

ProteinDF のための古典 MD エンジンの開発

恒川直樹、上野哲哉、佐藤文俊[†]

東京大学生産技術研究所[†]

【背景】

ポストゲノム時代に突入した現在、タンパク質の立体構造および機能の解析に注目が向けられている。そのタンパク質の研究では X 線構造解析・NMR・電子顕微鏡などを用いた原子レベルの分解能で立体構造決定がなされてきた。また、コンピュータシミュレーションによる立体構造予測やその物性解析がなされている。しかし、そのシミュレーションの多くはデータベース型の力場を基にした古典シミュレーションであり、これらの力場が提供するエネルギー精度は構造解析において十分なものではない。

自然界のタンパク質は生体内の環境下で機能している。温度 300K の生体溶液中でタンパク質はその自由エネルギー変化が 0.1eV 程のわずかな変化でその構造や機能を変化させる。このことは、タンパク質の解析において、エネルギー計算について許される誤差が 0.1eV 以下であることを意味する。既存のデータベース型の力場ではこの精度を保証していない。

この精度を保証するタンパク質の全電子状態計算を実現するソフトウェアとして ProteinDF [1]がある。この ProteinDF から得られる精密な電子状態を反映させた力場を取り込み、分子動力学シミュレーションを行うことにより、より精度の高いタンパク質の構造および機能の解析が可能となる。このことは生体高分子の基礎研究の発展のみならず、創薬や新素材としての開発のような生体高分子の応用にも大きな影響を与えるであろう。

タンパク質の大域的構造最適化や自由エネルギー計算、さらに動的特性を観測可能にする分子動力学法を、タンパク質の全電子状態計算を可能にした ProteinDF に組み込むことは明らかに有用であり、本研究ではその開発を行っている。

【力場】

タンパク質の古典分子動力学シミュレーションは何らかの力場（モデル）の下に力の計算を行

い、ニュートン運動方程式にのっとして、タンパク質の分子構造の時間変化を発展させていく方法である。このシミュレーションから得られる結果の正当性、つまり、精度はさまざまな要因によるが、その中でもそのシミュレーションに採用した力場の正当性（精度）に非常に大きく依存している。

原子レベルまで再現した既存のタンパク質の力場はデータベース型をしており、代表的なものとして AMBER が挙げられる。その AMBER の原子間相互作用はエネルギー関数で表すことができる。

$$U = \sum_{\alpha} U_{\alpha}^{bond} + \sum_{\beta} U_{\beta}^{angle} + \sum_{\gamma} U_{\gamma}^{dihedral} + \sum_{i,j} U_{ij}^{LJ} + \sum_{i,j} U_{ij}^{el},$$

$$U_{\alpha}^{bond} = k_{\alpha} (r_{\alpha} - r_{\alpha}^0)^2,$$

$$U_{\beta}^{angle} = k_{\beta} (\theta_{\beta} - \theta_{\beta}^0)^2,$$

$$U_{\gamma}^{dihedral} = V_{\gamma} [1 + \cos(n_{\gamma} \phi_{\gamma} + \delta_{\gamma})],$$

$$U_{ij}^{LJ} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6},$$

$$U_{ij}^{el} = \frac{q_i q_j}{\epsilon r_{ij}}.$$

最初の 3 項は化学結合長、結合角、二面角によるエネルギーを表す。4 項目は化学結合していない分子内および分子間の原子間に働く Lennard-Jones 相互作用であり、5 項目は静電相互作用を表す。AMBER 以外の多くのデータベース型の力場も同じようなエネルギー関数を採用している。

このエネルギー関数はいくつかのパラメータを含んでおり、そのパラメータはアミノ酸の種類や原子の種類によって値が異なる。このパラメータのデータベースが力場を実質的に決定している。このパラメータは量子化学計算（第一原理計算）および経験的に決定され、生体高分子の正しい構造が得られるように全体の整合性が取られ、大規模なタンパク質分子でも計算を可能にしている。

しかしながら、これらのパラメータはアミノ酸を単位とした量子化学計算を基礎にしており、たとえタンパク質内での位置が全く異なる場合でも、同じ種類のアミノ酸残基である場合、そ

Development of the Classical Molecular Dynamics Software for ProteinDF

[†]N. Tsunekawa, T. Ueno, and F. Sato, Institute of Industrial Science, University of Tokyo

