

PathGennie: rapid pathway generation via direction-guided adaptive sampling and kinetics

Dibyendu Maity, Suman Chakrabarty*

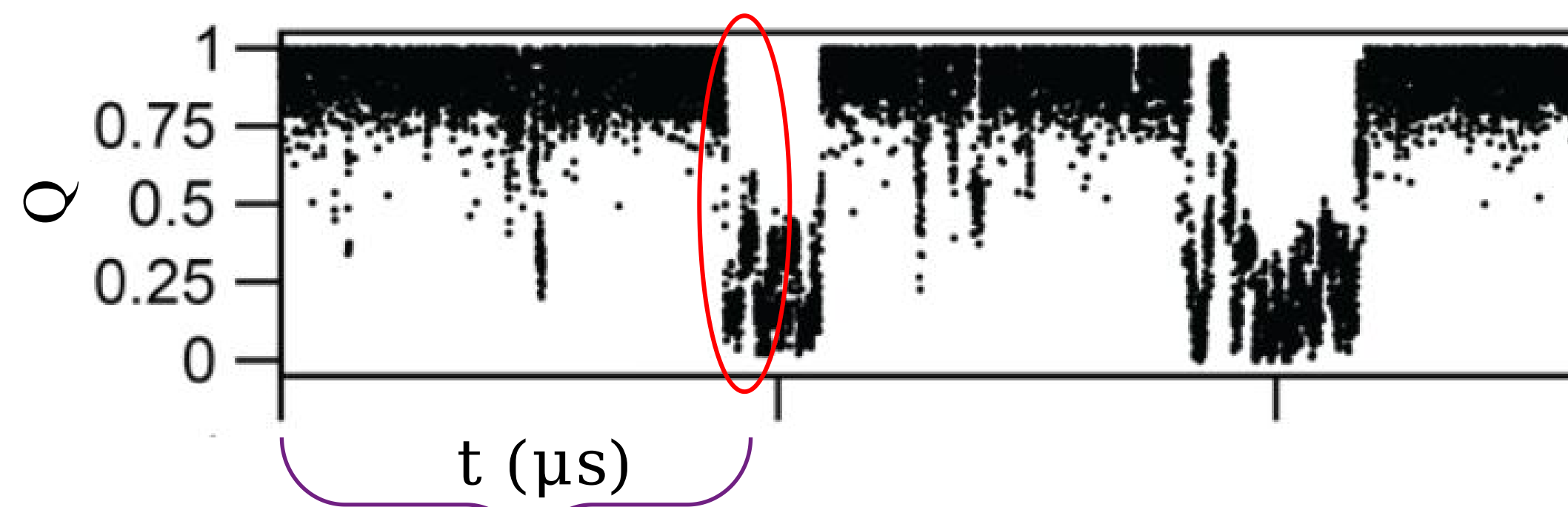
S.N. Bose National Centre for Basic Sciences

