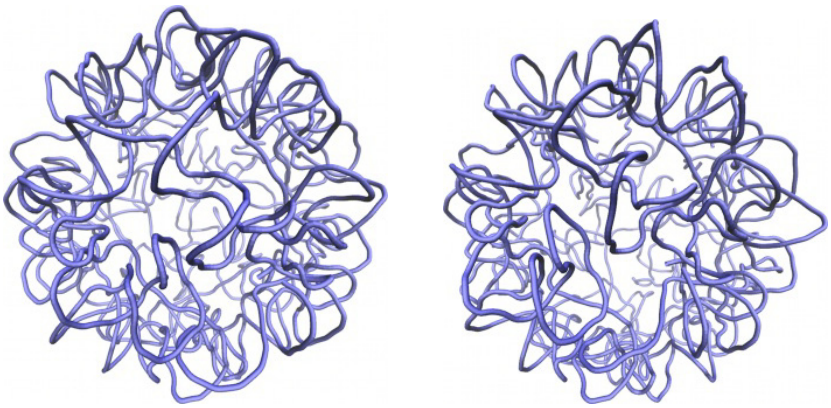
Supporting Information

for "Space Warping Order Parameters and Symmetry: Application to Multiscale Simulation of Macromolecular Assemblies" by Abhishek Singharoy, Harshad Joshi, Yinglong Miao and Peter J. Ortoleva

SI1. Comparison of DMS and NAMD simulations

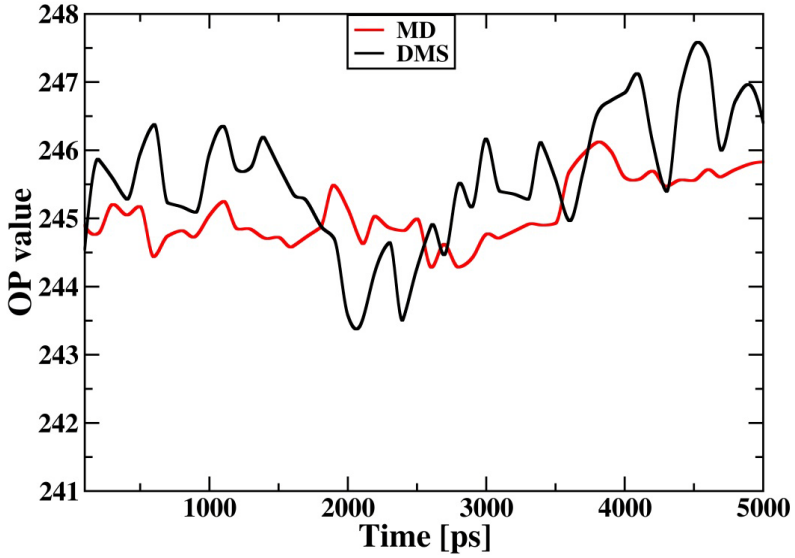
As shown in Fig. S1, STMV RNA in 0.3M $MgCl_2$ solution was found to maintain its three-dimensional structure through 25ns DMS simulation, indicating that the RNA is stable in the 2:1 electrolyte.



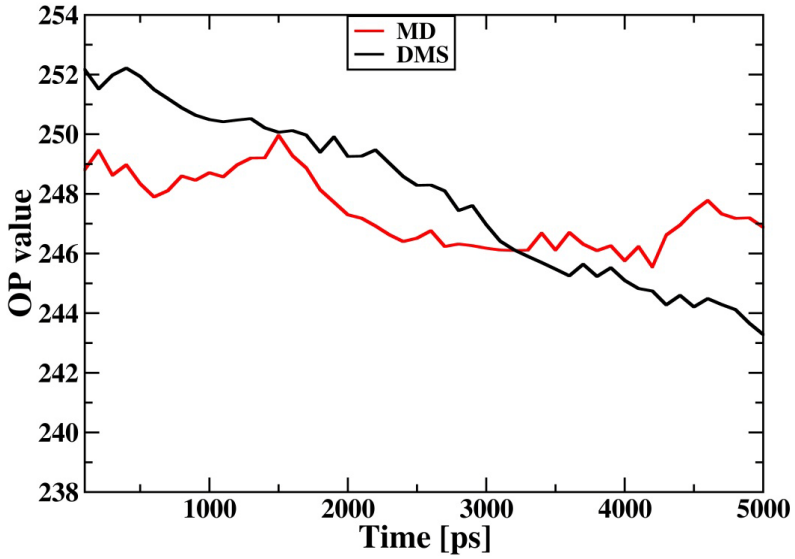
(a) (b)

Fig. S1 The three-dimensional structure of the STMV RNA in 0.3M $MgCl_2$ solution at (a) beginning and (b) end of 25 ns DMS simulation.

