agronomy

MDPI

Article

Probing the RNA Structure of a Satellite RNA of Cucumber Mosaic Virus Using SHAPE Method

Zhifei Liu [1,†], Xinran Cao [2,3,4,*,†], Chengming Yu [1] and Xuefeng Yuan [1,*]

[1] College of Plant Protection, Shandong Agricultural University, Tai'an 271018, China; wliuzhifei@163.com (Z.L.); ycm2006.apple@163.com (C.Y.)
[2] College of Plant Health and Medicine, Qingdao Agricultural University, Qingdao 266109, China
[3] College of Horticulture Science and Engineering, Shandong Agricultural University, Tai'an 271018, China
[4] Shouguang International Vegetable Sci-Tech Fair Management Service Center, Shouguang 262700, China
* Correspondence: xinran1001@163.com (X.C.); snowpeak77@163.com (X.Y.)
† These authors contributed equally to this work.

Abstract: Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) is a widely used technique for RNA structure analysis using N-methylisatoic anhydride (NMIA) treatment that has been proven to be applicable to different types of RNA templates. In this study, we performed the structural analysis of the viral satellite RNA of cucumber mosaic virus TA-Tb (satCMV TA-Tb) using the SHAPE method. In the preliminary experiment, we optimized the protocol of SHAPE method for analyzing satCMV TA-Tb by determining a suitable quantity of template RNA. This optimization effectively reduced the appearance of a large number of intense bands in the NMIA-untreated sample lane, along with a strong overall background signal that prevented the clear elucidation of the RNA structure. SHAPE analysis indicated the presence of non-structured, single-stranded flexible regions throughout satCMV TA-Tb with prominent flexible stretches located around nucleotide positions 145 to 200. The positions of these flexible regions were particularly consistent with a secondary structure of satCMV TA-Tb predicted by mfold software v.2.3, which consisted of five 5′- and 3′-proximal stem-loops and one internal large multi-branched stem-loop. Sequence alignment and secondary RNA structure prediction of other satCMV sequences that are phylogenetically the same group with satCMV TA-Tb also suggested the presence of 5′- and 3′-proximal stem-loop structures. Our data provide the structural basis for elucidating the mechanism by which satCMV TA-Tb regulates the pathogenicity and replication of its helper virus.

Keywords: SHAPE; RNA structure probing; satellite RNA; cucumber mosaic virus

check for updates

Citation: Liu, Z.; Cao, X.; Yu, C.; Yuan, X. Probing the RNA Structure of a Satellite RNA of Cucumber Mosaic Virus Using SHAPE Method. *Agronomy* **2023**, *13*, 1990. https://doi.org/10.3390/agronomy13081990

Academic Editor: Xueping Zhou

Received: 15 June 2023
Revised: 12 July 2023
Accepted: 26 July 2023
Published: 27 July 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

RNAs perform many fundamental biological roles, often through interaction with other RNAs and proteins [1]. RNAs usually form structures through internal base-pairing. These include secondary and tertiary structures, which are often difficult to characterize thoroughly based on sequence alone. Characterizing the structure of RNA structure could provide an understanding of how certain RNAs perform their specific biological functions. For example, single-stranded regions often serve as landing pads for proteins or enzymes. Loops have been shown to be critical in the recognition of small molecules as well as in controlling the formation of long-distance interactions within complex RNAs. Extended stem structures have been implicated in RNA-based diseases [2]. Therefore, a thorough understanding of the structures of RNAs is important. Computational methods can be used to predict RNA structures [3], although many factors prevent the accurate prediction of RNA structure based on sequence [4]. Furthermore, there are important short RNAs for which the prediction accuracy is very low [5].

Accurate RNA secondary structure information is essential for deducing structure-function relationships in RNA molecules. Local nucleotide RNA flexibility can be measured at the vast majority of positions using the selective 2′-hydroxyl acylation analyzed by the primer extension (SHAPE) method [6,7]. The SHAPE method enables the thorough examination of RNA structure because all RNA nucleotides, except for some post-transcriptionally modified RNAs, carry a 2′-hydroxyl group, which is targeted for modification by this method. The SHAPE method can be applied to various types of RNA for analyzing RNA structure in vitro as well as in vivo [2,8–11]. The principle of the SHAPE method is to analyze the base flexibility of target RNA via selective 2′ hydroxyl acylation [12]. By careful choice of SHAPE reagent and experimental design, nucleotide flexibilities can be compared to other sequence-based analyses [13]. The 2′-hydroxyl group of unpaired bases in the RNA conformation can be modified by electrophilic reagents, such as N-methylisatoic anhydride (NMIA) or other reagents [6,14]. The subsequent primer extension reaction terminates at the base modified by the electrophilic reagent, thus mapping the free bases in RNA [6,7,15].

A complete set of SHAPE experiments consists of sequencing reactions and SHAPE reactions. The sequencing reactions indicate base positions, while the SHAPE reactions include NMIA-added and NMIA-free reactions, which are used to probe RNA structure by mapping base flexibility. The NMIA-free reaction is a negative control, which presents background signals influenced by factors other than the NMIA reagent. The NMIA-added reaction is the RNA probing reaction and maps the NMIA-modified sites based on the termination site of reverse transcription, which indicates the position of unpaired bases in the target RNA. The structure of the target RNA is probed by comparing signals from the NMIA-added and NMIA-free reactions [6,13]. An NMIA-free reaction that results in weak or no signals is essential for SHAPE analysis. However, some factors can cause strong signals in the NMIA-free reaction. These factors include RNA quantity/integrity and special RNA structures, which may cause early termination during reverse transcription [6,16].

The SHAPE method has been widely used to analyze RNA structures from various organisms. Genome replication of viruses and viroids includes some important processes such as transcription and translation involving some cis-elements within special RNA structures [4,17–20]. Notably, some viruses are accompanied by satellite RNAs, which can regulate the virulence of their helper viruses through specific RNA cis-elements [21,22]. These satellite RNA cis-elements may form special RNA structures, so understanding RNA structure is essential for elucidating the mechanism by which satellite RNA regulates the replication of its helper virus. Our previous data showed that a satellite RNA of cucumber mosaic virus (satCMV) TA-Tb attenuated its helper viruses [23]. In order to investigate the underlying mechanism of attenuation, an understanding of the satCMV TA-Tb RNA structure is necessary. In this study, we attempted to analyze the RNA structure of satCMV TA-Tb using the SHAPE method. In order to reduce the signal background, some parameters were varied to optimize the SHAPE protocol. Our experiment showed that adjusting the quantity of template RNA in the SHAPE reaction can reduce background signals in the NMIA-untreated sample lane. The secondary RNA structure of the complete satCMV TA-Tb sequence provided by this study will be useful for further studies to elucidate how satCMV affects its helper virus.

2. Materials and Methods

2.1. Preparation of satCMV TA-Tb Transcripts

The plasmid pUC19-satCMV TA-Tb (containing the full-length sequence of GenBank No.: MF142365) was used as template to perform PCR with primers Tb-T7-5′linker-F (5′-attaatacgactcactatagggccttcgggccaagttttgtttgttggagaattg-3′) and Tb-3′linker-R (5′-gaaccgg accgaagcccgatttggatccggcgaaccggatcgagggtcctgtagaggaatg-3′). The PCR product was used as the template for the synthesis of satCMV TA-Tb transcripts in vitro.