Email: dibyendumaity1999@bose.res.in



Motivation: Idea

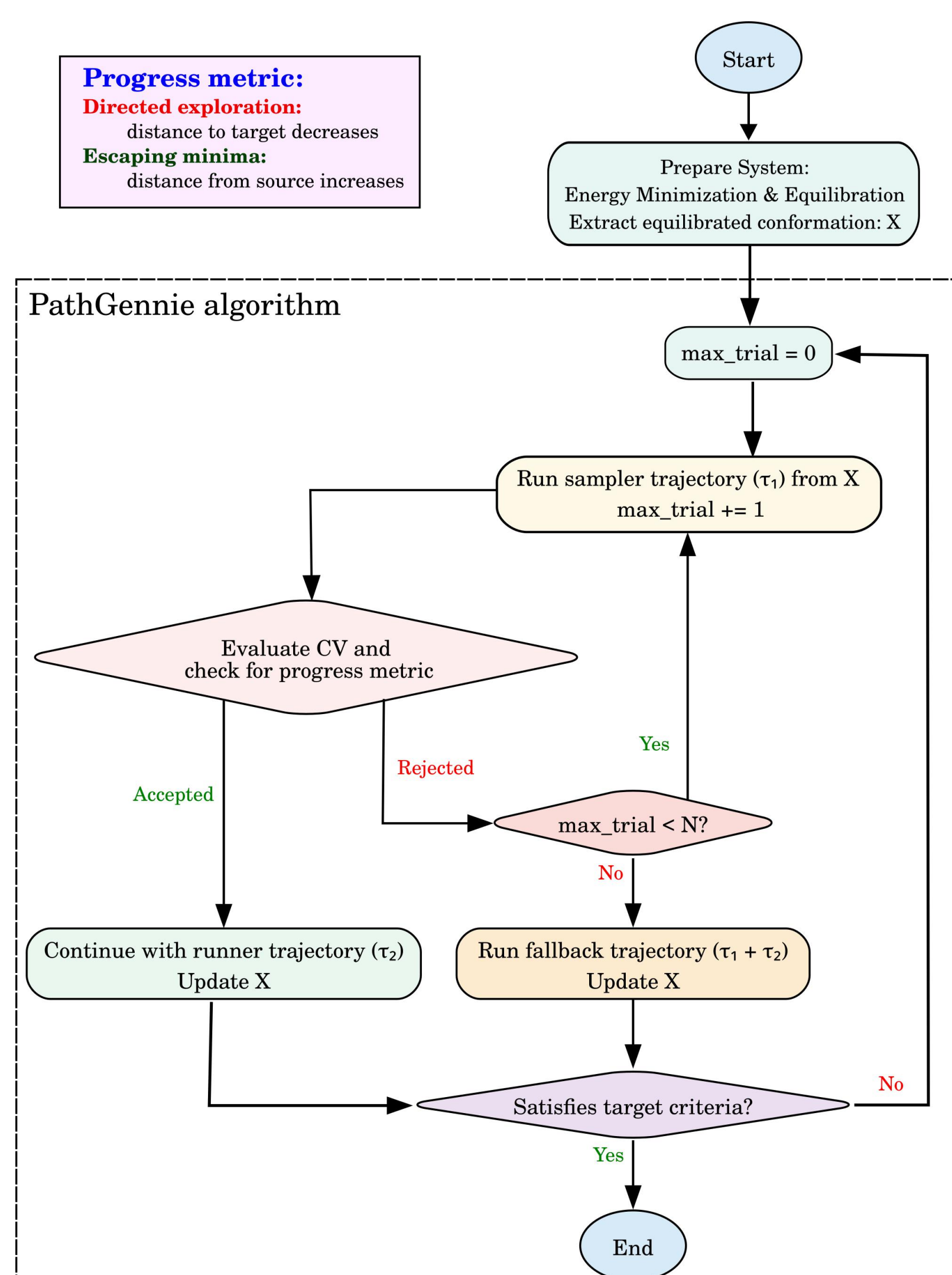
Rare Events: Long waiting time Short transition time



