

## Article

# Structural 3D Domain Reconstruction of the RNA Genome from Viruses with Secondary Structure Models

Simón Poblete <sup>1,2,\*</sup>  and Horacio V. Guzman <sup>3,\*</sup> <sup>1</sup> Instituto de Ciencias Físicas y Matemáticas, Universidad Austral de Chile, Valdivia 5091000, Chile<sup>2</sup> Chile and Computational Biology Lab, Fundación Ciencia & Vida, Santiago 7780272, Chile<sup>3</sup> Department of Theoretical Physics, Jožef Stefan Institute, SI-1000 Ljubljana, Slovenia

\* Correspondence: simon.poblete@uach.cl (S.P.); horacio.guzman@ijs.si (H.V.G.)

**Abstract:** Three-dimensional RNA domain reconstruction is important for the assembly, disassembly and delivery functionalities of a packed proteinaceous capsid. However, to date, the self-association of RNA molecules is still an open problem. Recent chemical probing reports provide, with high reliability, the secondary structure of diverse RNA ensembles, such as those of viral genomes. Here, we present a method for reconstructing the complete 3D structure of RNA genomes, which combines a coarse-grained model with a subdomain composition scheme to obtain the entire genome inside proteinaceous capsids based on secondary structures from experimental techniques. Despite the amount of sampling involved in the folded and also unfolded RNA molecules, advanced microscope techniques can provide points of anchoring, which enhance our model to include interactions between capsid pentamers and RNA subdomains. To test our method, we tackle the satellite tobacco mosaic virus (STMV) genome, which has been widely studied by both experimental and computational communities. We provide not only a methodology to structurally analyze the tertiary conformations of the RNA genome inside capsids, but a flexible platform that allows the easy implementation of features/descriptors coming from both theoretical and experimental approaches.



**Citation:** Poblete, S.; Guzman, H.V. Structural 3D Domain Reconstruction of the RNA Genome from Viruses with Secondary Structure Models. *Viruses* **2021**, *13*, 1555. <https://doi.org/10.3390/v13081555>

Academic Editor: Hironori Sato

Received: 28 June 2021

Accepted: 30 July 2021

Published: 6 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Keywords:** RNA secondary structure; RNA tertiary structure; RNA genome; STMV

## 1. Introduction

RNA genome is an essential part of RNA viruses. In particular, non-enveloped ones convey a direct relationship between capsid proteins and RNA molecules, which was the genesis of several research routes on the assembly, disassembly and delivery functionalities of RNA inside proteinaceous capsids [1–12]. RNA is a polyelectrolyte with highly versatile conformations; with self-associating base pairs, it creates a variety of three-dimensional circuitous structural motifs. The organization of the RNA genome is important for many aspects of virus assembly, particularly in steps where different regions of the genome, such as stems and loops, are thought to interact/bind with capsid proteins. However, despite the relevance of its structural organization, there is little known about how it is acquired: first, the formation of an unambiguous secondary structure as patterns of base pairs, followed by the more challenging question of how the tertiary structure is formed [13–20]. In recent years, several chemical probing tools [21–26] have tackled possible ways to measure/detect RNA secondary structures; nowadays, they offer (with a reasonable probability) very sophisticated schemes able to propose whole genomes packed inside virus capsids [27], cells [23], among other biological systems [20]. Those results provide valuable information of the secondary structures, which shall be used as a starting point for further computational reconstruction of three-dimensional structures [19,28–30]. The latter constitutes an even more challenging task by considering the conformational softness of RNA molecules and its direct implications in the structural heterogeneity of RNA interactions. Thus, the interactions and binding between capsid proteins inside its biological

environment can be drastically affected. In this context, besides the secondary structure provided by chemical probing methods, collecting further microscopic information from biophysical techniques [20,31] is crucial to interpret and improve 3D RNA computational modeling. Here, we propose a method that combines a coarse-grained model of RNA with a subdomain composition scheme to obtain the entire genome inside and outside proteinaceous capsids based on secondary structures from chemical probing and structural refinement by biophysical methods (X-ray crystallography, cryoelectron microscopy—CryoEM—among others). Our approach allows the rearrangement of large domains (in the order of hundreds of nucleotides) by using Steered Monte Carlo (SMC) simulations, which include a parsimonious supervision scheme, to generate three-dimensional models from the sole knowledge of the sequence and secondary structure. We tested our method with an extensively studied system, namely, the genome of the STMV. The latter was computationally tackled via molecular simulations; hence, specific structural microscopic arrangements can be also tested for the sake of evaluating our method feasibility [32–35].

The article comprises several sections, starting with a short description of the method, including the steps for the subdomain reconstruction process. Subsequently, we test the method with the input from the STMV secondary structures proposed by Weeks and co-workers [27], namely, in virio and in vitro models. In a further section, we tackle a particular subdomain of the STMV genome to locally disambiguate between the two most likely structures around a five-fold rotational axis, and briefly discuss the quantities to observe in order to distinguish the optimal structural arrangement. This work concludes with a discussion on how to further couple the present method with computational/theoretical and experimental approaches, and how to address possible issues in the reconstruction of large RNA structures.

## 2. Materials and Methods

### 2.1. Coarse-Grained Model

The SPlit and conQueR model (SPQR) is a coarse-grained representation [36] of RNA, originally designed for structure prediction [37]. Under this scheme, each nucleotide is treated as an anisotropic particle and a spherical bead, which stand for the nucleoside and the phosphate group respectively. The interactions among the coarse-grained nucleotides consist of excluded volume, canonical and non-canonical base pairs, stacking and base-phosphate interactions, with a backbone connectivity that specifies the sugar pucker and the glycosidic bond angle conformation, whereby the excluded volume interaction corresponds to an energy term that prevents the overlap between different coarse-grained beads. In the present work, however, we start with a reduced interaction set, which consists solely of the excluded volume and backbone connectivity, in order to reduce the simulation time, due to the system size. This model [37] has its own energy units, which is denoted by  $\epsilon_s$  in the following.

