Poly Adenosine Monophosphate Ribonucleic Acid Mediated Tobacco Mosaic Virus-like Rod Self-Assembly as an Alternative to Tobacco Mosaic Virus Rod Use in Future Nanotechnology Applications

By: Joshua Lucate, Chemistry Department, Mcgill University, Montreal

April, 2015

This is a thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Masters in Science

Table Of Contents

Abstract

Tobacco Mosaic Virus (TMV) is an RNA plant virus which in its native form is made up of an 18 nm diameter protein capsid that self-assembles around a single stranded RNA (ssRNA) genome. Due to its unique features TMV has attracted recent interest as a template for producing a variety of novel materials and devices with highly desired optical and electronic properties for use in a broad range of fields including memory storage and plasmonic applications. As TMV based technology develops significant challenges must be overcome in the future including infectivity and self-assembly issues. We propose using poly A RNA as an alternative to the native TMV RNA to form poly A containing TMV-like rods as poly A overcomes many of the possible future and current issues and furthermore is potentially more versatile than TMV RNA for use in the field of nanotechnology. Here we present evidence which suggest that poly A is indeed able to form TMV-like rods through interaction with TMV coat protein. Furthermore we present visual evidence using TEM and AFM to confirm that the rods formed using poly A as the nucleating RNA strand have the same shape and dimensions of TMV rods. The possibility and utility of tethering poly A containing TMV-like rods to a gold surface will also be discussed.

Tabac virus de la mosaïque (TMV) est un virus végétal à ARN qui, dans sa forme native est constitué d'un 18 nm protéine de capside de diamètre qui se auto-assemble autour d'un ARN simple brin génome. Grâce à ses caractéristiques uniques TMV a suscité l'intérêt récent comme modèle pour produire une variété de nouveaux matériaux et dispositifs avec des propriétés optiques et électroniques très recherchées pour une utilisation dans un large éventail de domaines, y compris le stockage de la mémoire et les applications plasmoniques. Comme la technologie à base de TMV développe d'importants défis doivent être surmontés dans le futur, y compris les questions de l'infectiosité et d'auto-assemblage. Nous proposons d'utiliser l'ARN poly A titre d'alternative à l'ARN de TMV natif pour former poly A contenant des tiges tels que TMV-poly A surmonte beaucoup du futur possible et les problèmes actuels et est

potentiellement plus polyvalent que l'ARN de TMV pour une utilisation dans le domaine de l'outre la nanotechnologie. Ici, nous présentons des preuves qui suggèrent que poly A est en effet capable de former des barres TMV-comme par l'interaction avec TMV protéine d'enveloppe. En outre, nous présentons des preuves visuelles à l'aide TEM et l'AFM pour confirmer que les tiges formé en utilisant poly A comme le brin d'ARN nucléation ont la même forme et les dimensions des tiges TMV. La possibilité et l'utilité d'atacher poly A contenant tiges TMV-comme d'une surface d'or sera également discutée.

Acknowledgements

I would like to state what a pleasure it has been for me to undertake this research. We have only seen the tip of the iceberg in terms of the effect which nanotechnology will have on our everyday lives and even now I absolutely believe that Tobacco Mosaic Virus will play a part in developing future technologies. I have been fortunate to be able to work in this exciting field and I hope that my contribution will inspire future researchers in this area. I would like to dedicate this thesis to my son, Jacob, every day with you brings so much joy to my life. I would like to start by thanking my partner in crime, Denise Ma, without whose support I would not have finished writing this thesis. A huge thank you has to go to my supervisor, Prof. Amy S. Blum who has been instrumental at every step along the way, from initial concept to troubleshooting to editing. Prof. Blum always believed in this research even when it wasn't looking promising and her belief inspired me to continue on the days when nothing was working and I wanted to light my bench aflame. I would also like to thank my evaluation committee, Prof. Gonzalo Cosa, Prof. Hanadi Sleiman and Prof. Karine Auclair not only for their useful suggestions concerning my research but especially for the critical role that they played in helping me develop my thought patterns. I would like to thank Dr. Katherine Castor for her help with the CD measurements as well as Prof. Sleiman for the use of her equipment. The DNA cubes used in this research were graciously constructed and purified by Dr. Chris Mclaughlin and Justin Conway. The mass spec measurements were carried out by Dr. Alexander Wahba. Finally, I would like to thank the members of the Blum group, especially Dr. Omar Zahr, who not only helped me in creating the TMV-CP production protocol used here but also helped me produce and purify protein on a regular basis. Last by not least I would like to thank Serene Bayram, also of the Blum group, for her help in producing and purifying some of the TMV-CP which was used in these experiments.

Chapter 1, Introduction to RNA Mediated Assembly of Tobacco Mosaic Virus

Introduction to Tobacco Mosaic Virus

Tobacco Mosaic Virus (TMV) is a plant virus which in nature affects mostly plants in the nightshade family, including tobacco and tomato. In its native form TMV is rod shaped and is derived from a single strand of RNA which is encapsulated by 2130 identical coat proteins (CP) self assembled into a right handed helix with a diameter of 18 nm, a length of 300nm and a 4 nm hollow inner channel[1]. The TMV helix has a pitch of 23 Angstroms with each period containing 16.333 CP with each CP binding to 3 nucleotides, making the TMV RNA 6395 nucleotides long[2]. A large amount of fundamental research has been done on TMV in the last 50 years making the behavior of the virus, the structure of the various components of the virus, as well as its self assembly properties and underlying mechanisms fairly well understood[1–8]. TMV has a rigid, robust and homogenous self assembled structure which stays virtually unchanged over a wide range of temperatures, pH and solution conditions. In addition to this stability, TMV has a large surface to volume ratio and a rather unique size and aspect ratio which make it applicable to a variety of potential applications. Due to these physical characteristics in conjunction with its well understood self-assembly properties TMV continues to become an increasingly attractive candidate for use in the emerging fields of Nanotechnology and Biotechnology.

The intact rod form of TMV, containing both RNA and CP is the most well known and currently the most studied. However, although the TMV RNA on its own is not very interesting, the CP itself is capable of forming various phases on its own, without the RNA, some of which also have interesting properties for nanotechnology applications. The TMV-CP aggregates to form different self-assembled structures depending on solution conditions. The primary dependence is on pH and ionic strength, as can be seen in the phase map in Figure 1.1. It is important to note that the phase map merely indicates the region at which one TMV state dominates over the others. Since the processes which govern TMV self assembly are

dynamic[5,6,9] , at any given state it is possible to observe the other states as a minor contribution to the overall ensemble, but they become rarer the further one is from the boundary.

1.1 Assembly Properties of TMV Coat Protein

Although there are other phases, the physiologically relevant species are the ones which will be described here since they are the only ones related to RNA mediated TMV rod formation. They are the rod, the disk, and small aggregated. The rod is the dominant species at low pH, and is nearly identical in structure to the native viral capsid. Under normal conditions the rod phase dominates below pH 6.5, although the pH at which the rod phase begins to dominate depends to a certain degree on ionic strength and temperature[5,10] The rods formed in the absence of RNA are also helical in nature and identical in diameter to those formed with RNA but the length of the RNA free rods are mostly random, with very little length control possible. At higher pH, extended assemblies which superficially look very similar to rods can be found at high ionic strengths, however these extended assemblies are not helical. Extended assemblies at higher pH and high ionic strength consist of stacked achiral disks and are thus planar. Helical assemblies can only be obtained with RNA or with a pH triggered switch which becomes activated by the addition of H^+ at around pH 6.5[5,11,12]. There is a second form of helical rod which can form only in the absence of RNA which contains 17.33333 CP per period of helix instead of 16.33333[13]. Not a lot is known about the second form, nor is it particularly useful however it does demonstrate that the added stability of TMV containing RNA does reduce the possible helical configurations of the CP.

The disk dominates at neutral pH as well as at higher ionic strength in solutions of higher pH. Importantly, the disk is the dominant state under normal physiological conditions which the TMV RNA would encounter in nature[6]. The TMV disk, or 20S protein is a planar, washer shaped 2 layer disk with an 18 nm diameter and a 4 nm hole at the center, with each layer containing 17 subunits. Both layers of the disk have the same orientation and thus the disk is polar[7,11,14–16]. The disk is by far the most controversial form of TMV-CP with many disagreements in the literature over the years not only as to whether it is the nucleating species in TMV rod

formation[1,4,6,7,17–21] but also over whether it is the primary species responsible for rod elongation[1,6,17,22–26]. The disk is on its own, is useless in that it must be converted to a different, helical form before it can protect RNA from degradation. Yet, it is a higher order structure, which is designed to be exactly 2 layers with larger assemblies prevented from forming under physiological conditions. From an evolutionary viewpoint such a structure likely would not exist unless it was critical to virus replication as if it were not at all involved in virus assembly it would then interfere with efficient virus replication.

Finally, at high pH and low ionic strength the CP exists as a mixture of low order aggregates, monomers, trimers, tetramers etc. and is often referred to as the 4s or A protein. The exterior of the CP is negatively charged at high pH, which is why the disk is electrostatically prevented from forming, except at higher ionic strength where the screening is sufficient to allow the CP to approach each other. Although disks can be made to form at higher pH by raising the ionic strength of solution, it is not clear whether the RNA can assemble rods or not as there are no reported cases of assembly above pH 7.5. Previously formed rods are stable above pH 7.5, but since the rate of formation drops off as the pH is raised,[27] if TMV could be assembled above pH 7.5 it would certainly proceed slowly.

Although pH and ionic strength have the largest impact on TMV-CP self-assembly, TMV-CP concentration also plays a minor role, with assembly favored in higher concentration environments. Temperature is another factor which has an effect on CP self assembly. Although temperature does not always have an effect on the state of the CP[5] it is still a critical component both to the self assembly of TMV and its CP as viral self assembly has been shown to be endothermic and is therefore driven by changes in entropy upon assembly.

The main driving force for the self assembly of the various states of the virus aggregation is hydrophobic interactions[6,9,10,12]. Thus, higher order structures form more easily at higher temperature. Experiments involving RNA and CP are usually done at room temperature, as at lower temperatures the CP has a preference to form A-protein[10,28]. It is still possible to obtain disks and rods at lower temperatures, however the range of pH, ionic strength and CP concentration over which the larger forms of CP can be obtained becomes

more limited, especially in the case of the disk[5,10]. Since the reaction between RNA and CP is endothermic it should proceed more effectively at higher temperatures. However, while the intact TMV rod containing the RNA is highly stable under almost any conditions[29,30], the CP on its own can begin to lose stability and precipitate at temperatures higher than 35 ^{o}C and thus the optimum assembly temperature is around 30 ^{o}C[12,31–33].

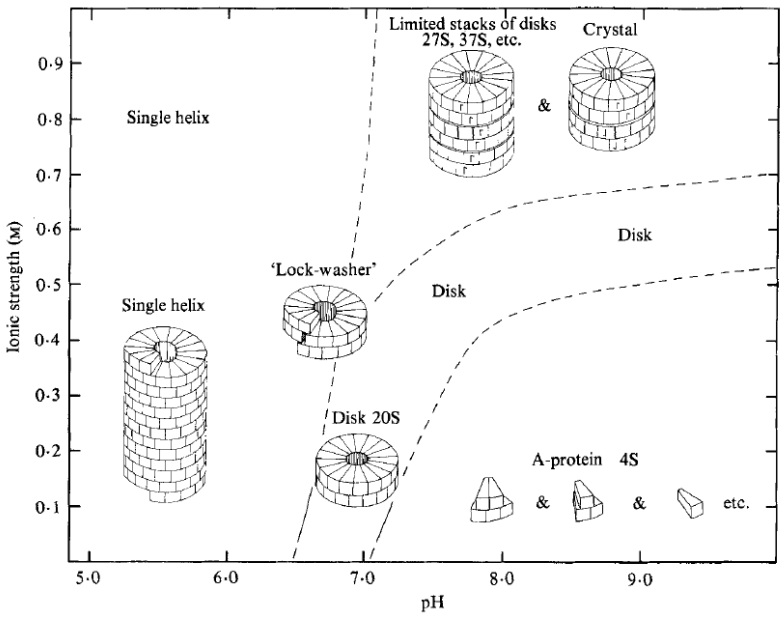


Figure 1.1, Adapted from Durham et al.[5], Phase diagram indicating the dominant form of TMV-CP at various pH and ionic strength. The rod is helical while the disk is planar. In the absence of the RNA at pH and ionic strengths where the rod form is not dominant only planar configurations of TMV-CP are possible, even at high ionic strength where planar stacked disks can look similar to rods. This behavior points to the existence of a H^{+} dependent switch centered between pH 6.6 and 7.2[5,12]. Figure 1.1 reprinted by permission from Macmillan Publishers Ltd: Durham, A. C. H.; Finch, J. T.; Klug, A. States of Aggregation of Tobacco Mosaic Virus Protein. *Nature. New Biol. 229*, 37–42. (1971)http://www.nature.com/nature-newbio/journal/currentdecade.html?year=1971&decade=1970

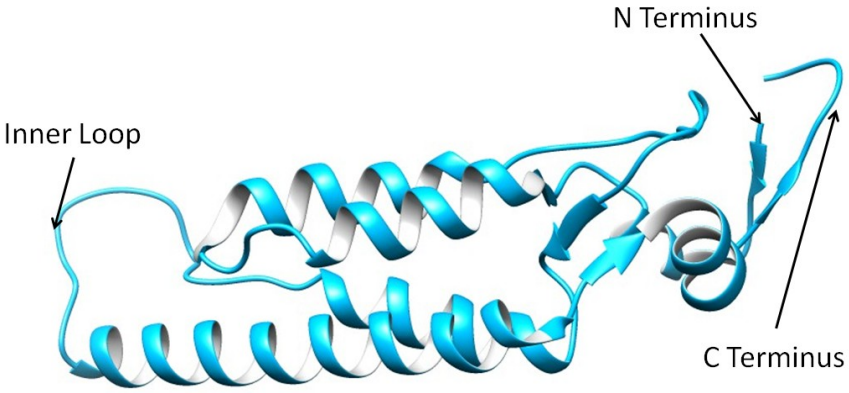
Overall, as can be seen in Figure 1.1, TMV-CP self assembly depends most strongly on pH as regardless of the other conditions, helical assemblies are only possible at lower pH or in the presence of RNA. All other extended forms of TMV at higher pH are non helical, thus there is a pH driven switch between helical and planar which varies between pH 6.6 and 7.2 depending on solution conditions. All the other factors which determine TMV self assembly simply adjust or limit the conditions for assembly. This switch is caused by an interaction known as Caspar pair formation. Caspar pair formation involves carboxylic acid groups on 2 adjacent

subunits which don't interact when TMV is in a planar configuration, but which do interact in a helical configuration. When both these groups are negatively charged they repel each other and destabilize the helical configuration. Hence the absence of helical forms of TMV at basic pH. However, when one of them becomes protonated they are able to hydrogen bond with each other, and therefore form a favorable interaction known as a Caspar pair[9]. This interaction stabilizes the helical forms of TMV and promotes TMV rod formation at lower pH in the absence of RNA.

1.2 Structure of the TMV-CP Subunit

The main structural components of the CP monomer are 4 alpha helices which run parallel to each other and end up perpendicular to the central hole in higher order TMV structures. Both the C and N termini are located on the same end, the end which ends up at the outer surface of the intact TMV rod. The hydrophobic residues are mainly located towards the end containing the C and N termini, and when the TMV-CP self assemble to form the disk structure it is these hydrophobic residues which are the driving force of the assembly and form what is known as the hydrophobic ring[9,14,34]. The other area of note is the RNA binding loop or inner loop. This loop is located on the opposite end of the where the C and N termini are located, and thus it ends up being pointed towards the inner channel upon self assembly of the A-protein into the disk state. It is this loop which is responsible for binding the RNA. The inner loop is somewhat unique among protein structures found in nature in that it is completely disordered in the disk state, presumably so that it can interact with the RNA, but highly structured in the intact virus which is needed for viral stability[35]. It is also this loop which is responsible for the specificity of the CP to react almost exclusively with the origin of assembly sequence (OAS). It accomplishes this by having quite a large amount of charge in this area, as the main driving force for RNA binding is charge cancellation.

A) Single TMV-CP



B) TMV Disk

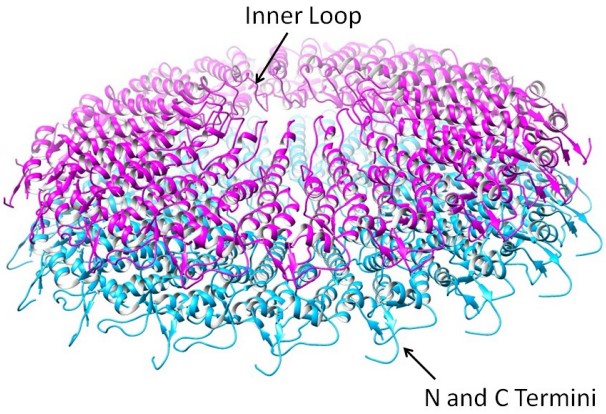


Figure 1.2, A) representation of the TMV-CP monomer, the N and C termini are located on the same face of the protein. The 4 alpha helices are the main structural components of the CP and provide the rigidity of the final TMV assembly. 34 individual CP come together to form B) the two layer TMV disk with the Termini all pointed towards the exterior surface of the structure, while the inner loop of each CP is pointed towards the inner channel. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco

1.3 Mechanism of RNA Assisted TMV Rod Assembly

Although TMV was first purified in 1935, the size and complexity of TMV made further analysis difficult. As a result, not much was known about TMV at that point. Techniques and tools used to analyse and understand the virus had to be created, and TMV was used to pioneer structural biology and virus assembly. As a consequence, it was not until 1955 that it was

discovered that TMV could be reassembled in solution from its constituent RNA and CP[36]. Even then it was not well understood why, as not much was known about the structure of TMV other than that it was helical in nature[37]. Many further attempts were later made to reconstitute the intact virus from its constituent CP and RNA, and although the assembly took place, the reaction often needed long reconstitution times. Considering its native biological environment inside plant cells, it was illogical for the reconstitution to proceed so slowly in nature, where slow encapsulation would put the viral genomic RNA at risk of enzymatic degradation. In 1971 Butler et. al reconciled the apparent incompatibility in re-assembly times that had been previously measured by showing that the speed at which the TMV RNA was protected by its viral capsid depended on solution conditions. Specifically, disk containing solutions formed TMV rods very quickly, while solutions containing smaller aggregates of TMV-CP formed rods much more slowly[4]. He therefore proposed that the disk is critical to fast reconstitution of TMV rods formed from its constituent RNA and CP. This idea eventually led to the development of the currently accepted mechanism for TMV rod formation.

In order to form TMV rods the CP must be converted from a planar configuration to a helical configuration. As mentioned previously this can be accomplished in the absence of RNA by lowering the pH of the solution. However, only RNA mediated TMV rod formation can take place at higher pH, usually between pH 6.5 and 7.3, as the RNA itself is able to nucleate the change in helicity through binding with the CP. In RNA mediated TMV rod formation there are 2 distinct assembly phases, nucleation and elongation. Nucleation is highly specific to the incoming RNA, with the CP able to recognize its own genome similarly to how the human immune system rejects and attacks any non-native objects which don't contain its own genome. Elongation is not specific and proceeds regardless of sequence to fully encapsulate the RNA[17].

1.3.1 Importance of the Origin of Assembly Sequence

TMV rod nucleation is a highly specific and delicate interaction, and only specific sequences of RNA are able to nucleate TMV rods. The region of the TMV RNA sequence responsible for TMV rod nucleation is known as the "origin of assembly" (OAS), since RNA

which does not contain this sequence will not be able to form TMV-like rods[17,22,38–41]. This region is located about 1000 bases from the 3' end of the genome[3,38]. The origin of assembly is about 100 bases long, and is comprised of a stem loop structure which exposes the critical sequence AAG AAG UCG to the incoming disk (Figure 1.3). This sequence is the first segment to bind to the TMV disk[17,38,42]. The appearance of G in every third base in the critical sequence within the OAS has been shown to be essential for TMV rod nucleation, while C bases generally don't interact favorably with TMV-CP. The exchange of A for C in the critical sequence within the OAS also was highly unfavorable for the nucleation of TMV rods[42]. Any modification of the OAS sequence which doesn't leave G in the 3rd position, such as changing the critical sequence or inserting an extra base into the OAS, is generally highly unfavorable, and results in significantly reduced or non-existent rates of nucleation of TMV rods[42]. The rest of the identity of the OAS sequence does not influence the rate of nucleation as long as the structural properties of the stem loop and spatial orientation of the critical sequence is maintained. An insertion into the OAS of an additional 3 bases significantly decreases the rate of nucleation, as the spatial orientation of the critical sequence is altered, but not to the extent of a single base insertion, as that ruins the phase of the critical sequence and essentially again destroys the needed property of having G in position 3[42].

Interestingly, swapping out the critical sequence for the sequence AAA AAA AAA resulted in over 1000 times less efficiency in nucleating TMV rods compared to the critical sequence,[42] which at face value would lead one to conclude that poly A is a poor nucleator of TMV-like rods. However, it is essential to realize that these experiments only test the ability of 9 A repeat bases to nucleate TMV rods within the constraints of a structured stem loop nucleation mechanism, while poly A seems to use a totally different mechanism. Although not much is known the mechanism of interaction between poly A and TMV CP, it may be somewhat incompatible with the structure of the stem loop, as poly A has been shown to assemble TMV-like rods on a similar time scale to native TMV RNA[27]. It is therefore likely that it is this incompatibility with the stem loop structure which leads to the poor nucleation by this critical sequence, and not an inherent problem with poly A in nucleating TMV rods. It will be

shown later that indeed poly A is quite an excellent nucleator of TMV-like rods and is even able to outperform the native genome under certain solution conditions.

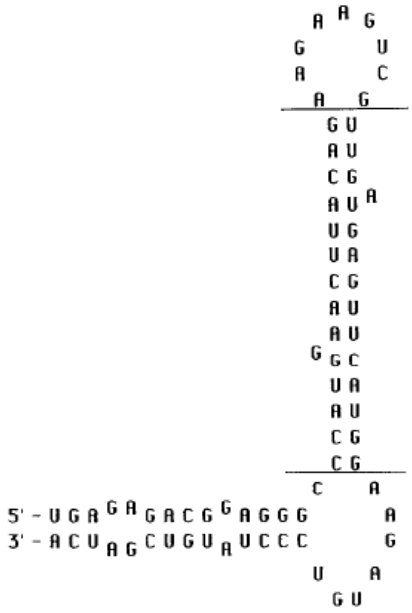


Figure 1.3, Secondary structure of the critical Stem loop structure responsible for TMV rod nucleation, The size and shape of the Stem loop structure as well as the identity of the 9 bases exposed at the top of the stem form the basis for the selectivity of the CP for its genome and are therefore critical to TMV rod nucleation. Adapted from Zimmern[38], Reprinted from *Cell*, *11*, Zimmern, D., The Nucleotide Sequence at the Origin for Assembly on Tobacco Mosaic Virus RNA, pgs 463–482, 1977 used with permission from Elsevier

1.3.2 TMV Rod Formation and Elongation

It would be expected that if physically possible, the OAS would be somewhere in the middle of the TMV RNA sequence as this would allow bi-directional TMV rod growth and thus faster encapsulation and better protection of the viral RNA in nature. Indeed this is the case, however the OAS is significantly closer to the 3' end of the RNA. As a result, the mechanism for rod elongation is different in each direction, with the growth towards the 5' end being significantly faster[6,17,43]. The 3D structure of the OAS, while critical to rod nucleation, also has the interesting effect of placing both ends of the TMV RNA in the same direction (Figure 1.4)[39,40,44]. During nucleation, the OAS binds to the first two layered disk, which alters the helicity of that disk from planar to helical. Thus, once the TMV rod has been nucleated and continues to grow, both ends of the RNA remain pointed in the same direction, with the 5' end pointed down the inner channel and joining the 3' end of the RNA[3,6].

While the structure, identity and location of the TMV OAS is not a point of contention, there has been significant disagreement over the precise mechanism of TMV rod assembly over the years. It is generally accepted that there is a need for a nucleating species which is capable of binding the entire OAS simultaneously, since smaller aggregates would be kinetically prevented from binding the OAS in a reasonable time frame[6]. However, there has been some disagreement not only over the identity of the nucleating species, but also over whether the disk or A-protein was the major species involved in rod elongation. In the mechanism proposed by Butler et al., which is based on turbidity measurements, only the planar disk is able to nucleate TMV rods[1,4,6,26]. However Namba disagreed with this mechanism, and suggested that instead nucleation was accomplished using a small helical aggregate such as the lock washer species[7,18–21]. Butler et al. also proposed that the disks are the major species used for elongation while others disagree and instead suggest that 4S protein is the main species used for rod elongation.

In support for TMV disks as the species involved in rod elongation, Butler et al. demonstrated through the use of high resolution TEM that the disk and not the proto-helix is the preferred species in the pH range where TMV rods are quickly nucleated[45]. Spatially, the RNA would struggle to easily gain access to the RNA binding region in the lock-washer species, as protein reconfiguration in helical assemblies significantly constricts the accessibility of that region relative to the disk, where the RNA binding region is completely accessible to solution in an "open jaws" configuration[14,46]. Thus, although Butler's data does not exclude the possibility that a proto helix could nucleate TMV, it does show that the disk at least nucleates rod formation much more efficiently under solution conditions normally encountered by TMV. Therefore, it is now generally accepted that the planar disk species and not the lock washer is required for nucleation of TMV and TMV-like rods[1,8,24,47].

The question of whether the disk is required for elongation is rather more complex and difficult to answer, because the various species of TMV-CP are in equilibrium with each other and are thus always present in solution and available for addition to a growing rod. There are conflicting results in the literature with Butler et al., Turner et al. and Schon et al. finding that

the disk was the major species added to growing rods using a combination of RNA banding experiments and turbidity measurements[4,17,26,26,27]. The banding experiments showed that the RNA was protected in batches of either 50 or 100 bases, corresponding to 1 or 2 layers of disk being bound to the RNA simultaneously[6,48], which points towards the disk being the major species contributing to TMV rod elongation. Turbidity experiments also clearly showed that elongation was significantly faster in solutions where the disk was the dominant species compared to solutions where A-protein was dominant[4,26]. Ohno et al. and Okada carried out very similar experiments, but found that the 4S or A protein was the major species contributing to rod elongation[24,25]. Although the question of which species is the major contributor to rod elongation is to this day still open to debate,[49] it has been found that a rather significant portion of the TMV RNA genome is only able to add the disk species[47]. Like all viruses, TMV is highly adaptable, and thus both 4S and disk protein are able to elongate TMV rods with comparable rates which depend on solution conditions[49]. The apparent contradictory findings in the literature is most likely due to the fact that the studies with differing results not only used different solution conditions but also used different strains of TMV-CP. RNA mediated TMV rod formation is driven by charge interactions. Therefore the rate of formation is quite sensitive to solution conditions, while even small changes in the amino acid sequence of the CP has been shown to have rather substantial effects on its self assembly behavior and stability[2,27,47,50–53]. Both of these effects are likely to affect the rate at which 4S protein is added to growing TMV rod differently to that of the rate of addition of disk.

Overall, these discrepancies point to the fact that TMV is highly adaptable and is likely able to eventually elongate with a variety of species and thus protect its RNA regardless of conditions. The fastest reconstitution times are achieved using the WT-TMV and the low ionic strength conditions used by Butler, in which the disk is the dominant form of TMV. Therefore under these conditions while some of the rod may elongation may be attributed to the 4S protein, the majority of the growth will be due to the addition of disk phase TMV-CP.

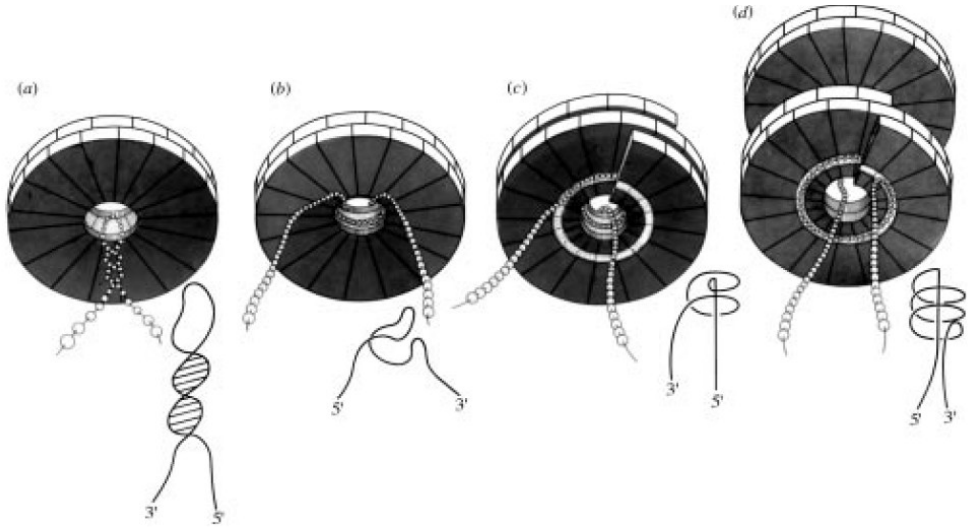


Figure 1.4, Illustration of Mechanism of Nucleation of TMV Rods, where the Stem loop structure of the OAS binds to the first 2 layer disk structure (a - b), thus changing the helicity of the disk from planar to helical (c). This binding results in both ends of the RNA being pointed in the same direction, with the 5 prime end pointed down the inner channel of the growing rod, rod growth in the 5 prime direction being accomplished through addition of further disks via the traveling loop mechanism (d), adapted from Butler[6], Reprinted from the *Journal of General Virology*, *65* (2), 253–279, Butler, P. J. G., The Current Picture of the Structure and Assembly of Tobacco Mosaic Virus, 1984, used with permission from The Society of General Microbiology

The mechanism put forward by Butler et al. therefore has rod elongation In the 5' direction accomplished via addition of disks, with each addition of disk dragging the RNA up through the inner channel to be exposed at the top of the growing rod, ready for the next disk to come in and bind to it (Figure 1.4). Thus, the mechanism for rod elongation in the 5' direction is known as the travelling loop mechanism[3]. As would be expected from the location of the OAS, TMV rod elongation towards the 3' end is substantially slower than in the 5' direction, and has been shown to proceed fastest when exposed to high concentrations of smaller aggregates of TMV-CP[54]. Also, banding experiments of RNA in the 3' direction did not show a large Fourier peak at around 100 bases, unlike RNA in the 5' direction which contained the 100 base Fourier peak[17]. Initially, the presence of the 5' RNA, which is looped back down the inner channel, could block addition of disk in the 3' direction without interfering with addition of 4S[6]. However, even when the 5' RNA is not present, the rate of 4S addition in the 3' direction is at least significantly faster than the rate of addition of disk[54]. Importantly, none of the available data proves that addition of disks in the 3' direction is impossible. It does mean that if disks can be added, the rate of addition of disk is substantially slower than in the 5' direction. The TMV

disk does have a vertical polarization to it[7,11,14–16], meaning that it is quite feasible that TMV disk would either only be able to add in one direction, or would add substantially more effectively in the 5' direction compared to the 3'.

In both directions, elongation of the TMV rod is sequence independent,[17] and aside from the OAS which is required for nucleation, any RNA sequence can be substituted for the rest of the TMV genome and TMV-like rods will still be formed[17,40,55]. However, there have not been any studies on whether the sequence of the RNA has an effect on the rate of rod elongation. Given that different trinucleoside diphosphates have varying affinities for the TMV-CP, with the AAG sequence having the highest affinity and a clear preference for a G base every 3rd base in the OAS[38,42,46,47,56], it is not an unreasonable proposition that the rate of elongation might be somewhat dependent on the RNA sequence. In addition to this, the 5' end of the TMV genome has a lower affinity for the CP compared to the rest of the genome, and as such, can be selectively disassembled in alkaline pH without removing the CP from the rest of the genome[57]. It therefore is certainly a possibility that there would be at least some dependence of the rate of TMV rod elongation on the RNA sequence, but further study would be needed to make any conclusive statements.

1.3.3 Relevance of calcium in TMV in nature and proposed calcium switches

TMV is a virus which has evolved not only to self assemble in vivo but also to infect plant cells and reproduce itself. Thus, while stability is very important for the RNA to be fully protected until it enters the cell, the RNA also has to be detached from the CP once at its target location or else no replication would occur. Therefore attachment of CP to RNA is designed to be a reversible process. In an ideal world, in order to maximize effectiveness of the virus it would be beneficial if it had 2 states: One where the RNA was strongly bound to the CP and therefore highly protected from various forms of damage; And a second where the CP was weakly bound to the TMV RNA which would allow for RNA transcription and therefore further replication of the virus. Indeed these states do exist and are accessed by binding or unbinding events which depend on solution conditions.

The binding of calcium has been shown to greatly increase the stability of TMV[28,58,59]. While the loss of calcium is one of the conditions necessary for the disassembly and replication of the virus to take place[7,8,50,60,61]. In plant cells, the inside of the cell has a higher pH and a very low calcium concentration of around 10^{-6} M, while outside the cells the calcium concentration is considerably higher at around 10^{-3} M, as is the concentration of H^{+} [7,8,50,61,62]. Therefore, calcium plays a critical role in the assembly and disassembly of TMV in nature, with calcium binding and neutral pH stabilizing the RNA - CP bond, while low calcium concentrations and high pH weaken the forces keeping the RNA and CP strongly associated to each other. Thus calcium and H+ concentration together act as a switch between the stable state and the weakly bound TMV replication state.

Under physiologically relevant conditions, between pH 6 and 9, TMV has two calcium binding sites which were first discovered through titration experiments[58,62,63]. Although these sites are also able to bind weakly to many cations, they bind divalent cations most strongly with the greatest affinity for calcium[58,59,62]. Out of the divalent cations tested, calcium also shows the steepest TMV binding response with only 1 mM Ca^{+2} sufficient to maintain around 85% of the TMV rods at pH 9 and 2^{o}C but all the TMV rods being completely disassembled under the same conditions but with the calcium concentration at 10 µM. In comparison, Mg^{+2} managed to stabilize about 50 - 60% under the same conditions at a concentration of 1mM but still retained ~20% of the original rods when the Mg^{+2} concentration was lowered to 10 µM. All monovalent cations tested resulted in the TMV rods being completely disassembled even at a concentration of 1 mM[59]. Thus calcium binding exhibits the sharp response curve required for switching quickly and effectively between the stable form and the weakly bound form of RNA. This calcium response is not only useful for TMV replication, but also has interesting consequences for TMV based nanotechnology.

Interestingly TMV-CP on its own is not capable of binding calcium, except at pH below 4.6[63,64], possibly because the RNA binding loop remains partially disordered unless it is bound to RNA, even in certain regions of the phase map where TMV-CP forms rods[35,64]. This idea is reinforced by the experiments of McMichael et al. where TMV-CP at around pH 5 was raised

very quickly to pH 6.5 and then exposed to a high (10 mM) concentration of calcium ion which resulted in RNA free TMV rods which were much more stable to disintegration when placed at $4^{o}C$ compared to magnesium, barium or potassium[28].

Not only does TMV bind calcium, but calcium is also able to displace H^{+} from the virus at physiological pH[62]. In addition to calcium ion, La^{+3}, Ba^{+2}, Mn^{+2}, Sr^{+2}, $Pb^{+2}$and Mg^{+2} are also able to displace protons with various success. In order for a single binding site to be able to bind protons, potassium, magnesium, calcium, cesium and lead, all of which have very different sizes and charge densities, the binding site must be flexible and adaptable to various binding partners. This simply would not be possible unless the binding site had multiple components, at least some of which were flexible and would therefore allow at least some form of binding to most cations, while providing much stronger binding to cations closer to the ideal size and charge density. Therefore the binding site would have the requirement of being negatively charged at physiological pH and also be somewhat flexible to be able to adapt itself to a variety of atomic sizes suggesting that at least some of the carboxalic acid residues are involved in cation binding events[62,65,66]. This was demonstrated to be the case when it became possible to get high resolution electron density maps from TMV fiber diffraction data with the presence of two calcium binding sites confirmed using this method[2,18,67]. As expected, these binding sites were not found in high resolution electron density maps created from the disk species[14,68].

1.3.4 Structure Details of TMV and RNA-Protein Interactions

Reconstructing the virus from x-ray techniques has proven difficult. Not only is the virus large and complex, but also the TMV rod cannot be crystallized, although it is possible to obtain crystals of the TMV disk, so fiber diffraction of well ordered TMV gels was used instead[1,14,61]. Additionally, it took several years for computing power to catch up to x-ray techniques[18,61]. The net result is that molecular details of TMV were only available many years after the self assembly properties were discovered.

As mentioned previously, the main structural components of the TMV-CP are 4 alpha helices which in the assembled virus run between the outer surface of the virus and the surface of the inner channel, perpendicular to both of those surfaces. These 4 helices give the TMV-CP its overall structure, and they are also responsible for its rigidity since those alpha helices do not go through changes in orientation during the conversion of disk to helix[13,14,61]. Although there are differences in side chain packing between the disk and rod state, the main difference between disk and rod is a change in the orientation of the entire subunit, with the CP being tilted relative to the central axis in the rod but not the disk resulting in different contact points between subunits, especially in the axial direction[2,7,18]. The helices are connected by a beta sheet which lies towards the outer edge of the assembled structure. Since there is no up or down in a cylindrical structure, an alternate naming scheme has been developed. The main helices, which are longer and run perfectly perpendicular to the inner and outer surfaces of the assembled virus, are known as the left and right radial helices (LR and RR respectively). The shorter helices, which can either be interpreted as sitting on top of or below of the radial helices, depending on the orientation of the viewer, are known as the left and right slewed helices (LS and RS respectively) as they are at a slight angle with respect to the radial helices and are also shorter in length[2]. By convention, protein residues are listed from N terminus to C terminus. In TMV this makes the N terminus residue 1 and the C residue 158. Thus in the disk state residues 20 - 32 lie in the LS helix, 38 - 48 make up the RS, 74 - 88 the RR and 114 - 134 the LR[14]. Although the orientation of the helices remains the same in the TMV rod, two of the helices are elongated in the rod compared to the disk. The LS helix is elongated by 1 residue and the LR is elongated by 4 residues[18].

The most critical residues in terms of TMV assembly are the hydrophobic residues and the charged residues. The hydrophobic residues are mainly located in the outer radius of the assembled TMV structures, and form a continuous ring of residues which interact via hydrophobic interactions, including at the CP interface, at a radius of 7.5 nm in assembled TMV structures. This ring is known as the "hydrophobic girdle" and is thought to be the driving force for TMV self assembly[9,14,34]. The charged residues are exclusively responsible for giving TMV self assembly its dependence on pH and ionic strength. They are mainly located in the inner region

of the TMV at low radius, but some are also located in patches throughout the TMV structure, usually at regions of the TMV-CP which come into contact with adjacent monomers[14,18].

Carboxylic acid groups were thought to play a critical role in mediating the assembly of the various states of TMV. Disks are thought to be prevented from forming at higher pH due to the repulsion between the negatively charged carboxyl groups[9]. The fact that disks do form at higher pH if the ionic strength is high enough suggests that this mechanism is correct, as in higher ionic strength solutions, the screening between the negative charges would more effective, which would then allow subunits to approach each other[5]. Carboxylic acids are also thought to be involved in the transition from disk to helix but in a more direct way. Due to changes in axial packing, the contact area between adjacent subunits in the helical rod configuration is slightly less than in the planar disk configuration[2,7,9,14]. Therefore, there must be another factor which favors the rod phase over the disk, other than hydrophobic interactions which are present in both cases. This interaction must also drive helical rod formation only at lower pH and not at higher pH. The interaction which accomplishes this was theorized as Caspar pair formation, which involves the formation of a hydrogen bond between two carboxylic acid residues located on neighboring subunits[9]. Indeed, models based on this principle can model the assembly behavior of TMV-CP well. Structurally, there seems to be at least one Caspar pair which is responsible for mediating the stability of the rod state. This has been identified as an interaction between Glu 50, located in the LS helix and Asp 77 from the RR helix of the axially adjacent subunit[2,50,61]. In addition to stabilizing the rod phase at low pH, these Caspar pair also de-stabilize the rod phase at high pH, since at high pH TMV in rod phase would place these two now negatively charged residues in close proximity to each other. This proximity causes substantial repulsion which has been shown, through mutational studies of Glu 50 and Asp 77 which remove the charges, to drive rod disassembly and TMV replication at higher pH[50,51,61].

Although Caspar pair formation seems to explain the rod to disk transition well, and in the case of the interaction between Glu 50 and Asp 77, Caspar pair formation seems to be the driving force, there are also secondary interactions between adjacent charged residues. The

formation of Caspar pairs is usually followed by secondary salt bridges to help stabilize the interaction, which is not unexpected, as in TMV most of the charged residues reside in a fairly small portion of the three dimensional structure[2]. In some cases, it is not simple Caspar pairs which interact, but the interaction of three or more residues containing carboxylic acid groups which are stabilized by a complex network of salt bridges and hydrogen bonding[2,51]. This type of network may explain the stabilization of a high concentration of negative charge in the "carboxyl cage" which include Glu 106, Glu 95, Glu 97 and Asp 109 which work together to provide a driving force for rod disassembly at high pH[2,51,61].

These types of networks, which usually involve positively charged Lys and Arg residues, can accommodate a larger a variety of situations because they are more complex and also more flexible due to side chain rearrangements, with the positively charged groups being particularly flexible in their conformations. This allows TMV-CP not only to bind RNA, but also to be bound to RNA and still be in the rod phase in different conformations depending on the environment. For example, Ca^{+2} has been shown to alter the interactions between RNA and TMV-CP[2,7,18,60]. In the absence of RNA, at low pH where the TMV-CP forms rods on its own, the negative charges on the phosphate backbone are replaced by 1 or more anions. In addition, there is likely some rearrangement of the charged groups in the RNA binding region to accommodate the anion(s)[13]. Thus, this network of charged residues is able to adapt to many slightly different conformations, partly due to the flexibility in the side chains of the positively charged residues, all of which exist in the rod state and therefore stabilize the rod state in a large variety of conditions.

Many of the charged residues at low radius are involved in RNA binding[2,18,60] while these charged residues are mostly involved with RNA binding and change conformation depending on environment. There are also some critical charged residues at the interfaces of the CP monomers which are critical for mediated TMV assembly[2,69,70]. Although the assembly process is driven by hydrophobic interactions, the process is mediated by electrostatic interactions[9,10,34,50,51]. If the TMV assembly process is thought of as the movement of a train, then the hydrophobic interactions can be thought of as the locomotive since they drive the

process, but then in that case, the electrostatic interactions would be the engineer. And just like an engineer cannot move a train on his own, he is still the one who controls the locomotive which does drive the train. Thus it is the electrostatic interactions which determine when the hydrophobic interactions are allowed to drive the assembly process. The electrostatic interactions do however drive disassembly as well as binding of the CP to RNA[2,51].

In order to understand details of how the RNA binds to the TMV-CP, it was first necessary to develop new techniques for generating and analyzing high resolution electron density maps, from which the precise protein conformations can be modeled accurately. Although it was possible to collect high resolution fiber diffraction data of TMV gels in the late 60's, due to problems with the computing power and the computer algorithms required to analyze such complex electron density maps, accurate details of TMV disks were not available until the late 70's[2,14,71,72]. The intact virus took longer, as not only is it larger, but unlike TMV disks, TMV rods including the RNA cannot be crystallized, and instead X-ray fiber diffraction was used to generate the electron density map[2,71,72]. Using this method, it took until the mid 80's to surpass the 4 Å barrier, and it took until 1989 to obtain a 3 Å resolution electron density map. Before the 4 Å barrier was reached, many different structures for the TMV-CP were proposed, but although they correctly identified the approximate binding position of the TMV RNA, the details of the interaction were shown later to be inaccurate, as was much of the other proposed structural detail[2,18,69,71,72].

Resolution is a critical part of getting accurate structures of complex bio-molecules, and a resolution of at least 4 Å in the electron density map is now considered necessary in order to get reasonable structures. Above this resolution, the determination of the structure depends strongly on the model used, and thus it is difficult to be confident on whether the determined features are actually representative of the virus structure, or if the features are simply artifacts created by the modeling[2,18,60,71,72,72]. Ideally, a resolution approaching 3 Å is best, as at this resolution the exact path of the backbone can be traced, and thus only the configuration of the side chains must be modelled, which leads to significantly fewer parameters in the modelling and thus much more accuracy[2,60]. Studies near 3 Å resolution are therefore the most reliable, as

they don't depend much on the type of modelling used to interpret the electron density map. As such, the highest resolution study of intact TMV is the 2.9 Å fiber diffraction based 1989 study of Namba et al.[2], which has been considered to be the structure of TMV under all conditions since publication[8]. Recently, a similar resolution electron density map for TMV which approaches 3 Å has been achieved using CryoTEM. This experiment found that there are actually small but significant differences in both the RNA and TMV-CP configuration depending on how much calcium is bound to the TMV, which would suggest that the exact structure of TMV is not static but instead depends on solution conditions[60].

The RNA binds to a specific region located at a radius of 4 nm from the central axis which contains several Arginine residues[2,69,70]. It was highly suspected well before a detailed structure of TMV was available that it was the Arginine residues which were binding the RNA, since out of the 158 amino acids found in TMV CP, only 25 of them are conserved among the various strains. These 25 residues include Arg 41, 90, 92 and 113, which are all located in the RNA binding region[2,69,70]. Since the inner channel surface is located at a radius of 2 nm, in the intact virus, the RNA is bound inside of the protein structure and is thus solution inaccessible[2,18]. The most critical interaction is a charge interaction between the negative phosphates on the RNA backbone and the positive Arginine residues. There are also secondary interactions which depend on solution conditions and are described fully below. These include mainly h-bonding and hydrophobic interactions, however, because the electron density signal generated by the RNA from high resolution structural studies is an average of the entire TMV genome, there is more uncertainty in exactly how each type of base interacts with the CP in each of the 3 possible binding positions compared with the smaller uncertainty on the protein conformation[2,18]. On average, each CP binds to 3 RNA bases, however each nucleotide interacts with an upper and a lower CP, and each individual CP will interact with a nucleotide below and above it, with the upper interactions being different from the interactions with the nucleotide below, and again dependent on solution conditions. This strengthens the axial bonds, and it is actually through reorganization of these bonds that the axial bonds are weakened when it is time for virus disassembly[2,60]. Since in the assembled rod the RNA is bound to two axially adjacent TMV subunits, the two layered disk is ideal for TMV rod nucleation, since it allows for

the initial RNA binding loop to immediately be fully bound through interaction with each of the two layers of disk.