Fig. S2 plots the timecourse of NAMD and DMS derived OP Φ_{100X} at 310K and 425K. The space warping OP was found to be comparable between the two simulations. This suggests the equivalence of overall structure evolution in DMS and NAMD simulations and validates the applicability of DMS over a biologically relevant range of temperatures.



(a)



(b)

Fig. S2 Comparison of NAMD and DMS derived OP trajectory at (a) 310 and (b) 425K for the Φ_{100X} suggesting equivalence of the overall structure evolution.

SI2. Temperature Dependence of Order Parameter Evolution

The need to add more OPs for capturing emergent collective modes is indicated by a systematic growth of the residuals $\bar{\sigma}_i$ (during the generation of all-atom ensembles) that

accompany the increase in temperature as shown in Fig. S3. These OPs , denoted $\vec{\phi}_{\underline{k}}^{\text{new}}$, are constructed from the population of residuals via

$$\vec{\phi}_{\underline{k}}^{\text{new}} = \frac{\sum_{i=1}^{N} m_i u_{\underline{k}i} \vec{\sigma}_i}{\sum_{i=1}^{N} m_i \{u_{\underline{k}i}\}^2} \tag{1}$$

It is found that some of these new OPs have coherent behavior and therefore there dynamics can be modeled via Langevin equations. Evolution of these OPs captures the collective modes that emerge due to increase in temperature.

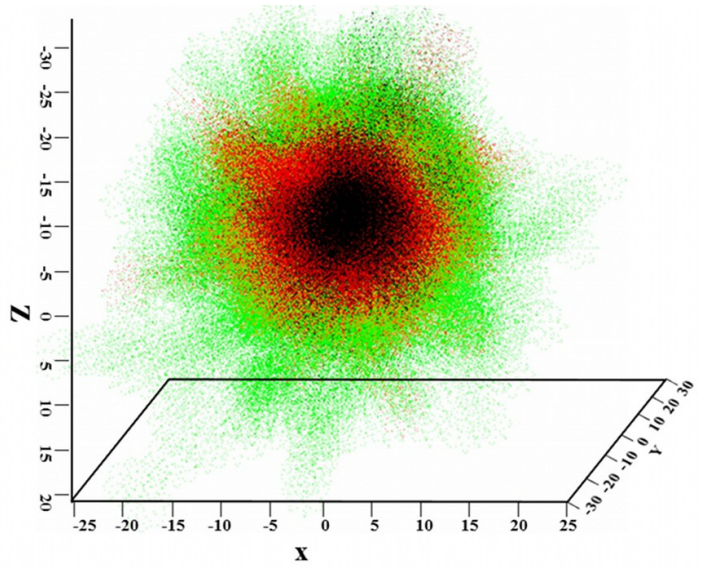


Fig. S3 Systematic growth of residuals $\vec{\sigma}_i$ accompanying increase in temperature from 310 (black) to 350 (red) to 400K (green). The growth indicates underlying collective motions that are not accounted for by the set of OPs initially chosen. As temperature increases a greater number of OPs are required to capture the emergent collective modes.

However, as temperature increases so does the diffusivities (Fig. 4a) and the noise $\vec{\xi}_{\underline{k}}$. Consequently, coherence in the OP dynamics decreases. Although collective modes are slow relative to characteristic fluctuations at a given temperature emerge, the timescale at which they change decreases with increase in temperature. To address this, the Langevin timesteps are

reduced as the temperature increases (see Table S2). However, at temperatures like 450-600K, the OP timescale approaches that of OP velocity autocorrelation function decay. This signifies that non-inertial Langevin equations (Eqn. 7) with the present set of OPs cannot be used. Thus, either additional space warping OPs are needed or inertial effects become important. To address the latter, the present multiscale formulation can be used when OP momenta are added to the coarse-grained description; the result would be a Fokker-Plank type equation for the joint probability of OPs and their momenta. These concepts are used to rationalize the choice of simulation methodologies used at various temperatures. Thus, DMS simulations are used for temperatures lower than or equal to 425K and NAMD is used for higher temperatures at which the use of non-inertial Langevin equations fail. Although, at high temperatures the space warping OPs are not sufficient for capturing the system evolution (e.g., OP momenta also contribute to the coarse-grained description), they remain key indicators of order and symmetry as demonstrated here for the RNA.

