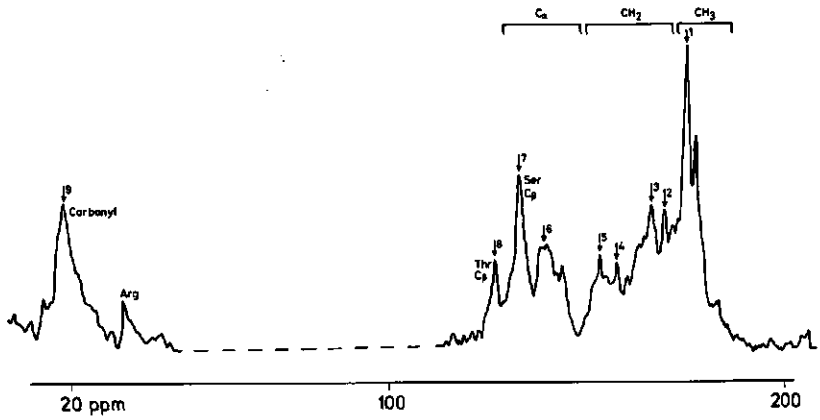


Fig. 2. Broad-band noise decouples ^{13}C NMR spectrum of oligomers of TMV protein at 90.5 MHz. The protein is enriched for about 12% with ^{13}C. The molecular weight is between 0.5 and 1 x 10^6. Conditions: conc. 40 mg/ml, temp. 30 °C, pH 7.3, 0.1 M-NaCl. The ppm scale was calculated refering to CS_2 as a standard. Arrows indicate positions where T_1 is determined. Because of a large background hump in the spectrum, the region between 50 and 100 ppm is omitted and the phases of the aliphatic, backbone and carbonyl/arg region have been adjusted separately.
Acquisition time plus pulse delay: 1 sec, number of accumulations: 4000, sensitivity enhancement: 40 Hz.

Table 2. Spin lattice relaxation times T_1 for ^{13}C at 25.2 and 90.5 MHz for various spectral positions[a]

Position[b]	1	2	3	4	5	6	7	8	9
T_1(s) at 25.2 MHz	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	-
T_1(s) at 90.5 MHz	0.22	0.26	0.27	0.42	0.37	0.34	0.27	0.35	3.0

[a] See caption of Fig. 2 for experimental conditions.
T_1 values are determined by progressive saturation and are uncorrected for the number of protons interacting with the observed ^{13}C nuclei.
The accuracy of the T_1 values is ± 10%.

[b] Numbers correspond to arrows in Fig. 2.

3.3.3 *^1H line shape analysis at 360, 100 and 60 MHz*

The line shape of an NMR spectrum is determined by the chemical shift σ and intrinsic line width $1/\pi T_2$ of the resonances. The spin-spin relaxation time T_2 for ^{13}C or ^1H is given by (16, 22)

$$1/T_2 = N\ D\ \tau_R, \qquad [6]$$

assuming that $\omega_C^2\ \tau_R^2 >> 1$ for ^{13}C and $\omega_H^2\ \tau_R^2 >> 1$ for ^1H. The term D is given by $0.2(\mu_o/4\pi)^2\gamma_C^2\gamma_H^2\hbar^2\ r_{CH}^{-6}$ for ^{13}C and $0.45(\mu_o/4\pi)^2\gamma_H^4\hbar^2\ r_{HH}^{-6}$ for ^1H.

In order to obtain additional information about the dynamic behaviour of the TMV coat protein it is of interest to know the line widths of all individual resonances in a ^{13}C and ^1H NMR spectrum. For the best resolved ^{13}C NMR spectrum, shown in Fig. 2, there is strong resonance overlap, preventing an accurate determination of line widths of all resonances. It is possible, to find an upper limit for the line width $1/\pi T_2$ of all components in a group of overlapping resonances, however, by considering their line shape. When $\Delta\sigma$ is the chemical shift range of a group of overlapping resonances, the line shape will generally be insensitive to T_2 if $\Delta\sigma >> 1/\pi T_2$. On the other hand, the line shape will be sensitive to T_2 and independent of $\Delta\sigma$ if $\Delta\sigma << 1/\pi T_2$.

To approach this situation in practice the chemical shift range must be made as small as possible. Since T_2 is frequency independent (see Eq. [6]), this can be realized experimentally by varying the measuring frequency and determining the point where the line shape is insensitive to $\Delta\sigma$. Compared to ^{13}C, the chemical shift range for ^1H is much smaller. Therefore the line shape of the aliphatic protons is measured at three different frequencies. In Table 3, the line width parameters Δ_{10}, Δ_{30} and Δ_{50} of the aliphatic line shape are presented at 360, 100 and 60 MHz. The definition of Δ_{10}, Δ_{30} and Δ_{50} is given in Fig. 3.

From Table 3 it can be seen that there is almost no change in the line width parameters when going from 100 to 60 MHz, indicating that the aliphatic line shape at 100 MHz is independent of $\Delta\sigma$ and thus has the largest sensitivity to variations of T_2, as compared to 360 MHz ^1H and 90.5 and 25.4 MHz ^{13}C line shapes. An upper limit for the line width $1/\pi T_2$ of a resonance in the aliphatic line shape is approximately equal to Δ_{50}, yielding 240 Hz for oligomers of TMV-protein at pD 7.7 with molecular weight 0.5 to 1 x 10^6. The line width parameters at pD 7.7 can be compared with those for oligomers with molecular weight 52,500 at pD 11.4 and oligomers at pD 8.2 (see Table 3). From these results it is clear

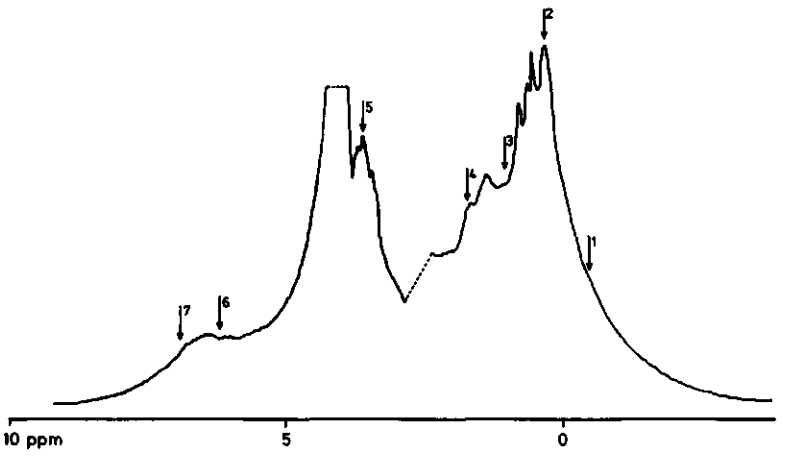


Fig. 3. 100 MHz ^{1}H NMR spectrum of oligomers of TMV coat protein in D_2O, with molecular weight between 0.5 and 1 x 10^4. Conditions: conc. 30 mg/ml, 0.2 M NaCl, temp. 5 °C, pD 11.4. No reference was used for the ppm scale. The spectral region at about 4.5 ppm is omitted because of the strong resonance of HDO. Δ_{10}, Δ_{30} and Δ_{50} are the line widths at 10%, 30% and 50% of the maximum height of the line shape of the aliphatic protons.
Acquisition time: 0.5 sec, number of accumulations: 10.000, sensitivity enhancement: 20 Hz.

Table 3. Variation of line width parameters with measuring frequency and molecular weight[a]

$\omega_H/2\pi$(MHz)	pD	T(°C)	Mol.wt.(x10^6)	Δ_{10}(Hz)[b]	Δ_{30}(Hz)[b]	Δ_{50}(Hz)[b]
360	7.6	34	0.5-1	1460	640	350
100	7.7	22	0.5-1	1180	460	240
60	7.7	34	0.5-1	980	440	280
100	8.2	5	0.05-0.5	1070	430	220
100	11.4	5	<0.1	930	360	220

[a] The accuracy of the line width parameters is within ± 10%.

[b] The line width parameters Δ_{10}, Δ_{30} and Δ_{50} are defined in Fig. 3. Spectra were run under equilibrium conditions for the protein oligomers. All protein concentrations are in the range 30 to 40 mg/ml in 0.2 M NaCl except for the sample recorded at 360 MHz which is in 0.1 M NaCl.

that the line width parameters are almost independent of a change in molecular weight. A similar behaviour of the line shape of the α-CH backbone resonances is found with variation of molecular weight, no accurate line width data are available however, because of partial overlap with the HDO and side chain resonances (see Fig. 3).

In the absence of internal motions and taking N = 1, Eq. [6] yields $1/\pi T_2$ = 40 Hz for a trimer with $\tau_R = 1.5 \times 10^{-8}$ sec, as calculated from Eq. [4]. For aliphatic protons an average value of N = 2 is assumed, resulting in $1/\pi T_2 \approx 80$ Hz. Taking into account the before mentioned assumptions in the calculation of τ_R and the uncertainty in the molecular weight of the protein oligomers at pD 11.4 (see Table 3), the agreement between the theoretical value of the line width $1/\pi T_2$ and the experimental value Δ_{50} (see Table 3) for the trimers is satisfactory. From the experimental results it can be concluded that each proton resonance of the backbone and side chain protons in oligomers of TMV protein has an upper limit of about 120 Hz for the line width $1/\pi T_2$ (with N = 1), independent of the molecular weight of the oligomers.

From the ratio of the D terms for ^{13}C and ^{1}H in Eq. [6] it can be calculated that the ratio of the line width $1/\pi T_2$ for ^{13}C and ^{1}H is approximately 0.5. This yields a calculated upper limit for the line width of a ^{13}C resonance of about 65 Hz (with N = 1). The chemical shift range of the overlapping resonances of the aliphatic carbons in the backbone and side chains is about 1.5 and 3.0 kHz, respectively (see Fig. 2). Thus $\Delta\sigma >> 1/\pi T_2$ and the line shape of the aliphatic carbons is mainly determined by chemical shift differences.

All ^{13}C and ^{1}H NMR measurements are performed in the pH range 7 to 10.8, where the integrated spectral intensity corresponds to the total number of observable nuclei (7). Line shapes can be determined with sufficient accuracy, since measurements are carried out by comparing the NMR spectra of disk- and trimer preparations of one and the same sample, the integrated intensity remaining the same. Therefore the possible effect of receiver dead time does not interfere with our measurements.

In the low pH range, where helical oligomers are formed (13), about 90% of the spectral intensity finally disappears, while the remaining ^{13}C or ^{1}H NMR spectrum seems to be unchanged. The implications of this effect are presently investigated.

3.4 DISCUSSION

3.4.1 Motion within the protein subunit

From ^1H T_1 measurements at 360 MHz presented above it is concluded that motions are present within the subunits of TMV-protein with correlation times smaller than 5×10^{-9} sec, calculated on the basis of Eq. [1] and assuming isotropic motions. It is impossible, however, that the rotational motion of the backbone and side chains within the subunit is isotropic, because of the definite secondary and tertiary structure of the protein (7). Therefore a model will be considered based on anisotropic internal motions in the presence of isotropic overall motion with a correlation time of 1.4×10^{-7} sec for the double-disk. Some models describing anisotropic internal rotational motions have been published (23-25). From the work of Woessner (23), T_1 is given by

$$1/T_1 = 0.6(\mu_o/4\pi)^2 N\gamma_H^4 \hbar^2 r_{HH}^{-6} \omega_H^{-2} [\tfrac{1}{4} \tau_R^{-1} + \tfrac{3}{4} \tau_i^{-1}], \qquad [7]$$

with

$$\tau_i^{-1} \equiv \tau_R^{-1} + \tau_g^{-1} \qquad [8]$$

assuming $\omega_H^2 \tau_R^2$ and $\omega_H^2 \tau_g^2 >> 1$. In Woessner's model the term τ_g is the rotational correlation time of a reorienting CH_3-group with random jumps about its symmetry axis. As is seen from Eqs. [7] and [8], for $\tau_R >> \tau_g$ the expression between brackets is given by $\frac{3}{4} \tau_g^{-1}$, so that T_1 is determined only by anisotropic internal motions. In the absence of internal motions the term between brackets is τ_R^{-1} and Eq. [7] is identical to Eq. [1]. Thus it can be seen that going from a model for isotropic motions to a model of anisotropic motions only, results in multiplying τ_R^{-1} in Eq. [1] by $\frac{3}{4}$ and taking $\tau_R = \tau_g$. For the double-disk with $\tau_R = 1.4 \times 10^{-7}$ sec and τ_g describing motions within the subunit which are faster than the calculated maximum correlation time of the protein subunit itself, corresponding to 5×10^{-9} sec, the approximation $\tau_R >> \tau_g$ is valid. This yields the result that in the case of anisotropic internal motion agreement can be obtained between the experimental value of T_1 at 360 MHz (see Table 1) and the T_1 given by Eq. [7] with $\tau_g \approx 1.6 \times 10^{-9}$ sec.

In a protein, random jumps over 3×120^o are possible only for protons in a limited number of CH_3-groups (24). For most protons of backbone and side chains these motions are likely to be restricted. Van Putte (26) has considered the

effect of restricted rotations in a model for rotational oscillations of methylene groups in a hydrocarbon chain. From this work it follows that T_1 is almost independent of the maximum angle of oscillation β_o for values larger than 40^o, whereas values of β_o smaller than 40^o strongly affect T_1. The value for β_o of 40^o is an upper limit, since smaller values of β_o, which increase T_1, can be compensated by smaller values of τ_g, which reduce T_1, provided $\omega_H^2 \tau_g^2 >> 1$. To see whether oscillation angles $\lesssim 40^o$ are reasonable for protons in a protein without irreversible distortion of the structure, an α-helix, constituting the most rigid part of the protein, is considered. Gō and Gō (27) have found that the dihedral angles in α-helices are fluctuating with an average amplitude of about 8^o. For an amino acid chain the total fluctuation angle will be larger, so that these kinds of motions may provide a mechanism, explaining the experimental values of T_1. Since the T_1 values of all protons in the TMV-protein are similar (see Table 1) this discussion applies to protons of backbone as well as side chains. It must be noted, however, that the resonances in the 360 MHz ^{1}H NMR spectrum overlap, so that the motional parameters derived from the experimental T_1 values will be an average.

In general it is found that by using the motional models of Woessner (23) and Van Putte (26) for ^{1}H, and Doddrell et al. (24) and Van Putte (26) for ^{13}C all other relaxation data in Table 2 and 3 can be explained semi-quantitatively in the same way as the 360 MHz ^{1}H relaxation data. Thus we conclude that the T_1 values of all nuclei in oligomers of TMV-protein are determined by fast restricted rotations with small amplitude within the subunits.

3.4.2 Motion of protein subunits in the oligomers

From ^{1}H 100 MHz experiments it is found that the line widths are independent of the molecular weight of the protein oligomers (Table 3). With the assumption of an isotropic motion, a rotational correlation time of 4×10^{-8} sec is calculated from the experimental line widths. The T_1 results discussed in the previous section, have led us to conclude to the presence of anisotropic motions of the nuclei within the subunits and for a correct description of the line widths these motions must be taken into account. It can be shown from the work of Woessner (23) that for $\omega_H^2 \tau_R^2 >> 1$ and $\tau_g << \tau_R$, T_2 is given by Eq. [6] with D reduced by a factor of 4. This reduction increases the discrepancy between the experimental value of Δ_{50} and the theoretical value of the line width by the same amount.

This line width is based on a lower limit of the rotational correlation time

τ_R of a trimer, while Δ_{50} is an estimated upper limit. Apart from experimental errors this suggests that the values of τ_R given by the Stokes-Einstein relation Eq. [4] are too small. This is in agreement with results obtained by Hallenga and Koenig (19), who have found that in the case that deviations from the Stokes-Einstein relation appear the experimental values for the correlation times are higher than calculated from this relation. There probably is another reason for the discrepancy noted before. So far, we have assumed the rotational model of Woessner (23) to be valid. In the case of restricted rotations (26) with $\beta_o \lesssim 40^o$ an increase of the theoretical line width is predicted also, reducing the discrepancy.

In the presence of anisotropic internal motions with $\omega_H^2 \tau_R^2 >> 1$ and $\tau_g << \tau_R$, the line widths depend on τ_R only. Since the line widths are independent of the molecular weight, τ_R is constant for all types of oligomers. This means that τ_R is not the correlation time for the overall rotational reorientation of these oligomers, such as a double disk. The independence of molecular weight of $(\pi T_2)^{-1}$ may be explained by assuming that motion simultaneously occurs within the subunit and of the subunit itself. This results in motions of C-H and H-H dipole-dipole vectors which are no longer sensitive to the oligomer reorientation in the solvent. This idea appears to be plausible because there are strong arguments favouring stability of protein subunit 3-D structure in solution (7), which is difficult to reconcile with large intra-subunit motions alone. More refined calculations are required to settle this problem. With the value of τ_R of 4×10^{-8} sec the assumption $\tau_g << \tau_R$ is valid, so that the discussion about T_1 still holds.

3.4.3 Intersubunit interactions

We must ask the question now, whether this can be understood from the thermodynamical point of view. Assemblage of TMV protein is an entropy-driven process (18). The free energy of association ΔG from trimers to double disks at 5 ^oC is -19.3 kJ.(mole trimer)$^{-1}$ with $\Delta H = +126$ kJ.mole^{-1} and $\Delta S = +519$ J.mole^{-1}.K^{-1} (18). For the association from monomer to trimer $\Delta G = 52.3$ kJ.(mole trimer)$^{-1}$ (18), while ΔH and ΔS also are positive (1). The high positive value of ΔS arises from the release of water from the protein surface upon association (18). From the high positive values of ΔH and ΔS, it can be concluded that no specific intersubunit interactions through hydrogen bonds and/or salt bridges are required for the association process. The driving force for association is the large entropy increase, which is counteracted by the positive ΔH.

The previous conclusion that our NMR data allow TMV subunits to have rotational freedom within the oligomers does not contradict the thermodynamic parameters. It merely indicates the absence of specific intersubunit interactions. In this paper we have not indicated in detail in which way subunit rotations may arise. We believe that in oligomers also translational degrees of freedom giving rise to effective subunit rotations may be present. Rotational and translational motions will increase the entropy and favour association.

3.5 ACKNOWLEDGEMENTS

The authors wish to thank Mrs. N.C.M. Alma-Zeestraten and Mr. L. Dorssers for helpful discussions and their valuable assistance with the NMR measurements. The technical assistance of Miss M. Boerjan with the preparation of TMV and the protein is gratefully acknowledged. We are indebted to the Department of Virology for hospitality and laboratory facilities. The NMR investigations at high magnetic field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

3.6 REFERENCES

1. D.L.D. Caspar, in "Adv. in Protein Chemistry 18, 37 (1963)".
2. G. Stubbs, S. Warren and K. Holmes, Nature 267, 216 (1977).
3. A.C.H. Durham, J. Mol. Biol., 67, 289 (1972).
4. A.C.H. Durham and J.T. Finch, J. Mol. Biol. 67, 307 (1972).
5. J.N. Champness, A.C. Bloomer, G. Bricogne, P.J.G. Butler and A. Klug, Nature, 259, 20 (1976).
6. K. Wüthrich, in "NMR in Biological Research: peptides and proteins", North Holland Publishing Company, Amsterdam, 1976.
7. J.L. de Wit, N.C.M. Alma-Zeestraten, M.A. Hemminga and T.J. Schaafsma, Biochemistry,submitted for publication,Chapter 2.
8. R. Lebermann, Virology, 30, 341 (1966).
9. L. Ornstein, Ann. N.Y. Acad. Sci., 121, 321 (1964).
10. W.R. Morrison, Anal. Biochemistry, 7, 218 (1964).
11. P.K. Glasoe and F.A. Long, J. Phys. Chem., 64, 188 (1960).
12. R. Freeman and H.D.W. Hill, J. Chem. Phys., 54, 3367 (1971).
13. S. Paglini and M.A. Lauffer, Biochemistry, 7, 1827 (1968).

14. A.C.H. Durham and A. Klug, J. Mol. Biol., 67, 315 (1972).
15. A.C.H. Durham, J.T. Finch and A. Klug, Nature, 229, 37 (1971).
16. N. Bloembergen, J. Chem. Phys., 27, 572 (1957).
17. R.A. Dwek, in "Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems", Clarendon, Oxford, 1973.
18. M.A. Lauffer, in "Entropy-Driven processes in Biology". Springer, New York, 1975.
19. K. Hallenga and S.H. Koenig, Biochemistry, 15, 4255 (1976).
20. Th. L. James, in "Nuclear Magnetic Resonance in Biochemistry", A.P. New York, 1975.
21. A. Kalk and H.J.C. Berendsen, J. Magn. Resonance, 24, 343 (1976).
22. E. Oldfield, R.S. Norton and A. Allerhand, J. Biol. Chem. 250, 6368 (1975).
23. D.E. Woessner, J. Chem. Phys., 36, 1 (1962).
24. D. Doddrell, V. Glushke and A. Allerhand, J. Chem. Phys., 56, 3683 (1972).
25. Y.K. Levine, P. Partington and G.C.K. Roberts, Mol. Phys., 25, 497 (1973).
26. K. van Putte, J. Magn. Resonance, 2, 23 (1970).
27. M. Gō and N. Gō, Biopolymers, 15, 1119 (1976).
28. T. Tao, Biopolymers, 8, 609 (1969).

