

Nanoscale Structure Determination of Murray Valley Encephalitis and Powassan Virus Non-coding RNAs

Tyler Mrozowich¹, Amy Henrickson¹, Borries Demeler^{1,4,5} and Trushar R Patel^{1,2,3,5*}

¹Department of Chemistry and Biochemistry, Alberta RNA Research and Training Institute, University of Lethbridge, 4401 University Drive, Lethbridge, AB, T1K 3M4, Canada.

²Department of Microbiology, Immunology and Infectious Disease, Cumming School of Medicine, University of Calgary, Calgary, T2N 1N4, AB, Canada. trushar.patel@uleth.ca.

³Li Ka Shing Institute of Virology and Discovery Lab, University of Alberta, Edmonton, AB, T6G 2E1, Canada.

⁴Department Of Chemistry And Biochemistry, University of Montana, Missoula, MT 59812, United States

⁵NorthWest Biophysics Consortium, University of Lethbridge, University of Lethbridge, 4401 University Drive, Lethbridge, AB, T1K 3M4, Canada

Corresponding Author: Trushar R. Patel, trushar.patel@uleth.ca

ORCID IDS:

Tyler Mrozowich - 0000-0002-1077-9682

Amy Henrickson - 0000-0003-3266-5202

Borries Demeler - 0000-0002-2414-9518

Trushar Patel - 0000-0003-0627-2923

Abstract: Viral infections are responsible for numerous deaths worldwide. Flaviviruses, which contain RNA as their genetic material, are one of the most pathogenic families of viruses. There is an increasing amount of evidence suggesting that their 5' and 3' non-coding terminal regions are critical for their survival. In this study, the 5' and 3' terminal regions of Murray Valley Encephalitis and Powassan virus were examined using biophysical and computational modeling methods. First, the purity of *in-vitro* transcribed RNAs were investigated using size exclusion chromatography and analytical ultracentrifuge methods. Next, we employed small-angle X-ray scattering techniques to study solution conformation and low-resolution structures of these RNAs, which suggested that the 3' terminal regions are highly extended, compared to the 5' terminal regions for both viruses. Using computational modeling tools, we reconstructed 3-dimensional structures of each RNA fragment and compared them with derived small-angle X-ray scattering low-resolution structures. This approach allowed us to further reinforce that the 5' terminal regions adopt more dynamic structures compared to the mainly double-stranded structures of the 3' terminal regions.

Keywords: Non-Coding RNA; Murray Valley Encephalitis Virus; Powassan Virus, Flavivirus; Analytical Ultra Centrifuge ; Small-angle X-ray Scattering ; Computational RNA Structure Modeling

1. Introduction

Family Flaviviridae are small, positive-sense single-stranded RNA viruses that replicate within host cells of arthropods and/or vertebrates. Flaviviruses include deadly viruses such as Murray Valley Encephalitis Virus (MVEV), Powassan Virus (PowV), Japanese encephalitis virus, Dengue virus, Zika virus, West Nile virus and Yellow Fever virus. The World Health Organization and the Centers for Disease Control both cite flaviviruses as a global health threat owing to the ease of transmission by mosquitoes and lack efficient therapeutic or immunoprophylactic strategies [1]. Given the magnitude and severity of disease from this class of viruses, there is a critical need for therapeutics; however, the limited understanding of the virus's replication and their complex interactions with the host cellular proteins through terminal region (TR) interactions hinders therapeutic development [2-4].

MVEV is a member of the Japanese encephalitis serological complex of flaviviruses and was first isolated in 1951 during the initial outbreak in Australia [5]. It is believed to be maintained at the top end of the Northern Territory and the North of Western Australia by a cycle involving mosquitoes (*Culex annulirostris*) and birds [6]. The last major outbreak was in 1974 where 58 cases were reported, 20% resulted in death [7]. Since 2011 there have been increasing instances of MVEV, including nine cases in 2011, three of which resulted in death [8,9]. PowV is the only member of the tick-borne encephalitis serogroup that is currently present in North America [10]. PowV has seen a drastic increase in the last 18 years, compared to the previous 40 with an increase of 671% [11]. It was first identified in Powassan, Ontario in 1958 [12] and has become endemic in the upper Midwest and the Northeastern United States, but cases have also been reported in eastern Russia [13]. The increase in PowV incidences over the last decade highlights the importance of conducting further research to understand the structure-function relationships of the viral genome to ultimately develop therapeutic approaches to combat such viral infections.

Flaviviral genomes consist of a single-stranded RNA molecule approximately 10-11k nucleotides long, depending on species. The RNA genome contains one single open reading frame (ORF) which is flanked by terminal regions (TRs) [1]. The ORF codes for a single polypeptide that is cleaved by a combination of viral and host proteases to produce three structural proteins and the seven non-structural proteins [14]. The flanking structural TRs are highly conserved across all flaviviruses and appear as complex folding sequences essential for viral replication [1,15-17]. Since these terminal regions are so highly conserved and intolerable to mutations/deletions, it is hypothesized, and subsequently shown, that these regions interact with the specific host proteins required for replication [18-20]. The important role of these terminal regions is also highlighted by their ability to cyclize, where the 3' TR interacts with the 5' TR [21,22]. This cyclization of the viral genome is essential for the correct positioning of the NS5 protein, produced by the viral genome [23], which is the viral RNA dependent RNA polymerase, which is required to produce (-) sense stranded RNA, which is used as the replication template for further production of the viral (+) sense RNA genome, facilitating viral replication [24]. Furthermore, flaviviral TRs also interact with a variety of host proteins [21,25,26]. Thus, it is essential to study the complex structure of the

TRs to further understand TRs-viral/host protein interactions and cyclization events, as they play an essential role in viral replication.

Although several attempts have been made to study RNA structures in general [27,28], detailed insight into the structural features of multiple flaviviral strains is lacking. Such detailed insight could provide a platform to develop novel antiviral therapies as well as further our knowledge about viral replication events. Therefore, using a combination of biophysical techniques and computational calculations, we have constructed 3-dimensional structures for the 5' and 3' TRs of both MVEV and PowV. This study provides insight into the organization of flaviviral TRs and provides a framework for the characterization of other flaviviral TRs, or TR segments or other non-coding RNAs.

2. Materials and Methods

2.1.1 RNA Preparation and Purification

We prepared cDNA sequences under the control of a T7 RNA polymerase promoter with two additional G nucleotides on the 5' end, followed by an XbaI restriction enzyme cut site (T[^]CTAGA). We designed our constructs for MVEV and PowV based on the Genbank sequences of KX229766.1 and EU670438.1, respectively. All RNA constructs used in the experiments are listed as follows:

1, MVEV 5TR 1-96nt

GGAGACGUUCAUCUGCGUGAGCUUCCGAUCUCAGUAUUGUUUGGAAGGAUCAUUGAUUAA
CGCGGUUUGAACAGUUUUUUGGAGCUUUUGAUUUCAAU

2, MVEV 3TR 10914-11014

GGCCUGGGAAAAGACUAGGAGAUUCUCGUCUAUUCACACAUCAGUCACAAGGCACCGAGC
GCCGAACACUGUGACUGAUGGGGGGAGAAGACCACAGGAUCUU

3, POWV 5TR 1-111

GGAGAUUUUCUUGCACGUGUGUGCGGGUGCUUUAGUCAGUGUCCGCAGCGUUCUGUUGAA
CGUGAGUGUGUUGAGAAAAAGACAGCUUAGGAGAACAAGAGCUGGGAGUGGUUU

4, POWV 3TR 10735-10839

GGCCCCAGGAAACUGGGGGGGCGGUUCUUGUUCUCCUGAGCCACCACCAUCCAGGCACAG
AUAGCCUGACAAGGAGAUGGUGUGUGACUCGGAAAAACACCCGCUU

Each RNA construct was *in-vitro* transcribed using T7 RNA polymerase then purified using a Superdex 200 increase (GE Healthcare) on an ÄKTA pure FPLC system (GE Healthcare) at 0.5mL/min. RNA Fractions were collected and pooled from the size exclusion chromatography (SEC) fraction collector. Pooled fractions were ethanol precipitated and resuspended in RNA buffer (50 mM Tris pH 7.5, 100 mM NaCl and 5 mM MgCl₂). The pooled fractions were analyzed using Urea-Poly Acrylamide Gel Electrophoresis (Urea-PAGE). 10 µL of ~500 nM RNA was mixed with 2 µL of RNA loading dye and loaded into a 1.0 cm well PAGE casting plate (BioRad). Urea-PAGE (11.2%) was run at 300V at room temperature for 20 minutes in 0.5 X TBE, followed by staining with Sybr Safe (ThermoFisher Scientific) and UV visualization. Prior to any experiment (AUC or SAXS) RNA was heated to 95 °C for 5 min and cooled passively to room temperature to facilitate refolding.

2.2 Analytical Ultracentrifugation (AUC)

The sedimentation velocity (SV) – AUC data for FPLC-purified RNA samples were collected using a Beckman Optima centrifuge and AN50-Ti rotor at 20 °C. Samples were loaded into Epon-2 channel centerpieces at 0.6 µM, 0.5 µM, 0.6 µM and 0.58 µM for 5' and 3' TRs of MVEV and PowV, respectively in 50 mM Tris pH 7.5, 100 mM NaCl and 5 mM MgCl₂ buffer and centrifuged at 35,000 revolutions per minute. Scans were collected at every 20 seconds. We used the UltraScan-III package [29] and San Diego Supercomputing Center on Comet as well as the Texas Advanced Computing Center on Lonestar-5 computing facility for data processing. The SV-AUC data were initially analyzed by means of two-dimensional spectrum analysis (2DSA) with simultaneous removal of time-invariant noise, the meniscus and position fitted [30], enhanced van Holde-Weischet protocol [31], genetic algorithm refinement [32] and further refined with Monte Carlo analysis [33]. The density and viscosity corrections for the buffer were estimated with UltraScan to be 1.0030g/cm³ and 1.0100 cP.