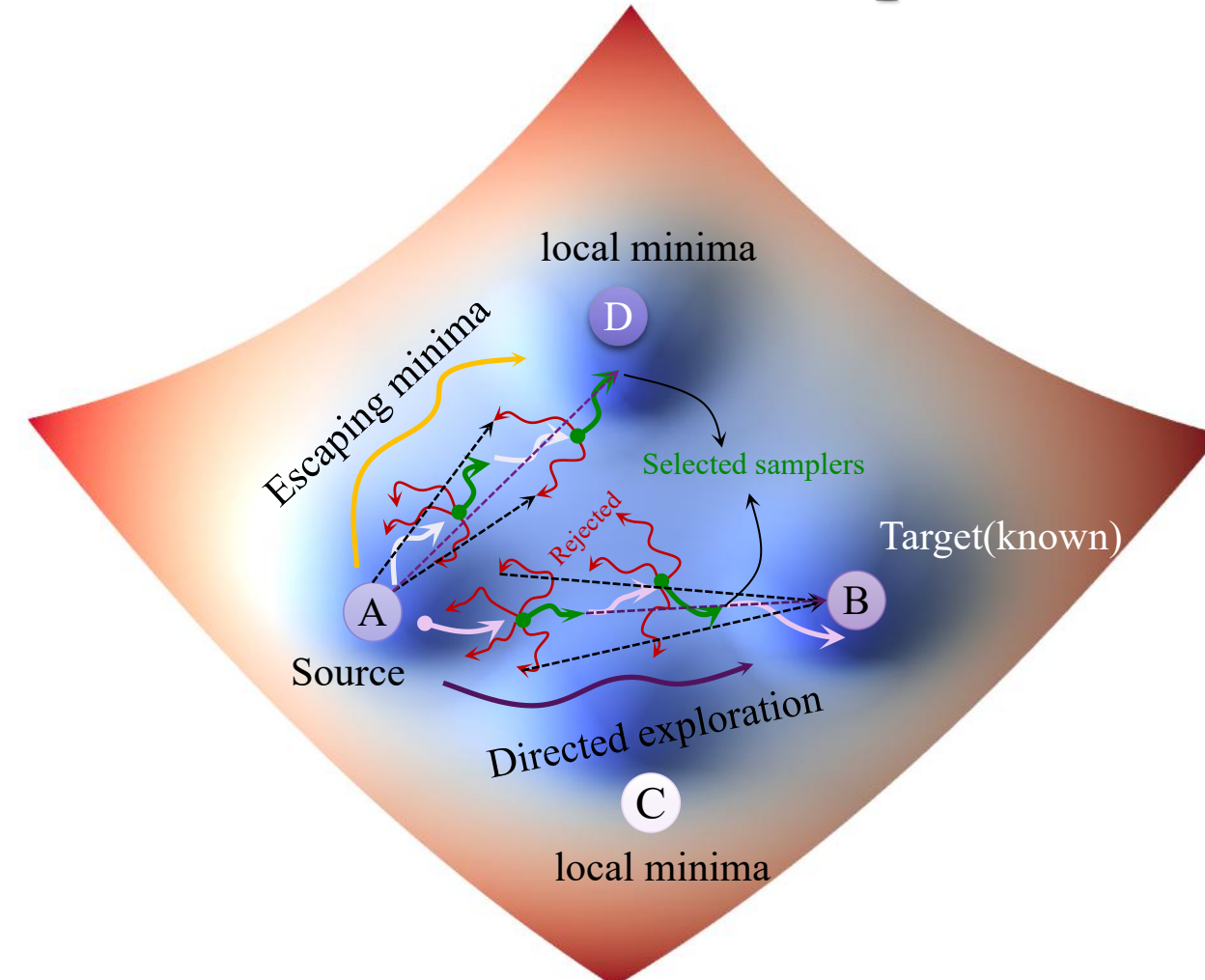
Dream: Can we just skip this part?!

o **Idea:** Use “monitored” “short” trajectories to selectively propagate in a desired direction: Magically finding sequence of events that leads to a successful transition

