

分子動力学シミュレーションによる医学・創薬に向けたタンパク質の構造解析

榮 慶丈,^a 西川直宏,^{a,b} 塚本修一朗,^a 鈴木孝禎,^c 岡本祐幸^{a,d,e,f}

Molecular Simulations of Protein Systems toward Drug Discovery

Yoshitake Sakae,^a Naohiro Nishikawa,^{a,b} Syuichiro Tsukamoto,^a Takayoshi Suzuki,^c and Yuko Okamoto^{a,d,e,f}^aDepartment of Physics, Graduate School of Science, Nagoya University; Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan:^bInstitute for Molecular Science; 38 Nishigo-Naka, Myodaiji, Okazaki, Aichi 444-8585, Japan: ^cGraduate School of Medical Science, Kyoto Prefectural University of Medicine; 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-0823, Japan:^dStructural Biology Research Center, Graduate School of Science, Nagoya University; Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan: ^eCenter for Computational Science, Graduate School of Engineering, Nagoya University;Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan: and ^fInformation Technology Center, Nagoya University; Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan.

(Received August 16, 2015)

Molecular simulations have been widely used in biomolecular systems such as proteins, DNA, *etc.* The search for stable conformations of proteins by molecular simulations is important to understand the function and stability of proteins. However, finding the stable state by conformational search is difficult, because the energy landscape of the system is characterized by many local minima separated by high energy barriers. In order to overcome this difficulty, various sampling and optimization methods for the conformation of proteins have been proposed. In this study, we propose a new conformational search method for proteins based on a genetic algorithm. We applied this method to an α -helical protein. We found that the conformations obtained from our simulations are in good agreement with the experimental results.

Key words—genetic algorithm; molecular simulation; protein folding

はじめに

タンパク質や DNA などの生体高分子の機能を理解する上で分子シミュレーションの手法は 1 つの有効な手段となってきた。その適用範囲は広く、電子状態を考慮した分子モデルによる酵素反応や、原子集団を粗視化した分子モデルによる分子モーターなどの巨大分子系を取り合うシミュレーションなど広範囲に及ぶ。ここでは特に古典力学を基礎とする全原子分子モデルによるタンパク質のシミュレーション

に注目する。ただしタンパク質などの巨大な分子を全原子分子モデルで取り扱う場合、その安定構造探索を効率的に行うのは一般に難しい。その大きな問題の 1 つは系が複雑であるため、系の持つ局所的な安定状態が膨大に存在することである。このため通常行うような定温の分子シミュレーションでは限られた計算資源内に十分な安定構造をサンプルできず、特にタンパク質の折り畳みシミュレーションなど、広範囲の自由エネルギー曲面を解析する場合には構造解析が困難である。この困難を克服するためにこれまで多くの研究者によって徐冷法や拡張アンサンブル法など様々なシミュレーション手法が開発されてきた。¹⁻⁷⁾ 加えてわれわれもまたいくつかのシミュレーション手法を開発してきた。⁸⁻¹²⁾ ここでは最近われわれが提案した遺伝的アルゴリズムを用いた手法を紹介する。

さらに後半では効率的な分子シミュレーション手法の適用例として簡単に 2 つの例についても紹介する。1 つは、アルツハイマー病の要因の 1 つとして

^a名古屋大学理学研究科 (〒464-8602 名古屋市千種区不老町), ^b分子科学研究所 (〒444-8585 愛知県岡崎市明大寺町字西郷中 38), ^c京都府立医科大学 (〒606-0823 京都市左京区下鴨半木町 1-5 京都府立医科大学下鴨学舎 3F 医薬品化学), ^d名古屋大学構造生物学研究センター (〒464-8602 名古屋市千種区不老町), ^e名古屋大学計算科学連携教育研究センター (〒464-8603 名古屋市千種区不老町), ^f名古屋大学情報基盤センター (〒464-8601 名古屋市千種区不老町)

*e-mail: sakae@tb.phys.nagoya-u.ac.jp

本総説は、日本薬学会第 135 年会シンポジウム S66 で発表した内容を中心に記述したものである。

知られているアミロイド β タンパク質のシミュレーションである。タンパク質の間違った折り畳みから、それらが凝集することで病気を引き起こすとされるフォールディング病は、実験はもちろんのことシミュレーション分野においても特に注目されているテーマである。そこでわれわれはアミロイド β ペプチド (amyloid β ; A β 25–35) を用いて濃度依存性に着目したシミュレーションを行い、その依存性による違いを初めてシミュレーションにより示した。2つ目はがん治療の標的分子として注目されているヒストン脱アセチル化酵素 (histone deacetylases; HDACs) の特異的阻害リガンド分子の結合自由エネルギーの評価である。最近実験により HDAC3 を特異的に阻害することが可能なりガンド分子 T247 が見つかった。これに対し、われわれのシミュレーション手法を用い、初めてその特異性を明らかにすることができた。