For the synthesis of RNA transcripts of satCMV TA-Tb in vitro, 0.2–1.0 μg of template DNA was mixed with 30.0 U T7 RNA polymerase (NEB), 2.0 μL 10× reaction buffer, 2.0 μL

20 mM rNTP, 1.0 µL 100 mM DTT, and 1.0 U RNase inhibitor. The reaction volume was made up to 20.0 µL with ddH_2O, and the reaction was incubated at 37 °C for 2.5 h. Satellite RNA transcripts were extracted using the phenol-chloroform-based method and then RNA was precipitated using 1/9 volume of sodium acetate and 2 volumes of 95% ethanol at −20 °C for 2 h. The RNA pellet was dissolved in 30.0 µL of ddH_2O. The RNA concentration was determined using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

2.2. Primer Labeling for the SHAPE Method

In order to analyze the RNA structure of satCMV TA-Tb (383 nt) by the SHAPE method, four reverse primers were used: Tb-shape 3′linker-R (5′-gaaccggaccgaagcccg-3′, position 269–383 nt); Tb-shape-334-R (5′-gccttagcctctccctgcgtg-3′, position 165–285 nt); Tb-shape-248-R (5′-gccgggttctgctagcaaactg-3′, position nt 85–200); and Tb-shape-140-R (5′-gcggctcgtgggctcccccaag-3′, position 5–95 nt). Primers were labeled at their 5′ end using [γ-32p]-ATP (PerkinElmer). For labeling, [γ-32p]-ATP (30 µCi) was added to 15.0 µL reverse primers (10 µM) and enzyme mix (5 U T4 PNK [NEB], 10 µL 10 × T4 PNK buffer, 71.0 µL ddH_2O) followed by incubating at 37 °C for 1.5 h. Labeled primers were centrifuged through illustra^{TM} Microspin^{TM} G-25 columns (GE HealthCare, Chicago, IL, USA) to remove dissociated radioisotopes. A RAP-RS1 isotope radiometer was used to monitor the isotope signal during the experiment.

2.3. Folding and Modification of RNA Templates

RNA folding: 2.0 µL of satCMV TA-Tb transcripts (1.0, 2.0, or 10.0 µM) was denatured at 95 °C for 3 min followed by cooling on ice for 2 min to open the RNA structure. RNA folding was induced by the addition of 3.0 µL 5× folding buffer (400 mM Tris-HCl pH 8.0, 800 mM NH_4Cl, 55 mM MgOAc) and 10.0 µL ddH_2O. The mixture was incubated at 37 °C for 20 min.

RNA modification: The folded RNAs were divided into two tubes (7.5 µL each). One was treated with 1.5 µL 100.0 mM NMIA (Sigma, Kawasaki, Japan), which was dissolved in dimethyl sulfoxide (DMSO) (Sigma), while the other was treated with 1.5 µL of DMSO as a control. Both reactions were incubated at 37 °C for 40 min. Both treated RNAs (modified and control RNA) were precipitated with 1/9 volumes of 3.0 M sodium acetate and 2.0 volumes of cold 95% ethanol. The modified RNA pellet was dissolved in 10.5 µL of 0.5 × TE buffer (10.0 mM Tris-HCl pH 8.0, 1.0 mM EDTA).

2.4. Primer Extension Reactions

Sequencing reactions: [γ-32p]-ATP labeled primers (3.0 µL) and 0.5 × TE buffer (8.0 µL) were added to 1.0 µL of untreated satCMV TA-Tb RNA (1.0 µM, 2.0 µM or 10.0 µM corresponding to the SHAPE reaction), and primer-template solutions were incubated at 65 °C for 5 min and 35 °C for 5 min, followed by incubation on ice for 1 min. Primer extension was initiated by the addition of 8.0 µL of sequencing ladder mix (4.0 µL 5 × SSIII first strand buffer, 1.0 µL 0.1 M DTT, 1.0 µL 10 mM dNTP, 1.5 µL 10 mM ddTTP, 0.5 µL Superscript III (Invitrogen, Waltham, MA, USA); preheated at 52 °C for 1 min) and incubation at 52 °C in a thermostatic metal bath for 1 h.

SHAPE reactions: [γ-32p]-ATP labeled primers (3.0 µL) were added to modified satCMV TA-Tb RNA (10.5 µL), and the primer-template solutions were incubated at 65 °C for 5 min and 35 °C for 5 min, followed by incubation on ice for 1 min. Primer extension was initiated by the addition of 6.5 µL SHAPE enzyme mix (4.0 µL 5 × SSIII First strand buffer, 1.0 µL 0.1 M DTT, 1.0 µL 10 mM dNTP, 0.5 µL Superscript III; preincubated at 52 °C for 1 min) and incubation at 52 °C for 1 h, resulting in the addition of 2′-O-adducts on the RNA template.

The samples of sequencing and SHAPE reactions were treated with 1.0 µL of 4.0 M NaOH and heated at 95 °C for 5 min to terminate the reaction. Then 10.0 µL of Gel Loading Buffer II (Invitrogen) or acid stop dye was added to samples followed by incubation at

95 °C for 5 min and cooling on ice for 5 min. The acid stop dye consisted of Tris-HCl (1.0 M) and stop dye in a ratio of 4:25. Stop dye consisted of 85% formamide, 3.0 mL 0.5 M EDTA pH8.0, 0.75 mL 20 × TBE, 0.25% (*w*/*v*) bromophenol blue, 0.25% (*w*/*v*) xylene cyanol tracking dye, and ddH_2O up to 30 mL.

2.5. Polyacrylamide Gel Electrophoresis (PAGE) and Autoradiography

Urea-polyacrylamide gels (8%) were prepared and a 6.0 µL sample was loaded per lane. Electrophoresis was carried out for 150–200 min at 2000 V (depending on the size of the template). The gels were dried using a Flat Plate Gel Vacuum Dryer (WD-9410). The gel was exposed to a phosphor screen (GE HealthCare) for 24–48 h, and scanned by a multifunctional laser scanning imager Typhoon FLA 9500 (GE HealthCare).

2.6. RNA Structure Analysis

The intensity of each band was quantified using the semi-automated footprinting analysis (SAFA) software v.1.1 [24]. After quantification, the effect of NMIA modification on the flexibility of each nucleotide was examined by subtracting the intensity of each band in the DMSO lane from that in the NMIA lane, producing a net value for each nucleotide. The top 10% of the net values were selected to generate a mean value, which corresponded to a relative flexibility value of 1. Subsequently, the relative flexibility values of each nucleotide were calculated by dividing their corresponding net values by the mean value. Nucleotides with a relative flexibility value of higher than 0.6 or ranging from 0.3 to 0.6 were adjudged to have medium to high, or low to medium flexibility, respectively. The mfold web server [24] was used to predict secondary structures for satCMV TA-Tb and the secondary structure diagrams of RNA were drawn using the RNA2Drawer softw v.1.1 are according to the protocol [25].

2.7. Phylogenetic and Sequence Analyses

Multiple amino acid sequence alignment of satCMV was calculated by MAFFT version 7 (https://mafft.cbrc.jp/alignment/server/, accessed on 11 July 2023) [26]. Trimming of poorly reliable regions in the alignments was carried out using trimAl version 1.3 (http://phylemon.bioinfo.cipf.es, accessed on 11 July 2023) [27]. The phylogenetic tree was constructed using the MEGA X software with the maximum-likelihood method [28]. Bootstrap analysis was performed using 1000 random samplings. DNAMAN ver 6 software (Lynnon Biosoft, San Ramon, CA 94583, USA) was also used for sequence alignment analysis of satCMV sequences.

3. Results

3.1. The Optimal RNA Template Quantity for the SHAPE Reaction Using satCMV TA-Tb Transcripts

The previously described SHAPE protocols stated that the minimum quantity of RNA template required for the SHAPE reaction is 0.2 pmol and the standard quantity is 1.0 pmol but did not specify an upper limit for the RNA template concentration [6]. When 1.0 pmol of RNA template of satCMV TA-Tb was used, the NMIA-treated ([+] NMIA) sample lane and NMIA-free ([−] NMIA) sample lane showed no signal (Figure 1A). As using 1.0 pmol of RNA template did not produce the desired results, we increased the quantity to 10.0 pmol. The result showed that many similar intense bands appeared in both the (−) NMIA and (+) NMIA lanes (Figure 1B). It is suggested that a high quantity of RNA template can ensure a sufficient SHAPE reaction signal. However, it was necessary to further optimize the SHAPE parameters to reduce the high signals in the (−) NMIA samples.