4. A ^{13}C AND ^{1}H NMR STUDY ON ROD-LIKE POLYMERIZATION OF TOBACCO MOSAIC VIRUS PROTEIN

J.L. de Wit, T.J. Schaafsma
submitted to FEBS Lett.

4.1 INTRODUCTION

NMR studies in biochemistry have mainly been restricted to proteins with molecular weights $<\sim$ 50.000 (1). Due to limited resolution, the amount of structural information which can be obtained from such studies, even at 360 MHz, is rather restricted (2).

NMR can also be used for the study of molecular dynamical properties of biological systems with molecular weights > 10^6. The observation of spectral features itself, for example of Tobacco Mosaic Virus (TMV) and TMV-protein (3,4), then contains information about the dynamical behaviour of nuclei within a protein subunit in such a biological system expressed in translational and rotational degrees of freedom.

Even in TMV, with a molecular weight of 42 x 10^6, small numbers of ^{13}C nuclei of different types of amino acids are detectable with conventional Fourier Transform ^{13}C at 90.5 MHz because of their high degree of mobility (3). For large biological systems ^{1}H NMR is inferior to ^{13}C NMR because of the much lower spectral resolution and the presence of a residual HOD resonance, even in 99,99% D_2O solutions. The sensitivity of natural abundance ^{13}C NMR is much lower than that for ^{1}H NMR, however. Therefore we have used 12% enriched TMV (3,5) in order to obtain ^{13}C spectra with sufficient signal to noise ratio.

This report is an extension of our earlier work, presenting the pH- and temperature-dependence of the rod-like polymerization in more detail.

4.2 MATERIALS AND METHODS

4.2.1 Tobacco Mosaic Virus

TMV was purified according to Leberman (6). TMV-protein was prepared following Durham (7) with slight modifications (4). All protein solutions used in our experiments had an $E_{280/252}$ ratio > 2.5 in 0.12 M-Tris HCl, pH 8.6 and showed a single boundary (3.8 S) in the analytical ultracentrifuge and a single band with polyacrylamide gel electrophoresis.

Analytical ultracentrifuge experiments were carried out using a Spinco model E ultracentrifuge equipped with Schlierenoptics, automatic speed and variable temperature control. S values are uncorrected for concentration, viscosity and temperature and have been determined with sedimentation velocity experiments at 22,000 or 40,000 rev./min.

TMV (strain Vulgare) was enriched up to 12%, using $^{13}CO_2$ as carbon source for photosynthesis in leaves of Nicotiana tabacum L. (Samsun NN'], inoculated with TMV. The enrichment procedure will be described elsewhere (8). Enrichment up to 12% represents an optimum at which no noticeable carbon carbon coupling is present (9]. $^{13}CO_2$ (90% enriched) and D_2O (99,9% enriched), were purchased from Bio Rad, NaOD (99% enriched) and DCl (> 99% enriched) were obtained from Merck. In D_2O solutions the normal pH meter readings have been taken (11).

4.2.2 Nuclear Magnetic Resonance

For ^{13}C measurements a Bruker SPX 360 supercon spectrometer in FT mode with quadrature detection was used, employing ∿ 5 W continuous wave ^1H decoupling power with ∿ 5 kHz bandwidth, an observe frequency of 90.5 MHz, and 10 mm sample tubes. A ^1H lock turned out to be unnecessary for our experiments. ^1H measurements were carried out on a Varian XL 100 NMR spectrometer in FT mode with an external ^{19}F lock, an observe frequency of 100 MHz and 5 mm tubes. On both NMR systems constant temperature was maintained with a modified temperature control system allowing long time averaging experiments (4). The receiver dead times of both spectrometers was ∿ 50 µsec.

4.3 RESULTS

4.3.1 pH-induced rod-like polymerization

The sedimentation coefficient of the protein polymers and oligomers at 7°C, pH 5.3 and 10.0 are > 40 S and 2,9 S respectively. In the sedimentation velocity experiments at pH 5.3 a distribution of polymers with sedimentation coefficients 40, 70, 80 and 90 S is found at 22,000 rev./min. Sedimentation coefficients < 40 S are absent at 40,000 rev./min. At pH 6.3 and 7°C, oligomers with sedimentation coefficients 15-18 S are present; below this pH and above this temperature oligomers with S < 15 have not been observed.

Fig. 1 represents a ^{13}C spectrum of TMV-protein at 7°C and pH 10.0 in 0.1 M

NaCl. The change in relative spectral intensities (measured by weighing the

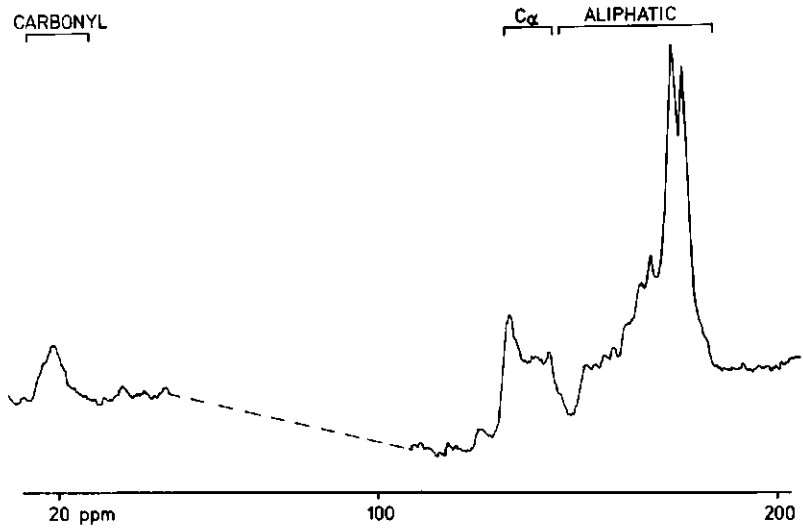


Fig. 1. Broadband ^{1}H decoupled 90.5 MHz ^{13}C NMR spectrum of TMV-protein taken at 7°C with a concentration of 30 mg/ml (1 ml total volume) and 17,000 accumulations; acquisition time: 0.5 s; sensitivity enhancement: 30 Hz. The ppm scale is referenced to CS_2 assuming 125.8 ppm for the β-Thr carbon position. Further conditions: 0.1 M NaCl, pH 10.00. The aromatic region is omitted because of an interfering instrumental effect.

carbonyl, α-C and aliphatic regions) as a function of pH is shown in Fig. 2. The spectral regions of which intensities have been compared are indicated. Within measuring error this change is identical for the different spectral regions.

Fig. 3 represents the drop in intensity for the aliphatic region of TMV-protein in D_2O solution, upon rod-like protein polymerization as detected with ^{1}H NMR. Although in D_2O the polymerization behaviour of TMV-protein is considerably different from that in H_2O (10) the rod-like polymerization may also be expected to be a linear condensation polymerization process. As has previously been shown (4) the ^{1}H line shape at 100 MHz is much more sensitive to line broadening effects than the ^{13}C line-shape. Upon rod-like polymerization again no broadening of ^{1}H line shapes is observed within experimental error. The absence of line broadening in the ^{1}H and ^{13}C spectra can not result from the receiver dead time of the NMR spectrometers, because the decrease in spectral

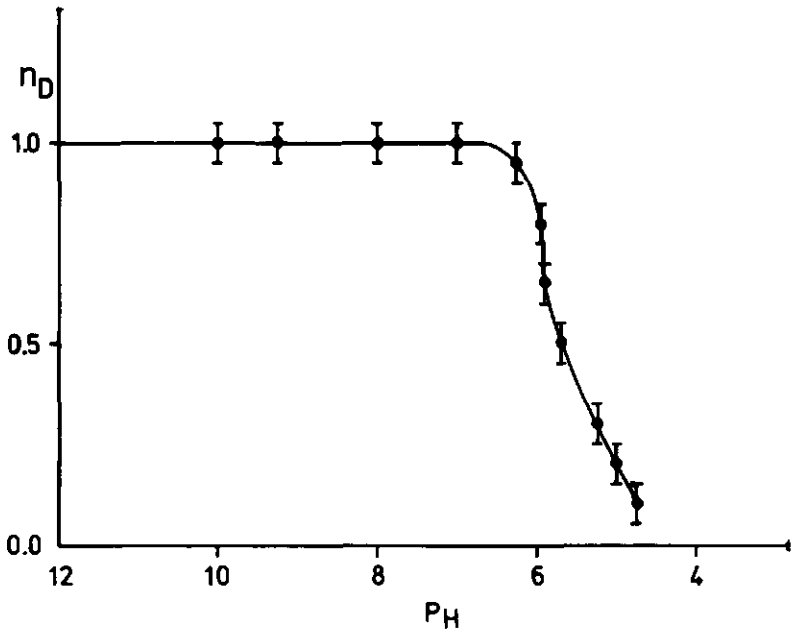


Fig. 2. The ^{13}C intensity of TMV-protein for different spectral regions as a function of pH. The alphanumerically labeled regions are indicated with arrows in fig. 1. Conditions: 7°C, 0.1 M NaCl, 15 mM sodiumphosphate and a protein concentration of 30 mg/ml. The NMR parameters are identical to those given in the caption of fig. 1. n_D is the detectable NMR intensity. The error in n_D arises from the presence of spectral noise.

intensity is identical for all spectral regions, independent of the line width close to the base line of a group of overlapping resonances. Absence of polymerization-induced line broadening therefore, cannot be due to a disappearance of spectral intensity through *gradual* line broadening beyond the limits set by the spectrometer.

The spectral intensity decrease upon lowering pH from 10 to 5.9 cannot be explained by a combination of limited acquisition time and increasing T_1 value, since spectra recorded with acquisition times 0.5 and 1 sec exhibit identical decrease of intensity with decreasing pH for the aliphatic region. Moreover, the intensity drop for non-protonated carbonyl carbons in ^{13}C NMR is comparable to that of the protonated carbons in the aliphatic region, despite the fact that the former have much longer T_1's than the acquisition times employed in these experiments (4).

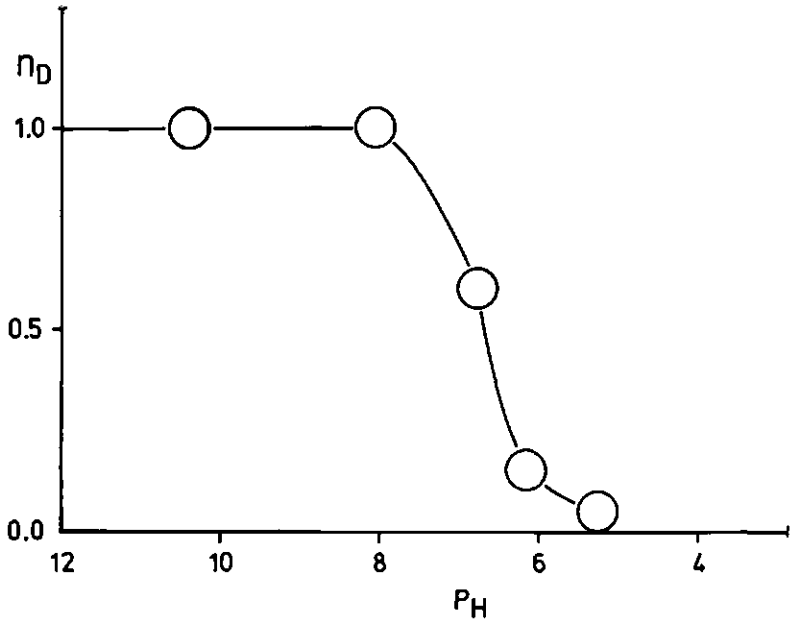


Fig. 3. The ^1H spectral intensity of TMV-protein of the aliphatic region as a function of pD. Conditions are; 5°C, 0.2 M NaCl and a concentration of ∿ 40 mg/ml in D_2O. The NMR parameters are: spectral width: 5000 Hz; accumulations: 10,000; sensitivity enhancement: 0.05 Hz and acquisition time: 0.5 s. Each graphical point represents the intensity of the largest group of resonances in the aliphatic region. n_D is the detectable NMR intensity. Experimental errors arise from the use of noise levels in the calculation of spectral intensities and from the fact that the pH values were measured outside the 5 mm NMR tube.

Figs. 2, 3 and the analytical ultracentrifuge data therefore provide evidence that the pH-dependent intensity of the observable part of the ^1H and ^{13}C spectra is due to those nuclei within the rod-like polymers which are mobile enough to eliminate important line broadening (3,4).

4.3.2 Temperature-induced rod-like polymerization

At pH values close to 6 rod-like polymerization can also be induced by increasing temperature (12). The ^{13}C spectral intensity change observed in the presence of such a polymerization process at different pH values and temperatures is presented in Table 1 for carbonyl carbons only; identical results are found

for the α-C and aliphatic regions. Since at high temperature all sedimentation coefficients (Table 1) are > 30 S (corresponding to a molecular weight of > 10^6 (15)) spectra should be rigorously broadened (3) in the absence of internal motion.

Table 1. ^{13}C spectral intensity for the carbonyl region[a] at different values for pH and temperature[b].

pH	7.2	6.3	6.1	6.0	5.9	5.6	5.3
0°C		1					
7°C	1		1	0.8	0.6	0.5	0.3[c]
15°C	1	1	0.8				
20°C	1	0.7[c]		0.5[c]	0.3[c]		
25°C	1	0.7[c]	0.6[c]				
30°C	1	0.7[c]			0.3[c]	0.3[c]	

[a] the carbonyl region is shown in fig. 1.
other conditions: conc. 30 mg/ml, 15 mM sodiumphosphate, 0.1 M NaCl.

[b] the accuracy of the measured spectral intensity is ± 0.05 relative units as determined from the spectral noise level.

[c] only components present with sedimentation coefficients > 30 S.

A comparison of the effect of temperature on the spectral intensity (Table 1) of large molecular weight polymers (> 30 S) in the range 20 - 30°C below pH 6.3 with the temperature effect on titration curves obtained by Butler et. al. (15) demonstrates that the absence of proton uptake observed by these authors upon increasing temperature in this range, is also reflected by the constant ^{13}C spectral intensity. The behaviour at pH 7.2 originates from polymerization to double disks and has been previously discussed (4). Similar results are obtained for ^1H NMR of TMV-protein in D_2O solutions. Both for ^{13}C and ^1H spectra no line broadening upon temperature-induced polymerization is observed. Thus, the observed resonances again must arise from internally mobile polymers, such that no important line broadening is observed upon increasing molecular weight.

4.4 DISCUSSION

Similar to what has been found for double-disk like polymerization (4), the

exchange rate τ_m^{-1} of protein subunits between rod-like polymers is too small to account for the molecular-weight-independent NMR spectra (3). For the present case, the Gibbs free energy of protein subunit binding within a polymer is $\Delta G = -76$ kJ (13), resulting in $\tau_m^{-1} < .35\ sec^{-1}$ (4,14), where τ_m is the residence time of a protein subunit within the oligomer.

We may assume that the secondary and tertiary structure of the protein subunits within the polymer in solution must be largely maintained. Therefore, it has previously been suggested (3,4) that within protein oligomers motional degrees of freedom of amino acid side chains and backbone (within each protein subunit) and of protein subunits (within each protein oligomer) must be considered.

Up to pH 5.3 (Fig. 1), spectra only differ from the high pH spectrum by a scaling factor within experimental error. Therefore, we propose that the main cause for the disappearance of spectral intensity upon pH-induced rod-like polymerization must reside in the loss of motional degrees of freedom of all nuclei in a protein subunit within the rod-like polymer. In this respect, our results do not allow to distinguish between motion within or of the protein subunits. Since no line broadening is observed upon lowering the pH, this loss of mobility must be abrupt, so that all carbon resonances of a protein subunit are either observed or have been broadened beyond detection. Similar to what has been found for double disk-like polymerization (4), the absence of line broadening upon rod-like polymerization in the pH range between 5.3 and 6.3 at temperatures > 20°C shows that the dynamics of the observable nuclei is independent of polymer molecular weight. Fig. 2 and Table 1 indicate that the change in spectral intensity is pH-controlled.

A two state model can describe this behaviour where in state N the nuclei in the protein subunits are "locked" and non-detectable and in state D the nuclei in the subunits have sufficient motional degrees of freedom to be detectable. The number of protein subunits in state D (n_D) depends on pH and is in equilibrium with the number of protein subunits in state N (n_N) according to the equation

$$n_D + xH^+ \rightleftharpoons n_N \qquad [1]$$

where x is the number of protons which must bind to induce a subunit transition from state D to N.

By plotting log $|n_D/(1-n_D)|$ against pH, the pK and x of eqn. [1] can be determined. From the resulting straight line, $x = 1.25 \pm 0.05$ and a pK of 7.0 ± 0.1 can be derived. Note that because $x = 1.25$ this pK does not simply represent the pK of the titratable groups within the protein subunits.

In agreement with other authors (16,17) Butler et.al. have shown (15) that there is an uptake of $\sim$ 3 protons/subunit, in particular by two titratable groups with raised pK's between 5.6 and 7.1 when TMV-protein is titrated from pH 8.5 to 5 over a wide temperature range and 0.1 ionic strength. An uptake of $\sim$ 1.3 protons/subunit is observed when temperature is increased from 5 to 20°C (15) at pH 6.3, 0.1 ionic strength. No further proton uptake is observed upon temperature increase to 30°C. Using a pK $\sim$ 6.6 (15) the fractions of protein subunits with zero, one and two protons added can be calculated to be .10, .26, and .64 resulting in an average uptake of 1.5 protons/subunit. Table 1 shows that the observable spectral intensity of 0.7 remains constant when temperature is further increased from 20 to 30°C. In view of the foregoing experimental results, the intensity of 0.7 shows that the addition of one proton has no effect on NMR intensity, in agreement with the calculated value of $x > 1$; the $n_D \rightarrow n_N$ transition, therefore, cannot be induced by the addition of the first of the $\sim$ 3 protons which is taken up when titrating from pH 8.5 to 5.0.

The pK's of the titratable groups cannot be deduced directly from our NMR results although it is clear that they must be between 5.6 and 7.0.

The two state model presented above, does not give a complete description of the titration behaviour of titratable nuclei in TMV subunits. At pH = 4.8 still $\sim$ 10% of the ^{13}C nuclei is observable which is about 10 times more than predicted from the model presented above. This remaining intensity may well originate from internal mobility, similar to that found for TMV (3). In both cases the α-C resonances have strongly diminished intensity.

4.5 ACKNOWLEDGEMENTS

The technical assistance of Miss M. Boerjan is gratefully acknowledged. We are indebted to the Department of Virology for providing laboratory facilities. The NMR investigations at high field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

4.6 REFERENCES

1. NMR in Biology (1977) ed. by Dwek, R.A., Campbell, I.D., Richards, R.E. and and Williams, R.J.P., Academic Press, London.
2. Wüthrich, K. and Wagner, G. (1975) FEBS Lett. 50, 265-268.

3. De Wit, J.L., Alma, N.C.M., Hemminga, M.A. and Schaafsma, T.J., Biochemistry, submitted for publication, Chapter 2.
4. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J. (1978) J. Magn. Resonance, in press, Chapter 3.
5. Egan, W., Shindo, H. and Cohen, J.S. (1977) Ann. Rev. Biophys. Bioeng. 6, 383-417.
6. Leberman, R. (1966) Virology 30, 341-347.
7. Durham, A.C.H. (1972) J. Mol. Biol. 67, 289-305.
8. De Wit, J.L., Alma, N.C.M. and Schaafsma, T.J., Biochim, Biophys. Acta, submitted for publication, Chapter 6.
9. Wüthrich, K. (1976) NMR in Biological research: peptides and proteins, North Holland publishing Company, Amsterdam.
10. Paglini, S. and Lauffer, M.A. (1968) Biochemistry 7, 1827-1834.
11. Glasoe, P.K. and Long, F.A. (1960) J. Phys. Chem. 64, 188-190.
12. Caspar, D.L.D. (1963) Adv. in Protein. Chem. 18, 37-121.
13. Lauffer, M.A. Entropy-driven processes in Biology (1975) Springer-Verlag, Berlin.
14. Dwek, R.A. (2975) Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems, p. 178, Clarendon Press, Oxford.
15. Butler, P.J.G., Durham, A.C.H. and Klug, A. (1972) J. Mol. Biol. 72, 1-18.
16. Durham, A.C.H., Vogel, D. and De Marcillac, G.D. (1977) Eur. J. Biochem. 79, 151-159.
17. Scheele, R.B. and Lauffer, M.A. (1967) Biochemistry 6, 3076-3081.

