

## Supporting Information (SI)

Saikat Dhibar<sup>1</sup>, Biman Jana<sup>1\*</sup>

### **Designing transferable transition state guided collective variable via interpretable machine learning model for enhanced sampling: A case study on polymer collapse transition**

<sup>1</sup>School of Chemical Science, Indian Association for the Cultivation of Science, Jadavpur, Kolkata-700032

Corresponding Email: [pcbj@iacs.res.in](mailto:pcbj@iacs.res.in).

Note S1:

We have run the NVT equilibration at 300K using Berendsen thermostat for a duration of 2 ns with a coupling constant of 0.1 ps. A 4 ns long NPT simulation was run at 300K temperature and the pressure at 1 bar using V-rescale thermostat and Parrinello-Rahman barostat.

Table S1: Sigma of different CVs for each polymer systems

System	Chosen CV for biasing	Sigma for WT-MetaD
C40	Rg	0.05
C40	LT-1	0.06
C40	TS-derived CV	0.06
C45	Rg	0.05
C45	LT-1	0.06
C45	TS-derived CV	0.07
C60	Rg	0.07
C60	TS-derived CV	0.08
C80	Rg	0.04
C80	TS-derived CV	0.12
C100	Rg	0.04
C100	TS-derived CV	0.10

Table S2: Bias factor for each system:

System	Biasfactor for WT-MetaD
C40	8
C45	8
C60	15
C80	30
C100	50

Figure S1: Sensitivity analysis of EN coefficients with shifted and narrower committor window values.

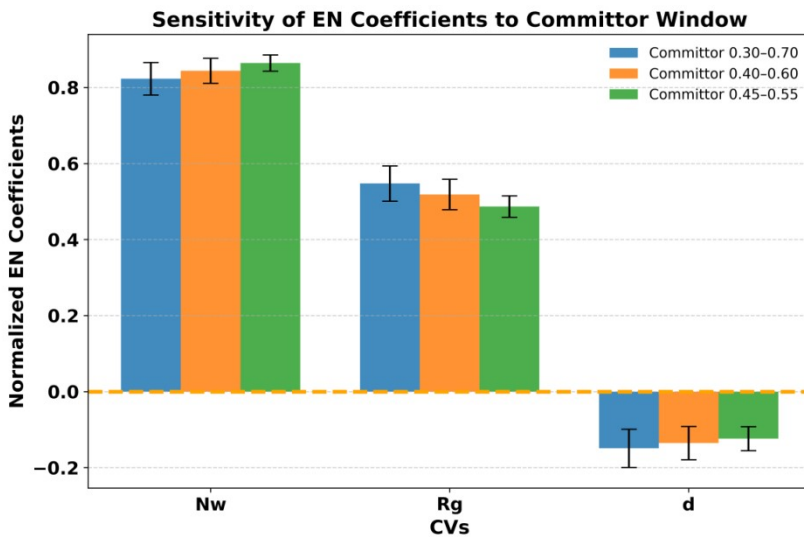


Figure S2: Normalized coefficients of different OPs in latent dimension-1 for C45 system. We have derived the coefficients of different order parameters from the unbiased MD run data processed through time-lagged denoising autoencoder

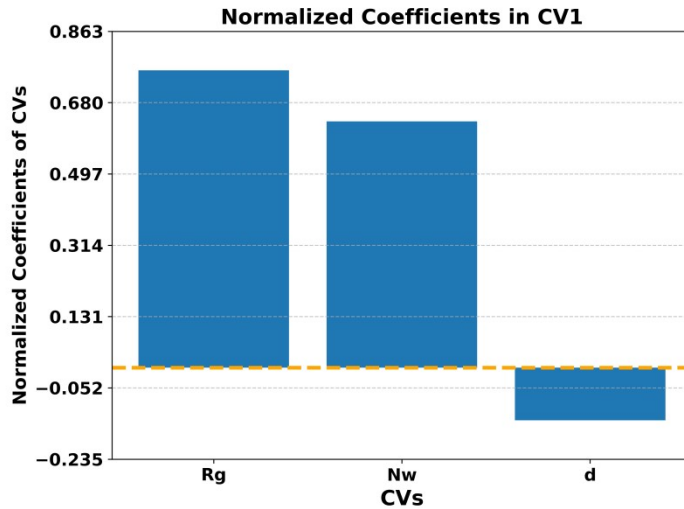


Figure S3: Reweighted FES along  $R_g$  and  $d$  for C45 system within 100 ns.

