デジタルアニーラ前処理空間探索と拡張アンサンブル法による環状ペプチド分子の高速安定構造探索

谷田 義明^{*}, 佐藤 博之, 眞鍋 敏夫, 寺島 千絵子

富士通, 〒212-0014 川崎市川崎区大宮町 1-5

e-mail: tanida.yoshiaki@fujitsu.com

(Received: July 30, 2021; Accepted for publication: September 29, 2021; Online publication: December 18, 2021)

Cyclization generally stabilizes the bioactive conformation of the peptide and increases its affinity for the target. However, since cyclic peptides frequently adopt multiple conformations in solution, the structural information is not fully understood in experiments, and the relationship between structure and function is not well understood. We demonstrate the practical possibilities of using a combination of a special purpose computing engine (Digital Annealer) and REST2 (replica exchange with solute tempering) simulation in "ab initio" structure prediction of macrocyclic peptides.

キーワード: Macrocyclic peptide, Digital Annealer, Lattice protein model, REST2, Structure prediction, Cluster analysis

1 目的

タンパク質間相互作用をターゲットにした創薬において、それらの境界面の形状・大きさのために低分子阻害剤の開発は困難である。近年、環状ペプチド分子は、タンパク質の表面に高い親和性と特異性をもって結合することから注目されている。しかしながら、環化による構造安定化に関わらず、水への溶解性が低く、溶液中で複数の構造をとることが多いため、実験手法でその構造を明らかにすることが難しい。このように溶液中での構造情報の不足が、環状ペプチド分子の新規設計に必要な配列、構造およびその機能の関係に対する理解の不十分な原因となっている。水分子を陽に扱った分子動力学シミュレーションは、溶液中での構造解明に対して非常に重要な手段である。この小論では、アミノ酸配列情報を格子プロテインモデルに置換して組合せ最適化問題に帰着させ、レプリカ交換法の初期構造依存性を低減させる効率的な立体構造予測手法を提案する。

2 方法

対象として、NMR 実験から 20 種の推定構造が報告された PLP-12 分子 Cyclo-(Phe-Val-Gly-Gly-Thr-Ser-Phe-Asp) [1] を用いた。通常、環状ペプチド分子の構造空間に対

Figure 1. Schematic view of the energy landscape

するポテンシャルエネルギー面 (Figure 1) は非常に複雑で、時間スケール、エネルギー障壁を超えるレートを考えると、通常の分子動力学シミュレーションでは十分にその空間を探索することができない。例えば、初期構造を A あるいは B から開始した場合ではその径路の困難さが異なる。そこで、前処理として探索空間の絞込みを行うために、格子点に C α 原子を配置し、原子の有無をイジングスピンのアップダウンに対応させる格子プロテインモデルで環状ペプチド分子を表現した。ここで、グリッド間隔を 3.8 Å に設定し、アミノ酸間ポテンシャルには TE-13 [2] を用いた。我々は、さらに環化条件のペナルティ項などを加えて系のスコア関数を定義した。そ

Figure 2. Structure with the minimal score value (left) and the all-atom structure after relaxation process (right).

の結果、初期構造を決定する問題は、このスコア関数を最小化する原子配置を求める問題に帰着された。この問題を高速に解くために富士通専用ハードウェア Digital Annealer (DA) を用いて、得られた極小解を初期アミノ酸 C α 原子位置に選んだ。続いて、C α 原子に側鎖を付加して構造緩和させることで初期構造を作成し、系のポテンシャルエネルギーを記述するために、点電荷(ESP電荷)と力場をアサインした。ここで、力場には、Amber99を基にバックボーンを構成する二面角のパラメーターを修正したものを用いた [3]。水分子は、TIP3Pモデルを使って表現し、ユニットセルを中性に保つために、Na⁺イオンをカウンターイオンとして追加した。このモデルに対して 298 K の分子動力学計算によって得られた熱平衡構造を REST2 シミュレーション [4] の初期構造とした。298 K から 800 K 相当の 7 レプリカを用意して 250 ns (平衡化 50 ns) の REST2 シミュレーションで構造アンサンブルを生成した。この時、各レプリカの交換頻度は ~ 0.2 であった。最後に、非階層型クラスタリングとして Daura アルゴリズムを用いて構造群を求めた。なお、全てのシミュレーションは GROMACS-2019.6 とプラグイ

