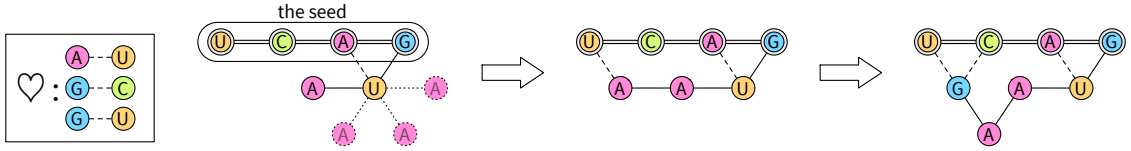
Efficient Universal Computation by Molecular Co-Transcriptional Folding

Cody Geary, Pierre-Étienne Meunier, Nicolas Schabanel,
Shinnosuke Seki

We investigate the algorithmic power of a previously overlooked phenomenon: the fact that molecules fold as they are transcribed (*co-transcriptional folding*). Although our questions are mostly fundamental, the model itself is inspired by the experimental work of Geary, Rothemund, and Andersen (*Science* 345:799-804, 2014), and could potentially be implemented using ssRNA Origami in the future.

We introduce a new theoretical model called *Oritatami* (i.e., folding in Japanese). The to-be-folded molecule is modeled as a sequence of beads (its *primary structure*) labelled by an alphabet Σ. Beads *attract* each other according to a symmetric relation $\heartsuit \subseteq \Sigma \times \Sigma$; e.g., in the case of RNA, we may set $\Sigma = \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{U}\}$ and $\heartsuit = \{(\mathtt{A}, \mathtt{U}), (\mathtt{C}, \mathtt{G}), (\mathtt{G}, \mathtt{U})\}$. A folded sequence is called a *conformation*. An Oritatami system consists in growing its initial conformation called *seed* (the *input* in algorithmic terms) by transcribing its primary structure and folding it on the triangular grid so as to form as many bonds as possible according to $\heartsuit$. Unlike conventional models of folding like Hydrophobic-Hydrophilic (HP) model, Oritatami systems do not fold the whole sequence at one time. Instead, it transcribes the next δ_t beads of the sequence, considers all possible ways to extend the current conformation by folding them, and determines their locations so as to maximize the number of bonds according to $\heartsuit$, where δ_t is a system parameter called the *delay time* which acts approximately as a temperature parameter (the larger it is, the less stable is the tail of the folding). For instance, an Oritatami system with delay time 1 settles a bead, once being transcribed, down in the unoccupied positions next to the preceding bead where maximum number of new bonds are formed. The figure bellow shows the various steps of a co-transcriptional folding of the sequence `UAAG` from the placed seed `UCAG` with a delay time 2 and the $\heartsuit$ attraction rule from above: note that the locations of the two `A`s change as the transcription goes on.



We study the simulation, prediction, and programming of Oritatami systems. First, we propose an efficient simulation of cyclic tag system (CTS) by an Oritatami system with delay time 3 whose primary structure is periodic. As a corollary, simulating Oritatami systems is **P**-hard. One can determine in polynomial time whether a given Oritatami system folds its sequence in an expected way. As for the programming, we first show that designing a set $\heartsuit$ of rules to fold a given sequence into given conformations is **NP**-hard. In contrast, if we can choose the sequence, then we can provide an algorithm to design an Oritatami system. The running time of this algorithm is linear in the length of the sequence but exponential in the number of target conformations.