2.3 Small angle X-ray scattering (SAXS)

Utilizing the B21 beamline at Diamond Light Source (Didcot, UK), small-angle X-ray scattering (HPLC-SAXS) data was collected as previously described [34]. Making use of an in-line Agilent 1200 (Agilent Technologies, Stockport, UK) HPLC connected to a flow cell, each purified RNA, 50 μL of $\sim 2.0 \text{ mg mL}^{-1}$ were injected into a buffer equilibrated Shodex KW403-4F (Showa Denko America Inc.) size exclusion column at a flow rate of $0.160 \text{ mL min}^{-1}$. Each frame was exposed to the X-rays for 3 s. Each sample peak region was integrated, using a combined total of approximately 10 datasets, buffer subtracted and then merged using Primus [35] or ScÅtter [36] as described previously [37]. The merged data were then analyzed using Guinier approximation to obtain the radius of gyration (R_g) and study homogeneity of samples [38]. We also performed dimensionless Kratky analysis [39] to investigate if the RNA molecules of interest are folded, as reviewed earlier [40]. Next, the pair-distance distribution ($P(r)$) analysis was performed using the program GNOM [41], which provided the R_g and the maximum particle dimension (D_{max} , the radius at which the $P(r)$ dependence approaches to zero). Information from the $P(r)$ plot was then used to generate the models using DAMMIN [42], with no enforced symmetry (P1), as described earlier [43]. The resulting models were averaged, and filtered to obtain a singular representative model using the DAMAVER package [44] as described previously [45,46].

2.4 Atomic Structures Calculations

MC-Fold [47] was used to predict alternative low energy secondary structures for 5' and 3' TRs of MVEV and PowV (Figure 1). We selected the lowest energy structures in each case as input files for MC-Sym that allows fragment-based reconstruction of 3-dimensional structures using known structures, as described earlier [47]. Briefly, 100 all-atom structures for each RNA were calculated, which were minimized using protocols implemented in MC-Sym. Scattering profiles were simulated with CRY SOL based on minimized structures and compared to experimental SAXS data...The minimized structures were subjected to CRY SOL to determine R_g and goodness-of-fit parameter (χ^2) [48]. The MC-Sym derived structures were ranked based on their χ^2 values and were aligned with low-resolution structures using program SUPCOMB [49].

3. Results

3.1. Purification of *in-vitro* transcribed RNA

The *in-vitro* transcribed RNAs were purified using a Superdex 200 increase column connected to the ÄKTA FPLC unit. Figure 2A presents an elution profile for all four RNA constructs, where peaks at ~8 mL represent elution of template plasmids used to *in-vitro* transcribe these RNAs. The RNAs of interest eluted at ~13 to 14.5 mL and the peaks ~11 mL suggest the presence of an oligomeric assembly of RNAs. The monodispersed fractions (at ~13 to 14.5 mL) were pooled together and subsequently analyzed using Urea-PAGE (Figure 2B). As demonstrated in Figure 2B, all four RNAs migrate similarly and are highly pure, devoid of any aggregation or degradation. These fractions were stored at 4 C until further experiments were carried out.

3.2. Homogeneity studies of RNA

Analytical Ultracentrifuge is a versatile technique to study the homogeneity of biomolecules in solution [50]. In order to further evaluate the purity of all four purified RNAs, we performed sedimentation velocity experiments at the concentration range of 0.5-0.7 μM , and processed the data using UltraScan [29] as described in the Materials and Methods section. Figure 3 presents the sedimentation coefficient distribution for MVEV 5' and 3' as well as PowV 5' and 3' RNA fragments. The sedimentation velocity analysis suggests that all four RNAs are mainly monodisperse with sedimentation coefficient values of 4.27 S for MVEV 5' TR, 4.30 S for MVEV 3' TR, 4.49 S for PowV 5' TR and 4.53 S for PowV 3' TR ($S=10^{-13}$ seconds), as summarized in Table 1. Note that the peaks at ~ 5.5 S suggest that all four RNAs form dimeric or high-order conformations in solution. The AUC data also suggest that PowV TRs with a M_w of ~38 kDa have a slightly higher sedimentation coefficient compared to the MVEV TRs with M_w of ~32 kDa (Table 1). Overall, these experiments indicate that all four RNAs are of suitable purity to perform HPLC-SAXS to determine solution structure.

3.3. Low-resolution structural studies of RNAs

Small-angle X-ray scattering allows solution structure studies of biomolecules and their complexes, albeit at low-resolution. Recent developments with instrumentation in terms of an HPLC unit connected inline with SAXS detection allows for data collection of monodispersed sample devoid of aggregated and degraded products [51,52]. As outlined in the Materials and Methods section, we collected selected HPLC-SAXS data for all four RNAs, followed by the selection of data from a monodispersed peak, buffer-subtraction, and merging of selected datasets. The merged SEC-SAXS data are presented in Figure 4A. Subsequently, the merged data were processed using the Guinier method (plot of $I(q)$ vs. (q^2)), which aids detection of purity and allows determination of the R_g from the data belonging to the low- q region [38]. Figure 4B presents the Guinier plots for 5' and 3' TRs of MVEV and PowV, where the linearity for low- q data demonstrates that indeed all four RNAs are monodispersed and devoid of any aggregation. Based on the Guinier analysis, we obtained R_g values of $29.01 \pm 0.22 \text{ \AA}$, $45.72 \pm 0.48 \text{ \AA}$, $34.65 \pm 0.19 \text{ \AA}$ and $35.84 \pm 0.13 \text{ \AA}$ for MVEV 5' TR, MVEV 3' TR, PowV 5' TR and PowV 3' TR respectively (see Table 1 for more details). Once we confirmed the monodispersity, we processed the raw SAXS data from Figure 4A to obtain dimensionless Kratky plots $(I(q)/I(0))*(q*R_g)^2$ vs $q*R_g$ that allow detection of the folding state of biomolecules [40,52]; for example, globular-shaped biomolecules in solution are observed with a well-defined maximum value of 1.1 at $q*R_g = \sqrt{3}$ [39]. However, the dimensionless Kratky plots for 5' and 3' TR of MVEV and PowV under

investigation demonstrate that all the samples are well-folded and extended in solution (Figure 4C).

Next, using program GNOM, we performed an indirect Fourier transformation to convert the reciprocal-space information ($\ln(I(q))$ vs. (q) , Figure 4A) into the real space electron pair-distance distribution function ($P(r)$, Figure 4D) to obtain R_g and D_{max} for all four RNAs. Note that unlike Guinier analysis, which is restricted to the data in the low- q region, the $P(r)$ analysis utilizes a wider-range of the dataset and aids reliable determination of R_g as well as of D_{max} . As outlined in Table 1, based on $P(r)$ analysis, we obtained the D_{max} of 100 Å, 150 Å, 105 Å, and 130 Å for MVEV 5' TR, MVEV 3' TR, PowV 5' TR, and PowV 3' TR respectively. Furthermore, the R_g values of 30.16 ± 0.11 Å, 46.16 ± 0.18 Å, 34.48 ± 0.06 Å and 38.39 ± 0.08 Å for MVEV 5' TR, MVEV 3' TR, PowV 5' TR and PowV 3' TR respectively were obtained. These values are very similar to those obtained from Guinier analysis indicating that the data are suitable for low-resolution shape reconstruction. The shape of the $P(r)$ plot is indicative of the solution conformation of biomolecules. For example, for a globular-shaped protein, we would have expected a bell-shaped $P(r)$ distribution curve with a maximum at $\sim D_{max}/2$ [53]. However, as presented in Figure 4D, all four RNAs display skewed bell-shaped curves with extended tails, suggesting their extended structures in solution.

To obtain a low-resolution solution structure of each RNA, we employed program DAMMIN [42] that utilizes simulated-annealing protocol and allows the incorporation of $P(r)$ data (*i.e.* D_{max} and R_g values as constraints). We calculated a total of 12 models for 5' and 3' TRs of MVEV and PowV and noted that individual models had an excellent agreement between the experimentally obtained scattering data and the calculated scattering data, as the χ^2 values in each case are ~ 0.8 (see Table 1). Next, we employed program DAMAVER to rotate, align and obtain an averaged filtered structure for each RNA [44]. In each case, the goodness of the superimposition of individual models was estimated by the overlap function—normalized spatial discrepancy (NSD). As presented in Table 1, the low NSD values suggest that 12 models in each case are highly similar to each other. The averaged filtered structures for the 5' and 3' TRs of MVEV and PowV are presented in Figure 5, which indicates that indeed, overall these RNAs have extended structures in solution.