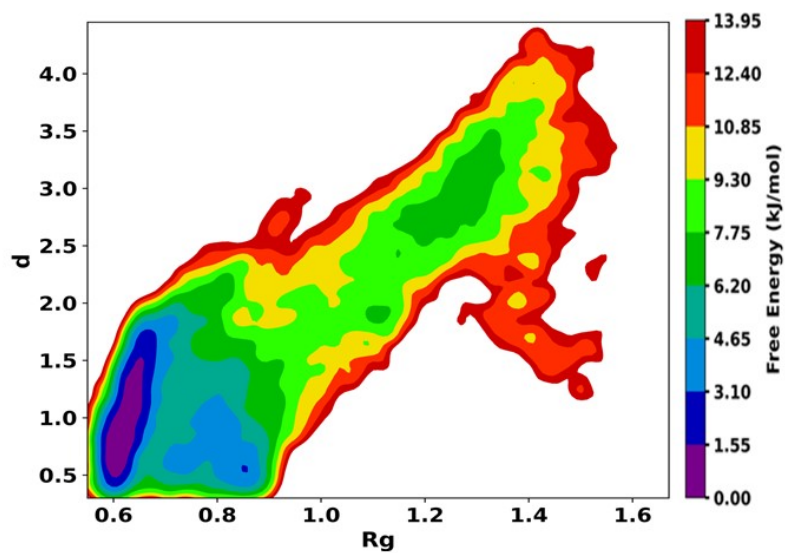


Figure S4: Reweighted FES along  $R_g$  and  $N_w$  for C40 system within 50 ns when biased (A) using TS derived CV, (B) using latent dimension-1 CV, (C) using  $R_g$ .

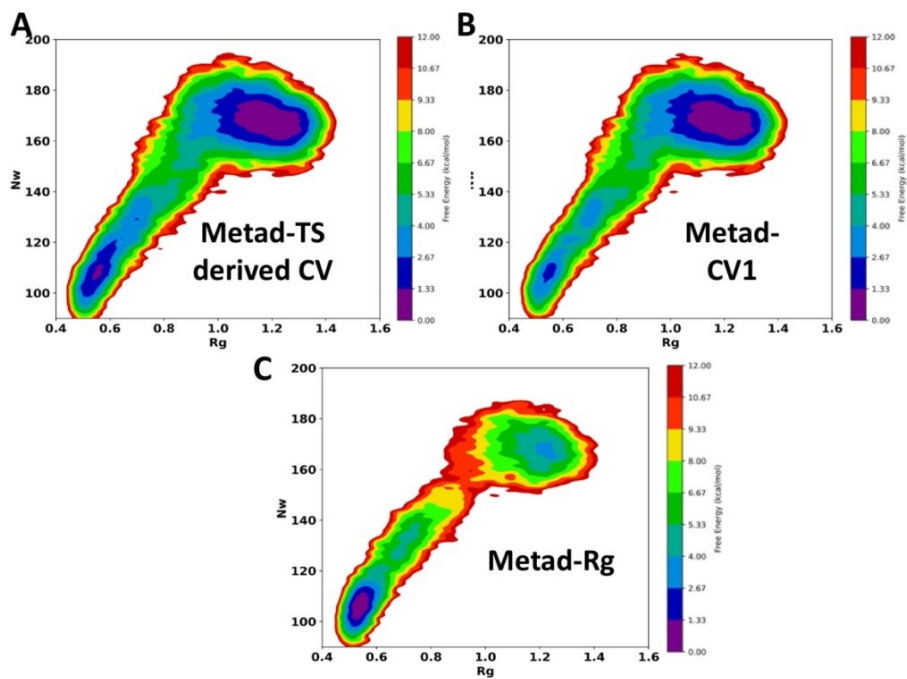


Figure S5: Reweighted FES along Rg and d biasing with suboptimal CV for C45 system within 100 ns

