Reinforcement Learning-Enabled Control of Single DNA Molecule Manipulation in Microfluidics

Fuhai Zhou [1] **Fushuai Wang** [2 3] **Nagaraj Nagalingam** [1] **Beatrice W. Soh** [1 3]

[1]*Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore* [2]*Department of Materials Science and Engineering, National University of Singapore, Singapore* [3]*Institute for Functional Intelligent Materials, National University of Singapore, Singapore*. Correspondence to: Beatrice W. Soh bwysoh@nus.edu.sg.

1. Introduction

The precise manipulation of small particles and biological samples has emerged as a powerful cornerstone across diverse engineering applications and fundamental science. However, in conventional feedback loops, effective manipulation typically necessitates a reduced-order physical model that characterizes the field-driven motion of particles within complex environments. Existing models remain confined to highly idealized settings, leaving a gap in handling complex nonlinear processes that has hindered the widespread adoption of automated manipulation. To address these limitations, we report the development of an automated 2D microfluidic platform integrated with a Reinforcement Learning (RL) agent. Using single DNA molecules as a model system, our approach achieves nontrivial shape control in a planar electric field, moving beyond traditional methods focused on simple particle trapping or linear stretching [1]. By learning to navigate the nonlinear coupling between the external field and the polymer's internal degrees of freedom, the agent can dynamically modulate the local electrokinetic environment in real time. This enables the successful execution of targeted bending task that was previously inaccessible, demonstrating a significant leap in the flexibility and intelligence of molecular manipulation, while offering a versatile strategy extensible to a wide array of automated manipulation applications.

2. Methodology

2.1 Problems setup

We formulate DNA conformational control as a reinforcement learning (RL) problem, where the agent learns to manipulate a single DNA molecule toward a desired target state. In this work, we consider two representative objectives: (1) regulating the DNA strectching length to a target range, and (2) steering the molecule into a target U-shaped conformation.

2.2 RL-based control framework

At each time step t, the current DNA coordinates are measured from the imaging system and converted into a compact observation vector containing geometric information of the molecule. Based on this observation, the RL agent outputs voltage actions to control the microfluidic flow field in real time. The overall framework is shown in Fig. 1.

We adopt Proximal Policy Optimization (PPO) [2] as the learning algorithm. The policy takes the observation s_t as input and predicts the control action

$$a_t = \pi_\theta(s_t),$$

where a_t denotes the voltage action applied at time step t. PPO is chosen because it provides stable optimization for continuous control and is well suited for sequential decision-making under noisy physical observations.

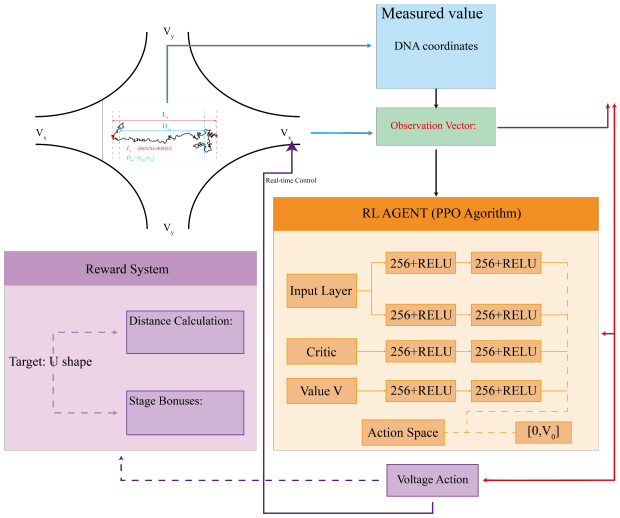

Fig. 1: Reinforcement learning framework for DNA conformational control

2.3 Reward design

For the target-length task, the reward is designed to encourage the DNA stretching length L_t to match a desired target L^*. A simple reward can be written as

$$R_{\text{length}} = -|L_t - L^*|,$$

so that higher reward is obtained when the measured DNA length is closer to the target.

For the U-shape control task, the reward contains three components. First, we measure the discrepancy between the current DNA conformation $X = \{x_i\}_{i=1}^N$ and the target shape $X^* = \{x_j^*\}_{j=1}^N$ using the Chamfer distance

$$D_{\text{ch}}(X, X^*) = \frac{1}{N}\sum_i \min_j \|x_i - x_j^*\|^2 + \frac{1}{N}\sum_j \min_i \|x_j^* - x_i\|^2.$$

A shape-matching reward is then defined as

$$R_{\text{shape}} = e^{-100 D_{\text{ch}}},$$

which increases as the DNA conformation becomes closer to the target U-shape. Second, an additional

concavity term $R_{\text{concavity}}$ is used to encourage the characteristic U-shaped bending. Third, a stage-based bonus is introduced to stabilize learning and reward intermediate progress:

$$R_{\text{stage}} = \begin{cases} 1, & D_{\text{ch}} < 0.05, \\ 0.5, & D_{\text{ch}} < 0.1, \\ 0, & \text{otherwise.} \end{cases}$$

The final reward for U-shape control is

$$R_{\text{Ushape}} = R_{\text{shape}} + R_{\text{concavity}} + R_{\text{stage}}.$$

This design allows the agent to learn both coarse global deformation and fine local shape adjustment.