3.4. Computational modeling of RNA structures

SAXS is an outstanding method that allows the determination of solution structures of biomolecules, albeit at low-resolution. Techniques such as X-ray crystallography and nuclear magnetic resonance spectroscopy, provide high-resolution structural information but have limited applications for RNA structural studies due to challenges obtaining high-quality crystals and RNA-labeling, respectively. An alternative to this approach is to employ computational modeling of RNA structures and screen those structures using experimental data. We employed the MC-Fold/MC-Sym pipeline developed by Parisien and Major [47]. For each RNA, we calculated 100 structures using the MC-Sym pipeline based on the secondary structure predicted using MC-Fold (Figure 1). Next, we used program CRY SOL [48] to calculate X-ray scattering profiles of 100 structures of each RNA – 5' and 3' TRs of MVEV and PowV. CRY SOL also allows the determination of R_g from these structures. We also calculated a χ^2 value between the X-ray scattering data of MC-Sym derived structures that were calculated using CRY SOL and experimentally collected X-ray scattering data. Figure 6 presents the R_g (y -axis left-hand side, solid black circles) and χ^2 values (y -axis right-hand

side, grey double circles) for each RNA system suggesting that the RNA molecules under investigation can adopt a variety of conformations. For example, in the case of 5' TR of MVEV, the distribution of R_g values range from $30 \pm 5 \text{ \AA}$, whereas the experimentally determined R_g for this construct is 30.16 \AA . Similarly, for PowV 3' TR, experimentally determined R_g is 38.39 \AA , which largely aligns with the distribution of R_g values of MC-Sym derived structures. On the other hand, the MC-Sym derived models for MVEV 3' TR mainly under-represent the R_g values when compared to experimental R_g of 46.16 \AA . In contrast, for PowV 5' TR, Mc-Sym derived structures have a higher R_g distribution ($\sim 45 \text{ \AA}$) compared to the experimentally determined R_g (34.48 \AA). Moreover, as presented in Figure 6, the χ^2 values for MC-Sym derived 3' TR of MVEV and PowV structures have better distribution around $1.5 \chi^2$, compared to MVEV and PowV 5' TRs.

3.5 Combination of Computational Modeling and Experimental Low-Resolution SAXS Structures

We selected ~ 10 MC-Sym derived structures with the lowest χ^2 values to further investigate if they align with the low-resolution averaged filtered models for each RNA system. Figures 7 and 8 present four of the best χ^2 fit MC-Sym derived structures for MVEV and PowV, indicating that these structures agree well with the low-resolution models based on SAXS data. Overall, MC-Sym derived structures for the 3' TRs of MVEV and PowV (Figure 8) indicate that they adopt highly extended structures mainly in the double-stranded regions. In contrast, the 5' TRs of MVEV and PowV are more dynamic with less double-stranded regions (Figure 7). Furthermore, these results indicate that the 5' TR structures have a higher content of un-paired/loop regions compared to the 3' TRs which contains mainly double-stranded regions.

4. Discussion

MVEV is endemic to northern Australia and has seen an increase in prevalence in the last decade. With extensive human development in areas affected, the likelihood of an outbreak is only increasing. Similarly, PowV is an under-studied flavivirus despite its increasing incidence and severe neurological effects. Patients infected with almost any flaviviruses, including PowV and MVEV, are only treated with symptomatic support. Treatment of the viral infection is nonexistent, resulting in a necessity for further investigation into these potentially deadly viruses. Therefore, further work is required to gain detailed insights into the viral life cycle to better prepare ourselves for potential future outbreaks.

Lowest energy secondary structure predictions (Figure 1) reveal that each RNA molecule varies slightly due to increasing energy, with the exception of PowV 5' TR, and to a lesser degree MVEV 3' TR. The structure with the lowest predicted energy of PowV 5' TR displays a large change in secondary structure with only a small change in energy (-103.07 vs -102.43). The secondary structure information for MVEV 3' TR indicates that an increase in energy could result in the formation of a second stem-loop with a small energy change (-104.44 to -103.29). This suggests that the RNA molecules can adopt multiple different conformations in solution.

We prepared all four RNAs using *in-vitro* transcription with T7 polymerase and purified using SEC (Figure 2B), which indicates that these RNAs are able to form dimer/oligomers in solution. The fractions indicating a homogenous preparation were pooled together and analyzed using Urea-PAGE (Figure 2B), where all four RNAs displayed a single band, indicating the absence of degradation. The resolution of the denaturing gel was not enough to discern the difference in molecular weights between MVEV (~32 kDa) and PowV (~37kDa) RNA constructs (Table 1). Therefore, we employed sedimentation velocity experiments using analytical ultracentrifuge which determines the composition and anisotropy of biomolecules in solution [50,54,55]. Figure 3 presents the sedimentation coefficient distribution of SEC-purified RNA preparation. As evident by the peaks between 5S and 6S, all these RNAs can form higher-order oligomers to a small degree, despite removing many of them by means of SEC purification. This is consistent with the HPLC data from SAXS (not shown) which also shows a minor amount of oligomerization for each RNA. Another explanation for these minor oligomerization peaks could be the RNA adopting different conformations, as was seen in the predicted secondary structure(s) of PowV 5' TR (Figure 1) where a small energy change was seen to cause the RNA to adopt a different conformation. This prediction could explain the increased peak that is evident at 5.5 S (Figure 3), for both PowV 5' TR and MVEV 3' TR. Additionally, the peak at ~ 4 S for MVEV 5' TR indicates minor degradation. More importantly, the differences in RNA size and shape are indicated by their sedimentation coefficient values (Table 1). For example, both the TRs of MVEV have a M_w of ~ 32 kDa and produce very similar S values (4.27 S and 4.3 S, for 5' and 3' TR, respectively). Furthermore, the S values for 5' and 3' TRs of PowV are 4.50 S and 4.53 S, this similarity correlates with the similarities of their M_w of ~ 37 kDa (Table 1).

Small-angle X-ray scattering (SAXS) has emerged as an excellent complimentary structural-biophysical method which aids solution structural studies of RNA, proteins, and their complexes, albeit at low-resolution [40,43,51,56-59]. Therefore, solution X-ray scattering is employed for biological systems where obtaining high-quality crystals or labeling of

biomolecules that presents challenges [21,43,45,57,58,60]. In this study, we employed an HPLC-SAXS set-up to collect scattering data from a preparation free of any aggregation or degradation. The monodispersed preparation was confirmed by Guinier analysis and shows excellent linearity of fit in the low- q region (Figure 4B). Furthermore, Guinier analysis also provided R_g of all four RNAs (based on the low- q region), which are highly similar to those calculated by means of $P(r)$ analysis (Figure 4D, Table 1). We also performed dimensionless Kratky analysis (Figure 4C) that demonstrates each RNA adopts an elongated structure, with a low amount of flexibility. The $P(r)$ distribution reveals a quick increase to the highest point, and then a gradual decrease where the electron-pair-distance approaches zero. A globular molecule would result in a gaussian-like distribution, which is not evident in any of the RNA, suggesting their elongated nature. The most extended molecule is MVEV 3' TR resulting in a D_{max} of 150 Å, even though its molecular weight is 31.57kDa (Table 1), lower than both 5' and 3' TR of PowV (Table 1). Low-resolution structure modeling (Figure 5) further confirms that MVEV 3' TR adopts a much thinner and elongated structure, suggesting that this RNA is almost entirely helical. This is also evident for PowV 3' TR which also displays an extended conformation. This elongation is consistent with the predicted secondary structure of flaviviral 3' terminal regions where the last ~100 nucleotides are suggested to almost entirely base pair into a large system [21,61-63]. Compared to the 3' TRs, low-resolution structures of 5' TRs of MVEV and PowV are less elongated. However, their low-resolution structures appear to be consistent with the predicted secondary structures of flaviviral 5' TRs, indicating that the 5' TRs adopt more dynamic structures [25,61-63]. Overall, MVEV and PowV TR structures are consistent with previously published low-resolution structures of flaviviral RNA such as the Zika, and West Nile virus [21,25,64].

SAXS can be combined with high-resolution structures/homology models of individual domains as well as with computational studies, to compliment predicted structures [21,40,57,60,65]. RNA structures of high resolution are hard to determine through crystallization and labeling studies mentioned above; therefore, an attractive alternative to calculating high-resolution structures is computational modeling. To this end, we employed the MC-Fold/MC-SYM pipeline to reconstruct all-atom structures and screen those using experimentally collected X-ray scattering data to identify structures that are likely adopted by these RNAs. Using CRY SOL [48], we calculated R_g for each structure for each TR, as well as back-calculated X-ray scattering data, and aligned them with experimentally collected scattering data (Figure 6). This analysis demonstrated that MVEV 3' TR and PowV 3' TR exhibit a strong correlation between the experimentally determined R_g and the calculated R_g from MC-Sym-derived structures. Similarly, for both of these RNAs, χ^2 values indicate agreement between scattering data that were experimentally collected and calculated from MC-Sym-derived structures, where the χ^2 values were close to ~1.5 for most structures (Figure 6). On the other hand, the χ^2 values for MC-Sym-derived structures for 5' TRs of MVEV and PowV have a much larger distribution (from ~1.5 to ~6). A similar trend was also observed for R_g for both 5' TRs where the MC-Sym derived structures displayed a wider distribution of R_g , suggesting that these RNAs could adopt multiple conformations. On the other hand, the correlation between R_g and χ^2 was far more pronounced in the more elongated TR RNAs - PowV 3' TR and MVEV 3' TR (Figures 5 and 6). Finally, we selected the ~10 MC-Sym-derived structures, which had the lowest χ^2 values when compared to experimental data and aligned them into the low-resolution SAXS envelopes. Selected groups of structures are presented in Figures 7 and 8, which indicate the computationally calculated structures agree well with the experimental data, especially for 3' TRs. For the 5'

TRs, we observe dynamic structures; however, they maintain overall, similar conformation in solution. Furthermore, the bulge regions in both 5' TRs also agree well with their secondary structures, suggesting that these secondary structures could be conserved and do not undergo conformational changes. An added benefit of these computationally derived structures is the addition of directionality of the RNAs (the black spheres represent the 5' end of each molecule in Figures 7 and 8), which is otherwise very challenging to decipher unless the molecule of interest is altered [66].