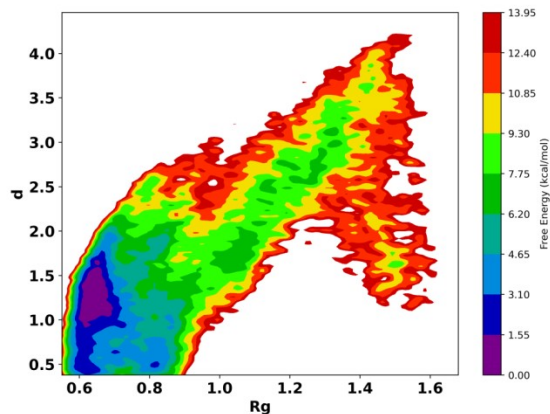


Figure S6: Normalized coefficients of different OPs near TSE for C60 system

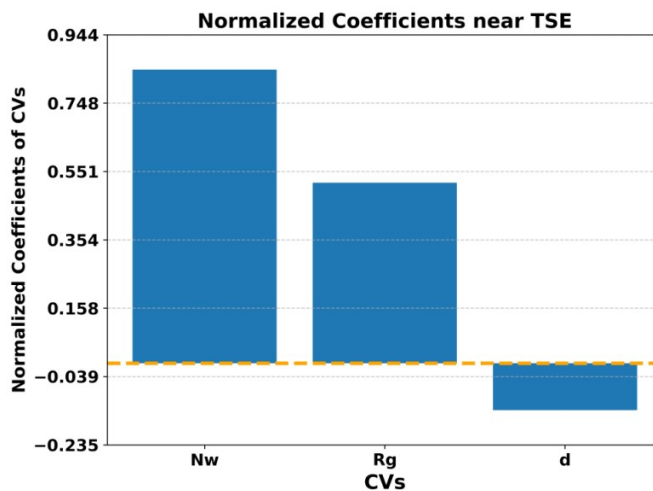


Figure S7: Trajectories biasing with (A) Rg and (B) TS-derived CV for C60 system.

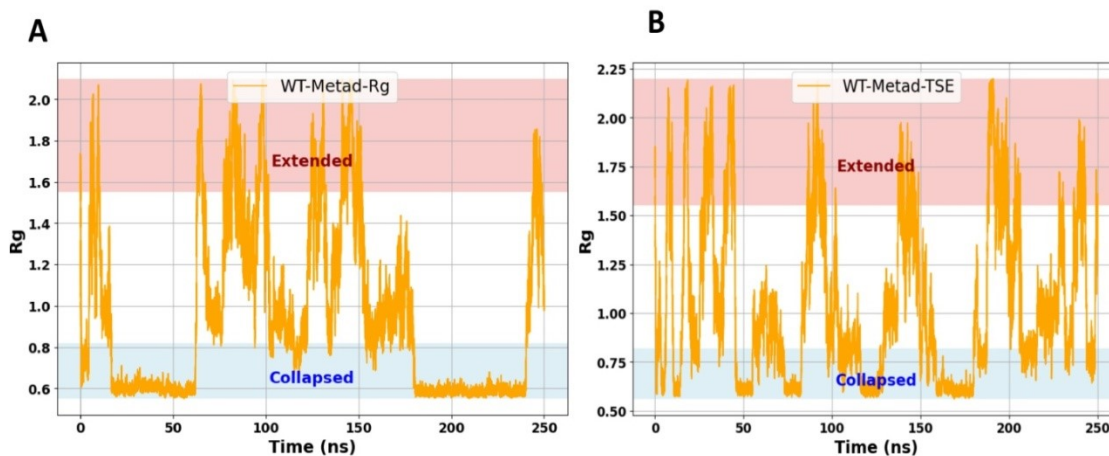


Figure S8: Reweighted FES along Rg and Nw for C60 system (A) within 250ns, (B) within 300 ns and (C) reweighted FES along Rgat different time interval. The metadynamics simulation was biased with Rg