In order to enforce a particular geometry on a set of nucleotides, we perform SMC simulations, which are constant temperature Monte Carlo (MC) simulations with additional energy restraints. These simulations minimize the  $\mathcal{E}$ RMSD metric between the simulated RNA fragment and a target structure: a quantity which has been shown to be suitable for comparing nucleic acids structures [17]. The  $\mathcal{E}$ RMSD between two structures is defined as a function of the set of vectors  $\{\mathbf{r}_{ij}\}$  and  $\{\mathbf{r}_{ij}^t\}$ , obtained from the simulated and target structures, respectively. Each of these vectors is calculated from the position of the nucleobase  $i$  with respect to a reference frame located at the origin of nucleobase  $j$  with its orientation. The rescaled vectors  $\tilde{\mathbf{r}} = (r_x/a, r_y/b, r_z/c)$ , with  $a = b = 5 \text{ \AA}$  and  $c = 3 \text{ \AA}$ , are used to define the vectors  $\mathbf{G}(\tilde{\mathbf{r}}_{ij}) = (\sin(\gamma\tilde{r}_x/\tilde{r}), \sin(\gamma\tilde{r}_y/\tilde{r}), \sin(\gamma\tilde{r}_z/\tilde{r}), 1 + \cos(\gamma\tilde{r}))\Theta(\tilde{r}_c - \tilde{r})/\gamma$ , where  $\gamma = \pi/\tilde{r}_c$ ,  $\Theta$  is the Heaviside function and  $\tilde{r}_c$  is a cutoff parameter. Focusing on a particular fragment denoted by  $s$ , composed of  $N_s$  nucleotides, the  $\mathcal{E}$ RMSD can be written as follows:

$$E_s = \sqrt{\frac{1}{N_s} \sum_{j,k} \left| \mathbf{G}(\tilde{\mathbf{r}}_{jk,s}) - \mathbf{G}(\tilde{\mathbf{r}}_{jk,s}^t) \right|^2} \quad (1)$$

which allows to define the restraining energy as follows:

$$U_{ss} = \frac{K_{ss}}{2} \sum_s E_s^2 \quad (2)$$

where  $K_{ss}$  is the corresponding harmonic spring constant [17]. In this manner, the relative positions and orientations of the nucleobases belonging to each fragment are enforced to have the values of the target template. In particular, such restraints can be used to impose a secondary structure or to enforce a global structural array between an arbitrary set of nucleotides. This procedure was also used for enforcing other motifs and backmapping RNA structures predicted by SPQR into an all-atom representation, by means of steered-molecular dynamics [37,38]. The value of the  $\mathcal{E}$ RMSD depends on the adimensional cutoff chosen. Therefore, each result involving the  $\mathcal{E}$ RMSD specifies this quantity.

An additional restraint can be exerted on the radius of gyration, which has the following harmonic form:

$$U_{RG} = \frac{1}{2} K_{RG} (R_G - R_G^0)^2 \quad (3)$$

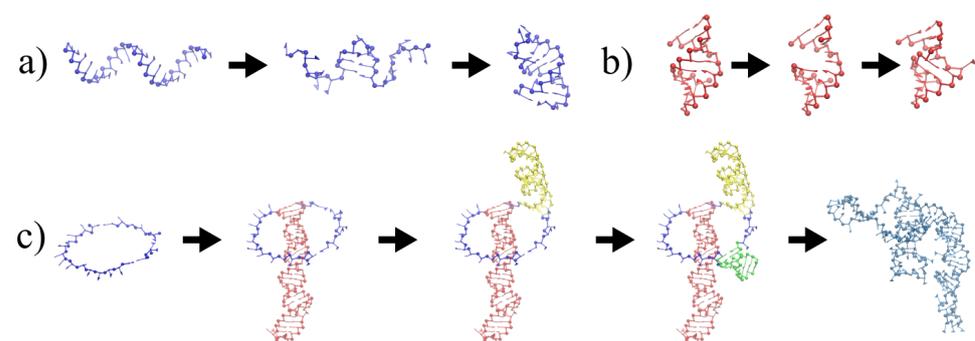
where  $K_{RG}$  is a harmonic spring constant,  $R_G$  is the radius of gyration and  $R_G^0$  is a target value for this quantity.

Finally, an important feature of the method is the use of soft repulsive or attractive energy terms, which are designed to put away colliding coarse-grained nucleotides or to reestablish broken bonds, which allow to easily relax assembled fragments into a more realistic conformation.

## 2.2. Fragment Assembler

Within our method, the secondary structure is given as an input, where we decompose the structure into two type of subdomains: junctions and stem loops. The latter can be composed of one or more sequential complementary double helices, which can be connected by internal loops or be closed by a hairpin.

We proceed to generate models for each stem loop as shown in Figure 1a. The hairpins, or whole stem loops, if they are short enough, are designed by applying secondary structure restraints on the closing pairs, starting from a fully extended conformation. This enforcement has to be performed in short steps made out of 3 pairs at a time to avoid the entanglement under the helix formation in the molecule.



**Figure 1.** Fragment assembly of STMV genome. (a) Reconstruction of a hairpin loop, (b) reconstruction of a stem loop. At the end, a short relaxation simulation ensures that the topology is realistic and that no clashes are present. (c) Assembly on template junction.

Similarly, the size of the stem is increased by exerting the secondary structure restraints on additional unpaired nucleotides, which belong to the hairpin loop. Here, we have empirically determined that the size of such a hairpin loop has to be at most 4 nucleotides. This procedure is necessary in order to prevent the entanglement of the hairpin with itself and with the stem. At the end, the artificially introduced additional pairs are released

from their restraint, and the structure is subsequently relaxed with the original secondary structure.

Each of these simulations is performed with temperature  $T$  such that  $k_B T = 3\epsilon_s$  for 100,000 MC sweeps. Each MC sweep consists of a MC trial move on each of the nucleotides.