5. Conclusions

In this study, we have *in-vitro* transcribed MVEV and PowV terminal RNA regions and performed their native purification, studied their homogeneity using analytical ultracentrifugation, determined their low-resolution structures and combined computational modeling to obtain high-resolution structural models. We demonstrated that the combination of SAXS and computational methods allow further in-sights into structural details of RNA molecules.

Author Contributions: Project was conceptualized by TM and TRP, who also performed SAXS data collection, and data analysis. TM prepared RNA samples. AH and TM collected AUC data. AH, TM and BD processed AUC data. All authors contributed in manuscript writing and figures preparation.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant, RGPIN-2017-04003 to TRP. BD is a Canada 150 Research Chair in Biophysics and TRP is a Canada Research Chair in RNA & Protein Biophysics. AUC calculations were performed at the San Diego Supercomputing Center on Comet (support through NSF/XSEDE grant TG-MCB070039N to BD) and at the Texas Advanced Computing Center on Lonestar-5 (supported through UT grant TG457201 to BD). We thank DIAMOND Light Source, UK for access to B21 Bio-SAXS beamline (BAG – SM22113).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fernández-Sanlés, A.; Ríos-Marco, P.; Romero-López, C.; Berzal-Herranz, A. Functional Information Stored in the Conserved Structural RNA Domains of Flavivirus Genomes. *Frontiers in Microbiology* **2017**, *8*, doi:10.3389/fmicb.2017.00546.
2. Chen, J.Y.; Chen, W.N.; Poon, K.M.; Zheng, B.J.; Lin, X.; Wang, Y.X.; Wen, Y.M. Interaction between SARS-CoV helicase and a multifunctional cellular protein (Ddx5) revealed by yeast and mammalian cell two-hybrid systems. *Archives of virology* **2009**, *154*, 507-512, doi:10.1007/s00705-009-0323-y.
3. Zhao, S.; Ge, X.; Wang, X.; Liu, A.; Guo, X.; Zhou, L.; Yu, K.; Yang, H. The DEAD-box RNA helicase 5 positively regulates the replication of porcine reproductive and respiratory syndrome virus by interacting with viral Nsp9 in vitro. *Virus Res* **2015**, *195*, 217-224, doi:10.1016/j.virusres.2014.10.021.
4. Li, C.; Ge, L.L.; Li, P.P.; Wang, Y.; Dai, J.J.; Sun, M.X.; Huang, L.; Shen, Z.Q.; Hu, X.C.; Ishag, H., et al. Cellular DDX3 regulates Japanese encephalitis virus replication by interacting with viral un-translated regions. *Virology* **2014**, *449*, 70-81, doi:10.1016/j.virol.2013.11.008.
5. A SURGICAL JUBILEE. *Medical Journal of Australia* **1955**, *1*, 799-800, doi:10.5694/j.1326-5377.1955.tb38253.x.
6. Skov, S.J.; Chikritzhs, T.N.; Li, S.Q.; Pircher, S.; Whetton, S. If you are registered with the eMJA, your username is the email address you have provided. AMA members can login using their AMA username and password. More info here.
7. Bennett, N. Murray Valley encephalitis: indeed a “mysterious disease”. *Vic Inf Dis Bull* **2008**, *11*, 94-107.
8. Niven, D.J.; Afra, K.; Iftinca, M.; Tellier, R.; Fonseca, K.; Kramer, A.; Safronetz, D.; Holloway, K.; Drebot, M.; Johnson, A.S. Fatal Infection with Murray Valley Encephalitis Virus Imported from Australia to Canada, 2011. *Emerg Infect Dis* **2017**, *23*, 280-283, doi:10.3201/eid2302.161161.
9. Knox, J.; Cowan, R.U.; Doyle, J.S.; Ligtermoet, M.K.; Archer, J.S.; Burrow, J.N.; Tong, S.Y.; Currie, B.J.; Mackenzie, J.S.; Smith, D.W., et al. Murray Valley encephalitis: a review of clinical features, diagnosis and treatment. *Med J Aust* **2012**, *196*, 322-326, doi:10.5694/mja11.11026.
10. Kazimírová, M.; Thangamani, S.; Bartíková, P.; Hermance, M.; Holíková, V.; Štibrániová, I.; Nuttall, P.A. Tick-borne viruses and biological processes at the tick-host-virus interface. *Frontiers in cellular and infection microbiology* **2017**, *7*, 339.
11. Fatmi, S.S.; Zehra, R.; Carpenter, D.O. Powassan virus—a new reemerging tick-borne disease. *Frontiers in public health* **2017**, *5*, 342.
12. McLean, D.M.; McQueen, E.J.; Petite, H.E.; MacPherson, L.W.; Scholten, T.H.; Ronald, K. Powassan virus: field investigations in northern Ontario, 1959 to 1961. *Canadian Medical Association Journal* **1962**, *86*, 971.

13. Anderson, J.F.; Armstrong, P.M. Prevalence and genetic characterization of Powassan virus strains infecting *Ixodes scapularis* in Connecticut. *The American journal of tropical medicine and hygiene* **2012**, *87*, 754-759.
14. Chambers, T.J.; Hahn, C.S.; Galler, R.; Rice, C.M. Flavivirus genome organization, expression, and replication. *Annu Rev Microbiol* **1990**, *44*, 649-688, doi:10.1146/annurev.mi.44.100190.003245.
15. Brinton, M.A. Replication cycle and molecular biology of the West Nile virus. *Viruses* **2013**, *6*, 13-53, doi:10.3390/v6010013.
16. Brinton, M.A.; Basu, M. Functions of the 3' and 5' genome RNA regions of members of the genus Flavivirus. *Virus Res* **2015**, *206*, 108-119, doi:10.1016/j.virusres.2015.02.006.
17. Brinton, M.A.; Fernandez, A.V.; Dispoto, J.H. The 3'-nucleotides of flavivirus genomic RNA form a conserved secondary structure. *Virology* **1986**, *153*, 113-121, doi:10.1016/0042-6822(86)90012-7.
18. Ariumi, Y. Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection. *Frontiers in genetics* **2014**, *5*, 423, doi:10.3389/fgene.2014.00423.
19. Li, C.; Ge, L.L.; Li, P.P.; Wang, Y.; Sun, M.X.; Huang, L.; Ishag, H.; Di, D.D.; Shen, Z.Q.; Fan, W.X., et al. The DEAD-box RNA helicase DDX5 acts as a positive regulator of Japanese encephalitis virus replication by binding to viral 3' UTR. *Antiviral research* **2013**, *100*, 487-499, doi:10.1016/j.antiviral.2013.09.002.
20. Tingting, P.; Caiyun, F.; Zhigang, Y.; Pengyuan, Y.; Zhenghong, Y. Subproteomic analysis of the cellular proteins associated with the 3' untranslated region of the hepatitis C virus genome in human liver cells. *Biochemical and biophysical research communications* **2006**, *347*, 683-691.
21. Deo, S.; Patel, T.R.; Chojnowski, G.; Koul, A.; Džananović, E.; McEleney, K.; Bujnicki, J.M.; McKenna, S.A. Characterization of the termini of the West Nile virus genome and their interactions with the small isoform of the 2' 5'-oligoadenylate synthetase family. *Journal of structural biology* **2015**, *190*, 236-249, doi:10.1016/j.jsb.2015.04.005.
22. Alvarez, D.E.; Lodeiro, M.F.; Luduena, S.J.; Pietrasanta, L.I.; Gamarnik, A.V. Long-range RNA-RNA interactions circularize the dengue virus genome. *Journal of virology* **2005**, *79*, 6631-6643, doi:10.1128/JVI.79.11.6631-6643.2005.
23. Villordo, S.M.; Gamarnik, A.V. Genome cyclization as strategy for flavivirus RNA replication. *Virus Res* **2009**, *139*, 230-239, doi:10.1016/j.virusres.2008.07.016.
24. Westaway, E.G. Flavivirus replication strategy. *Adv Virus Res* **1987**, *33*, 45-90, doi:10.1016/s0065-3527(08)60316-4.
25. Deo, S.; Patel, T.R.; Džananović, E.; Booy, E.P.; Zeid, K.; McEleney, K.; Harding, S.E.; McKenna, S.A. Activation of 2' 5'-oligoadenylate synthetase by stem loops at the 5'-end of the West Nile virus genome. *PloS one* **2014**, *9*, e92545, doi:10.1371/journal.pone.0092545.