Algorithm and features

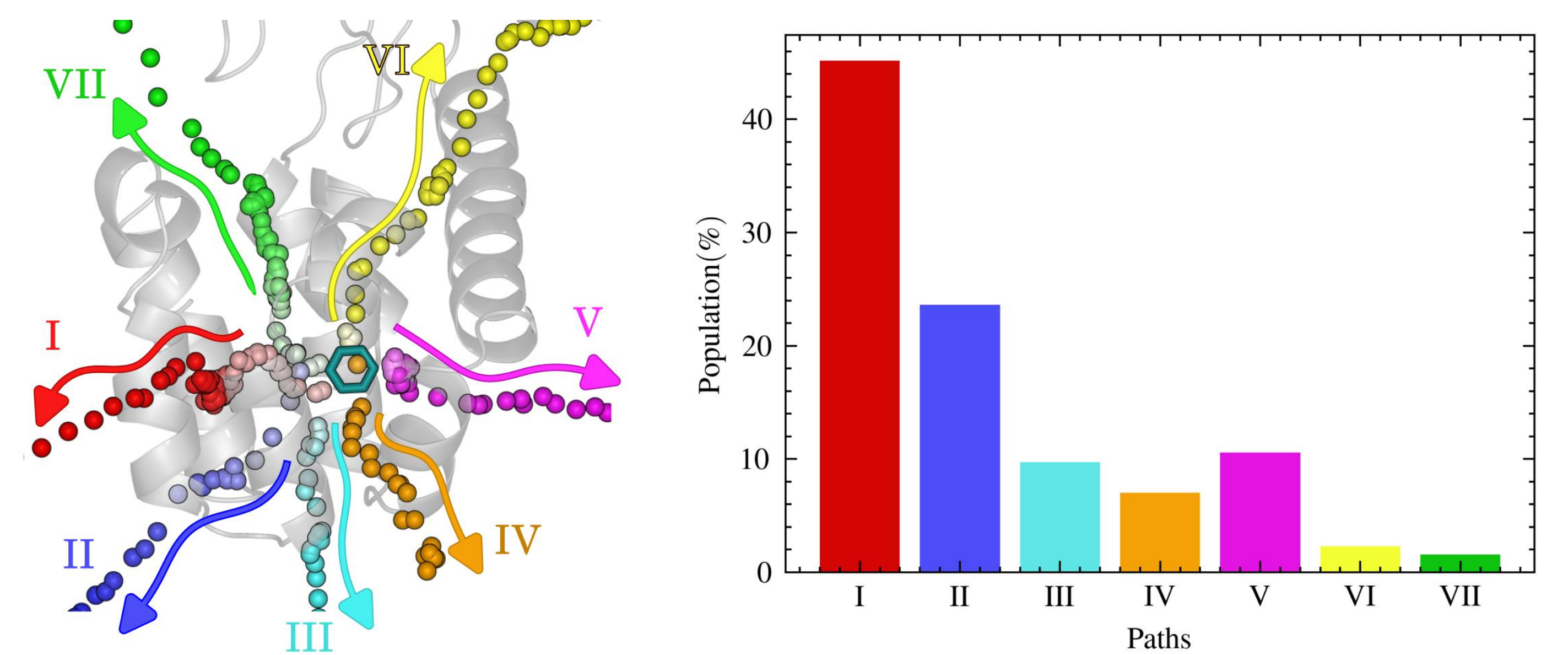


□ Two modes of exploration:



- ✓ Unbiased: No perturbation to hamiltonian or elevated temperature to the system
- ✓ Flexible CVs: Physical or data-driven
- ✓ Rapid: Tunable parameters; parallelizable; intuitive
- ✓ Versatile: From conformational transition to ligand unbinding