ン PLUMED-2.6.2 を用いて実行した。

3 結果と考察

配列情報から格子プロテインモデルに対して、得られたスコア関数極小解の構造を Figure 2 に示す。例えば、Phe-Val 重心間距離の TE-13 ポテンシャルは 6 Å 近傍にエネルギーミニマムが存在する。生成された全原子構造の重心間距離は 6.3 Å、C α 原子間距離は 3.7 Å であり、我々の格子モデルが実際の分子構造を表現できる最も粗いモデルとして適切であることが分かる。極小解の C α 原子位置に最も近い実験構造との RMSD は 1.9 Å であった。次に、REST2 による探索空間拡大の効果を通常の定温分子動力学計算(cMD)の場合と比較し、同時に NMR 実験で推定された構造の再現性について調べた。Figure 3 に、REST2 および cMD で得られた 20000 構造サンプルの Ser6 のバックボーン二面角 (ϕ, ψ) の頻度分布を示す。Figure 3 (a) 中の白丸は、NMR 実験から推定された 20 種の構造のものである。cMD では、NMR 実験の推定構造の存在する $\phi \sim 1$ rad の構造空間をほとんど探索できていないことが分かる。このことは、8 残基程度の環状ペプチド分子においても構造間のエネルギー障壁が高く、cMD では探索空間が不十分になってしまうことを示唆している。言い換えると、格子モデル極小解を初期構造にした場合でも拡張アンサンブル法と組み合わせる必要があることを意味している。実際、REST2 シミュレーションでは、20 種の実験構造を含む必要な構造空間を、十分な頻度でサンプリングできていることが分かる。

最後に、本手法による立体構造の予測について述べる。ここでは、Mainchain+C β の原子位置に関する RMSD に基づいたクラスタ分析を用いて、生成された構造アンサン

Figure 3. Comparison of the backbone dihedral angle distributions in Ser6. (a) REST2, (b) cMD; NMR results are also depicted as white circles in (a).

Figure 4. Representative structure of cluster analysis (green).
NMR structure (orange).

ブルの分類を行った。用いたカットオフ半径は1.7 Åで、その時のクラスターの総数は29であった。Figure 4に、最大頻度(46.1%)クラスターの代表構造を示す。この構造と20種の実験構造とのRMSDを計算したところ、4番目に登録された構造との間で0.47 Åと非常に良く一致した。なお、最も小さなRMSDは0.25 Åで、この構造は最大頻度クラスター内に存在した。これらのことから、溶液中の実験構造に十分近い構造を、出現頻度の高いクラスター群中の(代表)構造として予測可能であることが分かった。

4 まとめ

格子プロテインモデルを用いた前処理と拡張アンサンブル法を組み合わせることによって、溶液中の環状ペプチド分子の立体構造をアミノ酸配列構造情報から十分な精度で予測できることを示した。本手法を利用して新規創薬設計に有用な構造情報が得られることが期待される。

References

- [1] M. F. Fisher, J. Zhang, N. L. Taylor, M. J. Howard, O. Berkowitz, A. W. Debowski, et al., *Plant Direct*, **2**, e00042 (2018). DOI:10.1002/pld3.42
- [2] D. Tobi, R. Elber, *Proteins*, **41**, 40 (2000). DOI:10.1002/1097-0134(20001001)41:1<40::AID-PROT70>3.0.CO;2-U, PMID:10944392
- [3] H. Fujitani, A. Matsuura, S. Sakai, H. Sato, Y. Tanida, *J. Chem. Theory Comput.*, **5**, 1155 (2009). DOI:10.1021/ct8005437, PMID:26609625
- [4] G. Bussi, *Mol. Phys.*, **112**, 379 (2014). DOI:10.1080/00268976.2013.824126