Supporting Information for

"Enhancing biomolecular sampling with reinforcement learning: tree search molecular dynamics simulation method"

Kento Shin,[†] Duy Phuoc Tran,[†] Kazuhiro Takemura,[‡] Akio Kitao,[‡] Kei Terayama,[*,¶,§,‖] and Koji Tsuda[*,†,¶,⊥]

[†]*Graduate School of Frontier Sciences, the University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan*

[‡]*School of Life Sciences and Technology, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan*

[¶]*RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan*

[§]*Medical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, Kanagawa 230-0045, Japan*

[‖]*Department of Biomedical Data Intelligence, Graduate School of Medicine, Kyoto University, 54 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan*

[⊥]*Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Ibaraki 305-0047, Japan*

E-mail: kei.terayama@riken.jp; tsuda@k.u-tokyo.ac.jp

Phone: +81 (4)-7136-3983. Fax: +81 (4)-7136-3975

- **Supporting Data**: The files of 1UAO_extended.pdb and 1L2Y_extended.pdb are the extended structures of Chignolin and Trp-Cage, respectively.

- **Figure S1**: 2D plot of Chignolin structures and reactive trajectory in each trial onto the subspace spanned by distances of key hydrogen bonds, HB1 and HB2 of PaCS-MD and TS-MD

- **Figure S2**: The projected final snapshots of PaCS-MD and TS-MD to the first and second PCs the 18 NMR models in PDB id 1UAO.

- **Figure S3**: 2D plot of Chignolin structures obtained by PaCS-MD and TS-MD with iteration information

- **Figure S4**: 2D plot of Trp-cage structures obtained by PaCS-MD and TS-MD with iteration information.

- **Figure S5**: Examples of the TS-MD trajectories without penalization mapped to the free energy landscape.

- **Figure S6**: 2D plot of Chignolin structures of TS-MD without penalization with different settings of C.

- **Figure S7**: 2D plot of Trp-Cage structures of TS-MD without penalization with different settings of different C.