For larger stem loops, the procedure is illustrated in Figure 1b. A complementary double helix template is created from the first strand of the stem loop. Additional nucleotides are introduced in order to match the number of nucleotides on both strands in the internal loops, if needed. Then, the sequences are replaced by the original one and additional nucleotides are removed. The whole structure is finally relaxed to ensure the connectivity and to enable additional flexibility in the internal loops.

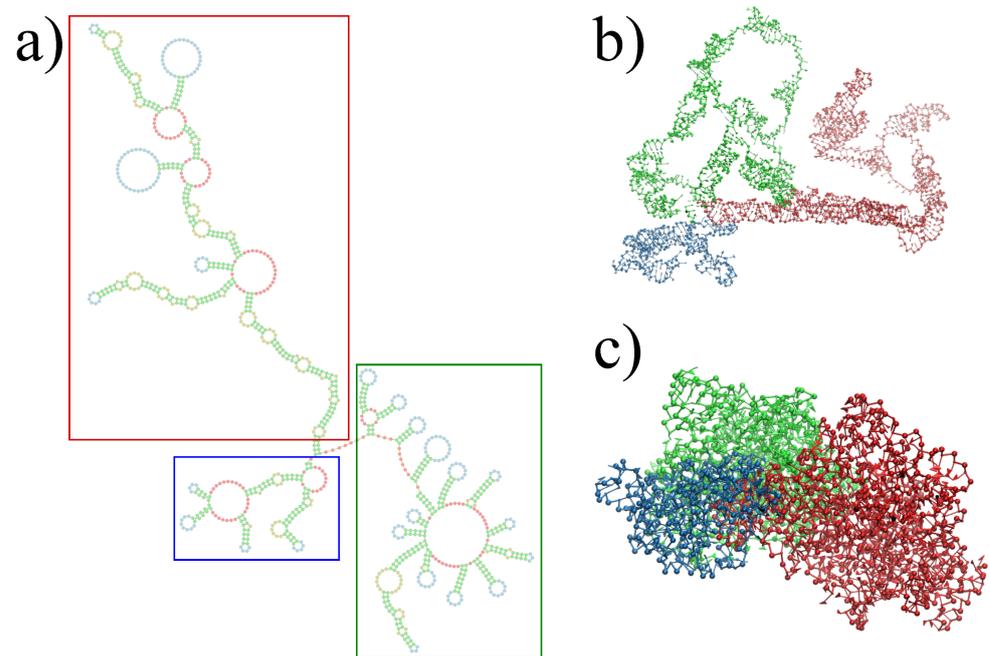
The junctions are generated as ring-like structures, which are composed of strands connected by a closing base pair, where a stem loop is anchored (see Figure 1c). A set of 10 decoys for each junction is generated, simulating by 400,000 MC sweeps at temperature  $9\epsilon_s/k_B T$ , under the imposition of a relatively large radius of gyration (between 40 and 90 Å, depending on the system size) and with  $K_{RG} = 20\epsilon_s/\text{Å}^2$ . Later on, the stem loops are placed on top of their respective closing base-pairs, checking for the possible formation of clashes and knots. Usually, more than one decoy satisfies these conditions, given the extended conformation of the junction ring. This procedure is extremely fast and efficient, given the short size of the junction loops, and it is useful for generating independent geometrical conformations, which are not easy to explore by MC simulations, given the size of the stem loops attached to them in the examined cases. Moreover, given that the in virio structure is likely to be suboptimal in terms of free energy, and that the structural restraints is likely to be applied on the stem loops, it is reasonable to relax and optimize the structure of the junctions in the subsequent step.

Finally,  $\mathcal{E}$ RMSD restraints can enforce the geometrical arrangement of the multiple stem loops predicted by the secondary structure model. For this, we make use of the X-ray crystal structure of STMV (PDB 4OQ9), where 30 stem loops are reported, comprising approximately 59% of the genome, but without specification of their sequence. Several efforts have been made in this direction, although they did employ different methodologies and different assumptions for secondary structure models [32,33,35].

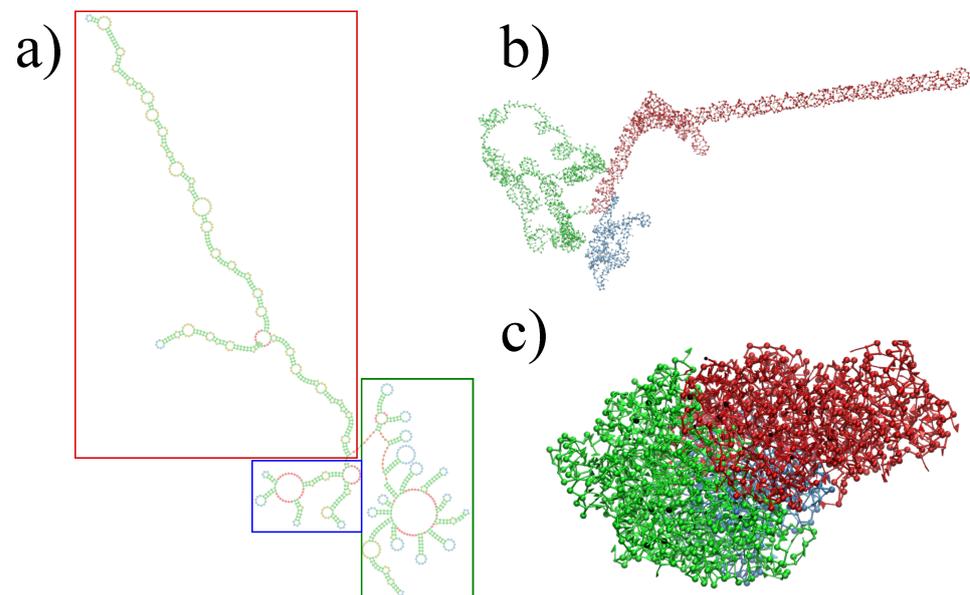
Note also that our fragment assembly procedure could, in principle, be applied to other RNA coarse-grained models [39–43]; however, the exact details on resolution, energy terms and/or other model parameters have to be individually evaluated and are beyond the scope of this article.