26. Meier-Stephenson, V.; Mrozowich, T.; Pham, M.; Patel, T.R. DEAD-box helicases: the Yin and Yang roles in viral infections. *Biotechnol Genet Eng Rev* **2018**, *34*, 3-32, doi:10.1080/02648725.2018.1467146.
27. Chen, Y.-L.; Lee, T.; Elber, R.; Pollack, L. Conformations of an RNA Helix-Junction-Helix Construct Revealed by SAXS Refinement of MD Simulations. *Biophysical Journal* **2019**, *116*, 19-30.
28. Chen, Y.; Pollack, L. Wiley Interdiscip. Rev. *RNA (New York, N.Y.)* **2016**, *7*, 512.
29. Demeler, B.; Gorbet, G.E. Analytical Ultracentrifugation Data Analysis with UltraScan-III. In *Analytical Ultracentrifugation: Instrumentation, Software, and Applications*, Uchiyama, S., Arisaka, F., Stafford, W.F., Laue, T., Eds. Springer Japan: Tokyo, 2016; 10.1007/978-4-431-55985-6_8pp. 119-143.
30. Brookes, E.; Cao, W.; Demeler, B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *European biophysics journal : EBJ* **2010**, *39*, 405-414, doi:10.1007/s00249-009-0413-5.
31. Demeler, B.; van Holde, K.E. Sedimentation velocity analysis of highly heterogeneous systems. *Analytical biochemistry* **2004**, *335*, 279-288, doi:10.1016/j.ab.2004.08.039.
32. Brookes, E.; Demeler, B. *Parsimonious Regularization using Genetic Algorithms Applied to the Analysis of Analytical Ultracentrifugation Experiments*; 2007; 10.1145/1276958.1277035pp. 361-368.
33. Demeler, B.; Brookes, E. Monte Carlo analysis of sedimentation experiments. *Colloid and Polymer Science* **2008**, *286*, 129-137, doi:10.1007/s00396-007-1699-4.
34. Meier, M.; Moya, A.; Krahn, N.; McDougall, M.; McRae, E.K.; Booy, E.P.; Patel, T.R.; McKenna, S.A.; Stetefeld, J. Structure and hydrodynamics of a DNA G-quadruplex with a cytosine bulge. *Nucleic Acids Research* **2018**, *46*, 5319–5331, doi:10.1093/nar/gky307.
35. Konarev, P.V.; Volkov, V.V.; Sokolova, A.V.; Koch, M.H.J.; Svergun, D.I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **2003**, *36*, 1277-1282, doi:10.1107/S0021889803012779.
36. Rambo, R. *ScÅtter, a JAVA-based application for basic analysis of SAXS datasets*, 2017.
37. Majdi Yazdi, M.; Saran, S.; Mrozowich, T.; Lehnert, C.; Patel, T.R.; Sanders, D.A.R.; Palmer, D.R.J. Asparagine-84, a regulatory allosteric site residue, helps maintain the quaternary structure of *Campylobacter jejuni* dihydrodipicolinate synthase. *Journal of structural biology* **2019**, 10.1016/j.jsb.2019.107409, 107409, doi:10.1016/j.jsb.2019.107409.
38. Guinier, A.; Fournier, G. *Small angle scattering of X-rays*; Wiley: New York, 1955.
39. Durand, D.; Vives, C.; Cannella, D.; Perez, J.; Pebay-Peyroula, E.; Vachette, P.; Fieschi, F. NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *Journal of structural biology* **2010**, *169*, 45-53, doi:10.1016/j.jsb.2009.08.009.
40. Patel, T.R.; Chojnowski, G.; Astha; Koul, A.; McKenna, S.A.; Bujnicki, J.M. Structural studies of RNA-protein complexes: A hybrid approach involving hydrodynamics,

- scattering, and computational methods. *Methods (San Diego, Calif.)* **2017**, *118-119*, 146-162, doi:10.1016/j.ymeth.2016.12.002.
41. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **1992**, *25*, 495-503, doi:doi:10.1107/S0021889892001663.
 42. Svergun, D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical journal* **1999**, *76*, 2879-2886.
 43. Reuten, R.; Patel, T.R.; McDougall, M.; Rama, N.; Nikodemus, D.; Gibert, B.; Delcros, J.; Prein, C.; Meier, M.; Metzger, S., et al. Structural decoding of netrin-4 reveals a regulatory function towards mature basement membranes. *Nature Communications* **2016**, *7*:13515, 1-17.
 44. Volkov, V.V.; Svergun, D.I. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* **2003**, *36*, 860-864, doi:doi:10.1107/S0021889803000268.
 45. Dzananovic, E.; Patel, T.R.; Chojnowski, G.; Boniecki, M.J.; Deo, S.; McEleney, K.; Harding, S.E.; Bujnicki, J.M.; McKenna, S.A. Solution conformation of adenovirus virus associated RNA-I and its interaction with PKR. *Journal of structural biology* **2014**, *185*, 48-57, doi:10.1016/j.jsb.2013.11.007.
 46. Patel, T.R.; Bernards, C.; Meier, M.; McEleney, K.; Winzor, D.J.; Koch, M.; Stetefeld, J. Structural elucidation of full-length nidogen and the laminin-nidogen complex in solution. *Matrix biology : journal of the International Society for Matrix Biology* **2014**, *33*, 60-67, doi:10.1016/j.matbio.2013.07.009.
 47. Parisien, M.; Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **2008**, *452*, 51-55, doi:10.1038/nature06684.
 48. Svergun, D.; Barberato, C.; Koch, M.H.J. CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **1995**, *28*, 768-773, doi:Doi 10.1107/S0021889895007047.
 49. Kozin, M.B.; Svergun, D.I. Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **2001**, *34*, 33-41, doi:doi:10.1107/S0021889800014126.
 50. Patel, T.R.; Winzor, D.J.; Scott, D.J. Analytical ultracentrifugation: A versatile tool for the characterisation of macromolecular complexes in solution. *Methods (San Diego, Calif.)* **2016**, *95*, 55-61, doi:10.1016/j.ymeth.2015.11.006.
 51. Brosey, C.A.; Tainer, J.A. Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Current opinion in structural biology* **2019**, *58*, 197-213, doi:10.1016/j.sbi.2019.04.004.
 52. Perez, J.; Vachette, P. A Successful Combination: Coupling SE-HPLC with SAXS. *Adv Exp Med Biol* **2017**, *1009*, 183-199, doi:10.1007/978-981-10-6038-0_11.
 53. Svergun, D.I.; Koch, M.H.J. Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **2003**, *66*, 1735-1782, doi:10.1088/0034-4885/66/10/r05.

54. Unzai, S. Analytical ultracentrifugation in structural biology. *Biophys Rev* **2018**, *10*, 229-233, doi:10.1007/s12551-017-0340-0.
55. Uchiyama, S.; Noda, M.; Krayukhina, E. Sedimentation velocity analytical ultracentrifugation for characterization of therapeutic antibodies. *Biophys Rev* **2018**, *10*, 259-269, doi:10.1007/s12551-017-0374-3.
56. Chen, Y.; Pollack, L. SAXS studies of RNA: structures, dynamics, and interactions with partners. *Wiley interdisciplinary reviews. RNA* **2016**, *7*, 512-526, doi:10.1002/wrna.1349.
57. Kim, D.N.; Thiel, B.C.; Morozowich, T.; Hennelly, S.P.; Hofacket, I.L.; Patel, T.R.; Sanbonmatsu, K.Y. Zinc-finger protein CNBP alters the 3-D structure of lncRNA Braveheart in solution. *Nature Communications* **2020**.
58. Patel, T.R.; Meier, M.; Li, J.; Morris, G.; Rowe, A.J.; Stetefeld, J. T-shaped arrangement of the recombinant agrin G3-IgG Fc protein. *Protein science : a publication of the Protein Society* **2011**, *20*, 931-940, doi:10.1002/pro.628.
59. Krahn, N.; Meier, M.; To, V.; Booy, E.P.; McEleney, K.; O'Neil, J.D.; McKenna, S.A.; Patel, T.R.; Stetefeld, J. Nanoscale Assembly of High-Mobility Group AT-Hook 2 Protein with DNA Replication Fork. *Biophysical journal* **2017**, *113*, 2609-2620, doi:10.1016/j.bpj.2017.10.026.
60. Dzananovic, E.; Astha; Chojnowski, G.; Deo, S.; Booy, E.P.; Padilla-Meier, P.; McEleney, K.; Bujnicki, J.M.; Patel, T.R.; McKenna, S.A. Impact of the structural integrity of the three-way junction of adenovirus VAI RNA on PKR inhibition. *PloS one* **2017**, *12*, e0186849, doi:10.1371/journal.pone.0186849.
61. Ochsenreiter, R.; Hofacker, I.L.; Wolfinger, M.T. Functional RNA Structures in the 3'UTR of Tick-Borne, Insect-Specific and No-Known-Vector Flaviviruses. *Viruses* **2019**, *11*, doi:10.3390/v11030298.
62. Liu, X.; Liu, Y.; Zhang, Q.; Zhang, B.; Xia, H.; Yuan, Z. Homologous RNA secondary structure duplications in 3' untranslated region influence subgenomic RNA production and replication of dengue virus. *Virology* **2018**, *524*, 114-126, doi:10.1016/j.virol.2018.08.018.
63. Kasprzak, W.K.; Shapiro, B.A. MPGAfold in dengue secondary structure prediction. *Methods Mol Biol* **2014**, *1138*, 199-224, doi:10.1007/978-1-4939-0348-1_13.
64. Zhang, Y.; Zhang, Y.; Liu, Z.Y.; Cheng, M.L.; Ma, J.; Wang, Y.; Qin, C.F.; Fang, X. Long non-coding subgenomic flavivirus RNAs have extended 3D structures and are flexible in solution. *EMBO Rep* **2019**, *20*, e47016, doi:10.15252/embr.201847016.
65. Reuten, R.; Nikodemus, D.; Oliveira, M.B.; Patel, T.R.; Brachvogel, B.; Breloy, I.; Stetefeld, J.; Koch, M. Maltose-Binding Protein (MBP), a Secretion-Enhancing Tag for Mammalian Protein Expression Systems. *PloS one* **2016**, *11*, e0152386, doi:10.1371/journal.pone.0152386.
66. Zettl, T.; Mathew, R.S.; Shi, X.; Doniach, S.; Herschlag, D.; Harbury, P.A.B.; Lipfert, J. Gold nanocrystal labels provide a sequence-to-3D structure map in SAXS reconstructions. *Sci Adv* **2018**, *4*, eaar4418, doi:10.1126/sciadv.aar4418.
67. Demeler, B. UltraScan A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. In *Modern Analytical*