- **Figure S8**: Typical examples of tree growths with large C and small C performed by TS-MD without penalization.
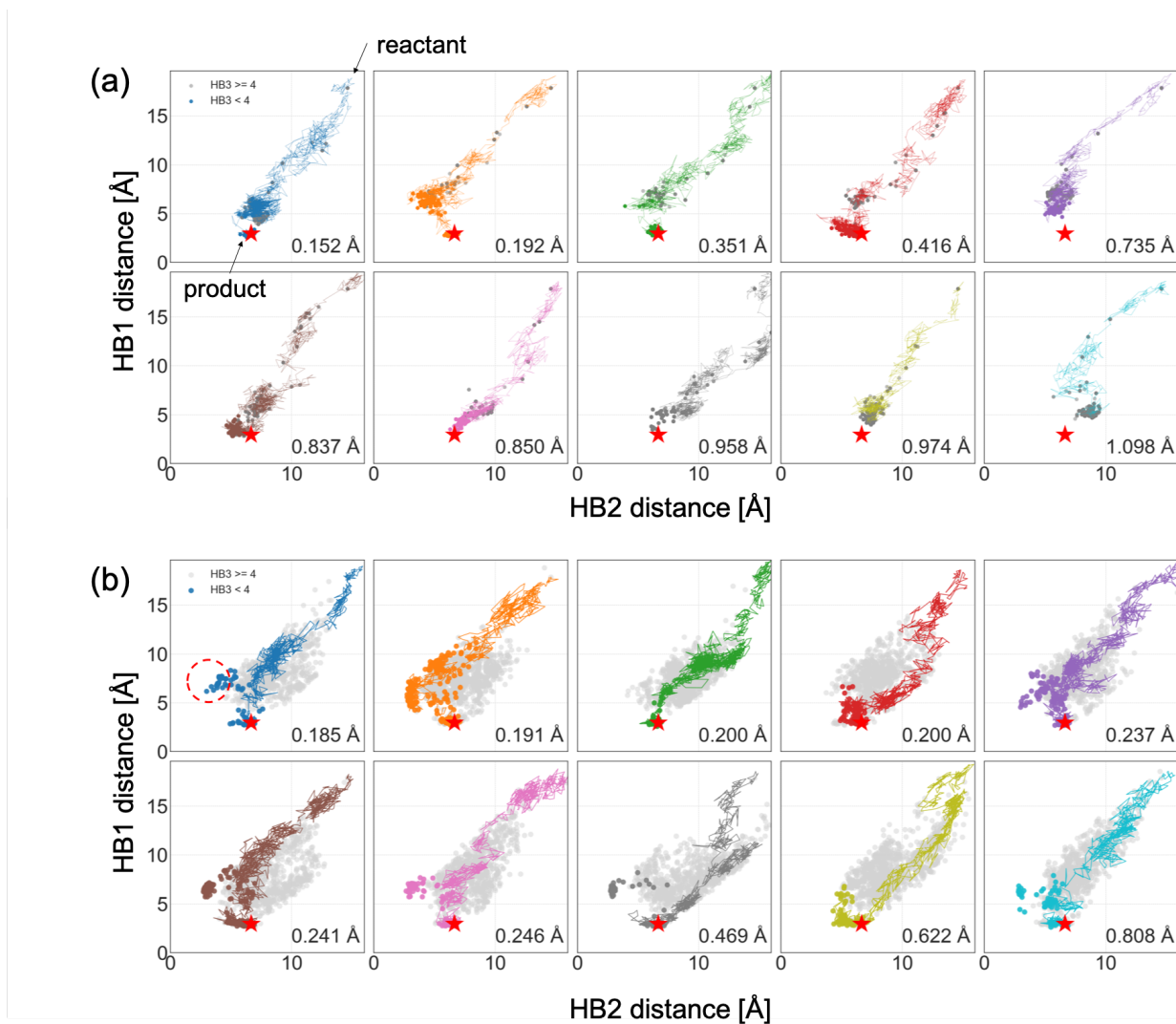
Figure S1: 2D plot of Chignolin structures (points) and reactive trajectory (a polygonal line) in each trial onto the subspace spanned by distances of key hydrogen bonds, HB1 and HB2 of (a) PaCS-MD and (b) TS-MD. Pale gray points