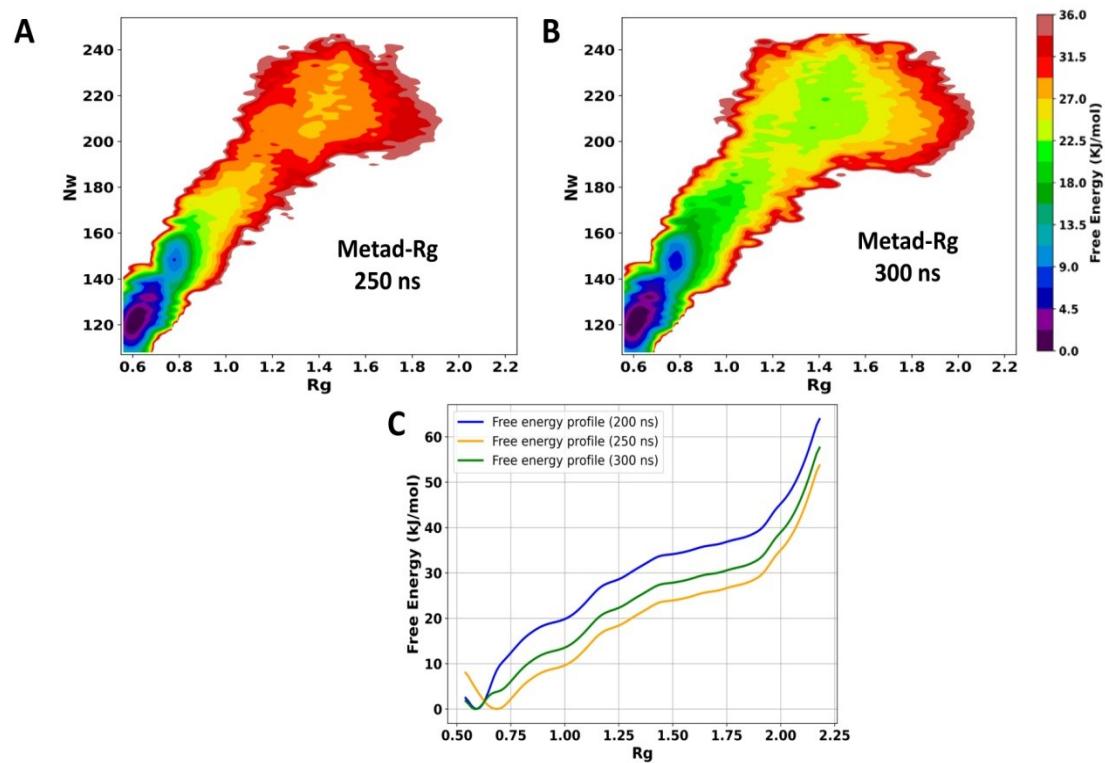


Figure S9: Reweighted FES along Rg and Nw for C80 system after (A) 150 ns, (B) 200 ns, and (C) 250 ns. Reweighted FES along Rg at different time intervals. The metadynamics was biased using TS-derived CV.