The details of how the RNA binds to the TMV-CP depend on solution conditions. In specific, they depend on the concentration of calcium present in solution[2,60]. In general, the viral RNA is rather contorted and seems to adopt a conformation not usually seen in RNA or DNA[2]. This may explain why no sequence of DNA was able to nucleate TMV rod formation, as DNA is less flexible than RNA and thus forms less secondary structure[44]. It is not clear whether it is the additional hydrogen bonding and base stacking interactions which would only be present in RNA since they involve the 2' hydroxy group on the ribose sugar, which make RNA able to nucleate rods but not DNA, or whether it is merely that the energy cost to place DNA in the required conformation is simply too high compared with that for RNA[2]. The bases themselves interact with the LR helix while the phosphates interact electrostatically with Arginine residues from the inner loop of the subunit directly axially adjacent (above or below depending on view). Therefore the RNA is actually bound to 2 axially adjacent subunits which helps to increase stability in the axial direction. The RNA is more strongly bound to the TMV CP under conditions of high calcium concentration (10^{-3} M), since axial bonds between adjacent subunits are then also strengthened through salt bridge formation which stabilizes the TMV rod in the presence of calcium.

Due to RNA signal averaging, there is some uncertainty in the precise dependence of sequence on resulting RNA conformation. However much of the RNA-CP interaction is determined from the shape of the CP which results in a fairly narrow range of electron density due to the RNA. Therefore it is possible for many of the RNA-CP interactions to be elucidated. There are no Aromatic residues in the direct vicinity of the TMV-RNA therefore no RNA - CP base stacking interactions are present[2,18]. The most critical interaction is electrostatic between the positively charged Arginine residues of the CP and the negatively charged phosphate groups of the RNA. The phosphate groups from the RNA bind to the inner loop of the axially adjacent subunit. At high calcium concentrations the primary electrostatic interactions involve phosphates 1 and 2, phosphate 1 binds strongly to Arginine 90, while phosphate 2 binds

strongly to Arginine 92. Phosphate 2 is also bonded to a proposed calcium ion. This calcium ion is also bonded to Asp 116 and the ribose hydroxyl group from nucleotide 1. Thus this calcium ion effectively shields phosphate 2 from Asp 116 which would otherwise form a considerable repulsive interaction with Asp 116 due to their close proximity. Asp 115 and 116 are also conserved throughout all strains of TMV and have been shown to be involved in virus disassembly[8,50,51]. There are also weaker, secondary electrostatic interactions which stabilize the RNA - CP bond under high calcium conditions including bonds between phosphate 1 and Arg 92, phosphate 1 and Arg 41 and phosphate 3 and Arg 92[2].

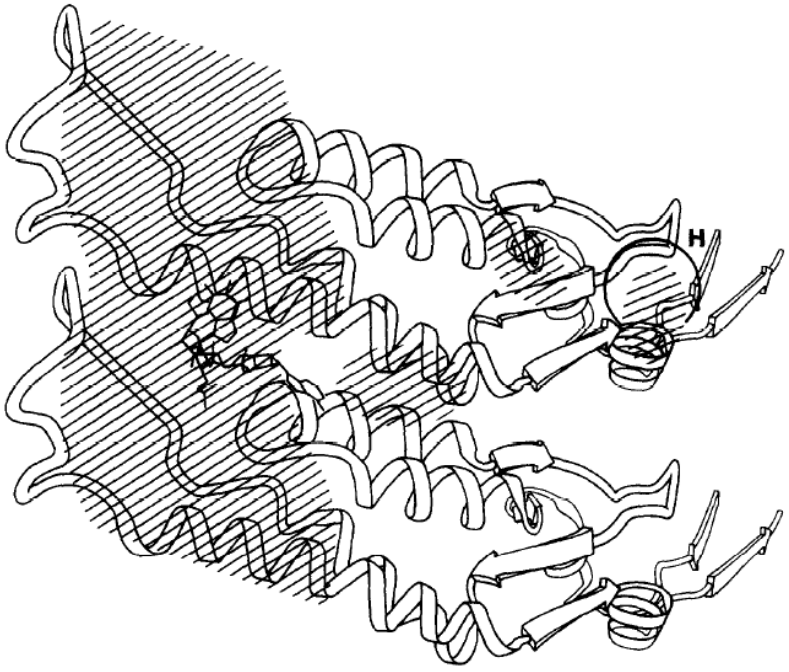


Figure 1.5, Adapted from Namba et al.[2] Diagram of TMV subunits from a TMV rod. The charged regions of the TMV are shaded with the RNA conformation shown binding 2 axially adjacent subunits. Reprinted from the *Journal of Molecular Biology*, 208 (2), 307–325, Namba, K., Pattanayek, R., Stubbs, G., Visualization of Protein-Nucleic Acid Interactions in a Virus: Refined Structure of Intact Tobacco Mosaic Virus at 2.9 Å Resolution by X-Ray Fiber Diffraction, 1989, used with permission from The Academic Press

Hydrogen bonding also plays a critical role in stabilizing the RNA - CP interaction and is especially important in giving TMV-CP its specificity for the OAS with several hydrogen bonds only able to form if a specific base is present. Under high calcium conditions if base 1 is a guanine it is able to form additional hydrogen bonds with Arg 122 and Asp 115. Although Asp 115 is conserved in all strains of TMV Arg 122 is not, however the positive charge on residue

122 is conserved in all strains of TMV[70]. Base 2 doesn't form any base specific hydrogen bonds but Base 3 is able to form a hydrogen bond with the main chain carbonyl group of Thr 89 if it is adenine[2]. Thr 89 is also one of the 25 residues which is conserved through all strains of TMV[70].

There are also less critical secondary interactions which stabilize the RNA - CP bond. As previously mentioned there are no base stacking interactions between the TMV-RNA and the TMV-CP, however base stacking between bases 1 and 3 is possible under high calcium conditions if bases 1 and 3 are purines[2]. Under these conditions, if the RNA is modeled as GAA a number of non specific interactions are able to take place due to the conformation of the RNA. The RNA bases are lying flat against the LR helix, with the hydrophobic face of base 1 interacting through van der Walls forces with a methyl group from Val 119 and base 3 interacting similarly with the main protein chain between Asp 116 and Ala 117. Base 2 is sandwiched between Ser 123 and Asn 127 from the LR helix and Asn33, Gln 34 and Thr 37 of the RS helix of the axially adjacent subunit which allow for non-specific interactions between the polar groups of base 2 and the CP. Of all the residues involved in non-specific RNA - CP interactions only Asp 116 and Ala 117 are conserved through all strains of TMV and they are likely conserved due to their interaction with the phosphate groups rather than their role here[70]. Although in the case of the sequence GAA these interactions would contribute to virus stability, the lack of conservation of these residues indicates that the base stacking and non-specific interactions are not critical to virus replication. Furthermore it would be possible for similar non-specific interactions to occur by using other residues with only small adjustments to the RNA conformation. Therefore the RNA likely makes minor, sequence dependent adjustments to minimize unfavorable non-specific interactions which would be very difficult to detect with a base averaged electron density map.

Overall the bases form favorable interactions with the LR helix while the phosphate groups interact with the Arginine residues from the axially adjacent subunit. As such the RNA acts as a link in the axial direction and the more strongly it is bound to both subunits the stronger the axial bonds and the more stable the intact virus will be. The RNA conformation in high calcium conditions is not only bound strongly to both CP subunits through the interactions

described above but the presence of the Calcium ion intersubunit salt bridges to form both in the axial and lateral directions which also increase virus stability[2,60]. Interestingly, the structure modeled under high calcium concentrations would provide axial stability even at high pH, since the groups which would become negatively charged at high pH and therefore repel each other, leading to rod dissasembly, are balanced by the calcium ion (figure 1.6).

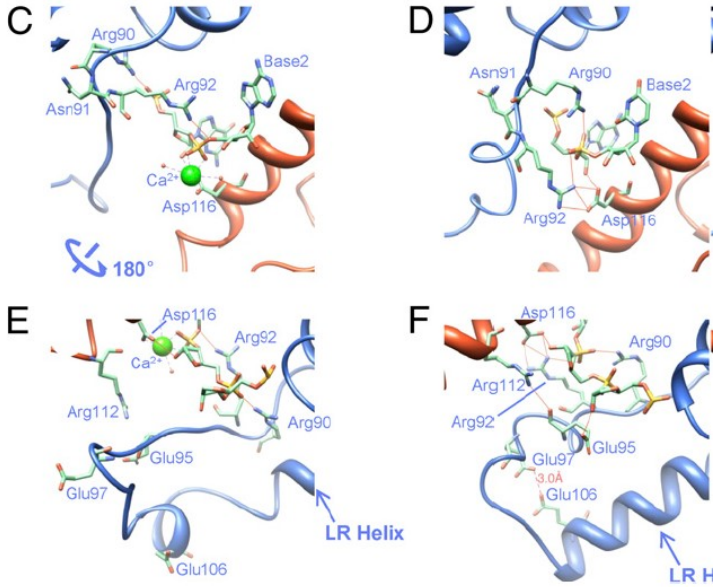


Figure 1.6, Adapted from Ge et al.[60] Comparison of RNA binding region with (C and E) and without (D and F) the presence of calcium. The calcium ion causes rearrangement of the TMV-CP residues which results in separation or shielding of residues which are negatively charged at high pH, thus leading to increased stability relative to the low calcium state. Reprinted from *PNAS*, 108 (23), 9637–9642, Ge, P.; Zhou, Z. H., Hydrogen-Bonding Networks and RNA bases Revealed by Cryo Electron Microscopy Suggest a Triggering Mechanism for Calcium Switches, 2011, used with permission from the Proceedings of the National Academy of Science of the United States of America

The second calcium binding site was not observed directly in a high resolution study, however titration experiments clearly show that more than 1 calcium is bound per CP, and that this calcium is able to displace protons, thus it must be a direct bond with the protein. Other than the binding site described above, there is only one other region which has the possibility to bind calcium, and this site was shown to bind lead quite well via modeling without any movement of the protein backbone, which is highly indicative that it could also bind calcium[67]. The site in question involves a lead ion bound to both oxygens from the carboxylic acid group in Glu 95 as well as the carbonyl oxygen from residue 102. There is also possible secondary interaction with the carbonyl oxygens of residues 101 and 104. In addition, although Glu 106

only interacts weakly with the lead ion in the model, a calcium ion would fit the binding site a bit better, so it is likely that in reality Glu 106 also interacts strongly with the calcium[67].

These two proposed calcium binding sites would only be present in the assembled virus, as one directly involves the RNA and the other requires a rather precise backbone conformation which is definitely not present in the disk and probably not present in pH triggered rod assemblies. The region in question is completely disordered in the absence of RNA, as NMR measurements show that although in the rod state that region is more ordered, the inner loop is still not as ordered as it is in the presence of RNA[35]. It would therefore be consistent with the experimental evidence that the TMV-CP on its own is not able to bind calcium at TMV assembly pH[64].

In the absence of calcium, the CP and RNA undergo rearrangement with an alternate network of interactions taking place[60]. Without the presence of calcium to balance the charge, Asp 116 and Base 2 repel each other (figure 1.6). Thus, Arg 92 instead forms a salt bridge with Asp 116, which has the effect of dragging Arg 90 into the position which Arg 92 occupies when at high calcium concentration. Arg 90 therefore interacts with Base 2 at low calcium concentration instead of Base 1. There are many secondary interactions which take place as a result of this rearrangement. For example, Arg 122 forms a salt bridge with Asp 88 instead of interacting with Base 1. Likewise, the LR helix is also extended further as a result, which causes Glu 95 to hydrogen bond both with Arg 112 and base 3, while Glu 97 and Glu 106 form a weaker hydrogen bond[60]. The net effect of this alternate low calcium charge network is to reduce strength of the RNA - CP interaction. Certain CP - CP interactions are also weakened while others are strengthened, which results in a TMV helix wherein pairs of turns are more strongly bound to themselves then to the next pair of turns. Thus the helix is partially like "beads on a string"[60]. The loss of the second calcium binding site has the effect of reducing CP - CP affinity in the lateral direction, which further weakens the strength of the TMV helix. In the alternate low calcium configuration there are several acidic residues in close proximity to each other which are only stabilized by a shared proton at lower pH. Removal of this proton would further destabilize the helical structure, which is likely why TMV is more stable at higher pH in a

high calcium environment[59] where the negative charges are still balanced by the calcium ions, than it is in a low calcium environment.

Interestingly the ideal sequence for CP binding as calculated from high resolution studies is different depending on whether or not calcium is bound. In a calcium free environment, A*G was identified as the ideal trimer[60]. This is what would be expected since the OAS has a preference for G in the 3rd position and A*G is repeated throughout the critical 9 base binding sequence with **G repeated throughout the OAS. Since TMV would be assembled in the cell, which is a low calcium environment, the OAS is the ideal nucleating sequence. However, in a high calcium environment, the ideal binding sequence becomes G*A, with having A in the 3rd position being highly desirable. Aside from the OAS, in the rest of the TMV RNA genome there is a preference to have A in the 3rd position[60]. This would make sense, as once the RNA is encapsulated and exits the cell, the assembled virus goes into a higher calcium concentration which would stabilize the RNA - CP interaction through this switching mechanism. This is therefore consistent with the known fact that calcium binding to TMV acts as a switch between the high stability state and the low stability state. In poly A although there is no G, there is still likely an energetic gain due to the adenine in the second position[47]. Importantly, there are no unfavorable interactions with A in any of the 3 positions. Therefore, in theory, poly A should get additional stability, even more than the native TMV genome, in a high calcium state, since poly A contains an A base in every 3rd position.

Due to the x-ray fiber diffraction method's reliance on heavy atoms, one of which is lead, which binds to TMV with similar affinity to calcium as it has similar size and charge density, the conditions were almost certainly in the pseudo high calcium regime, with lead likely stepping in as the preferred ion to bind to the calcium sites. Since calcium and lead ions are not quite identical, it is possible that the there would be some very small structural differences in the protein between calcium binding and lead, but they would be minor, likely with only a change in the bond distances and not a major change in conformation that could lead to different bonding partners. Given that the loss of calcium results in significant rearrangement of the CP and RNA, and that the same binding site in TMV is able to bind a variety of cations, it is

likely that smaller but noticeable rearrangement of the H-bonding and charge network occurs if other cations are bound. Therefore, the stability and therefore the precise conformation of the involved residues in TMV would be expected to respond to solution conditions much in the same way as the CP responds to aggregate to different states depending on solution conditions.

Although there have been no studies which have looked at the effect of various solution conditions on the precise conformation of the RNA binding region of TMV, the prediction that different solution conditions would lead to small differences in that region is supported by studies comparing TMV rods with and without RNA. It was found that when TMV-CP forms rods without the presence of RNA, the electron density which was provided by the RNA is replaced by anions, with only a few side chain re-arrangements leading to salt bridge formation as a consequence of the anion binding[13]. In other words overall the structure of the TMV-CP in the rod state is similar regardless of whether there is RNA present or not, with TMV in general being very adaptable in terms what kind of charge density it is able to bind due to being able to make small adjustments to the charge network located near the RNA binding region which would be required to bind these new ions.

1.3.5 Cooperative binding of TMV RNA

The ability of TMV-CP to simultaneously bind large amounts of RNA is useful not only for nucleation, where addition of smaller aggregates is thought to be kinetically unfavorable, but also for elongation and incorporation of less favorable segments of RNA. As such, there seems to be a preference for TMV RNA to be bound co-operatively by larger aggregates of TMV-CP. While three RNA bases bind per CP, there are certain base triplets which are more favorable than others[2,38,42,56]. However, TMV-CP is not able to bind short three base long RNA oligonucleotides, except near pH 5[46,56]. The reason for this is not clear, however it may be that since dislocation of a single CP from planar to helical configuration affects the interaction with all surrounding CPs, and the energy gain from binding to the RNA oligonucletotide is not sufficient to compensate for the energy lost by loss of contact with these adjacent surfaces. At lower pH, a CP can adopt a helical conformation on its own, and thus adjacent CP could adopt a

helical conformation upon binding of a three base RNA oligonucleotide. It is not possible for three base oligonucleotides to bind TMV-CP co-operatively. Thus, in order for the oligonucleotide to bind to the CP, each strand must be able to independently form a positive interaction with the CP, which is not possible at neutral pH. This further suggests that the binding of RNA to TMV-CP is a cooperative effort[46] without a strong driving force (as it needs to be a reversible process for virus replication). Thus, many bases must bind simultaneously, even with highly favorable sequences, in order for the interaction to be strong enough to engage the encapsidation reaction. It is possible that since there is a significant alteration of the CP orientation in the rod form vs. the disk, it is necessary for a significant percentage of the disk to be converted to helix before it becomes energetically favorable to form rods.

Although a three chain RNA oligomer is not able to nucleate TMV rod formation at TMV assembly pH, a six chain RNA is able to do so[46]. Thus, it seems that a six base interaction, which would allow four CP to adopt a helical configuration, is sufficient to overcome the loss of adjacent subunit surface interaction. Interestingly, when nine base oligomers were exposed to TMV, they were bound but did not attain full binding coverage, with only about 3/5 of the possible RNA binding sites occupied[46]. The reason for the nine base oligomer to only achieve 3/5 coverage while the six base oligomer achieved full binding coverage is not clear. However, it is possible that since RNA binding to the RNA binding track significantly reduces the mobility of the inner loop of the CP,[2,14,35] the larger nine base oligomers are simply too large to fully penetrate near previously bound CPs, and thus leave some of the sites unbound. In the disk, the inner loop is incredibly flexible in order to allow the RNA to pass through it and penetrate deeper into the CP structure to the RNA binding region. This is the reason why the inner loop is so unusually flexible, as the RNA must penetrate a further 2 nm after encountering the loop in order to reach the RNA binding region[2,18,35]. Once the protein is bound to the RNA the inner loop loses that flexibility and becomes highly structured[2,14,35], thus there is not much "wiggle room" for RNA to enter. TMV RNA is bound co-operatively at the same time, thus TMV never evolved to to bind RNA sequentially. As a result, it is possible that the six mer RNA is still small enough to be able to penetrate, while the nine is too large to achieve full coverage.

1.4 Potential Applications of TMV

Although the TMV disk and rod phase TMV-CP have interesting potential, in the field of nanotechnology the majority of recent research has focussed on the assembled TMV rod, usually containing the native genome. Higher order TMV structures such as the disk and rod are built using a single building block and as such have highly reproducible spacing between the CP subunits, as TMV has a highly rigid structure. Therefore TMV is often used as a template for organizing groups since the spacing between those groups is on the nm scale, and as such can be controlled to a degree which is highly challenging to achieve using alternative methods. Strict control of structure on the nanoscale often results in novel properties, which have the potential to form the basis for next generation devices. Although in some cases the TMV itself has a more direct role, the fact that it is able to play an organizational role makes it incredibly versatile going forward in terms of the number of potential applications which TMV based technology could be used for. TMV research has led to a wide variety of novel optical and electronic potential applications for TMV based devices including light harvesting[73,74], plasmonics[75,76], sensing applications[77–80], fabrication of memory storage devices[81–83], battery electrode construction[78,84,85], nucleation and growth of several types of nanoparticles[29,83,86,87], arrangement of nanoparticles[76,88–90] and growth of nanowires[82,91–95].

Although TMV was a major research topic for over 40 years, the focus of that extensive research was in elucidating fundamental TMV properties and self assembly characteristics. It wasn't until the late 1990's, when the idea of nanotechnology was being established, that it was considered that TMV might actually be a useful platform for future device construction. One of the first studies to take advantage of TMV's properties in this fashion was done by Shenton et al. who posited that since the self assembly of TMV is mediated by surface charge interaction, at least some of those solution accessible charge networks might be able to nucleate inorganic nanoparticles. As expected, TMV was indeed found to be capable of preferentially nucleating certain types of nanoparticles at its surface relative to nucleation in solution[79,86]. The surprising result from that study was that just as TMV is highly adaptable and able to assemble various forms depending on solution conditions, TMV was also found to be able to nucleate different

types of nanoparticles under different solution conditions[29,86,88]. Shenton et al. found that the main dependence on the type of nanoparticles produced was on the pH of the solution[86]. Iron oxide was nucleated at pH 9, CdS at pH 7 and PbS at pH 5. Changes in pH alter the self assembly behavior of the virus by altering the surface charge landscape of the TMV, which is likely what produced the observed pH dependence. The pH driven changes in conformation are suited to housing different precursor ions. In addition, it was found that the majority of the nanoparticles produced remained bound to the TMV afterwards, which would be expected if it was indeed the charged groups on TMV which were responsible for nucleation[86].

Dujardin et al. later confirmed that surface charge is indeed the mechanism through which TMV nucleates and binds these nanoparticles. In that study, site directed mutation was used to reduce the surface charge in the inner channel of the TMV. This work demonstrated that the WT TMV was able to nucleate silver nanoparticles in its inner channel, while the TMV mutant, with its reduced inner channel charge, was not able to nucleate silver nanoparticles under the same conditions[88]. The nucleation of of Pd nanoparticles was studied to confirm that TMV nanoparticle nucleation is done on the surface of the TMV and not in the bulk solution[79]. TMV nucleated Pd nanoparticles also showed reduced size and better monodispersity compared to nanoparticles nucleated without TMV present. Although this result cannot be absolutely generalized to the other types of nanoparticles, it is difficult to envision a mechanism other than the TMV surface charge, which would allow the nucleation to take place. As previously discussed, many of the charge interactions involve networks of charged residues which can adopt slightly different conformations under different conditions, and therefore bind to different ions[2,29,60]. This adaptability is not only helpful for the virus in nature but is also likely responsible for TMV's ability to nucleate many different types of nanoparticles.

1.4.1 TMV-CP modification Strategies

Mutation of TMV has become a common strategy to either enhance affinity to certain inorganic salt precursors or to create an affinity which is not present in WT TMV. In fact, most recent studies use modified TMV CP instead of the WT, as it is generally more adaptable to the purpose of the study when compared to the WT TMV. The engineered modification either

enhances a WT binding effect or introduces a unique chemical moiety which can then be used for purposes not possible with the WT CP. These CP modifications can be introduced through chemical modification of the CP or through genetic mutation of the CP with each strategy having benefits and drawbacks.

Mutation

Point mutation is the most common method of introducing binding/nucleation sites. This method modifies the genetics of the CP by swapping out a single amino acid with a similar but chemically unique amino acid. Since the protein capsid is assembled from a single coat protein, modification of a single amino acid residue can lead to a unique site with highly controlled three-dimensional spacing, which is still available for further chemical modification. Since only about 25 out of 158 of the residues are critical to TMV rod formation[70], there are many residues which can be treated in this fashion without having a major effect on TMV self assembly while introducing a targeted, uniquely reactive site to the CP[51,55,73,74,76,88,95]. A similar strategy is to add a single amino acid to the sequence at one of the termini. The advantage of the point mutation method is that since the binding site is introduced at the genetic level, each and every CP is guaranteed to possess the desired functionality, leading to precise and reproducible spacing between binding sites in the final TMV assembly. The downside is that although sites which will likely not seriously affect the TMV self assembly can be selected, we still do not understand enough about protein folding to be sure of what the full effects of changing the amino acid sequence will be. Some strains of TMV have significantly different AA sequence but are still effective, while other point mutations have led to stability or self assembly problems[24,52,53,70,96]. Therefore, each new mutant has to be fully characterized. Ideally, there will be only minor changes to the assembly and growth properties of the TMV-CP, and the purification of CP will remain unchanged, but in reality it is currently impossible to predict with 100% accuracy the full outcomes of protein sequence modification.

Chemical Modification of WT-TMV-CP

It is also possible to introduce binding sites through chemical modification of one of the surface sites of the virus[97–99]. The benefits of this method are that it is simpler and can usually

be prepared more quickly than a new mutant. Furthermore, since the WT protein is used as the base, purification of the chemically modified TMV is not an issue, and changes in self assembly properties are less likely using this method compared to mutation. The drawbacks are that usually it is challenging to find an ideal chemical reaction which will react with 100% of the available sites without producing any unwanted secondary side reactions. These secondary reactions, which introduce undesired modifications, in conjunction with the reduced yield lead overall to a reduction in the reliability of the spacing between the modified sites when compared to mutation derived modification. Therefore, for applications where maximum surface coverage is desired, chemical modification will usually perform better by introducing additional moieties but applications which require precise and reproducible spacing between functional groups are more suited to the mutation modification route.

The various modification strategies give TMV based research a broad range of potential applications. Commonly the TMV surface is modified in order to either improve its affinity for precursor salts or to create an affinity for precursor salts. For example, this was done in the case of Au, Pd and Ag nanoparticles which were nucleated and bound more effectively to the mutant TMV compared to the WT TMV, due to the increased affinity between TMV and precursor leading to an increased loading of the precursor near the virus surface[83,95]. Since in TMV based nanoparticle nucleation, the effect is based on charged amino acid groups, the resulting nanoparticles are usually then bound to the TMV. As such, TMV and TMV-CP can also be used to bind previously fabricated nanoparticles electrostatically[76,89]. Regardless of the method used to bind the nanoparticles, the nanoparticles are usually bound irreversibly to the CP as long as the TMV remains assembled[89].

1.4.2 TMV as a Template for Future Devices

As mentioned previously the charged, surface accessible groups of the TMV are able to preferentially nucleate and bind various types of nanoparticles through electrostatic interaction with the salt precursors. Using a combination of TMV-CP modification and solution conditions, not only can many types of nanoparticles/nanowires be produced, but also in some cases such

as with Cu it is possible to select the inner channel or the outer surface for Cu growth[91,94]. Such selection is possible because the amino acid makeup is different on the surface of the inner channel compared to the outer surface. In this fashion TMV has been shown to be able to nucleate and/or bind many types of nanoparticles including: CdS, PbS, Pt, Pd, Ag, Au which have promising optical and electronic properties which could be useful for future device construction[76,81,83,86–89,95]. For example, TMV decorated with Pt nanoparticles exhibit a novel electronic memory effect which could be used to as building blocks in next generation electronic devices[81]. Interestingly, it is seems that the effect is attained, not solely through the organization of the Pt nanoparticles by the TMV but also through the charge donor capacity of the embedded TMV RNA[81]. There are other examples of TMV's ability to donate electrons through interactions with the RNA bases and as such it is a rather unique template[100].

TMV can also be "activated" through binding of Pd precursor prior to introduction of metal precursors. This greatly increases the types of nanoparticles which can be nucleated and bound to the TMV, which in turn greatly increases the types of devices which can be made using TMV based nucleation and assembly. For example Cu, Ni and Co nanoparticles/nanowires bound/embedded in TMV can be created using this method, while TMV remains inert to those precursors without prior Pd activation[82,91,92,94]. Nanowires produced using this method have a diameter of only 3 nm. Furthermore, TMV based nanowires also have the ability to be an alloy of the metals, which further increases their potential for next generation electronic devices[82].

Some applications of TMV are based more directly on TMV and only require a few of its unique properties. One such application is TMV based battery electrodes, which take advantage of TMVs rigidity, small size and long aspect ratio. The construction of TMV based battery electrodes involves TMV deposited on a surface, with the virus axis perpendicular to the surface, and coated in nickel which leads to a vastly increased surface area compared to a flat surface due to the size and shape of the TMV, with the end result being a significant increase in electrode capacity[84,85].

TMV based technology has the potential to be applied in a broad variety of fields. Besides the previously mentioned applications, TMV has shown promise in magnetic

applications and future hard drive construction,[101,102] energy storage,[78] as well as in catalysis[78,103,104]. Furthermore, TMV has more recently been becoming more interesting in the emerging field of biotechnology with TMV based assemblies showing potential in producing enhanced vascular grafts ,high-relaxivity MRI contrast agents, and vaccines[105–108].

1.5 Introduction to Poly A

Poly A RNA is an RNA strand which is composed solely of adenine bases., Thus it can be easily synthesised without the need for the more involved and costly RNA synthesis methods required for a more complex RNA sequence which typically results in low yields and is normally constrained to produce only short RNA sequences reliably[109–112]. Poly A is able to form a double stranded helix with itself in acidic conditions[113,114]. However, near neutral pH Poly A is single stranded, but still exhibits a lot of secondary structure, and exists not as a random coil but as a mixture of coiled regions and helical regions[115–117]. In the work described herein, only near neutral conditions were used,so only the currently relevant single stranded form of poly A will be considered. The helix is right handed with a pitch of 2.54 nm and a radius of 1.07 nm[118,119]. There are nine residues per period, and the phosphates are on the outer edge of the helix while the adenine bases all point towards the center. The bases are held at an angle of about 66^{o} with respect to the helix axis, and interact with each other mainly through base stacking but also through interbase hydrogen bonding[119],[120].

On average, the helical regions contain 30 - 50 bases with coiled regions separating the helices[116,119,121,122]. The position and angle of the bases in the helical region lead to a very strong, characteristic CD spectrum which is temperature dependent, since only the helical regions contribute[123]. Helix formation is driven by enthalpy, with the helical conformation being only slightly more favorable than the coiled conformation at room temperature[115]. Unlike in most double stranded helices, in single stranded poly A, helix formation and denaturation is non-cooperative, and even the dimer but not the monomer of poly A shows considerable secondary structure[115,122]. In addition, the energy difference between the helical and the coiled form of poly A is not large. Thus the helix and coiled states are in dynamic equilibrium with each

other[115,117]. At 25 °C the rate constant for helix formation has been measured as 7.0 x 10^7 s^{-1} with an activation energy of 4.0 kcal/mole, while the reverse rate constant for coil formation was found to be 3.2 x 10^6 s^{-1} with an activation energy of 15.2 kcal/mole[117]. The coiled form is favored at higher temperature, as the formation of the coiled form is driven by entropy. Thus at lower temperatures, most of the Poly A will be in helix form, and at higher temperatures the coiled form dominates[123–125]. As the temperature is increased, the relative amount of helix in a given poly A strand will decrease, resulting in attentuated CD although the position of the peaks remains the same[123]. The enthalpy of helix formation for poly A is between 3 - 10 kcal/mole, with different methods yielding different results[125–128]. Regardless of the exact value of the enthalpy gained from helix formation, there is still a strong driving force for poly A to form the helix, as even at 90 °C there is still 9% of the base population which is still in the helical state[125].

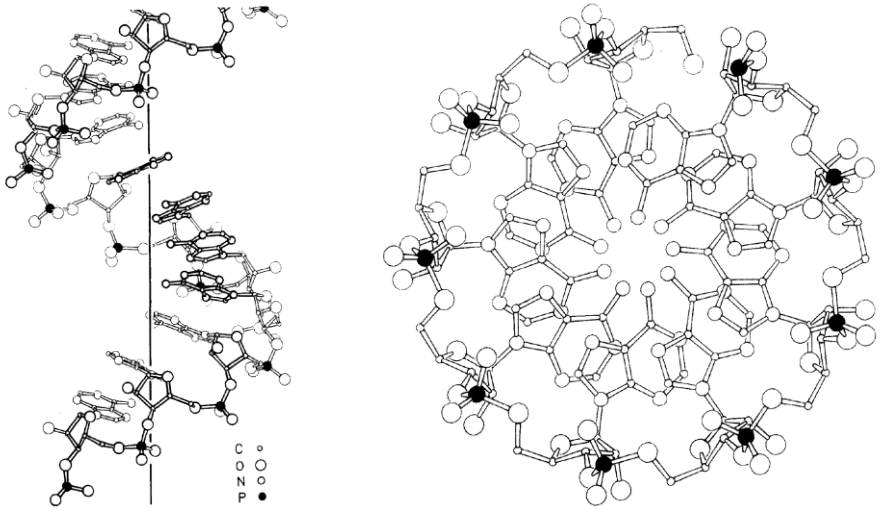


Figure 1.7, Structure of the helical portion of poly A viewed from the side (right) and from above (left). A period of the helix is made up from 9 nucleotides with the phosphates on the outer radius and the bases pointed towards the central axis. The bases are tilted and stack with each other. Adapted from Saenger et al.[118] Reprinted from the *Journal of Molecular Biology*, 93 (4), 529–534, Saenger, W., Riecke, J., Suck, D. A, Structural Model for the Polyadenylic Acid Single Helix, 1975, used with permission from The Academic Press

Although an enormous amount of research has been done on the interaction between TMV RNA and TMV-CP, very little has been done on the interaction between Poly A and TMV-CP. Poly A also nucleates TMV rod formation and has been shown to form rods spectroscopically[4,27,36,41], It will be demonstrated here for the first time, using visual evidence, that the rods formed have the same shape and size as those seen through the use of native

RNA genome. Although very little research has been done on Poly A mediated rod formation compared to the extensive research that has been done on native TMV RNA mediated rods, the results of the research which has been done is highly informative.

Fraenkel-Conrat and Singer were the first to study the interaction between TMV-CP and synthetic polynucleotides by pelleting the resulting complex and measuring the OD after redispersal[36]. Out of all the synthetic polymers which were tested, only Poly A, polynucleotides containing mostly A and possibly poly I (which is closely related to Poly A) are able to interact with CP with reasonable success[36]. The interaction of TMV-CP with synthetic polymers in general was found to have stability issues, with Poly G and Poly U barely able to form complexes with the CP. In those cases the interaction was weak and only occurred with a large excess of the poly nucleotide, with significantly less than expected yields of nucleoprotein material[36]. Poly C was not found to have any interaction with the CP. Poly I was able to form the expected amount of stable complexes with the CP under certain conditions, but the formed complex was not stable[36]. Poly A formed stable complexes with the CP but the rods begin to lose their stability above pH 8.2[36]. The stability of complexes between poly nucleotides with an excess of A bases was not tested.

The term complex must be used instead of rod, since in their experiments Fraenkel-Conrat and Singer defined rod as pelletable material which could be redisbursed and contained RNA. However Schon et al. have since shown that short poly A strands under certain conditions interact with the CP and are protected from nuclease attack, but do not have the same characteristics as the rod particle which is seen when CP interacts with the native genome[27]. To be protected from nuclease attack, the RNA would have to be at least somewhat inaccessible to solution. Therefore it is likely that the interaction occurs at or near the same RNA binding region, but with the interaction being weak does not lead to a change in conformation of the TMV-CP. Such an interaction would most likely result in material that was also pelletable. Thus, with the exception of poly A, although some of the other polynucleotides are able to interact with TMV-CP with varying success, the other polynucleotides have not been shown conclusively

to interact with the CP in a way which leads to a final product identical to the rod particle formed from native TMV RNA and CP.

Not surprisingly, when poly I and Poly A were compared, poly A was able to outcompete poly I for TMV-CP[36]. Interestingly, although the resulting rods were found to not be as stable as with the native genome, Poly A appears to nucleate TMV-like rods over a greater variety of solution conditions than the native genome is able to nucleate TMV rods[36], possiblly due to the fact that poly A most likely uses an alternative mechanism of rod nucleation which is more versatile than that of the native TMV genome's OAS. Certainly the interaction between Poly A and TMV-CP seems to be comparable to that of the native TMV RNA with CP as at pH 6.6 in phosphate buffer, where both poly A and TMV RNA are able to nucleate rods, the poly A was able to outcompete the native TMV RNA for CP and form poly A containing TMV-like rods preferentially to TMV rods[36].

TMV-like rods formed using poly A as the RNA binding strand have never been directly visually imaged using a high resolution technique such as TEM. However, unlike in the case with the other polynucleotides, the poly A TMV-CP interaction has been fully characterized as being identical to that between the TMV RNA and CP through optical techniques such as CD, fluorescence and UV difference spectroscopy and using CsCl density gradients to measure the amount of bound RNA per CP[27,41,129,130]. Any lingering doubt that poly A is indeed able to form TMV-like rods will be put to rest in the following chapter.

Although Poly A forms TMV-like rods which are identical to the rods formed from TMV-RNA, due to the small amount of research that has been done on Poly A, there are still many questions as to the mechanism through which Poly A nucleates TMV rods. Some information about Poly A nucleated rods is known. For example, Poly A has been shown to exhibit second order kinetics with respect to the concentration of disk, which has a 1st order dependence when interacting with the native genome[1,41]. A study done by Schon et al. found that four turns of helix, or two disk structures, were required to form stable proto rods which were capable of elongation, while only a single disk is necessary for rod nucleation in the rods formed using the native genome[27]. It was found that poly A which did not achieve at least four turns of helix were

unstable and could dissociate back into poly A and TMV disk[27]. Butler also performed a similar study which concluded that in order to form a stable proto helix using poly A, at least two disk structures must be added and changed from planar to helix conformation[41]. Schon et al. found another oddity unique to poly A mediated TMV-like rod formation. They found that base destacking in poly A strands could be made, under certain conditions, to precede the change in helicity of the CP and the increase in turbidity from rod formation[27]. Although different explanations are possible, one explanation would be that under certain solution conditions, once the four turn helix is formed, the remaining, unbound portion of the poly A strand "feels" the binding event which creates a helix destacking "wave" which is projected out ahead of the growing TMV-like rod. If this were the case it would suggest that it was the coiled region of the Poly A which is the nucleating region. This would make sense since at room temperature the coiled form is the more unstable of the two states,[117] and would therefore be more likely to form favorable interactions with the highly constraining TMV-CP.

Schon et al. also found that although native TMV RNA would finishing assembling all possible rods slightly faster than poly A under the conditions tested, a significant amount of the full length (5000 bases) poly A strands were completely protected from nuclease attack within 40s[27]. This is very different from the behavior of the native RNA with CP, where essentially all strands are nucleated very quickly and then addition of further protein takes longer with effectively all the RNA strands growing TMV rods simultaneously and the fastest full protection taking about 6 minutes[17,26]. Therefore, unlike in TMV, where rod elongation is the rate determining step in complete formation of TMV rods, in poly A mediated TMV-like rods, elongation is very fast, at least 6x faster than with TMV-CP, but nucleation is slow[27]. It would make sense that the TMV-RNA would be more effective at nucleating TMV rods, as it has been very carefully selected for in nature to be incredibly efficient at nucleating TMV rods, and occurs preferentially to rod elongation with 99% of strands nucleated within 1 minute under normal conditions and in 10 s under ideal conditions[26]. On the other hand Poly A does not have the ideal sequence for nucleating TMV rods, which is likely at least part of the reason why two nucleating disks are necessary to form a stable particle.

In terms of elongation however, the stem loop structure would put constraints on how quickly addition of protein could be completed, as the RNA has to travel up the loop and expose itself before it can bind to any further CP. Furthermore adenine is one of the most strongly binding bases to TMV-CP,[42,56] and while no studies have been done to measure the effect of base sequence on the rate of elongation of TMV rods, it is a definite possibility that since poly A has a strong interaction with the CP it could result in faster rod elongation. It is not clear from the currently available literature whether poly A is able to add protein quickly in both directions or if only the 3' end of poly A is able to nucleate TMV like rods. However, since there is no doubling back of the 5' end in poly A, which would likely at least interfere with elongation, it is possible that poly A is able to elongate quickly in both directions, unlike the native virus which elongates significantly faster in the 5' direction compared to the 3' [17,24,48,54].

Regardless of where and how poly A is able to nucleate TMV like rods since the stable protorod intermediary, elongation kinetics and nucleation kinetics of poly A mediated TMV rod formation differ fundamentally from those of native TMV RNA, the mechanism of assembly is most likely quite different. Furthermore, based on the known properties of poly A, there is simply no conceivable way for it to form any kind of stable stem loop structure. In addition the poly A sequence was found to be a poor nucleator of TMV-like rods when confined to using the stem loop mechanism, yet poly A manages to assemble TMV-like rods just the same. When these facts are considered in conjunction with one another there is simply no conceivable way that poly A could use a similar method of rod nucleation to that employed by the OAS.

Poly A of similar length to the native TMV genome gained full resistance to RNAse after about 5 minutes which is an overall rate which approaches that of the native genome. However a significant amount of the full length poly A gained full RNAse resistance after only 40 seconds. If nucleation took place at or near the 3' end this would constitute a rate of rod elongation which was over 6x that of the native RNA. Even considering that in poly A the RNA does not need to be pulled through the central channel, and assuming that the high affinity adenine base may lead to a reasonable increase in the rate of elongation, it is difficult to envision a rate difference of that magnitude, as in both cases RNA must be incorporated and the same protein

rearrangements must take place to form the helical rod structure. Given that there were a variety of different times at which various poly A strands become completely resistant to RNAse, and considering that all segments of the poly A are essentially equivalent, it seems more likely that the variance in the poly A rod completion times was due in part to a nucleation region within the poly A which was chosen at random. Unlike with OAS based rod assemblies, with poly A the 5' end would not interfere with elongation towards the 3' end of the poly A. If the increased affinity between poly A and the CP could significantly increase the rate of addition towards the 3', perhaps through the addition of disks, it would account for the fast resistance of poly A to nuclease without requiring a massive increase in the rate of elongation in the 5' direction. Furthermore since only early nucleation events which occurred somewhere near the middle of the poly A strand would finish in 40 seconds, most of the poly A would only be protected at a later time, thus being consistent with the experimental observations. Regardless of exactly how nucleation and elongation is achieved in poly A, the fact that there are such noticeable differences to OAS based assemblies makes poly A unique and thus likely means that there are some conditions under which only poly A is able to assemble TMV-like rods and not OAS based RNA and visa versa which make further poly A research paramount as it may eventually lead to new types of assemblies not possible with current TMV OAS technology.

1.6 Future of TMV based Technologies

TMV has already shown remarkable capacity to form the basis of components which could be used in next generation devices. However, there are many challenges to overcome before TMV based technology can be incorporated into actual devices on a broader scale. For most future applications having multiple components on the same structure would greatly increase the versatility and usability of TMV based devices. Although RNA mediated TMV rod assembly is more involved, as it requires additional steps, it is also more attractive for more complex TMV based assembly. It is attractive not only for the added stability to various working conditions which the RNA provides but also because it provides an additional degree of control,

not only over the length of the TMV rod but also over the final composition of TMV rods with multiple domains which will be required for future devices[131]. The attributes which make RNA mediated TMV rod assembly attractive can be potentially be further enhanced even further by using poly A in the place of native TMV RNA. Future TMV based devices should also have a macroscopic handle to allow for human control. As such the development of TMV rod tethering to a solid support, such as a surface, would provide the basis for a very versatile template which could be applied to a large number of applications. Furthermore a conductive surface would further increase the types of devices which the ideal TMV template could be used for and could still be used for nonconductive applications.

Therefore RNA mediated TMV rod assembly research is very useful for development of future technology. Using Poly A as the nucleating RNA instead of the native genome or other OAS based sequences would provide numerous advantages to future TMV based devices. Poly A is significantly cheaper and easier to produce than any RNA strands based on the origin of assembly. The length of the poly A strand used can easily and inexpensively be selected for, while both development and production of a desired length of OAS based RNA strand is costly and involved. Poly A has a unique assembly mechanism which, although is less specialized than OAS based RNA, is also more adaptable and therefore more likely to be useful for a larger variety of situations. Since poly A has absolutely no part of the native genome, TMV-like rods containing poly A are completely non-infectious, an important consideration since the native TMV genome is highly infectious and thus potentially dangerous to any local plant life[132].

One of the properties which makes TMV attractive is that large amounts of it can be produced cheaply, and each copy produced will be exactly the same, so any TMV based technology has the potential to be highly scalable at low cost[100]. TMV mutants tend to be more interesting, but unfortunately there is no guarantee that a desired mutant will respond well to production. In order to fabricate TMV based devices with multiple domains it would be necessary to further modify the TMV-CP and then reassemble in vitro the more complex TMV rods. Although the TMV RNA could also be purified, the length of that RNA would in most cases have to be altered which would again be a challenging and costly endeavour. Furthermore that

TMV RNA would lead to devices with infectious components and depending on the CP modifications it is possible that the TMV RNA would no longer be able to assemble the desired rod-like structure. Poly A seems to have a more adaptable nucleation mechanism, thus it is likely better able to handle currently unforeseen assembly complications. Although some of the issues with OAS based RNA can be overcome, only fully synthetic RNA such as poly A can overcome all of these issues simultaneously and cheaply, thus it is a very attractive candidate for use in the construction of future TMV based nanotechnology and is deserving of further study.

Poly A is also attractive from a fundamental viewpoint. Very little research has been done on Poly A nucleated TMV rods, so the mechanism is not well understood. A better understanding of the mechanism of poly A mediated TMV-like rod assembly may lead to further technological possibilities in the future. There is also currently an inherent uncertainty with regards to the RNA - CP interaction of individual base triplets in TMV as the high resolution techniques produce a TMV RNA genome weighted average RNA signal for each of the three CP binding positions. By using poly A mediated TMV rods to construct a high resolution electron density map the interaction between adenine bases and TMV-CP could be determined with a high degree of confidence and might even give additional insight into interactions between the CP and the other bases found in TMV RNA.

Given the potential of TMV based technology, it would be helpful to have as large an array of tools to confront potential roadblocks along the way. Poly A is a promising alternative to OAS based TMV-like rod production which has more attractive properties than OAS based RNA for incorporation into actual devices. In addition because the method of assembly is fundamentally different, poly A has the potential to present novel applications and to uniquely be applicable to potential situations where OAS based assemblies do not function properly.

The main issue with poly A is one of stability, with poly A rods shown to not be as stable as native TMV rods. This is strange as A bases have very favorable interaction with the CP while the native TMV RNA contains many C and U bases which do not interact particularly favorably with TMV. Thus logically, poly A should have greater stability than WT TMV but it does not.

However, there are a variety of methods of increasing the stability of poly A rods, many of which are quite simple, especially compared to the cost of modifying OAS based RNA sequences. Poly A likely is significantly more stable in solution conditions where calcium ions are bound to the TMV-like rod, as poly A contains an adenine in every 3rd position, and thus should gain a significant stability increase in high calcium conditions. The fact that synthetic RNA forms less stable TMV-like rods would suggest that having all bases of the same identity causes some sort of additional energy cost which is not present when more than one base is present. More detailed high resolution studies would be necessary to understand the reason behind the experimental results. However, one solution might be to use a random co-polymer with A and U in a 5:1 A:U ratio. It was found that such a synthetic RNA strand was able to assemble TMV-like rods but the stability of those rods was not investigated. Such an RNA strand would still have a majority of adenine and thus should interact just as favorably with TMV-CP as poly A, but might have additional stability if it is indeed "too much adenine" which is preventing poly A from achieving the same stability as the native TMV RNA. Finally, a strategy which also works for increasing native TMV stability but which should be particularly helpful in the case of poly A is CP mutation. By mutating out the charged residues responsible for TMV rod disassembly the formed rod or device would essentially be completely stable. In nature rod disassembly is crucial for Virus replication but such disassembly is not desired in nanotechnology applications thus by using the mutated CP, which has the same self-assembly characteristics as the WT, the rod assembly would essentially become irreversible and thus solve any stability issues, even in the case of poly A. Surely one or more of the above strategies would be sufficient to solve any potential poly A stability issues and thus it remains a highly viable option going forward in TMV based bionanotechnology.

References

(1) Butler, P. J. G. Self–assembly of Tobacco Mosaic Virus: The Role of an Intermediate Aggregate in Generating Both Specificity and Speed. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1999**, *354* (1383), 537–550.

(2) Namba, K.; Pattanayek, R.; Stubbs, G. Visualization of Protein-Nucleic Acid Interactions in a Virus: Refined Structure of Intact Tobacco Mosaic Virus at 2.9 Å Resolution by X-Ray Fiber Diffraction. *J. Mol. Biol.* **1989**, *208* (2), 307–325.

