AIMS Biophysics

http://www.aimspress.com/journal/biophysics/

AIMS Biophysics, 13(1): 94–110.
DOI: 10.3934/biophy.2026006
Received: 20 October 2025
Revised: 25 February 2026
Accepted: 02 March 2026
Published: 18 March 2026

Research article

Parallel tempering Monte Carlo simulation of met-enkephalin and WW-domain proteins

M. Rana[1], P. Singh[1,*], N. C. Verma[2] and S. Jain[3]

[1] Department of Physics, Medicaps University, Indore, India

[2] Israel Institute of Technology, Haifa, Israel

[3] Department of Physics, IPS Academy, Indore, India

* **Correspondence:** Email: priya.singh@medicaps.ac.in.

Abstract: Efficient sampling of low-energy conformational states remains a central challenge in protein folding simulations. In this study, we investigated the performance and limitations of standard Monte Carlo (MC) and parallel tempering Monte Carlo (PTMC) methods in exploring protein energy landscapes using two representative systems: the small peptide Met-enkephalin and the FBP28 WW-Domain protein. For Met-enkephalin, standard MC simulations achieved equilibrium at higher temperatures but showed insufficient sampling in low-energy regions. In contrast, PTMC significantly improved sampling efficiency across a wide temperature range, particularly in the low-energy regime, enabling robust exploration of the energy landscape. PTMC successfully identified the global minimum structure of Met-enkephalin with an energy of -11.23 kcal/mol and a root-mean-square deviation of 1.18 Å from the reference structure. The applicability of PTMC to larger proteins was further examined through extensive simulations of the WW-Domain, revealing pronounced energy fluctuations and folding–unfolding transitions at higher temperatures, while sampling at lower temperatures remained limited. The number of visits per unit energy range for two consecutive temperatures have also monitored to check the exchange rate and conformational sampling of the entire landscape. The average energy and specific heat with respect to temperature have also calculated to check the folding unfolding transition of WW-Domain. These results demonstrated the superiority of PTMC over standard MC for protein folding studies; however, they also highlighted challenges associated with larger protein systems.

Keywords: protein-folding; Monte-Carlo simulation (MCS); parallel-tempering (PT); energy landscape

1. Introduction

Proteins are large and complex biomolecules composed of one or more polypeptide chains [1, 2]. Each polypeptide is a linear polymer of amino acid monomers, where peptide bonds are formed between the carboxyl group of one amino acid and the amino group of the next [3, 4]. Proteins are synthesized by translating messenger RNA into a linear amino acid sequence, which initially exists in an unfolded or randomly coiled state [5–7]. These molecules subsequently fold into specific three-dimensional conformations, either globular or fibrous, to perform their biological functions. This transformation is known as protein folding [8–10]. Proteins can adopt a wide variety of conformations [11], each corresponding to a point in a highly complex energy landscape [12]. The study of such biological molecules has fascinated research for decades. Among modern approaches, computer simulations have emerged as a powerful tool to investigate protein folding [13–15] and conformational behavior [5]. Methods such as molecular dynamics (MD) and Monte Carlo (MC) simulations have been instrumental in providing molecular-level insights into various biological phenomena [8, 16, 17]. Despite their success, challenges remain in using these simulations for complex biological systems. One major difficulty lies in accurately sampling the energy landscape of a protein, which may involve either the potential energy or the free energy surface. These landscapes are often rugged, characterized by numerous local minima separated by small energy barriers [18].

Conventional MD or MC simulations can explore such landscapes, but they often suffer from sampling inefficiency. The outcome of a simulation may strongly depend on the initial conditions, and the system may become trapped in local minima, making it difficult to reach the global minimum or explore relevant low-energy states [18]. Classical enhanced-sampling methods [19–23] and [24] such as Metropolis MC, Replica-Exchange MC [25–28], Multi canonical MC, Wang–Landau sampling [29], histogram reweighting, and umbrella tempering improve exploration of complex energy landscapes by modifying acceptance rules, temperature exchange, or statistical weights to efficiently sample rare events and thermodynamic properties. Complementing these, simplified physical models (e.g., Go models), machine-learning–guided protein engineering, large-scale density functional theory, and emerging quantum annealing approaches address protein folding and design from different perspectives, ranging from reduced energy landscapes and data-driven sequence optimization to fully quantum-mechanical treatments and Ising-model–based optimization on quantum hardware.

In the present study, we investigate two protein systems: the Met-enkephalin peptide, a well-characterized pentapeptide, and the FBP28 WW-Domain protein [30–32]. This work critically evaluates the limitations of conventional MC simulations in modeling protein folding and exploring complex energy landscapes. Our analysis of both systems reveals key challenges, including insufficient sampling, slow convergence, and the difficulty of navigating rugged energy surfaces [18]. These shortcomings underscore the need for enhanced sampling strategies, particularly in low-energy regions, where methods such as parallel tempering MC (PTMC) have demonstrated clear advantages [33]. In our study of the Met-enkephalin protein, we applied both standard MC and PTMC simulations, finding that PTMC was markedly more effective at exploring low-energy structures [34]. PTMC not only enhances the likelihood of identifying the global minimum of a Protein Data Bank (PDB) structure but also enables more comprehensive sampling of the low-energy region compared to standard MC simulations [35]. In this work, we employ PTMC simulations to

investigate the energy landscape of the WW-Domain protein [36–38].