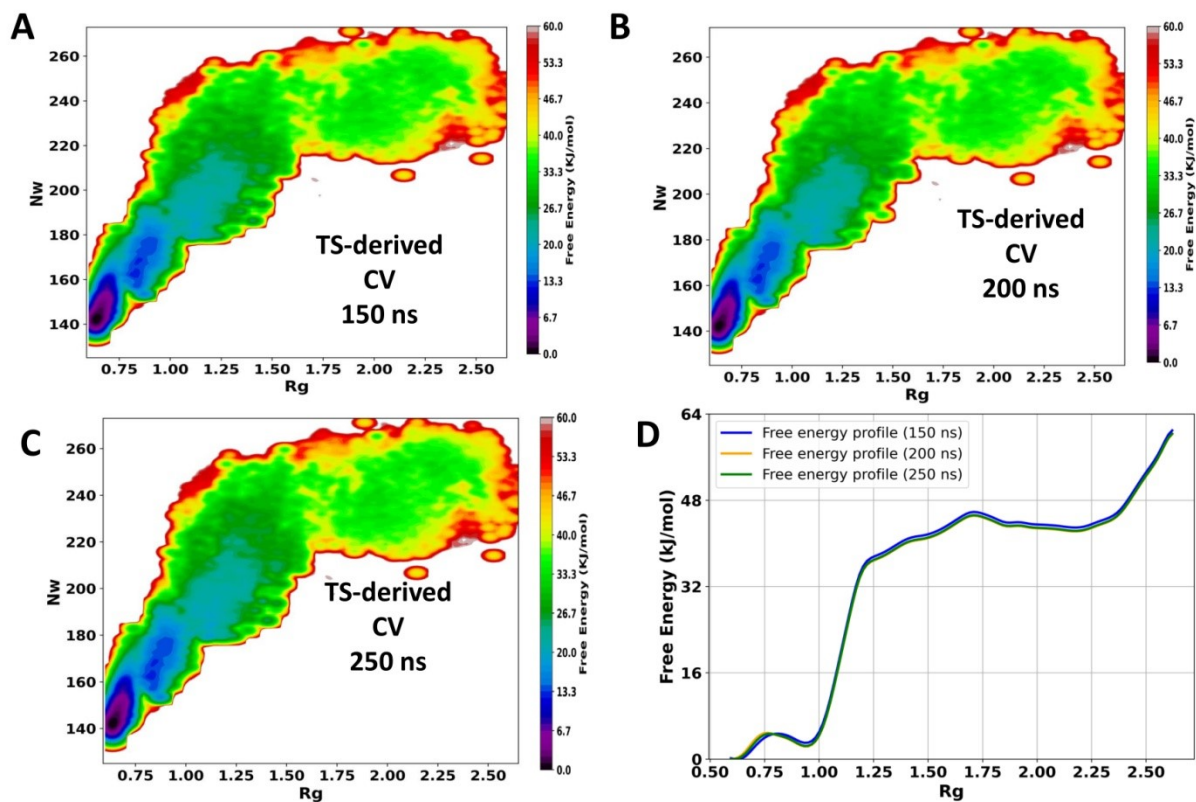


Figure S10: Free energy profile along Rg biasing with TS-derived CV for C80 system with estimated error from block analysis

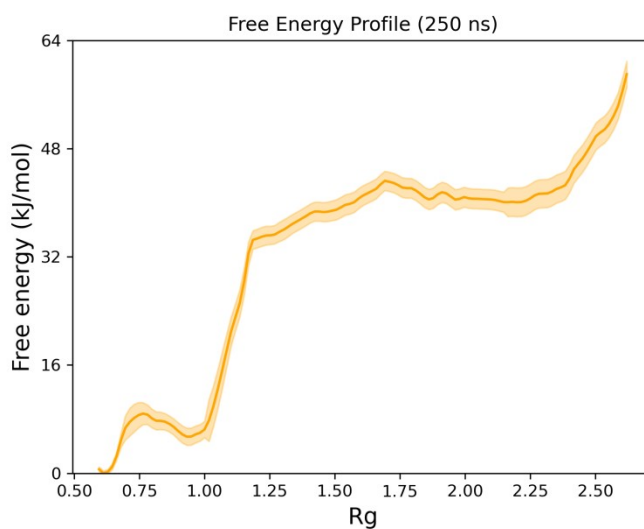


Figure S11: Reweighted FES along Rg and Nw for C100 system after (A) 150 ns, (B) 200 ns, and (C) 250 ns. Reweighted FES along Rg at different time intervals. The metadynamics was biased using TS-derived CV.