(3) Klug, A. The Tobacco Mosaic Virus Particle: Structure and Assembly. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1999**, *354* (1383), 531–535.
(4) Butler, P. J. G.; Klug, A. Assembly of the Particle of Tobacco Mosaic Virus from RNA and Disks of Protein. *Nature* **1971**, *229* (2), 47–50.
(5) Durham, A. C. H.; Finch, J. T.; Klug, A. States of Aggregation of Tobacco Mosaic Virus Protein. *Nature. New Biol.* **1971**, *229*, 37–42.
(6) Butler, P. J. G. The Current Picture of the Structure and Assembly of Tobacco Mosaic Virus. *J. Gen. Virol.* **1984**, *65* (2), 253–279.
(7) Caspar, D. L. D.; Keiichi Namba. Switching in the Self-Assembly of Tobacco Mosaic Virus. *Adv. Biophys.* **1990**, *26* (0), 157–185.
(8) Culver, J. N. Tobacco Mosaic Virus Assembly and Dissasembly: Determinants in Pathogenicity and Resistance. *Annu. Rev. Phytopathol.* **2002**, *40* (1), 287–308.
(9) Kegel, W. K.; van der Schoot, P. Physical Regulation of the Self-Assembly of Tobacco Mosaic Virus Coat Protein. *Biophys. J.* **2006**, *91* (4), 1501–1512.
(10) Durham, A. C. H.; Klug, A. Polymerization of Tobacco Mosaic Virus Protein and Its Control. *Nature* **1971**, *229* (2), 42–46.
(11) Durham, A. C. H.; Finch, J. T. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: II. Electron Microscope Observations of the Larger Polymers. *J. Mol. Biol.* **1972**, *67* (2), 307–314.
(12) Butler, P. J. G.; Durham, A. C. H.; Klug, A. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: IV. Control of Mode of Aggregation of Tobacco Mosaic Virus Protein by Proton Binding. *J. Mol. Biol.* **1972**, *72* (1), 1–18.
(13) Mandelkow, E.; Stubbs, G.; Warren, S. Structures of the Helical Aggregates of Tobacco Mosaic Virus Protein. *J. Mol. Biol.* **1981**, *152* (2), 375–386.
(14) Bloomer, A. C.; Champness, J. N.; Bricogne, G.; Staden, R.; Klug, A. Protein Disk of Tobacco Mosaic Virus at 2.8 |[angst]| Resolution Showing the Interactions within and between Subunits. *Nature* **1978**, *276* (5686), 362–368.
(15) Finch, J. T.; Leberman, R.; Chang, Y.-S.; Klug, A. Rotational Symmetry of the Two Turn Disk Aggregate of Tobacco Mosaic Virus Protein. *Nature* **1966**, *212*, 349–350.
(16) Finch, J. T.; Klug, A. Three-Dimensional Reconstruction of the Stacked-Disk Aggregate of Tobacco Mosaic Virus Protein from Electron Micrographs. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1971**, *261* (837), 211–219.
(17) Turner, D. R.; McGuigan, C. J.; Butler, P. J. G. Assembly of Hybrid RNAs with Tobacco Mosaic Virus Coat Protein: Evidence for Incorporation of Disks in 5'-Elongation along the Major RNA Tail. *J. Mol. Biol.* **1989**, *209* (3), 407–422.
(18) Namba, K.; Stubbs, G. Structure of Tobacco Mosaic Virus at 3.6 A Resolution: Implications for Assembly. *Science* **1986**, *231* (4744), 1401–1406.
(19) Correia, J. J.; Shire, S.; Yphantis, D. A.; Schuster, T. M. Sedimentation Equilibrium Measurements of the Intermediate-Size Tobacco Mosaic Virus Protein Polymers. *Biochemistry (Mosc.)* **1985**, *24* (13), 3292–3297.
(20) Raghavendra, K.; Salunke, D. M.; Caspar, D. L. D.; Schuster, T. M. Disk Aggregates of Tobacco Mosaic Virus Protein in Solution: Electron Microscopy Observations. *Biochemistry (Mosc.)* **1986**, *25* (20), 6276–6279.
(21) Raghavendra, K.; Adams, M. L.; Schuster, T. M. Tobacco Mosaic Virus Protein Aggregates in Solution: Structural Comparison of 20S Aggregates with Those near Conditions for Disk Crystallization. *Biochemistry (Mosc.)* **1985**, *24* (13), 3298–3304.

(22) Steat, D. E.; Turner, P. C.; Finch, J. T.; Butler, P. J. G.; Wilson, T. M. A. Packaging Ofrecombinant RNA Molecules into Pseudovirus Particles Directed by the Origin-of-Assembly Sequence from Tobacco Mosaic Virus RNA. *Virology* **1986**, *155*, 299–308.
(23) Fukuda, M.; Okada, Y. Elongation in the Major Direction of Tobacco Mosaic Virus Assembly. *Proc. Natl. Acad. Sci.* **1985**, *82* (11), 3631–3634.
(24) Okada, Y. Molecular Assembly of Tobacco Mosaic Virus in Vitro. *Adv. Biophys.* **1986**, *22*, 95–149.
(25) Ohno, T.; Takahashi, M.; Okada, Y. Assembly of Tobacco Mosaic Virus in Vitro: Elongation of Partially Reconstituted RNA. *Proc. Natl. Acad. Sci.* **1977**, *74* (2), 552–555.
(26) Butler, P. J. G. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: IX. Initial Stages of Assembly of Nucleoprotein Rods from Virus RNA and the Protein Disks. *J. Mol. Biol.* **1974**, *82* (3), 343–353.
(27) SCHÖN, A.; MUNDRY, K.-W. Coordinated Two-Disk Nucleation, Growth and Properties, of Virus-like Particles Assembled from Tobacco-Mosaic-Virus Capsid Protein with Poly (A) or Oligo (A) of Different Length. *Eur. J. Biochem.* **1984**, *140* (1), 119–127.
(28) McMichael, J. C.; Lauffer, M. A. A Specific Effect of Calcium Ion on the Polymerization-Depolymerization of Tobacco Mosaic Virus Protein. *Arch. Biochem. Biophys.* **1975**, *169* (1), 209–216.
(29) Knez, M.; Sumser, M. P.; Bittner, A. M.; Wege, C.; Jeske, H.; Hoffmann, D. M. P.; Kuhnke, K.; Kern, K. Binding the Tobacco Mosaic Virus to Inorganic Surfaces. *Langmuir* **2004**, *20* (2), 441–447.
(30) Holder, P. G.; Finley, D. T.; Stephanopoulos, N.; Walton, R.; Clark, D. S.; Francis, M. B. Dramatic Thermal Stability of Virus−Polymer Conjugates in Hydrophobic Solvents. *Langmuir* **2010**, *26* (22), 17383–17388.
(31) Matthews, R. E. F. Reconstitution of Turnip Yellow Mosaic Virus RNA with TMV Protein Subunits. *Virology* **1966**, *30* (1), 82–96.
(32) Fraenkel-Conrat, H.; Singer, B. Reconstitution of Tobacco Mosaic Virus 3. Improved Methods and the Use of Mixed Nucleic Acids. *Biochim. Biophys. Acta* **1959**, *33*, 359–370.
(33) Jockusch, H. The Role of Host Genes, Temperature and Polyphenoloxidase in the Necrotization of TMV Infected Tobacco Tissue. *J. Phytopathol.* **1966**, *55* (2), 185–192.
(34) Sturtevant, J. M.; Velicelebi, G.; Jaenicke, R.; Lauffer, M. A. Scanning Calorimetric Investigation of the Polymerization of the Coat Protein of Tobacco Mosaic Virus. *Biochemistry (Mosc.)* **1981**, *20* (13), 3792–3800.
(35) Jardetzky, O.; Akasaka, K.; Vogel, D.; Morris, S.; Holmes, K. C. Unusual Segmental Flexibility in a Region of Tobacco Mosaic Virus Coat Protein. *Nature* **1978**, *273* (5663), 564–566.
(36) Fraenkel-Conrat, H.; Singer, B. Reconstitution of Tobacco Mosaic Virus IV. Inhibition by Enzymes and Other Proteins, and Use of Polynucleotides. *Virology* **1964**, *23* (3), 354–362.
(37) Watson, J. D. The Structure of Tobacco Mosaic Virus: I. X-Ray Evidence of a Helical Arrangement of Sub-Units around the Longitudinal Axis. *Biochim. Biophys. Acta* **1954**, *13*, 10–19.
(38) Zimmern, D. The Nucleotide Sequence at the Origin for Assembly on Tobacco Mosaic Virus RNA. *Cell* **1977**, *11* (3), 463–482.
(39) Butler, P. J. G.; Finch, J. T.; Zimmern, D. Configuration of Tobacco Mosaic Virus RNA during Virus Assembly. *Nature* **1977**, *265*, 217–219.
(40) Hwang, D.-J.; Roberts, I. M.; Wilson, T. M. Expression of Tobacco Mosaic Virus Coat Protein and Assembly of Pseudovirus Particles in Escherichia Coli. *Proc. Natl. Acad. Sci.* **1994**, *91* (19), 9067–9071.
(41) Butler, P. J. G. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: VI. Assembly of the Nucleoprotein Rods of Tobacco Mosaic Virus from the Protein Disks and RNA. *J. Mol. Biol.* **1972**, *72* (1), 25–35.

(42) Turner, D. R.; Joyce, L. E.; Butler, P. J. G. The Tobacco Mosaic Virus Assembly Origin RNA: Functional Characteristics Defined by Directed Mutagenesis. *J. Mol. Biol.* **1988**, *203* (3), 531–547.
(43) Otsuki, Y.; Takebe, I.; Ohno, T.; Fukuda, M.; Okada, Y. Reconstitution of Tobacco Mosaic Virus Rods Occurs Bidirectionally from an Internal Initiation Region: Demonstration by Electron Microscopic Serology. *Proc. Natl. Acad. Sci.* **1977**, *74* (5), 1913–1917.
(44) Gallie, D. R.; Plaskitt, K. A.; Michael, T.; Wilson, A. The Effect of Multiple Dispersed Copies of the Origin-of-Assembly Sequence from TMV RNA on the Morphology of Pseudovirus Particles Assembled< I> in Vitro</i>. *Virology* **1987**, *158* (2), 473–476.
(45) Butler, P. J. G.; Bloomer, A. C.; Finch, J. T. Direct Visualization of the Structure of the “20 S” Aggregate of Coat Protein of Tobacco Mosaic Virus: The “disk” Is the Major Structure at pH 7.0 and the Proto-Helix at Lower pH. *J. Mol. Biol.* **1992**, *224* (2), 381–394.
(46) Turner, D. R.; Mondragon, A.; Fairall, L.; Bloomer, A. C.; Finch, J. T.; Van Boom, J. H.; Butler, P. J. G. Oligonucleotide Binding to the Coat Protein Disk of Tobacco Mosaic Virus. *Eur. J. Biochem.* **1986**, *157* (2), 269–274.
(47) Schuster, T. M.; Scheele, R. B.; Adams, M. L.; Shire, S. J.; Steckert, J. J.; Potschka, M. Studies on the Mechanism of Assembly of Tobacco Mosaic Virus. *Biophys. J.* **1980**, *32* (1), 313–329.
(48) Lomonossoff, G. P.; Butler, P. J. G. Location and Encapsidation of the Coat Protein Cistron of Tobacco Mosaic Virus. *Eur. J. Biochem.* **1979**, *93* (1), 157–164.
(49) Kraft, D. J.; Kegel, W. K.; van der Schoot, P. A Kinetic Zipper Model and the Assembly of Tobacco Mosaic Virus. *Biophys. J.* **2012**, *102* (12), 2845–2855.
(50) Culver, J. N.; Dawson, W. O.; Plonk, K.; Stubbs, G. Site-Directed Mutagenesis Confirms the Involvement of Carboxylate Groups in the Disassembly of Tobacco Mosaic Virus. *Virology* **1995**, *206* (1), 724–730.
(51) Lu, B.; Stubbs, G.; Culver, J. N. Carboxylate Interactions Involved in the Disassembly of Tobacco Mosaic Tobamovirus. *Virology* **1996**, *225* (1), 11–20.
(52) Hariharasubramanian, V.; Zaitlin, M.; Siegel, A. A Temperature-Sensitive Mutant of TMV with Unstable Coat Protein. *Virology* **1970**, *40* (3), 579–589.
(53) Rappaport, I.; Wu, J.-H. Release of Inhibited Virus Infection Following Irradiation with Ultraviolet Light. *Virology* **1962**, *17* (3), 411–419.
(54) Lomonossoff, G. P.; Butler, P. J. G. ASSEMBLY OF TOBACCO MOSAIC VIRUS: ELONGATION TOWARDS THE 3’-HYDROXYL TERMINUS OF THE RNA. *FEBS Lett.* **1980**, *113* (2), 271–274.
(55) Rego, J. M.; Lee, J.-H.; Lee, D. H.; Yi, H. Biologically Inspired Strategy for Programmed Assembly of Viral Building Blocks with Controlled Dimensions. *Biotechnol. J.* **2013**, *8* (2), 237–246.
(56) Steckert, J. J.; Schuster, T. M. Sequence Specificity of Trinucleoside Diphosphate Binding to Polymerized Tobacco Mosaic Virus Protein. *Nature* **1982**, *299* (5878), 32–36.
(57) Yi, H.; Rubloff, G. W.; Culver, J. N. TMV Microarrays: Hybridization-Based Assembly of DNA-Programmed Viral Nanotemplates. *Langmuir* **2007**, *23* (5), 2663–2667.
(58) Gallagher, W. H.; Lauffer, M. A. Calcium Ion Binding by Tobacco Mosaic Virus. *J. Mol. Biol.* **1983**, *170* (4), 905–919.
(59) Powell, C. A. The Effect of Cations on the Alkaline Dissociation of Tobacco Mosaic Virus. *Virology* **1975**, *64* (1), 75–85.
(60) Ge, P.; Zhou, Z. H. Hydrogen-Bonding Networks and RNA Bases Revealed by Cryo Electron Microscopy Suggest a Triggering Mechanism for Calcium Switches. *Proc. Natl. Acad. Sci.* **2011**, *108* (23), 9637–9642.
(61) Stubbs, G. Tobacco Mosiac Virus Particle Structure and the Initiation of Disassembly. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1999**, *354* (1383), 551–557.

(62) Durham, A. C. H.; Hendry, D. A. Cation Binding by Tobacco Mosaic Virus. *Virology* **1977**, *77* (2), 510–519.
(63) Shalaby, R. A. F.; Banerjee, K.; Lauffer, M. A. Ion Binding by Tobacco Mosaic Virus and Its Protein. *Biochemistry (Mosc.)* **1968**, *7* (3), 955–960.
(64) Gallagher, W. H.; Lauffer, M. A. Calcium Ion Binding by Isolated Tobacco Mosaic Virus Coat Protein. *J. Mol. Biol.* **1983**, *170* (4), 921–929.
(65) Caspar, D. L. D. Structure of Tobacco Mosaic Virus Radial Density Distribution in the Tobacco Mosaic Virus Particle. *Nature* **1956**, *177*, 928.
(66) Butler, P. J. G.; Durham, A. C. H. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus: V. Conservation of the Abnormally Titrating Groups in Tobacco Mosaic Virus. *J. Mol. Biol.* **1972**, *72* (1), 19–24.
(67) Pattanayek, R.; Elrod, M.; Stubbs, G. Characterization of a Putative Calcium-Binding Site in Tobacco Mosaic Virus. *Proteins Struct. Funct. Bioinforma.* **1992**, *12* (2), 128–132.
(68) Bhyravbhatla, B.; Watowich, S. J.; Caspar, D. L. D. Refined Atomic Model of the Four-Layer Aggregate of the Tobacco Mosaic Virus Coat Protein at 2.4-Å Resolution. *Biophys. J.* **1998**, *74* (1), 604–615.
(69) Durham, A. C. H.; Butler, P. J. G. A Prediction of the Structure of Tobacco-Mosaic-Virus Protein. *Eur. J. Biochem.* **1975**, *53* (2), 397–404.
(70) Altschuh, D.; Lesk, A. M.; Bloomer, A. C.; Klug, A. Correlation of Co-Ordinated Amino Acid Substitutions with Function in Viruses Related to Tobacco Mosaic Virus. *J. Mol. Biol.* **1987**, *193* (4), 693–707.
(71) Holmes, K. C.; Stubbs, G. J.; Mandelkow, E.; Gallwitz, U. Structure of Tobacco Mosaic Virus at 6.7 Å Resolution. *Nature* **1975**, *254* (5497), 192–196.
(72) Stubbs, G.; Warren, S.; Holmes, K. Structure of RNA and RNA Binding Site in Tobacco Mosaic Virus from 4-Å Map Calculated from X-Ray Fibre Diagrams. *Nature* **1977**, *267* (5608), 216–221.
(73) Miller, R. A.; Presley, A. D.; Francis, M. B. Self-Assembling Light-Harvesting Systems from Synthetically Modified Tobacco Mosaic Virus Coat Proteins. *J. Am. Chem. Soc.* **2007**, *129* (11), 3104–3109.
(74) Dedeo, M. T.; Duderstadt, K. E.; Berger, J. M.; Francis, M. B. Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. *Nano Lett.* **2010**, *10* (1), 181–186.
(75) Nie, Z.; Petukhova, A.; Kumacheva, E. Properties and Emerging Applications of Self-Assembled Structures Made from Inorganic Nanoparticles. *Nat. Nanotechnol.* **2009**, *5* (1), 15–25.
(76) Zahr, O. K.; Blum, A. S. Solution Phase Gold Nanorings on a Viral Protein Template. *Nano Lett.* **2012**, *12* (2), 629–633.
(77) Tan, W. S.; Lewis, C. L.; Horelik, N. E.; Pregibon, D. C.; Doyle, P. S.; Yi, H. Hierarchical Assembly of Viral Nanotemplates with Encoded Microparticles via Nucleic Acid Hybridization. *Langmuir* **2008**, *24* (21), 12483–12488.
(78) Gerasopoulos, K.; McCarthy, M.; Banerjee, P.; Fan, X.; Culver, J. N.; Ghodssi, R. Biofabrication Methods for the Patterned Assembly and Synthesis of Viral Nanotemplates. *Nanotechnology* **2010**, *21* (5), 055304.
(79) Manocchi, A. K.; Seifert, S.; Lee, B.; Yi, H. In Situ Small-Angle X-Ray Scattering Analysis of Palladium Nanoparticle Growth on Tobacco Mosaic Virus Nanotemplates. *Langmuir* **2011**, *27* (11), 7052–7058.
(80) Bruckman, M. A.; Liu, J.; Koley, G.; Li, Y.; Benicewicz, B.; Niu, Z.; Wang, Q. Tobacco Mosaic Virus Based Thin Film Sensor for Detection of Volatile Organic Compounds. *J. Mater. Chem.* **2010**, *20* (27), 5715–5719.
(81) Tseng, R. J.; Tsai, C.; Ma, L.; Ouyang, J.; Ozkan, C. S.; Yang, Y. Digital Memory Device Based on Tobacco Mosaic Virus Conjugated with Nanoparticles. *Nat. Nanotechnol.* **2006**, *1* (1), 72–77.

(82) Balci, S.; Hahn, K.; Kopold, P.; Kadri, A.; Wege, C.; Kern, K.; Bittner, A. M. Electroless Synthesis of 3 Nm Wide Alloy Nanowires inside Tobacco Mosaic Virus. *Nanotechnology* **2012**, *23* (4), 045603.
(83) Lim, J.-S.; Kim, S.-M.; Lee, S.-Y.; Stach, E. A.; Culver, J. N.; Harris, M. T. Formation of Au/Pd Alloy Nanoparticles on TMV. *J. Nanomater.* **2010**, *2010*, 1–6.
(84) Royston, E.; Ghosh, A.; Kofinas, P.; Harris, M. T.; Culver, J. N. Self-Assembly of Virus-Structured High Surface Area Nanomaterials and Their Application as Battery Electrodes. *Langmuir* **2008**, *24* (3), 906–912.
(85) Chen, X.; Gerasopoulos, K.; Guo, J.; Brown, A.; Wang, C.; Ghodssi, R.; Culver, J. N. Virus-Enabled Silicon Anode for Lithium-Ion Batteries. *ACS Nano* **2010**, *4* (9), 5366–5372.
(86) Shenton, W.; Douglas, T.; Young, M.; Stubbs, G.; Mann, S. Inorganic–organic Nanotube Composites from Template Mineralization of Tobacco Mosaic Virus. *Adv. Mater.* **1999**, *11* (3), 253–256.
(87) Manocchi, A. K.; Horelik, N. E.; Lee, B.; Yi, H. Simple, Readily Controllable Palladium Nanoparticle Formation on Surface-Assembled Viral Nanotemplates. *Langmuir* **2010**, *26* (5), 3670–3677.
(88) Dujardin, E.; Peet, C.; Stubbs, G.; Culver, J. N.; Mann, S. Organization of Metallic Nanoparticles Using Tobacco Mosaic Virus Templates. *Nano Lett.* **2003**, *3* (3), 413–417.
(89) Khan, A. A.; Fox, E. K.; Górzny, M. Ł.; Nikulina, E.; Brougham, D. F.; Wege, C.; Bittner, A. M. pH Control of the Electrostatic Binding of Gold and Iron Oxide Nanoparticles to Tobacco Mosaic Virus. *Langmuir* **2013**, *29* (7), 2094–2098.
(90) Lewis, C. L.; Lin, Y.; Yang, C.; Manocchi, A. K.; Yuet, K. P.; Doyle, P. S.; Yi, H. Microfluidic Fabrication of Hydrogel Microparticles Containing Functionalized Viral Nanotemplates. *Langmuir* **2010**, *26* (16), 13436–13441.
(91) Balci, S.; Bittner, A. M.; Hahn, K.; Scheu, C.; Knez, M.; Kadri, A.; Wege, C.; Jeske, H.; Kern, K. Copper Nanowires within the Central Channel of Tobacco Mosaic Virus Particles. *Electrochimica Acta* **2006**, *51* (28), 6251–6257.
(92) Tsukamoto, R.; Muraoka, M.; Seki, M.; Tabata, H.; Yamashita, I. Synthesis of CoPt and FePt3 Nanowires Using the Central Channel of Tobacco Mosaic Virus as a Biotemplate. *Chem. Mater.* **2007**, *19* (10), 2389–2391.
(93) Wnęk, M.; Górzny, M. Ł.; Ward, M. B.; Wälti, C.; Davies, A. G.; Brydson, R.; Evans, S. D.; Stockley, P. G. Fabrication and Characterization of Gold Nano-Wires Templated on Virus-like Arrays of Tobacco Mosaic Virus Coat Proteins. *Nanotechnology* **2013**, *24* (2), 025605.
(94) Zhou, J. C.; Soto, C. M.; Chen, M.-S.; Bruckman, M. A.; Moore, M. H.; Barry, E.; Ratna, B. R.; Pehrsson, P. E.; Spies, B. R.; Confer, T. S. Biotemplating Rod-like Viruses for the Synthesis of Copper Nanorods and Nanowires. *J. Nanobiotechnology* **2012**, *10* (1), 1–12.
(95) Lee, S.-Y.; Royston, E.; Culver, J. N.; Harris, M. T. Improved Metal Cluster Deposition on a Genetically Engineered Tobacco Mosaic Virus Template. *Nanotechnology* **2005**, *16* (7), S435.
(96) Jockusch, H. Two Mutants of Tobacco Mosaic Virus Temperature-Sensitive in Two Different Functions. *Virology* **1968**, *35* (1), 94–101.
(97) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. Dual-Surface Modification of the Tobacco Mosaic Virus. *J. Am. Chem. Soc.* **2005**, *127* (11), 3718–3723.
(98) Ma, D.; Xie, Y.; Zhang, J.; Ouyang, D.; Yi, L.; Xi, Z. Self-Assembled Controllable Virus-like Nanorods as Templates for Construction of One-Dimensional Organic–inorganic Nanocomposites. *Chem Commun* **2014**.
(99) Bruckman, M. A.; Kaur, G.; Lee, L. A.; Xie, F.; Sepulveda, J.; Breitenkamp, R.; Zhang, X.; Joralemon, M.; Russell, T. P.; Emrick, T.; et al. Surface Modification of Tobacco Mosaic Virus with "Click" Chemistry. *ChemBioChem* **2008**, *9* (4), 519–523.

(100) Fan, X. Z.; Pomerantseva, E.; Gnerlich, M.; Brown, A.; Gerasopoulos, K.; McCarthy, M.; Culver, J.; Ghodssi, R. Tobacco Mosaic Virus: A Biological Building Block for Micro/nano/bio Systems. *J. Vac. Sci. Technol. Vac. Surf. Films* **2013**, *31* (5), 050815.
(101) Kobayashi, M.; Seki, M.; Tabata, H.; Watanabe, Y.; Yamashita, I. Fabrication of Aligned Magnetic Nanoparticles Using Tobamoviruses. *Nano Lett.* **2010**, *10* (3), 773–776.
(102) Wu, Z.; Mueller, A.; Degenhard, S.; Ruff, S. E.; Geiger, F.; Bittner, A. M.; Wege, C.; Krill III, C. E. Enhancing the Magnetoviscosity of Ferrofluids by the Addition of Biological Nanotubes. *ACS Nano* **2010**, *4* (8), 4531–4538.
(103) Górzny, M. Ł.; Walton, A. S.; Evans, S. D. Synthesis of High-Surface-Area Platinum Nanotubes Using a Viral Template. *Adv. Funct. Mater.* **2010**, *20* (8), 1295–1300.
(104) Yang, C.; Manocchi, A. K.; Lee, B.; Yi, H. Viral Templated Palladium Nanocatalysts for Dichromate Reduction. *Appl. Catal. B Environ.* **2010**, *93* (3–4), 282–291.
(105) Bruckman, M. A.; Hern, S.; Jiang, K.; Flask, C. A.; Yu, X.; Steinmetz, N. F. Tobacco Mosaic Virus Rods and Spheres as Supramolecular High-Relaxivity MRI Contrast Agents. *J. Mater. Chem. B* **2013**, *1* (10), 1482.
(106) Smith, M. L.; Lindbo, J. A.; Dillard-Telm, S.; Brosio, P. M.; Lasnik, A. B.; McCormick, A. A.; Nguyen, L. V.; Palmer, K. E. Modified Tobacco Mosaic Virus Particles as Scaffolds for Display of Protein Antigens for Vaccine Applications. *Virology* **2006**, *348* (2), 475–488.
(107) Han, W.; Lin, Z. Learning from "Coffee Rings": Ordered Structures Enabled by Controlled Evaporative Self-Assembly. *Angew. Chem. Int. Ed.* **2012**, *51* (7), 1534–1546.
(108) Lin, Y.; Balizan, E.; Lee, L. A.; Niu, Z.; Wang, Q. Self-Assembly of Rodlike Bio-Nanoparticles in Capillary Tubes. *Angew. Chem.* **2010**, *122* (5), 880–884.
(109) Muller, S.; Wolf, J.; Ivanov, S. A. Current Strategies for the Synthesis of RNA. *Curr. Org. Synth.* **2004**, *1* (3), 293–307.
(110) Arunachalam, T. S.; Wichert, C.; Appel, B.; Müller, S. Mixed Oligonucleotides for Random Mutagenesis: Best Way of Making Them. *Org. Biomol. Chem.* **2012**, *10* (24), 4641.
(111) Beaucage, S. L.; Reese, C. B. Recent Advances in the Chemical Synthesis of RNA. In *Current Protocols in Nucleic Acid Chemistry*; Beaucage, S. L., Bergstrom, D. E., Herdewijn, P., Matsuda, A., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009.
(112) El-Sagheer, A. H.; Brown, T. New Strategy for the Synthesis of Chemically Modified RNA Constructs Exemplified by Hairpin and Hammerhead Ribozymes. *Proc. Natl. Acad. Sci.* **2010**, *107* (35), 15329–15334.
(113) Finch, J. T.; Klug, A. Two Double Helical Forms of Polyriboadenylic Acid and the pH-Dependent Transition between Them. *J. Mol. Biol.* **1969**, *46* (3), 597–IN27.
(114) Petrovic, A. G.; Polavarapu, P. L. Structural Transitions in Polyriboadenylic Acid Induced by the Changes in pH and Temperature: Vibrational Circular Dichroism Study in Solution and Film States. *J. Phys. Chem. B* **2005**, *109* (49), 23698–23705.
(115) Brahms, J.; Michelson, A. M.; Van Holde, K. E. Adenylate Oligomers in Single- and Double-Strand Conformation. *J. Mol. Biol.* **1966**, *15* (2), 467–488.
(116) Seol, Y.; Skinner, G. M.; Visscher, K.; Buhot, A.; Halperin, A. Stretching of Homopolymeric RNA Reveals Single-Stranded Helices and Base-Stacking. *Phys. Rev. Lett.* **2007**, *98* (15), 158103.
(117) Dewey, T. G.; Turner, D. H. Laser Temperature-Jump Study of Stacking in Adenylic Acid Polymers. *Biochemistry (Mosc.)* **1979**, *18* (26), 5757–5762.
(118) Saenger, W.; Riecke, J.; Suck, D. A Structural Model for the Polyadenylic Acid Single Helix. *J. Mol. Biol.* **1975**, *93* (4), 529–534.
(119) Suck, D.; Manor, P. C.; Saenger, W. The Structure of a Trinucleoside Diphosphate: Adenylyl-(3', 5')-Adenylyl-(3', 5')-Adenosine Hexahydrate. *Acta Crystallogr. B* **1976**, *32* (6), 1727–1737.

(120) Causley, G. C.; Johnson, W. C. Polynucleotide Conformation from Flow Dichroism Studies. *Biopolymers* **1982**, *21* (9), 1763–1780.
(121) Lin, J.; Kolomeisky, A.; Meller, A. Helix-Coil Kinetics of Individual Polyadenylic Acid Molecules in a Protein Channel. *Phys. Rev. Lett.* **2010**, *104* (15), 158101.
(122) Leng, M.; Felsenfeld, G. A Study of Polyadenylic Acid at Neutral pH. *J. Mol. Biol.* **1966**, *15* (2), 455–466.
(123) Causley, G. C.; Staskus, P. W.; Johnson, W. C. Improved Methods of Analysis for CD Data Applied to Single-Strand Stacking. *Biopolymers* **1983**, *22* (3), 945–967.
(124) Pörschke, D. The Dynamics of Nucleic-Acid Single-Strand Conformation Changes. *Eur. J. Biochem.* **1973**, *39* (1), 117–126.
(125) Richards, E. G. On the Analysis of Melting Curves of Stacked Polynucleotides. *Eur. J. Biochem.* **1968**, *6* (1), 88–92.
(126) Aida, M.; Nagata, C. Ab Initio Mo Study on Base Stacking: Adenine-Adenine Interaction in Single-Stranded Polyadenylic Acid (Poly A). *Chem. Phys. Lett.* **1982**, *86* (1), 44–46.
(127) Breslauer, K. J.; Sturtevant, J. M. A Calorimetric Investigation of Single Stranded Base Stacking in the Ribo-Oligonucleotide A7. *Biophys. Chem.* **1977**, *7* (3), 205–209.
(128) Filimonov, V. V.; Privalov, P. L. Thermodynamics of Base Interaction in (A)n and (A·U)n. *J. Mol. Biol.* **1978**, *122* (4), 465–470.
(129) Lanina, T.; Ledneva, R. K.; Bogdanov, A. A. A Circular Dichroism Study of the Interaction of Adenylic Acid Oligo- and Polynucleotides with TMV Protein. *FEBS Lett.* **1974**, *39* (2), 235–238.
(130) Ledneva, R. K.; Razjivin, A. P.; Kost, A. A.; Bogdanov, A. A. Interaction of Tobacco Mosaic Virus Protein with Synthetic Polynucleotides Containing a Fluorescent Label: Optical Properties of Poly (A, eA) and Poly (C,eC) Copolymers and Energy Migration from the Tryptophan to 1,N6-Ethenoadenine or 3,N4-Ethenocytosine Residues in RNP. *Nucleic Acids Res.* **1978**, *5* (11), 4255–4243.
(131) Geiger, F. C.; Eber, F. J.; Eiben, S.; Mueller, A.; Jeske, H.; Spatz, J. P.; Wege, C. TMV Nanorods with Programmed Longitudinal Domains of Differently Addressable Coat Proteins. *Nanoscale* **2013**, *5* (9), 3808.
(132) Alonso, J. M.; Górzny, M. Ł.; Bittner, A. M. The Physics of Tobacco Mosaic Virus and Virus-Based Devices in Biotechnology. *Trends Biotechnol.* **2013**, *31* (9), 530–538.

Chapter 2, Poly A Mediated Assembly of TMV-like Rods

2.1 Overview of techniques used to monitor assembly of TMV-like rods

Although poly A mediated TMV-like rod assembly has been demonstrated before, it is quite uncommon, and has never been proposed for use in nanotechnology applications. In previous literature, UV difference spectroscopy was used to monitor the assembly of RNA mediated TMV rod formation[1]. Resistance of the RNA to RNAse in conjunction with the expected UV difference curves and expected density of RNA bound TMV-like rods obtained from CsCl gradients was sufficient to conclude that the rods had formed. Direct visualization techniques for poly A mediated rod were not done, perhaps because most of this fundamental research was done between 1960 and 1980 and while TEM was available at that time, it perhaps would have been troublesome for the groups involved in TMV research since they had no experience with TEM. However it was seen to be quite useful to image the rods as not only did it give direct evidence of the assembled TMV-like rod but it also gave some additional insight into the mechanism of assembly.

The reaction between poly A and TMV-CP results in important conformational changes in both reactants. Although the conformational changes are more drastic for the poly A then for the TMV-CP, the alteration in configuration results in optical differences compared to the starting state in both the TMV-CP and the poly A (figure 2.1). These optical differences can not only be used to provide additional evidence of the TMV-like rod self-assembly but also allow for real time monitoring of the assembly reaction which are crucial for obtaining kinetic data of the reaction progress. The ability to monitor the assembly reaction using a method other than a direct visual method was particularly important given that it usually takes many hours, sometimes days between the preparation of the sample and the visual confirmation of the rod assembly. Initially, similar to what was done previously, UV difference spectroscopy was used to follow the poly A mediated TMV-like rod assembly in real time. It was found that due to the

setup of the experiment, which differed from that done in the literature, there was an inherent problem with obtaining high quality, noise free spectra.

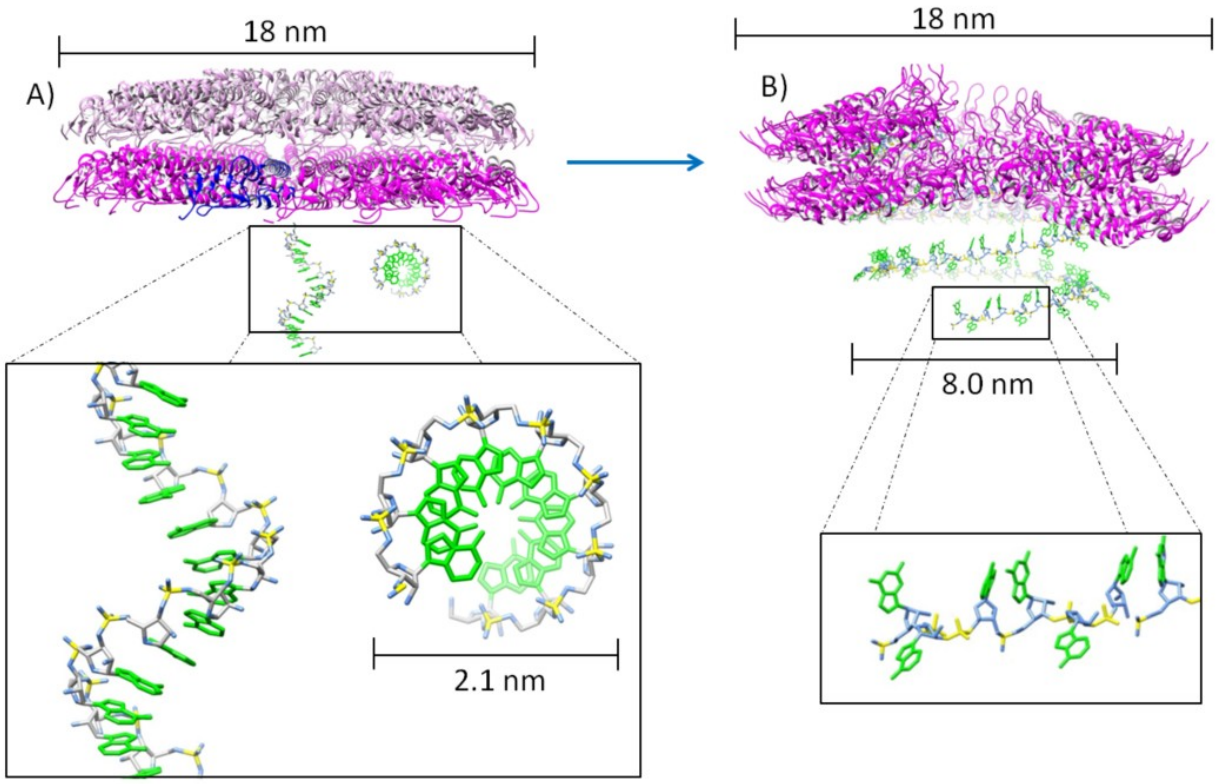


Figure 2.1, representation of the change in conformation of poly A and TMV-CP upon incorporation into a poly A containing TMV-like rod. A) Before reaction poly A exists as a single stranded helix (green) while at pH 6.8 TMV-CP (magenta) is primarily in the disk phase. B) Poly A containing TMV-like rod. The TMV-CP is removed from the bottom half of the rod for clearer viewing of the poly A conformation. Upon binding the poly A the TMV-CP switches from the planar configuration seen in the disk to the helical configuration seen in the rod. The poly A conformation is altered from a single helix with a diameter of 2.1 nm in which the adenines are all pointed to the central axis into a configuration with a diameter of 8.0 nm, with the bases (green) no longer pointed at the central axis. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco

2.2 UV/visible spectroscopy

In previous literature, they used a specialized dual cell cuvette which allowed for the protein signal to be blanked effectively[1]. Unfortunately, such a cuvette could not be obtained for this work, and as such, only poly A and not protein could be used for the blank. This was later found to the main limitation in obtaining high quality data. Even so, at a protein

concentration of ~ 1 mg/mL, although the signal was not large due to the low concentration of protein, the experimental noise from the protein was low enough that it was possible to record a similar signal as the one seen in the literature. Importantly, the observed signals were not observed when poly C was substituted for poly A, which we expected as poly C has not been shown to nucleate TMV-like rods.

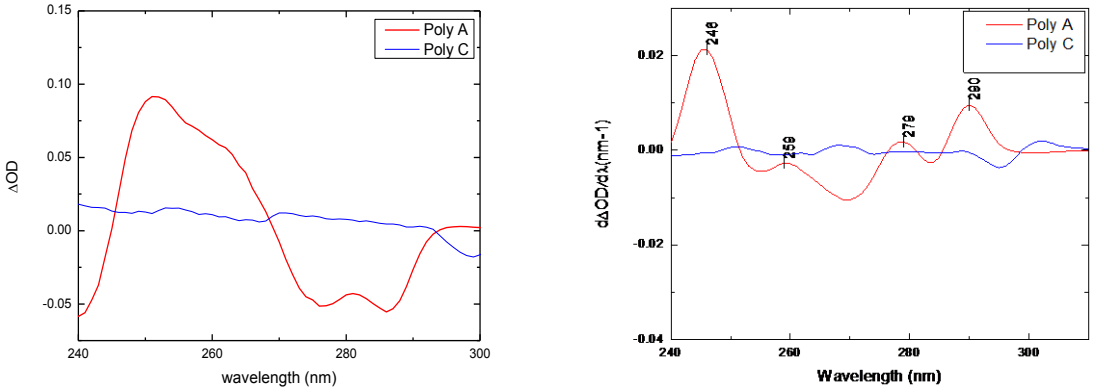


Figure 2.2, UV difference spectra (left) and derivative spectra (right) of before and after reaction of poly A and Poly C with TMV-CP. The poly C is not expected to interact with TMV-CP, and the lack of signal from the poly C is highly indicative that no reaction took place. On the other hand the poly A and TMV-CP did interact as can be seen from the signals in both the difference and the derivative UV spectra, which are comparable to what was seen previously in the literature[1].

In the UV difference spectra for poly A (red) and poly C (blue), the poly A signal observed was a bit different than any of the curves seen in the literature. However, in the literature the poly A which was used was purified. Therefore only a specific length of poly A with a reasonably narrow length distribution was used in an individual experiment. The signals were found to depend upon the length of poly A used[1]. In this work, a large distribution of poly A chain lengths were used, since the poly A was used as received, without further length purification. Although the curve was not an exact replica of any of the individual sizes, the observed spectra were more or less consistent with some weighted average of the curves from various sizes of poly A. Most importantly, the control reaction with poly C used in place of the poly A had no similar response, which indicates that there was an interaction which occurs only between the poly A and the TMV and not between the poly C and the TMV.

In general, the UV derivative spectrum is more informative than the difference spectrum because the derivative spectrum highlights slope changes, and can therefore be used to distinguish small changes and overlapping peaks. In this case, there are specific absorbance changes that should arise after TMV rods have been formed, which will appear as peaks in the derivative spectrum. Thus, in theory, the derivative spectrum can be used to identify whether a disk to helix TMV-CP transition has occurred. Furthermore, the derivative spectrum can also be used not only to determine whether or not RNA has been used for the assembly, but also to distinguish between poly A and native TMV-RNA as the TMV rod assembly initiator[1]. The critical and distinguishing absorbance changes arise from the observation that at RNA assembly pH values, the majority of the TMV used to assemble rods starts off in the planar disk phase. Since the formed rod is helical in nature, there is a change in helicity from planar disk to helical rod upon assembly. This subtle shift in the protein structure also has a small but measureable effect on the absorbance profile of the TMV. Although any kind of TMV-based assembly would likely lead to a positive delta OD signal, the derivative of that spectrum results in peaks which are unique to the method of TMV self assembly used.

The planar to helix transition results in a peak at 289 nm coupled with a minimum at 284 nm in the derivative spectrum. The appearance of these two features is an indication that TMV rods have been formed from TMV disks, with the difference in magnitude between the two features being related to the amount of rod produced. This signal comes from the TMV-CP, and as such, does not depend on the type of TMV rod assembly. Thus, even pH driven TMV rod formation would result in such a signal in the UV derivative spectrum. There is also a peak at around 250 nm which is present whenever RNA is used to nucleate the TMV rods, which can thus be used to differentiate between RNA mediated TMV rod assembly and pH driven TMV rod assembly. Finally and most importantly, at 278 nm there is a peak which is unique to poly A which is due to the de-stacking of the poly A bases which are stacked in helical configuration under neutral solution conditions, as described in Chapter 1[2,3].

Due to the setup of the experiment and the low concentration of protein used, the signals obtained in the UV derivative spectrum were not strong, but they did generally

resemble what was seen in the literature. The peak at 289 nm and minimum at 284 nm were present, as was the peak at around ~250 nm. The poly A de-stacking peak, while present at 278 nm, was small, but the small size was likely due to a combination of noise from the lack of a protein blank and the low concentration of reactants that were used. The fact that there was a TMV-CP planar to helix transition occurring at a pH where such a transition should not occur without the aid of a nucleating RNA strand was suggestive that TMV-like rods were formed, but that the experimental setup interfered somewhat with the 278 nm signal. The solution was also imaged visually, and contained structures tentatively assigned as TMV-like rods as will be discussed in later sections.

As can be seen in figure 2.2, there are clearly features which are present in the reaction between poly A and TMV-CP which are not occurring when poly C interacts with TMV-CP. Poly C is similar to poly A in that it also has some helical portions[4]. It therefore seems unlikely that the helical conformation in poly A is responsible for nucleating the TMV-like rod assembly reaction. Furthermore, for poly A under the conditions used, the helix is the more stable state[5,6]. The stability of the poly A helix likely interferes with TMV-like rod formation, as the poly A helix must be disassembled prior to incorporation into a TMV-like rod[1]. On the other hand, poly U has no helical nature but was found to be easier than poly C to incorporate into complexes with TMV-CP[4,7]. Poly C containing mixed polynucleotides were not found to interact favorably with TMV-CP while poly U containing mixed polynucleotides had significantly more success at interacting with TMV-CP[7]. It therefore seems likely that it is the coiled portion of poly A which is initiating the interaction with TMV-CP and not the helical region. It also seems likely that the main driving force for interaction of polynucleotides with TMV-CP is the strength of the interaction of the individual nucleotide with the CP and not the conformation of that polynucleotide as in general, C nucleotides do interact favorably with TMV-CP while U nucleotides interact more favorably then poly C but not as favorably as A or G nucleotides[8–10].

2.2.1 Critical problems with UV absorption techniques

In order to obtain a stronger signal, the reaction was tried again several times with higher concentration of reactants. However, as can be seen in figure 2.3, increasing the concentration of the poly A and TMV-CP only decreased the quality of the observed signals. As was mentioned previously, this was likely due to the fact that unlike in the previous study, TMV-CP could not be included in the blank as it would also form rods upon coming into contact with the poly A used in the blank. Like most proteins, TMV has a strong absorbance around 280 nm (Figure 2.4), which in particular seems to affect the poly A destacking peak at 278 nm. As the concentration was further increased, the TMV-CP absorption profile seemed to interfere with the entire spectrum.

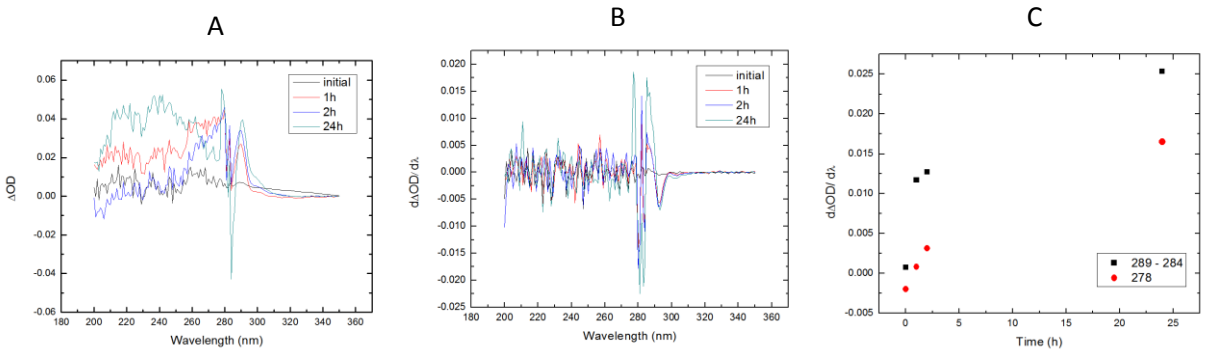


Figure 2.3, UV Difference (A) and Derivative (B) spectra time evolution at a TMV-CP concentration of ~ 3mg/mL. UV Difference spectra were corrected for light scattering. C) Evolution of the poly A base destacking signal (278 nm) and the TMV-CP helix formation signal (289 - 284 nm) of the UV derivative curves. Both signals are consistent with the formation of TMV-like rods, although the magnitudes of the signals were lower than expected.