2. Methods

2.1. MC simulation

MC methods aim to solve integrals by randomly sampling points in the phase space [39]. To understand the MC algorithm, it is essential to grasp how expectation values are calculated in statistical mechanics. The expectation value of a physical quantity A is computed by adding all possible states of the system, each weighted by a probability function [40, 41]. In many systems, this probability is given by the Boltzmann distribution, $w(r) = \exp(-\beta E_r)$, where, $\beta = 1/k_B T$, E_r is the energy of state r, and the partition function Z normalizes the distribution. Since, the number of possible states ω increases exponentially with system size, direct computation is impractical [1, 42]. MC simulations overcome this by sampling only the most probable configurations, especially those with high statistical weight, to efficiently estimate the expectation value. The Metropolis algorithm is employed to sample the configurational space efficiently [43, 44]. Given a current state r, a new state r' is proposed by applying a random perturbation. The move is accepted with probability [45]:

$$P_{\text{accept}} = \min\left(1, \frac{e^{-\beta E_{r'}}}{e^{-\beta E_r}}\right) = \min\left(1, e^{-\beta(E_{r'} - E_r)}\right) \tag{2.1}$$

If the move is accepted, the system transitions to r'; otherwise, it remains in r. This procedure ensures detailed balance and convergence to the equilibrium distribution. MC simulation is a powerful computational technique used to explore the conformational space of protein all-atom models by employing stochastic sampling methods. In this approach, random perturbations—such as changes in torsion angles, side-chain rotations, or atomic positions—are applied to the atomic coordinates of a protein structure. Each new conformation is evaluated using a defined energy function, often derived from molecular mechanics force fields, and accepted or rejected based on the Metropolis criterion, which allows for thermal fluctuations and helps the system escape local energy minima [43]. This probabilistic framework enables the efficient exploration of the protein's energy landscape, making MC simulations particularly useful for studying protein folding, conformational changes [19].

2.2. Parallel tempering simulation

PTMC, also known as replica exchange MC [46], is an advanced simulation technique used to efficiently explore the complex energy landscape of protein models [47]. In PTMC, multiple replicas of the protein system are simulated simultaneously at different temperatures using standard MC methods. Periodically, configurations between replicas at adjacent temperatures are exchanged based on a probability criterion that ensures detailed balance [48].

$$P_{\text{swap}} = \min\left(1, \exp\left[(\beta_i - \beta_j)(E_j - E_i)\right]\right) \tag{2.2}$$

The replica exchange criterion satisfies detailed balance with respect to the joint Boltzmann distribution of all replicas. As a result, each temperature stream samples the correct canonical ensemble corresponding to its assigned temperature.Therefore, thermodynamic averages evaluated at a simulated temperature are canonical ensemble averages at that temperature. This exchange allows

lower-temperature replicas to overcome high-energy barriers by occasionally sampling higher-energy configurations from the hotter replicas, thus improving the overall sampling efficiency. The method is particularly effective for protein systems, which often have rugged energy landscapes with many local minima, making conventional MC simulations prone to getting trapped. By enabling a more thorough exploration of conformational space, PTMC provides more accurate estimates of thermodynamic properties and helps identify relevant protein folding pathways and stable conformations [49].

PTMC belongs to the class of generalized statistical ensemble methods, in which multiple replicas simulated at different temperatures collectively sample a joint Boltzmann distribution while preserving canonical ensemble sampling at each individual temperature. This property ensures that thermodynamic observables evaluated at a given replica temperature correspond to proper canonical ensemble averages.

2.3. Thermodynamic averages and error analysis

PTMC simulations belong to the class of generalized-ensemble methods. In PTMC, multiple replicas are simulated simultaneously at fixed replica temperatures T_r, and periodic exchanges are attempted between neighboring temperatures. Each replica samples a canonical ensemble at its own temperature T_r.

In the present work, thermodynamic quantities at temperature T_r were computed using only configurations collected when the system was visiting that replica temperature. No cross-temperature reweighting was performed.

Let $E_{r,i}(T_r)$ denote the energy of the i-th configuration from trajectory r sampled at temperature T_r. The average energy for the r-th independent trajectory is given by

$$\langle E\rangle_r(T_r) = \frac{1}{N_r(T_r)} \sum_{i=1}^{N_r(T_r)} E_{r,i}(T_r), \tag{2.3}$$

where $N_r(T_r)$ is the number of sampled configurations at temperature T_r.

The final average energy was obtained by averaging over $M = 8$ statistically independent trajectories:

$$\langle E\rangle(T_r) = \frac{1}{M} \sum_{r=1}^{M} \langle E\rangle_r(T_r). \tag{2.4}$$

Because configurations within a single MC trajectory are correlated due to the Markov-chain nature of the sampling, individual configurations were not treated as statistically independent samples. Instead, thermodynamic quantities were first computed separately for each independent trajectory, and statistical uncertainties were estimated from fluctuations among the $M = 8$ independent runs.

The statistical uncertainty (standard error) of the average energy was estimated as

$$\sigma_E(T_r) = \sqrt{\frac{1}{M(M-1)} \sum_{r=1}^{M} [\langle E\rangle_r(T_r) - \langle E\rangle(T_r)]^2}. \tag{2.5}$$

The specific heat at constant volume was computed from canonical energy fluctuations within each trajectory:

$$C_{V,r}(T_r) = \frac{\langle E^2 \rangle_r(T_r) - \langle E \rangle_r^2(T_r)}{k_B T_r^2}, \tag{2.6}$$

where

$$\langle E^2 \rangle_r(T_r) = \frac{1}{N_r(T_r)} \sum_{i=1}^{N_r(T_r)} E_{r,i}^2(T_r). \tag{2.7}$$

The final specific heat was obtained by averaging over the independent trajectories:

$$C_V(T_r) = \frac{1}{M} \sum_{r=1}^{M} C_{V,r}(T_r), \tag{2.8}$$

with statistical uncertainty estimated as

$$\sigma_{C_V}(T_r) = \sqrt{\frac{1}{M(M-1)} \sum_{r=1}^{M} \left[C_{V,r}(T_r) - C_V(T_r)\right]^2}. \tag{2.9}$$

We note that generalized-ensemble simulations may alternatively be analyzed using reweighting approaches such as the weighted histogram analysis method (WHAM) [50, 51]. However, since thermodynamic quantities in the present work were evaluated directly at their respective replica temperatures without cross-temperature reweighting, such reconstruction was not required.

3. Details of the simulation

All-atom simulations were performed for two polypeptide systems using the ECEPP/3 (Empirical Conformational Energy Program for Peptides, version 3) [52] force field, which accounts for electrostatic interactions, hydrogen bonding, van der Waals interactions, and torsional energies. Solvent effects [53–56] were neglected to reduce computational cost and to enable a direct and consistent comparison of sampling efficiency between standard MC and PTMC methods. This approximation allows a focused assessment of algorithmic performance.

