Proceedings of ICBioMed 2026 Symposium: Biomedical Engineering and Smart Medicine
DOI: 10.54254/2753-8818/2026.HZ37332

Hydrophobic Patterns and Folding Thermodynamics-Kinetics in a 2D HP Model

Yiyuan Xie

Department of Artificial Intelligence, Macau University of Science and Technology, Macau, China
xyy060226@outlook.com

Abstract. Protein folding is a process in which the linear amino acid chain is transformed into a stable three-dimensional native structure which enables the protein to carry out its function. In several neurodegenerative diseases, including Alzheimer's disease (amyloid-β) and Huntington's disease (huntingtin), a pathological mechanism is the misfolding and aggregation of specific proteins. This study quantifies the independent effect of hydrophobic proportion and spatial distribution on conformational stability and conformational search efficiency using a fully controlled 20-residue two-dimensional (2-D) HP lattice model, and a standardized simulated annealing (SA) workflow. Results showed that the total hydrophobic content was the most important factor for thermodynamic stability and the relative performance of the two distribution patterns showed a significant crossover effect. However, only at low levels of hydrophobicity (less than 40%) do contiguous hydrophobic blocks outperform dispersed ones, with evenly dispersed residues forming more symmetric globular structures, being more energy stable and compact. This pattern-dependent reversal can account for the conflicting conclusions of the previous studies that are confined to a narrow hydrophobic range. All conclusions are only applicable to the present short-chain 2D lattice system and fixed annealing parameters, and cannot be directly compared with the real physical folding time for natural proteins or transferred to longer sequence systems.

Keywords: protein folding, HP model, hydrophobic pattern, simulated annealing, lattice model

1. Introduction

Protein folding is the evolution of a stable three-dimensional native structure of an amino acid linear chain that is needed to carry out its biological function. The correct folding is essential for the normal activity of the cell and protein misfolding and aggregation of certain proteins is the pathological basis for several neurodegenerative disorders, including, but not limited to, amyloid-β in Alzheimer's and huntingtin in Huntington's [1]. The experimental characterization of folding pathway at atomic resolution is technically challenging and time consuming, and simplified lattice models of folding have become a mainstream way to dissect general folding principles. Of these, the hydrophobic-polar (HP) model is the one that is able to account for the driving force of hydrophobic collapse, but the synergy of the arrangement of hydrophobic residues to determine thermodynamic stability and folding rates is still poorly understood.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.HZ37332

Structural behavior of the hydrophobic core that collapses can be reproduced in the hydrophobic-polar (HP) lattice model, which was first formalized by Lau and Dill [2], at minimal computational cost by reducing amino acids to binary residue types. The initial investigations in this area centered on the hydrophobicity of all the residues, as the main determinant of folding stability. On 2- and 3 dimensional lattices, using sequences of 10 to 36 residues, it was consistently found that an increase in the hydrophobic proportion leads to a decrease in minimum energy, and an increase in the compaction of the chain, and a monotonic relationship between overall hydrophobic content and thermodynamic stability was generally accepted [2, 3].

Later studies started to examine how the hydrophobic spatial orientation affects it, but because of the lack of control of the variables, the reported results are inconsistent. In particular, Shakhnovich demonstrated that hydrophobic contiguous sequences seem to fold into distinct [4], native-like structures in 3D lattice models, but he used only a fixed set of 27-residue sequences to demonstrate this and did not systematically explore the total hydrophobic proportion. Bornberg-Bauer reported that dispersed hydrophobic motifs promote early-stage chain collapse in HP lattice chains [5], but only at low hydrophobic ratios (around 30%) and no quantification of the final structural compactness was given. There are two important unaddressed factors which could account for these conflicting conclusions. First, most of the studies mentioned changes the sequence length and the case of the hydrophobic proportion at the same time, so it is not possible to separate the effect of the distribution pattern only. Second, the thermodynamic stability and the folding kinetics are hardly ever studied in the same simulation framework, meaning that pattern effects across studies are still not reliable.