- Ultracentrifugation: Techniques and Methods.*, Scott, D.J., Harding, S.E., Rowe, A.J., Eds. Royal Society of Chemistry (UK): 2005; pp. 210-229
68. Svergun, D.I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J Appl Crystallogr* **1992**, *25*, 495-503, doi:Doi 10.1107/S0021889892001663.
69. Darty, K.; Denise, A.; Ponty, Y. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **2009**, *25*, 1974-1975, doi:10.1093/bioinformatics/btp250.

Table 1. Biophysical parameters of MVEV and PowV non-coding RNAs.

Sample	MVEV 5' TR	MVEV 3' TR	PowV 5' TR	PowV 3' TR
Mw (kDa, sequence)	31.38	31.57	36.92	37.80
Sedimentation coefficient (S) [∇]	4.27 (4.08, 4.46)	4.30 (4.20, 4.41)	4.50 (4.42, 4.55)	4.53 (4.52, 4.53)
$I(0)^{\#}$	$0.0098 \pm 4.7 \times 10^5$	$0.019 \pm 1.4 \times 10^5$	$0.019 \pm 6.7 \times 10^5$	$0.034 \pm 8.2 \times 10^5$
$q \cdot R_g$ range	0.38 – 1.26	0.59 – 1.26	0.34 – 1.26	0.47 – 1.29
R_g (Å) [#]	29.01 ± 0.22	45.72 ± 0.48	34.65 ± 0.19	35.84 ± 0.13
$I(0)^{\#\Delta}$	$0.098 \pm 3.8 \times 10^5$	$0.018 \pm 8.1 \times 10^5$	$0.019 \pm 4.2 \times 10^5$	$0.034 \pm 8.9 \times 10^5$
R_g (Å) ^Δ	30.16 ± 0.11	46.16 ± 0.18	34.48 ± 0.06	38.39 ± 0.08
D_{max} (Å) ^Δ	100	150	105	130

χ^2^*	~0.78	~0.83	~0.78	~0.80
NSD [*]	0.71 ± 0.01	0.80 ± 0.02	0.85 ± 0.02	0.72 ± 0.01

\bar{v} - determined using SV-AUC analysis and UltraScan-III package [67]. Sedimentation coefficients obtained following genetic algorithm-Monte Carlo analysis. Measured value represents the mean. Bracketed values represent 95% confidence interval from Monte Carlo analysis.

- obtained from Guinier analysis [38].

Δ - determined using P(r) analysis using GNOM program [68].

* - values derived from DAMMIN [42] and DAMAVER [44] analysis.

List of Figures

Figure 1. Lowest energy predicted secondary structures of MVEV and PowV non-coding RNA regions calculated using VARNA [69]. ΔG values for the lowest energy structures for MVEV 5' TR, MVEV 3' TR, PowV 5' TR, and PowV 3' TR are -86.41, -104.44, -103.07, and -103.07 kcal mol⁻¹, respectively.

Figure 2. Purification of in-vitro transcribed 5' and 3' TRs of MVEV and PowV RNA. (A) SEC elution profile of purified RNAs where x-axis represents elution volume and y-axis represents absorbance at 260 nm; (B) Urea-PAGE demonstrating purity of the individual pooled RNA fractions from SEC peaks (~13 to 14.5 mL).

Figure 3. Homogeneity studies of MVEV and PowV non-coding terminal regions using SV-AUC, where x-axis represents sedimentation coefficient distribution. The SV peaks at ~4.5 S for each RNA represent monomeric fractions. Colour scheme is identical to Figure 2.

Figure 4. Characterization of MVEV and PowV terminal regions using SAXS. (A) A plot of scattering intensity (Log I(q)) versus scattering angle ($q = 4\pi\sin\theta/\lambda$) representing merged SAXS data for MVEV and PowV. (B) Guinier plots (plot of $\ln(I(q))$ versus q^2) representing the homogeneity of samples and allowing determination of R_g based on the low-angle region data. (C) Dimensionless Kratky plots ($I(q)/I(0) * (qR_g)^2$) versus $q.R_g$ for all four RNA samples demonstrating their extended structures. (D) Pair-distance distribution (P(r)) plots for all four RNA samples representing their maximal particle dimensions and allowing the determination of R_g from the entire SAXS dataset.

Figure 5. Determination of low-resolution structures using SAXS, which indicates that all RNA molecules adopt an extended structure in solution. For each terminal region RNA, the left and right panels represent 180° rotated view on the x-axis and y-axis, respective of the view presented in the middle panel. Dimensions represent the D_{max} .

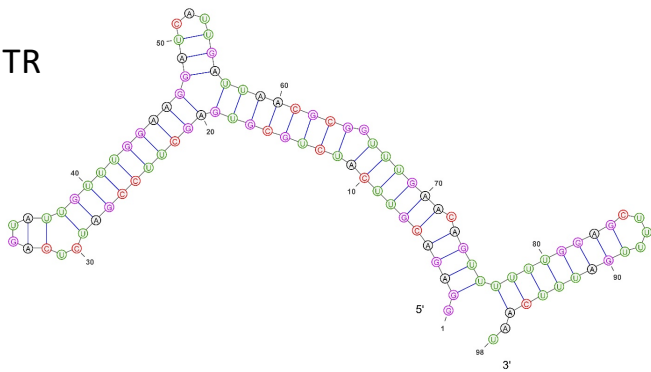
Figure 6. Screening of high-resolution structures calculated using MC-Sym [47] for MVEV 5' TR, MVEV 3' TR, PowV 5' TR and PowV 3' TR. For each RNA, the x-axis represents model numbers (total 100 models), whereas the y1-axis (solid black circles) and y2-axis (grey double circles) represent R_g (Å) and χ^2 values, respectively, calculated using CRY SOL package [48]. For each plot, the solid dark line corresponds to y1-axis and represents the value for experimentally determined R_g (Å) as presented in Table 1. The dotted line corresponds to y2-axis and represents a χ^2 value of 1.5. These plots indicate that MC-Sym derived structures represent a wide-range of conformations these RNAs can theoretically adopt.

Figure 7. High-resolution structures calculated using MC-SYM overlaid with low-resolution SAXS models. A) MVEV 5TR B) and PowV 5TR. Bottom panels in both cases represent a 180° rotation about the x-axis represented in the top panels. Black spheres represent 5' terminal region on each construct.

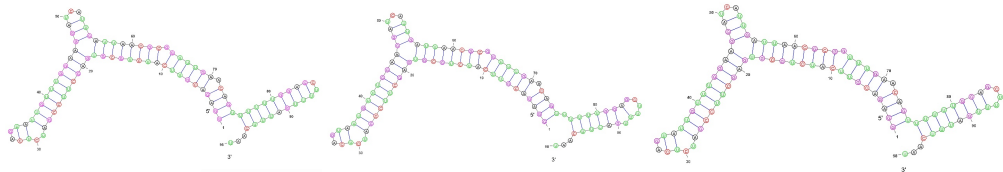
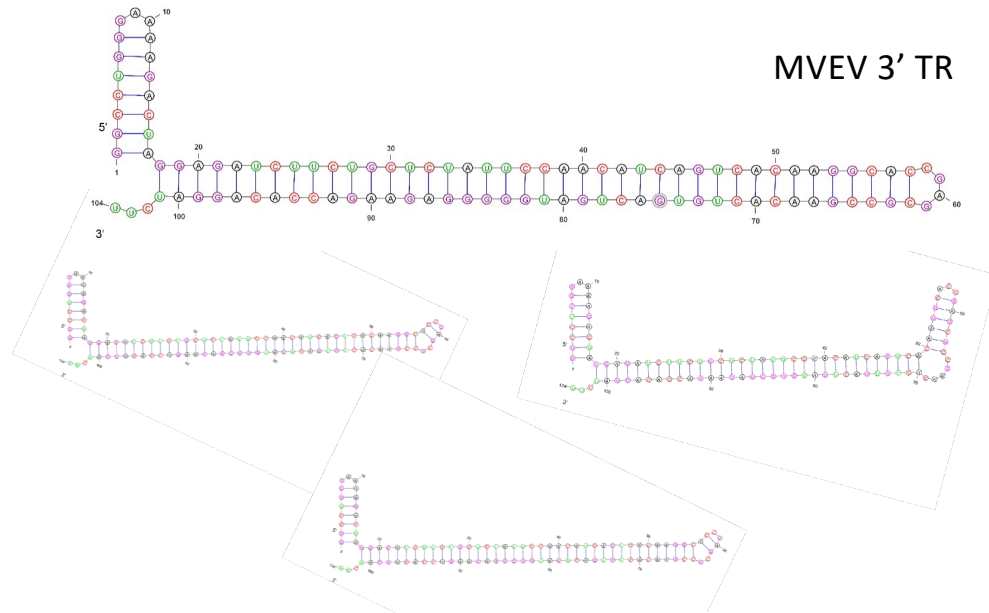
Figure 8. High-resolution structures calculated using MC-SYM overlaid with low-resolution SAXS models. A) MVEV 3TR B) and PowV 3TR. Bottom panels in both cases

represent an 180° rotation about the x-axis represented in the top panels. Black spheres represent 5' terminal region on each constructs.