遺伝的アルゴリズムを用いた構造探索方法

まずこのシミュレーション手法では複数の初期構造を用意し、これらに対し同時にシミュレーションを開始する。シミュレーションを実行している間、遺伝的アルゴリズムの操作である交叉と選択を一定シミュレーション時間(又はステップ数)毎に行う。交叉の操作ではシミュレーション中に新しく試行構造の提案をすることが目的である。そして選択の操作として、この試行構造がもし優れていれば、これを新しい初期構造としてシミュレーションを開始する。もしそうでなければ、元の構造をそのまま引き続きシミュレーションを再開する。

以下、この手法について具体的に説明する。遺伝的アルゴリズムではデータを遺伝子で表現した複数の個体を用意する。本研究手法の場合、個体は対象タンパク質の構造に対応し、通常これは偶数個用意する。その後、これら複数の構造について独立した(お互い相互作用しない)シミュレーションを同時に実行する。ある一定期間これらのシミュレーション実行後、遺伝的アルゴリズムの交叉の操作を用いて、シミュレーションで得られた構造とは一部異なる新しい試行構造を作成する。この遺伝的交叉の操作では、1点交叉や2点交叉法などいくつかの方法をこれまでに提案している。⁸⁻¹²⁾ここでは2点交叉について説明する。本手法の2点交叉では、2つのタンパク質構造内の一部の構造情報、すなわち二面

角の値を交換する (Fig. 1)。主な手順は以下の通りである。

(1) 個体数 M に対し、ランダムに $M/2$ ペアを選ぶ。

(2) 選ばれたあるペアについて、タンパク質内の全アミノ酸配列から一部の連続したアミノ酸配列 n をランダムに選ぶ。

(3) 選ばれたアミノ酸配列 n に含まれる二面角(すべての二面角若しくは主鎖の二面角のみ)の値をペア間で交換する。

通常選ばれる連続したアミノ酸配列 n は選ばれたペア毎に異なる。また、この手順で得る試行構造については、さらに以下に説明するような工夫をすることで、安定した試行構造を得る。遺伝的交叉によって上記の手順後得られる構造は通常高いポテンシャルエネルギーを持つ不安定な構造となる。このため、その後行う遺伝的選択の操作を行う前に、比較的安定構造にするための構造緩和処理を行う。構造緩和の過程では、交叉によって交換した二面角値について拘束をかけた短い平衡化シミュレーションを行う。すなわちこの拘束付き平衡化シミュレーションを実行することにより、交叉前の構造から交叉後の構造への変化を徐々に行うことができ、得られる試行構造の多くは自由エネルギー曲面上で少なくとも局所的には安定な構造となる。その後さらに、拘束なしの短い平衡化シミュレーションを行い、最終的に試行構造を得る。そして遺伝的選択の処理へと移る。遺伝的アルゴリズムの選択処理ではオリジナルの個体と交叉や突然変異によって変化した個体間から優れた個体を選び次世代の個体とする。本手法では優れた個体(タンパク質構造)の選択方法として構造のポテンシャルエネルギーについてメトロポリス判定を用いた。すなわち新しい個体を選択する確率として以下の式 $E_p \cdot (1)$ を用いる。

$$w(p \rightarrow c) = \min(1, \exp \{-\beta[E_c - E_p]\}) \quad (1)$$

ここで、 E_p 及び E_c はそれぞれ交叉前の構造、交叉後の構造のポテンシャルエネルギーである。 β は温度の逆数であり、 $\beta = 1/k_B T$ ($k_B T$ はボルツマン定数) で定義される。本手法ではこの交叉と選択の処理を一定シミュレーションステップ毎に繰り返すことで安定構造の探索を行う。