### 3. Results

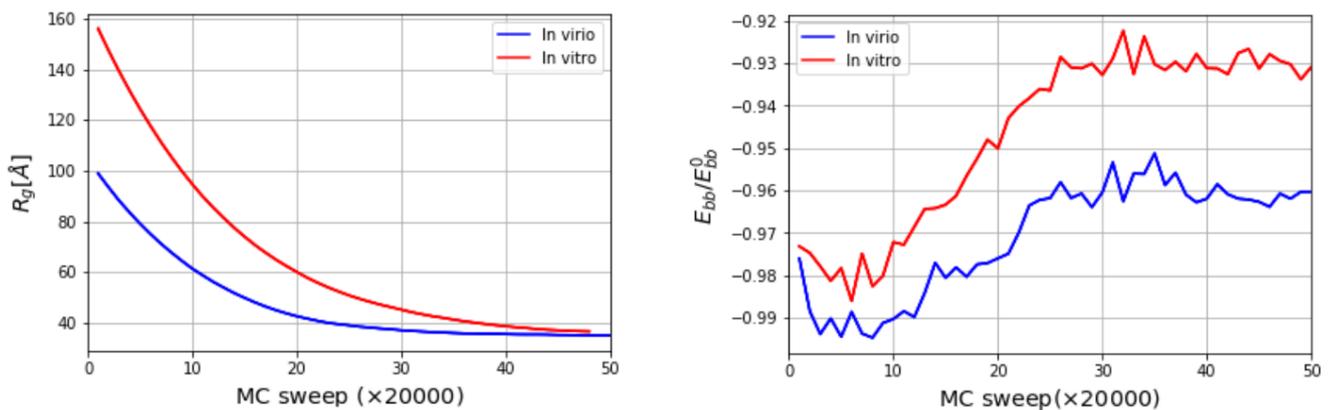
The three-dimensional in virio and in vitro models of the STMV genome are shown in Figures 2 and 3. Each decoy was minimally supervised, obtaining a refined structure with no broken bonds, nor clashes, nor knots between their components. Nevertheless, their extension is far from the dimensions inside a virus, reaching a radius of gyration of 99 Å for the in virio model, which yields a maximum distance of 182 Å between a phosphate group and the center of mass. Clearly, these values are well beyond the radius of the STMV, including its capsid, which is approximately 85 Å. A SMC simulation was performed for minimizing the radius of gyration for 1,000,000 MC sweeps ( $R_G^0 = 10$  Å,  $K_{RG} = 200\epsilon_s/\text{Å}^2$ ), which maintains the secondary structure restraints. We observed that both structures can reach a compact state in a similar amount of time, under the effect of a strong bias on the radius of gyration, as seen in Figure 4. Nevertheless, the backbone energy reaches a higher value for the in vitro structure, which suggests that a higher deformation of the molecule has to take place in order to satisfy the spatial restraint. The  $\mathcal{E}$ RMSD (with a cutoff of 100), which enforces the secondary structure, increases by around 5% in both cases, which expresses that the structure of the restrained stems is minimally affected.



**Figure 2.** Full reconstruction of the STMV in virio genome. (a) Secondary structure proposed by Weeks and co-workers [27], with colored rectangles distinguishing each domain: blue for 5' domain, red for central T domain and green for 3' domain. (b) Reconstruction after assembly and energy minimization and (c) after SMC simulation with a minimum radius of gyration.

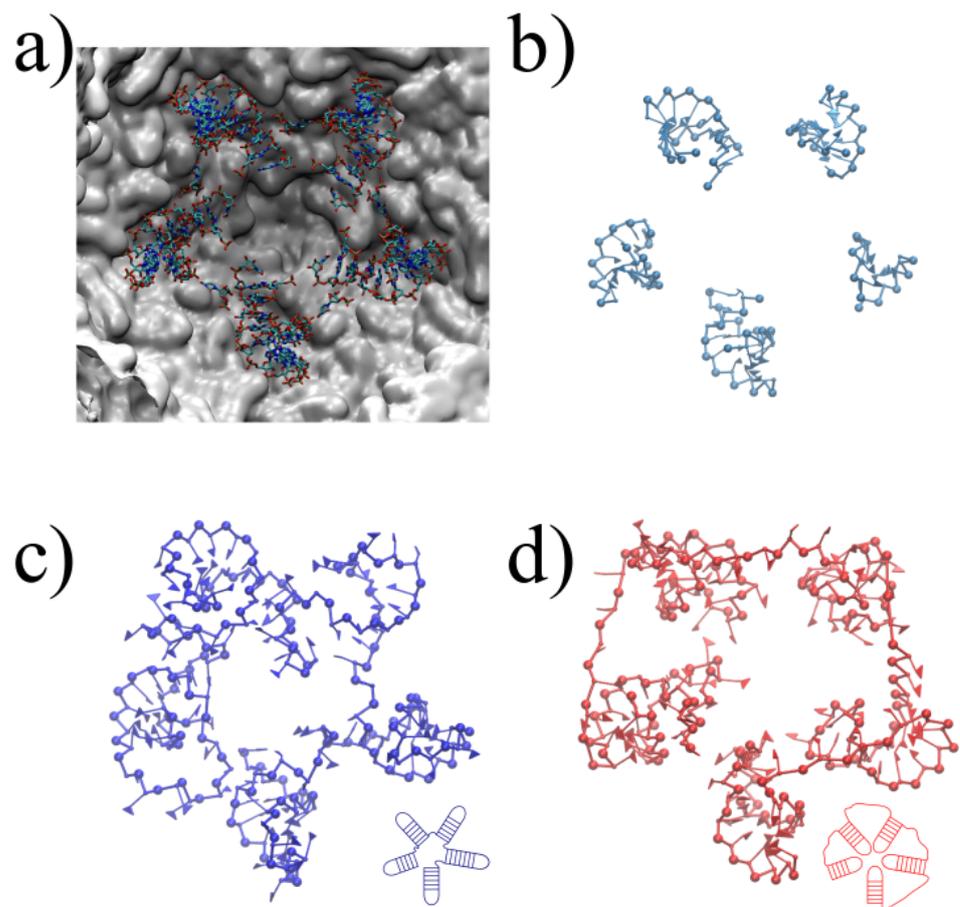


**Figure 3.** Full reconstruction of the STMV in vitro genome. (a) Secondary structure proposed by Weeks and co-workers [27], with colored rectangles distinguishing each domain: blue for 5' domain, red for central T domain and green for 3' domain. (b) Reconstruction after assembly and energy minimization and (c) after SMC simulation with a minimum radius of gyration.

