

Improving the model of the SARS-CoV-2 long-range interaction

Clustering the mutational profile of the long-range interaction formed state enabled us to refine the model of the long-range interaction in SARS-CoV-2. We predicted the structure of a 1,799 nt segment of the genome centered on the long-range interaction.

To verify the predicted structures, we performed SEARCH-MaP on the 1,799 nt segment, this time with shorter ASOs (15-20 nt) to reach single-stem precision. Each ASO targeted a single stem in the downstream portion of the interaction, and we measured the change in DMS reactivities of the FSE. ASOs targeting the inner stem and the outer half of the middle stem perturbed the DMS reactivities in exactly the anticipated locations. The other ASOs also perturbed the DMS reactivities at the anticipated location with additional other off-target perturbations that were likely due to the other parts of the structure rearranging as a consequence of the primary perturbation from the ASO. These results show that each stem in the refined model of the long-range interaction does exist and can be detected with SEARCH-MaP – sometimes cleanly and sometimes with other off-target effects. It appears that combining SEARCH-MaP data and computational modeling based on SEISMIC-RNA clustering out the long-range mutational profile offers better ability of finding long-range interactions than either technique alone. SEARCH-MaP compensates for the known difficulties in modeling long-range interactions computationally [CITE sources], while the predictions based on the long-range mutational profile found with SEISMIC-RNA can reveal the base pairs involved in the long-range interaction and compensate for secondary effects during SEARCH-MaP.

The long-range interaction inhibits the FSE pseudoknot

The refined model of the long-range interaction recapitulated the inner and middle stems of the previously proposed FSE-arch ?. The outer stem differed – rather than occur before the FSE, it was actually predicted to overlap the FSE, suggesting that it would be mutually exclusive with some structures in the FSE. Specifically, the bases predicted to form stem 2 of the pseudoknot – essential for frameshifting [CITE mutational studies] – were predicted to participate in the long-range interaction. This finding suggests that the long-range interaction would prevent the FSE pseudoknot from forming.

To test this hypothesis, we used one ASO that targeted specifically the six base pairs that would be mutually exclusive with the pseudoknot stem 2 and one ASO to disrupt alternative stem 1 (AS1), which competes with pseudoknot stem 1 ?. Because AS1 strongly outcompetes with the pseudoknot (such that it is the predominant structure even without the long-range interaction) ?, we expected that we would not detect the pseudoknot without the anti-AS1 ASO. Adding this ASO before folding the RNA should enable the pseudoknot to form unless another structure (e.g. the long-range interaction) outcompetes it. Adding this ASO after folding the RNA should enable the pseudoknot to form only if it can outcompete all other structures (e.g. the long-range interaction) that could inhibit it.

We predict the following results if the long-range interaction outcompetes the pseudoknot:

- Control: Roughly 50/50 mix of long-range formed and unformed structures; no detectible pseudoknot
- +Anti-AS1, before folding: Roughly 50/50 mix of long-range formed and unformed structures (but with unfolded AS1 and possibly folded PS1); no detectible pseudoknot in any long-range cluster
- +Anti-AS1, after folding: Roughly 50/50 mix of long-range formed and unformed structures (but with unfolded AS1 and possibly folded PS1); no detectible pseudoknot in any long-range cluster
- +Anti-long: Mix of long-range formed and unformed structures (formed will have a different structure and may be less

abundant than usual); no detectible pseudoknot - +Anti-long and +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure - +Anti-long and +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure

We predict the following results if the pseudoknot outcompetes the long-range interaction:

- Control: Roughly 50/50 mix of long-range formed and unformed structures; no detectible pseudoknot - +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure that is not long-range - +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure that is not long-range - +Anti-long: Mix of long-range formed and unformed structures (formed will have a different structure and may be less abundant than usual); no detectible pseudoknot - +Anti-long and +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure - +Anti-long and +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure

We predict the following results if the two structures are mutually exclusive and the first formed structure prevents the other from forming:

- Control: Roughly 50/50 mix of long-range formed and unformed structures; no detectible pseudoknot - +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure that is not long-range - +Anti-AS1, after folding: Roughly 50/50 mix of long-range formed and unformed structures (but with unfolded AS1 and possibly folded PS1); no detectible pseudoknot in any long-range cluster - +Anti-long: Mix of long-range formed and unformed structures (formed will have a different structure and may be less abundant than usual); no detectible pseudoknot - +Anti-long and +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure - +Anti-long and +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure

We predict the following results if the pseudoknot and the long-range interaction are not mutually exclusive:

- Control: Roughly 50/50 mix of long-range formed and unformed structures; no detectible pseudoknot - +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure that may be long-range - +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure that may be long-range - +Anti-long: Mix of long-range formed and unformed structures (formed will have a different structure and may be less abundant than usual); no detectible pseudoknot - +Anti-long and +Anti-AS1, before folding: Pseudoknot detectible in at least one alternative structure - +Anti-long and +Anti-AS1, after folding: Pseudoknot detectible in at least one alternative structure

We first used SEISMIC-RNA to cluster each replicate without ASOs with 1 to 8 clusters and found that they generated reproducible results up to 6 clusters, but not with more, suggesting that the RNA forms 6 predominant structural states. For each of these states, we computed the AUC-ROC versus a model of the FSE pseudoknot from another study ?. Only one state came close to agreeing with the pseudoknot (AUC = 0.88); it was the least abundant state in the ensemble (only 7%) and was also not consistent with the long-range interaction (AUC = 0.70).

When we blocked AS1 before folding the RNA, we found that it formed two structures: one (59% abundance) highly consistent with the pseudoknot (AUC = 0.96) but not the long-range interaction (AUC = 0.59), and another (41% abundance) more consistent with the long-range interaction (AUC = 0.80) than the pseudoknot (AUC = 0.75). We obtained similar, though slightly more pronounced, results when we added the anti-AS1 ASO after folding the RNA. These results show that blocking AS1 does enable the pseudoknot to fold, mutually exclusively with the long-range interaction.