Table S1. Input parameters for the NAMD simulations (both inside and outside DMS). *SampleSize denotes the number of independent short MD (1 ps) runs to generate all-atom ensembles for every Langevin timestep inside DMS.

Parameter	Values
Langevin damping	5
Timestep	1fs
fullElectFrequency	2fs
nonbondedFreq	1fs
Force-field parameter	par_all27_prot_na.prm
1-4scaling	1.0
Switchdist	10.0 Å
Cutoff	12.0 Å
Pairlistdist	20.0 Å
Stepspercycle	2
Rigid bond	Water
Initial box size	150Å x150Å x 150Å
No. of TIP3P water molecules	116902
No. of Mg^{2+} ion	474
SampleSize*	200

Table S2. DMS timesteps for RNA simulations at different temperatures.

Temperature (K)	DMS timestep (ps)	DMS simulation time [ns]	Benchmarked with NAMD [ns]
310	100	25	5
350	100	25	5
375	50	25	5
400	50	25	5
425	25	25	5
450	NA	--	25
475	NA	--	25
500	NA	--	25
600	NA	--	25

SI3. Dynamics of STMV RNA

The temperature-dependent behavior of the ensemble-averaged radius of gyration (Rg) of STMV RNA is plotted in Fig. S4, which shows a dramatic expansion transition at ~475K.

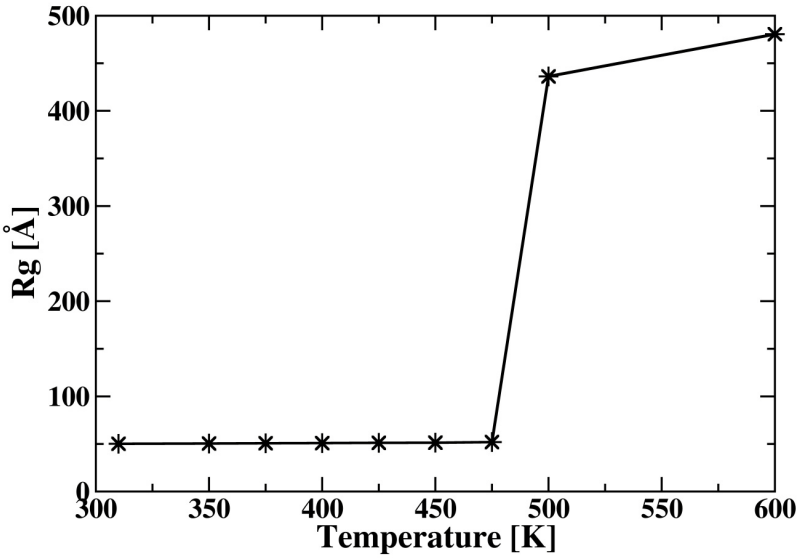
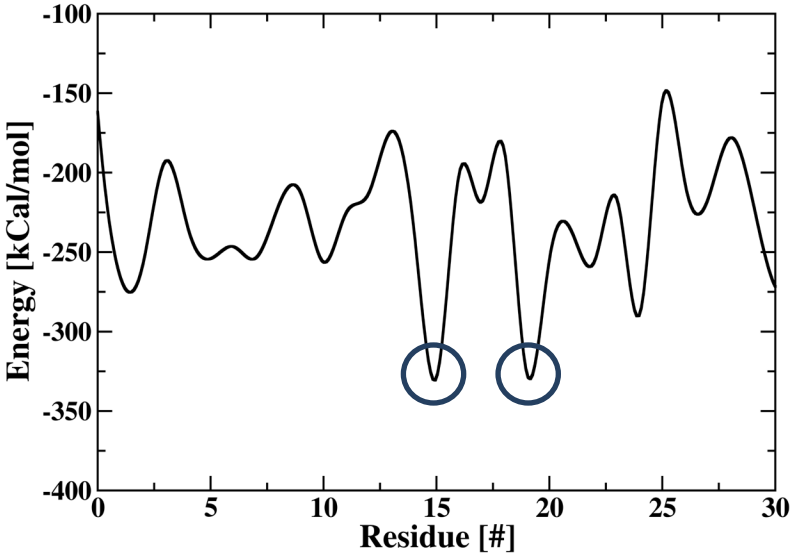