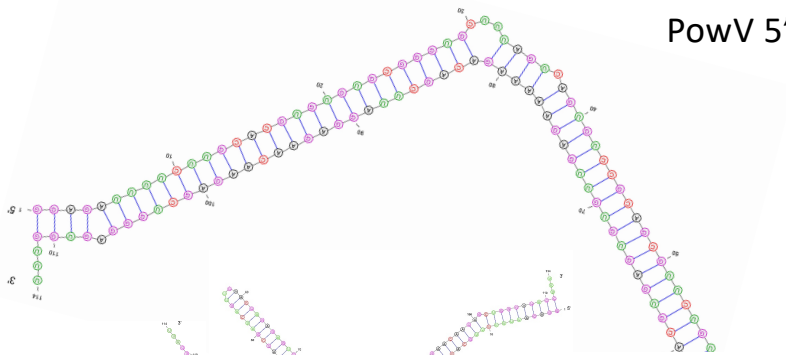
MVEV 5' TR



MVEV 3' TR



PowV 5' TR



PowV 3' TR

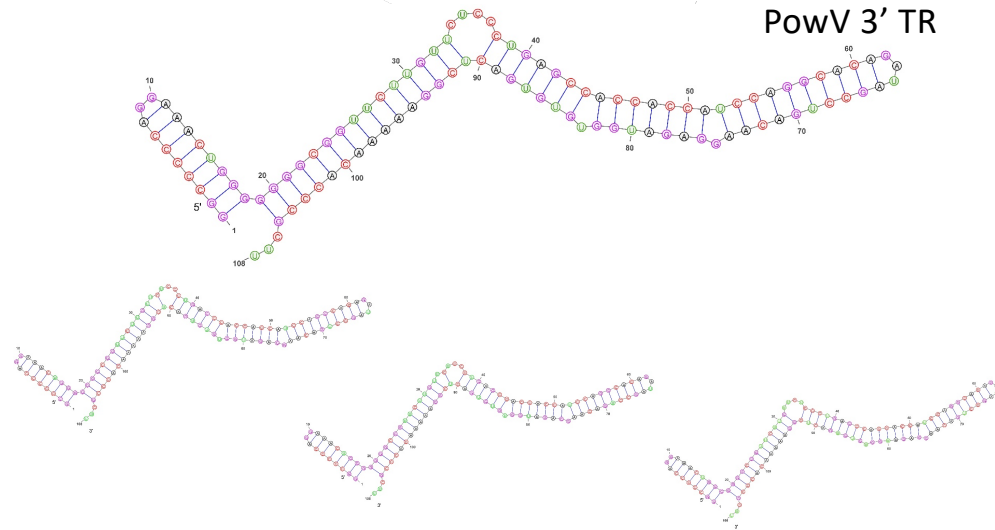
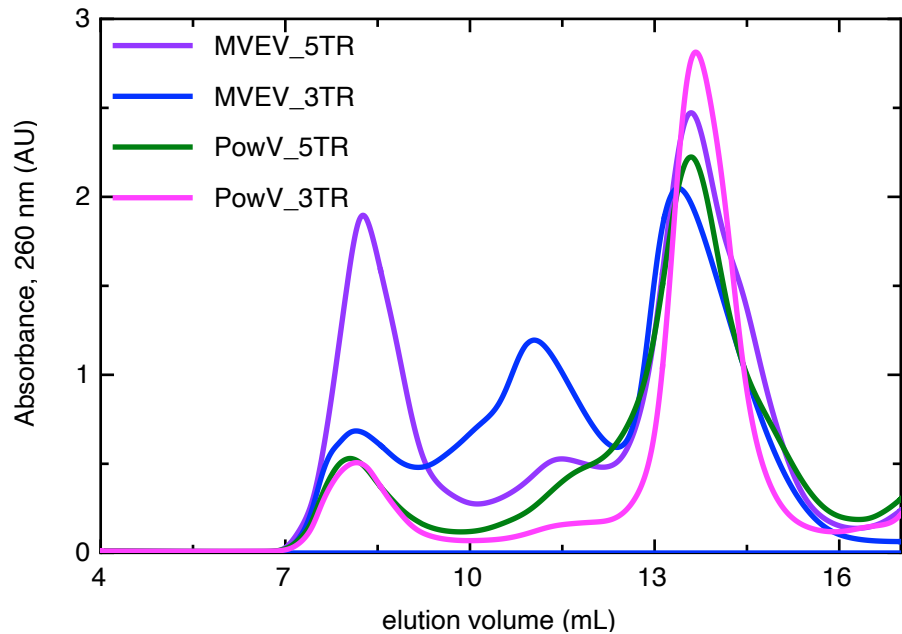
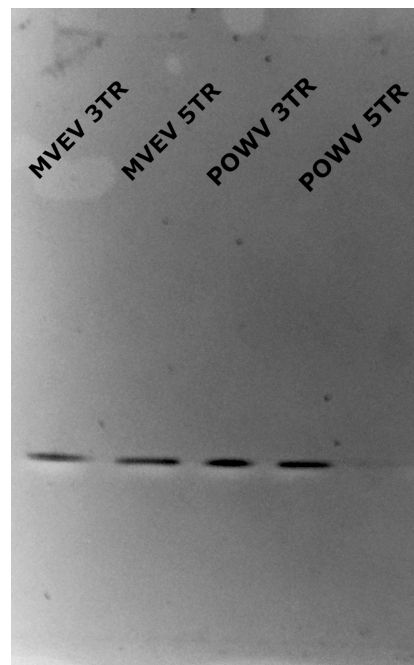


Figure 1

A**B****Figure 2**

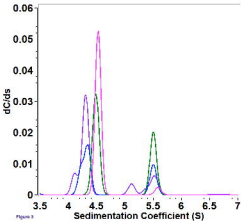


Figure 1

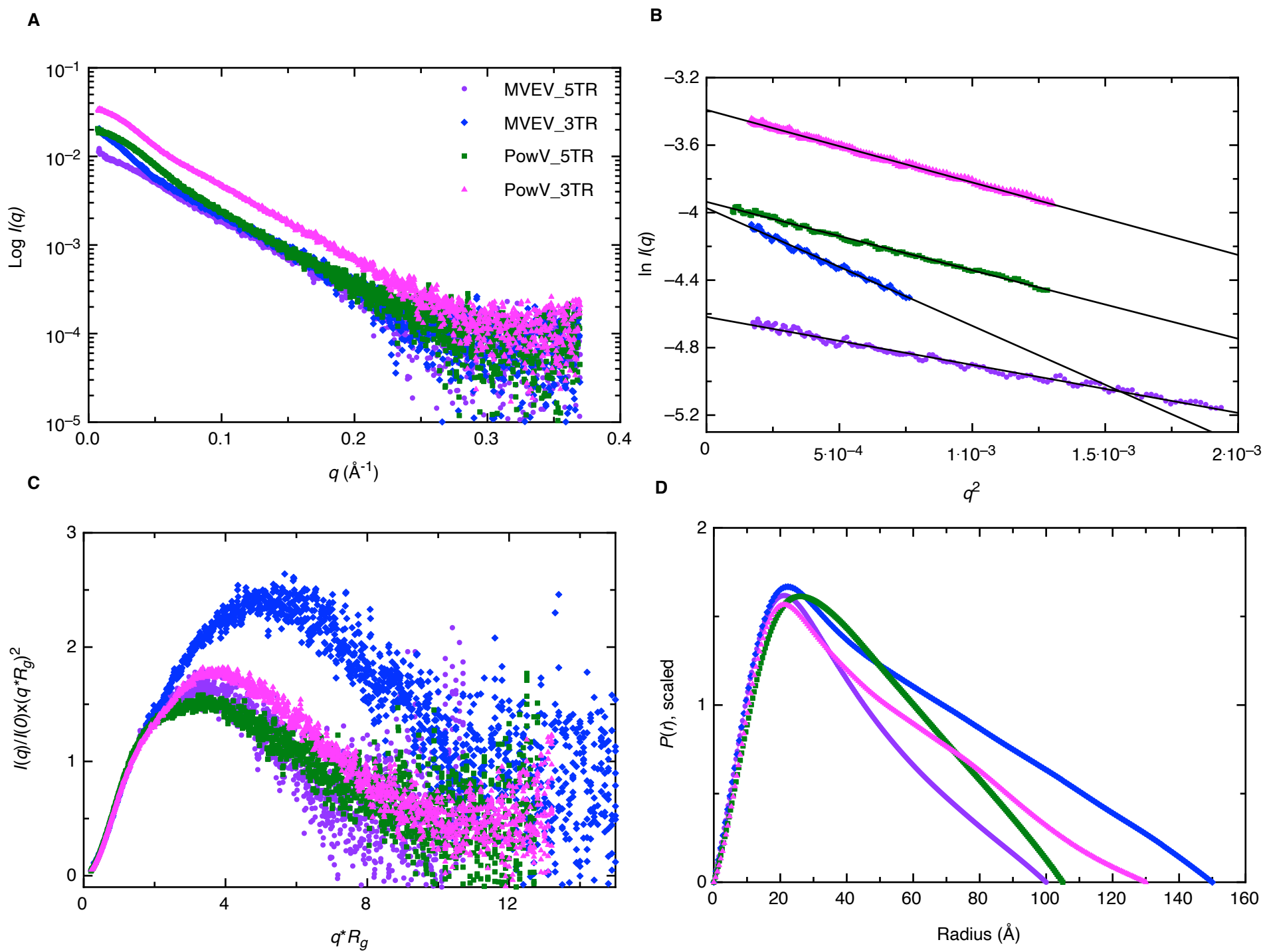
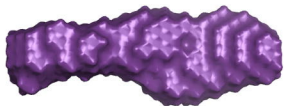
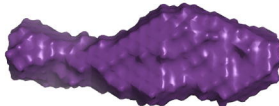