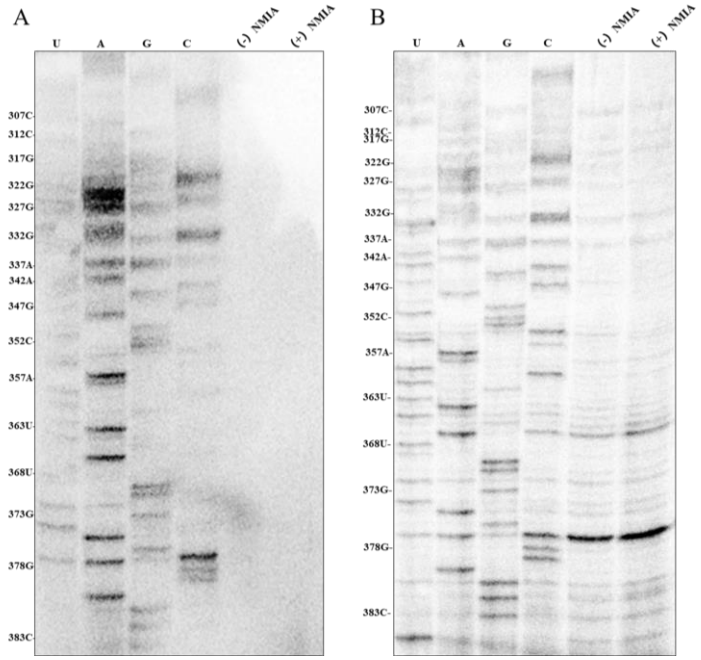


Figure 1. PAGE and autoradiography images from SHAPE analysis of satCMV TA-Tb RNA. (**A**) SHAPE reaction using 1.0 pmol RNA template. The detection range of the satCMV TA-Tb sequence is from nucleotide position 269 to 383. The primer used for reverse transcription was Tb-shape-334-R. U/A/G/C lanes represent four ladder sanger sequencing reactions which were performed using adenine, thymidine, cytosine, and guanine dideoxy terminating nucleotides, respectively. (−) and (+) NMIA lanes represent reactions in the absence and presence of the reagent, respectively. The gel was exposed to a phosphor screen for 48 h. (**B**) SHAPE reaction using 10.0 pmol RNA template. The gel was exposed to a phosphor screen for 12 h. Other conditions are similar to those described for (**A**).

In an attempt to reduce the strong background of the SHAPE reaction in the (-) NMIA lane of satCMV TA-Tb, the denaturation temperature of the RNA template before the reverse transcription reaction was increased from 65 to 95 °C, which is in accordance with the troubleshooting table in the SHAPE protocols [6]. The gel electrophoresis results showed that although the RNA template was denatured at 95 °C, the intense bands in the (−) NMIA lane did not diminish (Supplementary Figure S1, lane 3). In a further test, DMSO was replaced with ddH_2O in the RNA modification (−) NMIA reaction with a denaturation temperature of 65 °C. Intense bands also appeared in the (−) NMIA lane (Supplementary Figure S1, lane 1), indicating that DMSO is not responsible for the early termination of the reverse extension. In conclusion, these measures, i.e., changing the predenaturation temperature and the solvent, could not reduce the strong background signals in the (−) NMIA lane.

As the quantity of template RNA used in the SHAPE reaction with satCMV TA-Tb was 10.0 pmol, approximately 10 times the recommended quantity [6], we reduced the concentration to 2.0 pmol. The electrophoresis of the SHAPE reaction showed that the signal intensity of the bands in the (−) NMIA lane became weaker, and the termination signal of the SHAPE reaction was clearly visible in the (+) NMIA lane (Figure 2A). Quantification of band intensities clearly showed higher intensities in the (+) NMIA lane than in the (−) NMIA lane (Figure 2B) and relative nucleotide reactivities determined by subtracting the (−) NMIA intensities from the (+) NMIA intensities were obtained (Figure 2C). These results indicate that reducing the RNA template to a suitable quantity could diminish the intenseness of bands in the (−) NMIA sample lane, but too low a quantity of RNA template would lead to no visible signal. Therefore, for the structural analysis of satCMV TA-Tb, a 2.0 pmol RNA template is suitable. This observation suggests that for the SHAPE reaction

using short RNA that very easily forms a stable structure, the quantity of RNA template needs to be adjusted to obtain a clear reverse extension termination signal.

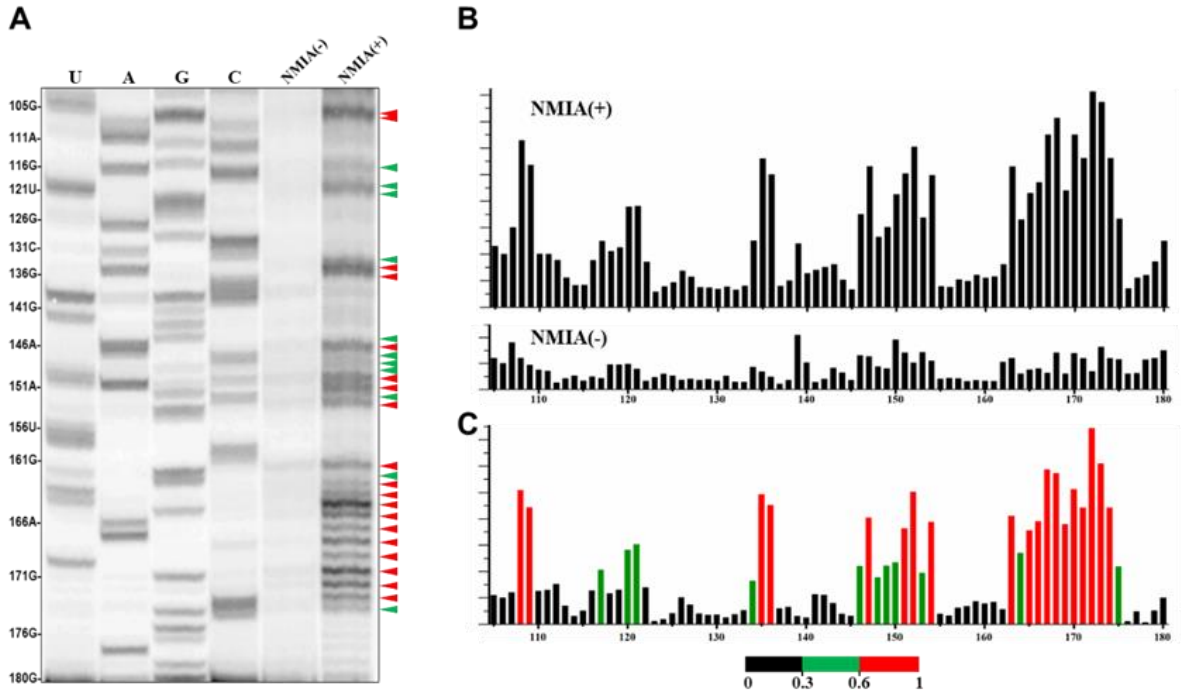


Figure 2. Result of SHAPE analysis of satCMV TA-Tb RNA with the appropriate quantity of RNA template. (**A**) PAGE and autoradiography image from the SHAPE experiment. The SHAPE reaction used 2.0 pmol RNA template. The detection range of the sequence is from nucleotide position 105 to 180. The primer used for reverse transcription was Tb-shape-248-R. U/A/G/C lanes represent four ladder sanger sequencing reactions which were performed using adenine thymidine cytosine and guanine dideoxy terminating nucleotides. Nucleotide numbers indicate the nucleotide position in the satCMV TA-Tb sequence. (−) and (+) NMIA lanes represent reactions in the absence and presence of the reagent, respectively. The nucleotides with high flexibility (reactivity > 0.6) are denoted by a red arrowhead, those with medium flexibility (reactivity from 0.3 to 0.6) are denoted by a green arrowhead, and those with low to no flexibility (reactivity < 0.3) are not marked. The gel was exposed for 12 h to a phosphor screen. (**B**) Band intensities in the (+) and (−) NMIA lanes from panel (**A**). The intensity of each band was quantified using the semi-automated footprinting analysis software SAFA v.1.1. (**C**) Relative nucleotide reactivity is determined by subtracting the (−) NMIA intensities from the (+) NMIA intensities. The coloring scheme for nucleotide reactivity is the same as is shown in (**A**), and those with low to no flexibility (reactivity < 0.3) are black.