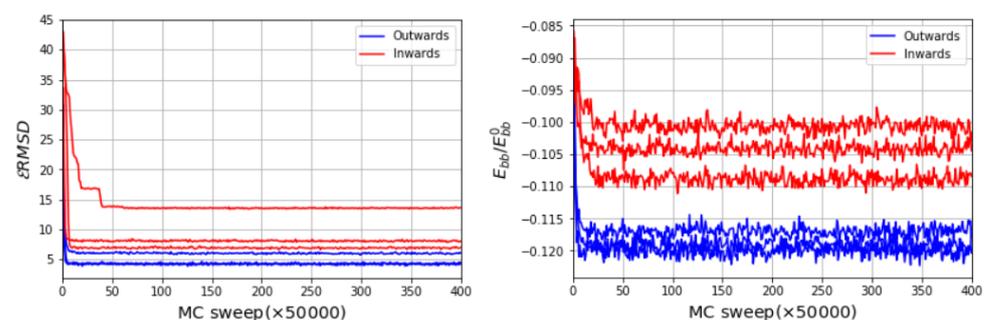


**Figure 4.** Radius of gyration  $R_g$  as a function of MC sweep (**left**) and normalized backbone energy (**right**).  $E_{bb}^0$  corresponds to the backbone energy of a single strand of the same size with a conformation taken from an A-form double helix.

An optimal determination of the 3D encapsidated structure of the RNA genome goes beyond the capabilities of the current RNA–protein interaction modeling, which was discussed and conceptualized under different hypotheses, specifically for virus assemblies [44–47]. Nevertheless, we aim to present a systematic and reasonable model, which brings a new methodology, to structurally interpret, explain and evaluate the feasibility of the chemically probed secondary structures and the further microscopic characteristics of confined RNA inside virus capsids. We illustrate the application of SMC simulations on a fragment of the 3' domain, which comprises nucleotides 772 to 897. This fragment is composed of five sequential hairpins, which shows resemblance to the secondary structure models proposed earlier [35]. It also serves as an initial benchmark to test the steering around a five-fold symmetry axis, as illustrated in Figure 5a, using the stems determined experimentally as templates for the SMC simulation. Despite the simplicity of this benchmark, note that the arrangement of the structure is not uniquely defined. Such heterogeneity relies mainly on two different scenarios, with the hairpins inward and outward to the pentameric center, as illustrated in Figure 5c,d. Starting from three different initial conformations relaxed after 50,000,000 MC sweeps, we impose the  $\mathcal{E}$ RMSD restraints to observe a final structure that resembles the template, depicted in Figure 5c. We monitor the backbone energy per nucleotide (Figure 6), in reduced units, where we find that, in fact, the additional stretching present in scenario two (inwards conformation) is well reflected in the backbone energy, regardless of the initial configuration. Moreover, the  $\mathcal{E}$ RMSD reaches considerably lower values, independently of the initial condition, which states that the inward scenario can become entangled more easily during the simulation. Once the energetically correct scenario under an efficient coarse-grained representation is determined, our method also offers a direct way of backmapping, should the microscopic application demand so.



**Figure 5.** Steered Monte Carlo simulations on five sequential stem loops. (a) X-ray structure (PDB:4OQ9) on five-fold symmetry axis, containing 5 stem loops of undefined sequence. (a) Stem loop templates for SMC simulation extracted from the PDB and brought to coarse-grained resolution. (c,d) Final coarse-grained conformations after steering, employing different target orientations consistent with experimental restraints illustrated at the bottom right of each configuration.



**Figure 6.**  $\epsilon\text{RMSD}$  as a function of MC sweep for inward and outward scenarios (left) and normalized backbone energy (right) normalized by a reference single strand in A-form structure. The used  $\epsilon\text{RMSD}$  adimensional cutoff is 20.

#### 4. Discussion

Our results show a structural 3D domain reconstruction of the RNA genome from viral secondary structures, and propose a way to control the quality of the model based on further microscopic information and/or modeling hypotheses. There are several routes and assumptions involved in the quality control of those 3D reconstructions. Here, we classify them in two main routes, namely, high-resolution microscopic information and novel computational/theoretical models. High-resolution microscopic information, such as

cryoelectron microscopy [48–50], probe RNA structures inside viruses. On one hand, those advanced microscopy techniques coincide with the ability of encapsidated RNA to form globally more than one structure inside the virus. On the other hand, localized subdomains can form partially ordered structures (with well-defined minimum energy structures), as is the case of the five-fold symmetry example. Nonetheless, many subdomains of such encapsidated RNA lie within the ‘dynamic structures’, which are commonly difficult to efficiently sample by pure experiments. Hence, the field appeals to computational models, which have been used, for example, to refine 3D RNA reconstructions with additional restraints on tertiary contacts given by SAXS experiments [51], or the reconstruction of the frameshift stimulating element from SARS-CoV-2 all the way down to atomistic resolution. Such studies were based on chemical probing and CryoEM [52]. Revisiting microscopic information with higher resolution structural information has resolved the formerly intractable structure Tetrahymena ribozyme [53]. The latter successful studies, among others currently in progress, brings us to a particular question: what if 3D domain reconstruction methods could also improve the current shortage of theoretical conceptualization and prediction for particular physical models? For example, in the case of protein shells encapsulating RNA, the precise origin of interactions between RNA and capsid proteins is still unsolved, although several theories have been brought to discussion [47]. To this end, we also prepared our method to be rapidly extended in this direction, especially for coupling diverse types of molecular interactions, such as electrostatic and van der Waals, among others, to candidate 3D reconstructed domains (more details in the following repository <https://doi.org/10.5281/zenodo.5035769>). Such an approach is complex by itself; however, the offered ability of our method to discern between molecular interactions can be used to test physical assumptions, and guide the interpretation of experiments and possibly the direction of new experiments. On top of this general method improvements, a robust parametrization of the underlying RNA-protein molecular interactions would consequently permit the modification and/or control of the capsid status from assembled to disassembled or vice versa, which are awaited features for further applications in nanomedicine.