Simulations were carried out using a locally modified version of the Simple Molecular Mechanics of Proteins (SMMP) package. In each MC step, a single dihedral angle was randomly selected and reassigned a value uniformly distributed between -180° and +180°.

To evaluate convergence and sampling efficiency, Met-enkephalin was first simulated using standard MC at four temperatures (200 K, 250 K, 300 K, and 350 K). Each simulation consisted of 1.5 million MC steps and results were averaged over 8 independent runs starting from different initial structures. Equilibration was assessed by monitoring total energy, and the lowest-energy conformations were compared with reported literature values. Energy histograms were also analyzed to examine convergence and low-energy sampling.

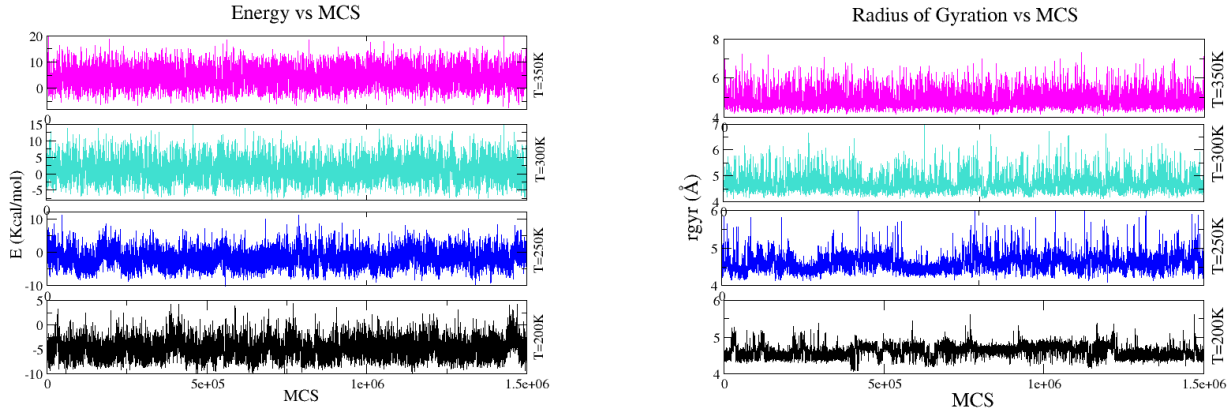
Following validation on Met-enkephalin, simulations were extended to the WW-Domain protein. Standard MC simulations showed poor convergence and insufficient sampling of low-energy states. Consequently, PTMC simulations were employed, which significantly enhanced sampling efficiency across temperatures. Results were averaged over 8 statistically independent trajectories starting from

different initial structures. Energy trajectories, replica exchanges between neighboring temperatures, and average energy and specific heat as a function of temperature were analyzed. The resulting energy–temperature profile captured the folding–unfolding transition of the WW-Domain, demonstrating the effectiveness of PTMC in exploring complex protein energy landscapes and accessing low-energy conformations [57, 58].

4. Results and discussion

4.1. Met-Enkephalin

Figure 1 presents the MC simulation data, where panel (a) shows the variation of total energy E(kcal/mol) with MC steps (MCS), and panel (b) illustrates the corresponding changes in the radius of gyration (Rgyr) with MCS. The energy versus MCS data was shown for 15 Lakh steps at the temperature range from 200 K, 250 K, 300 K, and 350 K. In Figure 1(a), each curve stabilizes around a constant mean value after initial fluctuations, confirming equilibration at each temperature. Higher temperatures show broader fluctuations, as expected, due to increased thermal energy, while lower temperatures exhibit narrower energy ranges. In Figure 1(b), Rgyr vs MCS shows the radius of gyration Rgyr fluctuations for Met-enkephalin at four temperatures (200 K, 250 K, 300 K, 350 K) over the course of the MC simulation. Across all temperatures, Rgyr remains within a relatively stable range, indicating that the system has reached and maintained structural equilibrium, with minor fluctuations reflecting thermal motion.



(a) **(b)**

Figure 1. MC simulation results for Met-enkephalin: (a) total energy as a function of MCS, and (b) Rgyr as a function of MCS.

The minimum-energy structure obtained from our simulations is presented in Figure 2. Figures 2(a) and 2(b) show the high-energy input structures, with energies of -1.2 kcal/mol and 4.03 kcal/mol, respectively. Figure 2(c) depicts the standard reference structure of Met-enkephalin [59], [34], while Figure 2(d) shows the minimum-energy structure from our work, compared against the lowest-energy structure obtained in the simulations. Our main objective is to explore the entire energy landscape and

also get the minimum energy structure of the larger size protein [18].

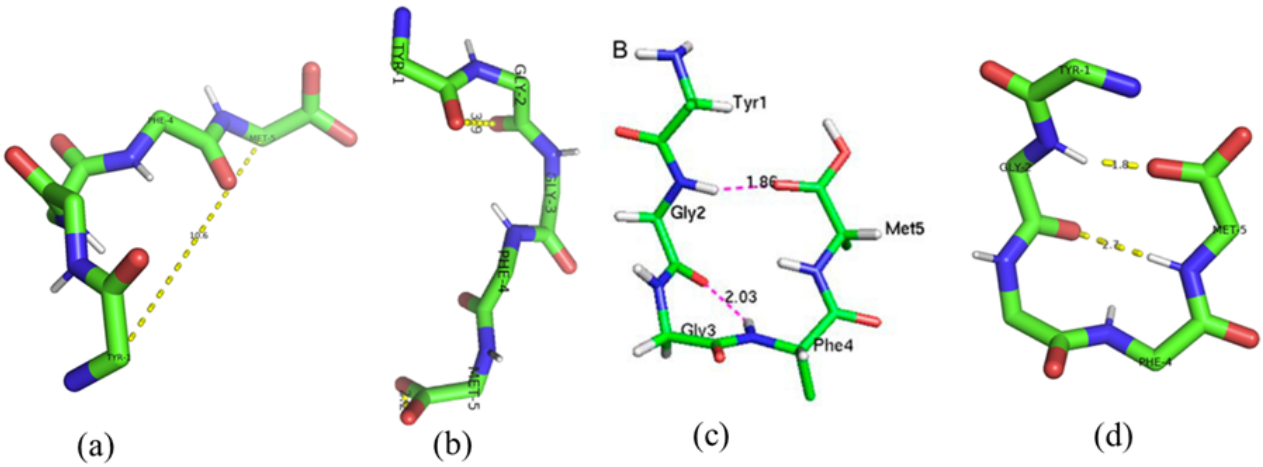


Figure 2. Structural configurations of Met-enkephalin: (a) and (b) Initial high energy Initial structures; (c) standard experimental minimum-energy structure; (d) minimum-energy structure obtained in the present work.