われわれはこの手法を1つのタンパク質、protein A¹³⁾ に適用した。protein A は 60 残基からなる

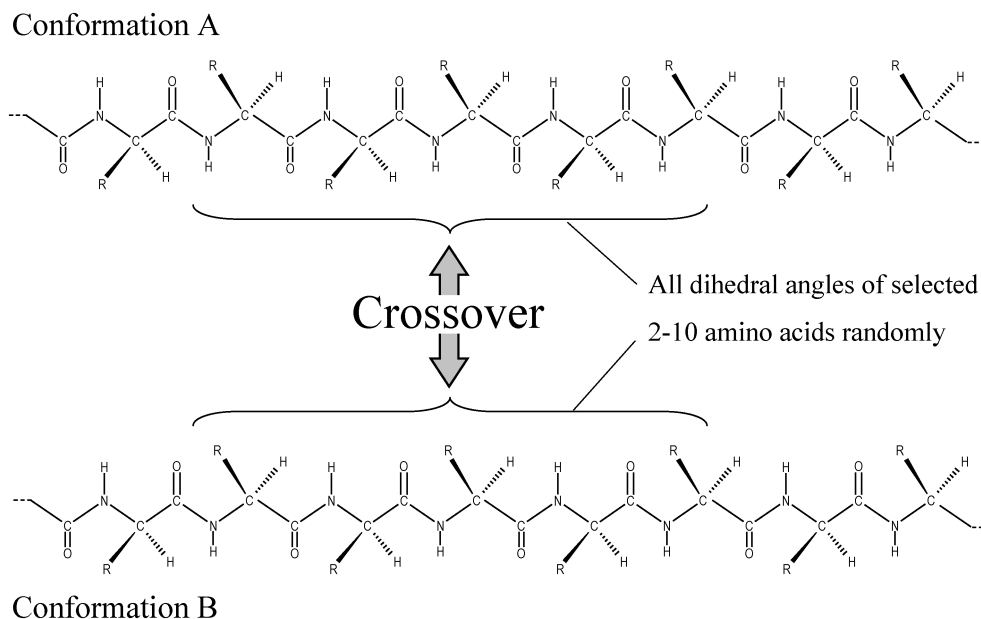


Fig. 1. Schematic Process of the Two-point Crossover Operation

In this process, all dihedral angles (in backbone and side chains) within the randomly selected n consecutive amino acids are exchanged between a pair of conformations.

タンパク質であるが、本計算では Gln10 から Ala55 の比較的不安定な末端部分を取り除いた 46 残基を用いた。シミュレーションプログラムとしては分子動力学シミュレーションプログラム AMBER12 を基に本研究手法の遺伝的交叉選択の処理を加えたものを用いた。時間ステップは水素結合距離について SHAKE¹⁴⁾ 法を用い、2.0 fs とした。系として 32 個体 ($M=32$) 用意し、各々 90 ns (90000000 ステップ) の計算を行った。またその間に、90 回の遺伝的交叉・選択の処理を行った。設定温度は 300 K とし、制御法としてランジュバンダイナミクスを用いた。非結合相互作用は 20 Å のカットオフ、溶媒効果として Generalized Born/Surface Area¹⁵⁾ による陰溶媒モデルを用いることで計算の効率化を図った。交叉処理として、選択するアミノ酸配列長 n は 10 から 20 残基間をランダムに選んだ。

結果として Fig. 2 に示すように実験構造と比較して R.M.S.D. (平均自乗偏差) 1.707 Å となる構造がシミュレーション中に得られた。さらに遺伝的アルゴリズムでは交叉や突然変異、選択を繰り返すことで (世代を経ることで) より優れた個体を見つける手法であるので、世代毎の変化を軸として、二次構造 (ここで α ヘリックスのみ) の形成割合、ポテンシャルエネルギー、R.M.S.D.、慣性半径の値を調べた (Fig. 3)。二次構造の形成割合を見る

と、全体として世代を経る毎にヘリックス構造の割合が増加している。実験構造 (PDB ID: 1BDD) の二次構造解析 (DSSP)¹⁶⁾ では 2 から 9, 17 から 29, 34 から 46 残基の間で α ヘリックスを形成している。シミュレーションで得られたヘリックスの形成率と比較すると [Fig. 3 (a)], おおよそ各ヘリックスの存在確率と一致していることがわかる。protein A 中のこれら 3 つの各ヘリックスの安定性の違いはこれまで多くの研究者によって議論され、いまだ解決していない問題の 1 つである。¹⁷⁻²¹⁾ われわれの結果からは、3 番目のヘリックスが最も安定であり、2 番目のヘリックスが他の 2 つのヘリックスと比較して不安定であった。他の解析結果、ポテンシャルエネルギー、R.M.S.D.、慣性半径では、世代変化の早い段階でそれぞれ大きく減少し、十分収束していることがわかる [Figs. 3 (b)-(d)].