In this paper, this work study the influence of hydrophobic proportion and hydrophobic distribution pattern on the folding thermodynamics and folding kinetics of a 2D HP model. It is a controlled experimental design, where the length of the sequences is the same for all and the hydrophobic proportion and the arrangement of the residues is systematically varied. Two questions about the sequences are examined first in this study, namely, do hydrophobicity increases always lead to increased thermodynamic stability, and is compactness and folding efficiency different between clustered and dispersed patterns. It also examines the influence of these sequence characteristics on the convergence properties in simulated annealing. The rest of the paper is structured as follows. A review of the relevant literature is included in Section 2. The model, sequence design and simulation method are described in Section 3. The result and discussion have been given in section 4. The paper is concluded in Section 5.

2. Literature review

2.1. Controversy over dominant drivers of thermodynamic stability

Early HP lattice studies converged on the notion that it is the overall hydrophobic proportion that is the most important factor for folding stability. Lau and Dill systematically tested 2D lattice sequences of various hydrophobic fractions ranging from 10-25 residues and reported that there exists a good linear correlation between the average minimum energy and hydrophobic fraction [2], with each increment of 10% corresponding to an approximately 1.2 energy unit drop in average minimum energy. Later 3-D lattice investigations, however, called into question the primacy of total composition. Shakhnovich investigated the 27-residue cubic lattice sequences with the same 44% hydrophobic content, and concluded that true stability of 44% hydrophobic sequences can be affected significantly by the spatial arrangement of their hydrophobic residues [4], as the sequences with clustered hydrophobic residues are 2.3 times more likely to adopt a unique native conformation

DOI: 10.54254/2753-8818/2026.HZ37332

than those with the same total hydrophobic content but dispersed hydrophobic residues. But it has been found that the results in these studies are inconsistent because the researchers did not use standard control of variables: for example, early 2D studies controlled for the proportion of variables but ignored the distribution pattern, and 3D studies controlled for the distribution pattern but did not test the proportion, without quantifying the relative contribution of the two.

2.2. Divergent conclusions on kinetic effects of hydrophobic patterning

The role of hydrophobic spatial arrangement on folding has been less well accepted, and conflicting results have been reported in various laboratories. Bornberg-Bauer analyzed 2D HP sequences of 20 residues [5], with ~30% hydrophobic content, and found that the presence of dispersed motifs decreased the number of collapse steps in early stages by 18% from clustered motifs, suggesting that several independent nucleation sites formed. Analyses using the framework of energy landscape theory, however, have led to the hypothesis that contiguous hydrophobic core regions lower the barriers to folding and result in more rapid convergence for sequences having a hydrophobic content of more than 50% [6, 7]. This paradox is, in part, due to a limited hydrophobic range in individual studies, with assertions of dispersed acceleration generally coming from a low hydrophobic set of sequences, and assertions of clustered kinetics coming from a high hydrophobic set of sequences. No study so far has systematically tested both distribution patterns along the length of the whole gradient of hydrophobic proportions.

2.3. Methodological variability confounding cross-study comparison

The reported pattern effects are further complicated by differences in the conformation search algorithms. All simulation studies of the HP lattice models depend on the use of heuristic algorithms, all of which have different capabilities to explore. NP-completeness of the minimum-energy search problem for HP lattice models was proven in [8]. Using genetic algorithms, Unger and Moult investigated the convergence of 2D HP sequences and found that [9], in contrast to the convergence of standard Monte Carlo sequences, clustered patterns need 30% more iterations than random ones for the same level of convergence. Though a simple heuristic algorithm, simulated annealing aims at balancing exploration and exploitation through the controlled use of temperature scheduling [10], it is important to note that most of the previous work on hydrophobic patterning did not keep annealing parameters the same for all sequence groups, thus adding more confounding variables. The authors use a single simulated annealing scheme for all sequences with the same parameters, minimizing algorithmic differences as a confounding factor when comparing distribution effects.