At higher concentration, as can be seen in figure 2.3, the UV difference and derivative spectra were less clear than in earlier attempts at lower concentration. Instead of the desired signals becoming larger and thus giving cleaner spectra, the spectra were dominated by noise, presumably from the larger, unblanked TMV-CP peak.

Although the higher concentration spectra looked messy, there was some progression in the spectra over time. This fact became clearer in figure 2.3c, where it can be seen that both the 278 nm poly A destacking signal and the 289 nm - 284 nm, TMV-CP helix formation signal were seen to increase over time in a similar fashion. The magnitude of both signals were less than expected and were not significantly greater than the background signal. If it had been

possible to include protein in the blank the background signal, especially around the main 280 nm protein absorption peak, would have been considerably better which would have provided a more favorable limit of detection and likely allowed even the relatively small poly A/TMV UV difference and derivative signals to be well resolved. Since such a blank was not possible any conclusions formed from these spectra must be viewed with caution. However, the fact that both signals were seen to increase in an exponential pattern, which is consistent with TMV assembly kinetics, would suggest that TMV rods were formed and it was merely a flawed experimental setup which prevented verification of this fact. This was found to be the case once the AFM and TEM samples of the solution were visualized.

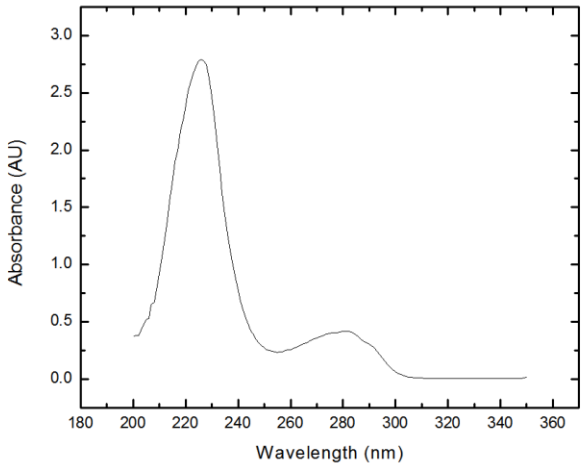


Figure 2.4, UV absorbance of 10x diluted TMV-CP. As can be seen there was a peak located at 280 nm which, because the CP could not easily be included in the blank, would likely distort the 278 nm poly A destacking peak even at moderate concentrations, and would likely distort all the relevant peaks at higher CP concentrations.

2.3 Circular dichroism

Since the UV spectra were not fully satisfactory as a method to monitor poly A template TMV assembly, an alternate optical technique was used to follow the poly A - TMV-CP interaction in real time. Circular dichroism (CD) was selected, as it has a variety of advantages over UV difference spectroscopy. Similar to UV difference spectroscopy, CD can also be used to distinguish between RNA mediated and pH driven TMV rod assembly. Furthermore, it is also possible to distinguish between poly A mediated TMV-like rod assembly and TMV RNA mediated rod assembly, as poly A and TMV RNA have significantly different CD spectra[11–13]. All

this can be accomplished while providing substantially more sensitivity in following the reaction progress compared with UV difference spectroscopy. Importantly, because of the method of data collection in CD, a mixable dual cell cuvette was not required in order to achieve this sensitivity. CD measures the difference between the samples ability to absorb left and right circularly polarized light. Since both will be equally absorbed by the TMV-CP, no blank is needed.

The region between 350 nm and 230 nm was chosen for CD investigations, simply due to the relative ease with which such spectra could be obtained. Both TMV-CP and poly A have a CD signal in the region between 240 nm and 300 nm, however the poly A signal was more useful for following the reaction progress as the TMV-CP signal in that region was weak. This is because proteins tend to have much stronger CD signals at wavelengths between 180 nm and 210 nm, while nucleic acids have stronger signals in the aromatic region (240 nm - 300 nm). Although it should also be possible to follow reaction progress by monitoring the stronger TMV-CP CD signal at shorter wavelengths, such monitoring would involve significant alterations to the experimental setup. Not only would the light source have to be changed, but also most buffers absorb strongly below 210 nm, and make CD data acquisition in this region additionally challenging. Since it was quickly discovered that the reaction could be followed well through the poly A helix CD signal, it became unnecessary to further investigate shorter wavelength CD measurements.

The limitation in using the poly A CD signal to monitor the reaction progress between poly A and TMV-CP is that poly A is only measureable in the CD spectrophotometer in the lower concentration regime as it is a low concentration technique. A further drawback is that although the disk and rod state could be differentiated using CD, detailed information about the state of the TMV was not possible in the wavelength range used. The limitations of the CD measurement techniques used did not prevent observation of a significant amount of TMV-like rods being formed.

A) B)

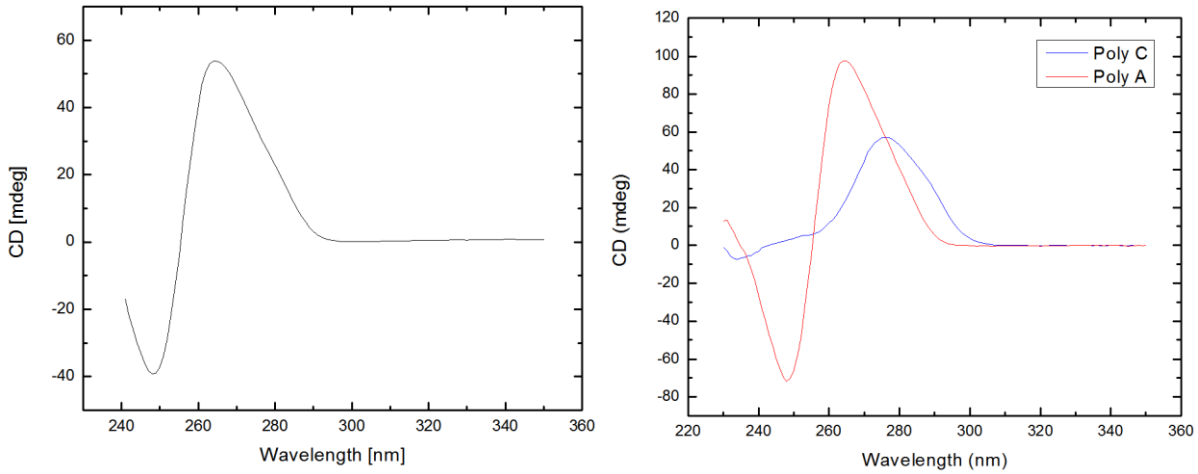


Figure 2.5, A) Characteristic CD signal of poly A with a maximum at 264 nm and a minimum at 248 nm, with the CD signal crossing the X axis at 255 nm. Only the helical portion of the poly A contributes to the CD signal. As such the signal is larger at lower temperatures, where helical content in increased, and lower at higher temperatures, where the percentage of helix is lower. B) Comparison of the CD signal of poly A and poly C

As can be seen in figure 2.5 poly A has a strong and characteristic CD signal. This was expected since at pH 6.8 a large percentage of poly A is in the single helix conformation. There is a strong positive peak at 264 nm as well as a negative peak at 248 nm. Only the helical regions of the poly A contribute to this signal hence any agent which denatures the poly A helix such as increased temperature will also attenuate its CD signal. Likewise any action which increases the percentage of helix, such as the introduction of a significant amount of magnesium ions[14], will increase the poly A CD signal. As figure 2.5b shows, at the same concentration, poly A has a significantly more intense and unique CD profile than poly C. This is due to the unusually strong propensity for poly A to form helical structures. While poly C also has some helical nature, it not only forms helices less readily than poly A, but poly C helices are also less structured than those for poly A. As such, the CD of poly C is somewhat similar to that from random sequence RNA, which often also have some regions of single helix[4,11,13,15,16].

2.3.1 CD spectra of TMV-CP

Although the TMV CD signal in the region studied is not nearly as strong as that from poly A, TMV does have a characteristic CD spectrum which can be seen in figure 2.6. The CD of

TMV-CP contains a multitude of connected peaks between 310 nm and 350 nm. In the experimental setup used, those peaks were not well resolved. This was due in part because those peaks are not strong and as such an extremely high concentration of TMV-CP is needed in order to resolve them[17]. The CD equipment used was also optimized for measuring nucleic acids and as such was not ideal for the measurement of weak protein signals. The beginning of an intense negative peak was also seen, which starts around 245 nm. However, since that peak is on the edge of the detection capacity of the experimental setup, the rest of the 245 nm signal is never seen in its entirety and the magnitude was sometimes unreliable.

Previous studies of TMV-CP have found that the peaks located between 310 nm and 350 nm increase in intensity as TMV-CP switches from planar to helical configurations[13,17,18]. Similarly in figure 2.6, although peak resolution is not ideal , there is a clear increase in the magnitude of the signals from the various peaks of TMV upon conversion of disk phase TMV (pH 6.8) to rod phase TMV (pH 5.5). It would therefore be expected that poly A mediated TMV-rod formation would produce a similar increase in signal between 350 nm and 310 nm due only to the change in helicity of the TMV-CP upon binding to poly A.

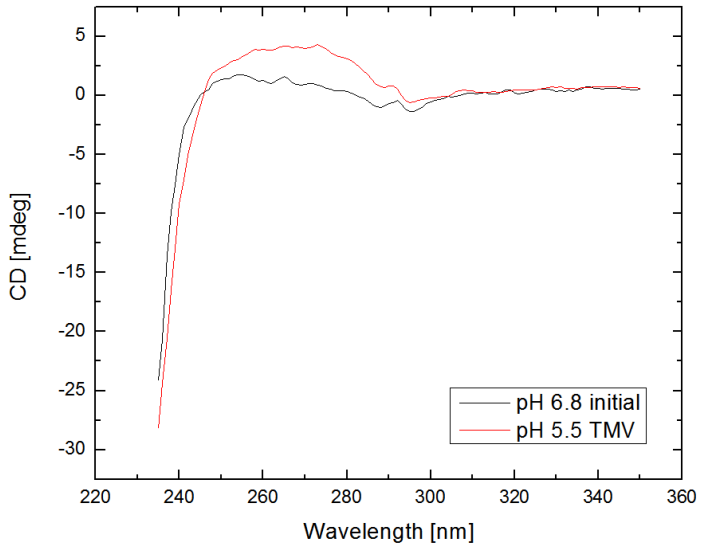


Figure 2.6, CD of TMV-CP at different pHs at the same concentration of TMV-CP of 5 mg/mL. pH 6.8 is in the disk phase of TMV-CP while at pH 5.5 TMV-CP exists mainly as helical rods. Upon conversion of disk to helix no new peaks were seen in the CD, however there was a general increase seen in the intensity of the peaks between 310 nm and 350 nm.

2.3.2 Real time monitoring of poly A mediated TMV-like rod formation

Poly A is added to a TMV disk solution at pH 6.8 and the reaction progress is monitored over time. Initially, the curve is very similar to that for pure poly A , with the TMV-CP having very little impact on the overall CD spectrum even though it is in excess. This is due to the significantly larger magnitude in signal from the poly A. However, as the reaction progresses, it can be seen in figure 2.7 that the poly A CD signal decreased dramatically in a similar way to what would happen to the CD signal if a poly A solution was chemically denatured. Therefore it can be concluded that during the interaction between the poly A and the TMV-CP, the poly A helix is being gradually disassembled. This would be expected if the poly A was being incorporated into growing TMV-like rods, as the poly A would be forced to adopt a configuration compatible with the RNA binding track of the TMV-CP. The configuration which TMV RNA adopts when bound to TMV-CP is considerably different to the helix which the poly A forms under the solution conditions used. The natural single stranded helix of poly A has a radius of 1.07 nm and a pitch of 2.54 nm[2,3], while RNA bound to TMV-CP is forced to adopt its period of 2.3 nm, and the RNA binding region forces any bound RNA to adopt a radius of 4 nm[19].

Particles deposited from the solution used in figure 2.7 were later imaged using TEM and AFM, and TMV-like rods were seen. Therefore it can be concluded from figure 2.7 that interaction of poly A with TMV-CP causes the helical regions of the poly A to be denatured and also causes TMV-like rods to form. Under the solution conditions used for measurement, only RNA mediated rod formation is possible. Thus, it would be logical to assume that the poly A must have been incorporated into the TMV-like rods. Since the poly A signal was gradually lost, it would suggest that the helical configuration of poly A is completely incompatible with the configuration which RNA bound to TMV-CP is forced to adopt. The CD spectrum obtained at 2h after the start of the reaction resembles more that of TMV-CP in the rod phase than that from poly A.

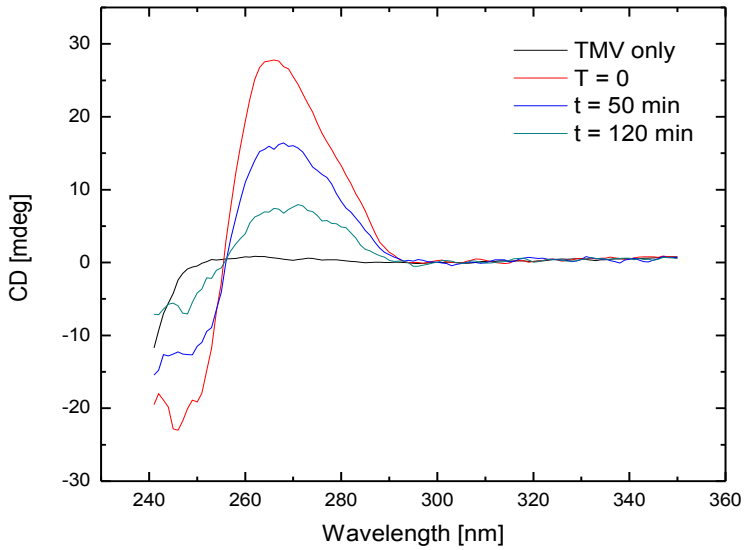


Figure 2.7, CD spectra of TMV-CP and poly A at various reaction times. Final concentration of TMV-CP was 5 mg/mL while the poly A was at .12 mg/mL. The CD signal from the poly A completely dominates that from the TMV-CP as such the initial spectrum was very similar to that of poly A. Over time the poly A signal was lost which would suggest that it was incorporated into TMV-like rods as nothing else present in solution had the capacity to denature the poly A helix.

Although the configuration of TMV RNA when bound to TMV-CP has been well established, there have been no studies to see if poly A adopts the same or merely a similar RNA conformation upon binding to TMV-CP. Figure 2.7 demonstrates that the CD signal from poly A after incorporation into the TMV is at least drastically smaller in intensity when compared to that produced by the poly A helical regions in single stranded poly A. Since the helix signal was essentially eradicated, it seems unlikely that the conformation of poly A inside of TMV-like rods has any similarity to that normally found in single stranded helical regions of poly A. Therefore it is more likely that poly A adopts a conformation similar to that which the native TMV RNA adopts.

2.3.3 Conformation of poly A upon incorporation into TMV-like rods

CD is often used to monitor changes in configuration of RNA and DNA because even small changes in configuration will have measureable differences in the CD spectrum.

Therefore, comparing the CD spectrum before and after reaction between TMV CP and poly A, it is possible to obtain some insight into the final configuration of the poly A. After 2 hours, the original poly A signal is lost as can be seen in figure 2.7. However, the poly A should still have some contribution to the overall CD spectrum, as in theory if it was incorporated into a TMV-like rod it would still be in a well defined, chiral configuration.

The final CD spectrum from figure 2.8 contains no trace of natural single stranded poly A helix. The poly A had not been destroyed, thus it must have been made to adopt a different configuration, as even a small amount of single stranded poly A would have some helical content and thus a measureable characteristic CD signal. In fact, the final CD spectrum of the poly A TMV-CP mixture was closer to resembling the curve seen in pH induced TMV-CP rods seen in figure 2.6 than it was to resembling the initial poly A TMV-CP mixture.

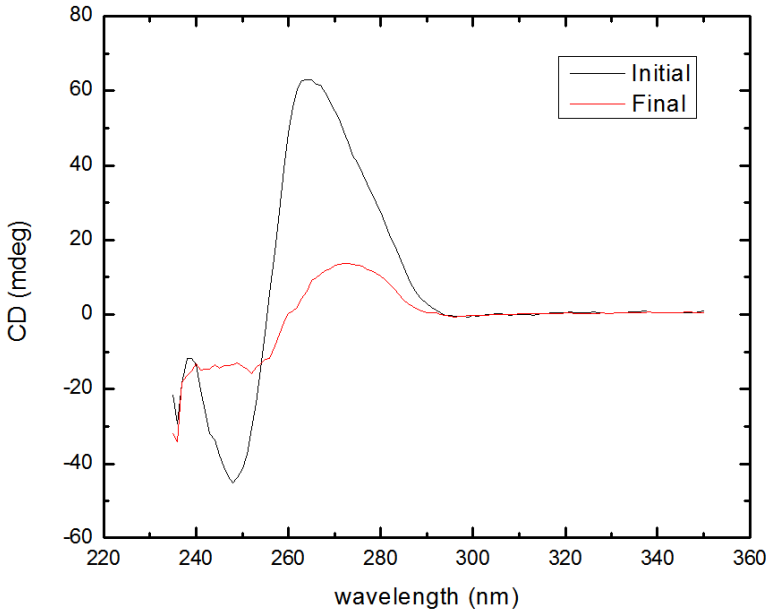


Figure 2.8, Comparison of CD spectrum of initial poly A - TMV-CP mixture (black) to the final CD spectrum of the poly A - TMV-CP mixture (red). Final poly A concentration was .25 mg/mL while final TMV-CP concentration was 5 mg/mL, both in pH 6.8 phosphate buffer. The initial CD curve resembled that of poly A helix, with a maximum at 264 nm and a minimum at 248 nm. The final curve was very different, with a maximum at 273 nm which was broader and significantly less intense. The final curve was more similar to the CD curve obtained from pH induced TMV rods then it was to poly A.

In order to get a clearer picture of the makeup of the final CD spectrum of the poly A TMV-CP mixture, it was compared to the TMV-CP spectrum in figure 2.9. As mentioned

previously, TMV rods and disks have slightly different peak intensities in their CD spectra, thus the final poly A TMV-CP mixture must be compared to pH induced TMV-CP rods and not the TMV-CP disk solution which was used to assemble the TMV-like rods.

As can be seen in figure 2.9, the pH induced TMV-CP rod spectra overlaps almost perfectly with that of the final poly A TMV-CP mixture. Although the majority of the TMV-CP peaks are swamped out by a non-TMV signal, the TMV-CP peak at 290 nm is not and could clearly be seen in figure 2.9 to be of the same magnitude as that from the pH induced TMV-CP rods. From figure 2.6 it can be seen that the 290 nm TMV-CP peak in the rod phase is significantly larger in magnitude than that from the same concentration of TMV-CP in the disk phase. Thus, while the 290 nm peak which is seen in the CD spectra is not sufficient on its own to conclusively state that TMV-like rods are produced from the interaction of poly A and TMV-CP, it would suggest that TMV-like rods were indeed formed.

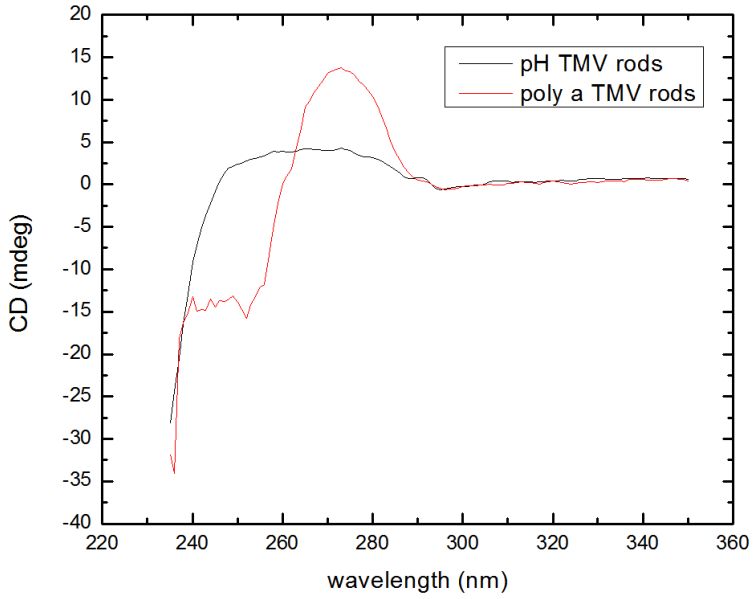


Figure 2.9, Comparison of final poly A TMV-CP mixture (red) to the same concentration (both at 5 mg/mL) of pH induced TMV-CP rods at pH 5.5 (black). The two spectra match up everywhere except in the region where an RNA signal would be expected. There is a peak at 273 nm and a minimum at 252 nm. Therefore it seems that there is an additional signal which cannot be attributed to the TMV-CP being in helical form rather than planar. The additional signal must be due to a poly A in a conformation very different from its original.

Although the final CD spectrum of the poly A mixture suggests that TMV-like rods are formed, it is also clear that some of the features in the spectrum of the final mixture cannot be assigned to signals from TMV CP in a rod-like configuration. Since the only other species which present in solution is poly A, it seems likely that the other contributor to the CD spectrum is the poly A. However, these features are also inconsistent with helical assemblies formed from poly A alone. Therefore it seems likely that the additional signals are due to poly A in some new chiral configuration. In order to simplify things, the TMV-CP signals can be subtracted from the initial and final TMV-CP poly A mixture, resulting in the spectra seen in figure 2.10.

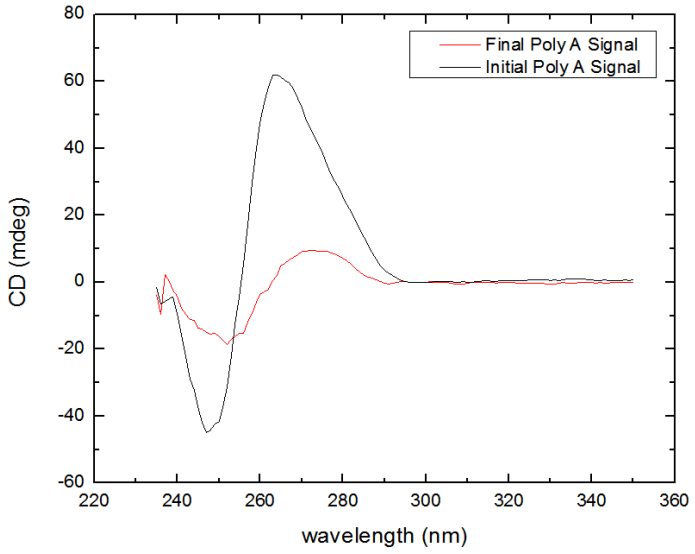


Figure 2.10, CD spectra of only the contribution of the poly A on the inital (black) and final (red) mixture of TMV-CP and poly A. The initial poly A signal is found by subtracting the CD signal of the TMV-CP solution used during the reaction from the inital poly A - TMV-CP CD spectrum. The final poly A spectrum is found by subtracting the CD spectrum obtained from pH induced TMV-CP rods from the final TMV-CP poly A mixture. The exact same solution was used for both curves, thus the concentration of poly A was identical in both cases. The CD contribution from the inital poly A is exactly as expected, with a peak at 264 nm and a minimum at 248 nm. The final poly A configuration produces a CD spectrum with a peak at 273 nm and a minimum at 252 nm. Not only is the shape of the curve altered in the final poly A CD spectrum, but also the intensity and identity of the peaks are substantially different.

As can be seen in figure 2.10, in which the same concentration of poly A in its initial configuration is compared to the final configuration after reaction with TMV-CP, the two CD spectra are quite different. Not only is the intensity of the final signal of the poly a substantially

lower, but the entire spectrum is no longer recognizable as the characteristic poly A CD signal. This kind of significant change in CD signal is indicative that the poly A has undergone a change in conformation. The large reduction of the intensity in general of the final poly A spectrum indicates that the new poly A configuration is not nearly as CD active as that of the initial poly A helix. The new configuration of poly A produces a peak at 273 in the CD spectrum as well as a negative peak at 252 nm. In the natural poly A helix, the positive peak is found to be significantly larger than the negative peak. However, in the new configuration of poly A the 252 nm negative peak is found to have a larger magnitude than the 273 nm peak. Poly A is able to bond with itself at pH below 5.5 to form a double stranded helix with itself. This form of poly A has a different CD spectra[5,20–22] than neutral, single stranded poly A. However, the CD spectra from double stranded poly A is more similar to the neutral poly A CD spectrum than that of the poly A contribution in the final poly A TMV-CP mixture. Although further confirmation is necessary, such a drastic change in the CD spectrum of the poly A surely indicates that the final poly A conformation adopted in the TMV-CP poly A mixture is one which cannot be obtained by any known form of poly A or by any simple ionic binding to poly A. Therefore, is it likely that that the poly A is incorporated into the RNA binding track of the TMV-CP, and formed poly A containing TMV-like rods.

If the poly A had been incorporated into a TMV-like rod, it would explain the observed CD spectrum. Poly A has a strong tendency to form helices. Cation binding can enhance the poly A CD signal by increasing the percentage of helix present in solution, however the only thing which reduces the magnitude of a poly A signal is denaturation of that poly A helix. The new configuration of poly A not only has a reduced intensity, but also the main peak is shifted from 264 nm to 273 nm, and the curve has a different shape than that of the original poly A spectrum, neither of which occur during poly A denaturation studies[5]. However, TMV-CP very much restricts the types of conformations which can be adopted by bound RNA[19]. Therefore, although some small configurational differences due to different RNA sequences may occur, they are unlikely to be significant changes as significant changes simply would not be physically able to be incorporated into TMV-like rods.

2.3.4 Comparison of TMV bound poly A CD signal to that of native TMV RNA

There have been CD studies of the various components of TMV in the spectral range utilized in this study. Not surprisingly, the TMV RNA is found to change conformation upon incorporation into a TMV rod. The interesting result from that study is that the signal from the TMV RNA which was incorporated into TMV rods was surprisingly similar to the one from the final poly A configuration of figure 2.10[13]. Although the helical TMV-CP signal is not taken into account in that study, since the CP was by far in excess, most of the change in signal must have been due to the RNA as only a small fraction of the TMV-CP would be in helical form. In addition, a negative peak was seen in that study around 253 nm which simply cannot be attributed to TMV-CP as it does not have a negative component at that wavelength in any phase. There is also a positive peak due to the TMV RNA somewhere between 270 nm and 280 nm. It is not clear where the maximum is as the curve includes some of the signal from the TMV-CP, however the shape of the curve is inconsistent with that which would be expected if only the TMV-CP was contributing[17]. Thus, some of the signal must be due to the contained TMV RNA.

Such a similar CD spectrum for poly A incorporated into TMV-like rods compared to that from TMV RNA incorporated into TMV rods suggests that in the process of forming the TMV-like rods, the poly A was bound to the RNA binding region of the TMV-CP. Furthermore, the configuration of RNA is extremely important in determining the CD spectra which that RNA will have. Although the identity of the bases also contributes, the conformation of the RNA has a very large impact on the final CD spectrum[15,16]. Therefore, the similarity of poly A spectrum to that of the TMV RNA would suggest that the conformation of any RNA which is incorporated into TMV or TMV-like rods would likely be very similar.

Although the CD data is consistent with the production of TMV-like rods during the reaction between poly A and disk phase TMV-CP, this is not sufficient evidence to be able to conclude with absolute certainty that poly A has indeed formed TMV-like rods. There are certainly other methods for getting poly A to go through helix denaturation. Although it is

difficult to come up with processes which would not only denature the poly A but would also produce an alternate RNA configuration, it cannot not be completely ruled out. Thus, an alternate method of confirmation of the production of TMV-like rods containing poly A is necessary.

2.4 Direct visualization of poly A mediated TMV-like rods

The production of TMV-like rods is confirmed by direct visualization using TEM and AFM. TEM is done with negative staining, in which a low electron density species, such as a protein, which therefore gives poor contrast in TEM, is stained with solution containing electron rich atoms. Thus, TMV being lower in electron density than the heavy stain shows up white or gray, while the surrounding areas, being rich in the heavy electron element shows up dark or black. In this study either uranyl acetate or phosphotungstic acid is used as the staining solution. As such, when TMV was negatively stained in TEM, the TMV rod and disk can be identified by its 18 nm diameter and also by the presence of the 4 nm inner channel. The channel appears dark, as it is filled with stain, and the rod itself appears white or grey in TEM, as the electron density in proteins is lower than in the negative stain.

AFM is also a useful tool for direct visualization of TMV rods. The interaction of the species in solution with the plating substrate is different than in TEM, which leads to some differences in the final images. These differences have the added advantage of giving a more complete picture of the actual state of the solution at the time of plating. In TEM, a thin film of solution was dried down onto the TEM grid after staining. The thin film is usually representative of the true relative concentration of species in solution. As such TEM provides an unbiased snapshot of the sampled solution. However, because a 3 dimensional area is being condensed onto a 2D surface through drying, objects will often lie on top of each other and perhaps distort each other in ways not seen in solution, and because of this condensation, it can sometimes be difficult to visualize individual objects. In contrast in AFM, since the plating is generally done on a freshly cleaved mica surface, the main interaction which attracts species in solutions is charge. Thus the strength of the interaction can be tuned by altering both the charge of the species in solution and that of the mica surface. This often leads to a disproportionate amount

of a given species on the surface as anything which can bind strongly or move more quickly to the surface will end up dominating in the resulting AFM image.

In tapping mode AFM, instead of bombarding the sample with high energy electrons, as is done in TEM, a cantilever with an extremely small tip attached to it is brought very close to the surface, and the tip deflection due to interaction with the sample is monitored using a laser impinging on a photodiode. This process can be used to detect differences in the heights of various surface bound objects, as well as to identify boundaries between various surface bound objects. Although AFM images are recorded as 3 dimensional, it is important to realize that only the height direction (Z) can be used for reliable measurement of object features. The width and length (x and y) dimensions measured in the AFM will be distorted due to tip broadening. Tip broadening occurs because all AFM tips have a non-zero diameter, and thus interact with the objects being studied at a variety of x and y distances from the center of the object in question. Although the AFM tip manufacturers do provide a range of tip diameters, the AFM tip is very small, usually made up of only a few thousand atoms. The tip is therefore very delicate and can both gain and lose matter over the course of a scan, essentially randomly changing the strength of the interaction between tip and the surface being studied and therefore changing the amount of tip broadening which is occurring. As such, even if the tip diameter was known precisely, it would still not be a reliable method for determining the x and y dimensions of surface bound objects. In this research, all AFM scans were done under dry conditions, thus even the Z coordinate would be expected to be distorted from its solution state by both charge interactions and drying effects.

Initially, direct visualization was challenging. The obtained TEM and AFM scans were inconclusive. As can be seen in figure 2.11, there was significant interference in the TEM image. Many TMV disks could be seen clearly and are measured to have an 18 nm diameter. However, any rods present in the image are more ambiguous. Identification of individual rods was very difficult. There were some rod-like objects present in figure 2.11. In fact, there appear to be many rods running in parallel, with the diameters of the proposed rods consistent with what would be expected from TMV rods and also the inner channel can possibly be visualized in

certain places. However, it is impossible to be conclusive as to how many rods are present and what are the various sizes of rods. It is also difficult to see where one rod ended and another began. The need for a negative stain meant that any objects which are in solution will also end up being visible in the TEM image. Therefore only an object's size and shape can be used to distinguish it from other species present in solution. Since the 2D film required for TEM is condensed from a 3D volume, and because all objects present in the film were stained, only a moderate concentration of total contaminants is required to interfere with conclusive identification of TMV-like rods.

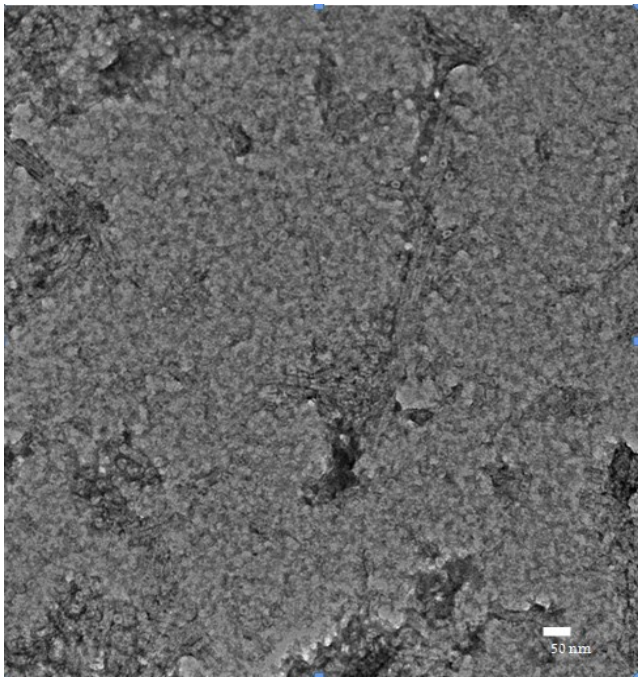


Figure 2.11, TEM of poly A TMV-CP mixture at pH 6.8. TMV-CP disks can be identified by their 18 nm diameter and 4 nm inner channel. There also appear to be TMV-like rods present which can be identified by their 18 nm diameter. However, the presence of rods cannot be concluded with certainty as it is not possible to resolve individual rods in the initial TEM specimens.

Although there are most likely rods present in figure 2.11, because they cannot be individually identified it is impossible to be certain that the objects in question are indeed TMV-like rods. Many of the objects have the correct diameter to be rods, and some also possibly appear to possess an inner channel. Objects which have diameters near that of TMV rods are rare, and none of them have inner channels. However, in the absence of clear, individually resolved rods without surrounding objects, it is not possible to rule out that some combination

of contaminants could have combined with the stain to appear similar enough to TMV rods to be confused for them at the resolution attained in figure 2.11.

AFM images of TMV-like rods are even more difficult to interpret. In AFM, since the probe is always perpendicular to the TMV rods, it is not possible to view the inner channel of any TMV rods present. Since the length and width are also distorted due to tip broadening, only the height and shape of an object can be used to identify it. It is therefore imperative that the objects being viewed under AFM be well resolved. Initial plating resulted in the objects presumed to be TMV rods being poorly resolved. As can be seen in figure 2.12, the objects presumed to be TMV-like rods could not be at all conclusively identified.

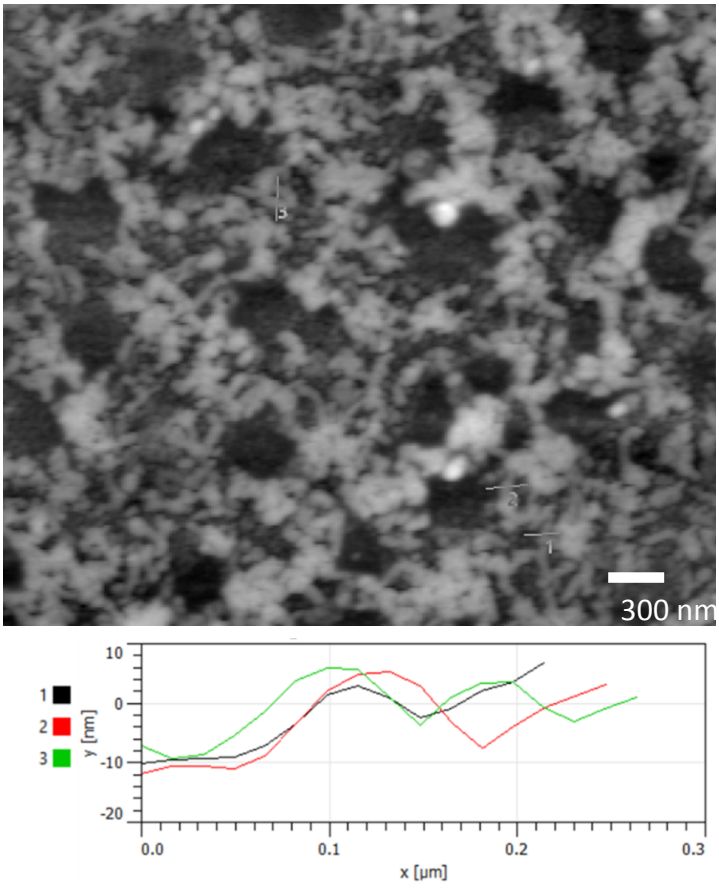


Figure 2.12, AFM of poly A TMV-CP after reaction. There are rod shaped species present in the AFM with 15 and 16 nm diameters which are consistent with what would be expected from TMV-like rods. However in a similar fashion to the initial TEM images, it is not possible to isolate the individual rods, and thus it cannot be concluded from this image with any certainty that TMV-like rods are produced.

The plating technique in AFM is fundamentally different to that of TEM, as in AFM plating the solution was not dried down onto the surface. In the AFM plating, some of the solution containing the presumably assembled poly A mediated TMV-like rods is exposed to a freshly cleaved mica surface. Electrostatic interactions draw negatively charged species to the mica surface, and then the surface is rinsed with de-ionized distilled water to remove any salts which would interfere with AFM imaging. Since species which are more strongly attracted to the mica surface are more likely to bind and stick, the population of species on the mica surface may not reflect their concentration in solution. AFM plating can be further complicated by multiple layers of deposited material, as this would result in an uneven starting point for AFM height measurements and thus make them unreliable. In the AFM image seen in figure 2.12, there appear to be features of the height expected for TMV rods and normally nothing else in solution is expected to have similar height characteristics. Nevertheless, not only is it impossible to get size distributions for the proposed TMV-like rods, but the fact that so many objects were deposited onto the mica makes resolution of individual rods impossible. As such, in a similar fashion to how the initial TEM images are not fully conclusive, the initial AFM images are at least equally inconclusive.

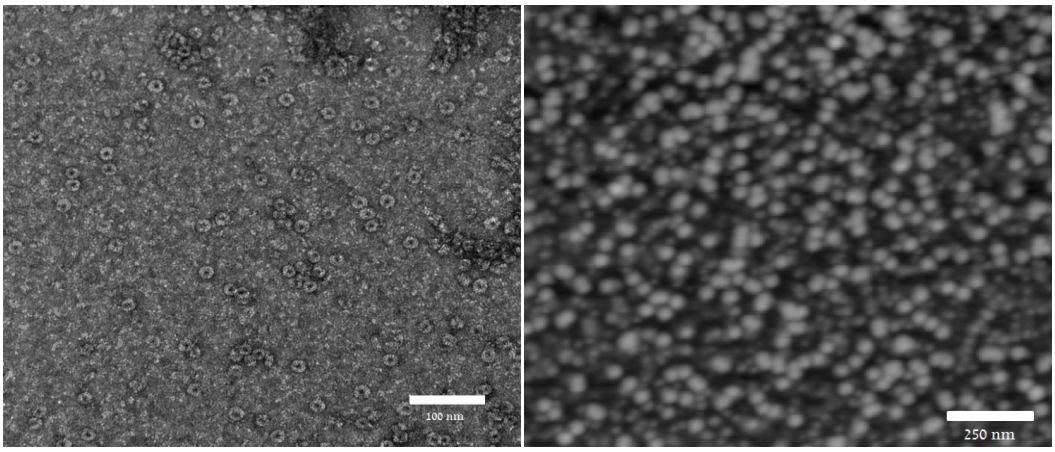


Figure 2.13, TEM (left) and AFM (right) of TMV-CP at pH 6.8. In the TEM (left) the TMV-CP disk can be identified by its 18 nm diameter and 4 nm inner channel while in AFM (right) the inner channel is not resolvable in the AFM scan. The TMV-CP disks are identified by their circular shape and 4 nm height. In the AFM scan shown here there is a double layer of TMV-CP disks, some being on top of others, which results in some brighter circles from the top layer of disks and some duller circles which form the bottom layer of disks. Multiple layers of disks is common in TMV-CP AFM images where the initial concentration of TMV-CP is high.

In contrast, initial TEM and AFM images of only TMV-CP at pH 6.8 result in well resolved TMV-CP disks. In TEM, the disks can be identified by their circular nature, 18 nm diameter and 4 nm inner channel. In AFM because the AFM tip normally is larger than the diameter of the inner channel, the inner channel of TMV-CP disks is rarely resolved. However TMV-CP disks are still able to be identified by their 4 nm height and circular profile.

Regardless of whether or not TMV-like rods are being formed, from figures 2.11, 2.12 and 2.13 it would appear that the introduction of poly A has significant consequences for the general interaction of TMV-CP with itself. The images suggest that either TMV-like rods are interacting with other rods and TMV-CP disks in large extended, loosely bound assemblies, or that poly A initiated a similar, large extended assembly which also contained some rod-like species. Since these extended assemblies are seen in both the TEM and the AFM images shown in figures 2.11 and 2.12, it is likely that those same assemblies are also present in solution before the samples are prepared for imaging. TMV-CP evolved in nature not to form extended assemblies under low ionic strength conditions at neutral pH. Most likely, that is the reason why TMV-CP can be imaged without interference. However, TMV-CP never evolved to be able to interact with poly A, and as such it is not unusual that poly A can be capable of initiating unexpected changes to the behaviour of TMV-CP in solution. The basis for such an interaction will be discussed in more detail in a later section.

As a comparison, pH induced TMV-CP rods were also imaged. Since no RNA is present, the mechanism for pH induced TMV-CP rods is quite different and generally requires substantially more time. In RNA mediated TMV rod assembly, the RNA acts as an assembly template and thus ensures that the formed helix is continuous. With pH induced TMV-CP rods there is no such agent, and as such the rods which are initially formed are full of imperfections known as "nicks". A 24 hour "annealing" period is required for the majority of the TMV-CP contained in the rods to rearrange itself into a continuous helix[23].

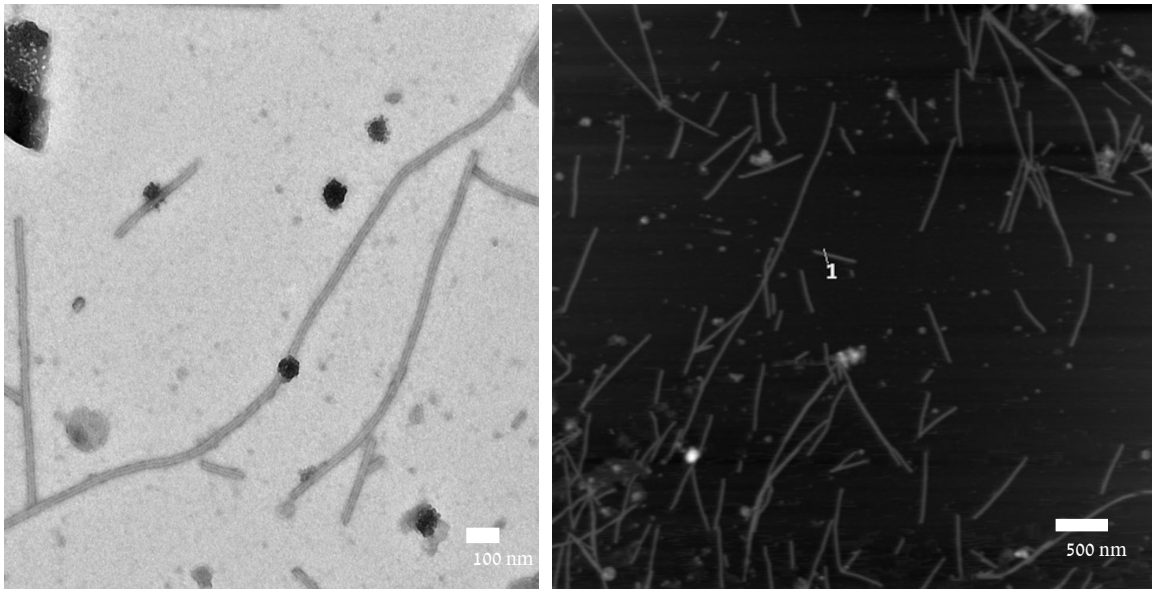


Figure 2.14, TEM (left) and AFM (right) of pH induced TMV-CP rods. TMV-CP rods are formed by dialyzing TMV-CP in pH 5.5 buffer for 24h. Very few disks are present, as TMV-CP rods are the dominant species in this area of the TMV-CP phase diagram. In both the AFM and TEM images the lack of disk species and the well resolved rods make conclusive identification of TMV-CP rods straight forward.

In figure 2.14, TMV-CP rods are Imaged with TEM and AFM without significant difficulty. The formed rods seem to be well dispersed and well resolved, with no evidence seen of the rods forming extended assemblies. No RNA is present to template the length of the rod, and as such the rods formed are seen to have a huge variety of sizes ranging from ~50 nm to several microns in length. While the shorter rods are rather rigid, most of the longer rods are slightly bent. Many of the rods in the 300 - 500 nm size range are not as rigid as would be expected if the rod contained RNA. This slight deviation from a perfect rod form is not unexpected and is due to the fact that the TMV-CP remains slightly less ordered in the absence of its RNA, even under acidic conditions[24].

Initial TEM and AFM results suggest that TMV-like rods are being produced as the result of the interaction between TMV-CP and poly A. However, no conclusive images were seen. Rather than modify the plating technique used for TEM and AFM analysis more plating was done instead. With the view that surely there must be some rods which would be reasonably resolved and not entangled in an ambiguous mess. Although such sample preparation is rare, it is occasionally possible to obtain conclusive TEM and AFM images of poly A mediated TMV-like

rods. Interestingly, only TMV-like rods of considerable length are viewed using this method. The technique utilized is not a solution to the problem of plating TMV-CP poly A mixtures, as such many of the TEM and AFM sample preparations still resulted in highly inconclusive structures. Therefore it is possible that there are smaller rods present in solution which are more apt to be tangled up in extended assemblies which are highly challenging to image and simply went unobserved in the few successful rounds of TEM and AFM sample preparations. The length of these rods measured using this method with TEM and AFM varies from about 400 nm - 800 nm.