5. MOLECULAR DYNAMICS OF TMV AND ITS PROTEIN OLIGOMERS STUDIED BY ^{13}C NMR

J.L. de Wit, T.J. Schaafsma
submitted to Eur. J. Biochem.

5.1 INTRODUCTION

^{13}C NMR spectra of 12% ^{13}C enriched Tobacco Mosaic Virus (TMV) and its protein oligomers have been reported previously (1,3). This rod-like plant virus belongs to the group of very large biomolecules and has a molecular weight of $\sim 42 \times 10^6$, 2200 identical protein subunits (molecular weight 17,500) and one RNA chain of 6600 nucleotides, so that using conventional high resolution Fourier Transform NMR, no signals are expected (2).

From spin-lattice relaxation times (T_1) and line shape analysis of the double disk- and rod-like oligomers and polymers, qualitative information has been obtained about the molecular dynamics of protein subunits and amino acid backbone and side chains within the protein subunits in the protein oligomers and polymers (1,3,4).

In the rod-like polymer molecular dynamics is sensitive to protonation of presumably carboxyl groups with anomalous pK values (4,5).

In this paper a more quantitative description of the molecular dynamics of protein subunits and amino acid chain (158 residues) is presented, based on Convolution Difference (CD) spectra (6) of virus and CD spectra and T_1 data of the rod-like polymer (4).

5.2 RESULTS

5.2.1 CD spectrum of TMV

In the virus 6% of the α-carbons and 17% of the aliphatic carbons are observable by normal FT NMR, because of short T_1 ($\lesssim$ 1 s) and long T_2 (1). In Fig. 1 the CD spectrum of 12% enriched TMV (7) is given. The TMV CD spectrum contains $\sim$ 80% of the observable α-carbons and $\sim$ 20% of the observable aliphatic carbons. These numbers have been obtained by eliminating all spectral components with line width > 2000 Hz (the line width at which in our case the CD intensity is reduced to 1%) from the normal ^{13}C virus spectrum (1) and comparing the integrated spectral intensity for the various regions of the resulting spectrum

Table 1. Assignments in observable and unobservable resonances[a] in a CD spectrum of TMV[b]

amino acid	α-C	β-C	γ-C	δ-C	ε-C
Phe	+[c]	+		+	+
Thr	+	+	+		
Ser	+[c]	+			
Ala	+	+			
Arg	+	+	+	+	
Lys	-	-	-		
Gln	+[e]	+	+		
Val	-	-	-		
Ile	-		+	+	
Leu	-	-	-	-	
Pro	-	-	-	-	
Glu	-	-	-		
Asp	+[e]	+			
Asn	-	-			

[a] Observable resonances are indicated with +; non-observable resonances with -; the total number of observed α-carbons is between 8 and 12. Trp and Tyr are unobservable. A number of carbons are omitted because of uncertainty in the assignments. The titration behaviour of the carbons of the amino acids in the pH range 5-10 was taken into account (9,10,11) in the assignments.

[b] Assignments have been made adopting spectral positions from James (8).

[c] The resonances at 137.7 ppm (Fig. 1), can be assigned to the sum of α-carbon resonances of Ser and Phe.

[e] Region 138.9 to 142.4 ppm, containing unresolved α-carbon resonances.

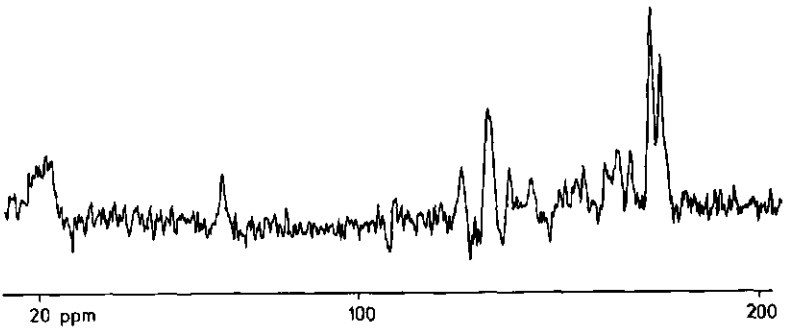


Fig. 1. ^{13}C convolution difference spectrum of a 60 mg/ml TMV solution in 1 mM sodiumphosphate pH 7.2, obtained from a ^{13}C spectrum (1) taken at 30°C under broadband ^{1}H decoupling with a 1 s acquisition time, and 50,000 accumulations.
The time constants used in obtaining the CD spectrum (6) are 0.005 and 0.011 s with K = 1, so that the two spectra are weighted equally. The ppm scale refers to CS_2 assuming 125.8 for the β-Thr carbon position.

with that of the normal spectrum. The assignment of the resolved peaks to individual amino acid resonances in this CD spectrum is summarized in Table 1.

The spectral positions used to obtain the data shown in Table 1, were based on computer simulations of ^{13}C spectra of urea-denatured TMV protein (7) and are similar to those published (8). Errors in the assignments because of resonance shifts may arise from pH effects and folding of the protein (9-12). Note that the original line widths of all observable resonances of the virus spectrum are < 300 Hz (1,3).

In the ^{13}C spectrum of TMV only resonances originating from nuclei with internal mobility are present. If the α-carbon of a particular amino acid is observable, the β-carbons and carbons further down the side chain also are observable. On the other hand if a CH_3 or CH_2 carbon is observable the α-carbon can still be absent from the spectrum.

In the region 50 to 100 ppm only two ε- and φ-carbons of phenylalanine are observed; δ-carbons of leucine (171 ppm) and proline (145.5 ppm) and γ-carbons of glutamic acid (159 ppm) are absent in the aliphatic region, indicating that no other carbons of these amino acids will be observable as shown in Table 1. Absence of these resonances can be interpreted to mean that the corresponding carbons are less mobile than those carbons of which resonances are present in the

CD spectrum. This is nicely illustrated when the virus CD spectrum at 30°C is compared with the CD spectrum of the double disk-like oligomer at 30°C in which for example δ-leucine carbons do occur. Previously, it was concluded that the ∿ 10 observed α-carbons in the normal virus spectrum (1) represent two α-threonine and ∿ 8 other amino acids. Because the observable CH_3, CH_2 and CH side chain carbons exceed the connected α-carbons with a factor of about 4, more than ten amino acids have mobile side chains (1).

5.2.1 Rod-like polymerization of TMV-protein

Rod-like polymerization of TMV-protein can be induced by lowering pH and/or increasing temperature at pH values below 7 (14,17). ^{13}C spectra taken at various stages of this polymerization process have been presented before (1,4). The CD spectra as a function of pH are given in Fig. 2. The spectral regions at which relatively large changes as a function of pH are observed have been indicated with arrows. At 7°C, pH 10.0 and 7.0 protein oligomers with sedimentation coefficients 2.9 and 4 S are found, respectively, with molecular weights $\lesssim$ 52,000 (1,15). The predicted upper limit for the ^{13}C line width in these oligomers with one proton attached to it, is 65 Hz (3). At 7°C, pH 5.3, sedimentation coefficients > 30 S are observed reflecting protein polymers with molecular weights > 10^7 (4).

The CD spectra (Fig. 2) have been analyzed by assuming that the most obvious shifts in a spectral region resulting from a change in pH can be attributed to carbons which titrate in that pH range. Peak 1 in Fig. 2A (pH 10.0) at 151.7 ppm shifts upfield to ∿ 152.5 ppm (Fig. 2B) and downfield upon a further decrease in pH to 151.7 ppm (Fig. 2C). The upfield shift of peak 1 from 151.7 to 152.5 ppm upon decreasing pH from 10.0 to 7.0 can be attributed to at least one δ-arginine carbon, due to the titration of an arginine amino side chain group in this pH range (9,10,11). In this respect the resonance position at pH 5.3 is unexpected.

A comparison of the position of peak 1 in Fig. 2A and C (pH 10.0 and 5.3) shows that the ionization state of the amino group of an arginine must be identical at these pH values, because of equal resonance positions and peak envelope, which differ from those observed in the CD spectrum at pH 7.0 (Fig. 2B). The arginine shift is upfield with pH decreasing from 10.0 to 7.0, implying that at high pH the ionization state must be $-NH_2$, whereas at pH 7.0 this must be $-NH_3^+$ (9,10,11). Peak 2, at 153.7 ppm (Fig. 2A) may be hidden under peak 1 in Fig. 2B because no titratable carbons at this spectral position in this pH range are present (8). Upon titrating to pH 5.3 peak 2 shifts upfield from 153.7 to

155 ppm (8). It is generally accepted that TMV-protein contains two carboxyl groups with anomalous pK values (5). Because the β-carbon of aspartic acid titrates in this region at ∿ 154 ppm, the upfield shift of peak 2 in the pH range 7.0 to 5.3 can be assigned to this amino acid. The origin of peak 3 at ∿ 165 ppm

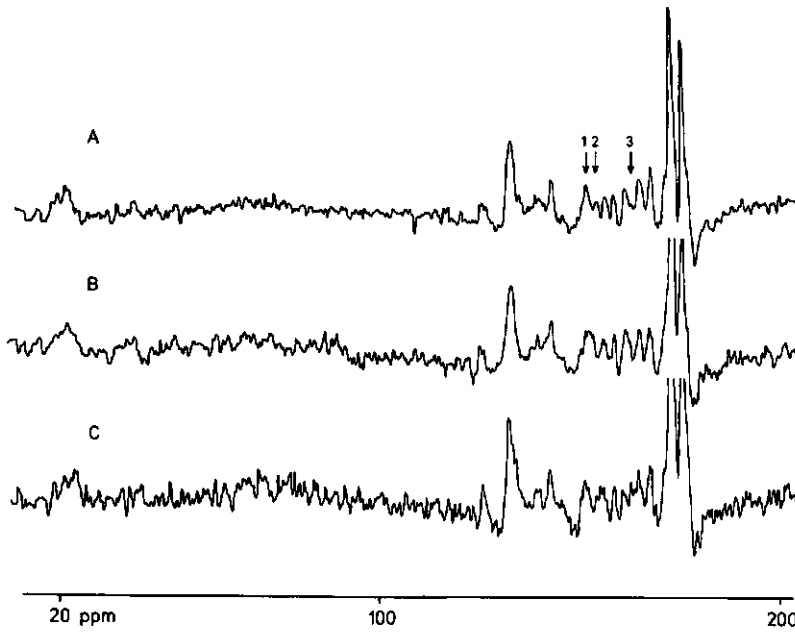


Fig. 2. ^{13}C convolution difference spectra of a 40 mg/ml TMV-protein solution in 0.1 M NaCl obtained from ^{13}C spectra taken at 7°C under broadband ^{1}H decoupling with a 0.5 s acquisition time and 17,000 accumulations. Other conditions are: A, pH 10.0; B, pH 7.0; C, pH 5.3.
The arrows indicate spectral regions with largest changes. The time constants used in obtaining the CD spectra (6) are 0.005 and 0.011 s with K = 1 so that the two spectra are weighted equally. The ppm scale refers to CS_2 assuming 125.8 for the β-Thr carbon position.

(Fig. 2C) is more difficult to assign, but since spectral positions of β-carbons of arginine and glutamic acid are in this region and an arginine is known to titrate at this pH, the presence and position of peak 3 may result from an interplay between pH-induced shifts of β-arginine and β-glutamic acid carbons. Other smaller spectral changes (Fig. 2) are within noise level and for this reason will not be discussed. Finally we note that the spectrum at pH 5.3 represents the detectable part of the carbons within the rod-like protein polymer (1,4).

In contrast with the result for double-disk like oligomers, at least two T_1's can be found for rod-like oligomers at pH 6, 7 and 30°C for ^{13}C at 90.5 MHz, the largest fraction having $T_1 \gtrsim 0.4$ s, whereas the smaller fraction has T_1 in the order of 1 s. The T_1 values are similar for all spectral regions and are ascribed to the presence of motional degrees of freedom of carbons within the protein subunits as described before (1,3,4).

5.3 DISCUSSION

5.3.1 Interpretation of T_1 data of double disk- and rod-like oligomers

T_1 data of approximately all aliphatic protons and carbons in TMV protein oligomers cannot satisfactory be interpreted using a correlation time τ_R (3,16) for oligomer reorientation in the T_1 equation for dipole-dipole relaxation. Therefore it has been proposed that the nuclei within the protein subunits are mobile (1,3). A relation between T_1 and one correlation time τ_g describing one motional degree of freedom about a carbon-carbon bond for ^{13}C, or about a proton-proton axis for ^1H has been given by Doddrell et al. (16) and Woessner (3) respectively (19).

In Table 2, T_1 normalized to interaction of an observable nucleus with one proton (N = 1) for the double disk-like oligomer, is presented at 100 and 360 MHz for ^1H and 25.2 and 90.5 MHz for ^{13}C. Table 2 also includes ranges of values of τ_g calculated for dipolar relaxation of ^{13}C (1,16) and ^1H (1,3,19). At each frequency two ranges of τ_g can be calculated for each T_1 (Table 2). The τ_g range between 2×10^{-11} and 1×10^{-10} s is assumed to be relevant. The choice of this range is supported by the following arguments: CH_3 groups have been found to rotate on this time scale in proteins (21); NOE effect on CH_3 groups at e.g. 90.5 MHz can only be explained using $\tau_g \lesssim 10^{-10}$ s (16); if at 360 MHz the slower τ_g range would apply, T_1 at MHz should be ~ 4 times smaller than actually observed. It is not necessary that τ_g describes full 2π rotation about a carbon-carbon axis (3); small angle (β_o) oscillations (3,18) yield similar T_1's without

Table 2. Spin-lattice relaxation times of the CH_3, and CH_2 groups in the double disk-like oligomer for ^{13}C at 25.2 and 90.5 MHz, and ^{1}H at 100 and 360 MHz[a].

Frequency (MHz)	25.2[c]	90.5[b]	100[c]	360[b]
T_1[d] (s)	0.3-0.4	0.5-0.8	0.5-0.7	0.7-1.4
τ_g (x 10^{-10} s)	<0.1	0.8-1.8	1.2-1.8	0.1-0.2
τ_g (x 10^{-10} s)	0.4-0.6	0.2-0.4	0.6-1.0	0.4-0.8

[a] For T_1 data and measuring conditions, see ref. 1,4. τ_g for ^{1}H data has been obtained from Woessner (19) and for ^{13}C data from Doddrell et al. (16) with $\tau_R = 10^{-7}$ s.

[b] T_1 determined by progressive saturation (27).

[c] T_1 determined by inversion recovery (27).

[d] T_1 scaled to one attached proton (1,3). The accuracy of the T_1 measurements is $\pm$ 10%, mainly due to the presence of oxygen (8) and scaling procedure (1,3).

changing τ_g as long as the rotation angle $\beta_o > 40^o$. The similarity in T_1 for backbone protons, α-carbons and amino acid side chain protons and carbons (3) at frequencies presented in Table 2 indicates that in the cases that β_o is identical for each nucleus, T_1 results in similar τ_g values for all observable nuclei.

A calculation of T_1, based on rotation about a single carbon-carbon bond according to the model of Doddrell et al. (1,3,16) does not reflect the real situation. The fact that all observable nuclei have a short T_1 implies that the motion, and thus the relaxation behaviour of each of these observable nuclei within the protein subunit is controlled by more than one τ_g parameter. The analysis of our measurements using such a model does not yield more relevant information, however, because the number of variables then exceeds the number of measurements. The calculated range for τ_g can be considered as a frequency-dependent weighted average (20) over several different τ_g^i's and rotation angles β_o^i (with i indicating rotation about a carbon-carbon bond), similar for all carbons and protons. Besides, from the work of Levine et al. (20), it can be shown that in the case of several τ_g^i's and assuming $\omega_H^2 \tau_R^2 > 1$, T_1 is frequency independent if $\omega_H^2(\tau_g^i)^2 < 1$ where ω_H is the measuring frequency.

With ω_H = 360 MHz, $\tau_g^i < 5 \times 10^{-10}$ s is calculated. This value can now be used as an upper limit because T_1 dependence on frequency is small (Table 2). Note that this upper limit is independent from the rotation angle β_o when $\beta_o > 40^o$ (18).

In view of the arguments published before (3,4) and presented above for the double disk-like oligomer it can be assumed that the behaviour of T_1 found for rod-like oligomers is also controlled by motions within a protein subunit. The suggested similarity of T_1 behaviour for the different spectral regions can be interpreted as indicated above. The spread in time constants could arise from different sections within a protein subunit or from differences in conformation of protein subunits. The first possibility cannot be ruled out, but other data, to be discussed below, favour the second suggestion.

5.3.2 Protein subunit mobility within the double-disk- and rod-like oligomers

NMR on double-disk and rod-like polymerization revealed within measuring error a molecular weight-(and thus τ_R) independent line shape behaviour. From the line width a phenomenological correlation time τ_p of $\sim 4 \times 10^{-8}$ s, describing this molecular weight independent line shape behaviour is calculated (1,3,4). Because the amino acid chain in the protein subunits within the oligomers in solution (1) is folded into a three dimensional structure (1,7), rotational motions of all nuclei within the protein subunits over large angles ($\sim 2\pi$) must be excluded (3). As shown before, motions within the protein subunit can be characterized by $\tau_g < 5 \times 10^{-10}$ s and rotation angles $\beta_o \sim 40^o$. Evidently, τ_p differs from this τ_g and τ_R, previously defined to describe motion of the oligomer. Can we give a physical meaning to the type of motion characterized by τ_p?

There are good reasons to believe that protein subunit motions within a protein oligomers of TMV exist and contribute to τ_p. Levine et al. (20) have found that for a carbon chain with full 2π rotation about carbon-carbon bonds and provided $\tau_g < \tau_R$, the relaxation of a particular carbon is τ_R-independent if there are at least five bonds between this carbon and a carbon fixed to the molecular frame (in our case: the oligomer). For the double disk-like oligomers τ_R is calculated to be $> 10^{-7}$ s (3). Even if we assume rotation about one carbon-carbon bond of the protein backbone, an unfolding of the α-helical regions, constituting > 30% of the protein backbone, would result, which is in disagreement with ORD data (25). This implies that large angle rotation about carbon-carbon bonds in the α-helix backbones cannot occur, so that mobility of the protein subunit itself must contribute to τ_p.

A τ_R-independent line shape implies a decoupling of the motion of a particular C-H dipolar vector within a protein subunit from the motion of the oligomer. This can be realized if the protein subunit motion is described by rotation diffusion around three mutually perpendicular axes (22,19) on a time scale given by the calculated value (4) $\tau_p \sim 4 \times 10^{-8}$. This model may be somewhat exaggerated, however. From the work of Levine et al. (20) it follows, that the presence of two degrees of freedom for rotation of the subunit weakens the τ_R-dependence of the line width with respect to the case where only one rotational degree of freedom is present. Noting that small angle ($\sim 40^{\circ}$) rotations are assumed to be present in the backbone, it is conceivable that the process characterized by τ_p is the result of the combined action of backbone mobility due to rotation about carbon-carbon bonds and subunit motion involving rotational motions about less than three (and probably two) mutually perpendicular axes; one of these rotational motions may be the result of a displacement of the subunit about the oligomer symmetry axis.

5.3.3 Mobility within TMV

Table 1 can be used to assign mobility to certain sections of the amino acid chain in protein subunits in TMV, if certain assumptions about the presence of rotational motions about carbon-carbon bonds within these protein subunits and about protein structure stability are made: according to Levine et al. (20) full 2π rotational motions about at least two carbon-carbon bonds can be assumed to interpret the linewidth of 300 Hz (1); all CH_3 groups exhibit fast rotation ($< 5 \times 10^{-10}$ s) because steric hindrance is weak or absent (21); the published molecular model showing distinct secondary and tertiary structure (23) in TMV is applicable to locate the α-helical sections; these sections are most stable and are assumed to have least internal mobility; if the α- or β-carbon of a residue is observed the remaining carbons of this amino acid are also observed; due to RNA binding, mobility can be excluded for amino acid residues $\sim$ 90 to 120 because they constitute the RNA binding site; the chemical shift inequivalence of amino acid side chain carbons, attached to observable α-carbons, is considerably reduced; motions of protein subunits within the virus can be excluded due to RNA binding (14) and in view of the distribution of carbon resonances in this virus spectra, deviating from the intensity-distribution of carbon resonances in the spectra of double disk- and rod-like oligomers, in particular w.r.t. the lack of C_α resonances in the virus spectrum. Furthermore, we note that the

number of mobile carbons within the protein subunit is at a minimum if all observable α-carbons (∿ 10) belong to one section of the amino acid chain. The analysis presented below is based on this minimum mobility approach. Turning now to Table 1, we observe that proline, tryptophane and tyrosine are absent in the CD spectrum (Fig. 1) so that the observed α-carbon can only be partly located in the last five amino acids of the C- and N-terminal (24) sections. Phenylalanines in non α-helix sections are Phe 10, 12, 35, 62, 67, or 144. Because the α-carbon of one phenylalanine is also observed, this amino acid must belong to a mobile section of the amino acid chain. Side chain carbons of leucine, lysine and glutamic acid which are connected to phenylalanines are not observed so that the observed phenylalanine is unlikely to be Phe 10, 12, 67 and 144. The observed phenylalanine is tentatively assigned as Phe 62 in the amino acid chain section 57-62 because Phe 35 is located in a small amino acid chain section 32-39 between the LS and RS helix (26). This is not in contradiction with the fact that Pro 63 is not observed, since this residue and Phe 62 need not have rotation about two carbon-carbon axes in common; the section 57-62 may be observed with Pro 63 still absent. The remaining observable side chain carbons may be those of Asp 64, Ser 65 and Asp 66.