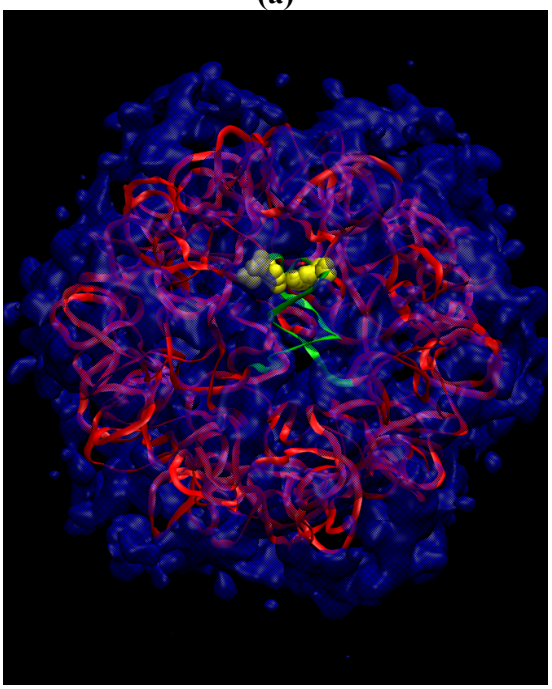
Fig. S4 Temperature-dependent behavior of the ensemble-averaged radius of gyration (Rg) of STMV RNA.

Fig. S5 shows the ensemble-averaged electrostatic energy of protein-bound RNA helix nucleotides at 310K, which indicates nucleotides 14 (ADE) and 19 (URA) are most stabilized by the protein core, and the structural snapshot of encapsidated RNA along with +10 KT/e isosurface of STMV capsid. The isosurface reflects the electro-positive character of the inner

capsid surface, thereby complementing the negatively-charged RNA. Nucleotides 14 and 19 are found to interact more with the capsid isosurface than the rest of the helix, i.e., yellow regions of the RNA overlap more with the isosurface than the green (see Fig. S5b). This indicates that nucleotides 14 and 19 are strongly bound to the positively-charged capsid inner surface and therefore are electrostatically more stable than other nucleotides of the RNA helix.



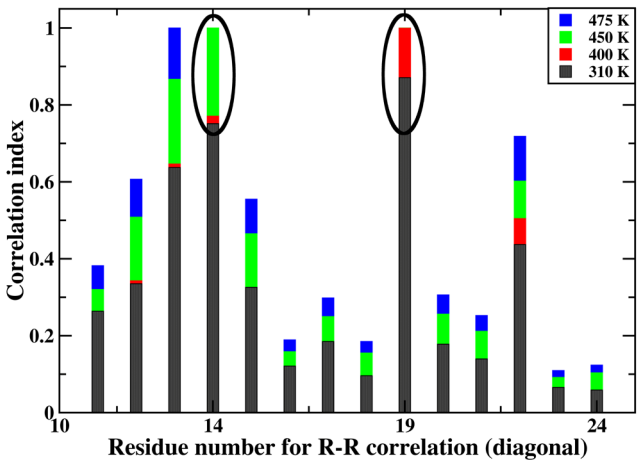
(a)



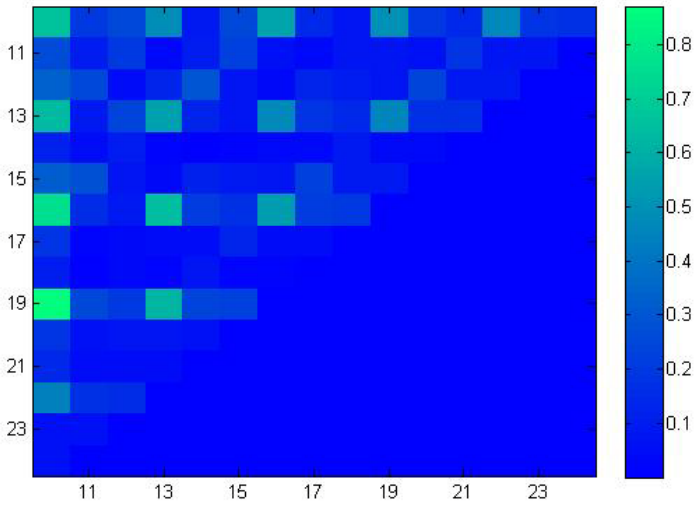
(b)

Fig. S5 (a) Ensemble-averaged electrostatic energy of protein-bound RNA helix nucleotides at 310K showing nucleotides 14 (ADE) and 19 (URA) are most stabilized by the protein core and (b) the structural snapshot of encapsidated RNA (red) along with +10 KT/e iso-surface (blue) of STMV capsid. A typical RNA helix is highlighted in green with nucleotides 14 and 19 shown in yellow sphere representation.