効率的な分子シミュレーション手法の適用例 その 1: $A\beta$ (25-35) 構造の密度依存性

アミロイド β ($A\beta$) はアミロイドーシス (アミロイド病) を引き起こすとされるアミロイドタンパク質の 1 つであり、アルツハイマー病の原因物質として知られている。脳に形成される老人斑の主成分であり、40-43 個のアミノ酸からなる。この $A\beta$ は重合し、アミロイド繊維となって神経細胞を傷害することが要因であると考えられている。また近年実

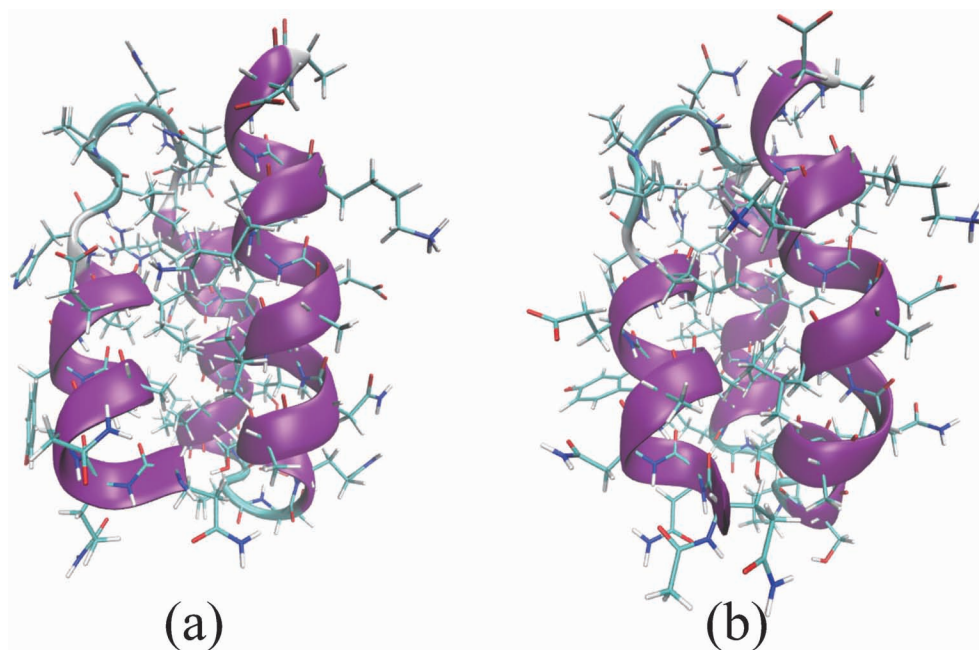


Fig. 2. Structures of Protein A

(a) PDB structure (PDB ID: 1BDD). (b) One conformation obtained from the GC simulation results, which has the lowest R.M.S.D. value (R.M.S.D. = 1.707 Å).