3. Methodology

3.1. Sequence design

To overcome three important methodological limitations of previous works, this study used a fully controlled variable design that allowed to separate the individual contributions of the hydrophobic proportion and spatial distribution pattern.

To avoid the chain length as a confounding factor, first, the sequence length was fixed at 20 residues. Early HP lattice studies (e.g., [2]) used sequences as short as 10-15 residues, that have very limited conformational space and are not reliable to capture distribution-dependent variations in folding; early short-sequence lattice studies also rarely reported kinetic measures like success rate of

DOI: 10.54254/2753-8818/2026.HZ37332

folding. Shakhnovich on the other hand employed 27-residue sequences in 3D lattices [4], which makes searching much more complicated in an exponential fashion, and makes it impossible to run sufficient replicates to perform statistical analysis. A 20-residue length was chosen as a compromise benchmark, as it offers sufficient conformational diversity to separate out the pattern effects, while being sufficient for the simulated annealing analysis to be performed multiple times, as done in previous kinetic simulation studies [5, 9].

Second, the four levels of hydrophobic proportion (25%, 40%, 55% and 70%) were selected to span the entire range of folding near the threshold to saturated hydrophobic collapse. Results from previous studies have shown conflicting results on the hydrophobic patterning, mainly because they only examined a limited range of hydrophobic content: Bornberg-Bauer studied sequences with moderate hydrophobic content of ~30% [5], and found that patterning with dispersion had a greater effect on faster collapse, while Chan and Dill studied sequences with moderate hydrophobic content between 40% and 60% [3], and emphasized the effect of total hydrophobic content. Systematic comparisons between patterns for low and high hydrophobicity have not been carried out. The four gradients in this work range between 5 and 14 H residues per chain, crossing the border between partial and full folding success rate and allowing for the quantification of the changing effect of patterns with overall hydrophobicity of the design.

Third, two patterns of canonical distribution were created with explicit generation rules, where three sequences were generated within each condition, which helped to avoid single-sequence bias.

•Clustered pattern: Contiguous block of all the H residues. Three replicates were prepared where the H block was replaced in the N terminal, middle and C terminal parts of the chain respectively, but the H count remained the same each time, thus controlling for positional effects.

•Dispersed pattern: Residues with H are distributed with as much evenness as possible along the chain to minimise contiguous H segments. Three replicates were created by rotating the beginning phase of the alternating H/P configuration, keeping the spacing between the parallel sequences the same.

The total number of sequences used was 24 (4 proportions × 2 patterns × 3 replicates); 10 independent simulation runs were conducted for each sequence, providing a total of 30 simulation runs (10 runs per condition) for statistical comparisons. This design eliminates pattern effects and the confounding effects of length and composition that are found in previous studies, and makes it possible to analyze thermodynamic stability and folding kinetics together.

3.2. 2D HP lattice model

A two-dimensional square-lattice HP model was used to do the folding simulations. The minimal coarse-grained framework was used to avoid the influence of confounding factors such as side-chain chemical details, electrostatic interactions, and to isolate hydrophobic interaction as the sole folding driving force and to enable a fully controlled experimental design to uncouple the contribution of hydrophobic proportion and spatial distribution. The locations of each amino acid residue are given by the integer lattice points (xi, yi), with i representing the index of the residue in the primary sequence. The consecutive residues along the chain must be spatially adjacent with a Manhattan distance of unity (must share a lattice coordinate): Astrict self-avoiding constraint: No two distinct residues can share the same lattice coordinate. Self-collision is checked by comparing all pairs of coordinates of non-consecutive residues in any candidate conformation. When two residues i, j (i and j are not adjacent) have the same x and y value, the conformation is considered to be invalid and is rejected before energy is calculated.

