Compactness and Apparent Circularization of ssRNA

A. Ben-Shaul*

Using the average maximum ladder distance <*MLD*> as a measure of the "compactness" of RNA secondary structure, we show that viral ssRNAs are consistently smaller than randomly permuted sequences of the same length and base composition. They are also smaller than natural, non-viral, ssRNAs, suggesting that viral RNAs are more compact owing to the evolutionary pressure to facilitate their packaging into small rigid protein capsids. We predict that the average <*MLD*>s of large non-viral ssRNAs scale as $N^{0.70}$, where N is the number of nucleotides, and – by mapping the secondary structures onto linear polymer models – argue that their radii of gyration, R_g, vary as $<MLD>^{0.50}$, and hence as $N^{0.35}$. We shall also discuss another – generic – property of linear ssRNA sequences, namely, the proximity of their 3' and 5' ends. We employ basic combinatorial and graph theory arguments to estimate the distance between the ends of linear sequences and argue that the secondary structures of linear and covalently-circularized RNAs are practically identical.

* Work in collaboration with A. Yoffe, P. Prinsen, A. Gopal, L. T. Fang, C. Knobler, and W. M. Gelbart