are structures with HB3 $\geq$ 4Å and colored points are ones with HB3 $<$ 4Å. The red star is the native structure (product). The minimum RMSD in each trial is written bottom-right. The area surrounded by the red dotted circle indicates the misfolded state where both HB2 and HB3 are formed, which corresponds to the structure surrounded by the same circle in Figure 6. In PaCS-MD, only one (orange) trial can sample the misfolded state. On the other hand, in TS-MD, six trials (blue, purple, brown, pink, gray and light blue) once visited the misfolded state but escaped from that and one (orange) trial found the reactive trajectory to the product passing through the misfolded state.
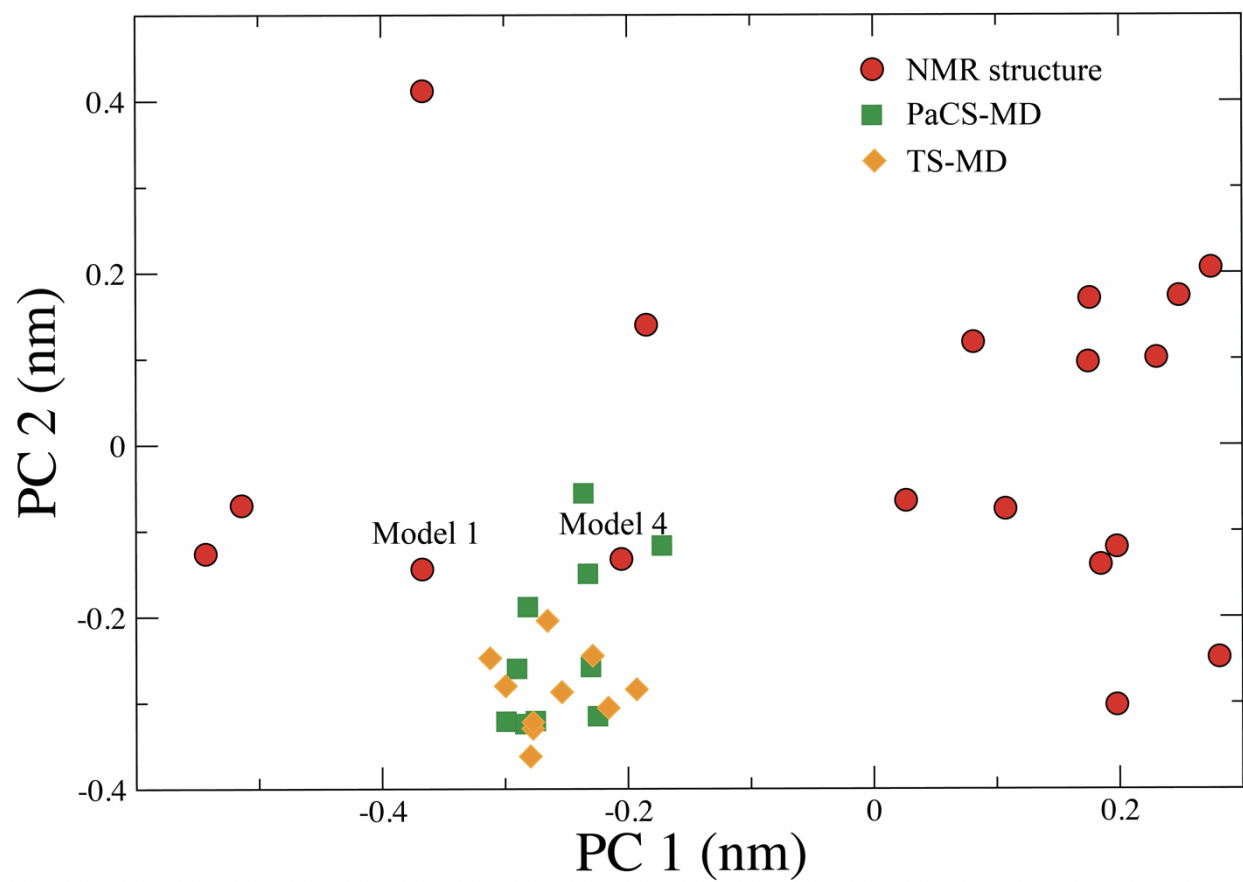