## 5. Conclusions

We have presented a new computational method for the 3D reconstruction of the complete RNA genome of virus particles. The RNA genome of the STMV was tested systematically with our algorithm, where we demonstrated the possibility of generating reasonable conformations of the three-dimensional structure of the RNA genome with coarse-grained efficient models (1 CG bead  $\approx$  30 atoms per nucleotide). This method includes several features, such as the radius of gyration topological constraints, or diverse restraints applied to RNA stems. In addition, our method can be easily extended and combined with further experimental information and/or computational/theoretical models. We have further shown an example of additional microscopic analysis with the five-fold hairpins distribution disambiguation for the STMV, based on the additional X-ray structural information. Finally, we have also discussed examples on how to not only further test the novel microscopic information, but to combine it with theoretical and semi-empirical models to elucidate further mechanisms of virus capsid assembly, disassembly and RNA delivery.

**Author Contributions:** H.V.G. and S.P. designed the research and wrote the paper; S.P. performed the MC simulations using the novel algorithm; H.V.G. supervised the research. Both authors have read and agreed to the published version of the manuscript.

**Funding:** H.V.G. acknowledges financial support from the Slovenian Research Agency ARRS (Funding No. P1-0055). S.P. acknowledges funding through the project FONDECYT Iniciación en Investigación No. 11181334.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data are available at the following Zenodo repository: <https://doi.org/10.5281/zenodo.5035769>.

**Acknowledgments:** We thank Rudolf Podgornik for illuminating discussions of the RNA and capsid protein biophysical interactions. Moreover, we want to acknowledge the work of Kevin Weeks and co-workers for providing chemical reactivity data of the STMV, which has inspired this work and further modeling ideas at a multiscale simulations meeting organized by the project SFB-TRR146 of the Deutsche Forschungsgemeinschaft. S. P. acknowledges Judit Lisoni for infrastructural support.