Figure 3 shows histograms of the frequency of sampled configurations per unit energy range for the Met-enkephalin protein at different temperatures (350 K, 300 K, 250 K, and 200 K) within the energy window −12.0 kcal/mol to 20.0 kcal/mol. The left panels correspond to MC simulations, while the right panels represent PTMC simulations.

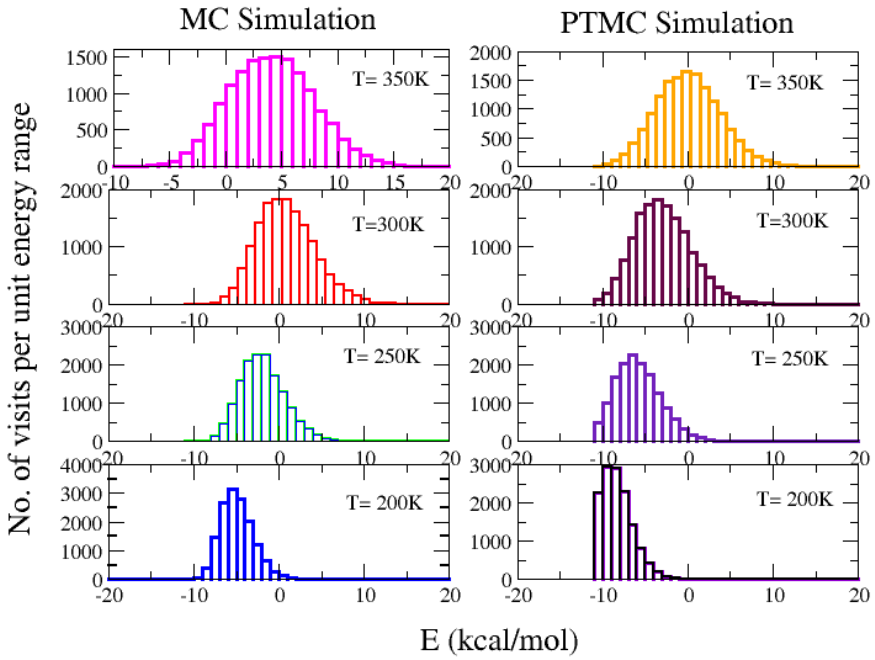


Figure 3. Comparison of MCS and PTMC methods for met-enkephalin peptide.

In MC simulations (left panels), the high-temperature distributions (350 K) are broad and symmetric, reflecting extensive sampling of high-energy states and good coverage of the energy landscape. At lower temperatures, the distributions become narrower and shift toward lower energies, showing convergence to stable, folded conformations with sampling concentrated near energy minima. In PTMC simulations (right panels), the system also explores a wide energy range at high temperature, similar to MC. However, at lower temperatures (250 K and 200 K), the distributions are more asymmetric and strongly shifted toward the low-energy region, with significantly better sampling of minimum-energy states. Unlike MC, which shows very few visits in the lowest-energy region, PTMC achieves sufficient sampling even at these states. This confirms that PTMC simulations reach the lowest-energy structures in fewer steps compared to standard MC simulations. For larger protein systems, where standard MC would require enormous computational effort to find the global minimum, PTMC provides a more efficient and reliable method for locating the global minimum structure.

Table 1 presents PT simulation data for the Met-enkephalin protein, including the lowest energy, end-to-end distance, and root mean square deviation (RMSD) values at different temperatures. The value of lowest energy decreases as temperature increases from 200 K to 350 K, reflecting reduced structural stability at higher temperatures. In the case of end-to-end distance, a steady increase with temperature indicates that the structure becomes more extended. A notable change occurs when comparing values below 300 K with those above 300 K, marking the folding–unfolding transition around this temperature. The RMSD values also increase with temperature, showing larger deviations from the reference structure and greater conformational flexibility. The most significant change is again observed across the 300 K boundary. Together, these structural quantities clearly suggest a major conformational change in Met-enkephalin near 300 K, consistent with a folding–unfolding transition.

Table 1. Results of PTMC simulations: lowest energy, end-to-end distance, and RMSD at different temperatures.

Temperature (K)	Lowest Energy (kcal/mol)	End-to-End Distance (mm)	RMSD (Å)
Standard Structure	-11.80	5.60	–
200	-11.23	5.81	1.18
250	-10.38	6.68	2.91
300	-8.31	9.15	5.36
350	-5.38	12.80	6.52

4.2. WW-Domain

To continue this work further, we have performed the PTMC simulation for WW-Domain protein [60]. In Figure 4, the total energy with MCS steps is shown using PTMC simulation. The simulation has run for 15 Lakh steps for temperature range of 500 K, 420 K, 350 K, and 280 K. Figure 4 shows the variation of energy versus MCS obtained from PTMC simulations of the WW-Domain protein at different temperatures. The four curves correspond to 500 K (blue), 420 K (pink), 350 K (green), and 280 K (yellow), arranged from top to bottom, respectively. At higher

temperatures of 500 K (blue) and 420 K (pink), the energy fluctuations are broad and cover a wider range [31] [29]. This indicates that the system explores higher-energy conformations, consistent with greater conformational flexibility and enhanced sampling of the energy landscape. At intermediate temperature (350 K), the fluctuations are moderate, showing that the system begins to stabilize while still sampling a variety of conformations. At lower temperature (280 K), the energy distribution is much narrower and shifted toward lower values, indicating convergence toward stable, low-energy structures associated with folded conformations. Overall, the plot demonstrates the effectiveness of PTMC simulations in sampling across a wide temperature range. The clear separation of energy levels across temperatures also supports the folding–unfolding transition behavior of WW-Domain.