Rotational motions within the virus cannot be restricted to the assigned amino acid residues since the number of observed CH_2 carbons (1) exceeds the maximum number which is calculated from Table 1. To be sure, the number of residues exhibiting rotational motions exceeds the number of observed residues. In order to obtain a somewhat clearer view on the extent of mobility in the backbone, we note that the number of α-carbons forced into rotational motion by adjacent observable carbons can be estimated by adding to the number of amino acids of each of the observable sections, a number of four (from connected amino acids which are also forced into motion and are non-observed with NMR), which results in > 20 mobile residues. From the upper limit for the linewidth (300 Hz) (1) for the observed carbons it can be calculated that the observed carbons must possess mobility which can be described by full rotations about two carbon-carbon bonds on a time scale smaller then 10^{-7}.

5.3.4 Rod-like polymerization

The interpretation of CD spectra of rod-like polymers cannot be based on the assumption given for the interpretation of the virus CD spectrum because in the former case the protein subunits themselves can have motional degrees of freedom.

A number of carbons may have rotational mobility in common with the virus although phenylalanine is absent in the CD spectra of rod-like polymers (Fig. 2).

The observable number of protein subunits within the rod-like oligomer, which has been demonstrated to remain constant upon addition of one proton (4), titrates anomalously (4,5) below pH 6.3. This proton uptake results in an upfield shift for at least one aspartic acid (Fig. 2). Furthermore, the upfield shift of peak 1 upon changing pH from 7.0 (Fig. 2B) to 5.3 (Fig. 2C) corresponds to a deprotonation reaction of at least one amino group ($NH_3^+ \rightarrow NH_2 + H^+$) of an arginine side chain which has been suggested to occur if such a charged group is forced into a hydrophobic environment (14,26). It can be conceived that the proton released from the arginine amino group is taken up by an acid amino acid residue which is adjacent to this arginine (23). The involvement of a lysine has been shown for another plant virus, Cowpea Chlorotic Mosaic Virus (25). In view of the results presented before, at least one aspartic acid, one arginine and possibly one glutamic acid are involved in the uptake of the first proton.

At pH values below 6.3 three types of subunits are present (4), two of which are observed with NMR. It is proposed that one type contains two protonated carboxyl groups with at least one aspartic acid and one arginine, all probably in a hydrophobic environment, whereas the other type has both carboxyl groups unprotonated, an arginine amino group being protonated in a hydrophilic environment. The presence of two types of protein subunits with different conformations may be an explanation for the fact that spin-lattice relaxation cannot be described by a single exponential. These conformational differences may be also reflected in line shape changes of e.g. β-Thr at 125.8 ppm.

It is reasonable to assume that the aspartic acid, arginine and the other acid amino acid residue belong to the so called carboxyl cage (23). The position of peak 3 at about 165 ppm in Fig. 2 suggests that this other residue may be a glutamic acid. Ascribing the observation of protein subunit resonances to mobility of the protein subunit itself implies that the complete carboxyl cage involving two protein subunits cannot be present in rod-like polymers.

5.4 ACKNOWLEDGEMENT

The authors wish to thank Mrs. N.C.M. Alma-Zeestraten for her valuable assistance with the NMR measurements. The technical assistance of Miss M. Boerjan with the preparation of TMV and its protein is gratefully acknowledged.

We are indebted to the Department of Virology for hospitality and laboratory

facilities. The NMR investigations at high magnetic field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

5.5 REFERENCES

1. De Wit, J.L., Alma, N.C.M., Hemminga, M.A. and Schaafsma, T.J. Biochemistry, submitted for publication, Chapter 2.
2. Torchia, D.A. and Van der Hart, D.L. (1976) J. Mol. Biol. 104, 315-321.
3. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J. (1978) J. Magn. Resonance, Chapter 3.
4. De Wit, J.L. and Schaafsma, T.J., FEBS Lett, submitted for publication, Chapter 4.
5. Durham, A.C.H., Vogel, D. and De Marcillac, G.D. (1977) Eur. J. Biochem. 79, 151-159.
6. Campbell, I.D., Dobson, C.M., Williams, R.J.P. and Xavier, A.V. (1973) J. Magn. Resonance 11, 172-181.
7. De Wit, J.L., Alma, N.C.M. and Schaafsma, T.J., Biochim. Biophys. Acta, submitted for publication, Chapter 6.
8. James, T.J. Nuclear Magnetic Resonance in Biochemistry, p. 246 (1975) Academic Press, New York.
9. Christl, M. and Roberts, J.D. (1972) J. Amer. Chem. Soc. 94, 4565-4573.
10. Gurd, F.R.N., Lawson, P.J., Cochran, D.W. and Wenkert, E. (1971) J. Biol. Chem. 246, 3725-3730.
11. Keim, P., Vigna, R.A., Marshall, C. and Gurd, F.R.N. (1973) J. Biol. Chem. 248, 6104-6113.
12. Feeney, J. in New Techniques in Biophysics and Cell Biology, Vol. 2, (1975) eds. Pain, R.H. and Smith, B.J., Wiley-Interscience, London.
13. This thesis, Chapter 7.
14. Lauffer, M.A. Entropy-driven processes in Biology, p. 72 et seqq. (1975) Springer-Verlag, Berlin.
15. Durham, A.C.H. (1972) J. Mol. Biol. 67, 289-305.
16. Doddrell, D., Glushke, V. and Allerhand, A. (1972) J. Chem. Phys. 56, 3683-3689.
17. Kaper, J.M. (1975) The chemical basis of virus structure, dissociation and reassembly, p. 387 et seqq., North Holland Publishing Company, Amsterdam.
18. Van Putte, K. (1970) J. Magn. Resonance 2, 23-34.

19. Woessner, D.E. (1962) J. Chem. Phys. 36, 1-4.
20. Levine, Y.K., Partington, P. and Roberts, G.C.K. (1973) Mol. Physics 25, 497-514.
21. Glushke, V., Lawson, P.J. and Gurd, F.R.N. (1972) J. Biol. Chem. 247, 3176-3185.
22. Woessner, D.E. (1969) J. Chem. Phys. 50, 719-721.
23. Stubbs, G., Warren, S. and Holmes, K. (1977) Nature 267, 216-221.
24. Durham, A.C.H. and Butler, P.J.G. (1975) Eur. J. Biochem. 53, 397-404.
25. Bancroft, J.B., Rees, M.W., Johnson, M.W. and Dawson, J.R.O. (1973) J. Gen. Virology 21, 507-513.
26. Ansevin, A.T., Stevens, C.L. and Lauffer, M.A. (1964) Biochemistry 3, 1512-1516.
27. Freeman, R. and Hill, H.D.W. (1971) J. Chem. Phys. 54, 3367-3377.

6. ^{13}C NMR AND ^{13}C ENRICHMENT OF TOBACCO MOSAIC VIRUS

J.L. de Wit, N.C.M. Alma, T.J. Schaafsma
submitted to Biochim. Biophys. Acta.

6.1 INTRODUCTION

6.1.1 Tobacco Mosaic Virus (TMV)

Tobacco Mosaic Virus (TMV) is a rod-shaped plant virus with one RNA chain of ∿ 6600 nucleotides protected by 2200 identical protein subunits (molecular weight 17,500) in a helical array. The total molecular weight is 42 x 10^6 (1). X-ray diffraction resulted in an incomplete molecular model for the protein structure (6). Even at 0.4 nm resolution, such a model could be constructed due to the large α-helix content and the approximately parallel course of the α-helix sections of the backbone.

6.1.2 Tobacco Mosaic Virus protein

Tobacco Mosaic Virus protein has the past decades been studied by a variety of techniques such as e.g. circular dichroism (2), fluorescence (5), sedimentation analysis (3), electron microscopy (4) and X-ray diffraction (9). A good understanding of TMV-protein polymerization behaviour has evolved from these studies (7,8). X-ray diffraction (9) indicates a similar folding of the amino acid chain within the protein subunit in the virus and double disk (6,9).

6.1.3 ^{1}H and ^{13}C Nuclear Magnetic Resonance

Recently it has been shown that both ^{1}H and ^{13}C NMR yield valuable information about the virus (10,13) and double disk- and rod-like protein oligomers (11,12,13). Usually NMR is assumed to be applicable only to small biological systems (e.g. proteins, nucleotides, lipids), with a molecular weight of ∿ 50,000 as an upper limit. Severe line broadening and lengthening of the spin-lattice relaxation time (T_1) is predicted for rigid systems having larger molecular weight (14). However, NMR spectra of TMV, double disk-like oligomers (molecular weight 0.6 x 10^6) and rod-like oligomers (molecular weight > 2 x 10^6) could be observed, thanks to rotational motion of protein subunits and/or

mobility in the protein subunit within the virus (10,13) and protein oligomers (10-13).

It has been shown that rotational motions within the protein subunits are present at NMR time scale (13). Following an analysis of T_1 data of virus, double disk- and rod-like oligomers (10-13), these motions are defined by correlation times τ_g describing rotations about all carbon-carbon bonds within a protein subunit of $\lesssim 5 \times 10^{-10}$ s with an upper limit of the average angle of rotation $\beta_o \sim 40^o$ for all observable carbons (28,13). For about 10% of all carbons within a protein subunit of the virus β_o must be 360^o for rotation about at least two carbon-carbon bonds, attached to the observable carbon nuclei (10,13).

Protein subunit mobility (translational, rotational) is observed in the double disk-like oligomer and is also present in the rod-like oligomer, depending on pH (12). This pH-dependence has been shown to involve protonation of two carboxyl groups of at least one aspartic acid, followed by a deprotonation reaction of the protonated amino group of an arginine side chain, both with anomalous pK's. The detectable part of the protein subunits within the rod-like oligomer, has been suggested to consist of two conformers which depend on the protonation state (12,13,16).

The model describing protein subunit mobility and the discussion about the presence of different protein subunit conformers is based on the generally accepted existence of a secondary and tertiary structure within the TMV protein subunit (10). In this paper the presence of this three dimensional structure in solution is supported using ^{13}C NMR and urea denaturation of protein subunits. In addition a comparison of ^{13}C NMR spectra of protein subunits of TMV strains Vulgare and U_2, computer simulation of the spectra of urea-denatured TMV protein and the difference between spectra of Vulgare and U_2 will be shown to favour the presence of secondary and tertiary subunit structure.

6.1.4 Enrichment with stable isotopes

Enrichment with stable isotopes, e.g. ^2H, ^{13}C and ^{15}N in combination with NMR has found wide application (17). Enrichment with these nuclei in algea, yeast and bacteria (15) can be carried out without much trouble. In plants, only ^{13}C and ^{15}N can be incorporated to a sufficiently high degree ($\sim$ 90%) without starvation (18,19), although isotope dilution causes increasing difficulty. ^2H enrichment without plant starvation can only be used up to $\sim$ 50%

isotope content, but growth is severely inhibited (De Wit, unpublished observations), similar to what is observed for animals (20,21). Some strains of algae and yeast must first be adapted to increasing $^{2}H_2O$ concentrations before a sufficient growth rate is reached in 99% $^{2}H_2O$ (22). Selective enrichments with ^{2}H, ^{13}C and ^{15}N are suitable for tracer studies, and for various NMR applications (23,24,17,15); non-selective enrichments with ^{13}C and ^{15}N are often benificiary if one wants to obtain NMR spectra with a reasonable signal to noise ratio.

It is for this reason that TMV has been non-selectively enriched to about 12%. Tobacco plants infected with TMV represent a very efficient system to incorporate ^{13}C with minor isotope dilution. The enrichment procedure and a simple method for the $^{13}CO_2$ ratio determination are also described in this paper.

6.2 MATERIALS AND METHODS

6.2.1 TMV purification and protein preparation

The purification of TMV strains Vulgare and U_2 was carried out according to Leberman (25). The protein preparation procedure used is similar to that described by Durham et al. (7), with minor changes described elsewhere (11). Protein preparations in 0.12 M Tris-HCl, pH 8.6, 5°C were checked and characterized with polyacrylamide gelelectrophoresis showing a single band (26), sedimentation analysis (7) with a single sedimentation coefficient $\sim$ 3.8 S and spectrometrically for Vulgare yielding $E_{280}/E_{250} > 2.5$. Compared to Vulgare, U_2 has a different aromatic amino acid composition, preventing a spectroscopic purity-check; purity of U_2 protein preparations has been determined with sedimentation analysis, showing a sedimentation coefficient of $\sim$ 3.8 S at pH 8.6 in 0.12 M Tris-HCl, 5°C.

6.2.2 ^{13}C NMR spectra at 90.5 MHz

^{13}C NMR spectra at 90.5 MHz have been obtained with a Bruker SPX 360 supercon spectrometer equipped with quadrature detection using 10 mm sample tubes and $\sim$ 5 W continuous wave broadband ^{1}H decoupling power. Sample pH was directly measured in the tube with a long 5 mm diameter pH electrode (Radiometer). Sample temperature was maintained with a modified version of the Bruker temperature accessory (11) and was measured with a thermometer in the tube, being positioned inside the probe. After equilibration of sample temperature, the tube with the

thermometer was quickly removed from the probe with the air lift for taking thermometer readings. The error expected for measuring sample temperature in this manner is estimated to be $\pm$ 2°C.

6.2.3 Computer simulations

Computer simulations of ^{13}C spectra of TMV-protein were performed using the known amino acid composition of TMV strains Vulgare and U_2. The starting positions for the chemical shifts were basically those of James (29) and were shifted within + or - 0.4 ppm from these positions for maximum overlap with the denatured TMV-protein spectrum. A DEC 10 computer program allows convolution of each individual carbon with a line shape function.

6.2.4 The preparation of 10-15% enriched TMV

The enrichment of TMV proceeds through assimilation of 90% ^{13}C enriched carbon dioxide by green leaves of the Samsun NN' variety of Nicotiana tabacum, infected with TMV. The amount of 90% ^{13}C enriched barium carbonate (Bio Rad) used for generating $^{13}CO_2$, was carefully chosen in relation to the wet weight of the total number of leaves, so that the final enrichment of TMV always ends up between 10-15%. This enrichment percentage does not lead to appreciable carbon-carbon spin-spin coupling which would decrease the resolution and signal to noise ratio without improving the spectral information (17). Tobacco plants of about six weeks age were given a dark period of one night in a temperature controlled room (25°C) in order to reduce the sugar content. The middle leaves of the plant are inoculated with TMV about two hours after starting the experiments, and are transferred in the dark to sterilized 100 ml beakers containing sterilized nutrient solution which were placed in a closed system of desiccators interconnected with butyl tubing which were also connected to a Kipp apparatus. The composition of the nutrient solution is equivalent to that of the Hoagland mineral salt solution and is completed with aureomycine (Serva) and fungizone (Squibb) to a concentration of 1 and 2 µg/ml, respectively, for the prevention of bacterial and mould growth; ferricitrate (0-5 µg/ml) is used as the iron source. In the desiccators high humidity is maintained through a water film on the bottom. After a second dark period of twelve hours, necessary for prolonged carbohydrate exhaustion, kinetine (Sigma) to a concentration of 1 µg/ml, is added to the nutrient solution to inhibit ageing of leaves. Immediately after

installing the leaves in the growth system, the desiccators are covered so that in the shortest possible time a maximum humidity is reached. During about 15 minutes the system is then flushed with a carbon dioxide-free mixture of 80% nitrogen and 20% oxygen. After closing the system from air, $^{13}CO_2$ (∿ 90% enriched) is released from barium carbonate, using a cylindrical funnel with pressure equalizing tube, by dropwise addition of hydrochloric acid (5 ml 2.5 N hydrochloric acid per gram barium carbonate) and admitted to the system (Kipp apparatus). After approximately 30 min., the time needed for diffusion of carbon dioxide into the system, the plants are illuminated by two 400 W sodium lamps (Philips SON 400 W, BSN 400 L02, S 50), at approximately 50 cm distance from the bottom of the desiccators. To prevent ageing it is crucial to keep temperature constant at ∿ 25°C, especially during the change from the dark to the light period. Growth of TMV is then continued as long as leaf degeneration is absent (large brown spots). After growth is stopped, leaves are frozen for storage. The average growth period is about four days whereas the average yield is 1-1.5 mg/g wet leaf under these conditions.

Approximately 10 g barium carbonate (∿ 90% enriched) is used for ∿ 100 mg enriched TMV. The relative low TMV yield compared to that for growth under normal conditions is caused by the short TMV multiplication period. A one day extension would probably double the amount of TMV, because TMV multiplication normally shows sigmoid behaviour with maximal multiplication at about 5 days. Furthermore, it has been found that TMV yield depends on light intensity and -spectrum, both of which were optimized by using 400 W sodiumlamps producing continuous illumination during the TMV multiplication period. Both assimilation and dissimilation rates (20 mg CO_2/hr/dm^2 leaf surface (34) and 4 mg CO_2/hr/g dry leaf weight (35) are high, so that after the addition of $^{13}CO_2$, a $^{13}CO_2/^{12}CO_2$ equilibrium is reached very quickly, the $^{12}CO_2$ originating from residual endogenous starch. The atmospheric $^{13}CO_2/^{12}CO_2$ ratio at any moment during the enrichment process can easily be determined. This is of importance, because this ratio directly controls the final TMV enrichment percentage.
A glas system is favourable since it is gas-impermeable, in contrast with most plastic chambers.

6.2.5 Infrared determination of the $^{13}CO_2/^{12}CO_2$ ratio

About 6 mg of TMV is completely burned in an oven at 1100°C under a continuous oxygen flow. The resulting mixture of gases is then led through a tube

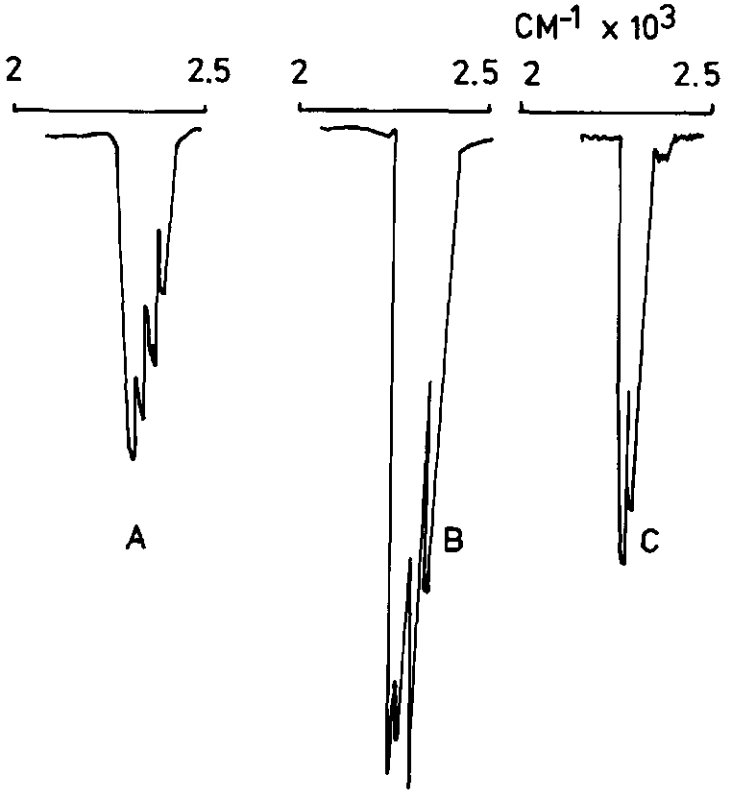


Fig. 1. Infrared absorption spectra of different mixtures of $^{13}CO_2$ and $^{12}CO_2$ taken against a reference gascell filled with a CO_2-free mixture of 80% N_2 and 20% O_2. A, standard mixture of 11.7% $^{13}CO_2$; B, 16% $^{13}CO_2$ obtained from enriched TMV-protein; C, atmospheric CO_2.

filled with potassium chlorate. For analysis, carbon dioxyde is precipitated as barium carbonate by flowing it through a carbon dioxide-free barium hydroxide solution. Hydrochloric acid is added dropwise so that the released carbon dioxide flows into an infrared gascel. The $^{13}CO_2/^{12}CO_2$ ratio is determined from the infrared carbon dioxide spectrum taken on a Hitachi double beam I.R. spectrometer (EPI-G3) against a gascel filled with the carbon dioxide free nitrogen/oxygen mixture.