DOI: 10.54254/2753-8818/2026.HZ37332

The energy function is based exclusively on hydrophobic-hydrophobic (H–H) contacts. A valid H–H contact is defined as a pair of hydrophobic residues that are nearest neighbors on the lattice (Manhattan distance = 1) but are not adjacent in the primary sequence ($|i-j|>1$). Consecutive H residues along the chain are not counted, as their adjacency is enforced by backbone connectivity and does not contribute to stabilizing the folded state. The total conformational energy is defined as: where N_{HH} is the total number of valid non-consecutive H–H contacts. Lower energy values correspond to thermodynamically more stable conformations. The energy calculation follows a standard pairwise traversal logic: all non-consecutive residue pairs are checked for lattice adjacency, and each valid H–H contact contributes −1 to the total conformational energy.

3.3. Simulated annealing algorithm

A simulated annealing algorithm was applied to search for the minimum-energy conformation of each sequence. This heuristic method is capable of escaping local minima via probabilistic uphill acceptance, making it suitable for the NP-hard HP folding optimization problem [8, 10].

3.3.1. Conformational update: pivot move

A complete conference paper should follow a logical structure to ensure clarity, coherence, and academic rigor. Candidate conformations are generated using a single-point pivot move, a standard local update strategy for lattice polymer simulations. The move is executed as follows:

• Randomly select a pivot bond index k (between residue k and residue " $k+1$ ", where " $1 \leq k < n$ ").

• Treat residue k as the fixed rotation center. Apply a rigid-body rotation (90°, 180°, or 270°, chosen randomly with equal probability) to all residues from " $k+1$ " to the end of the chain.

• Perform self-avoidance check on the updated conformation. If any residue coordinate overlaps, reject the move and retain the original conformation. If valid, proceed to energy evaluation.

The algorithm initiates each independent run from a randomly generated valid self-avoiding linear conformation. At each Monte Carlo step, the energy difference between the candidate conformation and the current conformation is calculated as:Allenergy-reducing moves(($\Delta E < 0$))areaccepted unconditionally. Energy-increasing moves are accepted with a probability following the Metropolis criterion:where T is the current system temperature. The temperature is decreased geometrically after each step according to Tt+1=α · Tt, where α is the cooling factor. The algorithm iterates through the full annealing schedule for each independent run, tracking the lowest energy encountered and the iteration step at which it is first reached. All runs follow an identical temperature decay protocol and terminate at the fixed maximum step count.

3.3.2. Parameter calibration and convergence validation

A systematic sensitivity testing and the convergence validation were made to set simulation parameters: initial temperature, cooling factor, 10, 000 steps/runs, 10 independent runs/sequences.

First, the initial temperature was selected as 10.0, and the rate of accepting uphill moves was 45%–58%, which is in the recommended range of 40%–60% for the simulated annealing optimization in all the sequence groups [10]. This is to have adequate conformational sampling at the early stage and to avoid being trapped at local minima.

Second, the final temperature is ~0.067 at 10, 000 steps per run and the uphill acceptance rate is below 1% at the end of the run, which is the confirmation of the system stabilization with the

DOI: 10.54254/2753-8818/2026.HZ37332

cooling factor α=0.995. Increase to 20, 000 steps led to minimal (less than 2%) average minimum energy gain; slower cooling rate α=0.999 resulted in 5 times more computation cost with minimal energy gain.

Thirdly, the accuracy of the algorithm was checked against exhaustive enumeration for 10-residue HP sequences, for which the algorithm found the known global minimum in 100% of the runs and 100% of the independent runs.

Finally, with 10 runs per sequence, the statistical reliability is more than sufficient: minimum energy standard deviation is < 0.3 energy units under all conditions and the CoV of convergence step is < 15%.

Detailed parameter settings are summarized in Table 1.

Table 1. Rules to format sections

Parameter	Value
Initial temperature	10.0
Cooling factor	0.995
Maximum steps per run	10000
Independent runs per sequence	10

3.4. Evaluation metrics

Thermodynamic and computational search dimensions were used to evaluate the folding performance.

• Average minimum energy: the mean of minimum energies for all the replicate runs under each condition.

• H–H contact count: computed directly from the energy "N" _"HH" "=-E" .

Compactness: is "1/" "R" _"g" (reciprocal of radius of gyration) – more condensed structures have larger values.