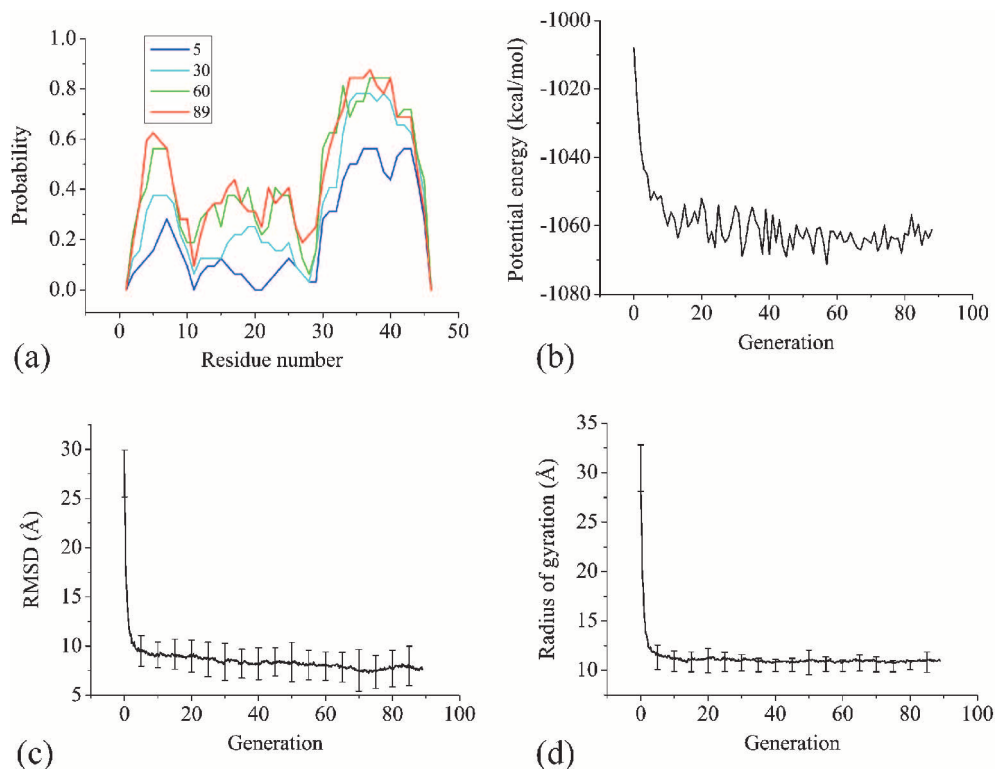


Fig. 3. The Fraction of Helix Conformations (a), the Average of the Potential Energy (b), the Average of the R.M.S.D. (c), and the Average of the Radius of Gyration (d) Obtained from the GC Simulation Results by Producing the Generation. The helix conformations are examined by using DSSP program.¹⁸⁾ The error bars in (c) and (d) stand for the S.D.

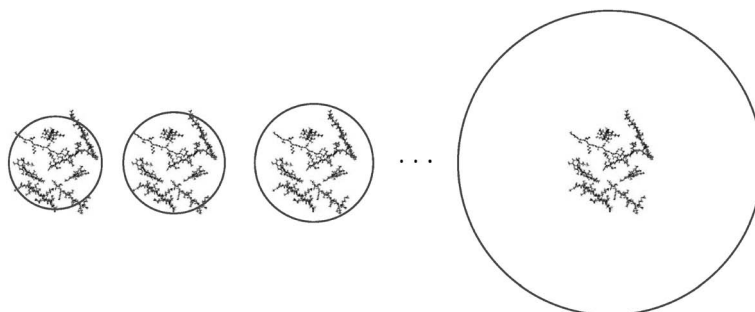


Fig. 4. The Initial Structure and the Constraining Sphere for Each Simulation

Each circle represents the boundary whether the constraint force applies to the fragments or not. Each picture corresponds as follows: The left-most picture shows the case for constraint radius of 27.5 Å, the second Fig. from the left shows the case for constraint radius of 30 Å, and the rightmost picture shows the case for constraint radius of 90 Å.

験によりアミロイド繊維形成の大きな要因の1つとしてAβの過飽和環境が示唆されている。²²⁾そこでわれわれは実験によって繊維化し毒性を有することが既に知られているAβ(25-35)ペプチドを用いて過飽和条件下、すなわちペプチドの高密度状態による分子動力学シミュレーションを行い、その構造形成過程を調べた。

具体的にはペプチドAβ(25-35)11残基を8分子配置した系についてレプリカ交換分子動力学シミュレーション^{6,7,23)}を行った。密度依存性を調べるために同条件のシミュレーションで、かつ密度の異なる系を9種類それぞれ行った(合計34.56μs)。ここで密度の変化として、Aβ(25-35)の動ける範囲(R_l)を27.5Åから90Åに変えている(Fig. 4)。これは約86倍の密度の差を持つ系を比較したことに対応している。

本研究では密度依存性の違いの指標として、シミュレーション中に現れた二次構造、特にβシート構造の形成率を比較した。二次構造判定にはDSSPを用い、系の数密度に対するβシート構造の形成率を調べた。なお数密度 N は以下のように定義した。

$$N = \frac{8}{4\pi R_l^3} = \frac{6}{\pi R_l^3} \quad (2)$$

結果として、ペプチドの数密度が高くなるほどβシート構造の形成率が高くなることが分かった(Fig. 5)。²⁴⁾この傾向は実験で示唆されたタンパク質の構造独自性とタンパク質濃度の関係性²²⁾とおおよそ一致した。