$^{12}CO_2$ strongly absorbs at 2349 cm^{-1} while the $^{13}CO_2$ absorption is shifted

to 2284.5 cm^{-1}. Both absorption peaks are split as shown in Fig. 1. A relation between the $^{13}CO_2/^{12}CO_2$ ratio and the ratio of the I.R. $^{13}CO_2$ and $^{12}CO_2$ peaks has been determined using the reference mixtures of known amounts of enriched and normal barium carbonate for which the same procedure is followed for the release of the $^{13}CO_2$ and $^{12}CO_2$ mixture from the enrichment experiment.

From a number of those experiments we have derived the following empirical equation

$$^{12}CO_2 : {}^{13}CO = 2.25 \times \left|\frac{h_C}{h_A} + \frac{h_D}{h_B}\right| \qquad [1]$$

where h_A, h_B, h_C, h_D are the peak heights presented in the spectrum of Fig. 1. The estimated error of the experimental isotope ratio is $\pm$ 5%.

6.3 RESULTS AND DISCUSSION

6.3.1 TMV ^{13}C NMR

Fig. 2 represents the ^{13}C NMR spectrum of TMV-protein at pH 11.7, 7°C and

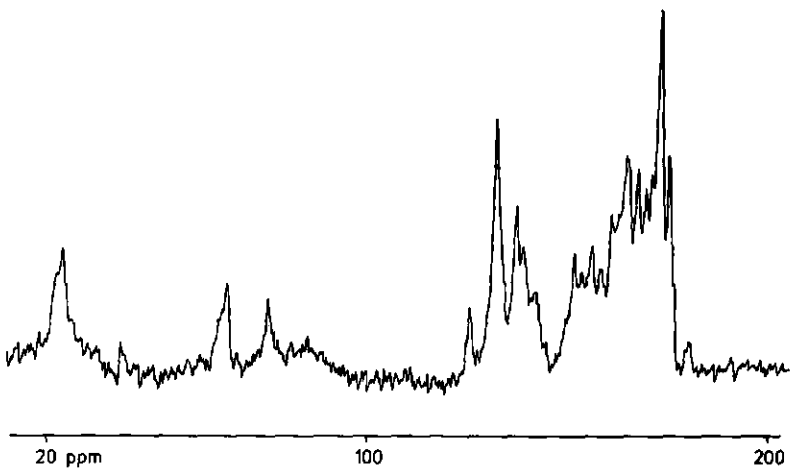


Fig. 2. Broadband ^1H decoupled 90.5 MHz ^{13}C spectrum of TMV-protein using quadrature detection with an acquisition plus delay time of 0,4 s; 17,000 accumulations; a sensitivity enhancement of 17 Hz. Further conditions are, concentration: 30 mg/ml; pH 11.7; temperature 7°C; 0.1 M NaCl. The ppm scale is referenced to CS_2 taking β-Thr at 125.8 ppm. The baseline is distorted due to spectral humps.

30 mg/ml. Under these conditions the sedimentation coefficient is $\lesssim$ 2.9 S, corresponding to a protein oligomer with a size between monomer and trimer. The peaks superimposed on the broad spectral hump have an average line width of 30 to 90 Hz. In the absence of internal motions the line width parameter T_2^{-1} under the condition that $\omega^2_{C,H}\,\tau^2_R > 1$ can be given for aliphatic carbons by the dipole-dipole relaxation equation (30),

$$T_2^{-1} = 0.2\ (\mu_0/4\pi)^2 N\gamma_C^2\gamma_H^2\hbar^2\tau_R r_{CH}^{-6} \qquad [2]$$

where T_2 is the spin-spin relaxation time, γ_H and γ_C are the gyromagnetic ratios of protons and carbons, respectively, r is the length of a C-H bond (0.109 nm), N is the number of protons interacting with an observable carbon and τ_R is the correlation time of the rotational motion of the TMV-protein oligomer and $\omega_{C,H}$ is the measuring frequency. For a sphere τ_R is given by (31).

$$\tau_R = M\bar{v}\eta/RT \qquad [3]$$

where $\bar{v}$ is the partial specific volume, M is the molecular weight and η is the viscosity. Assuming an average molecular weight of 35,000 under these conditions, τ_R is calculated to be 10^{-8} s for the dimer using Eqn. [3]. Eqn. [2] then yields a line width $(\pi T_2)^{-1}$ of 14 Hz for N = 1. Although the dimer geometry approaches that of a sphere, deviations from spherical symmetry may cause τ_R to be slightly larger than 10^{-8} s. Both experimentally observed line widths of the resolved ^{13}C resonances (Fig. 2) and upper limits for the ^1H line width presented elsewhere (11), exceed the calculated value of 14 Hz. This demonstrates that the variation in line width of all resolved resonances in the ^{13}C spectrum shown in Fig. 2 is determined by chemical shift inequivalence. A comparison of the ^{13}C spectrum shown in Fig. 2 with previously presented ^{13}C spectra of the trimer at pH 10.0, 7°C (12) shows a large decrease in line width of the resolved groups of resonances upon increasing pH from 10.0 to 11.7 which can now be interpreted to result from partial unfolding of the protein subunit as has also been shown from a spin label EPR study (33).

6.3.2 Urea denaturation of TMV-protein

The difference between the spectrum of TMV protein at pH 11.7 (Fig. 3A) with that of the same protein under denaturing conditions (Fig. 3B) can be explained by the presence of chemical shift inequivalence (32) due to secondary

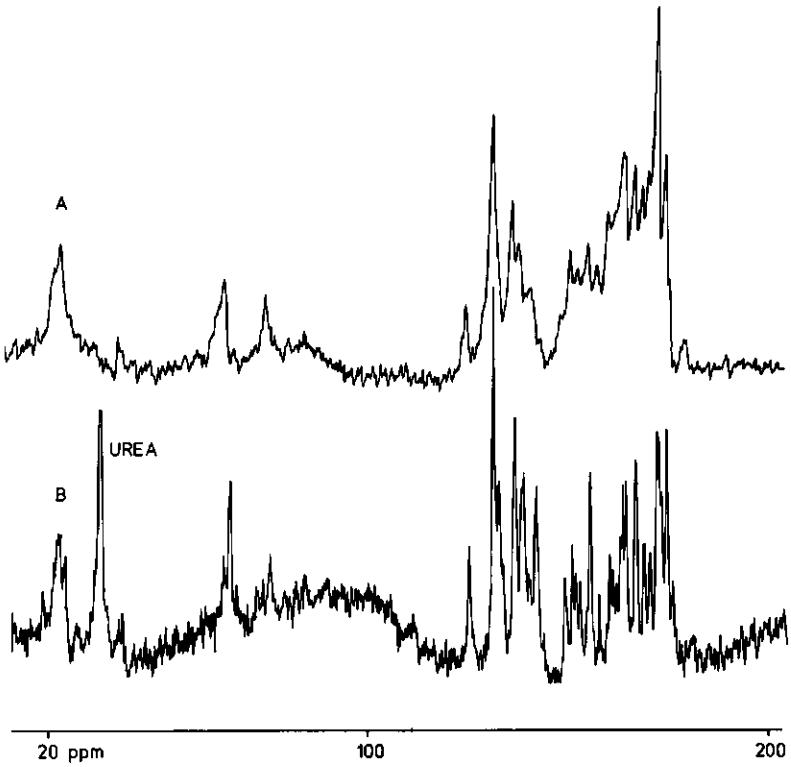


Fig. 3. Broadband ^{1}H decoupled 90.5 MHz ^{13}C spectra of TMV-protein using quadrature detection with an acquisition plus delay time of 0.4 s; concentration: 30 mg/ml; 0.1 M NaCl; pH 11.7; temperature 7°C. Further conditions are, for A: 17,000 accumulations; sensitivity enhancement of 17 Hz; for B: 4700 accumulations; sensitivity enhancement of 17 Hz; 6 M urea. The ppm scale is referenced to CS_2 taking β-Thr at 125.8 ppm. The baselines are distorted due to spectral humps.

and tertiary structure in native TMV protein. In Fig. 4A the 6 M urea denatured spectrum is shown; Fig. 4B represents the best fit of a computer simulation based on the known amino acid composition. The line widths used for this simulation are given in Table 1 and the resonance positions of the amino acids are essentially similar to those published (29), with slight deviations necessary for an optimum fit. Such deviations ($< \pm 0.4$ ppm) are attributed to residual chemical shift inequivalence and the effect of pH. A comparison of the 6 M urea-denaturated TMV-protein subunit spectrum (Fig. 4A) with the computer-simulated spectrum (Fig. 4B) shows that the former indeed arises from the fully denatured TMV-protein subunit. The observed differences between both spectra may arise from a number of

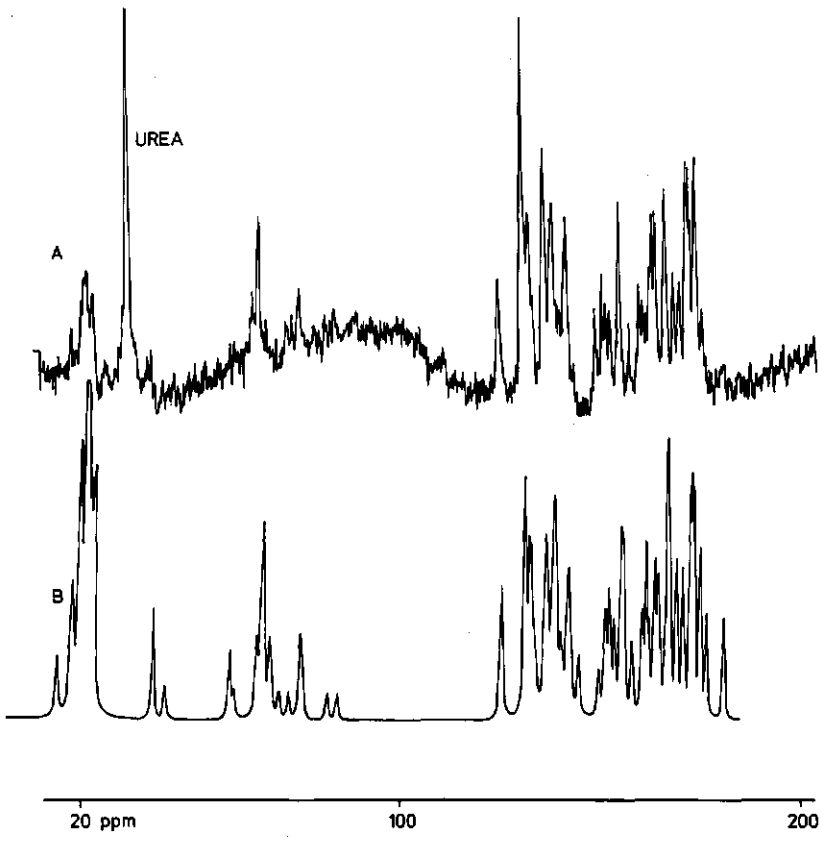


Fig. 4. A: broadband ^{1}H decoupled 90.5 MHz ^{13}C spectrum of TMV-protein using quadrature detection with an acquisition plus delay time of 0.4 s; concentration: 30 mg/ml; 0.1 M NaCl; pH 11.7; temperature 7°C; 6 M urea; accumulations, 4700; sensitivity enhancement, 17 Hz. B: computer simulation of the spectrum shown in A using the known amino acid composition of TMV strain Vulgare, line widths presented in Table 1 and spectral positions of James (29). The ppm scale is referenced to CS_2 taking β-Thr at 125.8 ppm.

causes such as: NOE effects, deviations from random ^{13}C distribution, residual

Table 1. Line widths used in the computer simulations shown in Fig. 4B.

type of carbon[a]	Line width[b] (Hz)
primary (CH_3)	45
secondary (CH_2)	50
tertiary (CH side chain)	60
tertiary (C_α)	70
quaternary	60
quaternary, carbamyl	50
quaternary, carbonyl	60

[a] The distinction in these type of carbons is based on the line widths which must be chosen for a best fit.

[b] In this simulation it is assumed that lines are lorentzian, the relaxation mechanism is dipolar and that each peak surface is proportional to the number of carbons at a certain spectral position.

chemical shift inequivalence in the spectrum of denatured protein, deviations from a uniform radio frequency field distribution and differences in internal mobility. The satisfactory agreement between the spectra of Figs. 4A and 4B indicates, however, that none of these causes makes a large contribution to the observed differences.

An estimation of the range of chemical shift inequivalence present in TMV-protein can be deduced from the experiments shown above to amount to at least $\pm$ 1 ppm.

6.3.3 Comparison of ^{13}C spectra of strains Vulgare and U_2

Additional evidence for the presence of secondary and tertiary structure in TMV-protein oligomers can be obtained by comparing ^{13}C NMR spectra of the protein oligomers of TMV strain Vulgare with those of U_2 as shown in Fig. 5A and B. A comparison of the computer simulated difference spectrum based on their amino acid compositions (Fig. 5C) with the spectral differences between the spectra in Fig. 5A and B reveals large deviations. The main cause for these deviations

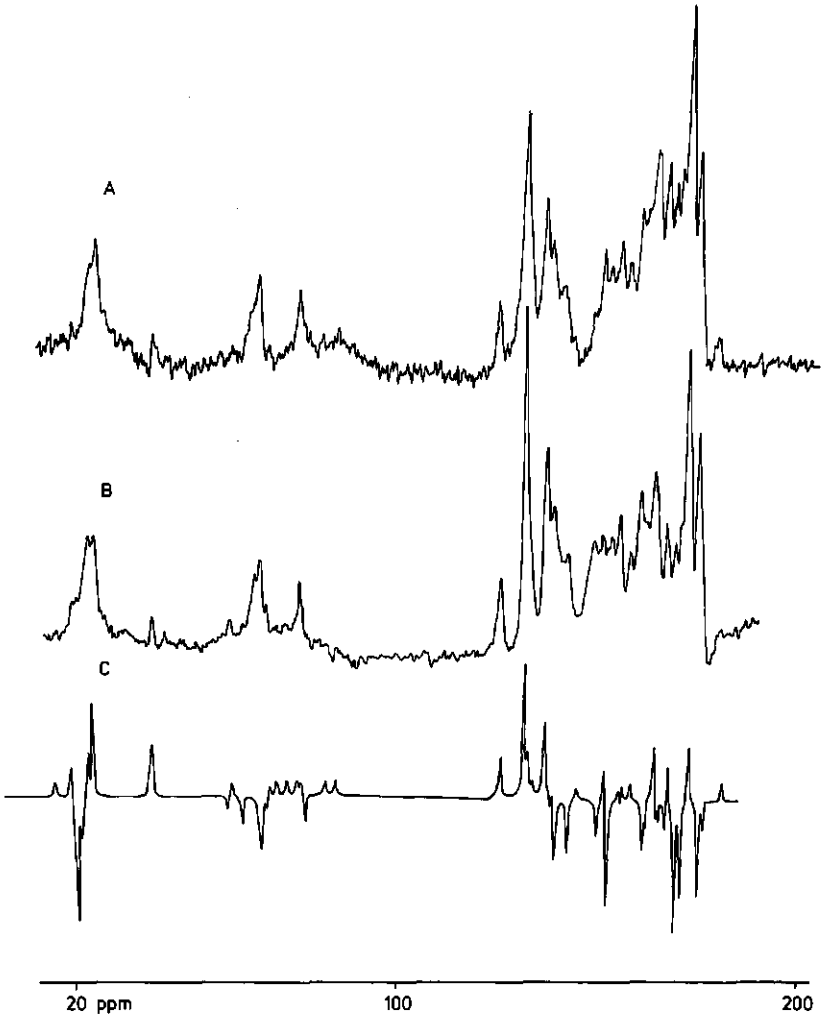


Fig. 5. Broadband ^{1}H decoupled 90.5 MHz ^{13}C spectra of TMV-protein; strains Vulgare and U_2 using quadrature detection with an acquisition plus delay time of 0.4 s; 0.1 M NaCl; pH 11.7; temperature 7°C; sensitivity enhancement, 17 Hz; accumulations 17,000. Further condiditons, for A; strain Vulgare; concentration 30 mg/ml. B: strain U_2; concentration 15 mg/ml; C: computer simulated difference spectrum between Vulgare (up) and U_2 based on their amino acid composition. The ppm scale is referenced to CS_2, taking β-Thr at 125.8 ppm.

must arise from variations in secondary and tertiary structure between Vulgare and U_2 because of their large differences in amino acid composition. This is an interesting demonstration of the sensitivity of ^{13}C NMR for small differences in protein structure in solution.

6.4 ACKNOWLEDGEMENTS

The technical assistance of Mr. H. Jongejan of the Department of Organic Chemistry and of Miss M. Boerjan is gratefully acknowledged. We are indebted to the Department of Virology for providing laboratory facilities.

The NMR investigations at high magnetic field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

6.5 REFERENCES

1. Lebeurier, G., Nicolaieff, A. and Richards, K.E. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 149-153.
2. Schubert, D. and Krafczyk, B. (1969) Biochim. Biophys. Acta. 188, 155-157.
3. Lauffer, M.A. Entropy-driven processes in Biology (1975) Springer Verlag, Berlin.
4. Durham, A.C.H. and Finch, J.T. (1972) J. Mol. Biol. 67, 307-314.
5. Guttenplan, J.B. and Calvin, M. (1973) Biochim. Biophys. Acta. 322, 294-300.
6. Stubbs, G., Warren, S. and Holmes, K. (1977) Nature, 267, 216-221.
7. Durham, A.C.H. (1972) J. Mol. Biol. 67, 289-305.
8. Caspar, D.L.D. (1963) Adv. Prot. Chem. 18, 37-121.
9. Champness, D.L.D., Bloomer, A.C., Bricogne, C., Butler, P.J.G. and Klug, A. (1976) Nature 259, 20-24.
10. De Wit, J.L., Alma-Zeestraten, N.C.M., Hemminga, M.A. and Schaafsma, T.J. Biochemistry, submitted for publication, Chapter 2.
11. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J. (1978) J. Magn. Resonance, in press, Chapter 3.
12. De Wit, J.L. and Schaafsma, T.J. FEBS Lett., submitted for publication, Chapter 4.
13. De Wit, J.L. and Schaafsma, T.J. Eur. J. Biochem., submitted for publication, Chapter 5.
14. Torchia, D.A. and Vander Hart, D.L. (1976) J. Mol. Biol. 104, 315-321.
15. Proc. of the first int. conf. on stable isotopes in chem., biol. and med. (1973) Argonne (eds. Klein, P.D. and Peterson, S.V.). Available through Nat. Techn. Info. Service U.S. Department of Commerce.
16. Butler, P.J.G., Durham, A.C.H. and Klug, A. (1972) J. Mol. Biol. 72, 1-18.
17. Wüthrich, K., NMR in Biological Research: peptides and proteins (1976) North Holland Publishing Company, Amsterdam.

18. Siegel, S.M., Halpern, L.A. and Giumarro, C. (1964) Nature 201, 1244-1245.
19. Kollman, V.H., Gregg, C.T., Hanners, J.L., Whaley, T.W. and Ott, D.G. (1973) Proc. of the first int. conf. on stable isotopes in chem., biol. and med., Argonne (eds. Klein, P.D. and Peterson, S.V.). Available through Nat. Techn. Info. Service U.S. Department of Commerce.
20. Manson, L.A., Carp, R.I., Defendi, V., Rothstein, E.L., Hartzell, R.W. and Krîtchevsky, D. (1960) Ann. New York Acad. Sci. 84, 685-694.
21. Morowitz, H.J. and Brown (1953) Natl. Bur. Standards Rept. No 2179.
22. Crespi, H.L., Conrad, S., Uphans, R.A. and Katz, J.J. (1960) Ann. New York Acad. Sci. 84, 648-666.
23. MacInnes, A.G., Walter, J.A., Wright, J.L.C. and Vining, L.C. (1976) in Topics in carbon-13 NMR spectroscopy, 125-172 (ed. Levy, G.C.) John Wiley and Sons, New York.
24. Blomberg, F., Maurer, W. and Rüterjans, H. (1976) Proc. Nat. Acad. Sci. U.S.A. 73, 1409-1413.
25. Leberman, R. (1966) Virology 30, 341-347.
26. Ornstein, L. (1964) Ann. N.Y. Acad. Sci. 121, 321-349.
27. Wittmann-Liebold, B. and Wittman, H.G. (1967) Mol. Gen. Genet. 100, 358-363.
28. Van Putte, K. (1970) J. Magn. Resonance 2, 23-34.
29. James, T.L. (1975) Nuclear Magnetic Resonance in Biochemistry, p. 246, Academic Press, New York.
30. Lyerla, J.R. and Levy, G.C. (1974) in Topics in carbon-13 NMR spectroscopy, 79-143 (ed. Levy, G.C.) John Wiley and Sons, New York.
31. Hallenga, K. and Koenig, S.H. (1976) Biochemistry 15, 4255-4264.
32. Feeny, J. (1975) New Techniques in Biophys. and Cell Biol. Vol. 2, 312-318 (eds. Pain, R.H. and Smith, B.J.) Wiley-Interscience, London.
33. Hemminga, M.A., v.d. Boomgaard, Th.P.A. and De Wit, J.L. (1977) FEBS Lett. 85, 171-174.
34. Zaitlin, M. and Hesketh, J.D. (1965) Ann. Appl. Biol. 55, 239-243.
35. Owen, P.C. (1957) Ann. Appl. Biol. 45, 327-331.