Computational search metrics: These are NOT real physical folding times, but rather a way of describing the difficulty of the algorithmic search.

• Folding success rate: a successful run is one that reaches a negative minimum energy, which will indicate that stable low-energy conformations are reached.

• Convergence step: the first Monte Carlo iteration when a run reaches its minimum energy and is thus a measure of the efficiency of the search algorithm.

A total of 240 simulation records (24 different sequences x 10 runs) were classified according to hydrophobic proportion and distribution pattern (30 runs per pattern), and statistically analyzed.

4. Results and discussion

4.1. Thermodynamic stability

The average lowest energy for both distribution patterns is seen to decrease systematically with increasing H proportion, from 25% to 70%, thus finding the thermodynamic stability of the folded conformation increased with increasing the H proportion, as shown in figure 1. At the 25% H level, clustered sequences reach a lower average energy than dispersed sequences. However, at 40%, 55% and 70% H levels, dispersed sequences achieve lower average lowest energies than clustered

DOI: 10.54254/2753-8818/2026.HZ37332

sequences. The energy gap between the two patterns is largest at 40% and 55% H, and narrows slightly at 70% H. Clustered sequences at 25% H content show the highest (least favorable) energy overall, while dispersed sequences at 70% H content reach the lowest energy.

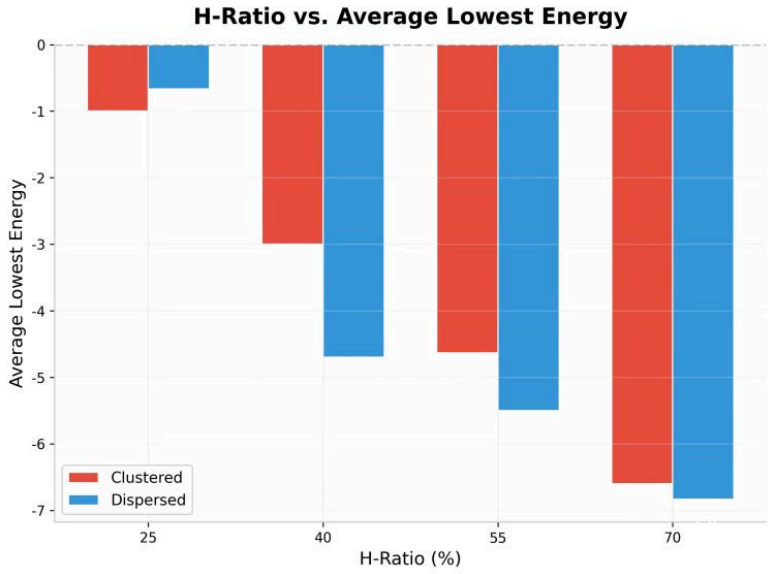

Figure 1. H-ratio vs. average lowest energy

The compactness comparison is presented in Figure 2. Overall, compactness shows an upward trend with increasing H proportion for both patterns. At the 25% H level, clustered sequences have slightly higher compactness (0.364) than dispersed sequences (0.344). However, at 40%, 55% and 70% H levels, dispersed sequences exhibit higher compactness values than clustered sequences. The largest difference appears at 55% H, where compactness is 0.397 for the clustered pattern and 0.474 for the dispersed pattern.