3.2. Structure Probing of Complete satCMV TA-Tb Sequence Based on Optimized SHAPE Method

To provide insight into the secondary RNA structure of satCMV TA-Tb, we used an optimized SHAPE method to investigate the flexibility of the individual nucleotides of the complete satCMV TA-Tb sequence. The quantity of RNA template was 2.0 pmol, in accordance with our results described above. The predenaturation temperature of the RNA template used prior to the reverse transcription reaction was 65 °C, and DMSO was used as the solvent. To profile the complete RNA structure of satCMV TA-Tb, four reverse primers, Tb-shape 3′linker-R, Tb-shape-334-R, Tb-shape-248-R, and Tb-shape-140-R were used in the reverse transcription reaction. The reactivity of each nucleotide to NMIA was visualized by phosphorimaging. As shown in Figure 3, the reactive nucleotides, which are denoted by red arrows (medium to high reactivity) and green arrows (low to medium reactivity) are likely to be located in non-structured, single-stranded flexible regions. Notably, prominent flexible stretches located around nucleotide positions 145 to 200 were detected by the SHAPE method. Prediction of the secondary structure of satCMV TA-Tb by mfold with

default conditions resulted in 30 predicted structures. The mfold seeks to bring the two ends together and most of the stable structures it generates do so. It should be noted that the structures are based on those the most stable free energy values under conditions that are not physiological for plants, or animals for that matter. Among those, five predicted structures showed the presence of flexible nucleotide stretches around nucleotides 145 to 200, in agreement with the SHAPE method assessment (Supplementary Figure S2). The positions of flexible nucleotide stretches in one particular predicted structure, which consisted of five 5′- and 3′-proximal stem-loops (SL1, SL2, SL4, SL5, and SL6) and one internal large multi-branched stem-loop (mbSL3, Figure 4), were highly consistent with those determined by the SHAPE method throughout the satCMV TA-Tb sequence. NMIA-reactive nucleotides were positioned at hairpin loops (hL2, hL3-1, hL3-3, hL3-4, hL3-5, hL4 and hL-5), internal loops (iL3, iL4, iL5-1, iL5-2, and iL5-3), multi-branched loops (mbL3-1, mbL3-2 and mbL3-3), bulged loops (bL3-1, bL3-2, bL4 and bL5) and single-stranded regions (ss1, ss2, ss3, and ss5), while only a few NMIA-reactive nucleotides were positioned at stem regions (Figure 4). Considering the low NMIA-reactive signal in hL1, hL 3-3, and hL6 as well as sequence complementarities between hL1 and hL3-3, and between hL1 and hL6, it is possible that the two pairs form kissing loops (kL1 and kL2, Figure 5). This information regarding secondary and tertiary structures of satCMV TA-Tb provided a structural basis for further mechanistic analysis by which satCMV affects virulence or replication of the helper virus.

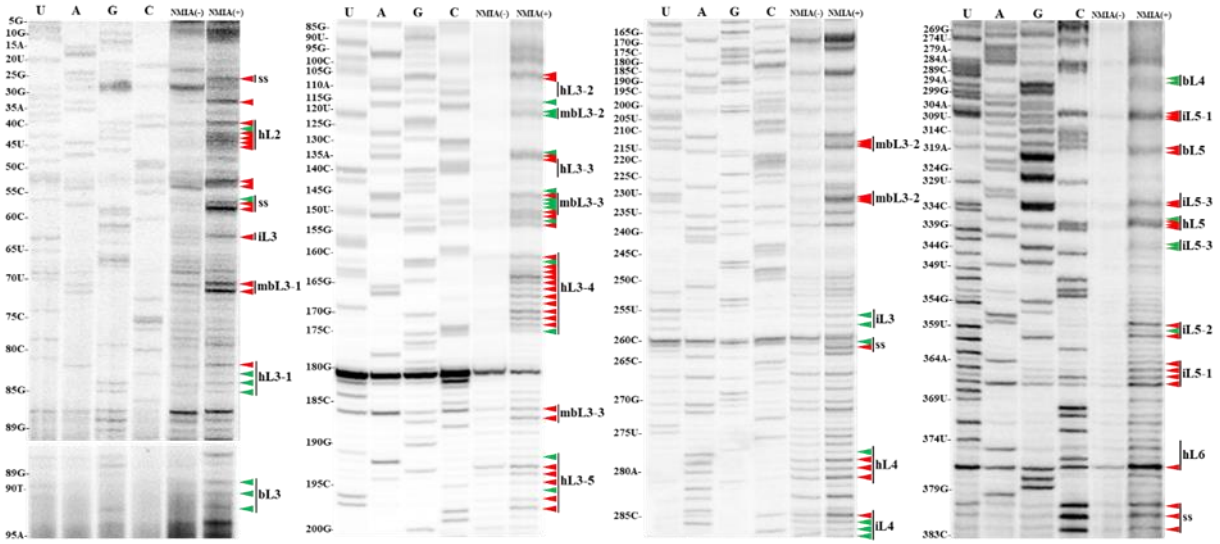


Figure 3. SHAPE analysis on the complete sequence of satCMV TA-Tb. PAGE and autoradiography images from the SHAPE experiment. The reaction used 2.0 pmol RNA template. The detection range of the sequence is from nucleotide position 1 to 383. The primers used for reverse transcription were Tb-shape-140-R, Tb-shape-248-R, Tb-shape-334-R, and Tb-shape 3′linker-R (from left to right). U/A/G/C lanes represent four ladder sanger sequencing reactions which were performed using adenine, thymidine, cytosine, and guanine dideoxy terminating nucleotides, respectively. (−) and (+) NMIA lanes represent reactions in the absence and presence of the reagent, respectively. The nucleotides with high flexibility (reactivity > 0.6) are denoted by a red arrowhead, those with medium flexibility (reactivity from 0.3 to 0.6) are denoted by a green arrowhead, and those with low to no flexibility (reactivity < 0.3) are not marked. Vertical lines to the right of arrowheads indicate the positions of flexible (single-stranded) regions predicted by mfold (see Figure 4). The gel was exposed for 12 or 24 h to a phosphor screen. U/A/G/C lanes are exactly one nucleotide longer than the corresponding NMIA lanes. hL, hairpin loop; bL, bulged loop; iL, internal loop; mbL, multi-branched loop; ss, single-stranded region.