7. A THERMODYNAMIC APPROACH TO THE MECHANISM OF ASSEMBLY AND DISSOCIATION OF TOBACCO MOSAIC VIRUS

7.1 INTRODUCTION

The assembly mechanism of Tobacco Mosaic Virus (TMV), a rod-shaped plant virus with a molecular weight of 42 x 10^6, consisting of RNA and protein, has long been an intriguing problem. In vitro, the assembly process can be distinguished in an initiation and elongation step (1). It is now agreed upon that the so-called double-disk constitutes the protein oligomer forming the initiation complex with a specific part of TMV RNA under nearly physiological conditions (2). The initiation site has been located at about 1000 nucleotides from the 3' RNA end (3,4). Many recognition mechanisms in biology (5) involve proteins interacting with specific regions of a RNA or DNA molecule, so that the study of the initiation process of TMV is of general interest. Electron microscopy observations have also resulted in a better understanding of the elongation process (3); the resolution of X-ray data (33) on TMV is sufficient to reveal the folding of RNA, relevant for this process. Models for the dissociation mechanism of TMV under physiological conditions have not grown beyond the level of suggestions.

Full understanding of the initiation and elongation step requires a description of the double disk-like oligomer and RNA in solution on molecular level. Recently it has been shown that in particular ^{13}C Nuclear Magnetic Resonance (NMR) yields valuable information on these processes. Large sections of the amino acid chain within the protein subunits (molecular weight 17,500) in TMV and its protein oligomers are mobile reflecting a thermodynamically unstable protein structure (7). From such NMR experiments, it can also be concluded that protein subunits in rod-like and double disk-like oligomers have rotational and/or translational degrees of freedom (7,8). Due its rigidity, the structure of native TMV, RNA cannot be elucidated by NMR.

With the aid of Convolution Difference (9) (CD) spectra, a more detailed location of the unstable regions within the protein subunit is presented. Taking into consideration TMV RNA structure in solution, we have constructed a model for the mechanism of assembly and dissociation of TMV based on NMR results and thermodynamics.

7.2 THERMODYNAMICS AND THE ROLE OF WATER

Polymerization in solution of TMV protein is a spontaneous endothermic process and thereby entropy-driven (10,11). The energy contribution of the entropy increase upon polymerization of TMV protein to double disk-like oligomers exceeds the positive enthalpy so that the Gibbs free energy change upon polymerization is negative (10). This is summarized by the equation,

$$\Delta G = \Delta H - T\Delta S, \qquad [1]$$

where ΔG is the Gibbs free energy change in $J.mole^{-1}$, ΔH the enthalpy change in $J.mole^{-1}$, T the absolute temperature in degrees Kelvin and ΔS the entropy change in $J.mole^{-1}.K^{-1}$. Many biological polymerization processes are entropy-driven and in this respect it has long been realized that water must play an important role (12).

In biological polymerization processes, which are controlled by the shielding of hydrophobic groups from water, the entropy contribution from water can be considered determinant (17). Water surrounding a hydrophobic boundary is structured and has decreased entropy compared to that of normal liquid water (14,13). Therefore, endothermic protein polymerization processes can be driven by an increase in water entropy through decreased contact between water and hydrophobic side chains of amino acid residues. For TMV protein polymerization, water release has been demonstrated (15).

An interesting implication of water being the determinant for entropy-driven protein polymerization is that the number of hydrophobic amino acid residues in a protein in contact with water upon polymerization must be made as small as possible. This can result in interactions between hydrophobic side chains of polymerizing protein subunits (16) without the formation of formal bonds. Such a polymerization process thus can occur in the absence of specific interactions such as salt bridges and hydrogen bonds between the proteins subunits. In this respect the entropy-driven protein polymerization has some resemblance with the association of lipids into micelles (17).

7.3 TMV-RNA

For the discussion of the structure of TMV RNA in solution it is useful to make a comparison with tRNA, which has been studied extensively (18,19). An

analysis of RNA crystal structure reveals maximal base stacking (20). Such stacking of nucleotide bases is considered to be the main driving force for the development of the RNA conformation in solution because the heat of formation of hydrogen-bonds between nucleotides in an RNA chain with intramolecular double-helix structure is comparable to that for binding between nucleotides and water (21,22).

Kinetic measurements show that the tertiary structure of $tRNA_2^{Glu}$ from *Escherichia Coli* is less stable than its secondary structure (23). It can further be shown that the experimental and theoretical ΔG (24) values for the formation of secondary structure in tRNA reasonably agree.

Let us now turn to the structural properties of viral RNA. A region of the RNA of TMV shows a tRNA-like structure since histidine is attached at the 3' end (25) by its amino acyl-tRNA synthetase. The regions of TMV RNA forming initiation complexes with the double-disk indicate that these regions are capable of forming hairpin-like secondary structures (26,27). The ΔG values of association into hairpins of TMV-RNA regions are expected to be in the range -23.4 to -133.1 kJ (26, 27), based on theoretical values for base-pairing (19).

The TMV RNA regions, discussed above, are known to have a high binding affinity to double disk-like protein oligomers (26,27). These regions are found at 700-1300 nucleotides from the 3' end of the RNA and constitute the site for initial RNA binding (4,26,27).

7.4 CONVOLUTION DIFFERENCE SPECTRA OF THE DOUBLE DISK-LIKE POLYMERIZATION

Following the same approach as for virus and rod-like polymers (7), we have also obtained convolution difference (CD) ^{13}C NMR spectra of the double disk-like oligomer, as shown in Fig. 1.

It has been found that the line width of all ^{13}C and ^{1}H resonances is independent of the increase in molecular weight upon double disk-like polymerization, and has an upper limit of 65 Hz (6,8) for ^{13}C. Furthermore, $\sim$ 20% of the carbons within the double disk-like oligomer show a decrease in line width upon increasing temperature (6). In Fig. 1A we present the CD spectrum of these carbons at 30°C and in Fig. 1B the CD spectrum of TMV protein trimers at 7°C. For comparison, in Fig. 1C the TMV CD spectrum (30°C) is shown. The line width of groups of resonances in the CD spectra of Fig. 1A and B decreases upon increasing temperature due to increased internal mobility affecting both chemical shift inequivalence and dipolar interaction.

In the disk CD spectrum at 30°C (Fig. 1A) more resonances are present then

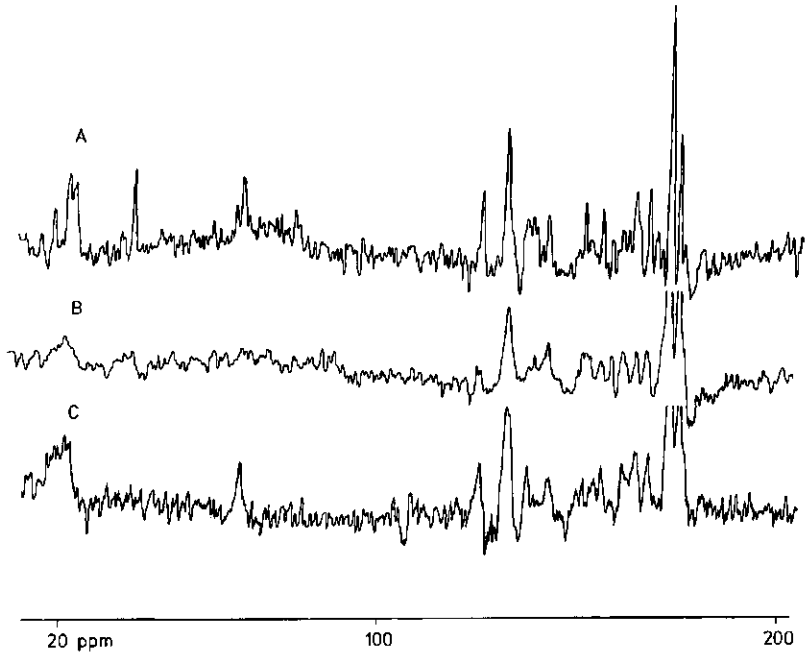


Fig. 1. ^{13}C convolution difference spectra of TMV and its protein oligomers obtained from ^{13}C spectra at 90.5 MHz take under broadband decoupling. A, acquisition time: 1 s; concentration: 40 mg/ml; temperature: 30°C; pH: 7.3 in 0.1 M NaCl; 4000 accumulations; CD time constants: 0.005 and 0.011 s with K = 1. B, acquisition time: 0.5 s; concentration: 40 mg/ml; temperature: 7°C; pH: 7.0 in 0.1 M NaCl; 17,000 accumulations; CD time constants: 0.005 and 0.011 s with K = 1. C, acquisition time: 1 s; concentration: 60 mg/ml; temperature: 30°C; pH: 7.2 in 1 mM sodiumphosphate; 50,000 accumulations; CD time constants: 0.005 and 0.011 s with K = 1. K is defined as a weighting factor for the substracting of the spectra. The ppm scale refers to CS_2 assuming 125.8 for the β-Thr carbon position.

observed in the virus CD spectrum (7) (Fig. 1C), i.e. those of δ-leucine, γ-glutamic acid, ε-arginine carbons. Phenylalanine ring carbons are observed in the region 50 to 100 ppm. This aromatic amino acid is absent in the CD spectrum at 7°C but was also observed at 30°C in the CD spectrum of the virus (7), suggesting temperature induced motion of a protein section; independent of the molecular weight (7) of virus or protein oligomers. In addition, the mobility of the phenylalanine side chain is independent of protein subunit mobility because in the virus, protein subunit mobility is absent due to protein/RNA interactions, whereas its resonances are still observed. Another difference between the CD

spectrum of the double disk-like oligomer (Fig. 1A) with both other spectra is the presence of the ε-arginine peak at 30°C.

7.5 DOUBLE DISK-LIKE OLIGOMER

The CD spectrum of the double disk-like oligomer (Fig. 1A) also includes resonances from phenylalanine. Noting that this amino acid is absent in the section 90-120, the carbons showing a decrease in line width upon increasing temperature (and molecular weight) cannot solely originate from amino acid residues in this section, which happens to have blurred electron density (28) in the X-ray micrograph. In agreement with assignments in the virus CD spectrum, the observed phenylalanine may be Phe 62 (7). We expect that the amino acids adjacent to Phe 62 are also present in the double disk CD spectrum at 30°C. The location of other mobile protein sections cannot be determined. An increased number of resonances in the double disk CD spectrum at 30°C (Fig. 1A) in comparison with the CD spectrum of the virus indicates the presence of more mobile sections, however. A difference between protein subunits in the virus (7) and in the double disk-like oligomer is the binding of RNA so that the arginine, glutamic acid and leucine carbons (Fig. 1A), only observed in the double disk CD spectrum may be located in the protein section which is capable of RNA binding (residues $\sim$ 90 to $\sim$ 120).

Previously, mobility within the double disk-like oligomer has been analyzed considering motion of the protein subunits in addition to motion of nuclei within the subunit (7). Then, salt bridges and hydrogen-bonds between protein subunits are absent on the NMR time scale (7). From this result we may conclude that the stability of the double disk-like oligomer is due to the positive ΔS (upon polymerization to double disk) arising from changes in the amount of water surrounding the oligomer (14). The positive ΔH may reflect mainly repulsion between protein subunits (one protein subunit has a negative charge of 2 units (30,31) at these neutral pH values), and is opposed by terms largely dominated by the entropy increase of water. Attractive v.d. Waals forces (29) between protein subunits need not be considered since ΔH increases even with a decrease of repulsion and distance between protein subunits upon lowering pH (1). This is shown through comparing rod-like and double disk-like polymerization (13).

In Fig. 2 hypothetical curves for the repulsion and attraction as a function of an arbitrarily intersubunit distance are presented. In this figure, curve I represents the sum of all (repulsive and attractive) energy terms arising from

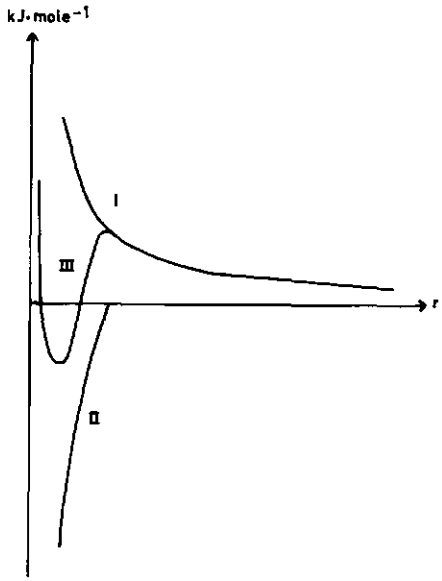


Fig. 2. I: Plot of ΔH, containing attractive and repulsive energy terms vs r, defining an arbitrary distance between protein subunits in the double disk-like oligomer. II: Plot of - TΔS vs r, in which ΔS contains the entropy terms arising from water shielding changes. III: A plot of the difference between I and II (ΔG = ΔH - TΔS) vs r. The curves I and II are hypothetical except for the inserted values for ΔH (+ 126 kJ.(mole trimer)$^{-1}$) and TΔS (156 kJ.(mole trimer)$^{-1}$) at r = 2.1 nm which represent the ΔH and ΔS for double disk-like polymerization (10). Curve II is also based on the assumption that water entropy contributions do not extend beyond a few water layers on the protein double disk surface.

v.d. Waals, coulomb, and short-range repulsive interactions. Curve II represents the water entropy contribution and is based on the assumption that only a few layers of water molecules are affected by the hydrophobic surface of the double disk. Curve III is the difference between I and II and is obtained by inserting the measured values for ΔH and ΔS (10) for the double disk-like polymerization, yielding the minimum value of ΔG. In this way, Fig. 2 visualizes, that the double disk stability arises from water entropy. In view of Fig. 2, we note that changes in the number of contacts between water and double-disk immediately affect the double-disk stability (Fig. 2): in particular, an upward shift of curve II in Fig. 2, according to Boltzmann's law, results in a wider distribution of equilibrium distances between protein subunits. Assuming ΔG > 0 for double-disks,

resulting in their complete destabilization, requires removal of watermolecules from ∿ 50% of the disk-surface. This can be estimated from the fact that removal of one of the protein subunit-surfaces from water yields an entropy change (10, 32) of ∿ 100 $J.K^{-1}$.

7.6 THE DOUBLE-DISK RNA INITIATION COMPLEX

X-ray data on TMV indicate that the bases are in contact with hydrophobic amino acid side chains of the protein subunits in a rather non-specific way (33). Three bases lie flat on the hydrophobic faces of the LR α-helix of one protein subunit which contains no aromatic amino acid, and/or they may be hydrogen bonded to adjacent amino acid side chains. The phosphate groups bind to three arginines of the RR helix of a second, adjacent protein subunit. Although a certain amino acid code involving the backbone hydrogen bonds has been suggested (34) for base recognition, this is not applicable here as follows from published X-ray data (33) showing that hydrogen bonding with the backbone is impossible. Base recognition also is unlikely to arise from hydrogen bonding between bases and side chains (35,33). Finally note that two out of three bases per protein subunit in the protein stack with adjacent bases, again indicating the non-specific nature of base-protein interaction.

There are also a number of arguments against the possibility that identical protein subunits, each capable of binding three nucleotides, can specifically accomodate a much longer sequence of nucleotides (26). It can safely be assumed, that binding based on sequence recognition can most efficiently occur with single strand (ss) sections of TMV-RNA. By comparing double disk binding to one of the polynucleotides (36) (e.g. poly A) with binding to TMV RNA, it can now be shown, that binding cannot occur with ss TMV RNA sections. The concentration of a TMV RNA region |ssRNA| capable of binding specifically to the double disk with binding energy ΔG_b^{RNA} can be calculated from [RNA] and from the Gibbs free energy for association (ΔG_{ass}) for the equilibrium ssRNA $\rightleftharpoons$ ds RNA (double strand RNA). Poly A (pA) under these conditions (pH ∿ 7, 20°C) is incapable of forming ds regions (45) so that the binding concentration of the pA region is given by [pA] with ΔG_b^{pA} defining the energy for binding pA to the double disk. Because pA has a considerably lower reaction velocity (36) for binding to TMV than TMV RNA,

$$\Delta G_b^{pA} >> \Delta G_b^{RNA} - \Delta G_{ass}, \qquad [2]$$

can now be obtained.

For the binding region of TMV RNA an average ΔG_{ass} of $\sim$ 3 KJ.(mole base pair)$^{-1}$ can be calculated (24). The difference in binding energy between sections of pA and RNA to double disk is represented by $\Delta G_b^{pA} - \Delta G_b^{RNA}$ and probably arises mainly from differences in hydrogen bonding between bases from pA and TMV RNA regions and protein. Since ΔG_{ass} is nearly optimal for TMV RNA, and the base-stacking energy to aliphatic chains in the protein presumably is lower than in ds TMV RNA and -finally-, since no particular triplet is repeated in the RNA sequence, it follows that

$$\Delta G_b^{pA} - \Delta G_b^{RNA} << -\Delta G_{ass}, \qquad [3]$$

in contradiction with | 2|. Consequently, the specificity of the protein double disk towards the RNA binding region cannot be assigned to a *sequence* of nucleotides recognized by a binding site on the double disk. Another interesting argument can be added; if a protein subunit preferentially binds to a particular triplet e.g. AAG it would bind much more effectively to a section of pA than to a chain containing less than two adenines on the first two positions of each triplet, in contrast with observation. For different homo polynucleotides this argument can be used for all triplets containing two bases in any position. From published data (26,27) it can be found that a number of TMV RNA regions (probably depending on experimental conditions) preferentially are encapsidated by the double disk; this indicates that the formation of the initiation complex may well be less specific than is usually assumed.

Summarizing, we now conclude, that the base-protein recognition in the TMV assembly process must be based on the presence of a secondary and/or tertiary structure in the RNA region capable of binding. A similar behaviour has also been suggested for RNA/coat protein interaction of the MS_2 phage (37).

The correlation times for rotational diffusion of water and double disk/RNA presumably are $< 10^{-10}$ s, and $< 10^{-7}$ s, respectively. This difference in time scale also reflects the difference in rate of translational diffusion for water and double disk/RNA, indicating that the rearrangement of water occurs at a much faster rate than the association (or dissociation) process. This is important to note, when considering the formation of the initiation complex. The interaction between RNA and double disk eventually leads to the formation of an intermediate (cage complex), in which the secondary and perhaps tertiary structure of the RNA region which must be specifically encapsidated is still intact. In the event of the cage complex formation, a number of water molecules are removed from that

protein surface which is approached by RNA. Destabilization of the double disk now occurs which results in its depolymerization mainly at the RNA boundary. Due to their hydrophobic surfaces the depolymerized protein subunits remain part of the cage complex, which retains its limited dimensions by forces causing the surface contact area between the cage-boundary and water to remain as small as possible.

Before the initiation complex can be formed a certain amount of activation energy is necessary for the formation of the cage complex (both RNA and double disk are negatively charged) as well as for the disruption of the tertiary and secondary RNA structure ($\sim$ 50 kJ) (2,19). During double-disk depolymerization, heat is released because of the positive ΔH (> 42 kJ.(mole subunit)$^{-1}$) (10) for protein subunit polymerization to trimer and double-disk oligomers. This energy can now be used for the formation of the cage complex and, more importantly, also for the disruption of the TMV RNA hairpin. It can now be understood that the specificity of RNA binding must be based on the proper geometry of the hairpin, such that enough water can be removed from the double-disk surface to destabilize it, and on the stability of the TMV RNA hairpin; it must be stable enough to have a long lifetime but not so stable that it cannot be molten. The molten RNA region now has lost its specific geometry so that the increase in water entropy upon shielding becomes the driving force for protein subunit binding in adjacent position with their arginines (33) to RNA phosphates. This assembly process is endothermic (10). The known proton uptake (38) upon assembly of TMV is necessary for neutralizing TMV RNA phosphates. It is conceivable that the flexible protein backbone regions further facilitate the RNA binding.