Figure S2: The projected final snapshots of PaCS-MD (green) and TS-MD (orange) to the first and second PCs the 18 NMR models (red) in PDB id 1UAO.
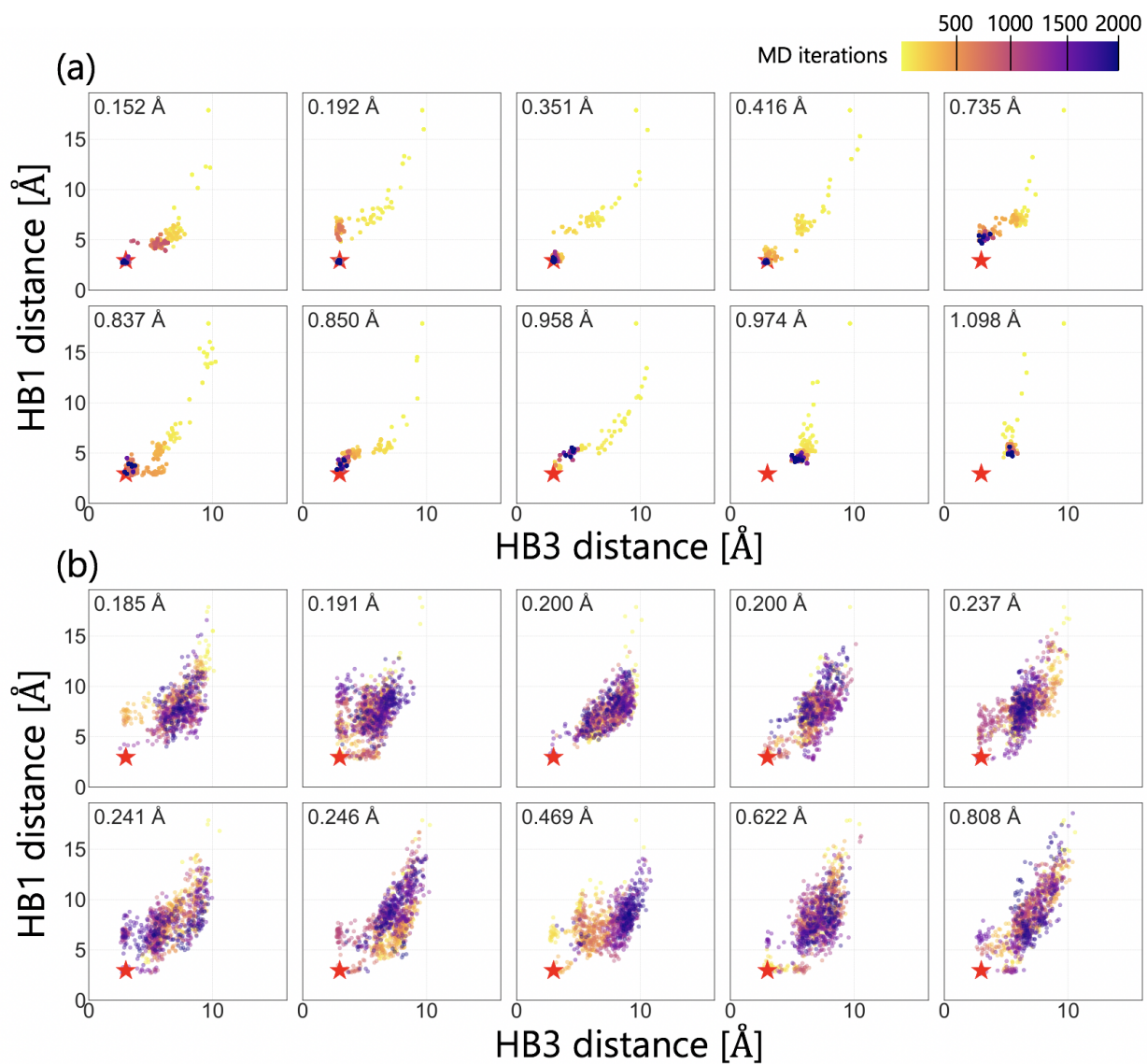
Figure S3: 2D plot of Chignolin structures obtained by PaCS-MD (a) and TS-MD (b) with iteration information. The color of each point indicates the number of MD iterations.
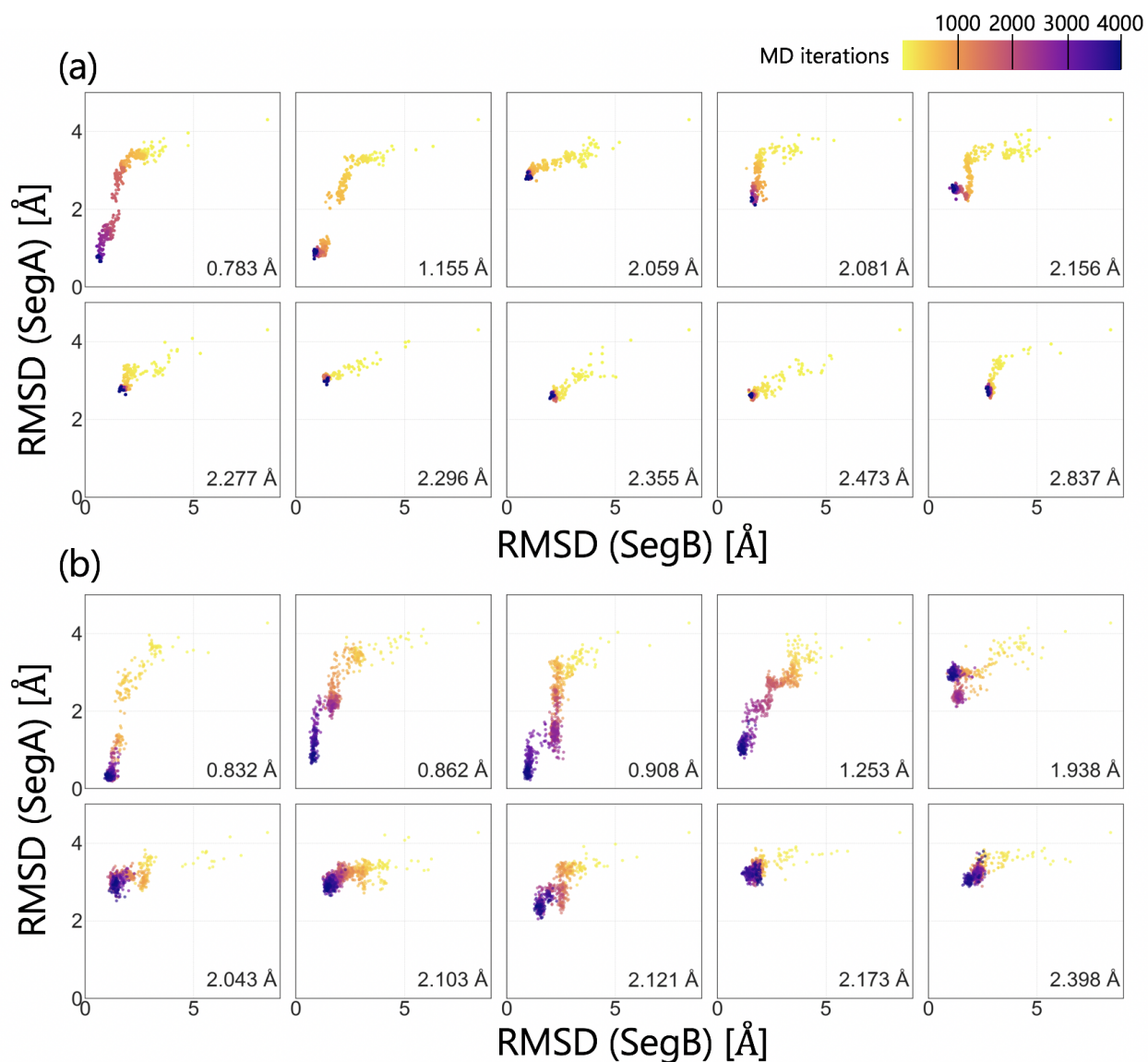