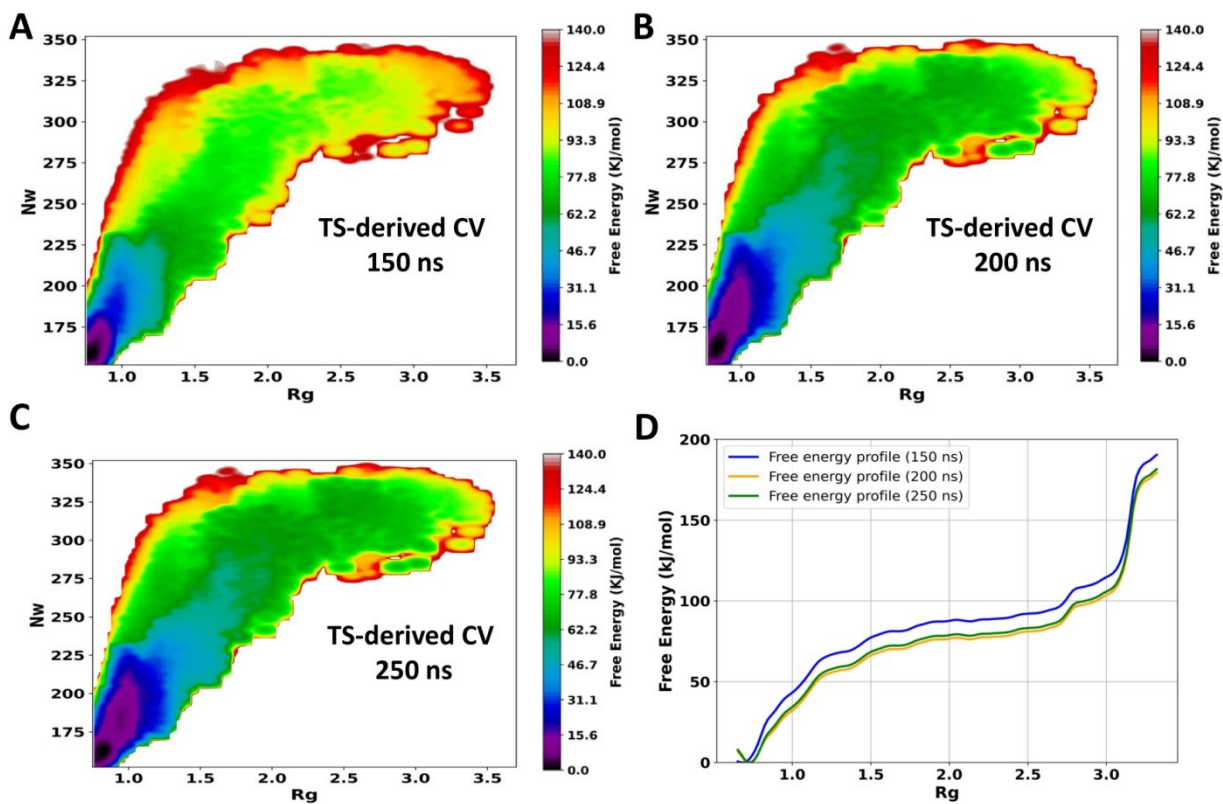


Figure S12: Free energy profile along Rg for C100 system with estimated error from block analysis

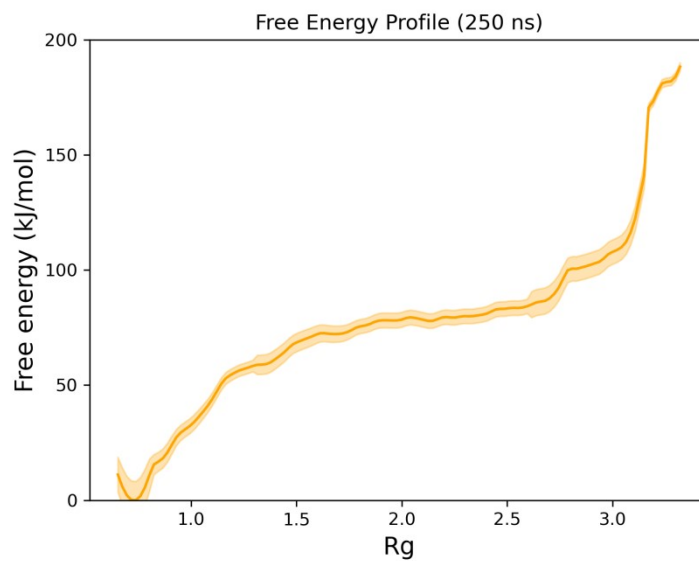


Figure S13: Free energy difference between extended and collapsed basin for C80 and C100 system in case of Rg and TS-derived CV

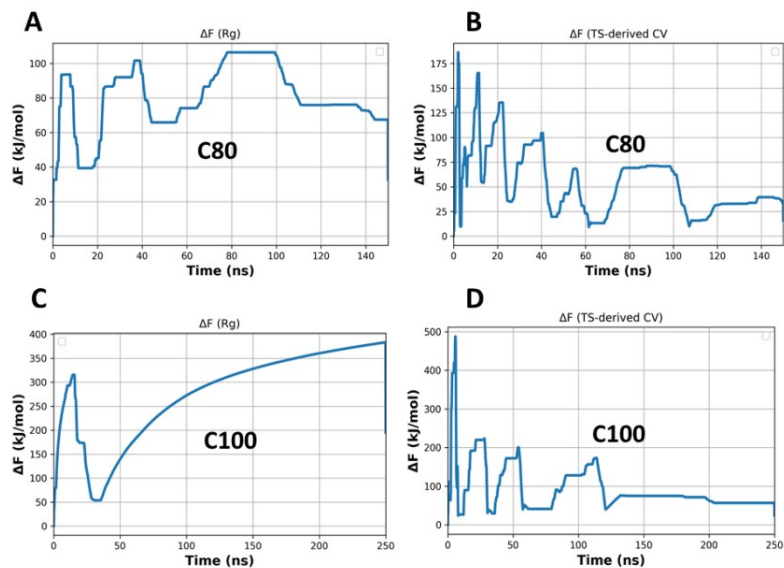


Figure S14: Trajectories biasing with only hydration OP in WT-Metad simulation.