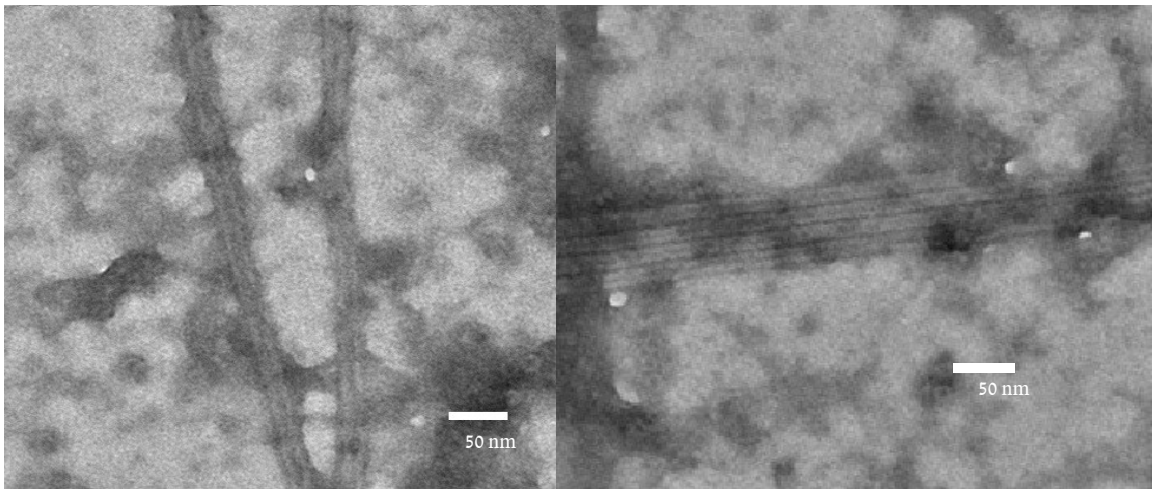


Figure 2.15, TEM of poly A mediated TMV-like rods. TMV-like rods are plated 2h after the start of the reaction between poly A and TMV-CP at pH 6.8. The rods which are clearly observable are quite long with sizes ranging from 400 nm to 800 nm. The inner channel is clearly visible and the helix can be seen to be continuous. Along with the 18 nm diameter of the TMV-like rods is sufficient to conclude with certainty that TMV-like rods are formed.

Clearly there are TMV-like rods in figure 2.15. There are also some disks present, but as can be seen in figure 2.15 the regions near the observed TMV-like rods seem to be mostly depleted of disks. Either many of the disks are used in the construction of the rods or the disks are sequestered elsewhere along with smaller rods. Furthermore, the absence of nicks in the TMV-like rods viewed in figure 2.15 would highly suggest that the rods contain poly A, as pH induced rods would require more than 2h to fully assemble without the presence of nicks. Interestingly there seems to be a preference for the formed rods to line up parallel to each other. This phenomenon is not seen in images of pH mediated TMV-CP rods from figure 2.14, however it has been previously reported in native RNA mediated TMV rod assembly[25]. The

extent to which the rods are aligned in figure 2.15 is larger than that which was previously observed, but this may be attributed to a combination of a slightly different pH used compared to previous studies, as well as the fact that the rods in figure 2.15 are longer than the 300 nm native TMV rods, and as such, have a greater area over which to interact with each other. Although it is unclear, it is possible that the reason why these rods are able to be directly imaged while most of the attempts are not is that the long rods which are assembled parallel to each other are able to disentangle themselves from extended assemblies or avoid forming them, while smaller, less oriented rods are not.

In contrast to the TMV-like rods seen in figure 2.15, a control from same solution of TMV-CP which was not exposed to poly A contains no rods as can be seen in figure 2.16. The majority of the solution is composed of TMV-CP disks which are identified by their 18 nm diameter and 4 nm inner channel.

In figure 2.15 there were some small rod like objects which are present. However, these are not small rods but actually stacked disks, which can be identified by the breaks at 90^{o} to the inner channel axis. As previously mentioned, these breaks are known as nicks and are common not only in stacked disks but also in the early stages of pH induced rod formation[23].

The AFM images which are recorded using the same method as that used in figure 2.14 also contain well defined TMV-like rods with a 15 nm height and characteristic TMV rod shape which are seen in figure 2.17. Similar to the TEM in figure 2.16, the TMV-CP solution used as a precursor to the TMV-like rods which are seen in figure 2.13 contain no rods which are visible in AFM scans.

In Figure 2.17, although rods are observed, there is also a large amount of TMV-CP disks present. The presence of disks is expected, since TMV-CP is in excess in the reaction mix. However, the apparent amount of disks is much larger in AFM scans compared to TEM images of the same solution. This is likely due to the fact that TMV-CP disks both interact highly favorably with the mica surface and are highly mobile in solution when compared to TMV-like rods. Thus, a full coverage of disks can be obtained in seconds, even after a TMV-CP disk solution has been diluted 1000X before exposure to the mica. Overall, the number of rods

which are able to be imaged is not as high as would be expected, even considering the length of the rods. Considering that the observation of these rods is a rarity, and is rather uncommon, it seems more likely that the other material present which was not identifiable consists of smaller rods mixed with disks, which would account for the lack of disks in the TEM image.

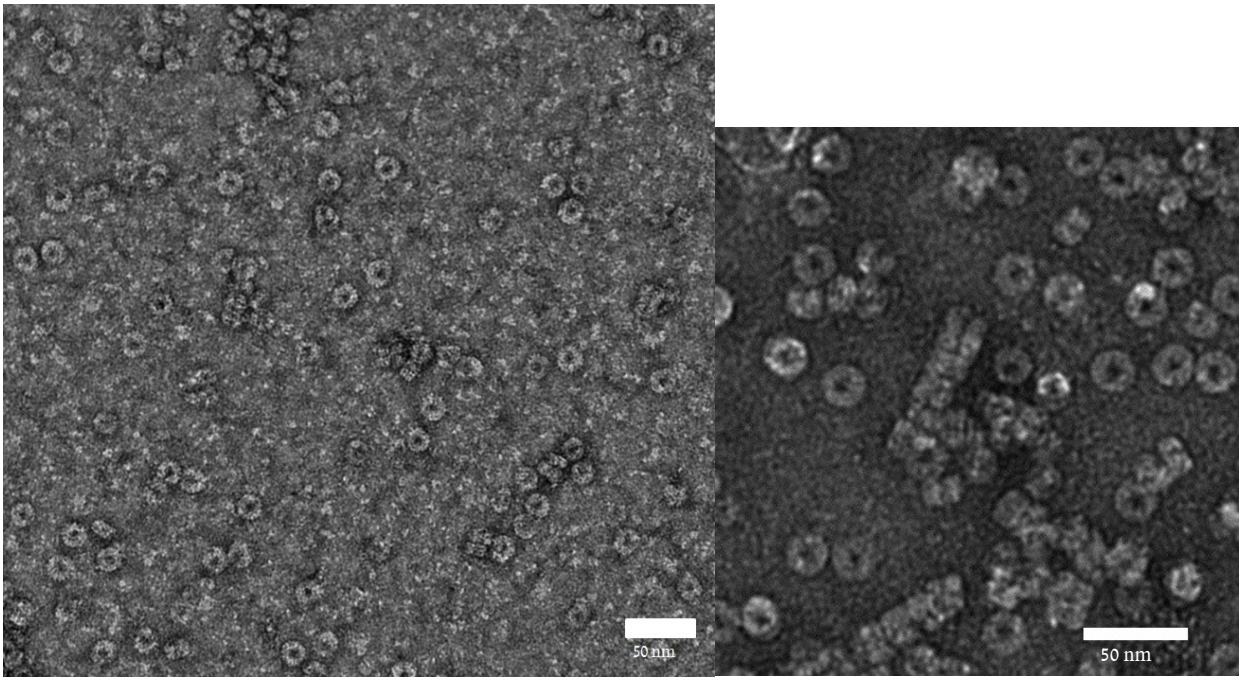


Figure 2.16, TEM of TMV-CP at pH 6.8, wider view (left) and higher resolution (right). Although there are some rod-like species present they are drastically shorter than the poly A mediated TMV-like rods seen in figure 2.14. Additionally the small rods seen are not continuous but contain many breaks known as nicks which can be seen in the TEM images.

When figures 2.14 – 2.17 are considered together, they provide conclusive evidence that poly A is able to mediate TMV-like rods from disk containing solutions of TMV-CP at pH 6.8. The rods observed in figures 2.15 and 2.17 are of fairly consistent length compared to those formed through the lowering of the pH from figure 2.14. TMV rods formed in the absence of RNA have no length template. Since the TMV-like rods which are observed in figures 2.15 and 2.17 have a much narrower length distribution than those seen in figure 2.14, they likely contain an RNA strand to template the length. Since the only RNA strand which is present in solution is poly A, and furthermore no rods are seen in the control figures, it is difficult to imagine a mechanism other than through poly A mediated TMV-like rod assembly which would produce these observed results.

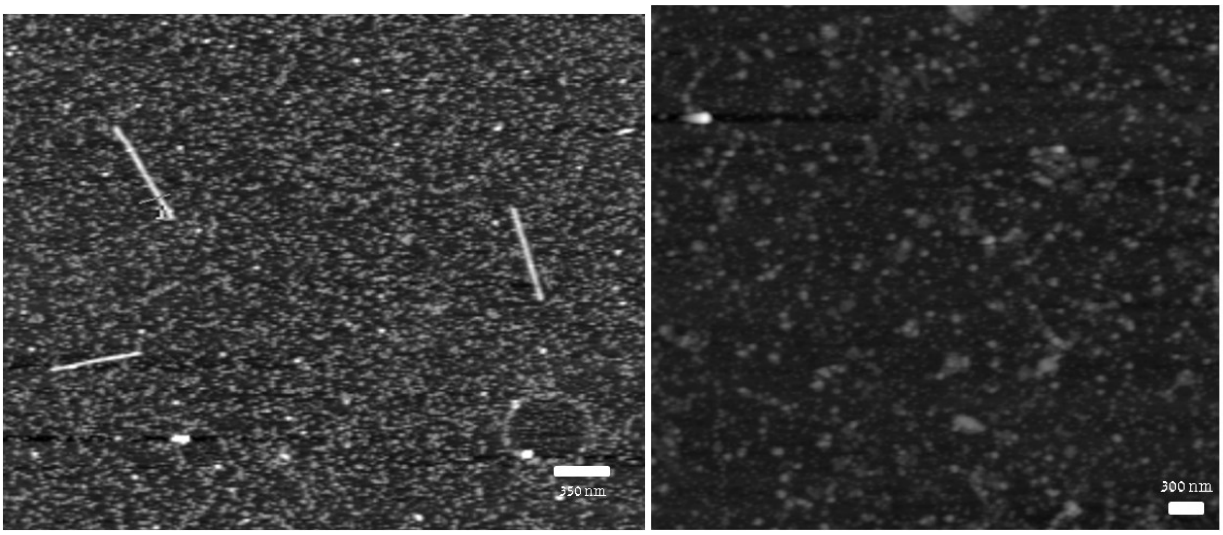


Figure 2.17, AFM of poly A induced TMV-like rods (left) and starting solution of TMV-CP at pH 6.8 (right). The TMV-like rods could be identified by their rod shape and their 15 nm height. Many disks are also present as expected since the disks are in excess. The TMV-CP solution which was used to assemble is plated onto freshly cleaved mica and the resulting AFM scan can be seen on the right. No rods are seen in the AFM scan thus the rods formed must have been due to interaction with the poly A.

As previously mentioned poly A mediated TMV-like rods seem to align in a parallel fashion even more strongly than what is typically seen for TMV rods formed using TMV RNA[25]. Since at pH 6.8, both the TMV-CP and the poly A are negatively charged, the formation of a TMV-like rod containing poly A should be even more highly charged and thus overall electrostatic interactions are not likely to be the explanation for the observed behavior. However as was discussed in chapter 1, the CP goes through some rearrangement during conversion from planar to helical. In addition, TMV rods have also been shown to exhibit pH dependent binding to inorganic salt precursors. The same groups which are responsible for this behavior are also capable of binding TMV rods together in a bundle to form a TMV superlattice through the binding of high concentrations of barium ions[26]. CP modification has also been shown to induce TMV-CP rods to form similar higher order structures[27]. It therefore seems reasonable to suggest that the outer surface of TMV and TMV-like rods, which contain both negatively and positively charged residues could interact with adjacent rods in order to produce the observed results. Since the rods would have to overcome overall electrostatic repulsion in order to come close enough with an adjacent rod to make a favorable interaction, there would only be a preference for aligned rods and not overall general alignment of the rods. The effect

would be further enhanced with longer rods, is likely to be concentration dependent, and explains the observed behavior of the TMV-like rods well. It could also be this same effect which leads to the extended poly A assemblies which have proven so difficult to image successfully. The fact that the alignment of TMV-like rods is not seen in AFM images is likely due to the difference in plating technique used for AFM sample preparation. The more complex technique likely gives more opportunity for the rods to be disentangled from each other, and would suggest that the interaction between the rods is a weak one as it is potentially overcome by fairly mild forces.

2.5 Examples of failed assemblies

In addition to the sample preparation difficulties which were initially encountered with poly A TMV-CP reaction mixture, the reaction itself seems to be frustratingly inconsistent, for no reason which was readily apparent at the time. The CD spectra, TEM and AFM of mixtures of TMV-CP and poly A RNA which fail to produce rods contain marked differences from the successful reactions and provide further evidence that indeed poly A mediated TMV-like rods are produced in the previous examples shown.

2.5.1 CD spectra over time of failed assemblies

Using CD, it is possible to identify a failed assembly of TMV-like rods from TMV-CP and poly A. As can be seen in figure 2.18, unlike with successful rod assemblies, there is very little change in the CD spectrum over time. There is a small decrease in the characteristic 264 nm poly A signal over time. As such, it cannot be ruled out that the mixture may produce a very small amount of TMV-like rods. However, the fact that the majority of the poly A remains in its single stranded helix, and that the decrease of the signal was so slow is inconsistent with that which would be expected from normal TMV-CP-poly A interaction kinetics. In control experiments where only the CD of the poly A is measured over time, a similar, slow decrease in signal intensity is sometimes observed and will be discussed in Chapter 3. Therefore it is more

likely that the reaction monitored in figure 2.18 simply failed to accomplish any assembly of TMV-like rods.

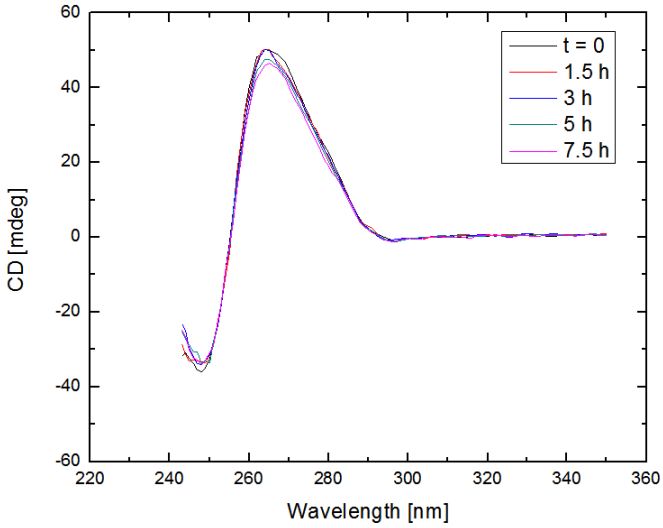


Figure 2.18, Time evolution of CD of mixture of TMV-CP at pH 6.8 and poly A which failed to assemble poly A mediated TMV-like rods. The final CD spectrum of the mixture is quite similar to the initial spectrum of the same mixture. Although the intensity of the characteristic poly A peak is slightly decreased in intensity, its location is not altered, and the slight decrease takes many hours. This would indicate that the majority of the poly A introduced to the TMV-CP solution remains in its natural helical form and did not interact with the TMV-CP.

2.5.2 TEM imaging of failed assemblies

While monitoring the CD spectra over time is the quickest way to be able to tell if a given reaction is producing TMV-like rods or not, the TEM and AFM images of the failed experiments are more conclusive, as clear images leave little to the imagination. Furthermore, when compared to the successful trials, the images of failed assemblies contrast greatly and look very different both from successful assemblies and from control images. Although staining and imaging poly A RNA proved to be difficult as it cannot always be identified, it is possible to view it occasionally in TEM both on its own and in experiments where the poly A helix CD signal is unperturbed.

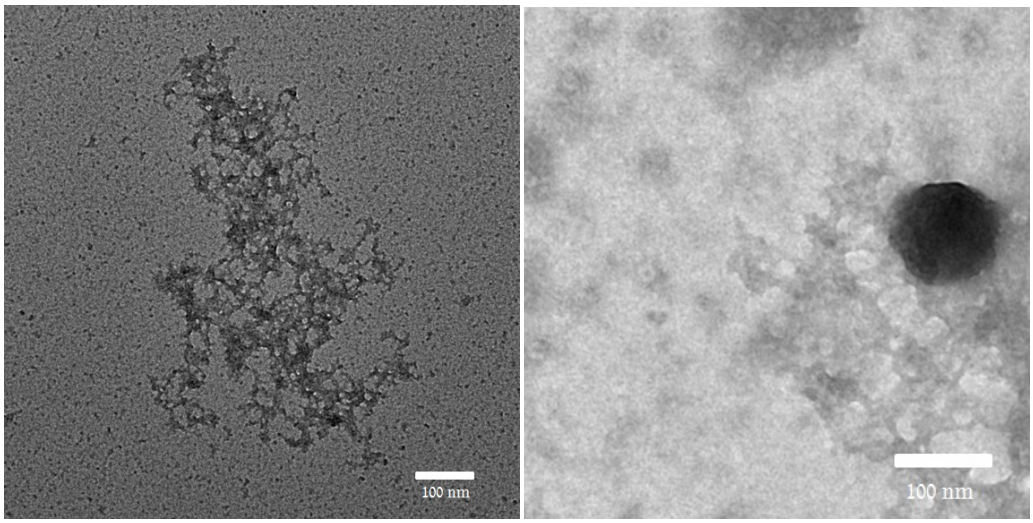


Figure 2.19, TEM of poly A (left) and failed poly A mediated TMV-like rod assembly (right). The poly A is the only species present in the left TEM image. Although the diameter of poly A is close to 1 nm, poly A seems to attract a large amount of stain to it, and as such shows up with a variable diameter usually between 7 - 10 nm. In the right image TMV-CP disks can be seen along with un-reacted poly A, which has the same form as it has in its free form.

As can be seen in figure 2.19, un-reacted poly A in the presence of TMV-CP and poly A which is alone in solution take on the same form in the TEM images. The interaction of the RNA with the stain is not always consistent, and presumably an excess of stain in the vicinity of a given poly A strand is necessary, as otherwise its small diameter would make it very difficult to visualize in the TEM setup utilized. As such it is not a particularly consistent method for showing that the poly A has remained in solution and has not been incorporated into a TMV-like rod. Visualizing the poly A is not as easy in the presence of TMV-CP disks, as they tend to sequester most of the stain, however as can be seen in figure 2.19, it is still sometimes possible to recognize it. Although there are some differences between the form of poly A in the control and the mixture samples, they both have the same general appearance, and nothing else resembling poly A is ever seen in a mixture sample. It is not possible to make any conclusions on the state of poly A in solution based on the TEM images seen, as only the dried poly A is ever imaged. This is due not only to the difficulty in distinguishing staining effects from the actual poly A conformation but also due to the fact that the drying process may have also altered the state of the poly A.

Although the TEM images are inconsistent at being able to detect un-reacted poly A due to staining difficulties, the AFM is not. At pH 6.8, poly A is highly negatively charged and as such is expected to interact strongly with a freshly cleaved mica surface. The poly A used tends to be quite long strands, and as such is quite obvious in the AFM. At pH 6.8, poly A is essentially a long strand of connected negative charges, some of which are arranged in a helical fashion while others are arranged more randomly. Compared to TMV-CP, poly A is very flexible. As such the interaction of poly A with the mica surface is more complex than that of TMV-CP, and takes up different configurations in different sample preparations. The negative phosphate groups are attracted to the mica surface, while they are repelled from each other, all while base stacking between A bases is possible. The various interactions will also be affected to varying extents by solution conditions. Therefore it should not be overly surprising that poly A is able to take different forms on the mica surface. No effort has been made to characterize what conditions would lead to a specific conformation of poly A, as all of them are easily identifiable and unique.

When the poly A is imaged using AFM, it is easily detectable as it forms large features consisting of an apparent extended network. As such, individual poly A strands are not distinguishable. As can be seen in the trace of figure 2.20, not only does the poly A differ from TMV rods in shape, but it also in regularity. The height of the poly A is inconsistent across its structure, while no such irregularity exists in TMV rods. It is not clear whether the observed structure consists solely of poly A, or if there are counter ions which are also present within the form assumed by the poly A. Poly A is a highly charged polymer, and similar to other nucleic acid chains, should be able to interact strongly with counter ions in solution. It would therefore seem possible and helpful for the poly A to build up a substantial concentration of counter ions around itself. This would be able to account for the fact that no discrete objects were seen, with the ions providing a bridge for different poly A strands to deposit together onto the mica. The rinsing technique used was optimized for TMV-CP plating and thus may be insufficient to desalt the poly A, and possibly explaining its appearance in figure 2.20.

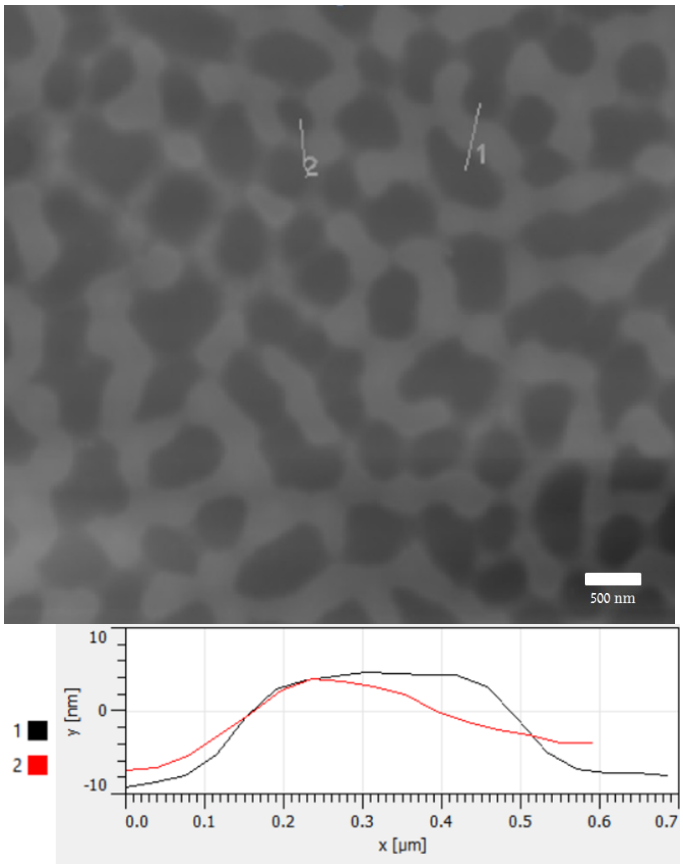


Figure 2.20, AFM image of poly A plated on freshly cleaved mica. The poly A seems to form some sort of extended network on the mica surface with feature heights around 14 - 15 nm but with shapes which are irregular and thus quite different and easily distinguishable from TMV rods.

Poly A has a very different appearance in AFM scans from any of the forms of TMV-CP. Furthermore, AFM scans of failed poly A mediated TMV-like rod assemblies show that in the presence of TMV-CP, the poly A takes a very similar form to that seen in figure 2.20. AFM therefore is a very conclusive technique for distinguishing between an assembled TMV-like rod and a reaction which fails to yield TMV-like rods. As can be seen in figure 2.21, not only are the TMV-like rods readily identifiable in successful reactions, but also the failed assemblies yielded poly A features which are highly distinguishable from both TMV-CP and poly A containing TMV-like rods. In figure 2.21 both reactants can be seen in their natural, unaltered form. The poly A forming an extended network on the mica and the TMV-CP disks being identified by the characteristic size and shape.

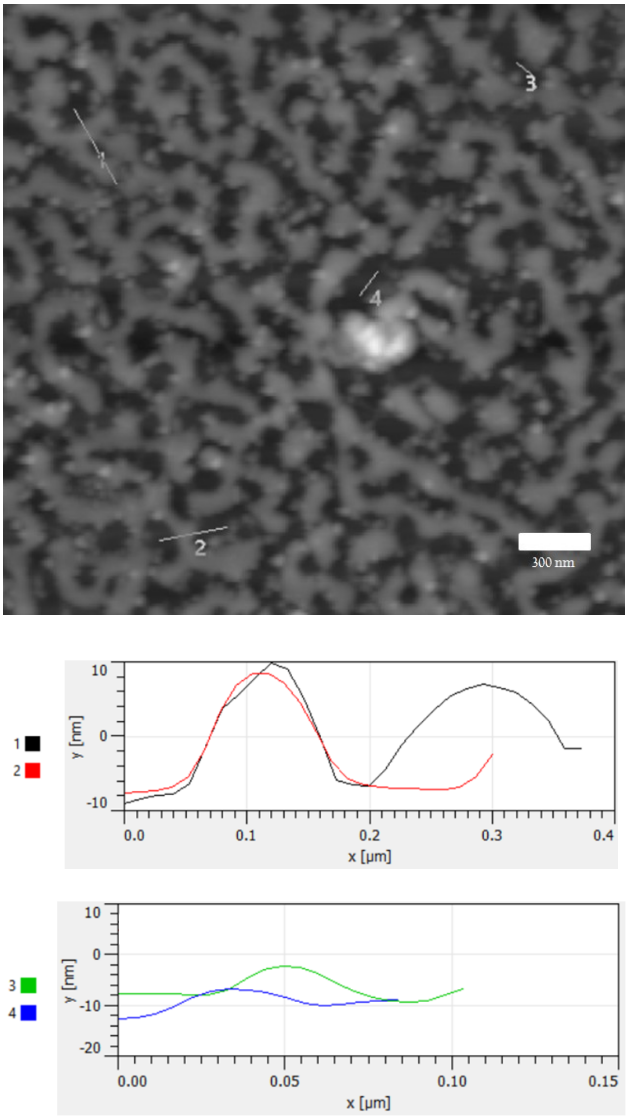


Figure 2.21, AFM of mixture of poly A and TMV-CP at pH 6.8 which failed to assemble TMV-like rods. Both the poly A and the TMV-CP disks can be clearly identified as being separate from each other with no TMV-like rods being observed. The poly A takes a form very similar to that seen in the poly A control AFM images albeit with slightly larger feature sizes of around 20 nm in height. The TMV-CP disks are identified by their 4 nm height and disk like shape.

Poly A is able to form more circular features in some AFM sample preparations. It is not clear what conditions affect the state of the poly A, but all the possible configurations of poly A on freshly cleaved mica result in features which are highly distinguishable from TMV-CP disks and rods as can be seen in figure 2.22. As such, no further attempts were made to determine the underlying factors of the effect. The heights and widths of the poly A objects seen in the AFM scans are correlated with larger diameter objects generally having a larger height. A

reasonably large distribution of spot sizes is observed, however it is not clear whether that represents an accurate distribution of the sizes of the poly A strands as a sizing gel was not done, and there is no guarantee that every spot represents a single poly A strand.

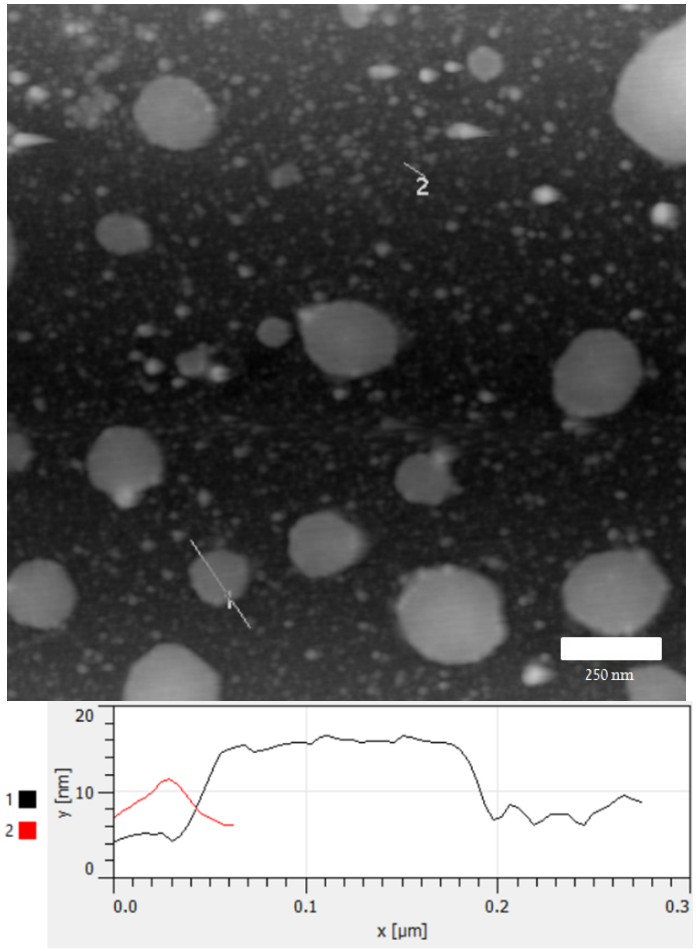


Figure 2.22, AFM of failed assembly reaction between TMV-CP and poly A in which the poly A assumes an alternate configuration on the mica. The poly A appears as large circles with 10 - 20 nm heights, with the height of the features dependent on the diameter of the poly A feature. As in Figure 2.21 the TMV-CP disks can be clearly identified by their 4 nm height and disk like shape. Although the poly A takes on different forms upon plating, all of them are quite different from any form taken on by TMV-CP and as such are easily distinguishable from the various forms of TMV.

Each of the techniques used to characterize the production of poly A mediated TMV-like rods is individually able to detect poly A mediated TMV-like rod formation. Although an individual technique could be inconclusive, on an individual experiment the various techniques never gave contradictory conclusions. In addition, both CD and AFM measurements are able to

clearly distinguish between a successful and an unsuccessful TMV-like rod assembly. The correlation of the results from different measurement techniques combined with the clear differences in the outcomes of those same experiments when the assembly fails provide excellent evidence that poly A mediated TMV-like rod formation occurs.

Overall it is clear that poly A mediated TMV-like rod formation is possible, and that those rods take on an identical form to both pH induced TMV-CP rods and native TMV rods. However, not only is there inconsistency in forming the TMV-like rods using poly A but also there is inconsistency in plating the formed rods in such a way that conclusive AFM and TEM images can be obtained. At the time there was no discernible reason for the inconsistencies, and if poly A mediated TMV-like rod formation is to be useful in future device construction it should be possible to reliably form the rod structures. Both of these problems were investigated and rectified in the following chapter.

References

(1) SCHÖN, A.; MUNDRY, K.-W. Coordinated Two-Disk Nucleation, Growth and Properties, of Virus-like Particles Assembled from Tobacco-Mosaic-Virus Capsid Protein with Poly (A) or Oligo (A) of Different Length. *Eur. J. Biochem.* **1984**, *140* (1), 119–127.

(2) Suck, D.; Manor, P. C.; Saenger, W. The Structure of a Trinucleoside Diphosphate: Adenylyl-(3', 5')-Adenylyl-(3', 5')-Adenosine Hexahydrate. *Acta Crystallogr. B* **1976**, *32* (6), 1727–1737.

(3) Saenger, W.; Riecke, J.; Suck, D. A Structural Model for the Polyadenylic Acid Single Helix. *J. Mol. Biol.* **1975**, *93* (4), 529–534.

(4) Seol, Y.; Skinner, G. M.; Visscher, K.; Buhot, A.; Halperin, A. Stretching of Homopolymeric RNA Reveals Single-Stranded Helices and Base-Stacking. *Phys. Rev. Lett.* **2007**, *98* (15), 158103.

(5) Brahms, J.; Michelson, A. M.; Van Holde, K. E. Adenylate Oligomers in Single- and Double-Strand Conformation. *J. Mol. Biol.* **1966**, *15* (2), 467–488.

(6) Dewey, T. G.; Turner, D. H. Laser Temperature-Jump Study of Stacking in Adenylic Acid Polymers. *Biochemistry (Mosc.)* **1979**, *18* (26), 5757–5762.

(7) Fraenkel-Conrat, H.; Singer, B. Reconstitution of Tobacco Mosaic Virus IV. Inhibition by Enzymes and Other Proteins, and Use of Polynucleotides. *Virology* **1964**, *23* (3), 354–362.

(8) Steckert, J. J.; Schuster, T. M. Sequence Specificity of Trinucleoside Diphosphate Binding to Polymerized Tobacco Mosaic Virus Protein. *Nature* **1982**, *299* (5878), 32–36.

(9) Turner, D. R.; Mondragon, A.; Fairall, L.; Bloomer, A. C.; Finch, J. T.; Van Boom, J. H.; Butler, P. J. G. Oligonucleotide Binding to the Coat Protein Disk of Tobacco Mosaic Virus. *Eur. J. Biochem.* **1986**, *157* (2), 269–274.

(10) Turner, D. R.; Joyce, L. E.; Butler, P. J. G. The Tobacco Mosaic Virus Assembly Origin RNA: Functional Characteristics Defined by Directed Mutagenesis. *J. Mol. Biol.* **1988**, *203* (3), 531–547.

(11) Antao, V. P.; Gray, D. M. CD Spectral Comparisons of the Acid-Induced Structures of Poly[d(A)]Poly[r(A)], Poly[d(C)], and Poly[r(C)]. *J. Biomol. Struct. Dyn.* **1993**, *10* (5), 819–839.
(12) Causley, G. C.; Staskus, P. W.; Johnson, W. C. Improved Methods of Analysis for CD Data Applied to Single-Strand Stacking. *Biopolymers* **1983**, *22* (3), 945–967.
(13) Inoue, H.; Kuriyama, K.; Ohno, T.; Okada, Y. Circular Dichroism and Sedimentation Studies on the Reconstitution of Tobacco Mosaic Virus. *Arch. Biochem. Biophys.* **1974**, *165* (1), 34–45.
(14) Kankia, B. I. Binding of Mg2+ to Single-Stranded Polynucleotides: Hydration and Optical Studies. *Biophys. Chem.* **2003**, *104* (3), 643–654.
(15) Self, B. D.; Moore, D. S. Nucleic Acid Vibrational Circular Dichroism, Absorption, and Linear Dichroism Spectra. I. A DeVoe Theory Approach. *Biophys. J.* **1997**, *73* (1), 339.
(16) Rizzo, V.; Schellman, J. A. Matrix-Method Calculation of Linear and Circular Dichroism Spectra of Nucleic Acids and Polynucleotides. *Biopolymers* **1984**, *23* (3), 435–470.
(17) Vogel, D.; Jaenicke, R. Conformational Changes and Proton Uptake in the Reversible Aggregation of Tobacco-Mosaic-Virus Protein. *Eur. J. Biochem.* **1974**, *41* (3), 607–615.
(18) VOGEL, D.; JAENICKE, R. Circular-Dichroism and Absorption Spectroscopic Studies on Specific Aromatic Residues Involved in the Different Modes of Aggregation of Tobacco-Mosaic-Virus Protein. *Eur. J. Biochem.* **1976**, *61* (2), 423–431.
(19) Namba, K.; Pattanayek, R.; Stubbs, G. Visualization of Protein-Nucleic Acid Interactions in a Virus: Refined Structure of Intact Tobacco Mosaic Virus at 2.9 Å Resolution by X-Ray Fiber Diffraction. *J. Mol. Biol.* **1989**, *208* (2), 307–325.
(20) Izquierdo-Ridorsa, A.; Casassas, E.; Gargallo, R.; Marqués, I.; Tauler, R. A Comparative Study of Polyelectrolyte Effects and Conformational Changes in Several Purine and Pyrimidine Homopolyribonucleotides. *React. Funct. Polym.* **1996**, *28* (2), 127–137.
(21) Finch, J. T.; Klug, A. Two Double Helical Forms of Polyriboadenylic Acid and the pH-Dependent Transition between Them. *J. Mol. Biol.* **1969**, *46* (3), 597–IN27.
(22) Vetterl, V.; Guschlbauer, W. Protonated Polynucleotide Structures: XI. Polyadenylic Acid at High Ionic Strength. *Arch. Biochem. Biophys.* **1972**, *148* (1), 130–140.
(23) Durham, A. C. H.; Finch, J. T. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: II. Electron Microscope Observations of the Larger Polymers. *J. Mol. Biol.* **1972**, *67* (2), 307–314.
(24) Jardetzky, O.; Akasaka, K.; Vogel, D.; Morris, S.; Holmes, K. C. Unusual Segmental Flexibility in a Region of Tobacco Mosaic Virus Coat Protein. *Nature* **1978**, *273* (5663), 564–566.
(25) Harrison, B. D.; Wilson, T. M. A. Milestones in Research on Tobacco Mosaic Virus. *Philos. Trans. R. Soc. B Biol. Sci.* **1999**, *354* (1383), 521–529.
(26) Wang, H.; Wang, X.; Li, T.; Lee, B. Nanomechanical Characterization of Rod-like Superlattice Assembled from Tobacco Mosaic Viruses. *J. Appl. Phys.* **2013**, *113* (2), 024308.
(27) Bruckman, M. A.; Soto, C. M.; McDowell, H.; Liu, J. L.; Ratna, B. R.; Korpany, K. V.; Zahr, O. K.; Blum, A. S. Role of Hexahistidine in Directed Nanoassemblies of Tobacco Mosaic Virus Coat Protein. *ACS Nano* **2011**, *5* (3), 1606–1616.

Chapter 3, Improved methods of formation of poly A mediated TMV-like rods, effect of calcium on TMV poly A assembly reaction and reaction kinetics

As was seen in Chapter 2, initially the reaction between the poly A and the TMV was unpredictable and often did not work with no apparent explanation. The development of a consistent poly A containing TMV-like rod is very important as it would be necessary for any future device construction. There are many factors which have an influence on the RNA mediated assembly of TMV and TMV like rods including ionic strength of solution, pH of solution, concentration of the TMV-CP, concentration of the RNA strand, temperature and the condition of the reactants. It is therefore not simple to discover which underlying factor or factors are causing the assembly inconsistencies. After many trials it became possible to gain some understanding of the underlying causes of the inconsistencies in reliably producing poly A mediated TMV-like rods. The most important turns out to be how the TMV-CP was produced. TMV-CP is grown using E-Coli cells, and depending on the details of the growth, the resulting TMV-CP could be of varying quality. Produced TMV-CP which is not of sufficient quality is unable to form TMV-like rods. Further complicating the matter is the fact that the quality of TMV-CP is not a simple yes or no situation. In some instances, increasing the CP concentration and slightly adjusting the ionic strength and pH of solution could help the odds of obtaining a successful assembly. Furthermore, the addition of calcium ions could overcome a moderate inadequacy in the quality of the TMV-CP, and allow for formation of poly A mediated TMV-like rods. However, the addition of calcium did not guarantee the assembly reaction would be successful. If the TMV-CP was of low quality even the addition of calcium was not enough to induce the assembly of poly A mediated TMV-like rods to take place.

3.1 Effect of calcium On Assembly of poly A Containing TMV-like Rods

As is seen in figures 3.1 and 3.2, in some cases the presence of a small amount of calcium ion in solution allows poly A to induce the formation of TMV-like rods. However, the same concentration of poly A and TMV-CP under identical conditions without added calcium shows no sign of rod formation even after 2 hours post-mixing (figure 3.2). The absence of changes in the spectra over time indicates that no reaction is occurring. Although 0.5 mM

calcium may seem like a very small amount, it is sufficient to allow two calcium ions to bind per TMV-CP at the current concentration ratio. It has been shown that TMV-CP has two calcium binding sites, but that these binding sites are only accessible at pH 6.8 to TMV in its rod form which has been assembled using RNA and not to the free TMV-CP[1–5]. TMV has previously been shown to sequester ions unusually strongly[6], especially calcium and magnesium. In fact, even before TMV was shown to be able to bind calcium ions, it was known that purified TMV would contain a significant amount of calcium and magnesium ions, and extensive dialysis, usually in the presence of EDTA, is required in order to remove them[7,8]. Therefore it should not be overly surprising that calcium could have such a dramatic effect on the reaction between poly A and TMV-CP. Especially considering that in theory the addition of calcium should stabilize poly A containing TMV-like rods to a greater extent than TMV rods since poly A has an A nucleotide in every third position[9].

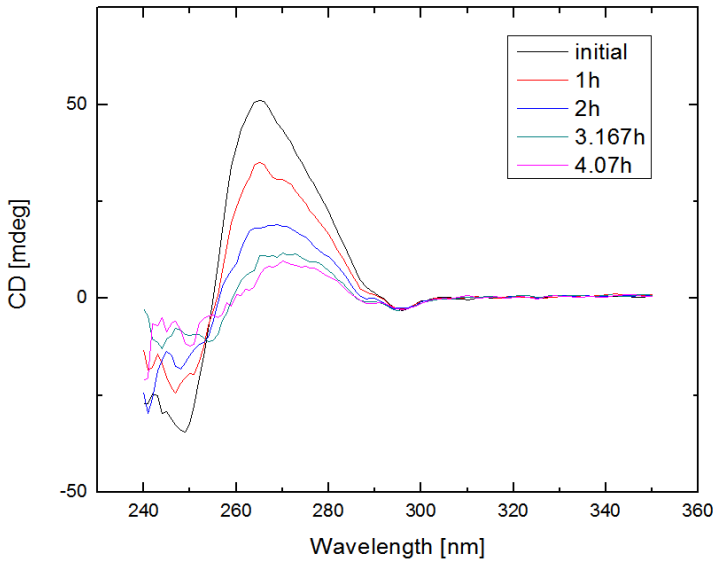


Figure 3.1, CD of Mixture of TMV-CP at pH 6.8 and poly A over time. TMV-CP concentration was 4.8 mg/mL while poly A concentration was .2 mg/mL. Also present in solution was .5 mM calcium chloride. Poly A mediated TMV-like rod assembly took place in the presence of calcium ion but not without the presence of calcium ion.

From figure 3.2 it is observed that not only did poly A fail to assemble TMV-like rods in the absence of small amounts of calcium ion, but that also the time evolution of the CD spectrum of the TMV-CP poly A mixture seen in Figure 3.1 could in no way be attributed to the interaction of poly A with calcium. If anything, the poly A helix signal in the presence of calcium

seen in figure 3.2 was slightly stronger than without. It would be expected that if calcium would interact with poly A, it would stabilize the helical state and thus increase the intensity of the CD signal from the poly A. However, it would be presumptuous to attribute the very slight increase in poly A helix CD signal intensity seen in figure 3.2 to an interaction between the poly A and the calcium ions as the slight increase could just as easily be explained by a small deviation in the pipette delivery volume. Furthermore previous studies would suggest that at the concentration of calcium used to aid the assembly, there should be no interaction between the poly A and the calcium. This is because magnesium interacts much more strongly with poly A than calcium does, but much larger concentrations of magnesium than those used here would be needed to observe an effect on the CD spectrum[10,11].

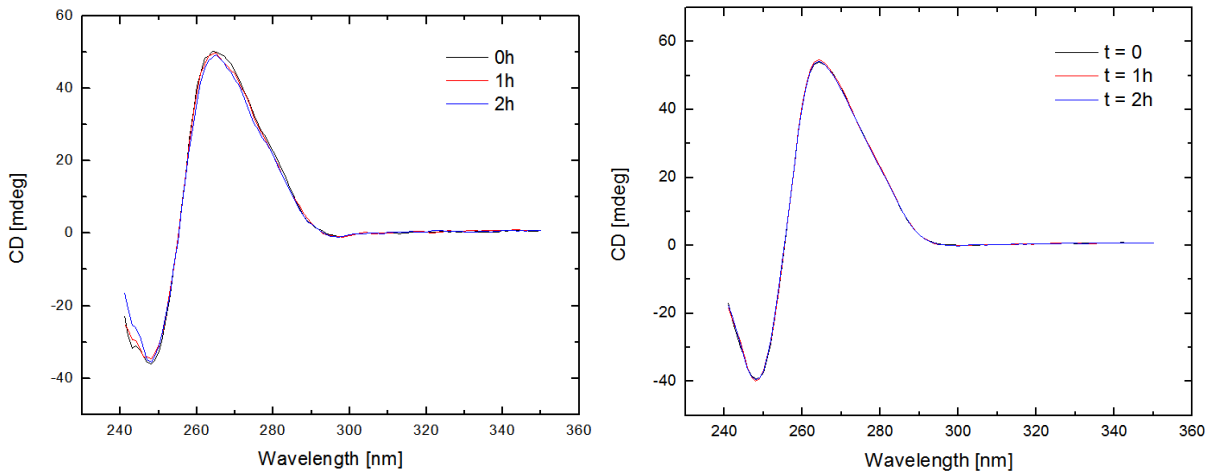


Figure 3.2, CD of mixture of TMV-CP and poly A over time (left). TMV-CP concentration was 4.8 mg/mL while poly A concentration was 0.2 mg/mL. The lack of time evolution of the CD spectrum of the mixture indicated that no TMV-like rods were formed with the batch of TMV-CP used in the absence of calcium. The concentration of calcium used of 0.5 mM was also not sufficient to produce the observed calcium effect in the absence of TMV-CP (right)

In figure 3.2, a small decrease in the poly A helix signal over time is observed. Such a small decrease could indicate a very slow reaction between the poly A and the TMV-CP. However, as was mentioned previously, such small decreases sometimes occur without explanation and do not result in TMV-like rods in the resulting TEM or AFM images. As such it is unlikely that the decrease in the CD signal could be attributed to TMV-like rod formation.

The addition of calcium ions does not guarantee that a given batch of TMV-CP will be able to form poly A mediated TMV-like rods, nor is it impossible to assemble TMV-like rods in the absence of calcium ions. However, the effect of calcium is consistent from batch to batch. Thus if a batch exhibits a need for the presence of calcium in order to assemble into TMV-like rods, it does so for every experiment in which the same batch of TMV-CP is used. The only exceptions to this rule were from experiments conducted on TMV-CP which was produced more than 2 weeks prior to the experiment. In those cases, TMV-CP which initially exhibited a need for the presence of calcium to assemble into rods was no longer able to be assembled, regardless of whether calcium was present or not. Such an effect was not limited to calcium aided assembly of TMV-like rods, but in general, TMV-CP is unlikely to assemble into TMV-like rods after two or three weeks post-purification.

3.2 Effect of other divalent cations on assembly of poly A containing TMV-like rods

Although TMV is able to bind multiple cations, it binds most strongly to calcium. In order to better understand the effect which allows poly A to nucleate TMV-like rods in the presence of calcium, other divalent cations were tested. Zn^{+2} and Mg^{+2} were used in place of the calcium to see if the effect was specific to calcium or if any divalent cation would work just as well. As can be seen in figure 3.3 it would appear that the effect is specific to calcium, although further study is required to be conclusive. Although only a single experiment was done with other cations, neither of them appears able to help poly A nucleate TMV-like rods. In both cases, there is some decrease in signal, however the decrease is slow, and importantly there is no change in profile in the central peak at 264 nm. The noticeable decrease in poly A helix signal over time was seen in previous experiments where some contaminating agent was present. Since no TMV-like rods were seen in the AFM and TEM, likely the slow loss of poly A signal was due to the contaminating agent which could be eliminated with a more careful preparation of further experiments.