Figure S4: 2D plot of Trp-cage structures obtained by PaCS-MD (a) and TS-MD (b) with iteration information. The color of each point indicates the number of MD iterations.
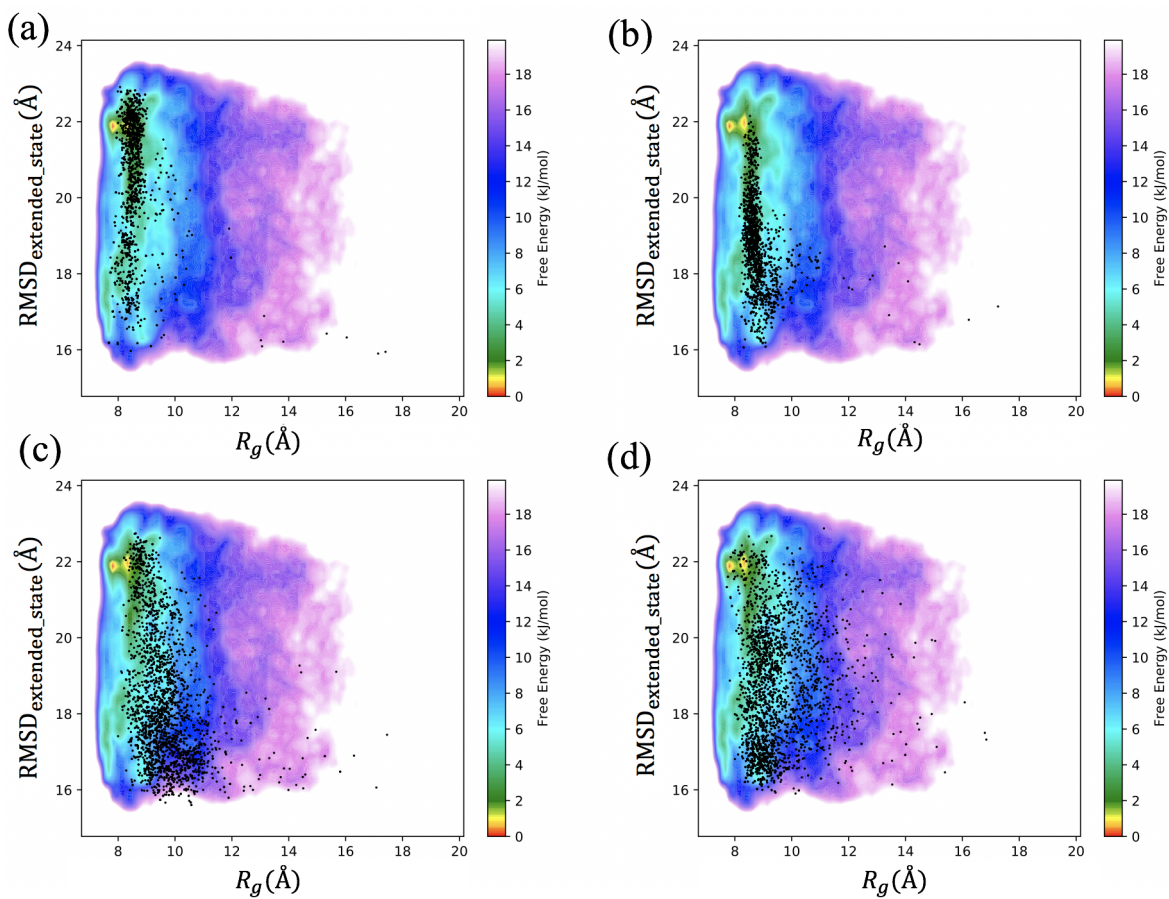
Figure S5: Examples of the TS-MD trajectories without penalization mapped to the free energy landscape with $C = 0.1$ ((a) and (b)) and $C = 0.4$ ((c) and (d)). R_g and $\mathrm{RMSD}_{\mathrm{extended\_state}}$ indicate the radius of gyration and RMSD to the extended state of Trp-cage.
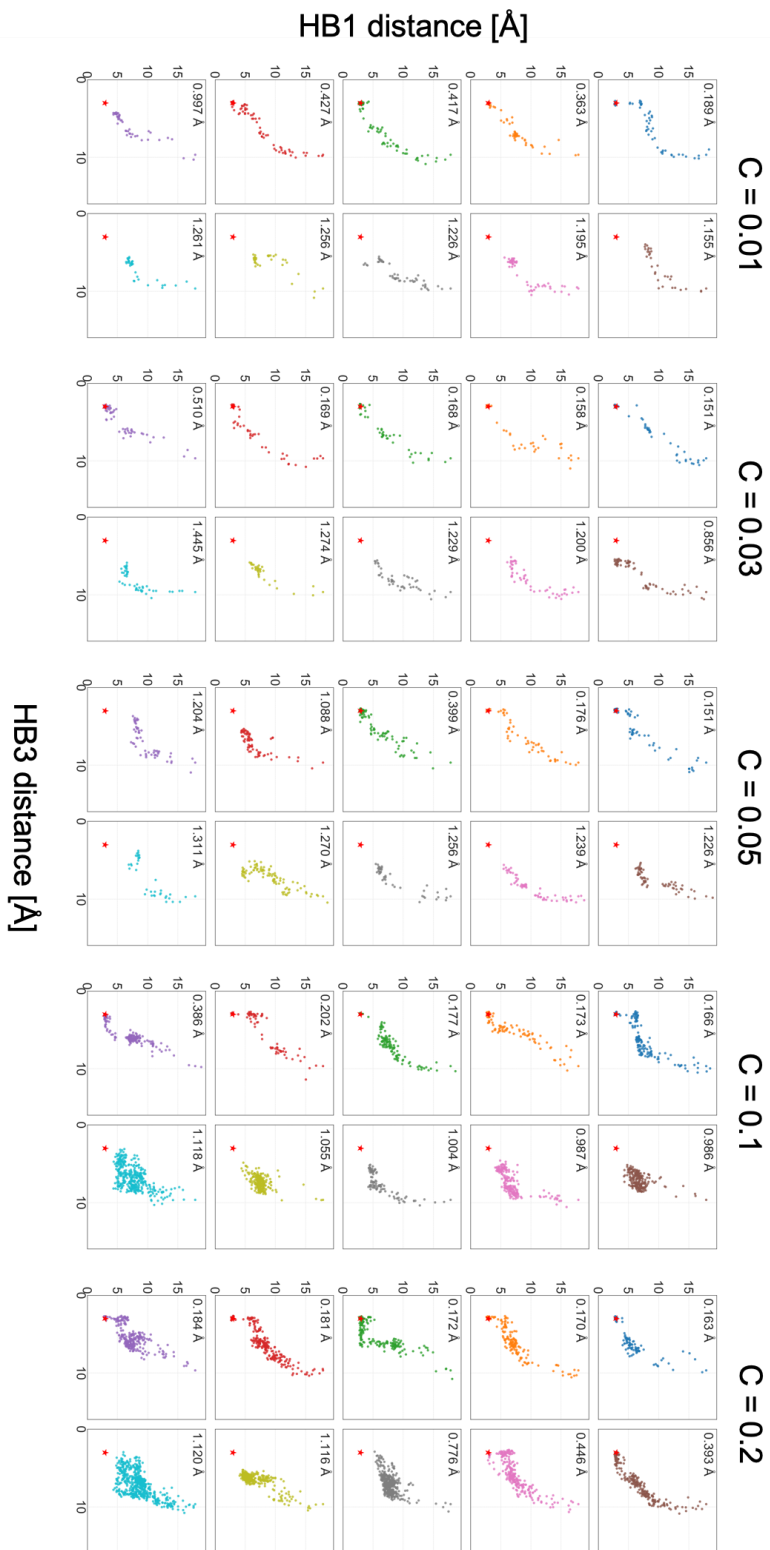
Figure S6: 2D plot of Chignolin structures of TS-MD without penalization with different settings of C (0.01, 0.03, 0.05, 0.1, 0.2). The red star is the native structure (product). The minimum RMSD in each trial is shown top-left.
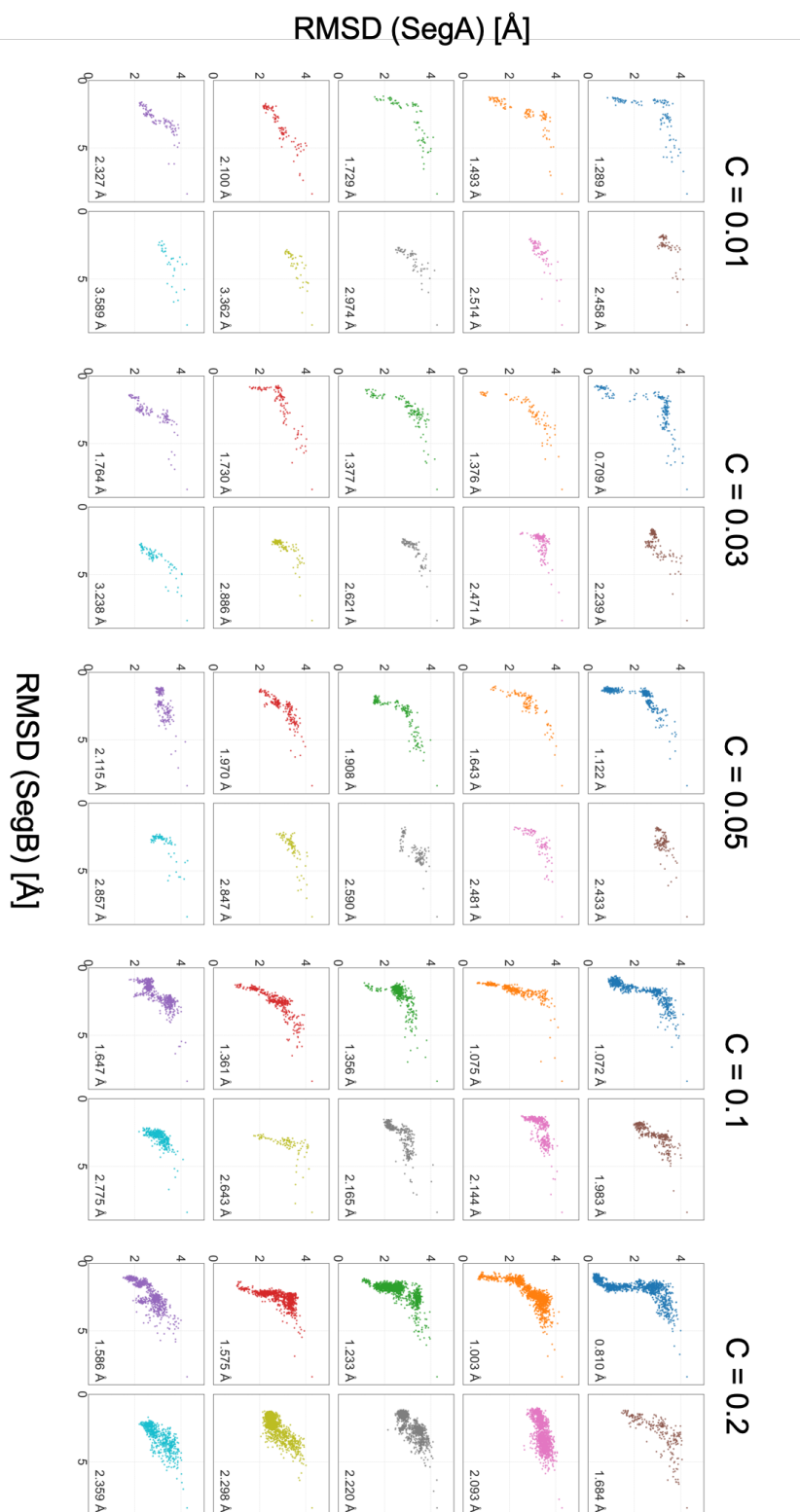