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

## References

1. Hu, Y.; Zandi, R.; Anavitarte, A.; Knobler, C.M.; Gelbart, W.M. Packaging of a Polymer by a Viral Capsid: The Interplay Between Polymer Length and Capsid Size. *Biophys. J.* **2008**, *94*, 1428–1436. [[CrossRef](#)]
2. Roos, W.H.; Gibbons, M.M.; Arkhipov, A.; Uetrecht, C.; Watts, N.R.; Wingfield, P.T.; Steven, A.C.; Heck, A.J.R.; Schulten, K.; Klug, W.S.; et al. Squeezing Protein Shells: How Continuum Elastic Models, Molecular Dynamics Simulations, and Experiments Coalesce at the Nanoscale. *Biophys. J.* **2010**, *99*, 1175–1181. [[CrossRef](#)]
3. Castellanos, M.; Pérez, R.; Carrasco, C.; Hernando-Pérez, M.; Gómez-Herrero, J.; de Pablo, P.J.; Mateu, M.G. Mechanical Elasticity as a Physical Signature of Conformational Dynamics in a Virus Particle. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12028–12033. [[CrossRef](#)] [[PubMed](#)]
4. Stockley, P.G.; Ranson, N.A.; Twarock, R. A New Paradigm for the Roles of the Genome in ssRNA Viruses. *Future Virol.* **2013**, *8*, 531–543. [[CrossRef](#)]
5. Garmann, R.F.; Comas-Garcia, M.; Koay, M.S.; Cornelissen, J.J.; Knobler, C.M.; Gelbart, W.M. Role of Electrostatics in the Assembly Pathway of a Single-Stranded RNA Virus. *J. Virol.* **2014**, *88*, 10472–10479. [[CrossRef](#)] [[PubMed](#)]
6. Hernandez-Garcia, A.; Kraft, D.J.; Janssen, A.F.J.; Bomans, P.H.H.; Sommerdijk, N.A.J.M.; Thies-Weesie, D.M.E.; Favretto, M.E.; Brock, R.; de Wolf, F.A.; Werten, M.W. et al. Design and Self-Assembly of Simple Coat Proteins for Artificial Viruses. *Nat. Nanotechnol.* **2014**, *9*, 698–702. [[CrossRef](#)]
7. Perlmutter, J.D.; Hagan, M.F. The Role of Packaging Sites in Efficient and Specific Virus Assembly. *J. Mol. Biol.* **2015**, *427*, 2451–2467. [[CrossRef](#)]
8. Beren, C.; Dreesens, L.L.; Liu, K.N.; Knobler, C.M.; Gelbart, W.M. The Effect of RNA Secondary Structure on the Self-Assembly of Viral Capsids. *Biophys. J.* **2017**, *113*, 339–347. [[CrossRef](#)] [[PubMed](#)]
9. Božič, A.L.; Micheletti, C.; Podgornik, R.; Tubiana, L. Compactness of Viral Genomes: Effect of Disperse and Localized Random Mutations. *J. Phys. Condens. Matter* **2018**, *30*, 084006. [[CrossRef](#)]
10. Viso, J.F.; Belelli, P.; Machado, M.; González, H.; Pantano, S.; Amundarain, M.J.; Zamarreño, F.; Branda, M.M.; Guérin, D.M.; Costabel, M.D. Multiscale Modelization in a Small Virus: Mechanism of Proton Channeling and its Role in Triggering Capsid Disassembly. *PLoS Comput. Biol.* **2018**, *14*, e1006082. [[CrossRef](#)] [[PubMed](#)]
11. Martínez, M.; Cooper, C.D.; Poma, A.B.; Guzman, H.V. Free Energies of the Disassembly of Viral Capsids From a Multiscale Molecular Simulation Approach. *J. Chem. Inf. Model.* **2019**, *60*, 974–981. [[CrossRef](#)] [[PubMed](#)]
12. Procyk, J.; Poppleton, E.; Šulc, P. Coarse-grained nucleic acid–protein model for hybrid nanotechnology. *Soft Matter* **2021**, *17*, 3586–3593. [[CrossRef](#)]
13. Brion, P.; Westhof, E. Hierarchy and Dynamics of RNA Folding. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113–137. [[CrossRef](#)]
14. Tinoco, I.; Borer, P.N.; Dengler, B.; Levine, M.D.; Uhlenbeck, O.C.; Crothers, D.M.; Gralla, J. Improved Estimation of Secondary Structure in Ribonucleic Acids. *Nature New Biol.* **1973**, *246*, 40–41. [[CrossRef](#)]
15. Tinoco, I.; Bustamante, C. How RNA Folds. *J. Mol. Biol.* **1999**, *293*, 271–281. [[CrossRef](#)]
16. Mathews, D.H.; Sabina, J.; Zuker, M.; Turner, D.H. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *J. Mol. Biol.* **1999**, *288*, 911–940. [[CrossRef](#)]
17. Bottaro, S.; Di Palma, F.; Bussi, G. The Role of Nucleobase Interactions in RNA Structure and Dynamics. *Nucleic Acids Res.* **2014**, *42*, 13306–13314. [[CrossRef](#)]
18. Miao, Z.; Westhof, E. RNA Structure: Advances and Assessment of 3D Structure Prediction. *Annu. Rev. Biophys.* **2017**, *46*, 483–503. [[CrossRef](#)] [[PubMed](#)]
19. Ponce-Salvatierra, A.; Astha; Merdas, K.; Nithin, C.; Ghosh, P.; Mukherjee, S.; Bujnicki, J.M. Computational modeling of RNA 3D structure based on experimental data. *Biosci. Rep.* **2019**, *39*, BSR20180430. [[CrossRef](#)] [[PubMed](#)]
20. Li, B.; Cao, Y.; Westhof, E.; Miao, Z. Advances in RNA 3D Structure Modeling Using Experimental Data. *Front. Genet.* **2020**, *11*, 1147. [[CrossRef](#)]
21. Weeks, K.M. Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.* **2010**, *20*, 295–304. [[CrossRef](#)] [[PubMed](#)]
22. Mustoe, A.M.; Al-Hashimi, H.M.; Brooks, C.L. Secondary structure encodes a cooperative tertiary folding funnel in the Azoarcus ribozyme. *Nucleic Acids Res.* **2015**, *44*, 402–412. [[CrossRef](#)] [[PubMed](#)]
23. Mustoe, A.M.; Lama, N.N.; Irving, P.S.; Olson, S.W.; Weeks, K.M. RNA base-pairing complexity in living cells visualized by correlated chemical probing. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24574–24582. [[CrossRef](#)]