From figure 3.3 it appears that Zn^{+2} and Mg^{+2} are not able to help nucleate TMV-like rods in the manner that calcium is able to since no loss of poly A CD signal is observed over the expected timescale. However, it is possible that they would be able to do so at a higher

concentration, since only one concentration was tested. Calcium is able to nucleate TMV-like rods at all concentrations of Ca^{+2} tested, 0.15 mM being the lowest. However, of all the divalent ions tested, calcium binds the strongest to TMV. It may be that Zn^{+2} and Mg^{+2} simply are not able to bind strongly enough to initiate the reaction. However, it is also possible that if at a higher concentration that a similar effect to that observed with calcium could be observed. Magnesium would have to be introduced at a concentration near 1 mM in order to achieve the same TMV binding coverage as calcium at 0.25 mM[12]. Presumably an even higher concentration of zinc would be required.

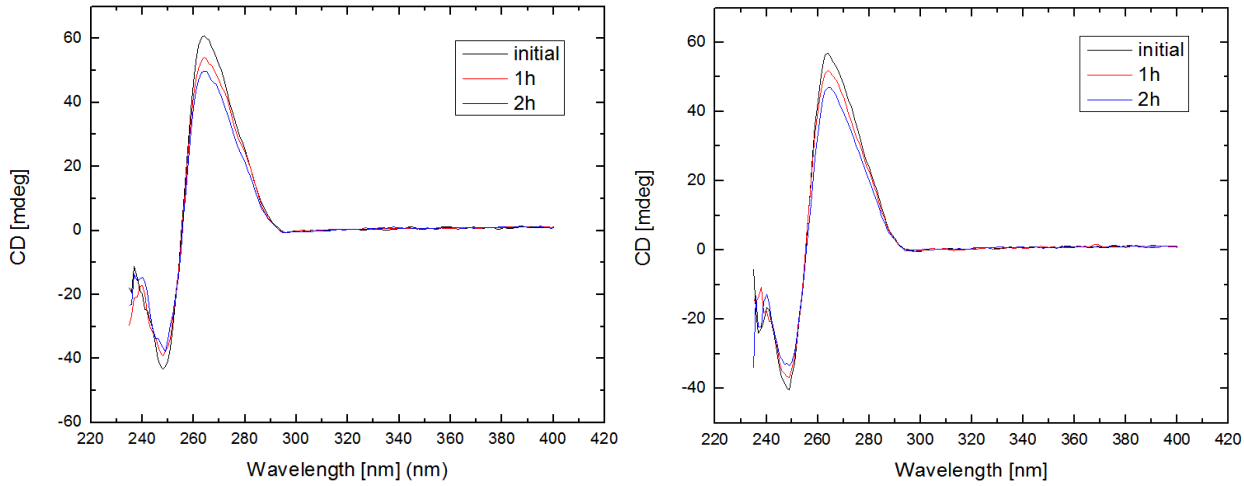


Figure 3.3, CD of poly A and TMV-CP at pH 6.8 in the presence of 0.25mM zinc (left) and magnesium (right) ions. In both cases the slow timescale and lack of change of profile of CD spectra suggests that no TMV-like rods are produced.

The other possibility is that no amount of magnesium or zinc would be able to help initiate TMV-like rod formation, and only the strength of the interaction between calcium and TMV allows calcium to help poly A form TMV-like rods. The main difficulty in using poly A to assemble TMV-like rods is in the nucleation step, as under normal solution conditions the OAS is significantly more efficient at nucleating TMV rods compared to poly A. Therefore it seems likely that calcium would help the reaction by aiding in the nucleating step. Since only a small portion of the TMV-CP would be involved in nucleation, and the majority of the CP would instead be involved in elongation of the TMV-like rod, such an interaction should proceed even at very low concentration of any ion for which the effect would be observed.

3.3 Reaction kinetics and their dependence on the presence of calcium ion

In order to better understand the effect of calcium and also to try and gain a better understanding of the mechanism behind poly A mediated TMV-like rod formation, the kinetics of the reaction were studied. This was done using the CD spectra of the poly A TMV-CP mixture, which can easily be monitored over the required timescales. The majority of the CD signal is due to the poly A. Therefore, to a first approximation, the loss in poly A signal is proportional to the production of TMV-like rods. The use of the CD technique requires low concentrations of RNA, and thus the signal responds linearly with increasing concentration and Beer-Lambert's law can be applied.

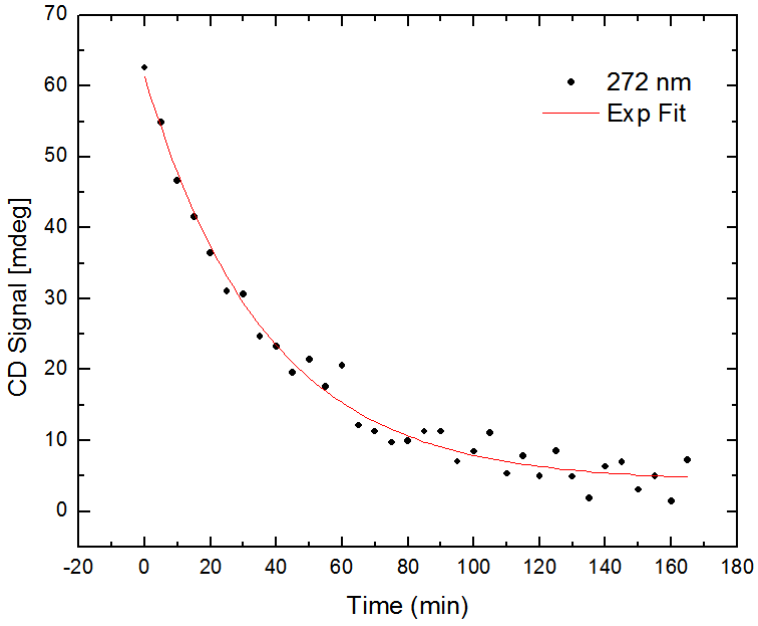


Figure 3.4, Decay of CD signal from poly A in a poly A TMV-CP mixture over time. The pH of solution was 6.8 and 0.25 mM calcium chloride was also present in solution. Final concentration of poly A was 0.3 mg/mL while that of TMV-CP was 5.5 mg/mL. The signal decreases in an exponential fashion which is consistent with normal virus kinetics. Approximately 50% of the poly A is used up in 30 minutes, which would correspond to a rate of addition of approximately 1 disk every 2 minutes to a growing rod in TMV, however such a description is not thought to be valid for poly A mediated TMV-like rod growth.

Figure 3.4 shows that the decrease in poly A signal over time from the mixture of poly A and TMV-CP is consistent with TMV kinetics, and thus can be fit to an exponential curve. Although the TMV-CP CD signal is non-zero, it is around 30 times smaller in magnitude than that from the poly A helix, and as such does not make a major contribution and can thus be ignored

without significantly changing the kinetic data extracted from the graph. Normally 260 nm is used for kinetic plots, as the spectrum of the assembled poly A containing rods have a CD signal of zero at that wavelength and thus make a convenient baseline. However in the example shown in figure 3.4 the 260 nm data is extremely noisy and gives very similar final outcomes, thus the 272 nm data is presented.

Using the kinetic data, the rate of addition of disks to poly A is calculated as approximately 1 disk every 2 minutes. This is considerably slower than the fastest recorded rate of addition of disks to TMV-RNA of 1 disk every 6 seconds, but also considerably faster than some reported completion times which require up to 24 hours to finish assembly. However, several factors must be considered. Most importantly is that because the mechanism of assembly of poly A mediated TMV-like rods is critically different to that of TMV rods, the calculated addition speed cannot be valid, since this rate assumes that all TMV-like rods nucleate at a similar time and then extend at the same rate. Unlike with TMV RNA, where all rods are nucleated before any of them finish, with poly A any rod which is nucleated finishes assembling very quickly, well before all the poly A strands have been nucleated. Thus, any bulk calculations which calculate the average rate of addition of disk are fundamentally flawed, as the majority of the poly A strands would simply be sitting un-nucleated in solution with only a small percentage of the overall poly A contributing to the average rate of disk addition.

Studies which could examine each strand individually would therefore be required in order to get a true measure of the rate of addition of disk to a growing poly A mediated TMV-like rod. Previous studies have pegged the rate of addition of TMV disks to a growing poly A rods at 6 times faster than that of TMV disks to TMV-RNA, which would correspond to 1 disk per second. However, the measurement of the rate was based on the number of completed rods 40 seconds after the start of reaction[13]. No earlier measurements were made, likely because the experimental design used at the time prevented faster measurement. As such, no upper bound could be placed on the speed of addition of disk to growing poly A containing TMV-like rod. Therefore it is a definite possibility that addition of TMV disks to growing poly A containing TMV-like rods could proceed even more rapidly than 1 disk per second. A speed of 1

disk per second would correspond to a 300 nm TMV-like rod being assembled in 60 seconds, with the rod being assembled even faster if the true speed of addition is faster than 1 disk per second.

Although not extensively studied, no AFM or TEM images of early events in solution ever demonstrated incomplete TMV-like rods being observed in this work. Conceivably it might be possible for partially complete rods to finish assembling while nucleation of new rods was suppressed during the drying of samples for imaging although this appears unlikely as interactions with the grid surface would be more likely to slow down any reactions occurring. Furthermore, even if that is the cause of the absence of incomplete TMV-like rods, it would still mean that addition of TMV disks to a growing TMV-like rod containing poly A occurs more quickly than what would be expected from native TMV. Since it was confirmed visually that the rate of addition of disk to a growing poly A containing TMV-like rod was quite rapid, it seems a definite possibility that unlike TMV-RNA, poly A is indeed able to add disks quickly in both the 5 prime and 3 prime directions. Although further study is required to be conclusive, it would be expected that if poly A was unable to quickly add TMV-CP in the 3 prime direction that some partially completed TMV-like rods would have been observed in the TEM or AFM images.

Usually in TMV kinetic plots there are 3 regions. Initiation, where the TMV-RNA is being nucleated and very few of the rods are large. Since not all the RNA has been nucleated yet in this region the overall rate of consumption of disks is not maximal. The second is the assembly region where almost all rods are growing, as such the rate of conversion of disk as well as the loss of free RNA are greatest in this region. The concentration of reactants is still high in this region and thus promotes assembly. The third region is the termination region, here the overall rate of consumption of disk again slows since the concentration of reactants has decreased. In addition many, but not all of the rods have been completed and can therefore no longer add disks. In figure 3.4 the initiation region is too fast to be recorded, an effect which usually occurs in solution conditions where TMV assembly occurs quickly. Although the mechanism which poly A uses to assemble TMV-like rods differs significantly from that which TMV-RNA uses to assemble TMV rods, as can be seen in figure 3.4 poly A containing TMV rods appear to form in a

fashion consistent with TMV kinetics, a result which has also been found previously[13]. Therefore the observed kinetics curve is likely due to the probability of nucleating a stable poly A containing TMV-like rod. This would mimic the curve seen in TMV assembly kinetic plots, as initially there would be a greater probability of nucleation which would eventually result in a greater rate of consumption of both TMV disks and free poly A RNA. At later times, the concentration of reactants is lower and thus the probability of nucleation is lower. In addition it has been shown previously that longer poly A chains may be nucleated more quickly than shorter ones[13,14], thus there may be a propensity to have more shorter, un-nucleated poly A chains at later reactions times.

The reaction plotted in figure 3.4 is complete after around 2.5 hours. The fastest recorded poly A assembly reaction was complete in 1 h[13]. Considering that the concentration used here is higher than that used in the study with the fastest completion time, the method used here did not yield a particularly fast production of TMV-like rods. The solution conditions used here were chosen for their reliability rather than their speed of production. TMV and TMV-like rod assembly kinetics depend strongly on ionic strength and pH of solution[13,15–18]. Furthermore faster assembly speed can be achieved if TMV-CP is in large excess[19]. However, in this work, although the TMV-CP was kept in excess, the ratio of reactants was kept closer to 1 to 1 in order to get clear spectra of the final products and also to simplify imaging of the TMV-like rods. Therefore, since significant improvements in the reliability of the reaction have been made it should be possible in the future, if faster assembly was desired, to move the assembly reaction to pH 7, at lower ionic strength and with a larger excess of TMV-CP which should result in faster assembly of TMV-like poly A containing rods.

The ultimate goal of this research was to provide an alternative to TMV-RNA based assembly of complex TMV rod based nanotechnology device components. In order to achieve this poly A would likely have to add disks which had different mutations and chemically augmented groups to the same rod, with a controlled spacing. The rate of addition of the various mutants, with their added modifications might be drastically different from each other which could be compensated for by modifying the relative concentration of the various types of

TMV disks in solution. However, if the rate of addition of a particular mutant became so high that it became locally depleted in the vicinity of the growing rod it would have the effect of disrupting the intended spacing of the final TMV-like rod. Therefore it may end up being beneficial in the future to be able to considerably slow the rate of addition depending on what final composition of TMV-like rod is desired. Therefore the ability of poly A to assemble TMV-like rods over a wide range of rates may end up being beneficial in the construction of future devices.

Although the effect of small amounts of calcium ion is in some cases drastic, in other cases it makes no difference. It was eventually determined that there is an underlying factor which was of even more importance than whether or not calcium was present in solution. The quality of the TMV-CP turned out to be essential. This fact was difficult to piece together since "the quality of the protein" is a somewhat abstract and not easily quantifiable. The issue is further clouded by the fact that the binding region of TMV-CP is in a separate location to the residues responsible for the rest of the self assembly properties of the TMV-CP. Thus since even low quality protein would exhibit normal self assembly behavior but would not assemble poly A mediated TMV-like rods, only the interaction of the CP with poly A could be used to quantify its "quality". As such "low quality CP" is defined as TMV-CP which cannot assemble into poly A containing TMV-like rods, even in the presence of calcium. "Medium quality CP" is defined as CP which could assemble into poly A containing TMV-like rods, but only in the presence of small amounts of calcium ion. "High quality CP" is defined as CP which could assemble into poly A containing TMV-like rods with or without the presence of calcium.

The factors which most influence the eventual quality of the CP were the temperature at which the CP was grown in the Tuner cells and the strength of the cell line which was used. A strong cell line would produce medium to high quality CP during production at 30^{o}C while a weak cell line would mostly produce low quality CP with the occasional medium quality batch of CP. Further confusing the matter is that the quality dropped over time with even high quality CP eventually deteriorating to low quality CP after a certain period of time, usually about 3 weeks after TMV-CP purification. The variation of strength of cell lines used coupled with the natural

variation of the quality of CP produced by a given cell line produces the apparent inconsistency of the poly A mediated TMV-like rod assembly reaction as well as demonstrated the calcium effect. To date all CP which has been grown at 18^{o}C has been able to assemble poly A containing TMV-like rods with or without the presence of calcium. The addition of calcium should still be beneficial to older high quality CP which may no longer be of sufficient grade to assemble into poly A containing TMV-like rods on its own, however other potential methods of elongating the lifespan of the TMV-CP will be discussed further.

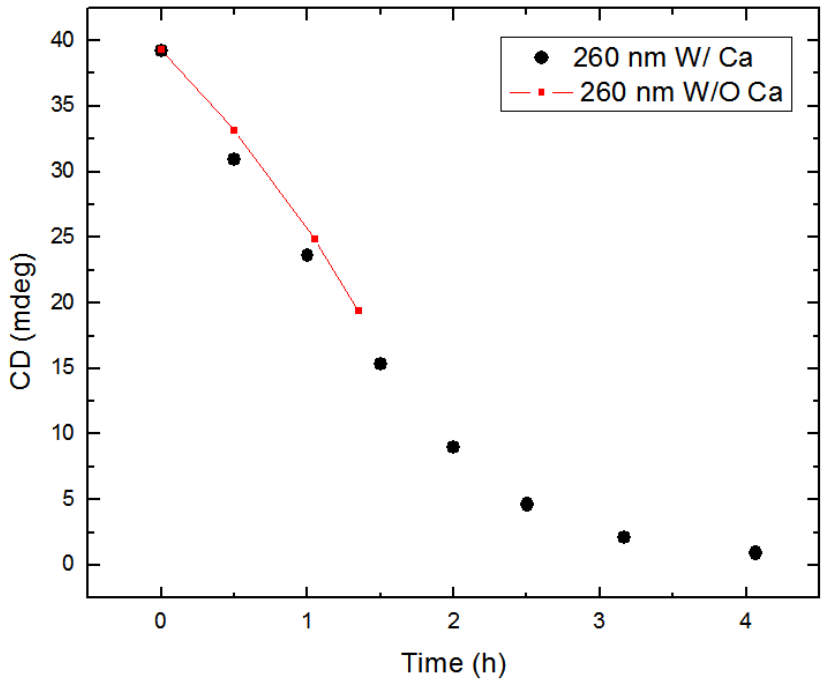


Figure 3.5, Reaction kinetics of high quality protein assembling into poly A containing TMV-like rods with (black) and without (red) 0.5mM calcium chloride present. The pH of solution was 6.8 and the concentration of poly A was 0.2 mg/mL while the concentration of TMV-CP was 4.5 mg/mL. In both cases the poly A CD signal at 260 nm is used to quantify the reaction progress. The presence of calcium does not appear to have an effect on the rate of reaction.

The addition of calcium seems to only determine whether or not the reaction will proceed, it doesn't seem to influence the rate of reaction. This can be seen in figure 3.5 where TMV-CP of the same concentration, with and without calcium, is exposed to poly A. As the figure shows, the rate of disassembly of the natural poly A helix follows the same curve in both instances. This suggests that poly A is used up at the same rate regardless of the presence of calcium. Although the reaction without calcium was not actively followed to completion the

reaction did go to completion as seen by CD, TEM and AFM taken the next morning. Given how closely the two curves line up it would be unlikely that the curve without calcium would deviate from that with calcium however that possibility can also not be ruled out at this point.

The effect which small amounts of calcium have on the poly A mediated assembly of TMV-like rods is dramatic. However, it also seems to have no influence over the rate of reaction. Upon further reflection this should not be surprising as calcium has been shown previously to be unable to bind to TMV-CP on its own under the solution conditions used in these experiments[20]. Furthermore the calcium doesn't seem to interact with the poly A as was seen in figure 3.2 and from previous experiments done on the interaction between poly A and calcium[10,11]. Therefore the effect which calcium has on the assembly must be due to interaction with a TMV-like rod, containing poly A as no other interaction could produce the observed effect.

There is very little known about the mechanism of poly A mediated TMV rod formation. It has been shown to be second order with respect to the disk concentration, unlike the native genome which is 1st order[13,14]. The native genome is first order because the rate depends on the interaction of the first disk with the hairpin loop in the OAS. However there is no such loop to initiate rod formation in poly A, so presumably there is an initial partially stable proto rod which is formed upon the interaction of the poly A with the first disk and then if a second disk comes into contact with that species during the lifetime of the intermediate complex then it is added and a stable proto rod which is capable of stable elongation is formed. If no second disk is encountered during the lifetime of the intermediate complex the reverse reaction would occur and the poly A and first disk would once again separate from each other. Calcium has been shown to be bound to TMV rods through a bridging interaction between TMV-CP and RNA[4]. Therefore a possible mechanism which would explain the observed behavior would be that the calcium was binding to the semi-stable poly A TMV disk intermediary proto rod and thereby stabilizing it. This would allow for the observed effect without altering the reaction kinetics. Only a very small amount of calcium ion would be required as it would only be

necessary for nucleation of the proto TMV-like rods and not elongation of them which occurs more easily and thus likely wouldn't require any help.

The fact that the calcium effect is observed is further evidence that the poly A is indeed incorporated into the TMV-like rod through interaction with the RNA binding track, and the observed behavior is not due to some other interaction between poly A and TMV-CP. As previously mentioned, the observed calcium effect could only be achieved through the calcium binding to the TMV-CP as monovalent ions are not able to produce the observed effect. Furthermore, initial experiments suggest that other divalent ions besides calcium are also not able to produce the same effect as calcium. Since calcium is known to only be able to bind to RNA containing TMV rods at the solution conditions under which the experiments were done, the poly A must have been previously bound to the RNA binding track of the TMV-CP prior to interaction with the calcium. The poly A would not necessarily have to be strongly bound to the CP. A weak interaction would likely bring the CP and poly A sufficiently close to each other to be strengthened by the interaction with a calcium ion.

Although the solution conditions used in figure 3.5 are identical to those which are used to capture the CD spectra used in figure 3.4, the reaction in figure 3.5 is substantially slower, taking about 4 h to complete. Since the reaction in figure 3.5 is slightly slower it is possible to view the end of the initiation region which is the reason why the shape of the curve in figure 3.5 is different from that seen in figure 3.4. In theory, the rate of reaction between poly A and TMV-CP should depend on the square of the concentration of TMV-CP disks, and should vary linearly with the concentration of poly A[13,14]. However, if this is true, the differences in rate of reaction between figures 3.4 and 3.5 should be substantially different, as they have both different poly A and TMV-CP concentrations. As measured the differences in rate are seen to be more closely related to the difference in poly A concentration than they are to the changes in TMV-CP concentration. TMV kinetic studies have found that when TMV rods are made from TMV-RNA, once the concentration of disk gets high enough, the increase in rate of formation of TMV rod due to an increase in disk concentration reaches a plateau[14]. In both figures 3.4 and 3.5 the concentration of CP is higher than that at which a plateau was seen. Although poly A

mediated TMV-like rod formation may not show the same behavior, logically at some point increasing the concentration of disks will not increase the number of collisions with growing TMV-like rods which are ready to add another disk. Therefore from figures 3.4 and 3.5 it would seem that a similar maximum disk threshold to the one which has been seen for TMV rods exists for poly A mediated TMV-like rods however a more thorough investigation of different conditions would be needed to be conclusive.

3.3 Possible causes of degradation of TMV-CP Over Time

The fact that the same cells could produce TMV-CP which exhibits such varying behavior in its interaction with poly A was initially puzzling. The CP is also observed to lose its ability to interact with poly A over time but not lose its self assembly behavior in the absence of poly A. The RNA binding region of TMV is somewhat special in that it is unusually flexible for a protein at room temperature[21]. This region is also not responsible for the self assembly properties of TMV-CP in the absence of RNA as those are governed mostly by the hydrophobic outer circle and the Caspar pair carboxylic acids. Therefore it was reasoned that this unusual region of the TMV-CP is the most likely cause of the production of CP which is not able to form TMV-like rods. There are two main possible culprits, both of which would explain the observed behavior of the TMV-CP. The most obvious would be if something in solution is binding to the high energy, inner loop of the CP over longer periods of time, with higher production temperature significantly accelerating the proposed binding reaction. The second possibility is that over time the unstable inner loop was adopting an alternate, semi-stable configuration which is incompatible with low energy, poly A assisted change from planar to helical configuration.

Either of these possibilities could potentially explain the observed behavior. Furthermore, at least 34 CPs are involved in the nucleating step of TMV rod formation, and in the case of poly A this number is increased to 68 CPs. In order for the interaction to be successful and produce a TMV-like proto rod the poly A must snake through all of the CPs and then cause a change in helicity in all the CPs simultaneously. In principle it would only take a single non-functional CP which would block the poly A's progress in each TMV-CP disk to be able to completely halt all nucleation of proto TMV-like rods by essentially causing an "RNA

blockade" in every TMV disk in solution. Therefore the majority of the overall CP could be unperturbed but even a reasonable fraction of corrupted CP would prevent all assembly of TMV-like rods. Without nucleation, no elongation could occur, and no rod formation would be observed. Mass spectrometry was used to try and distinguish between the two possibilities with anything bound to the TMV-CP expected to be visible as an increase in mass of a visible population of TMV-CP.

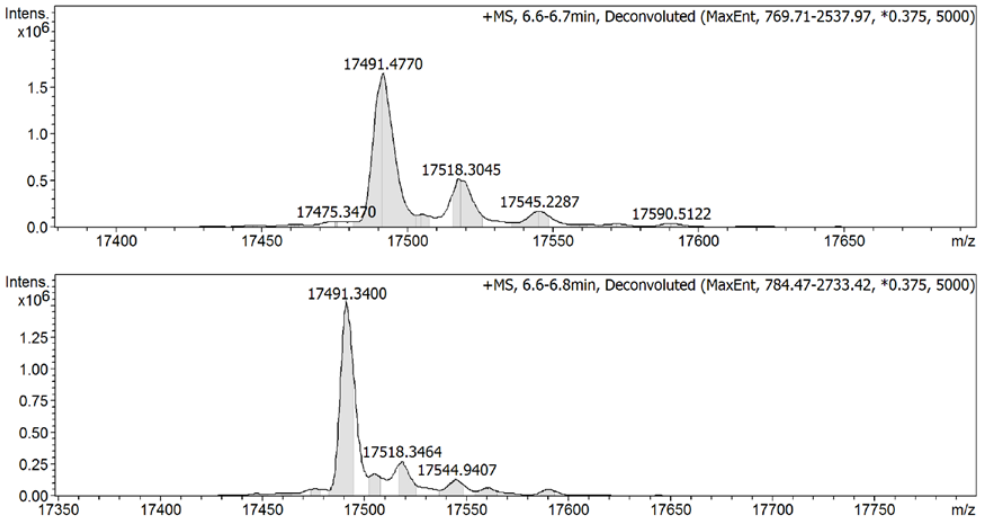


Figure 3.6, Mass spectrum of recently purified (5 days after purification) WT-TMV (Top) and 3 month old WT-TMV (Bottom). In both cases the obtained mass spectrum is almost identical with both samples containing all the same peaks.

It is seen in figure 3.6 that both aged and new TMV-CP not only have the same mass, but also have almost identical mass distribution. A binding event in the aged protein would result in a new peak in the mass spectrum. Figure 3.6 therefore indicates binding is unlikely since anything that binds so weakly that it is not detectible in the mass spectrum would likely also be able to be displaced by the poly A.

If the instability of protein conformation of the inner loop is indeed the cause of TMV-CP corruption then likely the lifetime of the produced TMV-CP could likely be increased by storing the CP under acidic conditions. Although this would not fully order the inner loop the way that RNA does, the inner loop has been shown to be more ordered under acidic conditions

compared to neutral conditions. Therefore that would likely at least increase the amount of time for which the purified CP could be used to help assemble poly A mediated TMV-like rods.

3.4 Improving imaging techniques

In addition to making the self assembly of poly A containing TMV-like rods a consistent process, improvements were also needed on the direct imaging techniques which are employed in this research. Although plating a multitude of samples would eventually result in some conclusive images, logically there must be a better way to achieve consistency of images of the formed poly A containing TMV-like rods.

As was seen in figure 3.7, while the addition of small amounts of calcium aid in assembling the TMV-like rods, the imaging of those rods cannot be reliably and reproducibly carried out directly from the assembly solution. Disks can be seen in the TEM and some rods are present, but as in previous TEM images taken in the absence of calcium ion, the resolution of the rods is poor and visualizing individual rods is very difficult. Essentially everything present in solution, TMV-like rods, TMV-CP disks and smaller aggregates of TMV-CP are all networked loosely together, and as such interfere with the imaging of the desired rods. This effect is even more apparent in the AFM from which it is only possible to conclude that something with the diameter of TMV rods is present along with a whole host of other, unidentifiable objects. The only thing which can be concluded from the AFM is that there are no obvious signs of free poly A. This is expected since the CD signal of the poly A helix is seen to go to zero as can be seen in Figure 3.1 which strongly indicates that the poly A is no longer present in solution.

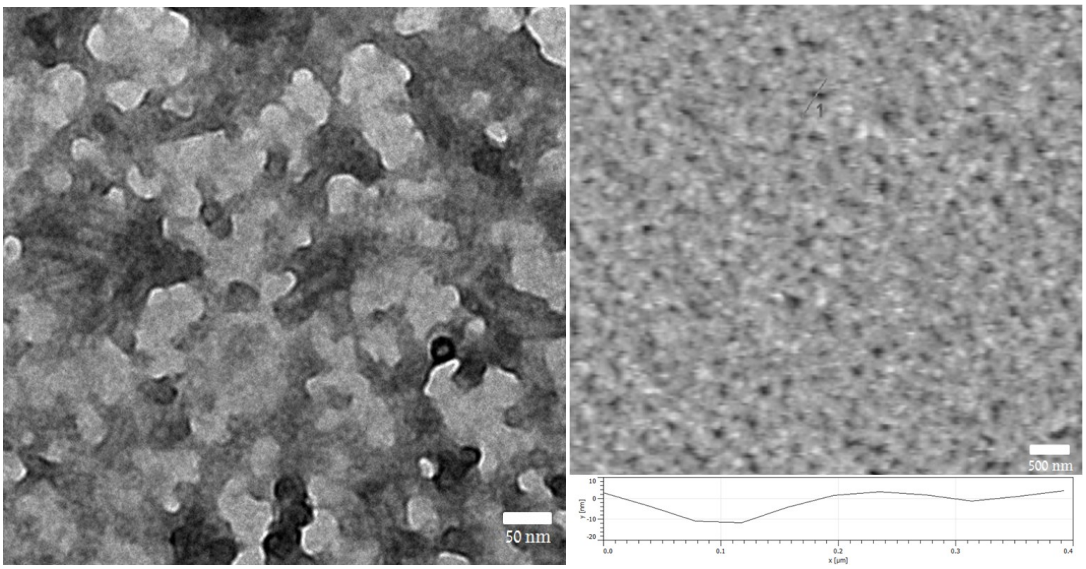


Figure 3.7, TEM (left) and AFM (right) of TMV-like rods formed in the presence of 0.25 mM calcium chloride. In the TEM (left) objects which resemble TMV-like rods can be seen to have their characteristic 18 nm diameter. In certain proposed rods the inner channel can be viewed but the rods are not clearly resolved. The AFM (right) of the same solution although there are many objects with the expected 16 nm height. The proposed rods are entangled in an extended network which not only prevents clear resolution of the rods but also in some cases prevents accurate heights from being easily measured.

3.4.1 Effect of dilution on imaging of TMV-like rods

In order to break up the loosely bound network of various forms of TMV and to also spread out the rods so that they can be individually imaged in AFM and TEM, the final solution containing the poly A mediated TMV-like rods is diluted by a factor of between 5 and 10 times in the exact buffer used for the self assembly reaction. As is seen in figures 3.8 and 3.9, not only does this have the effect of placing less material on the imaging substrate used, but it also seems to nullify the weak interactions between TMV-like rods and other forms of TMV-CP. As is seen in figure 3.8, the dilution allows for visualization of individual TVM-like rods. There are, as a consequence of the dilution factor, many less rods viewable per image than in non-diluted samples. However, multiple images can be combined to overcome this issue, and the fact that each rod is separate from other species means that clear and unambiguous assignment of formed TMV-like rods can be done.

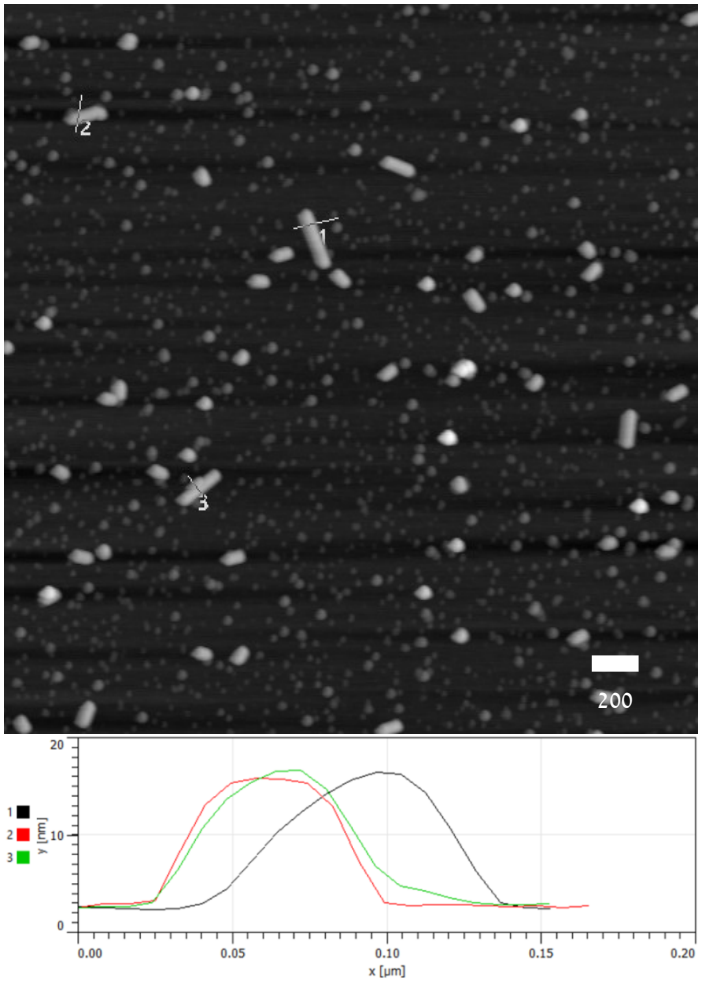


Figure 3.8, AFM of 8X diluted solution of poly A mediated TMV-like rods. The dilution has the effect of breaking up the previously imaged network into TMV-like rods which are individually resolved. The TMV-like rods can be identified from their 16 nm height and their shape. The TMV-like rods viewed here had lengths ranging between 70 and 250 nm.

Interestingly, as can be seen in figure 3.8 and 3.9, the rods imaged in TEM tend to be smaller in length than those seen in AFM images. This is likely due simply to a difference in plating technique. Smaller rods are more easily identifiable in TEM images compared to AFM images. Thus most of the perceived effect is likely due to the fact that TEM images can more easily obtain a higher resolution than AFM images and thus unambiguously image the smaller rods more easily. However, since the material plated in AFM must be attracted to the mica surface it is possible that the larger rods also are more easily bound to the mica surface.

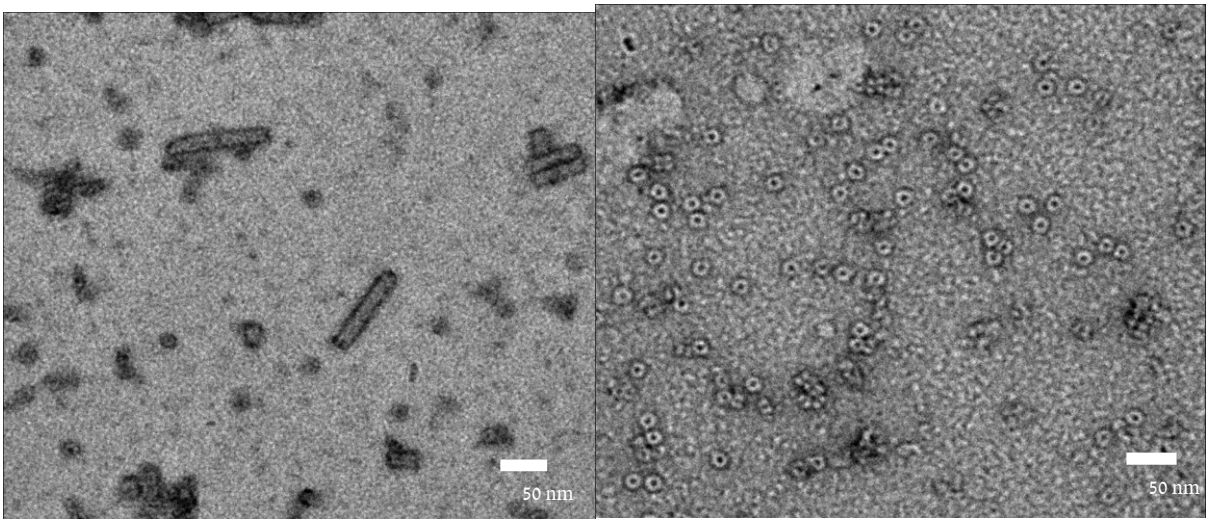
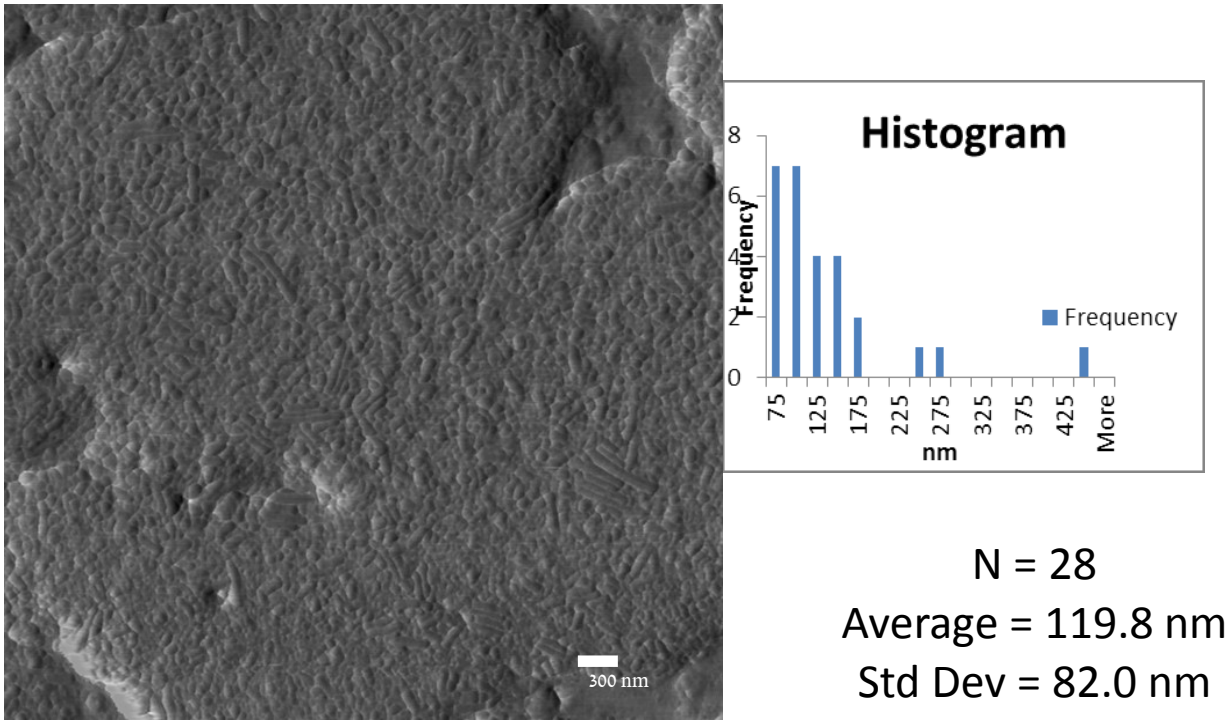


Figure 3.9, TEM of 8X diluted solution of poly A mediated TMV-like rods (left). The TMV-like rods can be identified by their 18 nm diamter as well as their inner channel. TEM of identically diluted TMV-CP at pH 6.8 (right). The dilution has the effect of breaking up any small TMV-CP rods or stacked disks into disks.

The dilution of the poly A mediated TMV-like rods results in substantially less material being trapped within the 3D volume which is collapsed to form the film used to create the TEM sample. In AFM sample preparation on freshly cleaved mica, the imaged items must be adsorbed to the surface. As such, the objects which are attached are dependent on the interaction between them and the solution. It was established previously that the TMV-like rods likely interact weakly with each other in solution and as such are deposited together (figure 2.11 and 2.12) which is the reason why they are difficult to resolve in both AFM and TEM. The fact that dilution allowed for the resolution of TMV-like rods in both AFM and TEM would therefore suggest that the weakly bound network of TMV-like rods and TMV-CP is broken up by the dilution of the sample, leaving only clearly viewable, individual TMV-like rods.

The rods viewed in figures 3.8 and 3.9 are significantly shorter than those from figures 2.14 and 2.16. Although the same poly A is used for both experiments, the poly A used to create the TMV-like rods seen in figures 3.8 and 3.9 has been significantly aged for those experiments while it was recently purchased for the earlier experiments. A certain amount of decay of the poly A strands in the form of breakup of larger poly A chains into multiple, shorter chains would therefore be expected considering the age of the poly A. This poly A degradation would be expected to result in a larger number of rods which are shorter and have a wider distribution of

rod sizes due to the larger distribution of poly A chain lengths. The other possibility is that the smaller rods are present in the earlier experiments but are simply not seen due to still being entangled elsewhere in solution and thus are not present in the clear AFM and TEM images.



Figures 3.10, AFM phase image of solution of poly A mediated TMV-like rods plated onto freshly cleaved mica without dilution. TMV-like rods ranging from 75 to 450 nm are observed with the average length being about 120 nm. The large observed size distribution is indicative that the poly A has likely been chopped from its original length.

AFM phase images are often used to more clearly view the boundary of objects, since phase images heavily emphasize edges. Plating of undiluted TMV-like rods almost always results in both AFM and TEM images being somewhat open to interpretation. However by simply preparing samples in bulk a clear image can sometimes be observed. Although dilution before plating is a consistently successful method, a clear non-diluted image can provide critical information. This is the case in figures 3.10, where the height data was unclear, but the phase image clearly shows the TMV-like rod boundaries. As can be seen, the TMV-like rods have roughly the same size distribution as those observed in figure 3.8. Although figures 3.10 does not rule out the possibility that smaller poly A chain lengths are also present in early experiments, it does suggest that the average poly A chain length decreases over time. There

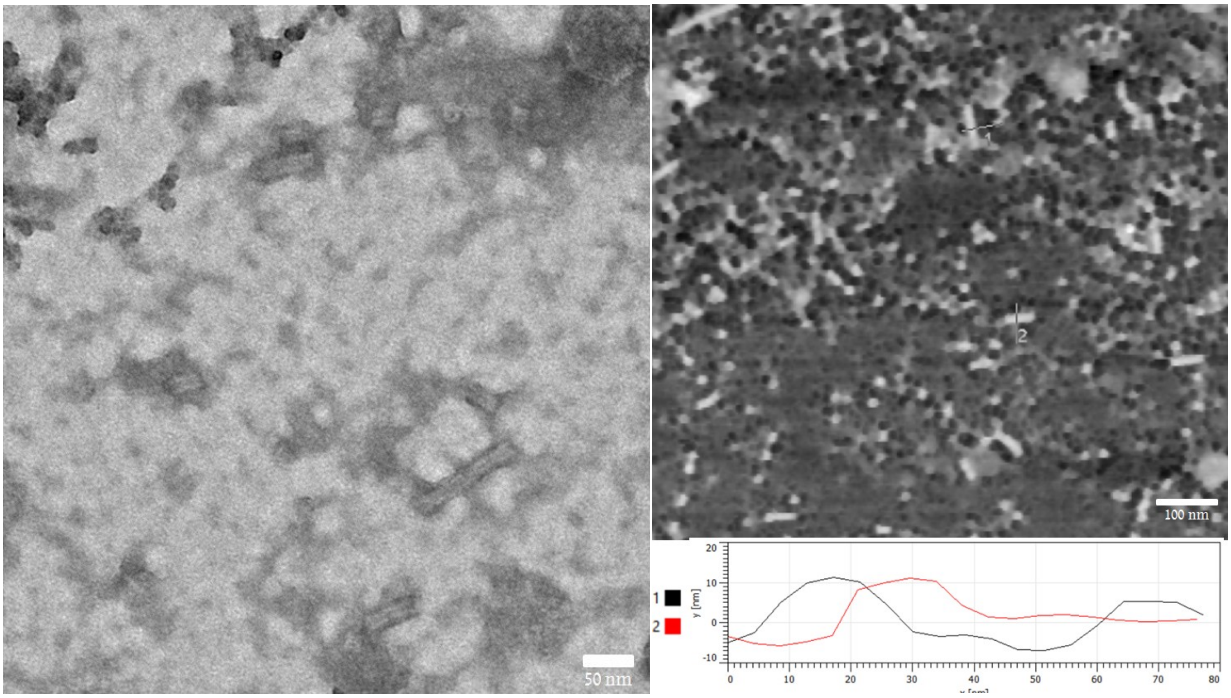
are also roughly 8 times more rods present in figures 3.10 compared to figure 3.8 which would suggest that the dilution fully breaks up weak networks formed initially in solution.

The average TMV-like rod length of 120 nm corresponds to a poly A chain containing 2600 bases. Although gel electrophoresis of the poly A was attempted both initially and again when the poly A was significantly older, the results were inconclusive. This is not totally unexpected as the strongly helical nature of poly A makes it notoriously difficult to determine its molecular weight using gel electrophoresis[22]. As such, the TMV-like rod length is likely a more accurate measure of the length of the poly A than gel electrophoresis, since TMV-CP rods of significant length simply are not able to form at the solution conditions used in this work and the length of rods formed in the presence of RNA are determined by the length of the RNA used[23–25].

As mentioned previously, the various states of the TMV-CP are in dynamic equilibrium with each other and therefore it is possible for all states of TMV-CP to exists, as a minority species, under any solution conditions. Normally the probability of viewing a given state of TMV-CP depends on the distance which one is from its expected location on the TMV-CP phase map. When the produced TMV-CP was first investigated for use in poly A experiments a strange phenomenon was discovered. After TMV-CP which had been purified using an anion exchange column was dialyzed directly from its 20 mM TEA (Triethanol Amine), pH 7.2 buffer directly into pH 6.8 phosphate buffer, the proportion of rods seen in the sample was unusually high. Normally at pH 6.8 a few small rods would be expected but the vast majority of the TMV-CP should be in the disk state. As can be seen in figures 3.11 and 3.12 this is not the case.

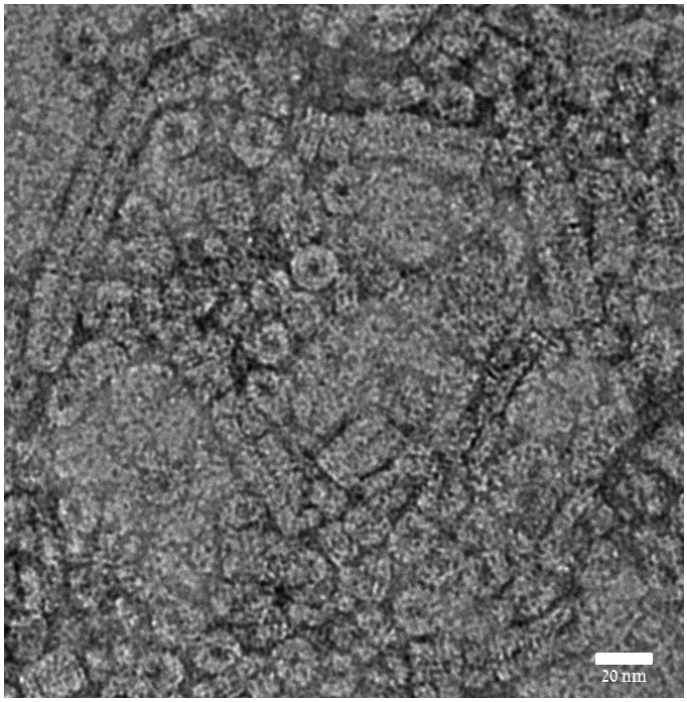
As can be seen in figures 3.11 and 3.12 many but not all of the TMV-CP rods which are observed contained nicks and imperfections which could potentially be used to distinguish them from poly A mediated TMV-like rods. Nevertheless the presence of these rods would significantly hinder the conclusive identification of poly A mediated TMV-like rods through direct visualization as distinguishing them from poly A mediated rods would likely prove to not be straight forward. Furthermore such rods would likely be too large to participate in poly A mediated TMV-like rod formation and would therefore make the effective concentration of

TMV-CP unknown. Due to the potential difficulties which such a starting solution of TMV-CP would create, a method for breaking up or at least reducing the size and amount of TMV-CP rods present at the assembly pH was needed.