The Figure 5 shows energy histograms for two consecutive temperatures obtained during the simulation. For each subplot, the x-axis represents energy (kcal/mol) and the y-axis represents the number of visits (frequency), i.e., how often configurations are sampled within a given energy range. Each plot overlays the energy distributions at two neighboring temperatures. A significant overlap between the histograms is observed, indicating efficient sampling and good exchange probability between consecutive temperatures. As the temperature increases, the distribution becomes slightly broader and shifts toward higher energies, reflecting enhanced exploration of higher-energy conformations. Specifically, we examined (i) the overlap of energy histograms between neighboring temperatures, (ii) the frequency of replica exchanges, and (iii) the ability of replicas to perform round trips between the lowest and highest temperatures. These indicators collectively demonstrate efficient mixing of configurations and effective exploration of the energy landscape, supporting the ergodic [13, 14] nature of the sampling.

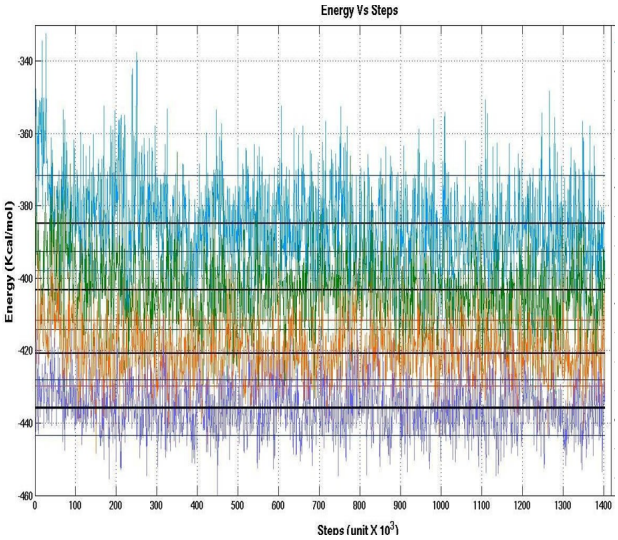


Figure 4. Energy versus MCS at different temperatures using PTMC simulation for WW-Domain protein. From top to bottom the temperatures are 500 K (Blue), 420 K (green), 350 K (orange), and 280 K (purple).

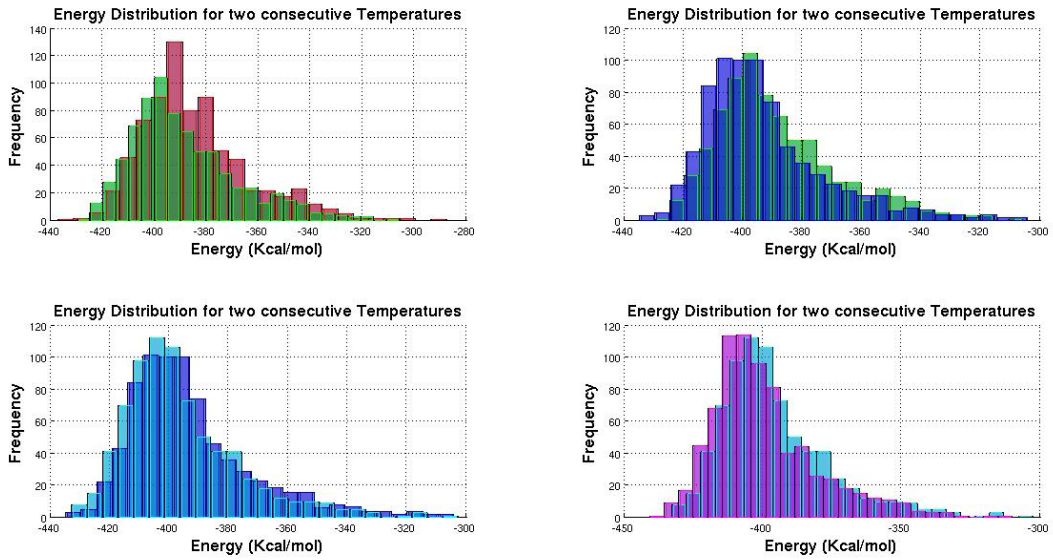
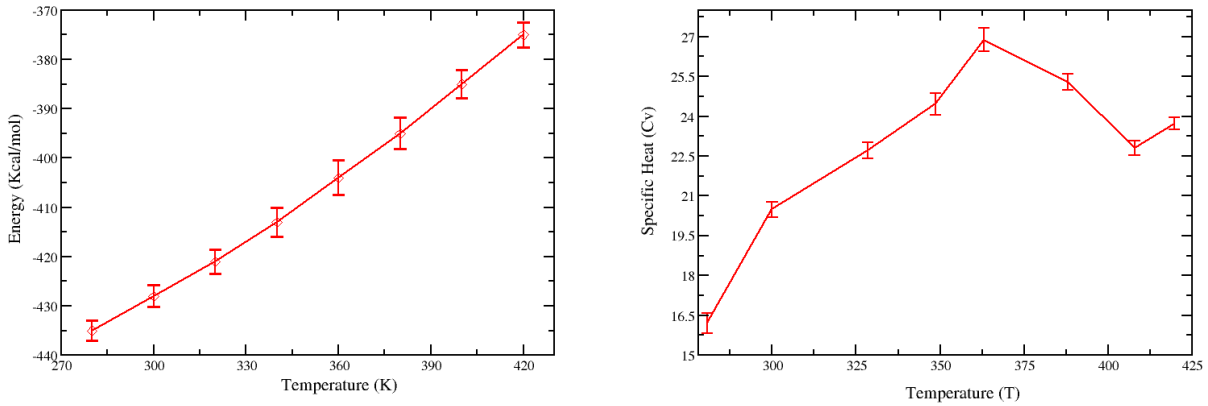


Figure 5. Frequency (number of visits) verses Energy (kcal/mol) for the two consecutive temperatures.



(a) **(b)**

Figure 6. Average energy E (left) and specific heat (right) of the WW-Domain protein as functions of temperature are obtained from parallel tempering simulations. The peak in identifies the folding–unfolding transition temperature. Error bars represent the standard error computed from eight independent trajectories, incorporating statistical inefficiency.