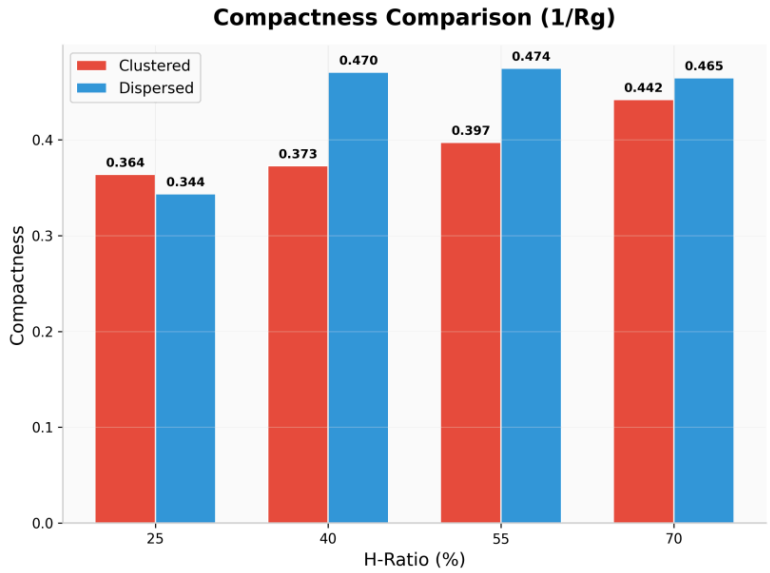

Figure 2. Compactness comparison (1/Rg)

DOI: 10.54254/2753-8818/2026.HZ37332

4.2. Computational search performance

Note that all metrics in this section describe the computational search performance of the simulated annealing algorithm, not the real-time folding kinetics of natural proteins. They illustrate the relative ease/difficulty of different sequences to adopt stable low energy conformations within a single optimization process.

The folding success rates of different groups are shown in Figure 3. The higher the proportion of H, the more likely it will be a success. There is a significant difference between the clustered and the dispersed pattern at the 25% H level, where the success rate is 100% and 33%, respectively. There is no difference between the two distribution patterns when the H proportion is 40% or higher, folding is 100% successful.

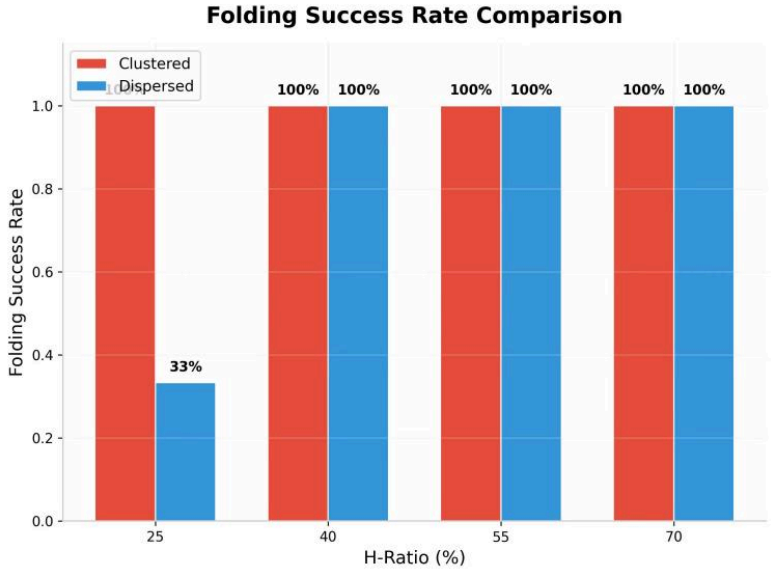

Figure 3. Folding success rate comparison

The energy convergence curves for four representative sequences, two of each distribution pattern and three hydrophobic proportion levels are shown in figure 4: for H25 Clustered, a successful run of H25 Dispersed, H40 Dispersed and H70 Dispersed. At the early iteration stage, all sequences have a relatively steep energy decay and eventually converge to a flat energy level. The number of steps needed for the sequence to converge (that is, to reach the lowest energy) varies significantly from one sequence to another. The fastest convergence is to H25 Clustered with a final energy of −1 around 80 Monte Carlo steps. The H25 Dispersed trajectory is successful in that it has more energy fluctuations and settles at −2 at about step 520. The minimum energy of H40 Dispersed is −5, which is approached at higher hydrophobic proportions, at about step 760, and of H70 Dispersed is −7, at about step 700.