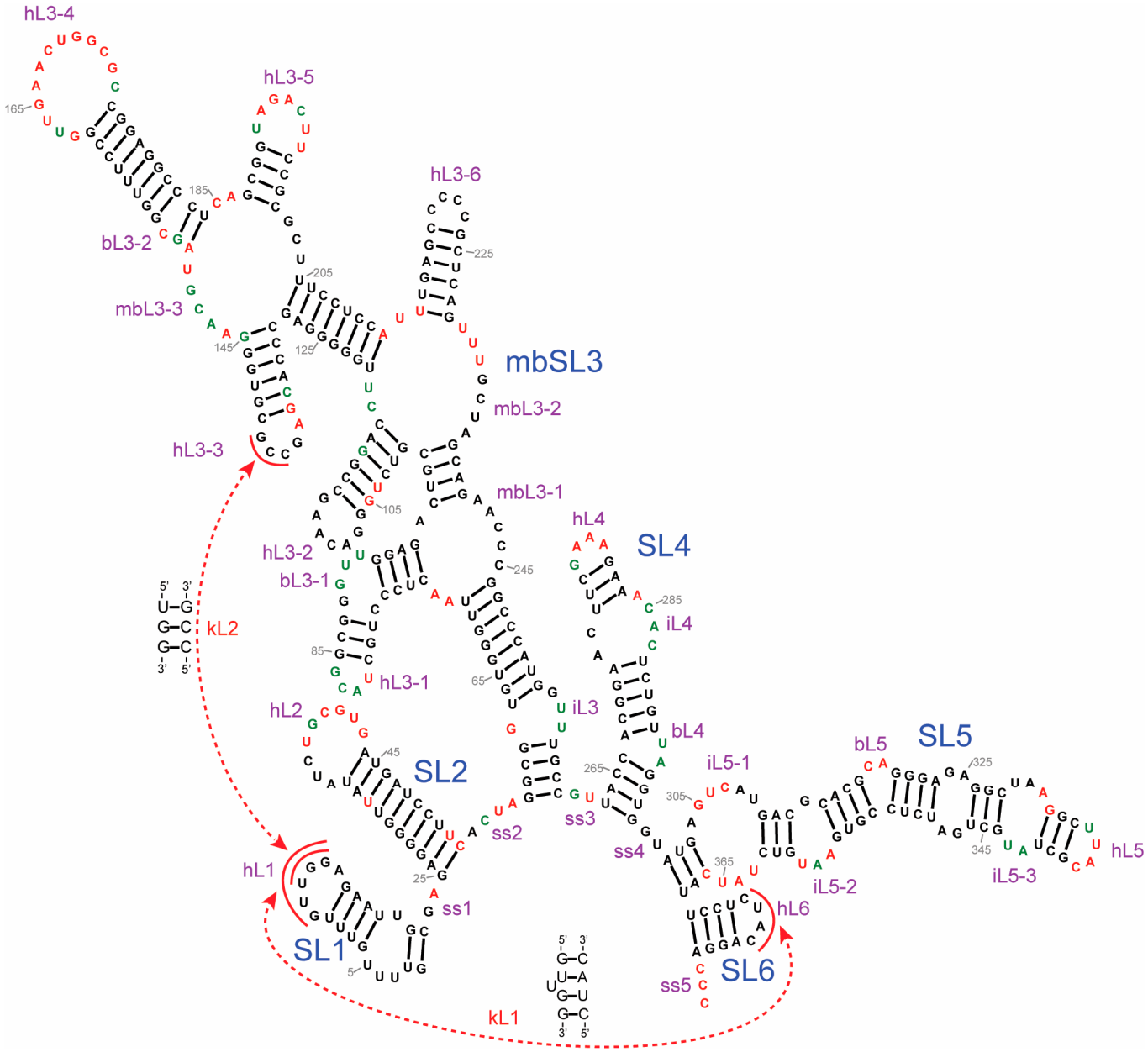


Figure 4. Superposition of relative nucleotide reactivities on a secondary structure model of satCMV TA-Tb predicted by mfold. The color scheme of the nucleotides is the same as that used to represent the reactivity (flexibility) of the nucleotides in Figure 4. SL, stem-loop; mbSL, multi-branched stem-loop; hL, hairpin loop; bL, bulged loop; iL, internal loop; mbL, multi-branched loop; ss, single-stranded region; kL, kissing loops.

3.3. Prediction of Secondary RNA Structures in Other satCMVs

To understand the relationships of satCMV TA-Tb with other satCMVs, a phylogenetic tree was constructed using satCMV TA-Tb and 37 satCMV sequences randomly selected from the sequences publicly available in the database. A partial sequence of CMV RNA3 was used as an outgroup. Phylogenetic analysis showed that satCMVs were clustered into two major groups (groups I and II) (Figure 5). The satCMV TA-Tb was placed in group II together with other satCMVs including satCMV Paf (Accession No. AB570295), N1-04 (JF918974), TSH (DQ249297), TFN (X65455), T43 (D10039), Pz (EF363688), Yn12 (EF363687), XJs1 (DQ070748), and F1 (AB072504), representing the larger satCMVs ranging from 368–405 nt [29,30]. Group I consisted of satCMVs with the usual 332–342 nt size

range and they were separated into two subgroups, I-A and I-B. Subgroup I-B contained pKN-sat (D28559), which was shown to have two insertion regions designated IS-1 and IS-2, relative to other satellite RNAs such as CARNA5 (NC002602) [31], Rs (AF451896) and 57 (D11051), which both contain insertion region IS-1 but not IS-2, while subgroup I-A contained satellite RNAs with no insertion region or only insertion region IS-2 such as Y-satRNA (D00542) (Supplementary Figure S3). Thus, the phylogenetic divergence of satCMVs in group I appears to be in part, due to the presence of insertion region IS-1. It was also noted that satCMVs in subgroup I-A formed two clusters containing either necrogenic or non-necrogenic satCMVs according to their known biological characteristics and/or the presence of necrogenic consensus sequences [32–37] (Figure 5).

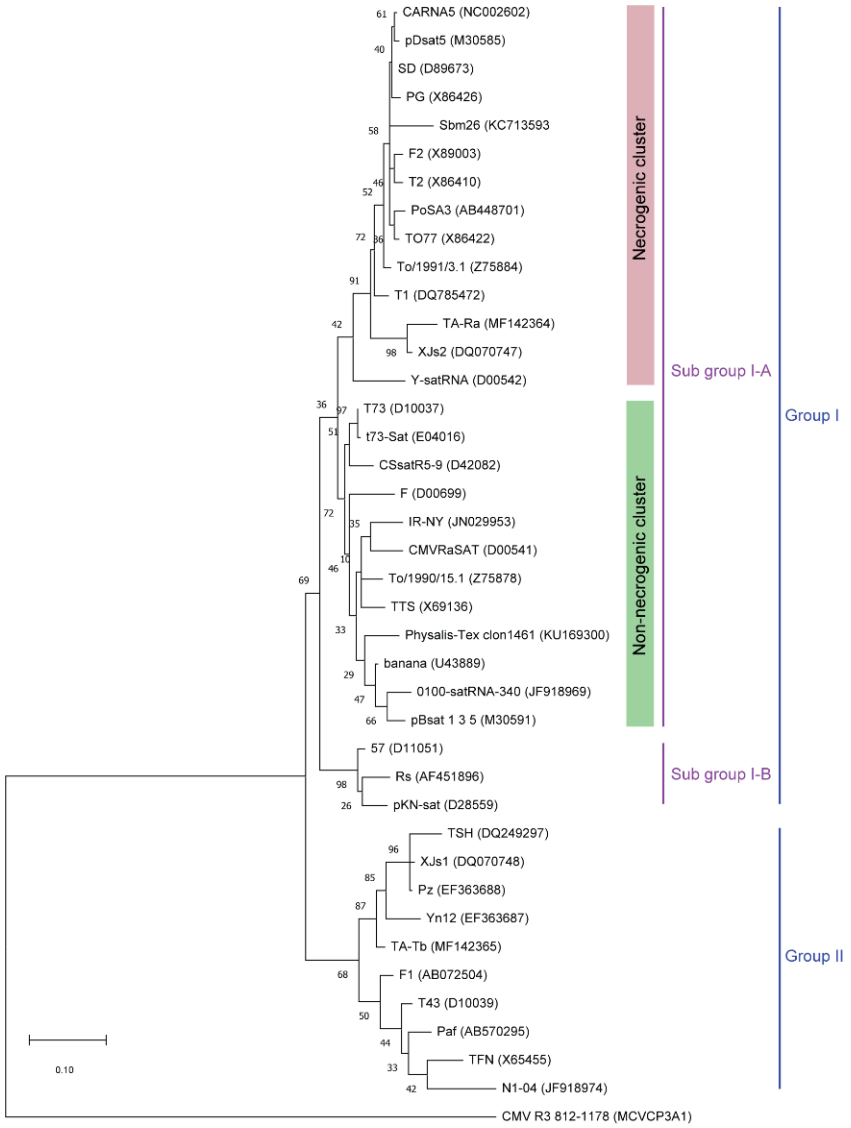


Figure 5. Phylogenetic relationships of satCMV TA-Tb with other selected satCMVs. The tree was constructed using the maximum-likelihood method. The satCMV names are followed by their accession numbers in parentheses. The numbers at the nodes indicate bootstrap support values calculated with bootstrap analysis using 1000 random samplings. The scale bar represents nucleotide substitutions per site.