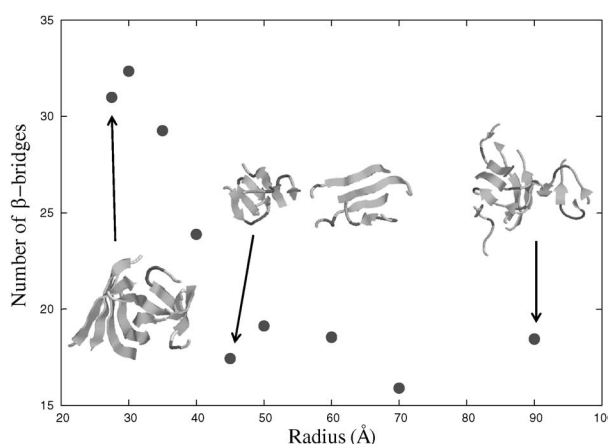


Fig. 5. The Concentration Dependency of β-Structures

The abscissa gives the constraint radius of each simulation, and the ordinate gives the average number of β-bridges included in the structures at 300 K. Three snapshots show typical structures gained by the simulations with constraint radius of 27.5 Å, 45 Å, and 90 Å.

効率的な分子シミュレーション手法の適用例 その2: HDAC3 阻害分子 T247 の選択性

生命活動として重要な役割を担う遺伝子の発現機構は、がんなどの重大な疾患を引き起こす要因になることが知られている。このため近年創薬などの医療分野において、この制御機構に着目した研究が行われている。その中でも特に近年注目を集めている研究分野として、遺伝子の後成的(epigenetic)な制御がある。これはDNAの塩基配列に直接手を加えずに行われる制御のことである。その代表的なものにヒストンのアセチル化と脱アセチル化がある。これらはヒストンに巻き付いているDNAとの結びつきをそれぞれ緩めたり、凝集させたりすることで遺伝子発現の促進、抑制をする。

われわれはこのうち脱アセチル化に着目した。脱アセチル化を触媒するタンパク質ファミリーはヒス

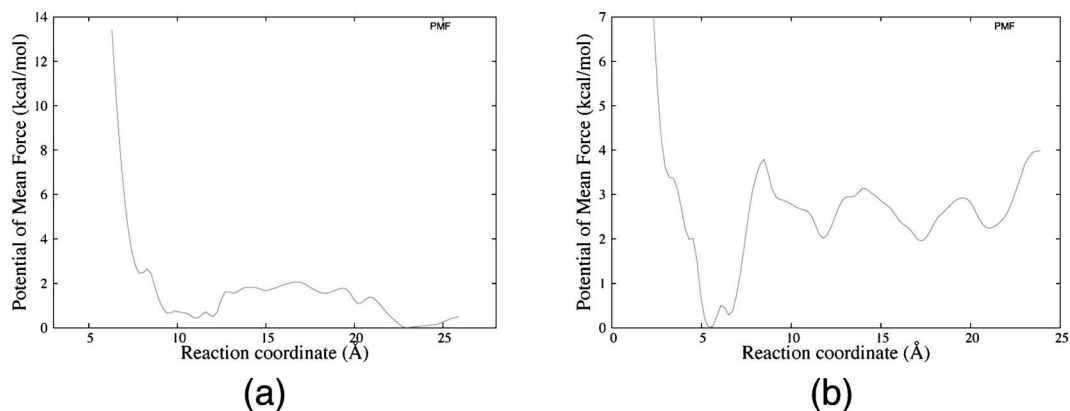


Fig. 6. Potential of Mean Force Profiles along the Reaction Coordinate (the distance from the active site) for HDAC2 (a) and HDAC3 (b)