24. Morandi, E.; Manfredonia, I.; Simon, L.M.; Anselmi, F.; van Hemert, M.J.; Oliviero, S.; Incarnato, D. Genome-scale deconvolution of RNA structure ensembles. *Nat. Methods* **2021**, *18*, 249–252. [[CrossRef](#)] [[PubMed](#)]
25. Marinus, T.; Fessler, A.B.; Ogle, C.A.; Incarnato, D. A novel SHAPE reagent enables the analysis of RNA structure in living cells with unprecedented accuracy. *Nucleic Acids Res.* **2021**, *49*, e34. [[CrossRef](#)] [[PubMed](#)]
26. Weidmann, C.A.; Mustoe, A.M.; Jariwala, P.B.; Calabrese, J.M.; Weeks, K.M. Analysis of RNA–protein networks with RNP-MaP defines functional hubs on RNA. *Nat. Biotechnol.* **2021**, *39*, 347–356. [[CrossRef](#)] [[PubMed](#)]
27. Larman, B.C.; Dethoff, E.A.; Weeks, K.M. Packaged and Free Satellite Tobacco Mosaic Virus (STMV) RNA Genomes Adopt Distinct Conformational States. *Biochemistry* **2017**, *56*, 2175–2183. [[CrossRef](#)]
28. Das, R.; Baker, D. Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14664–14669. [[CrossRef](#)]
29. Cheng, C.Y.; Chou, F.C.; Das, R. Chapter Two—Modeling Complex RNA Tertiary Folds with Rosetta. In *Computational Methods for Understanding Riboswitches; Methods in Enzymology*; Chen, S.J., Burke-Aguero, D.H., Eds.; Academic Press: Waltham, MA, USA, 2015; Volume 553, pp. 35–64. [[CrossRef](#)]
30. Magnus, M.; Antczak, M.; Zok, T.; Wiedemann, J.; Lukasiak, P.; Cao, Y.; Bujnicki, J.M.; Westhof, E.; Szachniuk, M.; Miao, Z. RNA-Puzzles toolkit: A computational resource of RNA 3D structure benchmark datasets, structure manipulation, and evaluation tools. *Nucleic Acids Res.* **2019**, *48*, 576–588. [[CrossRef](#)]
31. Schön, P. Atomic force microscopy of RNA: State of the art and recent advancements. *Semin. Cell Dev. Biol.* **2018**, *73*, 209–219. [[CrossRef](#)]
32. Freddolino, P.L.; Arkipov, A.S.; Larson, S.B.; McPherson, A.; Schulten, K. Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus. *Structure* **2006**, *14*, 437–449. [[CrossRef](#)]
33. Zeng, Y.; Larson, S.B.; Heitsch, C.E.; McPherson, A.; Harvey, S.C. A Model for the Structure of Satellite Tobacco Mosaic Virus. *J. Struct. Biol.* **2012**, *180*, 110–116. [[CrossRef](#)]
34. Cieplak, M.; Robbins, M.O. Nanoindentation of 35 Virus Capsids in a Molecular Model: Relating Mechanical Properties to Structure. *PLoS ONE* **2013**, *8*, e63640. [[CrossRef](#)]
35. Schroeder, S.J. Probing Viral Genomic Structure: Alternative Viewpoints and Alternative Structures for Satellite Tobacco Mosaic Virus RNA. *Biochemistry* **2014**, *53*, 6728–6737. [[CrossRef](#)]
36. Guzman, H.V.; Tretyakov, N.; Kobayashi, H.; Fogarty, A.C.; Kreis, K.; Krajniak, J.; Junghans, C.; Kremer, K.; Stuehn, T. ESPReso++ 2.0: Advanced Methods for Multiscale Molecular Simulation. *Comput. Phys. Commun.* **2019**, *238*, 66–76. [[CrossRef](#)]
37. Poblete, S.; Bottaro, S.; Bussi, G. A Nucleobase-Centered Coarse-Grained Representation for Structure Prediction of RNA Motifs. *Nucleic Acids Res.* **2017**, *46*, 1674–1683. [[CrossRef](#)] [[PubMed](#)]
38. Poblete, S.; Bottaro, S.; Bussi, G. Effects and Limitations of a Nucleobase-Driven Backmapping Procedure for Nucleic Acids Using Steered Molecular Dynamics. *Biochem. Biophys. Res. Commun.* **2018**, *498*, 352–358. [[CrossRef](#)]
39. Boniecki, M.J.; Lach, G.; Dawson, W.K.; Tomala, K.; Lukasz, P.; Soltysinski, T.; Rother, K.M.; Bujnicki, J.M. SimRNA: A Coarse-Grained Method for RNA Folding Simulations and 3D Structure Prediction. *Nucleic Acids Res.* **2015**, *44*, e63. [[CrossRef](#)] [[PubMed](#)]
40. Šulc, P.; Romano, F.; Ouldridge, T.E.; Doye, J.P.K.; Louis, A.A. A Nucleotide-Level Coarse-Grained Model of RNA. *J. Chem. Phys.* **2014**, *140*, 235102. [[CrossRef](#)] [[PubMed](#)]
41. Zhang, D.; Chen, S.J. IsRNA: An Iterative Simulated Reference State Approach to Modeling Correlated Interactions in RNA Folding. *J. Chem. Theory Comput.* **2018**, *14*, 2230–2239. [[CrossRef](#)] [[PubMed](#)]
42. Pasquali, S.; Derreumaux, P. HiRE-RNA: A High Resolution Coarse-Grained Energy Model for RNA. *J. Phys. Chem. B* **2010**, *114*, 11957–11966. [[CrossRef](#)]
43. Kerpedjiev, P.; Siederdisen, C.H.Z.; Hofacker, I.L. Predicting RNA 3D structure using a coarse-grain helix-centered model. *RNA* **2015**, *21*, 1110–1121. [[CrossRef](#)] [[PubMed](#)]
44. Perlmutter, J.D.; Hagan, M.F. Mechanisms of Virus Assembly. *Annu. Rev. Phys. Chem.* **2015**, *66*, 217–239. [[CrossRef](#)]
45. Buzón, P.; Maity, S.; Roos, W.H. Physical virology: From virus self-assembly to particle mechanics. *WIREs Nanomed. Nanobiotechnol.* **2020**, *12*, e1613. [[CrossRef](#)] [[PubMed](#)]
46. Schroeder, S.J. Perspectives on Viral RNA Genomes and the RNA Folding Problem. *Viruses* **2020**, *12*, 1126. [[CrossRef](#)]
47. Zandi, R.; Dragnea, B.; Travesset, A.; Podgornik, R. On Virus Growth and Form. *Phys. Rep.* **2020**, *847*, 1–102. [[CrossRef](#)]
48. Dai, X.; Li, Z.; Lai, M.; Shu, S.; Du, Y.; Zhou, Z.H.; Sun, R. In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature* **2017**, *541*, 112–116. [[CrossRef](#)]
49. Beren, C.; Cui, Y.; Chakravarty, A.; Yang, X.; Rao, A.L.N.; Knobler, C.M.; Zhou, Z.H.; Gelbart, W.M. Genome organization and interaction with capsid protein in a multipartite RNA virus. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 10673–10680. [[CrossRef](#)]
50. Tomezsko, P.J.; Corbin, V.D.A.; Gupta, P.; Swaminathan, H.; Glasgow, M.; Persad, S.; Edwards, M.D.; Mcintosh, L.; Papenfuss, A.T.; Emery, A.; et al. Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **2020**, *582*, 438–442. [[CrossRef](#)]
51. Chojnowski, G.; Zaborowski, R.; Magnus, M.; Bujnicki, J.M. RNA fragment assembly with experimental restraints. *bioRxiv* **2021**. [[CrossRef](#)]

- 
52. Rangan, R.; Watkins, A.M.; Chacon, J.; Kretsch, R.; Kladwang, W.; Zheludev, I.N.; Townley, J.; Rynge, M.; Thain, G.; Das, R. De novo 3D models of SARS-CoV-2 RNA elements from consensus experimental secondary structures. *Nucleic Acids Res.* **2021**, *49*, 3092–3108. [[CrossRef](#)] [[PubMed](#)]
  53. Kappel, K.; Zhang, K.; Su, Z.; Watkins, A.M.; Kladwang, W.; Li, S.; Pintilie, G.; Topkar, V.V.; Rangan, R.; Zheludev, I.N.; et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **2020**, *17*, 699–707. [[CrossRef](#)] [[PubMed](#)]