To illustrate the thermodynamic behavior [61, 62] and of the WW-Domain, Figure 6 represents the average energy and specific heat of the WW-Domain [63–65] protein as functions of temperature together characterize its folding–unfolding transition. The thermodynamic quantities shown were computed using configurations sampled at their respective replica temperatures during the parallel tempering simulation. Error bars represent statistical uncertainty estimated from statistically

independent trajectories, accounting for time correlations within each trajectory. At low temperatures, the average energy remains low, indicating that the protein predominantly occupies stable, folded conformations. With increasing temperature, the average energy increases smoothly, reflecting the gradual destabilization of native interactions and the onset of unfolding.

The specific heat in Figure 6(b) displays a clear peak at an intermediate temperature, reflecting enhanced energy fluctuations due to the coexistence of folded and unfolded states. This maximum defines the folding–unfolding temperature of the WW domain: the folded state dominates below it, while unfolded conformations prevail above it. The simulated peak occurs around 365–370 K, indicating a cooperative two-state transition. Experimental studies report melting temperatures of 330–345 K [63], although the simulated value is slightly higher, the overall single-peaked, and cooperative behavior is consistent with experimental observations.

We have also presented some WW-Domain protein structures. Figures 7(a) and 7(b) show the initial random structures, while Figure 7(c) represents the experimentally found Nuclear Magnetic Resonance (NMR) spectroscopy structure and Figure (d) illustrates the lowest energy structure of the WW-Domain. However, the lowest energy structure obtained from PT simulation does not match the standard NMR structure [31]. This may be due to the need for running a larger number of MC steps, as the size of the WW-Domain protein is nearly eight times greater than of the Met-enkephalin protein.

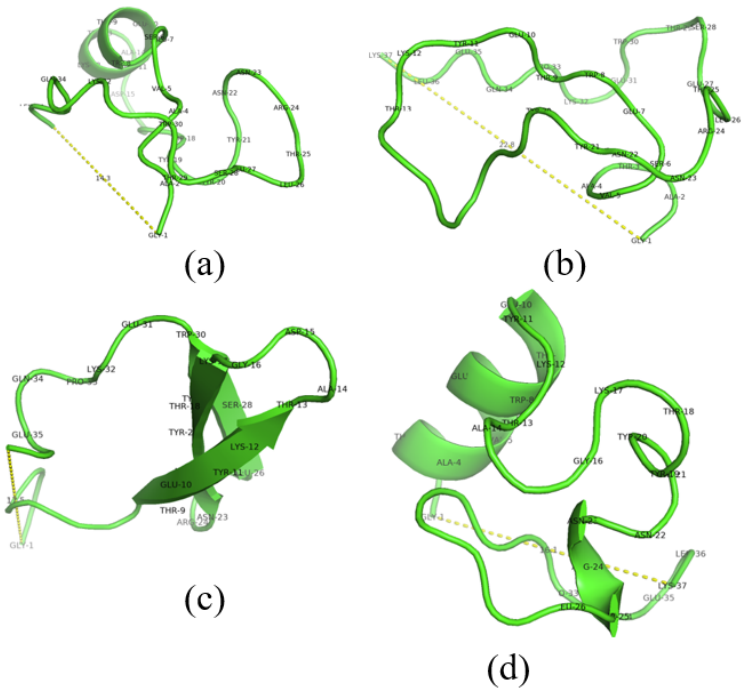


Figure 7. Initial structures (a) and (b) standard experimental structure (c), and minimum energy structures from this work (d) .

5. Conclusion

This study investigated the efficiency of conformational sampling in protein energy landscapes using Met-enkephalin and the WW-Domain as model systems. Standard MC simulations for Met-enkephalin reached equilibrium at higher temperatures but showed inadequate sampling in low-energy regions. In contrast, PTMC significantly enhanced sampling efficiency across temperatures, particularly in the low-energy regime, enabling robust exploration of the energy landscape and reliable identification of global minima. PTMC successfully identified the global minimum structure of Met-enkephalin with an energy of -11.23 kcal/mol and an RMSD of 1.18 Å from the reference structure.

Application of PTMC to the larger WW-Domain revealed active folding–unfolding transitions at higher temperatures, while sampling at lower temperatures remained limited due to increased landscape complexity. The temperature dependence of the average energy and specific heat provided clear signatures of structural transitions in both systems. The WW-Domain protein exhibits a pronounced peak in the specific heat, indicative of a cooperative folding–unfolding transition typical of small, single-domain proteins. The folding (melting) temperature was identified from the maximum of the specific heat, consistent with previous computational and experimental studies. These findings confirm the superiority of PTMC over standard MC for small peptides but indicate that further optimization of temperature schemes, replica numbers, and sampling strategies is necessary for accurate folding of larger protein systems. Future work will focus on improving PTMC protocols to better characterize larger protein domains. All reported thermodynamic quantities were obtained from canonical ensemble averages at the simulated replica temperatures with explicit statistical error estimation.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgment

I sincerely thank Dr. Rajamani Raghunathan from UGC-DAE-CSR Indore, for providing the computational facilities essential for this work.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

PS conceived and designed the study and established the methodology. MR performed the simulations at various stages of the work. MR, NCV, and SJ contributed to the data analysis. PS and

MR prepared the manuscript and reviewed the manuscript.