A similar mechanism as described above may be expected for the elongation process; then, it can be shown straightforwardly, that elongation employing double disks, proceeds more effectively, than with trimeric units (39,40). In view of E/M observations (3), the elongation step is thought to proceed from the RNA binding site to the 5' end after which the RNA segment between the binding site and the 3' end is incorporated. The latest proposed elongation mechanism implies that the unencapsidated 5' end must be pulled through the completed part of TMV before it is encapsidated (3,41). In our model, melting of RNA at its protein binding sites is energetically favoured. After the double disk falls apart and melts an RNA hairpin, we assume that the molten RNA hairpin adapts to a helix during the encapsidation by protein subunits. If this model is correct this results in both 3' and 5' ends protruding from different ends of the protein/RNA initiation complex. Electron micrographs showing the denatured 3' and 5' RNA

ends at one rod end (3,41) may be interpreted by assuming that the 5' end is folded back towards the 3' end while it is adsorbed into the protein coat of the partially completed virus rod.

7.7 DISSOCIATION OF TMV

The known procedures (42-44) which are used to dissociate TMV into protein and RNA in vitro involve rather rigorous steps (pH ∿ 11-12, detergents such as urea, sodium dodecylsulphate (SDS)). If it is assumed that pH values in plant cells under physiological conditions are approximately neutral, dissociation of TMV requires a detergent-like agent to be present in the plant cell.

TMV RNA can only enter the cytoplasm by passing through the cytoplasmic membrane containing lipids so that a better look at the influence of the lipid-like detergent, SDS, on TMV stability is interesting. The stripping of protein subunits from TMV by using 1% SDS has been known for quite some time (43). This can now easily be understood since TMV is not so stable as one would naïvely assume, due to the positive value of ΔH of > +42 kJ.(mole subunit)$^{-1}$ for interaction between the protein subunits so that the binding of each protein subunit in TMV is stabilized through interaction with three phosphate of RNA (∿ 21 kJ.(mole protein subunit)$^{-1}$) (10) and through the positive entropy term arising from water being shielded from the protein subunit surface. Because ΔH of protein-protein interaction exceeds ΔH for protein-RNA interaction, a change in ΔS, resulting in $\Delta G > 0$ for placing a protein subunit within the rod, induces a dissociation of the protein subunit. With this in mind it can be conceived that TMV falls apart just because of the hydrophobic environment when passing through the cytoplasmic membrane so that the released RNA can enter the cytoplasm.

7.8 ACKNOWLEDGEMENTS

The technical assistance of Miss M. Boerjan with the preparation of TMV and TMV protein is gratefully acknowledged. We are indebted to the Department of Virology for providing laboratory facilities. The NMR investigations at high field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO). One of the authors (J.L. d. W.) acknowledges helpful discussions with Dr. W. Norde.

7.9 REFERENCES

1. Kaper, J.M. (1975) The chemical basis of virus structure, dissociation, and reassembly, 410-418, North Holland Publishing Company, Amsterdam.
2. Butler, P.J.G., Bloomer, A.C., Bricogne, G., Champness, J.N., Graham, J., Guilley, H., Klug, A. and Zimmern, D. (1976) Proc. John. Innes Symp. 101-110.
3. Lebeurier, G., Nicolaieff, A. and Richards, K.E. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 149-153.
4. Otsuki, Y., Takebe, I., Ohno, T., Fukuda, M. and Okada, Y. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 1913-1917.
5. Nucleic Acid-Protein Recognition (1977) ed. Vogel, H.J., Academic Press, New York.
6. De Wit, J.L., Alma, N.C.M., Hemminga, M.A. and Schaafsma, T.J. Biochemistry, submitted for publication, Chapter 2.
7. De Wit, J.L. and Schaafsma, T.J. Eur. J. Biochem., submitted for publication, Chapter 4.
8. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J. (1978) J. Magn. Resonance, in press, Chapter 3.
9. Campbell, I.D., Dobson, C.M., Williams, R.J.P. and Xavier, A.V. (1973) J. Magn. Resonance 11, 172-181.
10. Lauffer, M.A. (1975) Entropy-driven Processes in Biology, Springer-Verlag, Berlin.
11. Stevens, C.L., Lauffer, M.A. (1965) Biochemistry 4, 31-37.
12. Freundlich, H. (1937) J. Chem. Phys. 41, 901-910.
13. Nemethy, B. and Scheraga, H.A. (1962) J. Phys. Chem. 66, 1773-1789.
14. Nemethy, G. and Scheraga, H.A. (1962) J. Chem. Phys. 36, 3401-3417.
15. Jaenicki, R. and Lauffer, M.A. (1969) Biochemistry 8, 3083-3092.
16. Lauffer, M.A. (1966) Biochemistry 5, 2440-2446.
17. Tanford, C. (1973) The Hydrophobic Effect, Wiley-Interscience, New York.
18. Kearns, D.R. (1976) Progr. in Nucl. Acid. Res. 18, 91-146.
19. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature 246, 40-41.
20. Kim, S.H. (1976) Progr. in Nucl. Acid. Res. 17, 182-213.
21. Graig, M.E., Crothers, D.M. and Doty, P. (1971) J. Mol. Biol. 62, 383-401.
22. Gralla, J. and Crothers, D.M. (1973) J. Mol. Biol. 73, 497-511.
23. Bina Stein, M., Crothers, D.M., Hilbers, C.W. and Shulman, R.G. (1976) Proc. Nat. Acad. Sci. U.S.A. 73, 2216-2220.

24. Crothers, D.M., Cole, P.E., Hilbers, C.W. and Shulman, R.G. (1974) J. Mol. Biol. 87, 63-88.
25. Litvak, S., Tarrago, A., Tarrago-Litvak, L. and Allende, J.E. (1973) Nature New Biol. 241, 89-90.
26. Guilley, H., Jonard, G., Richards, K.E. and Hirth, L. (1975) Eur. J. Biochem. 54, 145-153.
27. Zimmern, D. (1976) Phil. Trans. of the Royal Soc. of London B 276, 189-204.
28. Champness, J.N., Bloomer, A.C., Bricogne, G., Butler, P.J.G. and Klug, A. (1976) Nature 259, 20-24.
29. Visser, J. (1972) Adv. in Coll. and Interface Science 3, 331-363.
30. Paulsen, G. (1972) Zeitschrift für Naturforschung 27b, 427-444.
31. Caspar, D.L.D. (1963) Adv. in Prot. Chem. 18, 37-121.
32. Durham, A.C.H. and Klug, A. (1972) J. Mol. Biol. 67, 315-332.
33. Stubbs, G., Warren, S. and Holmes, K. (1977) Nature 267, 216-221.
34. Gursky, G.V., Tumanyan, V.G., Zasedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L. and Gottikh, B.P. (1977) in Nucleid Acid-Protein Recognition 189-215 (ed. Vogel, H.J.) Academic Press, New York.
35. Rich, A., Seeman, N.C. and Rosenberg, J.M. (1977) in Nucleic Acid-Protein Recognition 362-373 (ed. Vogel, H.J.) Academic Press, New York.
36. Butler, P.J.G. and Klug, A. (1971) Nature New Biol. 229, 47-50.
37. Gralla, J., Steitz, J.A. and Crothers, D.M. (1974) Nature 248, 204-208.
38. Srinivasan, S. and Lauffer, M.A. (1970) Biochemistry 9, 193-200.
39. Butler, P.J.G. and Klug, A. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 2950-2953.
40. Richards, K.E. and Williams, R.C. (1973) Biochemistry 12, 4574-4581.
41. Butler, P.J.G., Finch, J.T. and Zimmern, D. (1977) Nature 262, 217-219.
42. Durham, A.C.H. (1972) J. Mol. Biol. 67, 289-305.
43. May, D.S. and Knight, C.A. (1965) Virology 25, 502-507.
44. Buzzell, A. (1960) J. Am. Chem. Soc. 82, 1636-1641.
45. Kallenbach, N.R. and Bermans, H.M. (1977) Quart. Rev. of Biophys. 10, 137-237.

8. NMR ON LARGE BIOLOGICAL SYSTEMS: *E. COLI.* RIBOSOMES, TOBACCO MOSAIC VIRUS AND COWPEA CHLOROTIC MOTTLE VIRUS

J.L. de Wit, L.C.J. Dorssers, T.J. Schaafsma.

8.1 INTRODUCTION

Within the framework of this Thesis, it is interesting to define the scope of the application of NMR to large biological systems in relation to other non-destructive spectroscopic methods. There is a growing interest in such methods which can be used in solution to obtain information at molecular resolution about structure and dynamics of these systems, e.g. chloroplasts, ribosomes, viruses, phages, membranes, mitochondria and cells. The interest for these methods comes from the fact that one can approach the *in vivo* situation. Methods which are primarily suitable to obtain structural information such as X-ray diffraction (1) and electron microscopy (2), have limited possibilities in solution, however (3). For studies in solution more adequate methods may come from optical spectroscopy e.g. circular dichroism (4), fluorescence (5), phosphorescence (6) and laser Raman spectroscopy (7). They are dependent on the presence of chromophores, however, and measurements are often difficult due to interfering light scattering because of solution turbidity. Since the development of a completely different detection method (8), the latter problem has essentially been solved. Optical Detection of Magnetic Resonance (ODMR), as a combination of optical spectroscopy and magnetic resonance, has been applied to the study of the photosynthetic unit and its components (9). The scope of Electron Spin Resonance (ESR) has recently been extended to large biological systems through the introduction of Saturation Transfer (ST) ESR using nitroxide spin labels (11,10). Introducing a nitroxide spin label always can complicate interpretations because the spin label may disturb its environment, however.

Nuclear Magnetic Resonance (on nuclei such as ^{19}F, ^{15}N, ^{2}H, ^{1}H, ^{13}C, ^{31}P) covers a large range of possible non-destructive applications in solution which - except for their low sensitivity as compared to the methods mentioned above and limited resolution which can be overcome by selective stable isotope enrichment - have almost infinite possibilities in small biological systems (12,13). It is not necessary to assume that only solid state NMR methods such as

^{13}C induction spectroscopy are suitable for large biological systems (14,15,16). Using this method in apparently rigid systems like ivory, Schaefer et al. (17) found internal mobility on a time scale of $\sim 10^{-6}$ s which is close to that necessary for conventional FT NMR in solution.

The possibility for using conventional FT NMR in large biological systems critically depends on the presence of internal mobility. According to the Stokes-Einstein equation (18) a large biomolecule such a Tobacco Mosaic Virus (TMV) reorients slowly and from the use of dipolar relaxation equations (19) it follows that the spin-lattice relaxation time (T_1) becomes excessively long ($\sim$ 2000 s (28)) and the spin-spin relaxation time (T_2) so short that NMR resonances should be broadened beyond detection (20).

We have investigated if internal mobility is a property of other large biological systems than TMV. We will discuss ribosomes and Cowpea Chlorotic Mottle Virus (CCMV) to some extent, representing a class of biomolecules of large molecular weight.

8.2 RESULTS AND DISCUSSION

Because TMV has a large molecular weight, the NMR spectral intensity is a direct measure of the number of internally mobile nuclei (18,21), all rigid nuclei being unobservable. The number of mobile nuclei in other biological systems can also be determined in this way, the results being summarized in Table 1. Although molecular weights of these systems are smaller than that for TMV ($2\text{-}5 \times 10^6$ instead of 42×10^6), this procedure is justified since T_1 of the observable resonances is < 1 s and line widths are < 300 Hz.

Cowpea Chlorotic Mottle Virus (CCMV) is a spherical plant virus (molecular weight 4.6×10^6 and $\sim$ 14 nm diameter) consisting of 180 identical protein subunits (molecular weight 19,400) and four different RNA chains distributed over three particles (25). 100 MHz ^1H spectra of the CCMV virus are similar to those of large TMV protein oligomers. The spectral intensity of side chain resonances of CCMV, presented in Table 1, originates from internally mobile nuclei.

This internal mobility again may arise from mobility of the nuclei within the CCMV protein subunits and/or protein subunit mobility within CCMV particles. T_1's of the observable resonances are < 1 s, suggesting the presence of small angle rotational motions of the nuclei within the protein about carbon-carbon and/or carbon-proton bonds, similar to that observed for TMV (18).

TABLE 1.

Spectral intensities of backbone and side chains nuclei in protein in biological systems.

system	backbone	side chain
TMV[a]	0.07	0.17
TMV-protein rod (pH=5.3 7°C)[a]	0.1	0.1
TMV-protein rod (pH=6.2 20°C)[a]	0.65	0.65
CCMV[b]	+[e]	0.3
E.Coli ribosomes[a]	0.27	0.3
other systems[b] (phages and viruses and proteins)	d	0.05-0.3

[a] Measured as integrated spectral intensity both for ^{13}C and ^{1}H;
[b] Measured as integrated spectral intensity with ^{1}H NMR and corrected for the residual HOD;
[c] Measured by ^{1}H NMR at 100 MHz for side chain protons;
[d] Hidden under the HOD peak;
[e] Backbone is present but hard to quantify.

E. coli 70 S ribosomes (molecular weight $\sim 2.7 \times 10^6$) consist of a so so called 50 and 30 S particle with molecular weight 1.7×10^6 and 1×10^6 respectively. The 50 S particle contains 65% RNA divided over two chains (5 and 23 S) with molecular weight 0.04×10^6 and 1.1×10^6 respectively and 34 different protein subunits with molecular weights in the range 7,000 to 31,000. The 30 S particle contains 65% RNA in one chain (16 S) with molecular weight 0.6×10^6 and 21 different protein subunits with molecular weights in the range 11,000 to 31,000. The exact number of protein subunits interacting with RNA in both particles is unknown. Present knowledge about the organization and way of functioning of this intensively studied organelle is rather limited (22).

Table 1 summarizes intensities of a spectrum of 12% ^{13}C enriched *E. coli* ribosomes (see Figure 1). These spectral intensities have been determined by comparison with those of spectra of solutions with known concentrations of hemoglobine, TMV-protein, and denatured ribosomes.

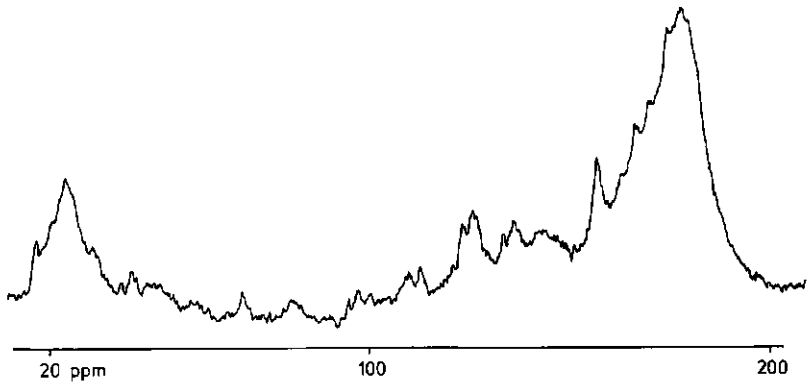


Fig. 1. Broadband noise decoupled ^{13}C NMR spectrum of *E. Coli* ribosomes at 90.5 MHz. The ribosomes are ∿ 12% ^{13}C enriched. Conditions: concn. 38 mg/ml, temp 5°C, pH = 7, 10mM MgCl , 60 mM NH CL, 3mM β- Mercaptoethanol. The ppm scale was referenced to CS assuming 125.8 ppm for β-Thr. Acquisition time plus pulse delay: 1 sec; number of accumulation 18,700; sensitivity enhancement: 33 Hz.

From ^{1}H NMR at 100 MHz it is found that the observable protons in the ^{1}H NMR spectra have line widths < 300 Hz and T_1 < 1 s. From this figure, a line width of < 150 Hz can be deduced for ^{13}C using the equations for dipolar relaxation (21) for ^{1}H and ^{13}C. From the line width $\tau_c < 5.10^{-8}$s can be calculated using the equations for dipolar relaxation (19). Using the molecular weight of the ribosome in the absence of internal mobility, a line width of > 4 kHz is predicted from the Stokes-Einstein and dipolar relaxation equations, which leads us to conclude that the observable nuclei must be internally mobile.

The considerations used in previous Chapters when discussing internal mobility in TMV and its protein oligomers cannot be used because of lack of data on protein-RNA and protein-protein interactions and tertiary structure of the protein subunits in the ribosomes. The spectrum of Fig. 1. reveals a carbonyl region (10-40 ppm), nucleotide resonances (40-100 ppm), ribose resonances (100-130 ppm), C_α resonances (130-150 ppm), and aliphatic resonances (150-200 ppm). Although the ribosomes purified according to Fahnestock et al. (23) retain their normal activity, these RNA resonances could be a result of RNA breakdown during ribosomal preparation.

Finally, we have carried out 100 MHz ^{1}H measurements on Brome Mosaic Virus (BMV), Broad Bean Mottle Virus (BBMV), Alfalfa Mosaic Virus (AMV), Cowpea Mosaic Virus (CpMV), a number of T- viruses, MS2 virus, f2 virus, Q_{β} virus, ϕX174 virus and Adenovirus. The results are summarized in Table 1, indicating that in all cases an appreciable of ^{1}H nuclei are observed and therefore are internally mobile.

Along the same lines as found for TMV, the observation of ^{1}H resonances is a clear indication that ^{13}C NMR can also be fruitfully applied to a large host of biosystems, since they all have internal mobility as a common property. We have not not carried out ^{13}C measurements on the collection of viruses mentioned above, however.

Morrison et al. (24) studied ribosomal proteins subunits obtained from *E. coli* ribosomes with ^{1}H NMR. Our results indicate that when selective ^{13}C enrichment is used (e.g. a single ^{13}C protein subunit in non-enriched ribosomes) a variety of these protein subunits can be studied in their natural environment.

8.3 ACKNOWLEDGEMENTS

We are indebted to the Department of Molecular Biology and Virology for hospitality and laboratory facilities Dr. B.M. Verduin kindly provided us with CCMV, BMV and BBMV. The NMR investigations at high magnetic field (8.4 T) were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

8.4 REFERENCES

1. Stubbs, G., Warren, S. and Holmes, K. (1977) Nature 267, 216-221.
2. Lake, J.A., Sabatini, D.D. and Nonomura, Y. (1974) Ribosomes (eds Nomura, M., Tissières, A., Lengyel, P.) 543-557.
3. Holmes, K.C. (1976) Trends Biochem. Sci. 1, 183-184.
4. Schubert, D. and Krafczyk, B. (1969) Biochim. Biophys. Acta 188, 155-157.
5. Taniguchi, M. and Taniguchi, T. (1975) Biochim. Biophys. Acta, 1-17.
6. Konev, S.V. (1967) Fluorescence and Phosphorescence of Proteins and Nucleic Acids, Plenum Press, New York.
7. Thomas, G.J.Jr., Prescott, B., McDonald-Ordzie, P.E. and Hartman, K.A.

(1976) J. Mol. Biol. 102, 103-124.
8. Maugh, T.H.II. (1977) Science 188, 38-40.
9. Van der Bent, S.J., Schaafsma, T.J. and Goedheer, J.C. (1976) Biochem. Biophys. Res. Comm. 71, 1147-1152.
10. Thomas, D.D., seidel, J.C., Hyde, J.S. and Gergely, J. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 1729-1733.
11. Hemminga, M.A., de Jager, P.A. and de Wit, J.L. (1977) Biochem. Biophys. Res. Comm. 79, 635-639.
12. Wüthrich, K. (1976) NMR in Biological Research peptides and proteins North-Holland Publishing Company, Amsterdam.
13. Dwek, R.A. (1973) Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems, Clarendon Press, Oxford.
14. Pines, A., Gibby, M.G. and Waugh, J.S. (1973) J. Chem. Phys. 569-690.
15. Urbina, J. and Waugh, J.S. (1974) Proc. Nat. Acad. Sci. U.S.A. 71, 5062-5067.
16. Schaefer, J., Stejskal, E.O. and Buchdahl, R. (1977) Macromolecules 10, 384-405.
17. Schaefer, J. and Stejskal, E.O. (1976) J. Am. Chem. Soc. 98, 1031-1032.
18. De Wit, J.L., Alma, N.C.M., Hemminga, M.A., Schaafsma, T.J. Biochemistry submitted for publication, chapter 2.
19. Lyerla, J.R. and Levy, G.C. (1974) in Topics in Carbon-13 NMR Spectroscopy I, 79-143, Wiley-Interscience, New York.
20. Torchia, D.A. and Van der Hert, D.L. (1976) J. Mol. Biol. 104, 315-321.
21. De Wit, J.L., Hemminga, M.A. and Schaafsma, T.J. (1978) J. Magn. Resonance, in press, chapter 3.
22. Ribosomes (1974) eds. Nomura, M., Tissières, A. and Lengyel, P., Cold Spring Harbor Laboratory.
23. Fahnestock, S., Erdmann, V. and Nomura, M. (1974) in Methods in Enzymology, XXX, 559-560 (eds Moldave, K. and Grossman, L.) academic Press, New York.
24. Morrison, C.A., Bradbury, E.M. and Carrett, R.A. (1977) FEBS Lett. 81, 435-439.
25. Bancroft, J.B. (1976) Adv. Virus Res. 16, 99.