のパラメータ値は同じ値を利用することになる。実際のタンパク質では構造やタンパク質内の位置によって、アミノ酸残基の電子状態は異なり、明らかに既存の力場とは食い違いが生じている(図1、表1参照)。すなわち、タンパク質の全電子状態計算はこのような問題点を解消する唯一の方法である。

本発表では ProteinDF から得られた精度の高い物理量を反映させた力場による分子動力学シミュレーションプログラムについて報告する。

【分子動力学法】

古典分子動力学法はニュートン運動方程式を数値的に解くことにより系の時間発展を再現する。運動方程式(微分方程式)は差分方程式に置き換えられ、差分方程式を解くことによりシミュレーションは実行される。本研究では安定性が高いとされている速度ベルレ法を採用した。この差分法は2次のシンプレクティック差分法の一つである。

また、系の温度の拘束条件を入れた差分方程式により温度一定のアンサンブルを生成することが可能となる。これにより、希望温度のシミュレーションが容易になる。

分子動力学シミュレーションにより、タンパク質の大域的な構造変化が得られる。これは、今まで特定の構造のみの電子状態計算にとどまることなく、より正確な構造最適化を可能にする。当日は、従来の方法と本方法の結果の比較について報告する。

【謝辞】

本研究は文部科学省リサーチ・レボリューション計画(RR2002)、ITプログラム「戦略的基盤ソフトウェアの開発」の支援の下に行なわれました。

【参考文献】

[1] F. Sato, Y. Shigemitsu, I. Okazaki, S. Yahiro, M. Fukue, S. Kozuru, H. Kashiwagi, *Int. J. Quantum Chem.* **63**, 245 (1997)

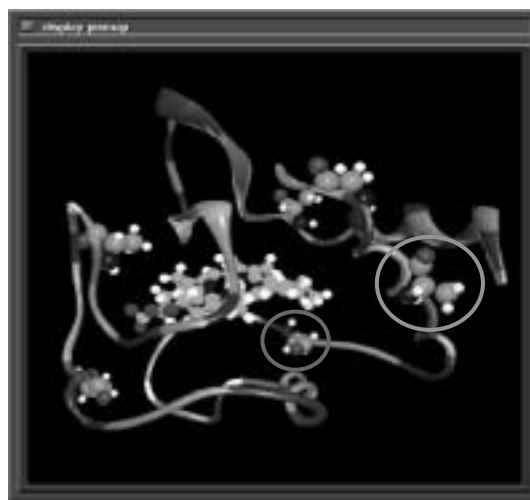


図1： シトクロムc上の83番目のアラニンと96番目のアラニンの位置

	ALA15	ALA43	ALA51	ALA83	ALA96	ALA101	ALA parm94
N	-0.4820	-0.5236	-0.4759	-0.5203	-0.4889	-0.4735	-0.4157
C α	-0.0941	-0.1405	-0.1075	-0.0903	-0.0955	-0.1010	0.0337
C	0.2254	0.2771	0.2057	0.2198	0.2409	0.2497	0.5973
O	-0.3562	-0.3706	-0.3363	-0.3673	-0.3673	-0.3887	-0.5679
C β	-0.6150	-0.5116	-0.5652	-0.5912	-0.6137	-0.6123	-0.1825
H	0.4107	0.4350	0.3329	0.3799	0.4018	0.4016	0.2719
H α	0.2457	0.1971	0.2642	0.2405	0.2448	0.2400	0.0823
H β 1	0.2590	0.2132	0.2573	0.2478	0.2275	0.2453	0.0603
H β 2	0.2134	0.1791	0.2411	0.2101	0.2755	0.2404	0.0603
H β 3	0.2457	0.2625	0.2179	0.2294	0.2514	0.2227	0.0603
合計	0.0526	0.0177	0.0342	-0.0416	0.0765	0.0241	0.0000

表1： シトクロムcの全電子計算によるアラニン(ALA)の形式電荷とparm94のアラニン点電荷
既存のデータベース型の力場では電荷は、タンパク質の構造や位置の違いがあっても、同じ残基の種類であれば変わらないが、タンパク質全体での全電子計算によって得られた形式電荷は位置によって変わってくる。例えば、シトクロムcの83番目のアラニンと96番目のアラニンでは形式電荷が0.1もの差が生じる。電荷の0.1の違いが10先に与える影響はほぼ0.1eV(=3kcal/mol)にもなる。