References

1. Skolnick J, Kolinski A (1989) Computer simulations of globular protein folding and tertiary structure. *Annu Rev Phys Chem* 40: 207–235. https://doi.org/10.1146/annurev.pc.40.100189.001231
2. Markwick PRL, McCammon JA (2011) Studying functional dynamics in biomolecules using accelerated molecular dynamics. *Phys Chem Chem Phys* 213: 20053–20065. https://doi.org/10.1039/C1CP22100K
3. Corey RB, Pauling LC (1953) Fundamental dimensions of polypeptide chains. *Proc R Soc Lond B Biol Sci* 141: 10–20. https://doi.org/10.1098/rspb.1953.0011
4. Deng C, Wu J, Cheng R, et al. (2014) Functional polypeptide and hybrid materials: precision synthesis via alpha-amino acid N-carboxyanhydride polymerization and emerging biomedical applications. *Prog Polym Sci* 39: 330–364. https://doi.org/10.1016/j.progpolymsci.2013.10.008
5. Alberts B, Johnson A, Lewis J, et al. (2002) *Molecular Biology of the Cell,* 4 Eds., New York: Garland Science.
6. Faure G, Ogurtsov AY, Shabalina SA, et al. (2016) Role of mRNA structure in the control of protein folding. *Nucleic Acids Res* 44: 10898–10911. https://doi.org/10.1093/nar/gkw671
7. Lipovsek D, Pluckthun A (2004) In vitro protein evolution by ribosome display and mRNA display. *J Immunol Methods* 290: 51–67. https://doi.org/10.1016/j.jim.2004.04.008
8. Szilagyi A, Kardos J, Osvath S, et al. (2007) *Handbook of Neurochemistry and Molecular Neurobiology,* New York: Springer.
9. Onuchic JN, Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* 14: 70–75. https://doi.org/10.1016/j.sbi.2004.01.009
10. Daggett V (2006) Protein folding–simulation. *Chem Rev* 106: 1898–1916. https://doi.org/10.1021/cr0404242
11. Lumry R, Eyring H (1954) Conformation changes of proteins. *J Phys Chem* 58: 110–120. https://doi.org/10.1021/j150512a005
12. Straub JE, Thirumalai D (1993) Exploring the energy landscape in proteins. *Proc Natl Acad Sci USA* 90: 809–813. https://doi.org/10.1073/pnas.90.3.809
13. Whitfield TW, Bu L, Straub JE (2002) Generalized parallel sampling. *Phys A: Stat Mech Appl* 305: 157-171. https://doi.org/10.1016/S0378-4371(01)00656-2
14. Kirkpatrick TR, Thirumalai D (1987) Dynamics of the structural glass transition and the p-spin-interaction spin-glass model. *Phys Rev Lett* 58: 2091. https://doi.org/10.1103/PhysRevLett.58.2091
15. Onuchic JN, Luthey-Schulten Z, Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem* 48: 545–600. https://doi.org/10.1146/annurev.physchem.48.1.545
16. Bernetti M, Bertazzo M, Masetti M (2020) Data-driven molecular dynamics: a multifaceted challenge. *Pharmaceuticals (Basel)* 13: 253. https://doi.org/10.3390/ph13090253

17. Perilla JR, Goh BC, Cassidy CK, et al. (2015) Molecular dynamics simulations of large macromolecular complexes. *Curr Opin Struct Biol* 31: 64–74. https://doi.org/10.1016/j.sbi.2015.03.007

18. Wales DJ (2018) Exploring energy landscapes. *Annu Rev Phys Chem* 69: 401–425. https://doi.org/10.1146/annurev-physchem-050317-021219

19. Singh P, Sarkar SK, Bandyopadhyay P (2014) Wang-Landau density of states based study of the folding-unfolding transition in the mini-protein Trp-cage (TC5b). *J Chem Phys* 141: 015103. https://doi.org/10.1063/1.4885726

20. Frantz DD, Freeman DL, Doll JD (1990) Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: applications to atomic clusters. *J Chem Phys* 93: 2769–2784. https://doi.org/10.1063/1.458863

21. Marinari E, Parisi G (1992) Simulated tempering: a new Monte Carlo scheme. *Europhys Lett* 19: 451–458. https://doi.org/10.1209/0295-5075/19/6/002

22. Earl DJ, Deem MW (2005) Parallel tempering: theory, applications, and new perspectives. *Phys Chem Chem Phys* 7: 3910–3916. https://doi.org/10.1039/b509983h

23. Okamoto Y (2004) Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J Mol Graph Model* 22: 425–439. https://doi.org/10.1016/j.jmgm.2003.12.009

24. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314: 141–151. https://doi.org/10.1016/S0009-2614(99)01123-9

25. Cheng X, Cui G, Hornak V, et al. (2005) Modified replica exchange simulation methods for local structure refinement. *J Phys Chem B* 109: 8220–8230. https://doi.org/10.1021/jp045437y

26. Srinivasaraghavan K, Zacharias M (2009) Folding of Trp-cage mini protein using temperature and biasing potential replica-exchange molecular dynamics simulations. *Int J Mol Sci* 10: 1121–1137. https://doi.org/10.3390/ijms10031121

27. Lyman E, Ytreberg FM, Zuckerman DM (2006) Resolution exchange simulation. *Phys Rev Lett* 96: 028105. https://doi.org/10.1103/PhysRevLett.96.028105

28. Liu P, Kim B, Friesner RA, et al. (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Phys Rev Lett* 102: 018101. *Proc Natl Acad Sci USA* 102 13749–13754. https://doi.org/10.1073/pnas.0506346102

29. Singh P, Sarkar SK, Bandyopadhyay P (2011) Understanding the applicability and limitations of Wang–Landau method for biomolecules: Met-enkephalin and Trp-cage. *Chem Phys Lett* 514: 357–361. https://doi.org/10.1016/j.cplett.2011.08.053

30. Takahashi A (2021) *Handbook of Hormones,* 2 Eds., Amsterdam: Elsevier. https://doi.org/10.1016/C2019-1-01019-4

31. Macias MJ, Gervais V, Civera C, et al. (2000) Structural analysis of WW domains and design of a WW prototype. *Nat Struct Biol* 7: 375–379. https://doi.org/10.1038/75144

32. Crane JC, Koepf EK, Kelly JW, et al. (2000) Mapping the transition state of the WW domain beta-sheet. *J Mol Biol* 298: 283–292. https://doi.org/10.1006/jmbi.2000.3665

33. Hansmann UH (1997) Parallel tempering algorithm for conformational studies of biological molecules. *Chem Phys Lett* 281: 140–150. https://doi.org/10.1016/S0009-2614(97)01198-6

34. Zhan L, Chen JZY, Liu WK (2006) Conformational study of met-enkephalin based on the ECEPP force fields. *Biophys J* 91: 2399–2404. https://doi.org/10.1529/biophysj.106.083899

35. Sindhikara D, Spronk SA, Day T, et al. (2017) Improving accuracy, diversity, and speed with prime macrocycle conformational sampling. *J Chem Inf Model* 57: 1881–1894. https://doi.org/10.1021/acs.jcim.7b00052