Sequence alignment of ten satCMV sequences belonging to group II showed a high degree of sequence conservation (Figure 6A). However, it appeared that many sequence variations occur in the regions corresponding to the stem regions of the stem-loops pre-

dicted for satCMV TA-Tb (Figure 6B), suggesting that other satCMVs may have different secondary RNA structures with that of satCMV TA-Tb. Notably, the SL6 structure (see Figure 4) may be conserved due to a high sequence similarity in the 3′ end regions of these satCMVs (Figure 6A). Secondary RNA structure prediction of the portions of satCMV Paf, N1-04, TSH, and Pz sequences that correspond to regions of SL1 and SL2 structure in satCMV TA-Tb using mfold showed that the 5′-proximal region of these satCMVs may also form two stem-loop structures (SL1 and SL2), although they appeared to have different conformation with those predicted for satCMV TA-Tb (Figure 6B). Thus the 5′- and 3′-proximal regions of this satCMV group may similarly form stem-loop structures.

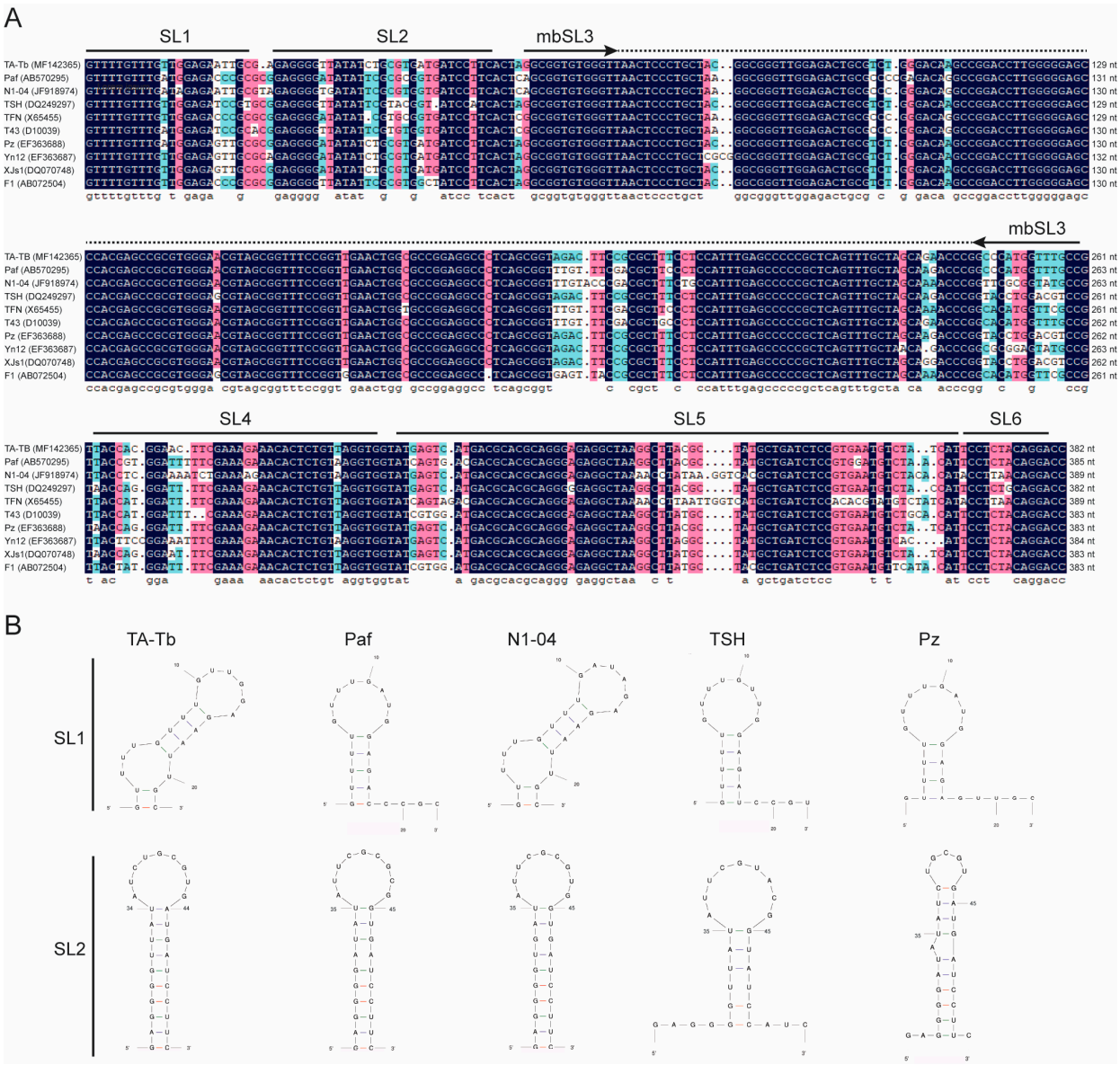


Figure 6. Sequence alignment and secondary RNA structure prediction of satCMVs. (**A**) Sequence alignment of satCMV TA-Tb and phylogenetically closely related satCMVs. The satCMV names are followed by their accession numbers in parentheses. The sequence length (nt) is indicated at the right side of the alignment. The positions of the regions predicted to form stem-loop structures in satCMV TA-Tb are indicated with lines and dashed lines. SL, stem-loop; mbSL, multi-branched stem-loop. (**B**) Secondary RNA structure prediction of the portions of satCMV Paf, N1-04, TSH, and Pz sequences that correspond to regions of SL1 and SL2 structure in satCMV TA-Tb.

4. Discussion

In this study, the problem of the appearance of intense bands in the (−) NMIA lanes in the SHAPE analysis for small satellite RNA structures was solved by adjusting the quantity of template RNA, thus enabling more accurate probing of RNA structure. Notably, the requirement for adjusting the quantity of template RNA was not mentioned in other previous studies using the SHAPE method. The tendency of satellite RNA to form stable structures poses a challenge to SHAPE reaction optimization. Our research group also used the SHAPE method to analyze longer RNAs from the intergenic region of CMV (1000 nt in length). The quantity of RNA template was 1 pmol, which resulted in a clean background signal and, therefore, the structure could be analyzed. Another study using the SHAPE method for the structural analysis of long non-coding RNA showed that, even up to 25 pmol of RNA template, there was no obvious signal in the (−) NMIA lane [38]. This suggests that for short RNAs that very readily form stable structures, the appropriate quantity of RNA template in the SHAPE reaction should be determined experimentally. Although we tested increasing the predenaturation temperature to disrupt the stable RNA fold, the structure may partially reform at the reaction temperature (65 °C). We then tried to reduce the strong background signal by changing other reaction conditions; we found that the production of strong bands in (−) NMIA lanes is not caused by either RNA degradation or structural effects of reverse transcriptase termination in the RNA SHAPE reaction. In fact, the satellite RNA itself has a short sequence, which can readily form a stable structure that is not easily disrupted. Finally, we determined that the concentration of the RNA template is the most critical factor.

We speculate that the background signals may be related to non-specific amplification. The specificity of the hybridization of nucleic acids is affected by many factors, such as the mismatch rate. In addition, the specificity of hybridization is influenced by the ratio of template RNA to primer. It is recommended that the optimal amount of primer is 5–10 times that of the RNA template [39]. By calculation, the concentration of primers used in our experiment was approximately equal to that of the RNA template (10 pmol). Thus, the quantity of primers used is much lower than the reported optimal ratio for hybridization. When the template RNA was reduced, the non-specific signal in the negative control diminished significantly. This is possibly due to the primer and template concentrations reaching the appropriate ratio, at which the hybridization specificity of the reverse reaction is enhanced and the non-specific hybridization and extension are reduced.