Figure 4

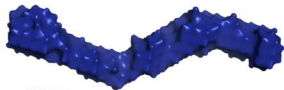
MVEV 5'TR



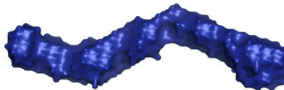
100 Å



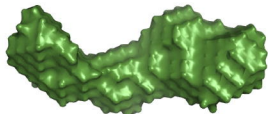
MVEV 3'TR



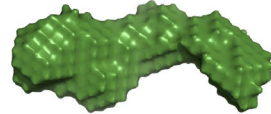
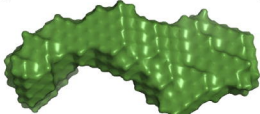
150 Å



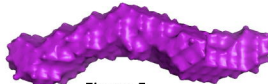
PowV 5'TR



105 Å



PowV 3'TR



130 Å

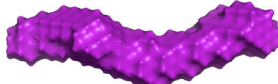
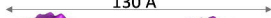


Figure 5

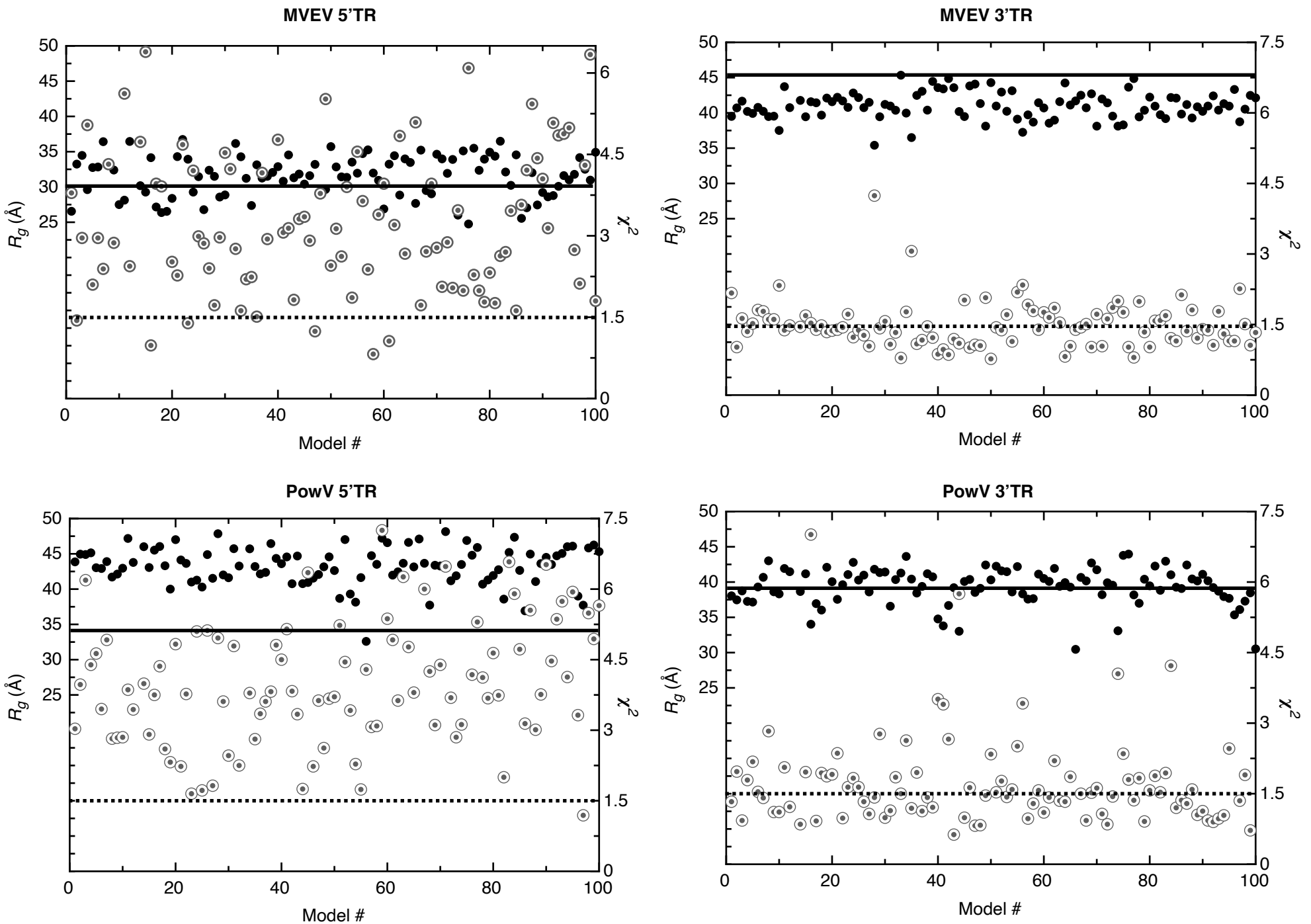


Figure 6

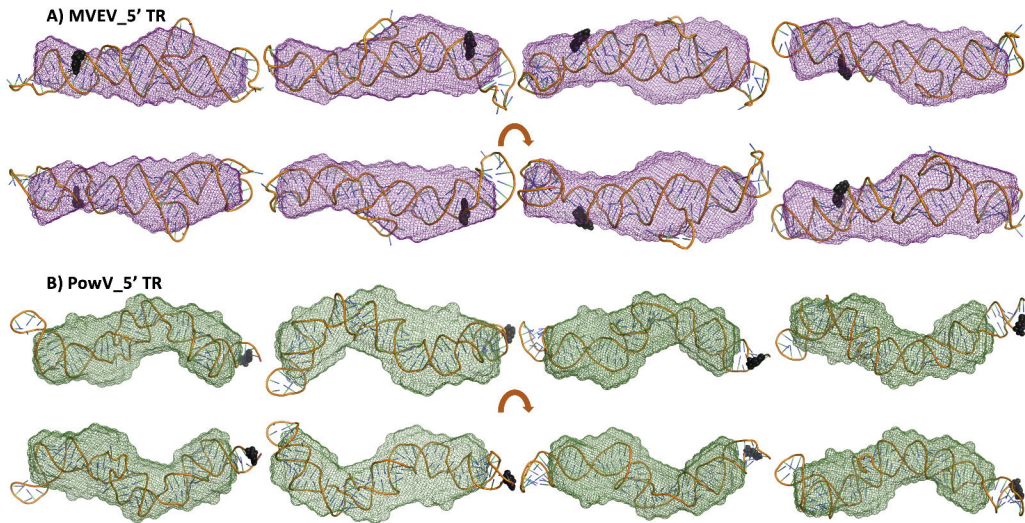
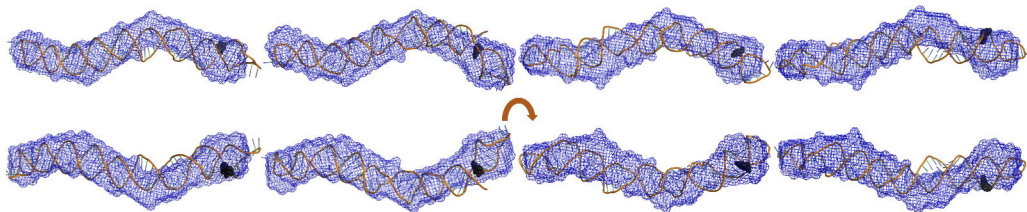


Figure 7

A) MVEV_3' TR



B) PowV_3' TR

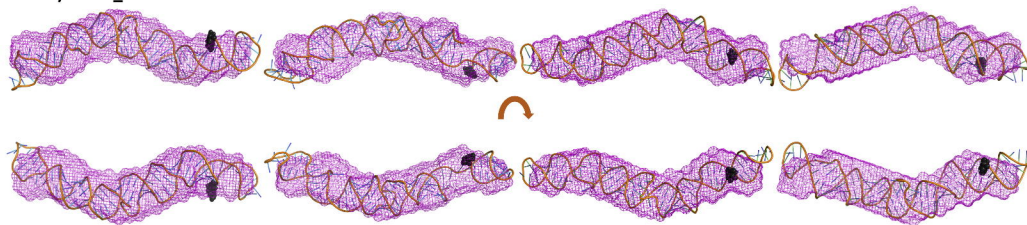


Figure 8