LIST OF ABBREVIATONS

AMV	Alfalfa Mosaic Virus
BMV	Brome Mosaic Virus
BBMV	Broad bean Mottle Virus
β_o	oscillation angle
CCMV	Cowpea Chlorotic Mottle Virus
CD	convolution difference
CpMV	Cowpea Mosaic Virus
γ_H	gyromagnetic ratio of a proton
γ_C	gyromagnetic ratio of a carbon
D	dipolar interaction parameter
ds	double strand
Δ	change
E	extinction coefficient
ESR	electron spin resonance
FT NMR	fourier transform nuclear magnetic resonance
G	Gibbs free energy
ΔG_b^{pA}	Gibbs free energy for binding pA to double disk
ΔG_{ass}	Gibbs free energy for association from single to double strand RNA
ΔG_b^{RNA}	Gibbs free energy for binding RNA to double disk
ΔG_a	Gibbs free energy for activation
η	viscosity
H	enthalpy
$\hbar$	Planck's constant/2π
HR	high resolution
J	Joule;scalar coupling constant
$J(\omega)$	spectral density function
K	constant
k	Boltzmann's constant
kJ	kiloJoule
μ_o	magnetic permeability

M	molecular weight
N	number of nuclei
n_D	number of detectable protein subunits
n_N	number of non-detectable protein subunits
ORD	optical rotation dispersion
pA	poly A
polyA	polyadenylic acid
ppm	parts per million
R	gas constant
RNA	ribonucleic acid
r_{CH}	carbon-proton distance
r_{HH}	proton-proton distance
σ	chemical shift
$\Delta\sigma$	chemical shift range
S	sedimentationcoefficient in Svedberg units
s	second
SDS	sodium dodecyl sulfate
S/N	signal to noise ratio
ss	double strand
ST	saturation transfer
τ_c	correlation time
τ_g	correlation time describing rotation of a tetrahedral group about its carbon-carbon bond
τ_g^i	τ_g with i specifying the carbon-carbon bond
τ_p	correlation time for internal mobility of a protein
τ_R	correlation time for rotational motion of a sphere
τ_m	residence life time
τ_i	correlation time
T	absolute temperature; Tesla
T_1	spin-lattice relaxation time
T_2	spin-spin relaxation time
$1/\pi T_2$	lorentzian line width at half height
TMV	Tobacco Mosaic Virus
Tris	Tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
$tRNA_2^{Glu}$	glutamic acid transfer RNA
$\bar{v}$	partial specific volume
W	Watt
ω_H	larmor frequency of a proton
ω_C	larmor frequency of a carbon

APPENDIX I

TABLE I

Amino acid sequence for the protein of TMV Vulgare

Number	Residue	Number	Residue	Number	Residue
-	Acetyl	27	Cys	54	Pro
1	Ser	28	Thr	55	Ser
2	Tyr	29	Asn	56	Pro
3	Ser	30	Ala	57	Gln
4	Ile	31	Leu	58	Val
5	Thr	32	Gly	59	Thr
6	Thr	33	Asn	60	Val
7	Pro	34	Gln	61	Arg
8	Ser	35	Phe	62	Phe
9	Gln	36	Gln	63	Pro
10	Phe	37	Thr	64	Asp
11	Val	38	Gln	65	Ser
12	Phe	39	Gln	66	Asp
13	Leu	40	Ala	67	Phe
14	Ser	41	Arg	68	Lys
15	Ser	42	Thr	69	Val
16	Ala	43	Val	70	Tyr
17	Trp	44	Val	71	Arg
18	Ala	45	Gln	72	Tyr
19	Asp	46	Arg	73	Asn
20	Pro	47	Gln	74	Ala
21	Ile	48	Phe	75	Val
22	Glu	49	Ser	76	Leu
23	Leu	50	Glu	77	Asp
24	Ile	51	Val	78	Pro
25	Asn	52	Trp	79	Leu
26	Leu	53	Lys	80	Val

Number	Residue	Number	Residue	Number	Residue
81	Thr	107	Thr	133	Ile
82	Ala	108	Leu	134	Arg
83	Leu	109	Asp	135	Gly
84	Leu	110	Ala	136	Thr
85	Gly	111	Thr	137	Gly
86	Ala	112	Arg	138	Ser
87	Phe	113	Arg	139	Tyr
88	Asp	114	Val	140	Asn
89	Thr	115	Asp	141	Arg
90	Arg	116	Asp	142	Ser
91	Asn	117	Ala	143	Ser
92	Arg	118	Thr	144	Phe
93	Ile	119	Val	145	Glu
94	Ile	120	Ala	146	Ser
95	Glu	121	Ile	147	Ser
96	Val	122	Arg	148	Ser
97	Glu	123	Ser	149	Gly
98	Asn	124	Ala	150	Leu
99	Gln	125	Ile	151	Val
100	Ala	126	Asn	152	Trp
101	Asn	127	Asn	153	Thr
102	Pro	128	Leu	154	Ser
103	Thr	129	Ile	155	Gly
104	Thr	130	Val	156	Pro
105	Ala	131	Glu	157	Ala
106	Glu	132	Leu	158	Thr

The sections 21 to 31, 40 to 53, 79 to 96, 103 to 112 and 114 to 134 are in the α-helix configuration (Stubbs, G., Warren, S. and Holmes, K.(1977) Nature, 267, 216 - 221.).

This listing is obtained from the Centre du Documentation de C.N.R.S.

APPENDIX 2

TABLE 2

Chemical shifts of carbon nuclei in ppm relative to CS_2.

carbon nucleus[a]	chemical shift[b]	carbon nucleus[a]	chemical shift[b]
Ile C_{δ}	181,3	Cys C_{β}	155,6
Ile C_{γ_2}	177,7	Asp C_{β}	155,2
Ala C_{β}	176,3	Lys C_{ε}	153,4
Val C_{γ_2}	176,0	Leu C_{β}	153,1
Val C_{γ_1}	175,0	Arg C_{δ}	152,1
Thr C_{γ}	174,1	Gly C_{α}	150,3
Leu C_{δ_2}	171,7	Pro C_{δ}	146,9
Lys C_{γ}	170,9	Ala C_{α}	143,5
Leu C_{δ_1}	170,6	Cys C_{α}	142,0
Arg C_{γ}	168,7	Asn C_{α}	141,9
Pro C_{γ}	168,5	Asp C_{α}	141,4
Leu C_{γ}	168,5	Leu C_{α}	140,0
Ile C_{γ_1}	167,7	Arg C_{α}	139,6
Lys C_{δ}	166,6	Gln C_{α}	139,3
Trp C_{β}	166,6	Lys C_{α}	138,9
Gln C_{β}	166,0	Glu C_{α}	138,9
Arg C_{β}	165,0	Trp C_{α}	139,6
Glu C_{β}	164,8	Phe C_{α}	137,6
Pro C_{β}	163,6	Tyr C_{α}	137,4
Lys C_{β}	162,7	Ser C_{α}	137,3
Val C_{β}	162,6	Ile C_{α}	134,5
Gln C_{γ}	161,6	Val C_{α}	134,1
Glu C_{γ}	159,1	Pro C_{α}	131,9
Asn C_{β}	157,6	Thr C_{α}	133,7
Tyr C_{β}	156,8	Ser C_{β}	131,6
Phe C_{β}	156,0	Thr C_{β}	125,8
Ile C_{β}	155,8	Trp C_{γ}	84,0

carbon nucleus[a]	chemical shift[b]	carbon nucleus[a]	chemical shift[b]
Trp C_{ζ_2}	80,9	Tyr C_{γ}	64,9
Tyr $C_{\varepsilon_1+\varepsilon_2}$	77,4	Phe $C_{\delta_1+\delta_2}$	64,0
Trp C_{ε_3/η_2}	74,4	Phe $C_{\varepsilon_1+\varepsilon_2}$	63,5
Trp C_{ε_3/η_2}	73,4	Tyr C_{δ_2}	62,4
Trp C_{δ_3/δ_1}	70,8	Tyr C_{δ_1}	62,4
Trp C_{δ_3/δ_1}	67,7	Phe C_{γ}	56,4
Trp C_{δ_2}	66,2	Trp C_{ε_2}	56,4
Phe C_{ζ}	65,6	Tyr C_{ζ}	38,2
		Arg C_{ζ}	36,1

[a] Names of carbon nuclei are according to the IUPAC IUB rules.
[b] Identical to those of James, T.J. (1975) Nuclear Magnetic Resonance in Biochemistry 246-247, Academic Press, New York.

REFERENCES

1. Nuclear Magnetic Resonance in Biochemistry: Applications to Enzyme Systems, Dwek, R.A. (1975) Clarendon Press, Oxford.
2. Lyerla, J.R.Jr. and Levy, G.C. (1974) Topics in Carbon-13 NMR spectroscopy Vol. I, ed. by Levy, G.C., Wiley-Interscience, New York.
3. Hallenga, K. and Koenig, S.H. (1976) Biochemistry 15, 4255-4264.
4. Tao, T. (1969) Biopolymers 8, 609-632.
5. Noack, F. (1971) in NMR, Basic Principles and Progress, Vol. 3, 83-144, eds Diehl, P., Fluck, E. and Kosfeld, R., Springer-Verlag, Berlin.

APPENDIX 4

The treatment of internal mobility in this Thesis has mainly been based on equations derived for internal rotation of a methyl group superimposed on isotropic overall motion.

The spin-lattice ($1/T_1^{HH}$) and spin-spin ($1/T_2^{HH}$) dipolar relaxation rates for a ^1H - ^1H pair are given by (1),

$$\frac{1}{T_1^{HH}} = 0.3(\mu_o/4\pi)^2 \frac{\gamma_H^4 h^2}{r_{HH}^6}\left[J(\omega_H) + 4J(2\omega_H)\right] \qquad [1] \ ,$$

$$\frac{1}{T_2^{HH}} = 0.15(\mu_o/4\pi)^2 \frac{\gamma_H^4 h^2}{r_{HH}^6}\left[3J(0) + 5J(\omega_H) + 2J(2\omega_H)\right] \qquad [2] \ ,$$

$J(\omega_H)$, $J(2\omega_H)$ and $J(0)$ can be given by,

$$J(\omega) = 0.25\left[\frac{\tau_c}{1 + \omega^2\tau_c^2} + \frac{3\tau_i}{1 + \omega^2\tau_i^2}\right] \qquad [3] \ ,$$

with $$\tau_i^{-1} = \tau_c^{-1} + \tau_g^{-1} \qquad [4] \ ,$$

in this case τ_g is the internal rotational correlation time describing random jumps of the ^1H - ^1H vector about the symmetry axis of the methyl group.

$1/T_1^{CH}$ and $1/T_2^{CH}$ for dipolar relaxation of a ^{13}C - ^{1}H pair is given by (2),

$$\frac{1}{T_1^{CH}} = 0.1(\mu_o/4\pi)^2 \frac{\gamma_H^2\gamma_C^2h^2}{r_{CH}^6}\left[\frac{1}{9}\chi_1 + \frac{8}{27}\chi_2 + \frac{15}{27}\chi_3\right] \quad [5],$$

$$\frac{1}{T_2^{CH}} = 0.05(\mu_o/4\pi)^2 \frac{\gamma_H^2\gamma_C^2h^2}{r_{CH}^6}\left[\frac{\phi_1}{9} + \frac{8\phi_2}{27} + \frac{16\phi_3}{27}\right] \quad [6],$$

where

$$\chi_i = \frac{\tau_i}{1 + (\omega_C-\omega_H)^2\tau_i^2} + \frac{3\tau_i}{1 + \omega_C^2\tau_i^2} + \frac{6\tau_i}{1 + (\omega_H+\omega_C)^2\tau_i^2} \quad [7],$$

$$\phi_i = \chi_i + 4\tau_i + \frac{6\tau_i}{1 + \omega_H^2\tau_i^2} \quad [8],$$

and

$$i = 1,2 \text{ and } 3 \text{ with } \tau_1 = \tau_c \quad [9],$$

$$\frac{1}{\tau_2} = \frac{1}{\tau_c} + \frac{1}{6\tau_g} \quad [10],$$

$$\frac{1}{\tau_3} = \frac{1}{\tau_c} + \frac{2}{3\tau_g} \quad [11],$$

τ_g in this case is the correlation time describing reorientation of the ^{13}C-^{1}H pair about its symmetry axis.

REFERENCES

1. Woessner, D.E. (1962) J. Chem. Phys. 36, 1-4.
2. Doddrell, D., Glushke, V. and Allerhand, A. (1972) J. Chem. Phys. 56, 3683-3689.

SUMMARY

This Thesis describes the application of conventional ^{13}C and ^{1}H high resolution Fourier Transform Nuclear Magnetic resonance (HR FT NMR) to Tobacco Mosaic Virus (TMV) and its protein oligo- and polymers and some other large biological systems. The rod-like TMV consists of 2200 identical protein subunits protecting one RNA chain (molecular weight 42 x 10^6). The most important protein oligo- and polymers are the trimer (molecular weight $\sim$ 50,000), rod-like polymer (molecular weight > 10^6) and the double disk-like oligomer (molecular weight $\sim$ 0.6 x 10^6). This study could be carried out because these large biomolecules exhibit internal mobility. Apart from rotation of TMV and protein oligo- and polymers themselves several types of such internal mobility can be distinguished: rotational motions about carbon-carbon bonds in the polypeptide chain (backbone and side chains) within the protein subunits characterized by a rotational diffusion correlation time $\tau_g < 5 \times 10^{-10}$s, assumed to correspond to small-amplitude rotation extending over $\sim 40^o$; translational and rotational motions of protein subunits within the protein oligomers about one or (more probably) two mutually perpendicular axes with a correlation time $\tau_p < 4 \times 10^{-8}$s; temperature dependent rotational motions over a full 2π angle of both backbone and side chain about at least two carbon-carbon bonds of the section 57 to 62 of the polypeptide chain in the virus and of the section 57 to 62 and in the section 90 to 120 in the double disk-like oligomer on a time scale < 10^{-7}s. The section 90-120 is known to constitute the RNA binding site. From the effect of proton binding to the rod-like polymers on internal mobility, it turns out that at least one aspartic. one arginine and probably a glumatic acid are involved in the anomalous titration behaviour of TMV protein rod-like polymers. These amino acids probably belong to the so called carboxyl cage which, based in our results, is expected to be hydrophobic. After the addition of the first proton the carboxyl cage is still incomplete, a result which follows from the fact that protein subunits with one bound proton, are still detectable.

Based on the finding that protein subunits are mobile, it is shown that the stability of the double disk-like oligomer solely arises from entropy

increase upon shielding of hydrophobic protein surfaces from water during the protein polymerization process to double disks, no specific protein-protein interactions being present. The positive ΔH and ΔS for this polymerization process led us to conclude that, if enough water is removed from the double disk protein surface, the double disk destabilizes and dissociates. From this conclusion a model for the assembly of TMV from double disk and TMV RNA has evolved.

The first step in the assembly process, the formation of the initiation complex, is based on the double disk which specifically recognizes an RNA region. Summarizing, the model implies, that recognition takes place with an intact RNA hairpin; that the double disk dissociates at least at that surface which is approached by the RNA hairpin because of diminished water contact; that the heat released during double disk dissociation, because of the positive ΔH, is available for melting the RNA hairpin; that the initiation complex then can be completed and, finally, that the specificity arises from size and stability of the RNA hairpin, i.e. its secondary structure. It is reasonable to suppose that initiation is completed with both 3' and 5' end of RNA protruding from different sides of the initiation complex. The elongation process can be described similar as the formation of the initiation complex.

Also a simple model for TMV dissociation, under physiological conditions in protein and RNA upon entering the plant cell, is presented. It is suggested that TMV dissociates when passing the cytoplasmic menbrane.

In a small excursion to other large biological systems (plant viruses, phages, ribosomes) we show that the type of expi rements described in this Thesis can be extended to many large biological systems. In this way proteins can be studied in their natural environment, close to the *in vivo* situation. Finally this Thesis shows that it is relatively simple to enrich these systems with stable isotopes against low material costs. For ^{13}C NMR measurements TMV was enriched with ^{13}C up to 12-15%.

SAMENVATTING

Dit proefschrift beschrijft de toepassing van ^{13}C en ^{1}H kernspin resonantie (NMR) aan Tabaksmozaiek virus (TMV) en zijn eiwit oligo- en polymeren. Het staafvormige TMV bestaat uit 2200 identieke eiwit eenheden die het RNA beschermen. Het totale molekuulgewicht van het virus is 42×10^6. De belangrijkste eiwit oligo- en polymeren zijn de trimeer (molekuulgewicht $\sim$ 50,000), staafvormige polymeer (molekuulgewicht $> 10^6$) en het "double disk" vormige oligomeer (molekuulgewicht $\sim 0.6 \times 10^6$). Dit onderzoek was mogelijk omdat zowel virus als eiwit oligo- en polymeren intern beweeglijk zijn. Behalve de rotatie van TMV en eiwit oligo- en polymeren om hun eigen assen, kan men meerdere typen interne beweeglijkheid onderscheiden in deze systemen: rotaties om de koolstof-koolstof bindingen van de polypeptide keten correlatietijd voor rotatiediffusie $\tau_g < 5 \times 10^{-10}$s corresponderend met rotatie over hoeken van $\sim 40^o$; translatie en rotatie diffusie van eiwit eenheden in de eiwit oligomeren om één of- wat meer waarschijnlijk is twee onderling loodrechte assen met eencorrelatietijd $\tau_p < 4 \times 10^{-8}$s; temperatuur-afhankelijke rotaties over een hoek van 360^o om tenminste twee koolstof-koolstof bindingen in sekties van de polypeptide keten (skelet en zijketens) van aminozuur 57 tot 62 in het virus en van 57 tot 62 en in de sektie 90 tot 120 in de "double disk" vormige oligomeer op een tijdschaal $< 10^{-7}$s. De sektie 90 tot 120 is bekend als de RNA bindingsplaats.

Uit het effect van proton-binding aan het staafvormige polymeer op de interne beweeglijkheid kunnen we laten zien dat tenminste één aspargine zuur, één arginine en waarschijnlijk een glutamine zuur betrokken zijn bij het afwijkende titratiegedrag van staafvormige eiwit polymeren. Deze aminozuren bevinden zich waarschijnlijk in de "carboxyl cage" die naar de resultaten doen vermoeden onvolledig en hydrophoob is.

De eiwiteenheden in de "double disk" oligomeer blijken beweeglijk te zijn zodat de stabiliteit van dit oligomeer berust op de entropie toename gedurende het eiwit polymerisatie proces tot "double disks" waarbij hydrofoob eiwit oppervlak uit water verdwijnt. Specifieke eiwit-eiwit interakties zijn zijn dan ook afwezig. De positieve ΔH en ΔS, voor dit polymerisatie proces

geven aan dat als voldoende water het kontakt verliest met het "double disk" eiwit opppervlak, de double disk instabiel wordt en dissocieert. Deze konklusie vormt de basis voor een model voor assemblage van eiwit en TMV RNA.

De eerste stap in het assemblage proces, het ontstaan van het initiatie complex, berust naar reeds bekend was, op het herkennen van een specifiek stuk RNA door de "double disk". De assemblage houdt dan, kort samengevat, in dat deze herkenning plaatsvindt aan een intacte RNA hairpin; dat de double disk in ieder geval dissocieert aan de kant waar de RNA hairpin nadert; dat de warmte die bij double disk dissociatie vanwege de positieve ΔH vrijkomt, beschikbaar is om de RNA hairpin te smelten; dat dan het initiatie komplex gevormd kan worden en tenslotte dat de specificiteit berust op de grootte en stabiliteit van de RNA hairpin. Men kan redelijkerwijs veronderstellen dat het initiatie komplex ontstaat met het 3' en 5' RNA uiteinde aan verschillende kanten van dit komplex. De elongatie kan op dezelfde manier beschreven worden als het ontstaan van het initiatie komplex.

Gebruik makend van de in dit proefschrift beschreven ideeën over eiwit-eiwit interaktie wordt een simpel model voor TMV dissociatie onder fysiologische omstandigheden voorgesteld, berustend op het feit dat het mantel-eiwit dissocieert wanneer TMV het cytoplasmatisch membraan passeert. Een klein uitstapje naar andere biologische systemen (planten-virussen, fagen, ribosomen) laat verder zien dat de in dit proefschrift beschreven toepassingen ook voor andere grote systemen van nut kunnen zijn met als belangrijk voordeel, dat eiwitten dan in hun natuurlijke omgeving die de *in vivo* situatie benaderd bestudeerd kunnen worden. Dit proefschrift beschrijft verder, dat verrijking met stabiele isotopen gemakkelijk is en wat betreft de isotoopkosten niet duur hoeft te zijn. De ^{13}C NMR experimenten zijn uitgevoerd met 12-15% ^{13}C verrijkt TMV.