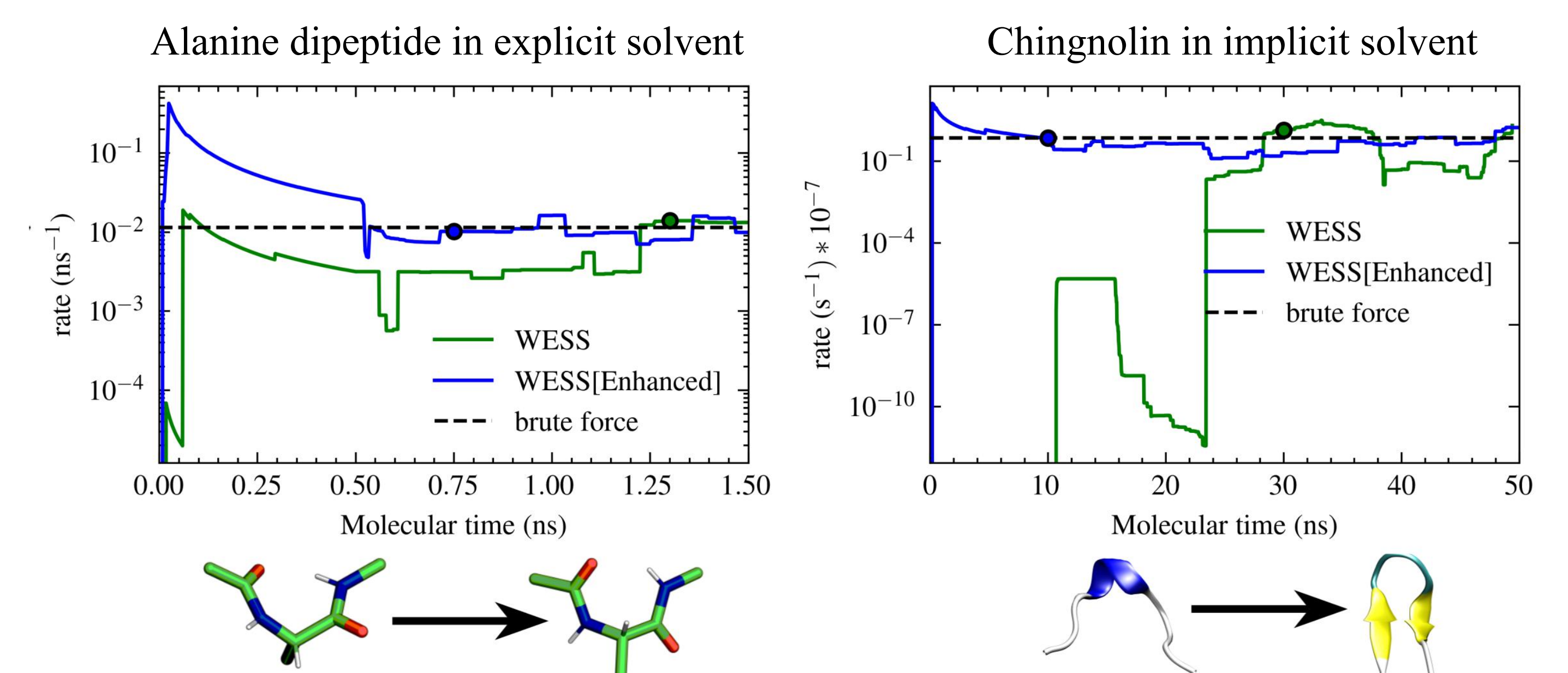
□ Quick generation of unbinding path ensemble for T4L L99A



- ✓ Data-driven CV: PCA on synthetic data: ensemble of ligand-protein poses
- ✓ PathGennie operated on high-dimensional PC space (top 4 components in this case)
- ✓ Millisecond events captured in hundreds of picoseconds
- ✓ Path ensembles preserve relative populations as in τ RAMD simulations

PathGennie + WE: Hybrid approach for enhanced efficiency

✓ Bins in weighted ensemble (WE) are prefilled with states collected from trajectory generated using PathGennie at 0th iteration



Aggregated simulation time to get converged rate:

Alanine dipeptide: P_{II} → α_L
 ✓ Regular WE: 1.7 μ s
 ✓ Enhanced WE: ~600 ns

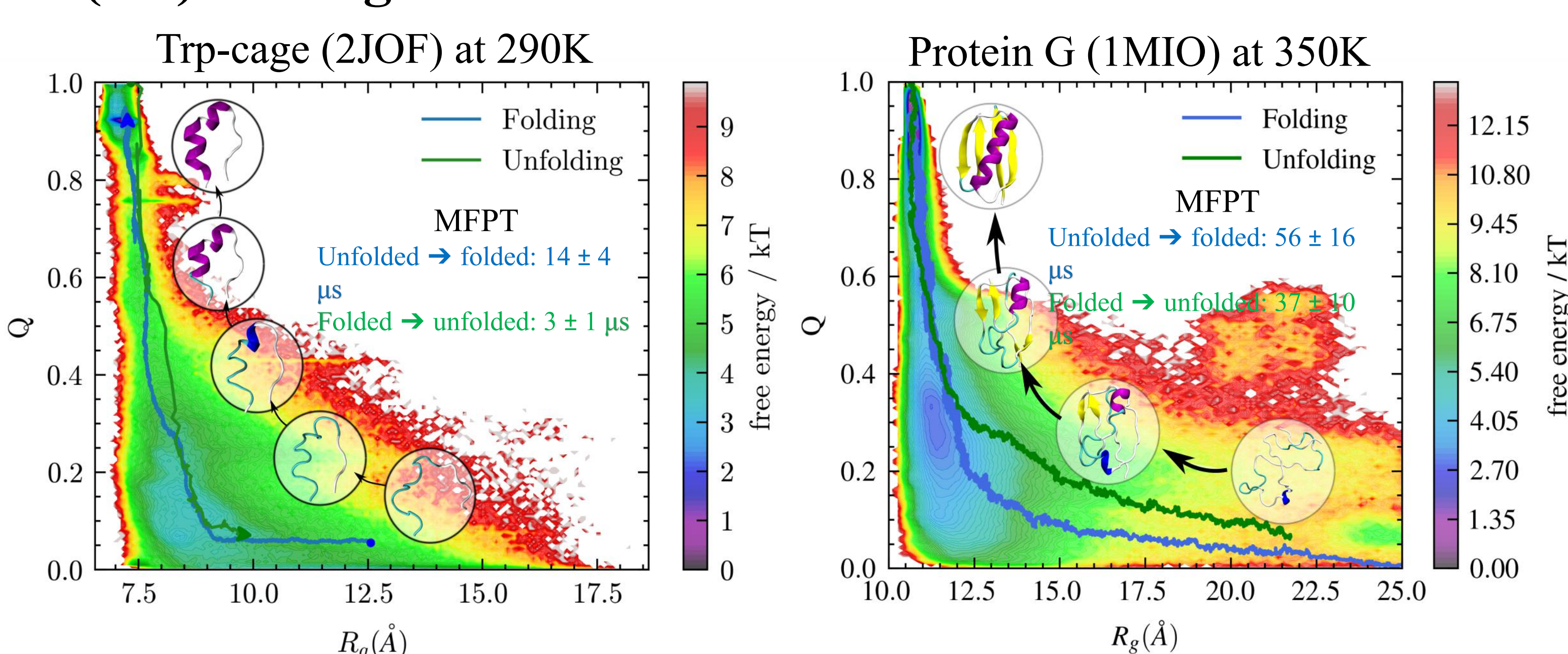
CLN025: Unfolded → folded
 ✓ Regular WE: 1.4 μ s
 ✓ Enhanced WE: ~350ns

Results: Ultrafast generation of reactive trajectory across different test cases

Systems	(N, τ_1 , τ_2) (ps)	Reactive length (ps)	Total length (ps)	Wall-Clock Time
T4L L99A + Benzene (Unbinding)	(25, 0.01, 0.05)	98 ± 22	228 ± 56	~2 min
Abl-kinase + Imatinib (Unbinding)	(25, 0.02, 0.03)	184 ± 49	639 ± 224	~4 min
Trp-cage (Unfolding)	(15, 0.02, 0.08)	150 ± 33	618 ± 139	~1.5 min
Protein G (Unfolding)	(15, 0.01, 0.09)	176 ± 67	312 ± 128	~3 min
Trp-cage (Folding)	(50, 0.01, 0.04)	336 ± 95	430 ± 122	~5 min
Protein G (Folding)	(50, 0.01, 0.04)	612 ± 121	1593 ± 359	~15 min