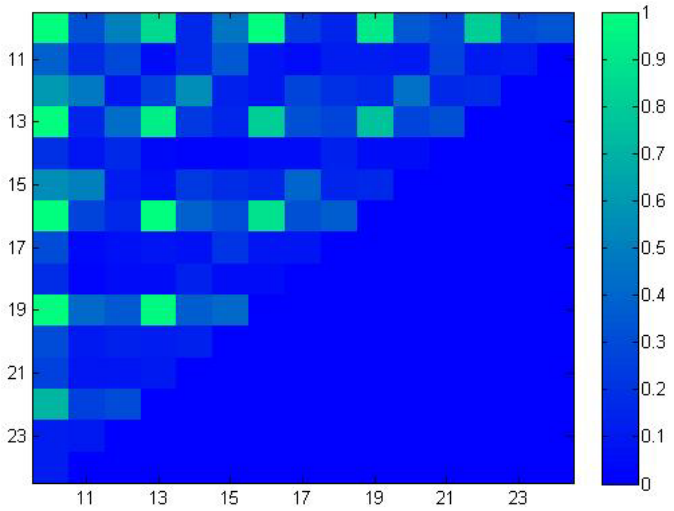
Nucleotides 14 and 19 in STMV RNA are identified as thermo-labile nucleotides (Fig. 6a) as they express pronounced self-correlation at temperatures much lower than the transition zone. Fig. S6a plots diagonal terms of the absolute nucleotide-nucleotide correlation in the vicinity of the thermo-labile centers. For comparison, correlations at four different temperatures are presented. The plot shows that residue 14 and residue 19 exhibit maximum correlations at temperatures below phase transition temperature (310 and 400 K). Near the transition zone (around 475 K), neighboring nucleotides become strongly coupled to those of the thermo-labile centers, i.e., nucleotides 14 and 19. Thus, off-diagonal correlations increase as the temperature is raised from 310 and 475 K. This implies that at temperatures below transition zone, thermo-labile nucleotides predominantly express motions; while near the transition temperatures neighboring nucleotides also participate in these motions.



(a)



(b)



(c)

Fig. S6 (a) The diagonal terms of the absolute nucleotide-nucleotide correlation between nucleic acids in the vicinity of the thermo-labile centers, i.e., nucleotides 14 and 19. (b) The absolute nucleotide-nucleotide correlation at 301K and (c) 475K near the transition zone.

Fig. S7 plots the average deviations of the nucleotide dihedral angles from their initial values for the 30 RNA nucleotides as in protein-bound STMV at 310K and 400K, and the protein-unbound (i.e., free) at 400K. The average value of angle deviations is significantly larger for the protein-unbound RNA than for the protein-bound state. This underscores the role of capsid proteins in stabilizing the RNA structure at multiple temperatures.

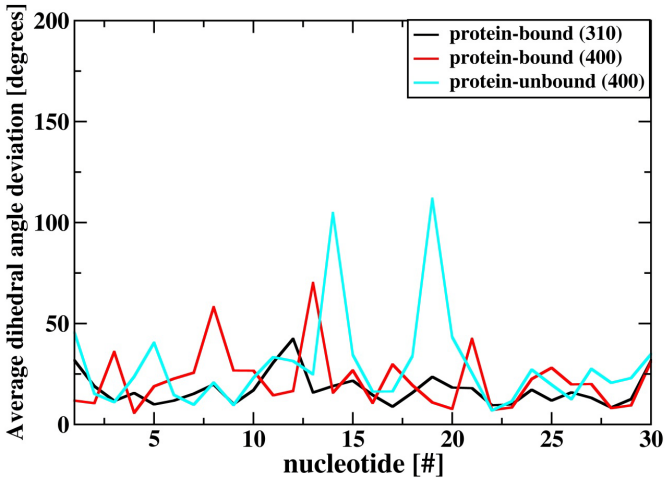


Fig. S7. Average deviations of the nucleotide dihedral angles from their initial values for the 30 nucleotides protein-bound STMV RNA at 310K (black) and 400K (red), and the protein-unbound (i.e., free) at 400K (cyan).