36. Nguyen H, Jager M, Moretto A, et al. (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc Natl Acad Sci USA* 100: 3948–3953. https://doi.org/10.1073/pnas.0538054100

37. Maisuradze GG, Liwo A, Scheraga HA (2010) Relation between free energy landscapes of proteins and dynamics. *J Chem Theory Comput* 6: 583–595. https://doi.org/10.1021/ct9005745

38. Beccara SA, Skrbic T, Covino R, et al. (2012) Dominant folding pathways of a WW domain. *Proc Natl Acad Sci USA* 109: 2330–2335. https://doi.org/10.1073/pnas.1111796109

39. Weinzierl S (2000) Introduction to Monte Carlo methods. https://doi.org/10.48550/arXiv.hep-ph/0006269

40. Santana R, Larranaga P, Lozano JA (2008) Protein folding in simplified models with estimation of distribution algorithms. *IEEE Trans Evol Comput* 12: 418–438. https://doi.org/10.1109/TEVC.2007.906095

41. Suruzhon M, Bodnarchuk MS, Ciancetta A, et al. (2022) Enhancing ligand and protein sampling using sequential Monte Carlo. *J Chem Theory Comput* 18: 3894–3910. https://doi.org/10.1021/acs.jctc.1c01198

42. Berg BA, Neuhaus T (1991) Multicanonical algorithms for first order phase transitions. *Phys Lett B* 267: 249–253. https://doi.org/10.1016/0370-2693(91)91256-U

43. Bhanot G (1988) The Metropolis algorithm. *Rep Prog Phys* 51: 429. https://doi.org/10.1088/0034-4885/51/3/003

44. Rahman A, Saikia B, Gogoi CR, et al. (2022) Advances in the understanding of protein misfolding and aggregation through molecular dynamics simulation. *Prog Biophys Mol Biol* 175: 31–48. https://doi.org/10.1016/j.pbiomolbio.2022.08.007

45. Noe F (2008) Probability distributions of molecular observables computed from Markov models. *J Chem Phys* 128: 244103. https://doi.org/10.1063/1.2916718

46. Koneru JK, Reid KM, Robustelli P (2025) Performing all-atom molecular dynamics simulations of intrinsically disordered proteins with replica exchange solute tempering. https://doi.org/10.48550/arXiv.2505.01860

47. Predescu C, Predescu M, Ciobanu CV (2005) On the efficiency of exchange in parallel tempering Monte Carlo simulations. *J Phys Chem B* 109: 4189–4196. https://doi.org/10.1021/jp045073+

48. Ozboyaci M, Kokh DB, Corni S, et al. (2016) Modeling and simulation of protein–surface interactions: achievements and challenges. *Q Rev Biophys* 49: e4. https://doi.org/10.1017/s0033583515000256

49. Wei G, Xi W, Nussinov R, et al. (2016) Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chem Rev* 116: 6516–6551. https://doi.org/10.1021/acs.chemrev.5b00562
50. Kumar S, Rosenberg JM, Bouzida D, et al.(1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13: 1011–1021. https://doi.org/10.1002/jcc.540130812
51. Chodera JD, Swope WC, Pitera JW, et al. (2007) Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J Chem Theory Comput* 3: 26–41. https://doi.org/10.1021/ct0502864
52. Nemethy G, Gibson KD, Palmer KA, et al. (1992) Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides. *J Phys Chem* 96: 6472–6484. https://doi.org/10.1021/j100194a068
53. Bashford D, Case DA (2000) Generalized Born models of macromolecular solvation effects. *J Comput Aided Mol Des* 14: 129–152. https://doi.org/10.1146/annurev.physchem.51.1.129
54. Zhou R (2003) Free energy landscape of protein folding in water: explicit vs. implicit solvent. *Proteins* 53: 148–161. https://doi.org/10.1002/prot.10483
55. Garcia AE, Onuchic JN (2003) Folding a protein in a computer: an atomic description of the folding/unfolding of protein A. *Proc Natl Acad Sci USA* 100: 13898–13903. https://doi.org/10.1073/pnas.2335541100
56. Okur A, Wickstrom L, Layten M, et al. (2006) Improved efficiency of replica exchange simulations through use of a hybrid explicit/implicit solvation model. *J Chem Theory Comput* 2: 420–433. https://doi.org/10.1021/ct050196z
57. Walter JC, Barkema GT (2015) An introduction to Monte Carlo methods. *Physica A* 418: 78–87. https://doi.org/10.1016/j.physa.2014.06.014
58. Evans DA, Wales DJ (2003) The free energy landscape and dynamics of met-enkephalin. *J Chem Phys* 119: 9947–9955. https://doi.org/10.1063/1.1616515
59. Perez JJ, Villar HO, Loew GH (1992) Characterization of low-energy conformational domains for Met-enkephalin. *J Comput Aided Mol Des* 6: 175–190. https://doi.org/10.1007/BF00129427
60. Dias AMGC, Teixeira GDG, Barbosa AJM, et al. (2025) Design and evolution of a synthetic small protein scaffold based on the WW domain. *Protein Sci* 34: e70164. https://doi.org/10.1002/pro.70164
61. Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. *Phys Rev Lett* 63: 1195–1198. https://doi.org/10.1103/PhysRevLett.63.1195
62. Gallicchio E, Andrec M, Felts AK, et al. (2005) Temperature weighted histogram analysis method, replica exchange, and transition paths. *J Phys Chem B* 109: 6722–6731. https://doi.org/10.1021/jp045294f
63. Cobos ES, Iglesias M, Ruiz-Sanz J, et al. (2009) Thermodynamic characterization of the folding equilibrium of the human nedd4-WW4 domain: at the frontiers of cooperative folding. *Biochemistry* 48: 1404–1412. https://doi.org/10.1021/bi9007758

64. Beccara S, Skrbic T, Covino R, et al. (2012) Dominant folding pathways of a WW domain. *Proc Natl Acad Sci USA* 109: 2330–2335. https://doi.org/10.1073/pnas.1111796109

65. Chai YL, Chin KH, Hansmann UHE, et al. (2003) Parallel tempering simulations of HP-36. *Proteins* 52: 436–445. https://doi.org/10.1002/prot.10351

AIMS Press