As a whole, the more H the sequence contains, the lower will be the final sequence energy; however, the energy of the final sequence will require more iterations to achieve a stable state because a higher number of H residues will increase the conformational search space. The dispersed pattern exhibits higher search volatility and later convergence point than the clustered pattern at the matched 25% hydrophobic content, reflecting the fact that the dispersed pattern has less probability of successful folding at low hydrophobic content. The sequences that contain more hydrophobic residues tend to need more iterations to reach their minimum energy, which suggests that, for the

DOI: 10.54254/2753-8818/2026.HZ37332

same annealing schedule, sequences with high hydrophobic content may be more difficult and/or slower to search than those with fewer hydrophobic residues.

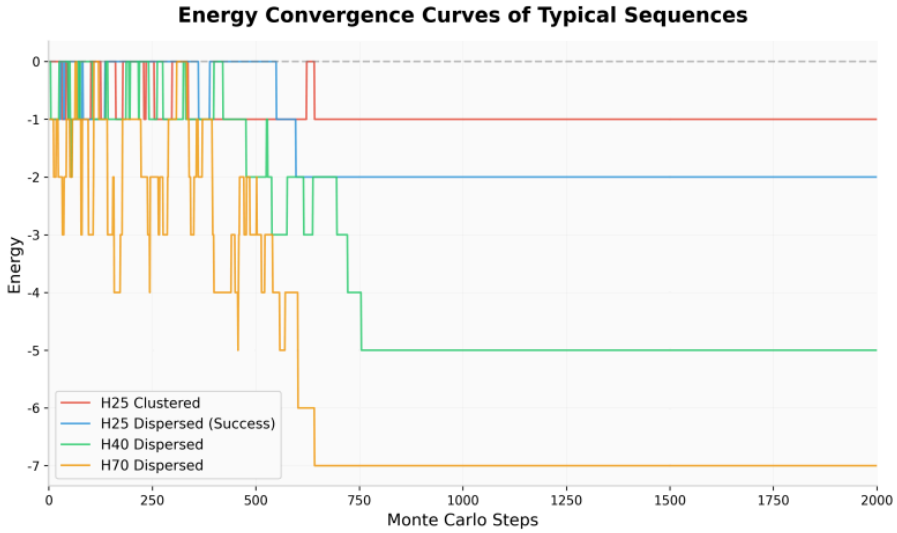

Figure 4. Energy convergence curves of typical sequences

4.3. Mechanistic inference based on measurable metrics

The underlying mechanism leading to distribution pattern effects is not observed directly but based on three validated measurable metrics: folding success rate, structural compactness and characteristics of convergence trajectory.

4.3.1. Native state structural divergence

With a hydrophobicity of 25%, the clustered pattern gives 100% success of folding, the dispersed pattern gives only 33%. This gap suggests that the contiguous hydrophobic blocks are more easily able to form a stable hydrophobic core and thus are able to drive chain collapse more easily when H is in short supply. In dispersed sequences, scattered hydrophobic residues can only form isolated contact pairs, and hence do not form a continuous stabilizing core, with a partially extended overall conformation.

Although the total H–H contact potential is comparable for the two patterns, they show different symmetry of the structures at the 55% hydrophobic level. The clustered arrangement has a tendency to create an asymmetric conformation with the hydrophobic residues grouped together and the polar residues pointing out, which results in a higher radius of gyration. The tertiary structure is an extended one which converts to a more symmetrical globular one in which hydrophobic residues are buried and polar residues are exposed at the surface. This uniformity of packing gives rise to a higher compactness of 19.4% for the dispersed pattern, which is to be expected given the fact that it is at a lower proportion that shows the lowest minimum energy.

4.3.2. Nucleation pathway differences in conformational search

Based on the features of convergence trajectories, two distribution patterns have different nucleation paths in the conformational search process. The clustered sequences exhibit a large early stage energy drop, and small fluctuations with a single-core nucleation mechanism, i.e., the pre-connected

DOI: 10.54254/2753-8818/2026.HZ37332

hydrophobic block serves as an inherent nucleation seed, and can rapidly stabilize and encourage the subsequent chain folding. Large energy fluctuations are observed in the dispersed sequences during the early and middle stages; the transition period is longer before a final stable state is reached. This is consistent with the multi-site nucleation mechanism: the hydrophobic residues are dispersed and form several tiny contact clusters independently by conformational sampling that finally coalesce into a stable full-sized core. This requires more search iterations and more uncertainty, and is reflected in the low hydrophobic level of the dispersal sequences in the low success rate and later convergence point.