3. Results

3.1 Experimental validation of length control

We first consider a simple control task, where the objective is to regulate the DNA stretching length to a desired target L^*. The RL agent is first trained in a simulated environment, and the learned policy is then transferred to real experiments for validation. This setup enables us to evaluate whether the proposed framework can bridge the gap between simulation and physical systems. As shown in Fig. 2, the top row presents the voltage control signal and the resulting DNA length over time. Despite frequent switching of the control inputs, the agent consistently maintains the DNA stretching within a relatively narrow range, indicating that the learned controller achieves stable regulation under experimental noise and uncertainty.

The fluorescence microscopy snapshots in the bottom row further illustrate the DNA conformations during control. As shown, the molecule remains in a partially stretched state rather than fully extended. Due to entropic elasticity, the DNA chain naturally tends to coil; however, the agent adapts its control strategy according to the instantaneous molecular configuration. At different time points, the molecule exhibits consistent and controllable conformational responses.

These results demonstrate that the learned policy is not limited to the simulation environment but can generalize to real experimental conditions. Overall, this provides strong evidence for the feasibility of reinforcement learning-based real-time control of single-molecule DNA systems.

3.2 U-shape control and learning dynamics

We next consider a more challenging task: steering the DNA molecule from an initially straight conformation into a target U-shape. Due to the complexity of real experiments, we first study this task in a simulated environment.Figure 3a shows a clear progression from a nearly straight state to a bent intermediate conformation and finally to a pronounced U-shape. This result suggests that the learned controller performs a smooth multi-stage deformation process rather than relying on abrupt or random actions.

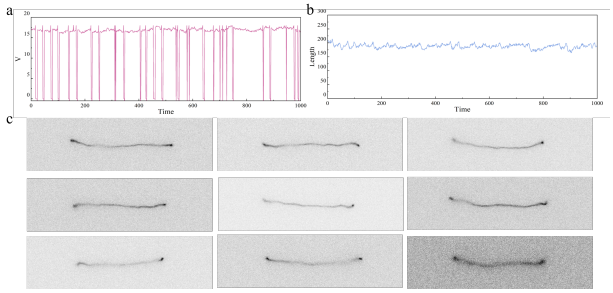

Fig. 2: Experimental validation of RL-based DNA control. a) Voltage control signal. b) DNA stretching length over time. c) Fluorescence images of DNA under control in experiments, showing that the stretching length is regulated around the desired target L^*.

Figures 3b and 3c show the two control action components, V_x and V_y, during one control episode. The action trajectories exhibit clear structure and coordination, suggesting that the agent has learned a meaningful control strategy adapted to the molecular deformation process. Figure 3d shows the reward evolution over time. The reward increases substantially during the middle stage of the episode, indicating progressive alignment with the target shape. Together, these results show that the RL agent can successfully learn nontrivial shape control policies for DNA molecules.

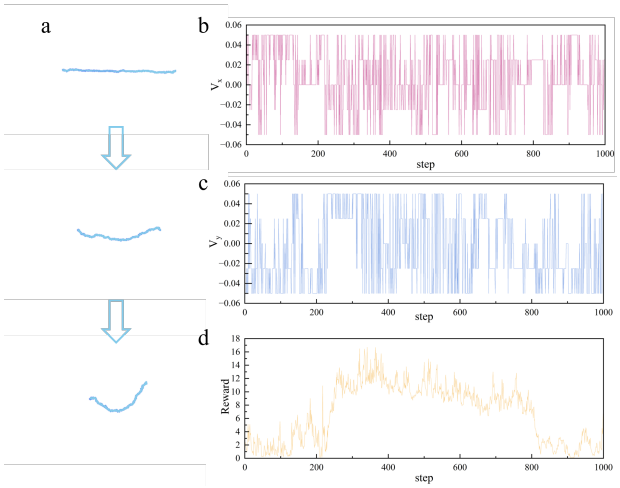

Fig. 3: U-shape control and policy dynamics. a) Evolution of DNA conformation from an initially straight configuration to a target U-shape. b),c) Learned control actions along two directions, V_x and V_y. d) Reward trajectory during the control episode.

Acknowledgments

The computational work for this work was partially performed on resources of the National Supercomputing Centre, Singapore. Fushuai Wang was supported by a research scholarship from Prof. Kostya

Novoselov at the National University of Singapore. The authors acknowledge funding from the National University of Singapore Presidential Young Professorship start-up grant (Grant No. A-0010260-00-00).

References

[1] Beatrice W Soh, Zi-En Ooi, Eleonore Vissol-Gaudin, Chang Jie Leong, and Kedar Hippalgaonkar. Automated electrokinetic stretcher for manipulating nanomaterials. *Lab on a Chip*, 23(16):3716–3726, 2023.

[2] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.