Figure S7: 2D plot of Trp-Cage structures of TS-MD without penalization with different settings of different C (0.01, 0.03, 0.05, 0.1, 0.2). The minimum RMSD in each trial is written bottom-right.
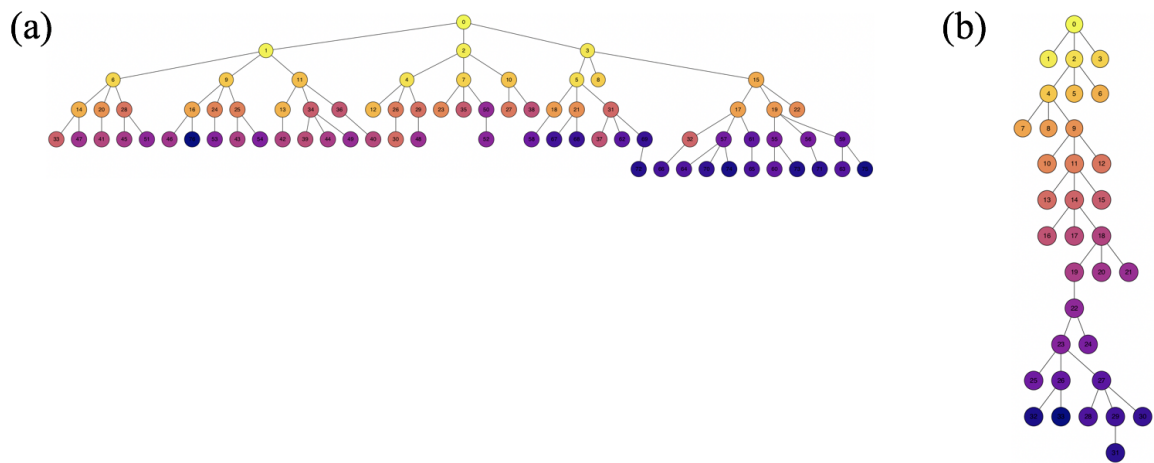
Figure S8: Typical examples of tree growths with large C (a) and small C (b) performed by TS-MD without penalization. The values of C were 0.0001 and 1.