Figures 3.11, TEM (left) and AFM (Right) of TMV-CP dialyzed directly after purification into pH 6.8 phosphate buffer. As can be seen both disks and rods are present in solution with the concentration of rods significantly above what would be expected for a TMV-CP under these conditions.

As was mentioned previously the number and size of the rods present in figures 3.11 and 3.12 would not normally be expected from TMV-CP under these solution conditions. It therefore seems likely that at some point during the purification of the TMV-CP the rods had formed and required a driving force to break up the unusual amount of rods present. The purification of the CP was done at 4^{o}C, a temperature which is not normally associated with the rod state of TMV-CP. However, during purification the TMV-CP would also likely be exposed to concentrations of CP which are simply impossible to achieve in solution which may explain why not only did the TMV-CP rods form during purification but also persisted in solution conditions where they would normally be expected to slowly break up.



Figures 3.12, Close-up of TEM image taken of TMV-CP which was directly dialyzed from 20 mM TEA buffer, directly after purification of the CP, into pH 6.8 phosphate buffer. Both TMV-CP disks and rods are present, with many of the rods having dislocations and knicks consistent with rods formed without RNA bound to it.

In order to "reset" the CP, for all experiments presented here the purified TMV-CP was dialyzed to pH 8.9, at 4^{o}C, prior to dialyzing into the assembly buffer at pH 6.8. This has the effect of breaking up all of the rods seen in figures 3.11 and 3.12 which therefore makes visual identification of poly A mediated TMV-like rods possible using TEM and AFM. A second method for breaking up the TMV-CP rods was later discovered. When TMV-CP which is dialyzed directly after purification into pH 6.8 buffer was exposed to the same 10X dilution conditions which were used to break up the loose network formed by the poly A mediated TMV-like rods the TMV-CP rods were no longer visible in the TEM as can be seen in figure 3.13.

Although interesting, the dilution method of TMV-CP rod breakup is not very practical as concentration of CP was more difficult than a simple dialysis. However, it does provide further evidence, not only that poly A mediated TMV-like rods could be reproducibly formed but also that the poly A was bound to the TMV-like rod and therefore stabilized it from breakup due to rapid dilution. The dilution therefore would breakup any TMV-CP rods which did not contain

poly A but not the poly A containing TMV-like rods. Evidently the TMV-CP purification process induces the CP into the rod state. While a simple dialysis to pH 6.8 does not provide enough driving force to break up the TMV-CP rods over a reasonable time frame, it seemed that both dialysis to pH 8.9 as well as extreme dilution did provide such a driving force.

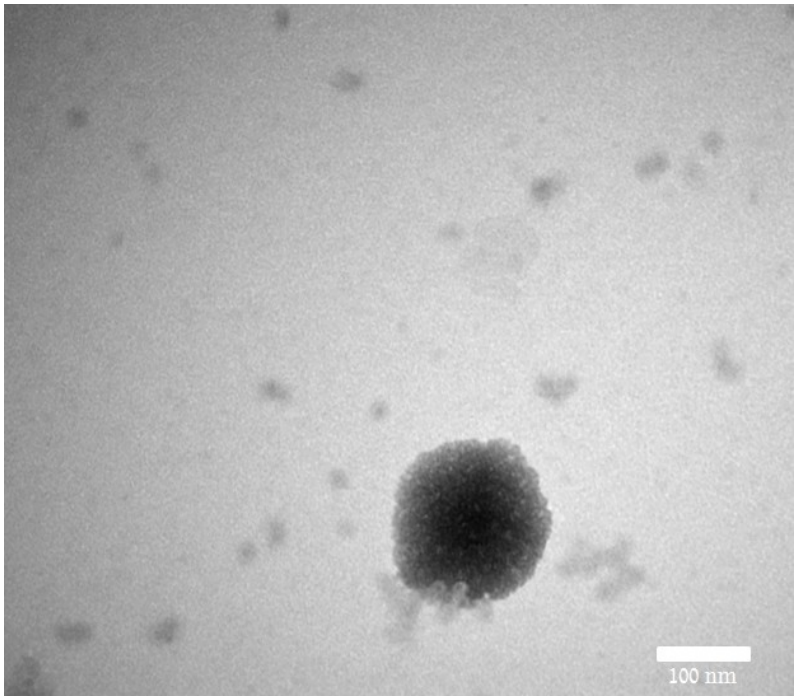


Figure 3.13, TEM of recently 10X diluted TMV-CP which was initially dialyzed directly after purification into pH 6.8 phosphate buffer. The rapid dilution breaks up the TMV-CP rods which are initially present after dialysis. Even considering drying and plating time after dilution the rods must break up within 20 - 25 minutes after initial dilution. The same pH 6.8 phosphate buffer is used for dilution.

In general dilution seems to disturb the equilibrium of the TMV-CP and drive it back towards the disk state. In poly A mediated TMV-like assembly reactions where the TMV-CP had been previously dialyzed to pH 8.9 prior to dialysis into the assembly buffer as can be seen in figure 3.14 the TMV-CP is almost exclusively made up of disks. However there are some small rods present. Most were smaller than 50 nm in length, however, a few closer to 100 nm in length could sometimes be seen. In order to initiate the assembly reaction between the poly A and the TMV-CP, the two solutions must be mixed. Since the poly A solution contains absolutely no TMV-CP it essentially dilutes the TMV-CP. These dilutions are kept to a minimum, usually consisting of a drop in concentration of stock TMV-CP of no more than 5%. The TMV-CP which is to be used as the control solution is exposed to the exact same dilution by introducing

assembly buffer to achieve the same dilution as the one done in the assembly reaction. The effect of this small dilution is to break up the small rods which had formed. It was seen in figure 3.14 that by the time the control solution is plated, usually about 4 hours after the slight dilution, the small amount of TMV-CP rods initially present in the stock solution is no longer present. Only disks and stacks of disks are present which is why the control images are unusually devoid of TMV-CP rods, a few of which would normally be expected at pH 6.8.

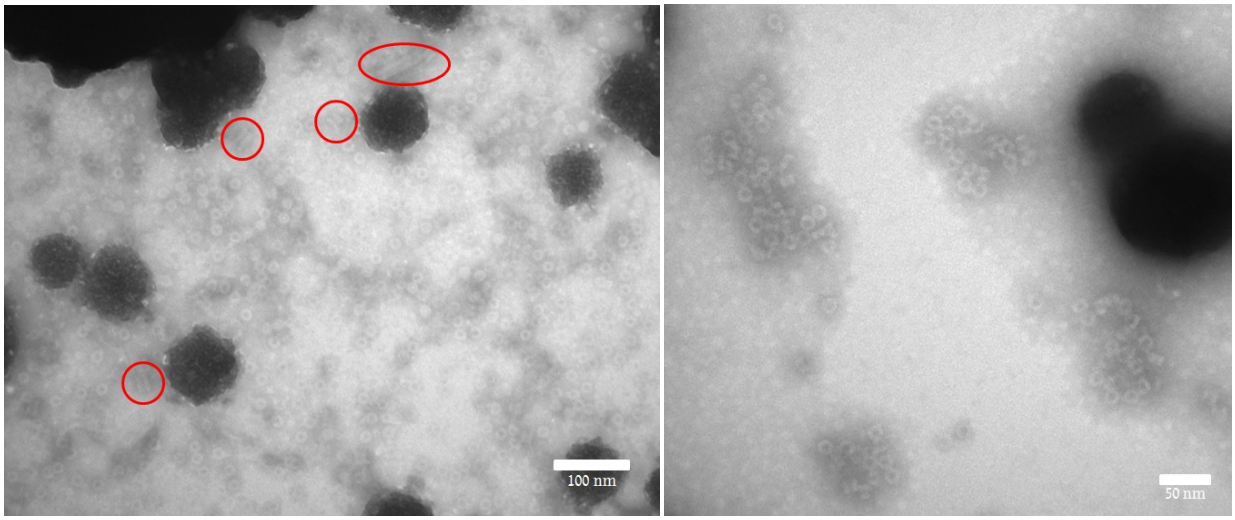


Figure 3.14, TEM of TMV-CP before (left) and after (right) a 5% dilution in the concentration of CP. The TMV-CP rods have been circled in red in order to make them clearer. Initially there are a small amount of rods which are mostly under 50 nm in length, however about 4h after the slight dilution there are no obvious rods of significant size present in the resulting TEM image.

The slight dilution seems to be sufficient to break up what little rods are present and makes identification of poly A mediated TMV-like rods even clearer than was hoped. Although there are likely still stacked disks present after the slight dilution there are no rods seen of length greater than 30 nm. It is not clear whether or not the breakup of these rods occurs relatively quickly or not, as the time dependence of the rod breakup was not investigated. However, a simple 5% dilution is able to break up rods over several hours, thus it should not be surprising that the 10X dilution from figure 3.13 would be able to break up the TMV-CP rods in minutes. While poly A containing rods would be more stable and thus protected from even an extreme dilution, TMV-CP rods have no such protection and seem to be relatively easily dismantled using dilution.

References

(1) Shalaby, R. A. F.; Banerjee, K.; Lauffer, M. A. Ion Binding by Tobacco Mosaic Virus and Its Protein. *Biochemistry (Mosc.)* **1968**, *7* (3), 955–960.

(2) Gallagher, W. H.; Lauffer, M. A. Calcium Ion Binding by Tobacco Mosaic Virus. *J. Mol. Biol.* **1983**, *170* (4), 905–919.

(3) Durham, A. C. H.; Hendry, D. A. Cation Binding by Tobacco Mosaic Virus. *Virology* **1977**, *77* (2), 510–519.

(4) Namba, K.; Pattanayek, R.; Stubbs, G. Visualization of Protein-Nucleic Acid Interactions in a Virus: Refined Structure of Intact Tobacco Mosaic Virus at 2.9 Å Resolution by X-Ray Fiber Diffraction. *J. Mol. Biol.* **1989**, *208* (2), 307–325.

(5) Pattanayek, R.; Elrod, M.; Stubbs, G. Characterization of a Putative Calcium-Binding Site in Tobacco Mosaic Virus. *Proteins Struct. Funct. Bioinforma.* **1992**, *12* (2), 128–132.

(6) Adiarte, A. L.; Lauffer, M. A. The Donnan Effect in Tobacco Mosaic Virus and Its Components. *Arch. Biochem. Biophys.* **1973**, *158* (1), 75–83.

(7) Loring, H. S.; Fujimoto, Y.; Tu, A. T. Tobacco Mosaic virus—A Calcium-Magnesium Coordination Complex. *Virology* **1962**, *16* (1), 30–40.

(8) Brakke, M. K.; Van Pelt, N. Influence of Bentonite, Magnesium, and Polyamines on Degradation and Aggregation of Tobacco Mosaic Virus. *Virology* **1969**, *39* (3), 516–533.

(9) Ge, P.; Zhou, Z. H. Hydrogen-Bonding Networks and RNA Bases Revealed by Cryo Electron Microscopy Suggest a Triggering Mechanism for Calcium Switches. *Proc. Natl. Acad. Sci.* **2011**, *108* (23), 9637–9642.

(10) Kankia, B. I. Binding of Mg2+ to Single-Stranded Polynucleotides: Hydration and Optical Studies. *Biophys. Chem.* **2003**, *104* (3), 643–654.

(11) Manzini, G.; Xodo, L. E.; Fogolari, F.; Quadrifoglio, F. Secondary Structure Effects on the Interaction of Different Polynucleotides with Ca2+. *Biopolymers* **1990**, *30* (3-4), 325–333.

(12) Powell, C. A. The Effect of Cations on the Alkaline Dissociation of Tobacco Mosaic Virus. *Virology* **1975**, *64* (1), 75–85.

(13) SCHÖN, A.; MUNDRY, K.-W. Coordinated Two-Disk Nucleation, Growth and Properties, of Virus-like Particles Assembled from Tobacco-Mosaic-Virus Capsid Protein with Poly (A) or Oligo (A) of Different Length. *Eur. J. Biochem.* **1984**, *140* (1), 119–127.

(14) Butler, P. J. G. Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: VI. Assembly of the Nucleoprotein Rods of Tobacco Mosaic Virus from the Protein Disks and RNA. *J. Mol. Biol.* **1972**, *72* (1), 25–35.

(15) Butler, P. J. G.; Klug, A. Assembly of the Particle of Tobacco Mosaic Virus from RNA and Disks of Protein. *Nature* **1971**, *229* (2), 47–50.

(16) Otsuki, Y.; Takebe, I.; Ohno, T.; Fukuda, M.; Okada, Y. Reconstitution of Tobacco Mosaic Virus Rods Occurs Bidirectionally from an Internal Initiation Region: Demonstration by Electron Microscopic Serology. *Proc. Natl. Acad. Sci.* **1977**, *74* (5), 1913–1917.

(17) Fukuda, M.; Okada, Y. Elongation in the Major Direction of Tobacco Mosaic Virus Assembly. *Proc. Natl. Acad. Sci.* **1985**, *82* (11), 3631–3634.

(18) Schuster, T. M.; Scheele, R. B.; Adams, M. L.; Shire, S. J.; Steckert, J. J.; Potschka, M. Studies on the Mechanism of Assembly of Tobacco Mosaic Virus. *Biophys. J.* **1980**, *32* (1), 313–329.

(19) Kraft, D. J.; Kegel, W. K.; van der Schoot, P. A Kinetic Zipper Model and the Assembly of Tobacco Mosaic Virus. *Biophys. J.* **2012**, *102* (12), 2845–2855.

(20) Gallagher, W. H.; Lauffer, M. A. Calcium Ion Binding by Isolated Tobacco Mosaic Virus Coat Protein. *J. Mol. Biol.* **1983**, *170* (4), 921–929.

(21) Jardetzky, O.; Akasaka, K.; Vogel, D.; Morris, S.; Holmes, K. C. Unusual Segmental Flexibility in a Region of Tobacco Mosaic Virus Coat Protein. *Nature* **1978**, *273* (5663), 564–566.
(22) MacLeod, M. C. Uncertainty in the Determination of the Molecular Weight of Poly (A)-Containing RNA. *Anal. Biochem.* **1975**, *68* (1), 299–310.
(23) Rego, J. M.; Lee, J.-H.; Lee, D. H.; Yi, H. Biologically Inspired Strategy for Programmed Assembly of Viral Building Blocks with Controlled Dimensions. *Biotechnol. J.* **2013**, *8* (2), 237–246.
(24) Turner, D. R.; McGuigan, C. J.; Butler, P. J. G. Assembly of Hybrid RNAs with Tobacco Mosaic Virus Coat Protein: Evidence for Incorporation of Disks in 5'-Elongation along the Major RNA Tail. *J. Mol. Biol.* **1989**, *209* (3), 407–422.
(25) Hwang, D.-J.; Roberts, I. M.; Wilson, T. M. Expression of Tobacco Mosaic Virus Coat Protein and Assembly of Pseudovirus Particles in Escherichia Coli. *Proc. Natl. Acad. Sci.* **1994**, *91* (19), 9067–9071.

Chapter 4 Development of Strategies for Eventually Obtaining Surface bound Poly A Containing TMV-Like Rods

It was shown in the previous chapters that poly A mediated TMV-like rods could be produced reproducibly. Although further study is required before Poly A can be used to assemble a more complex TMV-like rod containing multiple domains, it was undoubtedly shown to be a promising alternative template to TMV-RNA based TMV rod technology. As discussed previously, poly A would be more likely than TMV-RNA to be able to handle the additional constraints which would be placed onto the assembly due to the TMV-CP modifications and additions. While that would be sufficient for TMV based applications which remain in solution, many potential TMV applications must be incorporated into larger, macroscopic devices to allow for human manipulation. Such incorporation requires a handle between the very small nano scale TMV-like rod assembly and the macroscopic scale of human hands. One possible solution to this issue is to have the TMV-like rod tethered to a solid surface in order to gain control over its location. Tethering of a poly A containing TMV-like rod to a gold surface has the added benefit of tethering the TMV-like rod to a functional surface and thus introduces the possibility of incorporating TMV based components into microelectronic circuits and other such devices.

Thus, this chapter reports on methods suitable for tethering poly A to a gold surface. In the native genome, different portions of the TMV-RNA have different affinities for the TMV-CP. Therefore, one end of the TMV-RNA can be partially exposed while keeping the rest of the TMV rod assembled. This exposed end of the TMV-RNA can then be used to bind that RNA and therefore the assembled rod to a surface[1]. Such a strategy would not be possible with pure poly A, as every portion has the same affinity to the TMV-CP. Therefore, in order to create surface bound TMV-like rods, poly A must be bound to the surface before assembly takes place. TMV-like rod assembly requires that the poly A be free to interact with the TMV-CP. However, like all nucleic acids the poly A has a natural tendency to be attracted to and become immobilized by a gold surface. This would almost certainly prevent the poly A from being able to interact with the TMV-CP and thus no TMV-like rod formation would be possible. As such in order to produce

surface bound TMV-like rods containing poly A the poly A must not only be bound to the gold surface but also be bound such that only the tethered end is able to directly interact with the gold surface.

Several methods will be described which can potentially be used to construct a platform which would be able to both bind to the free poly A and then shield it from the gold surface, and also allow for the surface bound poly A to be incorporated into TMV-like rods. Initially, single stranded (ss) DNA strands were deposited onto a gold surface with the intention that they could be later ligated to poly A. The attraction of the non-specifically bound portion of the DNA strand to the gold surface is a well known problem, and as such, gold bound thiolated DNA has a tendency to "lie down" on the gold surface and therefore becoming solution inaccessible[2]. A common solution to this problem is to "backfill" the single stranded thiolated DNA with 6-Mercaptohexanol (MCH). Thiolated ssDNA contains a 6 carbon spacer between the nucleotides and the sulfur atom. As such, it is compatible with monolayers of MCH which theoretically should not displace ssDNA bound to gold through a thiol linkage. However since the gold -thiol bond is stronger than the non-specific interactions with the rest of the ssDNA chain, MCH preferentially binds to the gold and forces the ssDNA to be free to interact in solution[2–6].

The effect that MCH has on the AFM features of thiolated DNA bound to a gold surface was investigated. A 100 mer thiolated ssDNA strand with a random sequence was chosen. Since the DNA is only present to later be bound to the poly A and keep the poly A from interacting with the gold surface, its sequence is of little importance. Although the theoretical, fully stretched length of the ssDNA was over 30 nm, it is seen in figure 4.1 that the feature sizes are well below that length. Single stranded DNA is substantially less stiff than dsDNA, which is reflected in its short persistence length of about 2.5 nm. This is much less than that for dsDNA which is about 50 nm[7–9]. Therefore, the fact that no features above about 8 nm are seen in the AFM image is not surprising. Furthermore all AFM scans were done with dried, gold bound ssDNA samples rather than in solution. Although there was no basis for comparison, the drying

of the samples would most likely have the effect of reducing the feature sizes rather than increasing them.

Figure 4.1 shows that the gold bound ssDNA was seen as mountain shaped features. While the AFM in theory has the capacity to detect individual strands of ssDNA, it is impossible to know how many strands are represented by each mountainous feature in the AFM image. This is due to the fact that sufficiently closely bundled strands of ssDNA could also show up in an AFM image as a single feature. The current level of understanding of the possible conformations of thiolated ssDNA when bound to a gold surface is not detailed enough to be able to form any conclusions about the ssDNA surface density based on the AFM image shown in figure 4.1. A large variety of shapes are observed, with some features being taller and skinnier and others being shorter and wider. The majority of the features are between 3 and 3.5 nm in height. Given that many of the shortest features are also quite narrow, it is possible that the larger spots consist of more DNA strands then the smaller spots. An equally likely possibility is that the variation in the volume of the observed DNA spots is due to a both a variation in the exact conformation of the DNA upon drying convoluted with a variation in how individual DNA strands or groups of strands interact with the AFM tip.

AFM height images of mercaptohexanol (MCH) deposited onto a gold surface were also recorded. Unlike with the thiolated ssDNA, the MCH features in the AFM image are much more consistent. While it is also not possible from figure 4.2 to conclude the number of MCH molecules which make up a single AFM image feature, each feature is more or less identical. The feature height varies between about 1.2 to 1.6 nm, which is consistent with the expected length of an MCH molecule

The MCH features are more consistent than the thiolated ssDNA, which is consistent with the idea that due to its increased complexity and length compared to the MCH, the ssDNA strand can adopt a much larger variety of conformations which presumably leads to a larger variety of AFM height feature sizes.

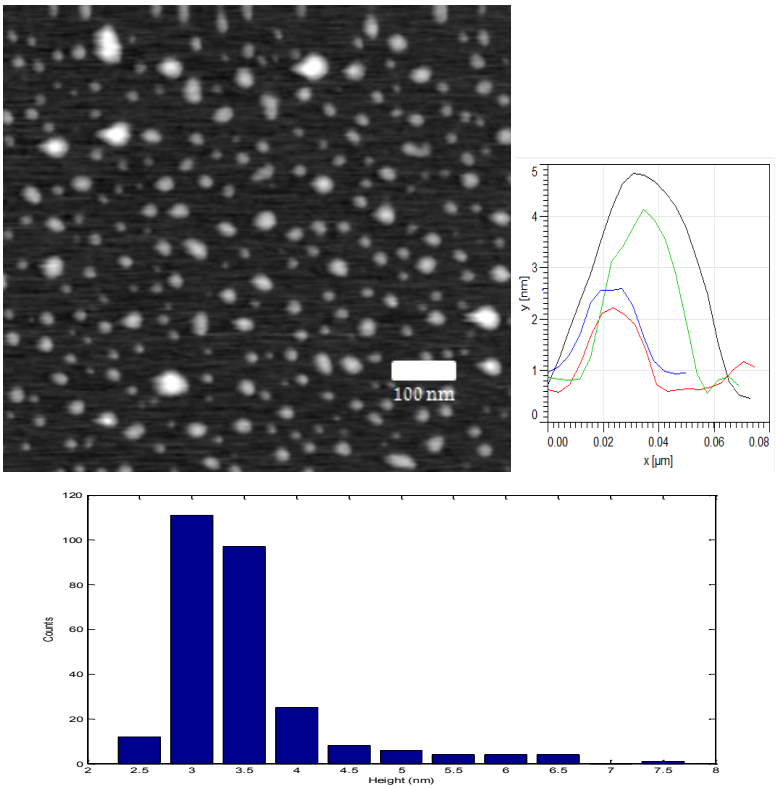


Figure 4.1, AFM height image of a random sequence, 100 base ssDNA on a gold surface. A large variety of feature heights and aspect ratios are observed with some of the objects quite tall while others are more spread out and are relatively short. The average height of the observed features is 3.5 nm.

Exposing thiolated ssDNA bound to a gold surface to mercaptohexanol has been shown to increase the hybridization efficiency of the bound DNA, presumably by preferentially orienting the ssDNA away from the gold surface, making it available for hybridization to its complementary strand[2–6]. Using AFM it was investigated whether the addition of MCH would have any effect on the feature height and shape seen in the resulting AFM. As a comparison, the reverse order was also investigated, with MCH being deposited first and the thiolated DNA second. Although all the AFM scans are done on dried samples and not on gold surfaces in solution, it is expected that any significant changes in the orientation of the ssDNA in solution due to the mercaptohexanol should also be visible after drying.

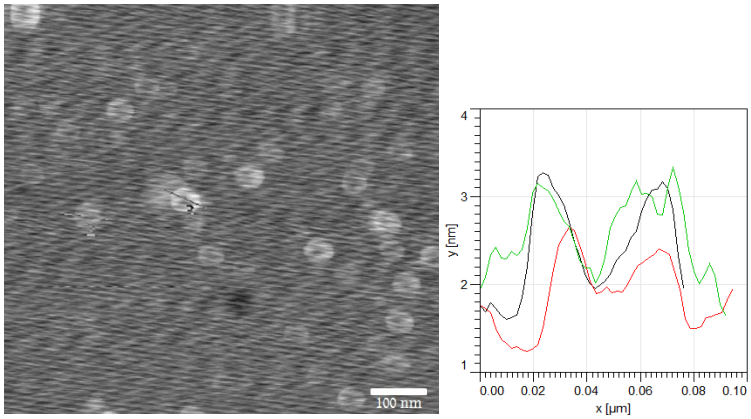


Figure 4.2, AFM of Mercaptohexanol (MCH) on gold surface. The surface coverage is not particularly complete since MCH only contains 6 carbons and as such has limited ability to form a self assembled monolayer (SAM).

Figure 4.3 shows that pre-treatment of the gold surface with mercaptohexanol results in a significantly lower density of features seen in the AFM height image. Furthermore, the heights of the recorded features are also drastically less than those observed in the ssDNA image from figure 4.1. Many of the feature heights are between 1.5 nm and 2nm. Since the MCH height is to be up to 1.6 nm, some of the features were likely due to MCH molecules. By pre-treating the gold with MCH, the thiolated ssDNA essentially had a reduced surface area to bind to as it would be unlikely to be able to displace MCH with great success. Therefore the majority of the bound MCH would be expected to still be bound to the gold and as such could account for the smaller features which are present. The reduced density of features could similarly be explained by the pre-treatment of MCH as it would be expected to significantly reduce the available gold binding area for thiolated ssDNA.

As can be seen in figure 4.3b, the features which are smaller than 1.5 nm tend to be rather broad and contain many small peaks in their traces. Those features are most likely due to mercaptohexanol, while the ssDNA peaks tends to have a single peak and only varies in size and width. As such, while it is not possible to make conclusive assignments to feature heights which are 1.5 nm, the shape of the trace could be used to make a reasonable guess as to the identity of that feature. No features are seen which are larger than 4 nm. This would suggest that the pre-treatment of the gold with MCH either limited the possible conformations of DNA which are able to bind to the gold, or forced the DNA to adopt the conformations leading to shorter AFM feature heights.

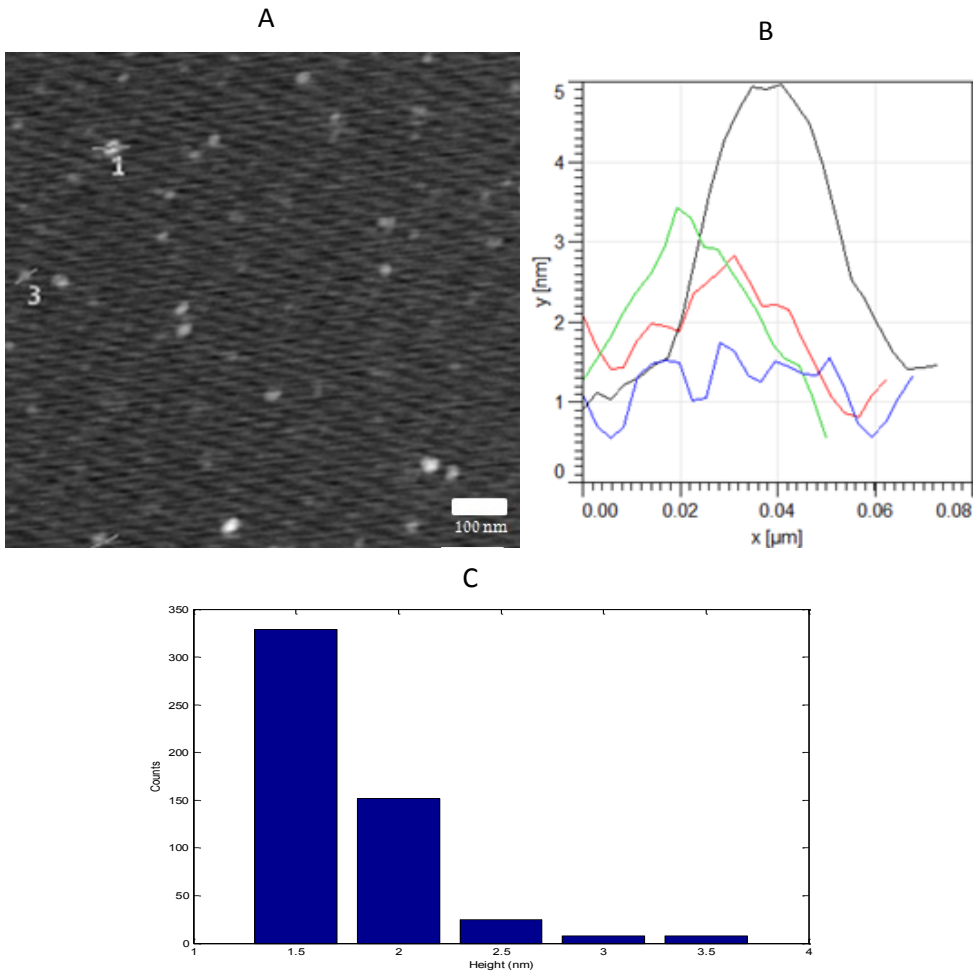


Figure 4.3, A) AFM height image of thiolated ssDNA deposited onto gold after the surface had been previously treated with MCH. The density of features was greatly reduced from what was seen in the AFM images with only ssDNA. B) Trace of selected features in the AFM image, many of the features could not be distinguished with certainty from an MCH molecule as many of the feature heights are less than 2 nm. C) Height distribution of features in the AFM scan. The majority of the features are under 2 nm and as such could easily have been MCH. Of the features which are 2 nm or more, which could be attributed to ssDNA, most of them are not large, with the largest feature seen to be 3.5 nm.

In figure 4.4 the density of features is not very high compared to that from the thiolated ssDNA without any MCH treatment seen in figure 4.1. However, post treatment with MCH results in a return of the taller AFM features, with not only feature heights of 6 and 7 nm being seen, but also corresponding to a much larger percentage of the overall features than is seen in figure 4.1. This would suggest that post treatment with MCH encourages the thiolated ssDNA to adopt more upright conformations. The largest height observed is around 7 nm compared to 7.5 nm which is observed for untreated ssDNA. However this does not necessarily mean that the untreated DNA has the largest AFM feature. In AFM, the measured height is a height

relative to the surrounding surface. When an object is measured relative to bare gold, the height can be attributed to the feature height. However when more than one species is present, it is possible that the baseline measurement is made from the second feature. Thus in figures 4.3 and 4.4, since MCH is present, the measured DNA height may in some cases have been reduced by being surrounded, but not on top of an island of MCH molecules which would have the effect of reducing the measured height of the DNA in the AFM image. From the traces seen in figure 4.4b it is seen that the height to width ratio of the AFM features does not have very much variety compared to those observed in figure 4.1.

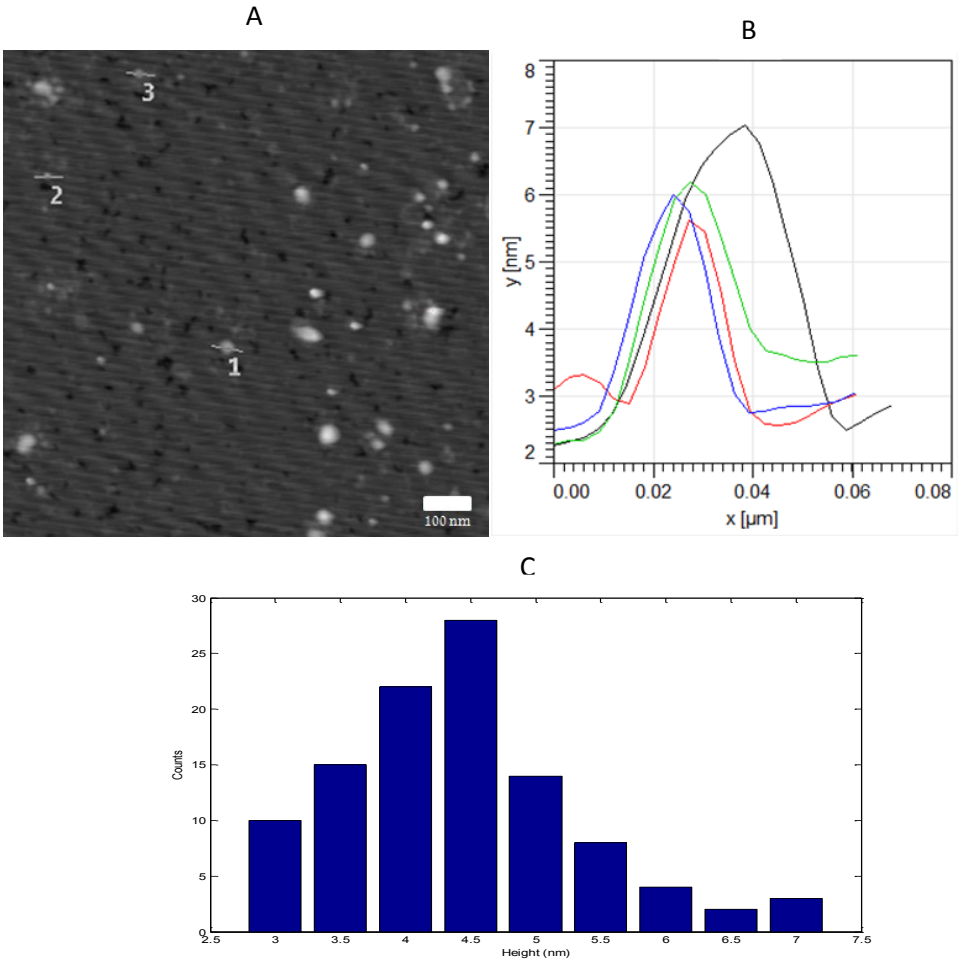


Figure 4.4, A) AFM height image of thiolated ssDNA deposited on a gold surface and then treated or "backfilled" with MCH. The density of AFM features is not particularly high compared to figure 4.1. B) Traces of selected AFM features C) size distribution of AFM features recorded in AFM scan

Backfilling a gold surface which has been bound to thiolated ssDNA with MCH has been shown previously to dramatically increase the hybridization efficiency of the bound strand when exposed to its complementary DNA strand[2,3]. In figure 4.4, backfilling such a surface with MCH results in less AFM features overall which have a much more consistent height to width ratio and also tend to be taller. The increased hybridization efficiency was attributed to an MCH induced change in configuration of the bound ssDNA to a more solution accessible configuration. Therefore ssDNA features in dried AFM which have large heights and a large height to width ratio can be correlated to ssDNA configurations which are free in solution to hybridize with high efficiency, and thus would be likely to also be free to bind to poly A.

The thiolated ssDNA results are promising for providing a platform which would be able to assemble TMV-like rods with gold bound poly A. Also RNA bound to similarly prepared surfaces has been shown to be able to adopt complex shapes[10,11]. However, in order to assemble TMV-like rods with surface bound poly A, not only must the chosen platform be able to shield the poly A from the gold surface but it must also be able to shield the TMV-CP from the gold surface. Furthermore, since not only the poly A but also the TMV-CP is negatively charged at assembly pH, it would be expected that the assembly of such TMV-like rods would require a significant build up of charge density in order to form near a surface. As such even though backfilling with MCH provided sufficient shielding for hybridization of DNA it would likely prove highly challenging for it to provide sufficient shielding to allow for surface bound poly A to help assemble TMV-like rods.

Double stranded DNA nanostructures are significantly more rigid than ssDNA[12]. Furthermore they have been shown to be able to achieve a substantially higher surface coverage on a gold surface compared to thiolated ssDNA[13]. Therefore double stranded DNA nanostructures are more likely to be able to provide a platform capable of assembling surface bound poly A containing TMV-like rods as not only would they allow the assembly to be carried out further from the gold surface, but they would also provide excellent shielding of the TMV-CP from the gold surface as they are negatively charged.

The first double stranded (ds) DNA nanostructure investigated was a tetrahedron shaped structure. Such a structure has been shown to preferentially orient the face containing 3 thiol groups towards the gold surface[13]. The tetrahedron is constructed from 4 ssDNA strands, with 1 being longer than the rest, thereby providing a single stranded portion of the structure which is preferentially oriented away from the gold surface. In principle, that single stranded portion could be used to either hybridize or ligate poly A, and thus provide an anchor which would be expected to strongly attach poly A to the gold surface with its triple thiol-gold bonds while also keeping the poly A solution accessible by raising it and shielding it from the gold surface. The DNA tetrahedron developed by Pei et al. shows promise for use in surface assisted poly A mediated TMV-like rod assembly as a very high surface coverage is possible in that study and it was found that the single stranded portion is able to hybridize efficiently even in the presence of 50% blood serum[13].

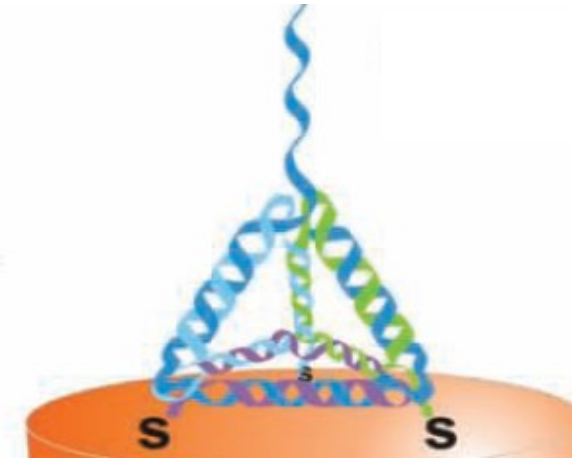


Figure 4.5, Adapted from pei et al.[13], Schematic of dsDNA tetrahedron on a gold surface. Single stranded portion is pointed away from the gold surface. Reprinted from *Adv. Mater.* 22 (42), 4754–4758, Pei, H.; Lu, N.; Wen, Y.; Song, S.; Liu, Y.; Yan, H.; Fan, C., A DNA Nanostructure-Based Biomolecular Probe Carrier Platform for Electrochemical Biosensing, 2010, used with permission from John Wiley and Sons

The same tetrahedron as the one used in the study done by Pei et al. is constructed using the method outlined by Pei et al.[13]. The dsDNA assembly is analyzed using a 9% Acrylamide gel. Pei et al. claimed to assemble the tetrahedron with an 85% yield, however

based on the gels done here using the same assembly procedures, an example of which is presented in figure 4.6, a yield of 85% is likely misleading. While the quality of the gel done is purely qualitative and not rigorous, it can clearly be seen that the majority of the DNA is not able to penetrate into the gel and is stuck in the wells. The bands corresponding to the ssDNA which is used to assemble the tetrahedron are placed in a blue box in figure 4.6 and are seen to run significantly faster than the assembled structure with the longer ssDNA running slightly slower than the other DNA strands. Heavier ssDNA would be expected to run more slowly as its increased size makes travel through the gel more difficult. The highly structured form of 3D DNA nanostructures dramatically decreases the ability of the 3D structures to travel through the gel, beyond what would be expected of ssDNA of equivalent mass. Therefore, although it is not confirmed, the bands in figure 4.6 which are highlighted in red are most likely the properly assembled dsDNA tetrahedron. Their position relative to that of the ssDNA is similar to what was seen in the gels done by Pei et al. [13].

Although further study is required to rule out the presence of properly assembled DNA tetrahedron in the non-penetrating material, given the distance which the tetrahedron band was able to travel through the gel, it seems unlikely that a significant amount of properly assembled tetrahedron would be stuck in the non-penetrating band. Since objects of similar size and bulk to the tetrahedron should be able to penetrate the gel, the non-penetrating material likely consists of larger DNA assemblies. Three dimensional dsDNA have additional difficulty traveling through gels due to their more complex structure. Therefore the non-penetrating DNA would likely only have to be 2-3 times larger than the tetrahedron to become unable to penetrate the gel. The gels presented by Pei et al. did not include the wells. Production of non-penetrating material is a common event in three dimensional dsDNA nanostructure assembly. Given that the exact same sequences and procedure are used here, and Pei et al. did not present their gels in their entirety, it seems more likely that they produced similar non-penetrating products which were simply not considered relevant. Even if Pei et al. managed a yield of tetrahedron 3X higher that that presented in figure 4.6, it would still not be near an 85% yield, with respect to the starting strands of ssDNA. It therefore seems unlikely that an 85% yield, with respect to the percentage of starting strands which are converted to

properly assembled tetrahedron, could be achieved using this method. Although Pei et al. do not specify how they calculated their yield, they likely used the intensity of the light seen in the tetrahedron gel band relative to the rest of the band. Since their wells are cut off, and the non-penetrating DNA is not considered, it would be possible to achieve an apparent 85% yield using this method of calculation. Without further clarification this yield would be deceptive as one would assume that 85% of the starting strands were converted to properly assembled tetrahedron. However based on figure 4.6, the percentage of starting strands which are converted into properly formed tetrahedron is very low, and therefore indicates that the method used to assemble the tetrahedron is unlikely to be capable of achieving a true 85% yield.

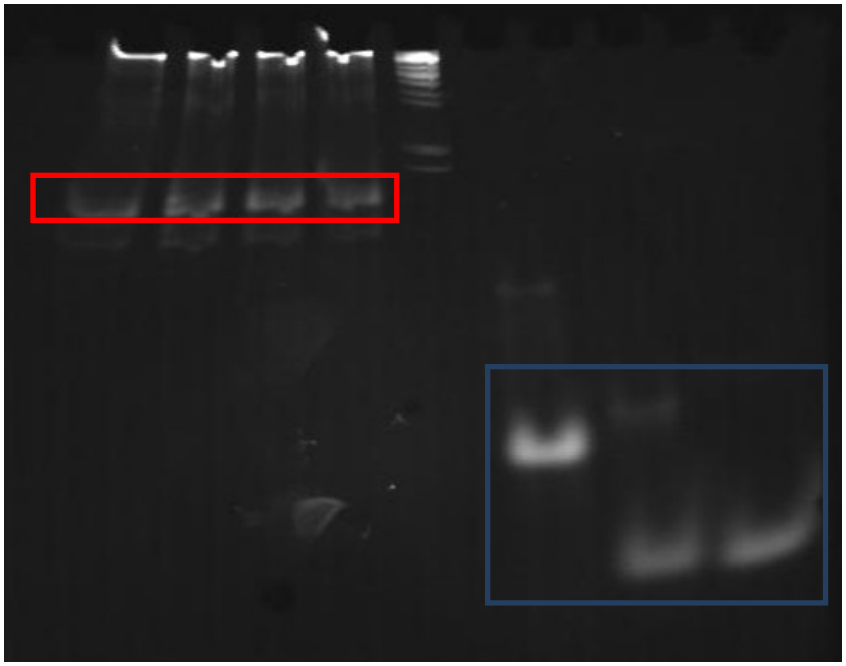


Figure 4.6, 9% Acrylamide gel stained with . Wells from left to right were: Assembled Tetrahedron, Assembled Tetrahedron, Assembled Tetrahedron, Assembled Tetrahedron, large DNA ladder, Strand A, Strand B, Strand C. The mobility of the single stranded DNA is significantly greater than the assembled products. The ssDNA bands have been highlighted in blue, Strand A, which is longer than the rest is less mobile than Strands B and C since it is larger. All the wells which contain assembled DNA tetrahedron contain a large amount of material which does not penetrate the gel. The band from the properly assembled dsDNA tetrahedron is highlighted with a red rectangle.

In order to assemble a dsDNA tetrahedron, the structure forces dsDNA to adopt bending angles of 70.5^{o}, a highly unfavorable angle for dsDNA. The assembly reaction is a 1 pot reaction containing all 4 strands. Therefore during the assembly reaction a partially completed

tetrahedron, which for example, is missing only strand B could hybridize with a free strand B, or it could also hybridize with another partially assembled tetrahedron which has a partially un-hybridized B strand. The second option would result in extended assemblies, as some of the other strands could similarly hybridize to other partially completed tetrahedrons. These extended assemblies would likely be tetrahedron-like as parts of them would be properly assembled but other parts would be bridged to another tetrahedron-like assembly. They could be extremely large, since there is no theoretical limit to their size, but given the low concentration of starting strands used, they would more likely be kinetically limited to 5 or 6 equivalents of tetrahedron by weight. A properly formed ds DNA Tetrahedron should contain a few hundred bases, which although it would go slowly should penetrate a 9% Acrylamide gel as is seen in figure 4.6. However the extended assemblies would very quickly grow to a thousand or more bases with an extended and complex geometry which would likely exponentially decrease the gel permeability of such structures. These extended assemblies are the most likely explanation for the identity of the non-penetrating material in the gel well seen in figure 4.6.

Kinetically, the tetrahedron should be the preferred product of the assembly, since the strands are at low concentration. However, the 70.5^{o} angle is an unfavorable one for the type and sequence of DNA used which likely makes the extended assemblies the thermodynamically favored products, as some of the angle strain could be relieved using that reaction pathway. The sharp angle is likely the reason why the majority of the product observed in the gel is extended tetrahedron-like assembly rather than properly formed dsDNA tetrahedron. The corners of dsDNA nanostructures are known to be points of strain and are the main cause of improperly assembled DNA nanostructures. When the technique was first developed, it was common practice to cut the wells of the gels to "ignore" the poorly assembled non-penetrating DNA. However, since then, strategies have been developed to significantly reduce or completely remove the angle strain in the DNA nanostructures by introducing organic molecules in place of DNA at the vertices. These organic molecules can either be highly flexible or are chosen because they already contain the desired angle of the DNA nanostructure. Alternatively 4 single stranded T bases can be used to reduce the angle strain. All these strategies substantially reduce the angle strain and have lead to nanostructure yields close to

100% with respect to the starting strands and as such the wells are presented in the gels and can be seen to be either free or almost free of non-penetrating DNA[14–17].

Despite the inability to efficiently form properly assembled dsDNA tetrahedron the formed products were investigated using AFM. Pei et al. obtained impressive results without further purification of their assembled products, which as previously discussed is likely dominated by improperly assembled extended tetrahedron-like assemblies. Therefore, the resulting assembled solutions were deposited without further purification in the hopes that similarly impressive surface coverage could be achieved. Although the main product is seen in figure 4.6 to not be the properly assembled dsDNA tetrahedron, the extended structures would likely still be functional as their design would mean that even tetrahedron-like assemblies would have thiol groups preferentially located at opposite ends from the free single stranded portion of the DNA tetrahedron. No complimentary strand for that single stranded portion is included. Therefore, regardless of the form of the rest of the assembly, that strand should always be single stranded. Furthermore, the thiols would be expected to preferentially bind to the gold surface. Thus, even in tetrahedron-like assemblies, the single stranded portion should be pointed away from the surface and shielded from the gold surface, even if it is not done to the same extent as would be the case in the properly assembled dsDNA tetrahedron.

The unpurified assembly solution is investigated by deposition both onto a gold surface and also a freshly cleaved mica surface. The theoretical height of the double stranded portion of a DNA tetrahedron is just under 5 nm. However, since the thiolated strands contain a 6 carbon spacer, this would increase the expected height on gold to almost 6.5 nm since the face containing the spacers would be preferentially facing downwards. These estimates are without taking into consideration the single stranded portion which is expected to also be able to interact with the AFM tip. The thiolated face would be expected to preferentially be located at the bottom when deposited onto a gold surface, but not on a mica surface.