As satCMVs usually do not encode a protein, the nucleic acid sequence and/or RNA structure of satCMVs are thought to be the factors that are responsible for its viability and that directly affect the pathogenicity or replication of the helper virus [40,41]. For example, satCMV Y was shown to induce yellow symptoms in tobacco by downregulating the mRNA of tobacco magnesium protoporphyrin chelatase subunit I through the generation of small RNAs from the satellite that is complementary to the mRNA [42]. Previously, the SHAPE method was used to analyze the partial sequence of satCMV T1 [21]. To our knowledge, our study presents the first probing of the RNA structure of a complete satCMV sequence using the SHAPE method. Notably, our SHAPE analysis on satCMV TA-Tb showed high consistency with one particular RNA structure predicted by mfold, which consisted of five 5′- and 3′-proximal stem-loops and one internal large multi-branched stem-loop (Figure 4). Furthermore, secondary RNA structure prediction of other satCMV sequences that are phylogenetically the same group with satCMV TA-Tb suggested the presence of 5′- and 3′-proximal stem-loop structures (Figure 6). Formation of a stable stem-loop (hairpin) structure at 5′ end of satCMVs has been predicted by some studies [31,43–45]. It was demonstrated that possible interaction between the 5′-untranslated region of CMV RNA2 and the hairpin loop of satCMV T1 were not responsible for the suppression of CMV RNA2 accumulation by satCMV T1 [21]. A three-way branched secondary structure predicted and confirmed by SHAPE data to form in the internal region of satCMV T1 was shown to be critical for the helper virus inhibition and satCMV T1 viability in the plant [21]. Previously, a study also employed the combination of a prediction program and

the SHAPE method to investigate the structure of a plant viral satellite RNA. A series of distinct stem-loop structures including the tertiary interactions were predicted to form in 5′-terminal, central, and 3′-terminal regions of satellite RNA of cymbidium ringspot virus (sat-Cym), and mutational analysis indicated the critical role of these structures for the accumulation of sat-Cym [46]. The 3′-untranslated region (UTR) of plant umbraviruses is known to contain 3′ cap-independent translation enhancers (3′ CITEs) characterized by the formation of distinct secondary structures such as the barley yellow dwarf virus-like translation element, the panicum mosaic virus-like translational enhancer, and the T-shaped structure [20,47,48]. Supported by SHAPE analysis, recent studies identified the presence of a conserved hairpin in the 3′ UTRs located upstream of 3′ CITEs in umbraviruses and mutational analysis indicated that this conserved hairpin is critical for viral accumulation and efficient translation of gRNA and sgRNA2 [49].

5. Conclusions

Many sequence regions within viral RNA genomes or non-coding RNA species associated with plant viruses commonly function through the formation of particular RNA structures. The SHAPE method corroborated with mutational analysis seems to be reliable for the deeper study of RNA structures with important biological roles. Further mutational analyses are necessary to elucidate the role of secondary and tertiary RNA structures in satCMV TA-Tb for regulation of the accumulation and pathogenicity of helper virus as well as the viability of satCMV TA-Tb.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy13081990/s1, Figure S1: PAGE and autoradiography image from SHAPE analysis of satCMV TA-Tb RNA with different predenaturation temperatures and solvents. Figure S2: Different secondary structures of satCMV TA-Tb predicted by mfold. Figure S3: Sequence alignment of several satCMVs belonging to sub group I-A and I-B.

Author Contributions: Investigation, Z.L. and X.C.; Resources, C.Y.; Writing—original draft, Z.L.; Writing—review & editing, X.C. and X.Y.; Funding acquisition, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (32072382, 32001867, 31872638) and the Natural Science Foundation of Shandong Province (ZR2020QC129).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

Acknowledgments: We thank Ida Bagus Andika for critically reading of the manuscript.

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

References

1. Sharp, P.A. The centrality of RNA. *Cell* **2009**, *136*, 577–580. [CrossRef] [PubMed]
2. Spitale, R.C.; Flynn, R.A.; Torre, E.A.; Kool, E.T.; Chang, H.Y. RNA structural analysis by evolving SHAPE chemistry. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 867–881. [CrossRef]
3. Washietl, S.; Will, S.; Hendrix, D.A.; Goff, L.A.; Rinn, J.L.; Berger, B.; Kellis, M. Computational analysis of noncoding RNAs. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 759–778. [CrossRef] [PubMed]
4. Deigan, K.E.; Li, T.W.; Mathews, D.H.; Weeks, K.M. Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 97–102. [CrossRef] [PubMed]
5. Ding, F.; Sharma, S.; Chalasani, P.; Demidov, V.V.; Broude, N.E.; Dokholyan, N.V. Ab initio RNA folding by discrete molecular dynamics: From structure prediction to folding mechanisms. *RNA* **2008**, *14*, 1164–1173. [CrossRef]
6. Wilkinson, K.A.; Merino, E.J.; Weeks, K.M. Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): Quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **2006**, *1*, 1610–1616. [CrossRef]
7. Merino, E.J.; Wilkinson, K.A.; Coughlan, J.L.; Weeks, K.M. RNA structure analysis at single nucleotide resolution by selective 2 ′-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **2005**, *127*, 4223–4231. [CrossRef]
8. Leppek, K.; Das, R.; Barna, M. Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 158–174. [CrossRef]