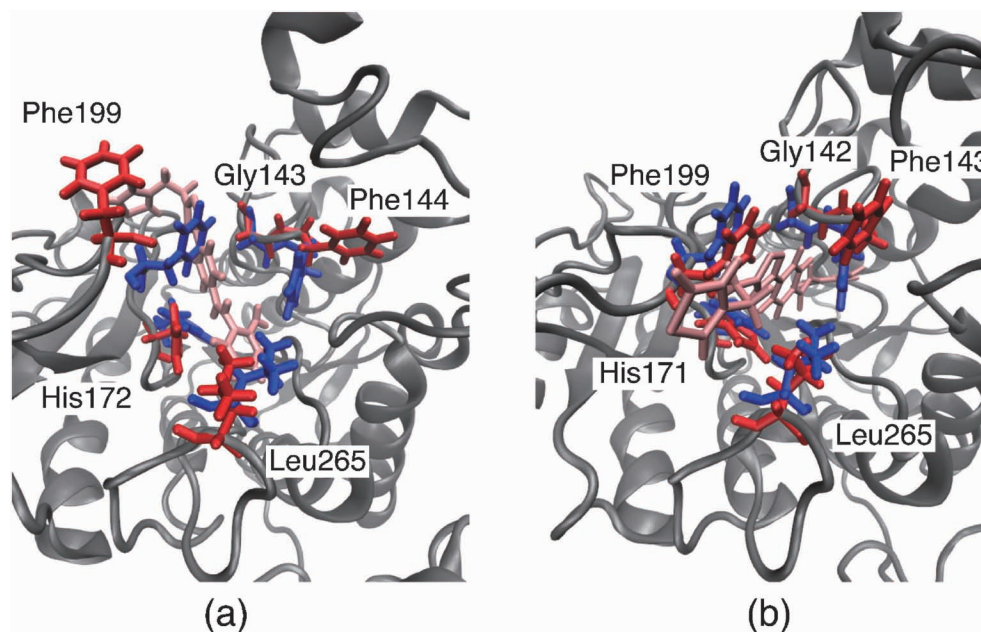


Fig. 7. The Positions of the Five Amino Acids, Which Are Referred to as the Lipophilic Tube,²⁷⁾ for HDAC2 (a) and HDAC3 (b). Red, blue, and magenta stand for the structures obtained from the simulations, the X-ray structures, and the structures of T247 obtained from the simulation, respectively.

トン脱アセチル化酵素 (HDACs) と呼ばれる。最近、これら HDACs のうち HDAC3 だけを選択的に阻害することが可能なリガンド分子 T247 が発見された。²⁵⁾ そこでわれわれはこの T247 の機能を分子シミュレーションによって検証した。具体的には、HDAC3 と同クラスに属する HDAC2 の 2 種類の酵素に対し、T247 との結合自由エネルギーについてアンブレラサンプリングレプリカ交換分子動力学法²⁶⁾を用いて見積もった。この手法を用いるとタンパク質の活性部位とリガンド間との距離を反応座標としたときの、タンパク質-リガンド結合構造を効

率よくサンプリングすることが可能となる。本研究では HDAC2, HDAC3 それぞれについてリガンド分子 T247 と活性部位間の距離を反応座標とした自由エネルギー値を計算した (Fig. 6)。

その結果 HDAC3 では T247 が結合している状態が最も安定となったが、HDAC2 では T247 は活性部位付近にあまり安定に存在できないことが分かった。この安定性の違いは、活性部位への入り口にある疎水性残基²⁷⁾の構造が、HDAC2 では実験構造と比較して、シミュレーションから得られた構造が大きくずれていること (Fig. 7) が要因の 1 つである

と考えられる。

まとめ

本内容ではタンパク質などの複雑な系のための効率的な構造探索の提案と、既存の構造探索手法を特に医療分野に関連のある2つの事例について紹介した。提案した手法は遺伝的アルゴリズムの交叉と選択によって、広い構造空間の探索を行う。ここでは1つの適用例として protein A の構造探索結果を示し、結果として実験構造に近い構造を得ることができた。またこの手法はアルゴリズム上高い並列化効率を持つ利点を有する。後半では構造探索手法の適用例を紹介した。1つはアルツハイマー病の原因物質と考えられている $\alpha\beta$ のペプチドについて、特に密度依存性に注目したシミュレーションを行い、高密度になるほど β シート構造の数が増加することを示した。2つ目はヒストン脱アセチル化酵素について、実験で選択性を持つことが知られているリガンド分子 T247 と酵素との結合自由エネルギーを見積もった。結果として実験によるリガンド分子の選択性を裏付ける傾向となり、さらにその要因が活性部位近くの疎水性アミノ酸であることを分子レベルで予想した。

ここでは医学や創薬などの医療分野に関連する分子シミュレーションの構造解析手法の紹介と適用例を示した。しかしながら現状ではまだ分子シミュレーションが医療分野へ十分貢献できていない。今後も引き続きより効率的なシミュレーション手法の開発及び様々な生体分子への適用を行っていきたい。

謝辞 本研究を遂行するにあたり、有益な御助言と御支援を賜りました名古屋大学石井克哉教授、同志社大学の廣安知之教授、三木光範教授に深く御礼申し上げます。また、この研究は自然科学研究機構計算科学研究センター及び名古屋大学情報基盤センター [HPCI システム利用研究課題 (課題番号: hp130108)] の計算資源によって行われました。なお、本研究の一部は日本学術振興科学研究費補助金新学術領域生命分子システムにおける動的秩序形成と高次機能発現 (No. 25247071) の援助の下行われました。