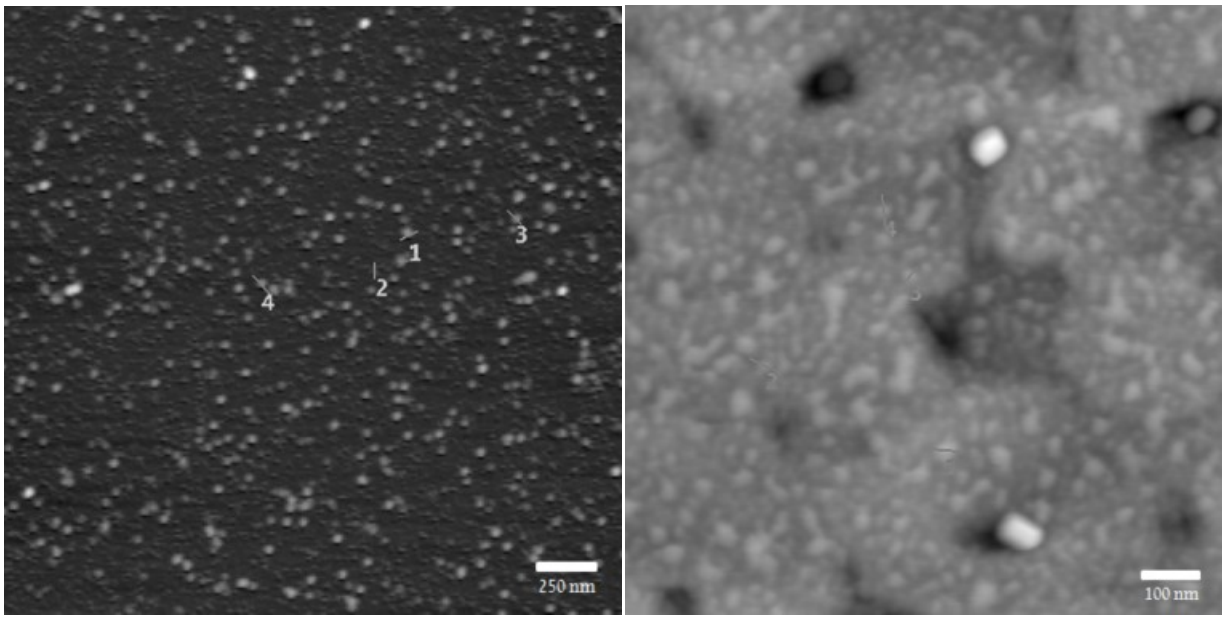


Figure 4.7, AFM height image of non-purified dsDNA Tetrahedron assembly solution deposited onto mica (left) and gold (right). A large variety of feature heights are seen in the mica AFM. A high density of features are seen on both surfaces. Feature sizes from the mica AFM image are fairly consistent in area with both taller and shorter features generally taking up a similar surface area.

Both the AFM done on the DNA tetrahedron solution deposited on the mica surface and that which was deposited on the gold surface result in similar features. The feature heights observed in figure 4.7 are fairly consistent in width and differ mostly in height. Most widths observed are around 50 nm. Using quartz crystal microbalance measurements, Pei et al. calculated that their average distance between surface bound DNA nanostructures is 4 nm. If a similar density is achieved in figure 4.7, it would mean that multiple tetrahedrons are deposited closely together with large separations between clusters of deposited DNA nanostructures. Large areas which are free of DNA are observed. Since tip broadening would only have the effect of shrinking the observed free surface and not of enlarging it, the true distance between AFM features is likely even larger than what is observed in the AFM images. Although the AFM feature widths are fairly consistent, it is not possible to distinguish between the features due to single tetrahedron and tetrahedron-like assemblies and groups of tetrahedron and tetrahedron-like assemblies grouping and depositing together in a consistent fashion.

The maximum amount of properly constructed dsDNA tetrahedrons which would fit within a single AFM feature would be about 20. However as is seen in figure 4.7, although a

large variety of shapes of features is observed, their areas are fairly consistent. The variation in shapes may be due to the fact that based on the gels, many types of tetrahedron-like assemblies would be expected and they would likely have varying shapes. Since the features are too far apart to be connected through bound DNA, it would place an upper limit on the amount of tetrahedra which could have been incorporated into a single extended tetrahedron-like assembly. When imaging small objects in AFM, the majority of the apparent width is due to tip broadening. As such, if the small objects double in diameter, the resulting width in the AFM would only increase by a small amount.

The AFM height distribution plots from figure 4.8 are particularly interesting. Since the thiol groups would be expected to be preferentially surface bound only on the gold substrate and they are connected to the DNA through a carbon spacer, a slightly higher feature height is expected for the tetrahedron on gold, compared to the height on mica. Instead the average height observed on the mica surface is larger than that on the gold surface. Importantly, a significant fraction of the observed Tetrahedron feature heights on the mica are larger than the expected 6 nm, while almost none of the observed feature heights from the gold deposited sample have heights of 6 nm or greater. Interestingly in both gold and mica AFM images a few feature heights of 12 nm are observed but almost none are observed between 7 and 12.

The most frequent height observed on the mica bound tetrahedron is 2.5 nm, while the most frequent height observed from the gold bound tetrahedron AFM image is 3.5 nm. This difference is consistent with the expectation that the surface containing the thiol groups would result in an increase in feature heights and would be expected to be preferentially pointed towards the gold but not the mica surface. In spite of the fact that the most common height is larger on the gold bound tetrahedron measurements, the average height is larger on the mica AFM images. This is consistent with the tetrahedron generally interacting more strongly with the gold surface compared to the mica surface. A stronger interaction would result in greater distortion of the dsDNA in close proximity to the gold surface which would decrease the measured feature height.

AFM feature heights of greater than 6 nm are rare in both AFM images but are considerably more rare in the case of the gold bound tetrahedron. Feature heights of larger than 6.5 nm could not be explained by a single layer of properly assembled Tetrahedron. However on both surfaces investigated, feature heights of 12 nm are observed. That height corresponds almost exactly to double the expected height of a single tetrahedron. While those may have been stacked tetrahedrons, the rinsing done prior to measurement suggests that those features are due to extended tetrahedron-like assemblies. The geometry of the extended assemblies must have meant that obtaining a height of 12 nm is more likely than 9 or 10 nm. Since DNA hybridization is an all or nothing reaction it seems likely to be the cause of the height quantization which is observed on both surfaces. The fact that no 18 nm heights are seen would suggest that no tetrahedron like assemblies containing 3 layers are formed. This would place a maximum limit to the size of the tetrahedron like assemblies of between 10 - 15 tetrahedron equivalents, as larger assemblies should be able to form a 3rd layer and therefore achieve 18 nm heights in the AFM images.

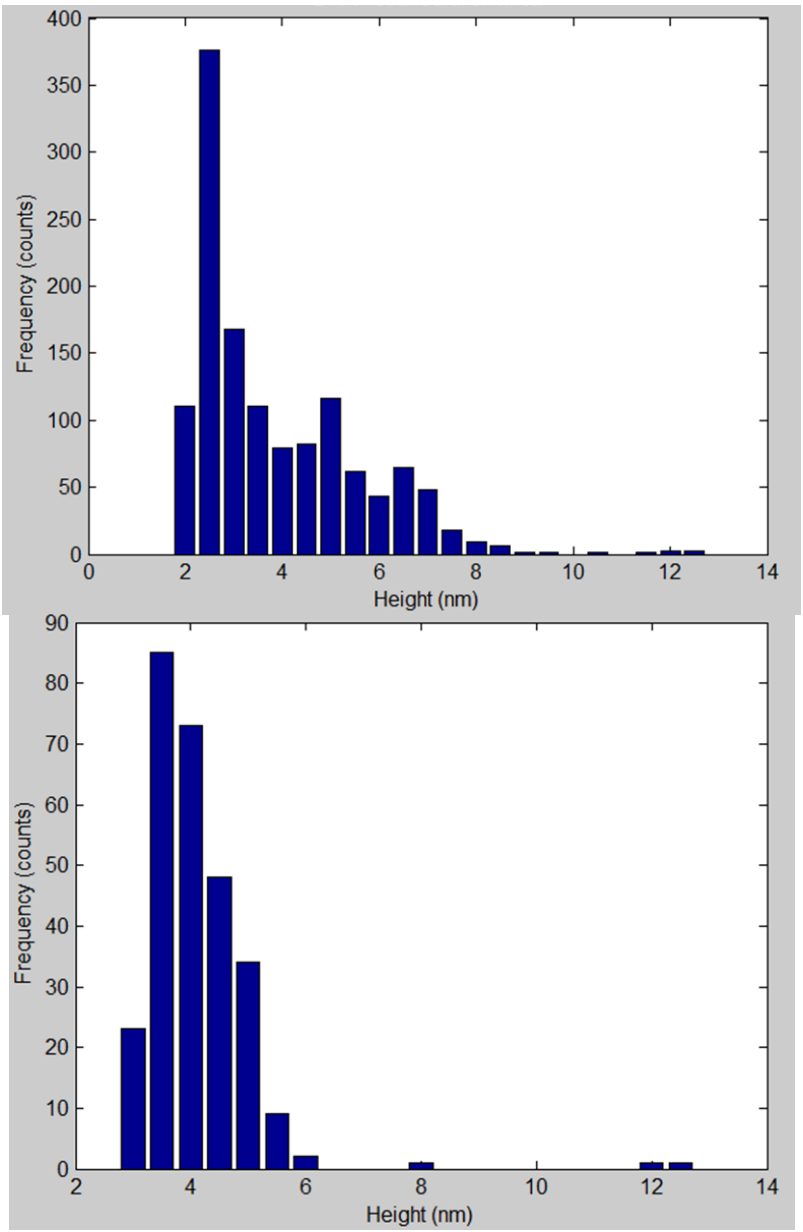


Figure 4.8, Height statistics from AFM features of unpurified dsDNA Tetrahedron solution deposited onto mica (top) and gold (bottom). In both AFM height images a variety of heights are observed. Although the most common height observed is higher in the case of the gold surface, the mica surface contains a greater proportion of taller features.

The surface coverage achieved using the unpurified dsDNA tetrahedron is very good. However, the inability to properly assemble a high percentage of the desired structure likely leads to large distributions in the placement of the single stranded portion of the DNA. In order to achieve a consistently free single stranded portion of the DNA nanostructure which would be completely removed from the gold surface an alternative DNA nanostructure was investigated

for use as the poly A anchor. A dsDNA cube has only 90^{o} angles and as such can be more easily synthesized with a high yield of properly assembled product.

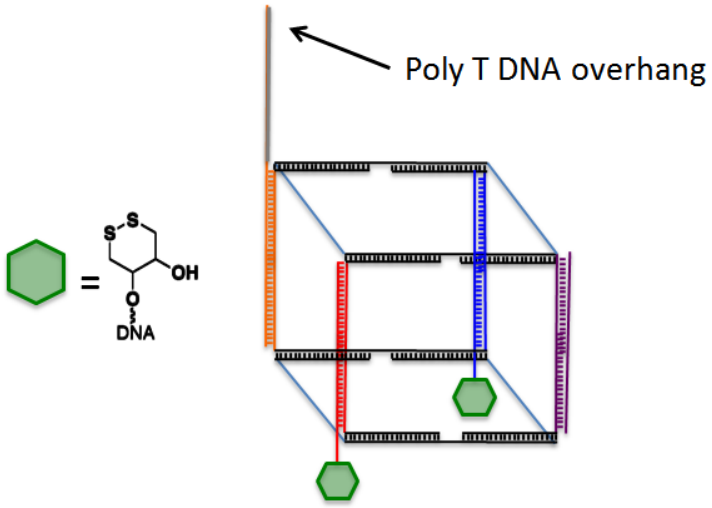


Figure 4.9, Scheme of double stranded DNA cube. Once the dithiol bond is broken the cube can be bound to a gold surface through 4 thiol-gold bonds. Design of cube adapted from Mclaughlin et al.[16]

The DNA cube outlined in figure 4.9 was purified and deposited onto a gold surface, rinsed, dried and then imaged using AFM. The theoretical height of such a DNA cube should be about 7 nm. However although features of around 5 nm are seen in figure 4.10 their density is very low. The cube would be expected to interact with the gold surface, and therefore a feature height of 5 nm would be reasonable. However, a variety of heights are seen. Although unpurified DNA nanostructures would be expected to have varying heights, purified DNA nanostructures should all be of exactly the same dimension. Therefore, although they may interact with the gold surface to a varying degree, they should be fairly consistent. However it was seen in figure 4.10 that not only do the feature heights vary from about 3 to 5 nm, but also the observed features are rather diffuse and not sharp as would be expected from well assembled DNA nanostructures.

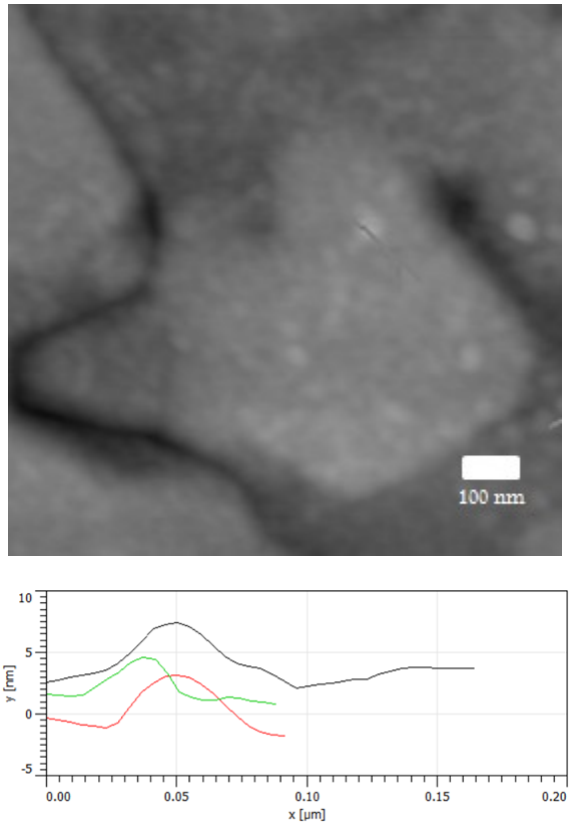


Figure 4.10, AFM height image of DNA cube deposited onto a gold substrate. Feature heights between 3 and 5 nm are observed with a low density, however the features are not sharp as would be expected from a rigid DNA cube as such the features could not be conclusively identifed as DNA cube.

Further confusing the image is the control gold surface, which was exposed to the exact solution conditions and preparation but without any DNA present. The control image is no longer flat but contains a surface roughness of 2 - 3 nm. These apparent height features are due not to any DNA deposition but either to the interaction of the gold surface to solution or the deposition of part of the reaction buffer onto the gold surface.

The AFM phase image in figure 4.12 of the same scan as figure 4.11 could be used as a further indication as to whether or not DNA is deposited on the surface. This is because DNA should interact very differently with the AFM tip compared to the gold surface or other bound molecules, and as such should produce a high contrast in the phase AFM phase image. The contrast due to DNA would be expected to be substantially sharper than that which was seen in figure 4.12.

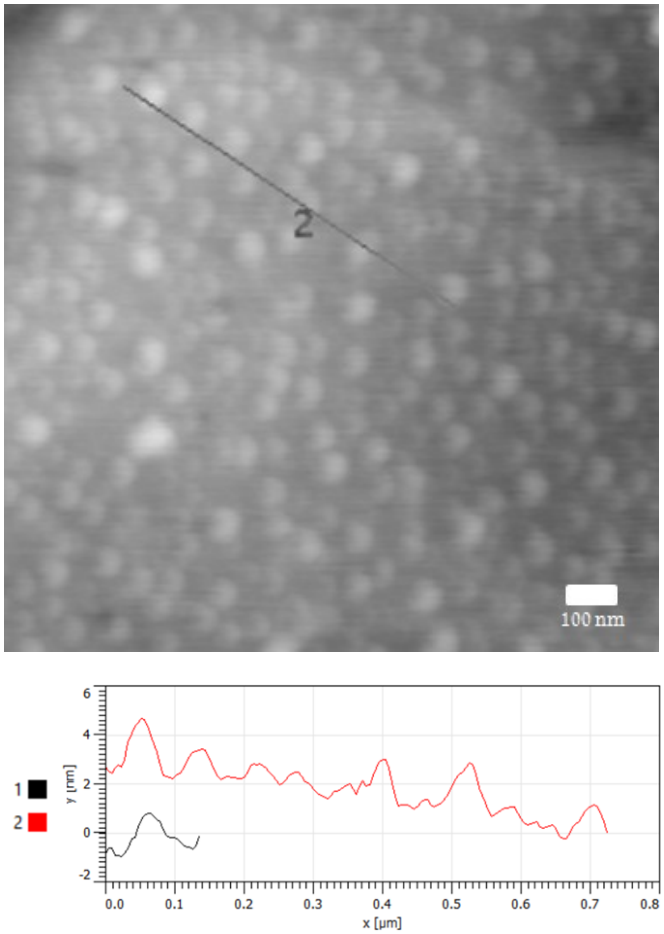


Figure 4.11, AFM height image of gold surface exposed to the same preparations and solution conditions which are used in the DNA cube deposition but without the DNA being present. After exposure to the deposition solution the gold is no longer flat but has an apparent surface roughness of 2 - 3 nm. Thus only features larger than 2 - 3 nm could be conclusively identified.

The deposited cube was not fully double stranded as is seen in figure 4.9. Considering the strength of the interaction between the tetrahedron and the gold surface it seems likely that such a cube might be collapsed when deposited on the gold surface. This would result in significantly smaller AFM height features, would make it difficult to detect, and could potentially have explained the apparent poor surface coverage seen in figure 4.10. Therefore, the cube was fully rigidified and all complementary strands were introduced to make the cube fully double stranded. The dsDNA cube was first deposited on mica in the hope that the AFM tip would interact similarly with the cube regardless of the substrate. This assumption would allow the dsDNA cube to be fully characterized on the less rough mica surface and thereby allow for easier identification of that cube on the gold surface.

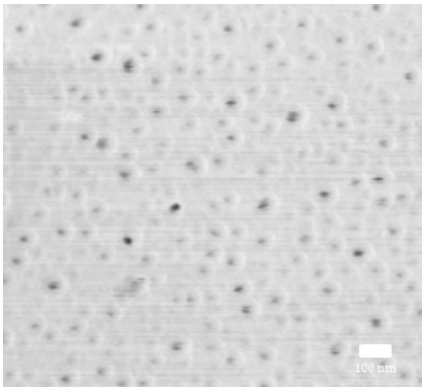


Figure 4.12, AFM phase image of bare gold subjected to the deposition conditions but without any DNA present in solution.

It is seen in figure 4.13 that the height of the cube which was observed on mica is between 5 and 6 nm. Although AFM of purified dsDNA cube was done, the density of cube is too high in those images to get clear data, and as such the unpurified sample was more informative. Both the height and phase signals from the dsDNA cube are sharp and consistent with what would be expected for a properly assembled, rigid dsDNA cube. The majority of the deposited material is not 3 dimensional and has a height of just over 1 nm. That height combined with large length to width ratio of the observed features is consistent with what would be expected from double stranded DNA. The way some of the strands are deposited looks similar to a corner of a dismantled and flattened cube thus they are likely incompletely assembled cube.

Fully rigidified dsDNA cube which was deposited onto gold is able to be identified in a similar fashion to the method used on the mica surface despite the increased surface roughness of the gold. The AFM features of both the height and phase images are seen in figure 4.14 to be sharp, which is what is expected of the rigid dsDNA cubes. The observed heights re slightly less than those observed on the mica, with the heights ranging from 4 - 5 nm. However, considering that the cube would be expected to interact more strongly with the gold than with the mica, a slight lowering would be expected and had been seen previously with the dsDNA tetrahedron.

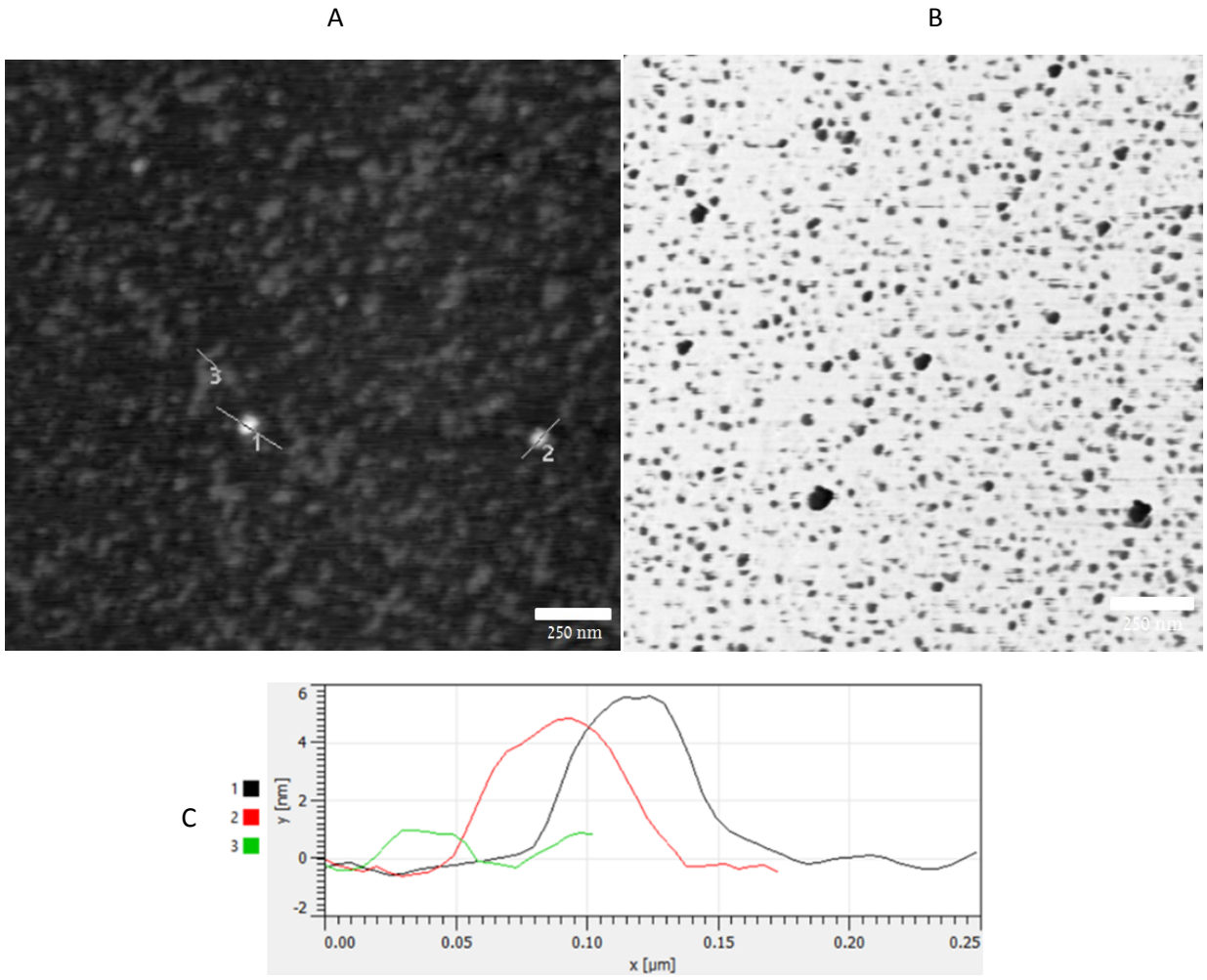


Figure 4.13, A) AFM height image of unpurified dsDNA on mica B) AFM phase image, the phase images of the dsDNA cubes are sharper and of greater contrast then the poorly assembled DNA. C) height trace of cubes (red and black) with heights of about 5-6 nm which is consistent with the expected height of 7 nm. The surrounding material had heights of just over 1 nm (green) which is consistent with double stranded DNA which is not assembled into a 3 dimensional structure.

The surface coverage observed in figure 4.14 is very poor. Although the dsDNA cubes are well resolved, they are not numerous. Prior to deposition, the dithiol was activated using TCEP, which breaks up the dithiol bond into single thiols which can then be bound to the gold surface. However it is seen in figure 4.14 that the achieved density of dsDNA cube is essentially the same regardless of whether the DNA is activated with TCEP or not. Considering the achieved surface coverage in the tetrahedron studies. This would suggest that TCEP is significantly more effective at cleaving linear dithiol bonds compared to cyclic dithiol bonds. This would be logical, as the cyclic bonds could easily reform once broken while once the linear dithiols are separated they are free to float away from each other in solution. Therefore in

order to bind to the gold the cyclic dithiols would have to be broken just before coming into proximity to the gold surface or else they would be prone to simply leave again.

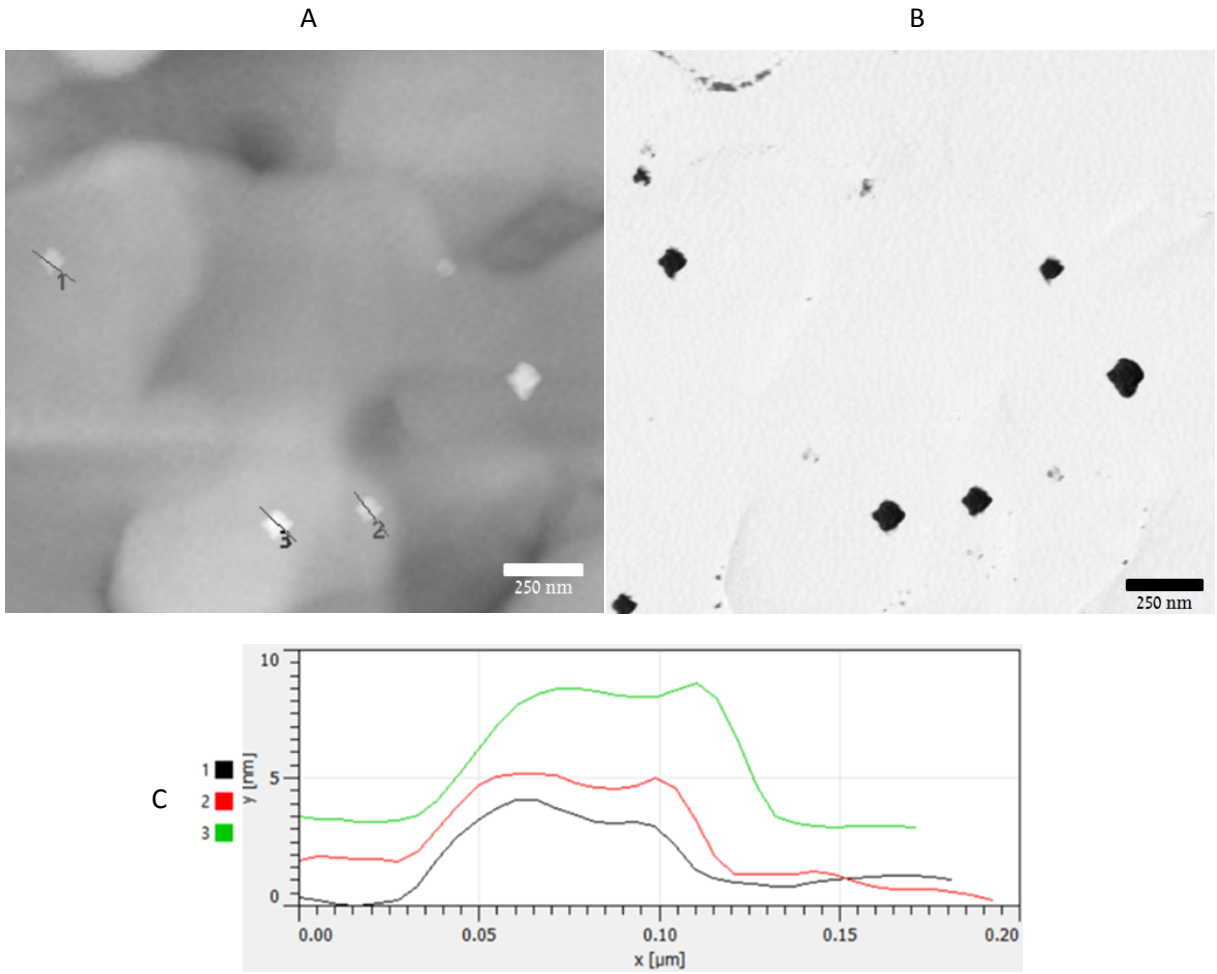


Figure 4.14, A) AFM height image of deposited purified dsDNA cubes. The observed heights are between 4 and 5 nm. B) the phase signal from the dsDNA cubes is substantially higher in contrast to the surrounding material and also corresponds to the exact locations where the height increase is observed which is consistent with the expected signal from a rigid dsDNA cube. C) Trace of cubes. The heights varies between 4 and 5 nm but are very regular in their apparent occupied area.

While the fully rigidified dsDNA cube seems to provide a more consistent structure which could be assembled with higher yields compared to that of the dsDNA tetrahedron, the linear dithiols used in the construction of the dsDNA tetrahedron seem to interact significantly more favorably with the gold surface compared to the cyclic dithiol. Therefore an ideal structure for anchoring poly A to a gold surface would likely be the dsDNA cube modified such that linear dithiols were protruding from 1 face instead of the cyclic dithiols

which were used. Based on the experiments done here, such a structure should be very consistent in keeping the single stranded portion away from the gold surface, while also achieving a high surface density. Shielding of the single stranded portion could be further increased by backfilling with MCH after depositing the modified dsDNA cube on the gold surface.

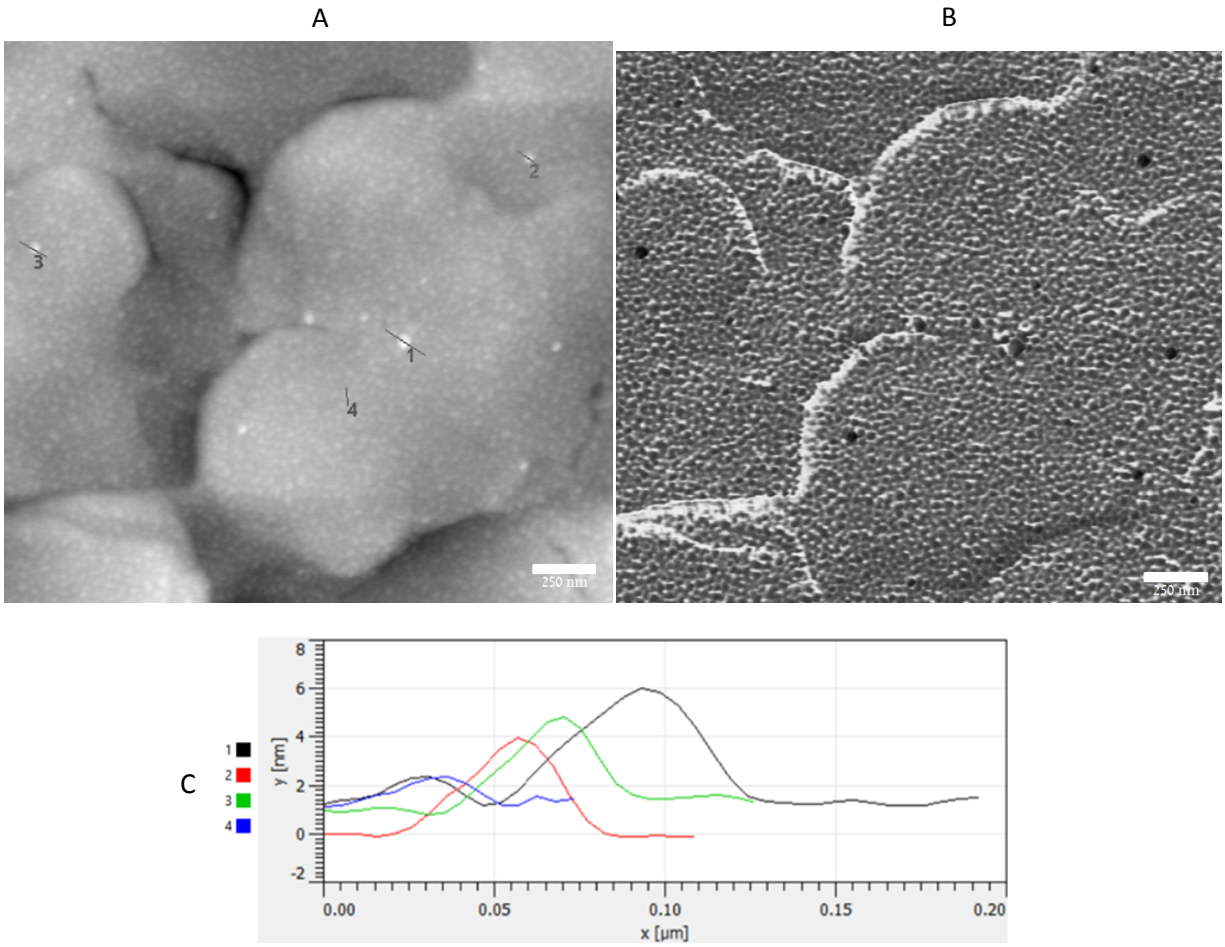


Figure 4.15, A) AFM height image of rigid cube deposited on gold which was not activated using TCEP, the observed heights of features are still between 4 and 5 nm as is observed in the activated deposition. B) The AFM phase image also shows greater contrast and sharp features at locations where dsDNA cube had been deposited. C) the trace of the dsDNA cubes showed heights of between 4 and 5 nm while the surrounding area (blue) had heights of only 1 - 2 nm.

References

(1) Yi, H.; Nisar, S.; Lee, S.-Y.; Powers, M. A.; Bentley, W. E.; Payne, G. F.; Ghodssi, R.; Rubloff, G. W.; Harris, M. T.; Culver, J. N. Patterned Assembly of Genetically Modified Viral Nanotemplates via Nucleic Acid Hybridization. *Nano Lett.* **2005**, *5* (10), 1931–1936.

(2) Levicky, R. L.; Herne, T. M.; Tarlov, M. J.; Satija, S. K. Using Self-Assembly To Control the Structure of DNA Monolayers on Gold: A Neutron Reflectivity Study. *JACS* **1998**, No. 120, 9787–9792.

(3) Herne, T. M.; Tarlov, M. J. Characterization of DNA Probes Immobilized on Gold Surfaces. *J. Am. Chem. Soc.* **1997**, *119* (38), 8916–8920.

(4) Steel, A. B.; Herne, T. M.; Tarlov, M. J. Electrochemical Quantitation of DNA Immobilized on Gold. *Anal. Chem.* **1998**, *70* (22), 4670–4677.

(5) Zhang, Z.-L.; Pang, D.-W.; Zhang, R.-Y.; Yan, J.-W.; Mao, B.-W.; Qi, Y.-P. Investigation of DNA Orientation on Gold by EC-STM. *Bioconjug. Chem.* **2002**, *13* (1), 104–109.

(6) Seetharaman, S.; Zivarts, M.; Sudarsan, N.; Breaker, R. R. Immobilized RNA Switches for the Analysis of Complex Chemical and Biological Mixtures. *Nat. Biotechnol.* **2001**, *19*, 336–341.

(7) Chi, Q.; Wang, G.; Jiang, J. The Persistence Length and Length per Base of Single-Stranded DNA Obtained from Fluorescence Correlation Spectroscopy Measurements Using Mean Field Theory. *Phys. Stat. Mech. Its Appl.* **2013**, *392* (5), 1072–1079.

(8) Tinland, B.; Pluen, A.; Sturm, J.; Weill, G. Persistence Length of Single-Stranded DNA. *Macromolecules* **1997**, *30* (19), 5763–5765.

(9) Lu, Y.; Weers, B.; Stellwagen, N. C. DNA Persistence Length Revisited. *Biopolymers* **2002**, *61* (4), 261–275.

(10) Xiao, Y.; Lou, X.; Uzawa, T.; Plakos, K. J. I.; Plaxco, K. W.; Soh, H. T. An Electrochemical Sensor for Single Nucleotide Polymorphism Detection in Serum Based on a Triple-Stem DNA Probe. *J. Am. Chem. Soc.* **2009**, *131* (42), 15311–15316.

(11) Fan, C.; Plaxco, K. W.; Heeger, A. J. Electrochemical Interrogation of Conformational Changes as a Reagentless Method for the Sequence-Specific Detection of DNA. *Proc. Natl. Acad. Sci.* **2003**, *100* (16), 9134–9137.

(12) Moses, S.; Brewer, S. H.; Lowe, L. B.; Lappi, S. E.; Gilvey, L. B. G.; Sauthier, M.; Tenent, R. C.; Feldheim, D. L.; Franzen, S. Characterization of Single- and Double-Stranded DNA on Gold Surfaces. *Langmuir* **2004**, *20* (25), 11134–11140.

(13) Pei, H.; Lu, N.; Wen, Y.; Song, S.; Liu, Y.; Yan, H.; Fan, C. A DNA Nanostructure-Based Biomolecular Probe Carrier Platform for Electrochemical Biosensing. *Adv. Mater.* **2010**, *22* (42), 4754–4758.

(14) McLaughlin, C. K.; Hamblin, G. D.; Aldaye, F. A.; Yang, H.; Sleiman, H. F. A Facile, Modular and High Yield Method to Assemble Three-Dimensional DNA Structures. *Chem. Commun.* **2011**, *47* (31), 8925.

(15) Lo, P. K.; Karam, P.; Aldaye, F. A.; McLaughlin, C. K.; Hamblin, G. D.; Cosa, G.; Sleiman, H. F. Loading and Selective Release of Cargo in DNA Nanotubes with Longitudinal Variation. *Nat. Chem.* **2010**, *2* (4), 319–328.

(16) McLaughlin, C. K.; Hamblin, G. D.; Hänni, K. D.; Conway, J. W.; Nayak, M. K.; Carneiro, K. M. M.; Bazzi, H. S.; Sleiman, H. F. Three-Dimensional Organization of Block Copolymers on "DNA-Minimal" Scaffolds. *J. Am. Chem. Soc.* **2012**, *134* (9), 4280–4286.

(17) McLaughlin, C. K.; Hamblin, G. D.; Sleiman, H. F. Supramolecular DNA Assembly. *Chem. Soc. Rev.* **2011**, *40* (12), 5647.

Conclusions and Future Work

In comparison to TMV-RNA, poly A is a lower cost, more versatile alternative to TMV-RNA for production of TMV rod based nanotechnology. Unlike OAS based rod assembly, due to the fact that poly A is fully synthetic there is no chance of unexpected plant infectivity issues. TEM, AFM and CD measurements presented here all independently indicate that poly A does induce TMV-like rod formation in solutions of WT-TMV-CP at pH 6.8. The main issues with poly A are not fundamental but due to problems with stability and consistency. These are the most likely reasons why poly A has thus far been largely ignored as an option for producing TMV-like rods even though it may turn out to have more potential than TMV-RNA for production of TMV rod based nanotechnology.

The consistency issues with poly A mediated TMV-like rods have been resolved here through development of an alternative TMV-CP production protocol, with CP growth at $18^{O}C$ found to produce CP which was consistently capable of TMV-like rod self-assembly. While overall concentrations of TMV-CP of about .5 mg/mL were found to be ideal for resolution of individual rods in TEM and AFM images. Furthermore the presence of small amounts of calcium during assembly aided the production of TMV-like rods in CP which otherwise was unable to assemble (figures 3.1, 3.2, 3.8 and 3.9). Although Calcium ion is not needed to assemble high quality, recently grown TMV-CP and was not observed to influence the rate of reaction (figure 3.5) of high quality WT-TMV-CP into TMV-like rods, it may prove to be useful to help assemble more challenging TMV-like rods such as rods assembled at a gold surface or TMV-CP mutants which are especially difficult to assemble. TMV rods assembled using OAS based RNA may not be able to take advantage of a similar boost in self-assembly ability due to the presence of calcium since the OAS was evolved to assemble TMV rods in the absence of calcium ions.

Stability problems with poly A containing TMV-like rods should be similarly resolvable. Insights into RNA-CP interaction and its dependence on triplet sequence could be obtained by studying high resolution electron density maps of poly A containing TMV-like rods which would not suffer from base averaging. These insights may lead to improved design of poly A based

RNA strands with greater stability upon incorporation into TMV-like rods. One such candidate is a mixed polynucleotide containing mostly adenine but also small amounts of guanine. The stability of such a polynucleotide has not been determined but seeing as the TMV genome doesn't suffer from a similar lack of stability it seems likely that a simple solution exists and simply needs to be discovered. Removal of some of the negatively charged residues responsible for rod disassembly through point mutation would also increase the stability of poly A containing TMV-like rods.

Direct visual evidence through TEM and AFM images leave little doubt that poly A induces TMV-like rod formation with the same dimensions and shape as TMV rods (figures 2.15-17, 3.8-10). However, since the interaction between poly A and TMV-CP has not been thoroughly characterized it could be argued based only on the direct images that the poly A simply induces the rod formation and is not bound to the RNA binding track of the TMV-CP. Although no direct evidence was collected here to suggest that this was the case all of the indirect evidence suggested that the produced TMV-like rods do indeed contain poly A which is bound to the RNA binding track of the TMV-CP. TMV-like rods produced over only 2 hours would be expected to contain nicks and imperfections. The fact that none are seen suggest that the rods were formed through interaction with RNA. The lack of free, visible poly A in the AFM and TEM images would also suggest that the poly A is incorporated into the interior of the formed TMV-like rods. Furthermore the CD contribution of the poly A upon completion of the TMV-like rod assembly reaction (figure 2.11) was fairly similar to that from TMV-RNA contained within TMV rods. Since the CD spectrum of RNA depends strongly on the conformation of that RNA it suggests that the poly A was in a similar configuration to TMV-CP bound TMV-RNA. Finally the resistance of the poly A mediated TMV-like rods to dilution compared to pH induced TMV-CP rods demonstrates that the poly A imparts considerable stability to the TMV-like rod (figures 3.9, 3.11 and 3.13) which would very likely only be possible if it were bound to the RNA binding track of the TMV-CP.

No evidence of multiple nucleation sites was seen in the AFM and TEM images. Multiple nucleation sites on a single RNA strand would be expected to cause non-linear rod-like

assemblies. the lack of multiple nucleation sites on poly A strands would be consistent with the currently held view that nucleation of poly A strands is slow and occurs over a slower time period than rod elongation.

No new peaks were observed in the mass spectrum of recently purified TMV-CP in comparison to three month old CP and also no peaks were removed (figure 3.6). The mass spec. data would suggest that the loss in assembly ability of TMV-CP over time is not caused by a strong binding event. The inner loop is the most likely portion of the TMV-CP to be responsible for the effect since it is by far the most reactive portion of the TMV-CP and is absolutely critical for binding RNA. A semi stable conformation of the inner loop which is incompatible with reaction with poly A is the most likely candidate but a more thorough investigation would be required to determine the cause of the observed effect.

Although no poly A was successfully tethered to a gold surface in this study, significant progress has been made towards achieving a gold tethered poly A containing TMV-like rod. The best candidates investigated were the dsDNA nanostructures which will likely be able to provide the level of shielding required for the poly A mediated TMV-like rod assembly reaction to take place near the surface. The best surface coverage was achieved using the DNA tetrahedron (figure 4.7), however, problems with dna self assembly of that system meant that a range of actual structures and effective heights was achieved. The variation in heights would likely result in a variation in ability to shield the TMV-CP from the gold surface during assembly. On the other hand the fully rigidified DNA cube achieved a very consistent structure but failed to achieve the high surface coverage of the tetrahedron (figure 4.14 and 4.15). This was likely due to the cyclic dithiol bond which did not seem to be effectively broken using TCEP. Based on these investigations the ideal tethering dsDNA nanostructure by modifying the DNA cube such that linear dithiols could be used instead, the same overall binding strength could be achieved but with the types of surface coverage which was seen in the tetrahedron system. This would result in a consistent height of single stranded poly T DNA which could then be hybridized to a poly A strand. The consistent height and high surface coverage would likely provide sufficient shielding for poly A to be bound to the gold surface and still be available for TMV-like rod self-

assembly. Backfilling with MCH would provide even more protection for the TMV-CP and could be used to enhance the shielding for the proposed poly A tethering DNA nanostructure if further protection from the gold surface was needed.

Finally, poly A can be used to enhance our fundamental understanding of what factors influence the rate of elongation of TMV and TMV-like rods using various forms of TMV-CP. This fundamental knowledge could be important when designing complex TMV rod based heterostructures with multiple domains as various mutations and modifications will undoubtedly result in a variety of rates of addition to the growing rod. By studying the mechanism of poly A mediated TMV-like rod self-assembly it should be possible to gain insight into what effect mutations and solution conditions will have on TMV and TMV-like rod assembly and thereby aid in designing appropriate mutations and modification strategies to effectively produce the desired TMV-like rod based structure. Studying the reaction kinetics under a greater variety of solution conditions and reactant concentrations should help in furthering our understanding of the mechanism of assembly of poly A mediated TMV-like rods.

Appendix A, Materials and Methods

Production and purification of TMV-CP

WT TMV plasmid was obtained from Norclone, it was produced by returning the 123C mutant plasmid obtained from Prof. Matt Francis back into the WT plasmid. TMV-CP was grown in terrific broth using Tuner(DE3)pLyS competent cells (Novogen). 10 µM IPTG was used to induce TMV-CP expression. High quality TMV-CP was produced by expressing at 18^{o}C for 12 - 14 hours while medium to low quality TMV-CP was produced by expressing at 30^{o}C for 24 hours. After cell lysis ammonium sulfate was used to precipitate the protein and the resulting pellet was re-dispersed in 20 mM TEA, pH 7.2 and purified using a CLB6 Anion exchange column (Purchased from GE) through incorporation of a 300 mM sodium chloride gradient.

Poly A mediated TMV-like rod self-assembly reaction

Unless otherwise stated, purified TMV-CP was dialyzed against 25 mM Tris, .1% Sodium Azide buffer pH 8.9 at 4 ^{o}C for 24 h with at least 1 buffer change before dialyzing to the desired reaction conditions. If needed TMV-CP was concentrated using millipore Centrifugal Filter units (3kDa MW cutoff, Fisher). The same centrifugal filter units were used to concentrate the poly A dissolved in the reaction buffer. All TMV-CP preparation was done at 4^{o}C and then allowed to come to room temperature over 24 hours before mixing with poly A.

Optical Measurements

CD spectra were taken with a Jasco J-810 spectrapolarimeter and UV spectra were measured using a Cary 100 Bio UV/Vis Spectrophotometer.

Atomic Force Microscopy

Recently cleaved Mica substrates were used to plate reaction solution. Droplets were quickly wicked dry using whatman filter paper (#1) after deposition and thirty second rinsing in barnstead water. All AFM measurements were taken using a Cypher AFM.

TEM Imaging

A droplet of the desired solution was deposited on the TEM Grid for 3 min. after which it was stained with a 2.4% phosphotungstic acid solution or a 1.5% Uranyl Acetate solution. TEM images were obtained using a Phillips CM200 TEM.

DNA Deposition

DNA deposition was done by depositing 1 μM of the desired DNA structure for 4 hours. DNA was activated with a 10x molar excess of TCEP. Deposition buffer was 10 mM Tris, 1mM EDTA buffer pH 7.4, 1M sodium chloride.