In summary, hydrophobic proportion sets the upper limit of thermodynamic stability achievable: the number of possible H–H contacts; distribution pattern modulates the efficiency of the conformational search and the final structural compactness: nucleation pathways and conformational symmetry.

5. Conclusion

In the context of 20-residue 2D HP lattice models with fixed parameters of simulated annealing, the results quantify the separate and joint contributions of hydrophobic proportion and spatial distribution on conformational stability and search efficiency. The maximum attainable thermodynamic stability is defined by the total hydrophobic content by setting a lower limit for the minimum possible energy. By spatially arranging the modulation of nucleation pathways and final structural symmetry, non-monotonic changes in the compactness and the difficulty of the computational searches are observed.

A crossover effect is observed over the proportion gradient tested which is dependent on the pattern. The energy results show that contiguous hydrophobic blocks are not as efficient as evenly dispersed ones except for the 25% hydrophobic concentration, when contiguous hydrophobic blocks not only reduce energy, but are also more compact. At hydrophobic content of 40% and higher, dispersed residues adopt more symmetric globular shapes, which allows them to be compact, low-energy. This result is useful in reconciling some of the previously conflicting conclusions regarding the effects of spatial patterning, which came from studies that only examined either a small range of hydrophobicities or modeled lattices of different size.

Three obvious limitations exist to these conclusions. First, they are only valid for 20-residue short chains; longer sequences might acquire some domain-level folding properties which decrease or flip the effects of patterns. Second, there are only two extreme distributions; and the optimum level of hydrophobic clustering between these extremes on intermediate distributions is not quantified. Third, all the measures of search efficiency depend on the fixed annealing schedule used in this work; they cannot be simply equated with actual physical folding time, nor can they be directly translated to other optimization algorithms.

These results are a controlled 2D HP 20-residue benchmark for comparing annealing schedules, move sets or hydrophobicity gradients on the same sequence set. Follow-up validation to longer chains, three dimensional lattices or real protein systems has to be targeted.

References

[1] Dill, K. A. (1985). Theory for the folding and stability of globular proteins. *Biochemistry, 24*(6), 1501–1509.
[2] Lau, K. F., & Dill, K. A. (1989). A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules, 22*(10), 3986–3997.
[3] Chan, H. S., & Dill, K. A. (1991). Origins of structure in globular proteins. *Proceedings of the National Academy of Sciences of the United States of America, 88*(3), 785–789.

DOI: 10.54254/2753-8818/2026.HZ37332

[4] Shakhnovich, E. I. (1994). Proteins with selected sequences fold into unique native conformation. *Physical Review Letters, 72*(3), 390–393.
[5] Bornberg-Bauer, E. (1997). Chain growth algorithms for HP-type lattice proteins. In *Proceedings of the 1st Annual International Conference on Computational Molecular Biology (RECOMB '97)*, 47–55.
[6] Onuchic, J. N., Luthey-Schulten, Z., & Wolynes, P. G. (1997). Theory of protein folding: The energy landscape perspective. *Annual Review of Physical Chemistry, 48*, 545–600.
[7] Socci, N. D., Onuchic, J. N., & Wolynes, P. G. (1996). Protein folding mechanisms and the multidimensional folding funnel. *Proteins: Structure, Function, and Genetics, 24*(4), 493–505.
[8] Berger, B., & Leighton, T. (1998). Protein folding in the hydrophobic-hydrophilic model is NP-complete. *Journal of Computational Biology, 5*(1), 27–40.
[9] Unger, R., & Moult, J. (1993). Genetic algorithms for protein folding simulations. *Journal of Molecular Biology, 231*(1), 75–81.
[10] Kirkpatrick, S., Gelatt, C. D., Jr., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671–680.