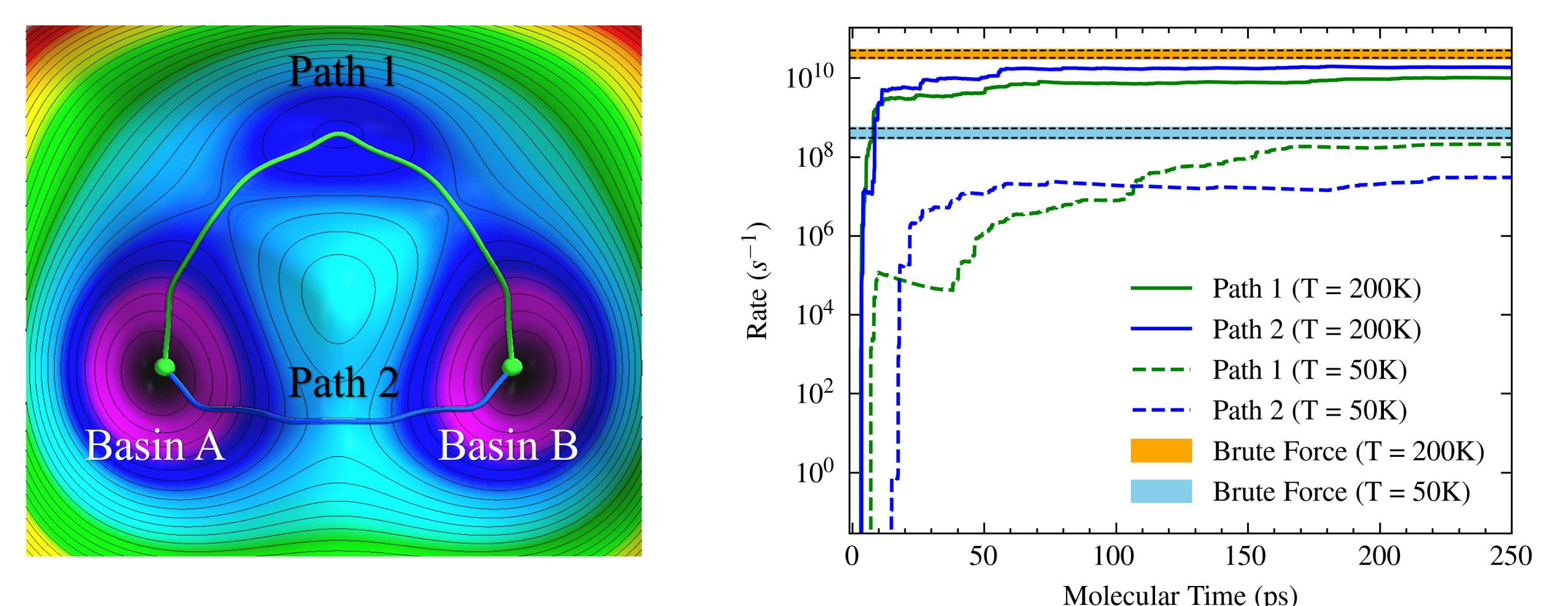
*Total length includes all moves (productive and unproductive) and reflects the simulation's overall cost, while wall-clock time is the average time to generate a reactive trajectory on a single NVIDIA RTX 2080 Ti GPU.

□ (Un)Folding of fast folders:



- ✓ Physical CV: Fraction of native contacts, RMSD (or R_g)
- ✓ PathGennie generated paths follow the minimum free energy pathway!

Path resolved kinetics: Contribution of each path towards the rate of transition



- ✓ Cross-over of path dominance at different thermodynamic condition resolved using PathCV as reaction coordinate for WE and reference path generated using PathGennie

Take home messages!

- ✓ Using PathGennie we can generate ensemble of realistic transition paths quickly
- ✓ Integrating with kinetic models to enhance overall efficiency
- ✓ Code freely available @ github.com/teamsuman/PathGennie