利益相反 開示すべき利益相反はない。

REFERENCES

- 1) Mitsutake A., Sugita Y., Okamoto Y., *Biopolymers*, **60**, 96–123 (2001).
- 2) Berg B. A., Neuhaus T., *Phys. Lett. B*, **267**, 249–253 (1991).
- 3) Berg B. A., Neuhaus T., *Phys. Rev. Lett.*, **68**, 9–12 (1992).
- 4) Lyubartsev A. P., Martsinovski A. A., Shevkunov S. V., Vorontsov Velyaminov P. N., *J. Chem. Phys.*, **96**, 1776–1783 (1992).
- 5) Marinari E., Parisi G., *Europhys. Lett.*, **19**, 451–458 (1992).
- 6) Hukushima K., Nemoto K., *J. Phys. Soc. Jpn.*, **65**, 1604–1608 (1996).
- 7) Hukushima K., Takayama H., Nemoto K., *Int. J. Mod. Phys. C*, **7**, 337–344 (1996).
- 8) Hiroyasu T., Miki M., Ogura M., Proceedings of the 44th Institute of Systems, Kyoto, May 2000, pp. 113–114.
- 9) Hiroyasu T., Miki M., Ogura M., Okamoto Y., *J. IPS Japan*, **43**, 70–79 (2002).
- 10) Hiroyasu T., Miki M., Ogura M., Aoi K., Yoshida T., Okamoto Y., Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2003), Orlando, FL, July 2003, pp. 117–122.
- 11) Sakae Y., Hiroyasu T., Miki M., Okamoto Y., *J. Comput. Chem.*, **32**, 1353–1360 (2011).
- 12) Sakae Y., Hiroyasu T., Miki M., Ishii K., Okamoto Y., *J. Phys.: Conf. Ser.*, **487**, 012003 (2014).
- 13) Gouda H., Torigoe H., Saito A., Sato M., Arata Y., Shimada I., *Biochemistry*, **31**, 9665–9672 (1992).
- 14) Ryckaert J. P., Ciccotti G., Berendsen H. J. C., *J. Comput. Phys.*, **23**, 327–341 (1977).
- 15) Onufriev A., Bashford D., Case D. A., *Proteins*, **55**, 383–394 (2004).
- 16) Kabsch W., Sander C., *Biopolymers*, **22**, 2577–2637 (1983).
- 17) Zhou Y., Karplus M., *Nature*, **401**, 400–403 (1999).
- 18) Alonso D. O. V., Daggett V., *Proc. Natl. Acad. Sci. USA*, **97**, 133–138 (2000).
- 19) García A. E., Onuchic J. N., *Proc. Natl. Acad. Sci. USA*, **100**, 13898–13903 (2003).
- 20) Sato S., Religa T. L., Daggett V., Fersht A.

- R., *Proc. Natl. Acad. Sci. USA*, **101**, 6952–6956 (2004).
- 21) Shao Q., Gao Y. Q., *J. Chem. Phys.*, **135**, 135102 (2011).
- 22) Yoshimura Y., Lin Y., Yagi H., Lee Y.-H., Kitayama H., Sakurai K., So M., Naiki H., Goto Y., *Proc. Natl. Acad. Sci. USA*, **109**, 14446–14451 (2012).
- 23) Sugita Y., Okamoto Y., *Chem. Phys. Lett.*, **314**, 141–151 (1999).
- 24) Nishikawa N., Sakae Y., Okamoto Y., Proceedings of Computational Science Workshop 2014, Tsukuba, August 2014, Vol. 5, 011020.
- 25) Suzuki T., Kasuya Y., Itoh Y., Ota Y., Zhan P., Asamitsu K., Nakagawa H., Okamoto T., Miyata N., *PLoS ONE*, **8**, e68669 (2013).
- 26) Sugita Y., Kitao A., Okamoto Y., *J. Chem. Phys.*, **113**, 6042–6051 (2000).
- 27) Bressi J. C., Jennings A. J., Skene R., Wu Y., Melkus R., Jong R. D., O’Connell S., Grimshaw C. E., Navre M., Gangloff A. R., *Bioorg. Med. Chem. Lett.*, **20**, 3142–3145 (2010).