9. Spitale, R.C.; Flynn, R.A.; Zhang, Q.C.; Crisalli, P.; Lee, B.; Jung, J.-W.; Kuchelmeister, H.Y.; Batista, P.J.; Torre, E.A.; Kool, E.T. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **2015**, *519*, 486–490. [CrossRef]
10. Lucks, J.B.; Mortimer, S.A.; Trapnell, C.; Luo, S.; Aviran, S.; Schroth, G.P.; Pachter, L.; Doudna, J.A.; Arkin, A.P. Multiplexed RNA structure characterization with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11063–11068. [CrossRef]
11. Ding, Y.; Tang, Y.; Kwok, C.K.; Zhang, Y.; Bevilacqua, P.C.; Assmann, S.M. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **2014**, *505*, 696–700. [CrossRef] [PubMed]
12. Siegfried, N.A.; Busan, S.; Rice, G.M.; Nelson, J.A.; Weeks, K.M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **2014**, *11*, 959–965. [CrossRef] [PubMed]
13. Mortimer, S.A.; Weeks, K.M. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc.* **2007**, *129*, 4144–4145. [CrossRef] [PubMed]
14. Smola, M.J.; Rice, G.M.; Busan, S.; Siegfried, N.A.; Weeks, K.M. Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* **2015**, *10*, 1643–1669. [CrossRef] [PubMed]
15. Wilkinson, K.A.; Merino, E.J.; Weeks, K.M. RNA SHAPE chemistry reveals nonhierarchical interactions dominate equilibrium structural transitions in tRNAAsp transcripts. *J. Am. Chem. Soc.* **2005**, *127*, 4659–4667. [CrossRef]
16. Wilkinson, K.A.; Vasa, S.M.; Deigan, K.E.; Mortimer, S.A.; Giddings, M.C.; Weeks, K.M. Influence of nucleotide identity on ribose 2′-hydroxyl reactivity in RNA. *RNA* **2009**, *15*, 1314–1321. [CrossRef]
17. Fischer, J.W.; Busa, V.F.; Shao, Y.; Leung, A.K. Structure-mediated RNA decay by UPF1 and G3BP1. *Mol. Cell* **2020**, *78*, 70–84.e6. [CrossRef]
18. Geng, G.; Yu, C.; Li, X.; Yuan, X. A unique internal ribosome entry site representing a dynamic equilibrium state of RNA tertiary structure in the 5′-UTR of Wheat yellow mosaic virus RNA1. *Nucleic Acids Res.* **2020**, *48*, 390–404. [CrossRef]
19. Poller, W.; Dimmeler, S.; Heymans, S.; Zeller, T.; Haas, J.; Karakas, M.; Leistner, D.-M.; Jakob, P.; Nakagawa, S.; Blankenberg, S. Non-coding RNAs in cardiovascular diseases: Diagnostic and therapeutic perspectives. *Eur. Heart J.* **2018**, *39*, 2704–2716. [CrossRef]
20. Gao, F.; Kasprzak, W.K.; Szarko, C.; Shapiro, B.A.; Simon, A.E. The 3′ untranslated region of Pea enation mosaic virus contains two T-shaped, ribosome-binding, cap-independent translation enhancers. *J. Virol.* **2014**, *88*, 11696–11712. [CrossRef]
21. He, L.; Wang, Q.; Gu, Z.; Liao, Q.; Palukaitis, P.; Du, Z. A conserved RNA structure is essential for a satellite RNA-mediated inhibition of helper virus accumulation. *Nucleic Acids Res.* **2019**, *47*, 8255–8271. [CrossRef]
22. Kwon, S.-J.; Chaturvedi, S.; Rao, A. Repair of the 3′proximal and internal deletions of a satellite RNA associated with Cucumber mosaic virus is directed toward restoring structural integrity. *Virology* **2014**, *450*, 222–232. [CrossRef] [PubMed]
23. Cao, X.; Liu, S.; Yu, C.; Li, X.; Yuan, X. A new strategy of using satellite RNA to control viral plant diseases: Post-inoculation with satellite RNA attenuates symptoms derived from pre-infection with its helper virus. *Plant Biotechnol. J.* **2019**, *17*, 1856. [CrossRef] [PubMed]
24. Laederach, A.; Das, R.; Vicens, Q.; Pearlman, S.M.; Brenowitz, M.; Herschlag, D.; Altman, R.B. Semiautomated and rapid quantification of nucleic acid footprinting and structure mapping experiments. *Nat. Protoc.* **2008**, *3*, 1395–1401. [CrossRef]
25. Johnson, P.Z.; Kasprzak, W.K.; Shapiro, B.A.; Simon, A.E. RNA2Drawer: Geometrically strict drawing of nucleic acid structures with graphical structure editing and highlighting of complementary subsequences. *RNA Biol.* **2019**, *16*, 1667–1671. [CrossRef] [PubMed]
26. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]
27. Tárraga, J.; Medina, I.; Arbiza, L.; Huerta-Cepas, J.; Gabaldon, T.; Dopazo, J.; Dopazo, H. Phylemon: A suite of web tools for molecular evolution, phylogenetics and phylogenomics. *Nucleic Acids Res.* **2007**, *35*, W38–W42. [CrossRef]
28. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547. [CrossRef] [PubMed]
29. Fraile, A.; Garcia-Arenal, F. Secondary structure as a constraint on the evolution of a plant viral satellite RNA. *J. Mol. Biol.* **1991**, *221*, 1065–1069. [CrossRef] [PubMed]
30. Garcia-Arenal, F.; Palukaitis, P. Structure and functional relationships of satellite RNAs of cucumber mosaic virus. *Satell. Defective Viral RNAs* **1999**, *239*, 37–63.
31. Hidaka, S.; Hanada, K. Structural features unique to a new 405-nucleotide satellite RNA of cucumber mosaic virus inducing tomato necrosis. *Virology* **1994**, *200*, 806–808. [CrossRef] [PubMed]
32. Masuta, C.; Takanami, Y. Determination of sequence and structural requirements for pathogenicity of a cucumber mosaic virus satellite RNA (Y-satRNA). *Plant Cell* **1989**, *1*, 1165–1173. [PubMed]
33. Crescenzi, A.; Grieco, F.; Gallitelli, D. Nucleotide sequence of a satellite RNA of a strain of cucumber mosaic virus associated with a tomato fruit necrosis. *Nucleic Acids Res.* **1992**, *20*, 2886. [CrossRef] [PubMed]
34. Grieco, F.; Lanave, C.; Gallitelli, D. Evolutionary dynamics of cucumber mosaic virus satellite RNA during natural epidemics in Italy. *Virology* **1997**, *229*, 166–174. [CrossRef]
35. Devic, M.; Jaegle, M.; Baulcombe, D. Cucumber mosaic virus satellite RNA (strain Y): Analysis of sequences which affect systemic necrosis on tomato. *J. Gen. Virol.* **1990**, *71*, 1443–1449. [CrossRef]

36. Stanković, I.; Vučurović, A.; Zečević, K.; Petrović, B.; Nikolić, D.; Delibašić, G. Characterization of cucumber mosaic virus and its satellite RNAs associated with tomato lethal necrosis in Serbia. *Eur. J. Plant Pathol.* **2021**, *160*, 301–313. [CrossRef]
37. Sleat, D.E.; Palukaitis, P. Site-directed mutagenesis of a plant viral satellite RNA changes its phenotype from ameliorative to necrogenic. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 2946–2950. [CrossRef]
38. Wang, D.; Yu, C.; Liu, S.; Wang, G.; Shi, K.; Li, X.; Yuan, X. Structural alteration of a BYDV-like translation element (BTE) that attenuates p35 expression in three mild Tobacco bushy top virus isolates. *Sci. Rep.* **2017**, *7*, 4213. [CrossRef]
39. Boorstein, W.R.; Craig, E.A. [25] Primer extension analysis of RNA. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 1989; Volume 180, pp. 347–369.
40. Simon, A.E.; Roossinck, M.J.; Havelda, Z. Plant virus satellite and defective interfering RNAs: New paradigms for a new century. *Annu. Rev. Phytopathol.* **2004**, *42*, 415–437. [CrossRef]
41. García-Arenal, P.P. Cucumoviruses. *Adv. Virus Res.* **2003**, *62*, 241–323.
42. Shimura, H.; Pantaleo, V.; Ishihara, T.; Myojo, N.; Inaba, J.-i.; Sueda, K.; Burgyán, J.; Masuta, C. A viral satellite RNA induces yellow symptoms on tobacco by targeting a gene involved in chlorophyll biosynthesis using the RNA silencing machinery. *PLoS Pathog.* **2011**, *7*, e1002021. [CrossRef] [PubMed]
43. Rodriguez-Alvarado, G.; Roossinck, M.J. Structural Analysis of a Necrogenic Strain of Cucumber Mosaic Cucumovirus Satellite RNAin Planta. *Virology* **1997**, *236*, 155–166. [CrossRef] [PubMed]
44. Bernal, J.J.; García-Arenal, F. Analysis of the in vitro secondary structure of cucumber mosaic virus satellite RNA. *RNA* **1997**, *3*, 1052–1067.
45. REZAIAN, M.A.; WILLIAMS, R.H.; SYMONS, R.H. Nucleotide sequence of cucumber mosaic virus RNA 1: Presence of a sequence complementary to part of the viral satellite RNA and homologies with other viral RNAs. *Eur. J. Biochem.* **1985**, *150*, 331–339. [CrossRef] [PubMed]
46. Ashton, P.; Wu, B.; D'Angelo, J.; Grigull, J.; White, K.A. Biologically-supported structural model for a viral satellite RNA. *Nucleic Acids Res.* **2015**, *43*, 9965–9977. [CrossRef]
47. Ilyas, M.; Du, Z.; Simon, A.E. Opium poppy mosaic virus has an Xrn-resistant, translated subgenomic RNA and a BTE 3′ CITE. *J. Virol.* **2021**, *95*, 10–1128. [CrossRef]
48. Gao, F.; Simon, A.E. Differential use of 3′CITEs by the subgenomic RNA of Pea enation mosaic virus 2. *Virology* **2017**, *510*, 194–204. [CrossRef]
49. Bera, S.; Ilyas, M.; Mikkelsen, A.A.; Simon, A.E. Conserved Structure Associated with Different 3′ CITEs Is Important for Translation of Umbraviruses. *Viruses* **2023**, *15*, 638. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.