

利用課題名 大規模 GPU 計算による光合成細菌の細胞膜システムの

全原子シミュレーション

<sup>1</sup>関嶋政和、<sup>1</sup>遠藤敏夫、<sup>2</sup>ジム フィリップス、<sup>2</sup>クラウス シュルテン

1 東京工業大学

2 イリノイ大学

邦文抄録(500 字程度)

紅色光合成細菌において、光合成は chromatophore（色素体）と呼ばれる細胞膜の窪みで行われる。色素体に含まれる膜タンパク質の種類は少なく、光を収集する LH1 アンテナ複合体 (LH1)、LH2 アンテナ複合体 (LH2) や光反応中心 (RC)、そして bc1 複合体などがある。近年、LH1 や LH2 の活発な自己組織化によって色素体の膜の窪み形状が決められるということが知られている。

本研究では、chromatophore の自己組織化と膜の形状の決定のメカニズムを原子レベルで理解するために、原子間力顕微鏡によって得られた情報をもとに組み立てたモデルに対し、TSUBAME2.0 の 700 ノードを用いてノード内 2CPU、3GPU をフルに使用し分子動力学シミュレーションプログラムを行った。

**Keywords:** 分子動力学シミュレーション、膜タンパク質、光合成、自己組織化

1. 研究の背景、意義、研究目的

植物は、太陽光を浴びると、根から吸収した水と空気中から吸収した二酸化炭素をもとに炭水化物と酸素をはき出しており、この働きが光合成として知られている。すなわち、光エネルギーを化学エネルギーに変換を行っている。太陽から放射されたエネルギーは、太陽定数が  $1366\text{W/m}^2$  であることから地球全体が受け取っているエネルギーは  $174\text{PW}$  と求めることが出来る。このうち 47% が地表に到達しており、人口光合成により太陽光エネルギーを化学エネルギーとして貯蔵し、必要なときに必要なエネルギーを取り出せる新エネルギー系を構築する研究・開発が進められている[1,2]。

光合成細菌は、湖沼等の嫌気的な水界に生息している。単一の光化学系をつかった酸素非発生型の光合成を行い ATP を合成しており、遺伝子操作の自由度が高く、エネルギー変換機構の基本原則、光合成の最小ユニット、光合成器官の進化等を解明するための基礎研究に重要である[3]。

紅色光合成細菌の光捕集系は図1に示すように、

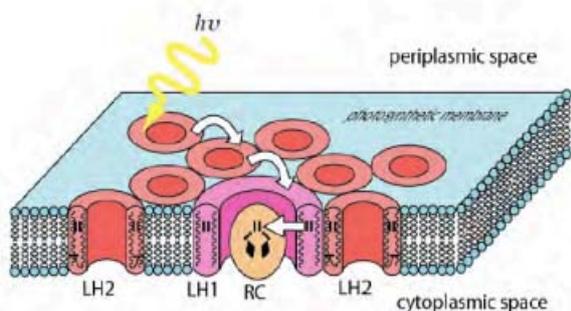


図 1: 紅色光合成細菌の光捕集複合体と反応中心の模式図[4]

光合成は光合成器官クロマトフォア

(chromatophore)と呼ばれる細胞膜の窪みで行われる。クロマトフォアに含まれる膜タンパク質の種類は少なく、光を収集する LH1 アンテナ複合体 (light harvesting complex 1)、LH2 アンテナ複合体 (light harvesting complex 2) や光反応中心 RC (reaction center)、そしてシトクロム bc1 複合体 (cytochrome bc 1 complex) などがある。LH2 を持たない *Rhodospirillum (Rsp.) rubrum* のような菌種もあるが、全ての紅色細菌には LH1 が存在する[4]。

近年の実験によって LH1 や LH2 の活発な自己組織化によってクロマトフォアの膜の窪み形状が決められるということが知られてきている。すなわち、これらの膜タンパク質の自己組織化によってどのように膜の形状が変化するかを知ることは、光合成のメカニズムの理解にもつながると考えられる。

本研究では、この自己組織化と膜の形状の決定のメカニズムを原子レベルで理解するために、原子間力顕微鏡によって得られたイメージをもとに組み立てられたモデル[5,6]をもとに、その分子動力学シミュレーションを行った。このモデルはタンパク質や脂質膜、周囲の水分子をあわせるとおよそ 2000 万原子になる巨大な系であり、TSUBAME2.0 を用いた大規模な実験により、膜とタンパク質複合体の相互作用の解析が可能になると考えられる。

2. 分子動力学法

分子動力学法 (Molecular Dynamics: MD) は古典力学に基づき分子の振る舞いを求めるシミュレーション方法である。システムを構成する各原子について、微小時間ごとに周囲の原子との相互作用力を求め、Newton の運動方程式  $ma = F$  を数値積分法で解き系の時間発展を求める。生体分子

## 様式 グランドチャレンジ大規模計算制度 利用成果報告書（英語版）

を対象にした分子動力学法は分子のダイナミクスの解析、構造最適化、タンパク質の結合部位予測などに応用されている[7]。

分子動力学法では原子間の相互作用は分子力場を用いて計算される。

$$U(\mathbf{q}) = \sum_{bonds} k^b (b - b_0)^2 + \sum_{angles} k^\theta (\theta - \theta_0)^2 + \sum_{torsions} V_n [1 + \cos^2(n\phi - \phi_0)] + \sum_{i,j \in atoms} \epsilon_{ij} \left[ \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \right] + \sum_{i,j \in atoms} \frac{q_i q_j}{r_{ij}} \quad (1)$$

式(1)は代表的な力場の1つである AMBER 力場 [8]であり、第一項は化学結合の長さ、第二項は化学結合の角度、第三項は化学結合による二面角、第四項は van der Waals 力、第五項は静電相互作用力を表す。このうち van der Waals 力と静電相互作用力は化学結合によらず離れた原子間に働く力であるため、原子数の 2 乗で計算量が増加する。実際に多くの場合分子動力学シミュレーションの実行時間の 9 割以上がこれら 2 種類の力の評価に費やされる。

求めた相互作用力から運動方程式を用いて次のシミュレーションステップでの原子の座標と速度を計算する。分子動力学法で用いられるシミュレーションの時間幅は通常 1~2 フェムト秒程度の非常に小さい時間である。科学的に意義のある結果を得るためにはナノ秒やマイクロ秒のオーダーの計算が必要となるため、シミュレーションを数百万ステップ以上実行する必要がある。したがって分子動力学法は計算コストが高く、その高速化が強く求められている計算の1つである。

### 3. NAMD

NAMD[9]は Illinois 大学の Theoretical and Computational Biophysics Group が公開している分子動力学シミュレーションプログラムである。同大学の Parallel Programming Laboratory で開発されている Charm++[10]で記述されており、特に巨大分子系のシミュレーションで高い性能を發揮するよう設計されている。

NAMD ではシミュレーション空間の分割と相互作用力計算の分割によってスケーラビリティを向上させる工夫がなされている。シミュレーションのはじめにシミュレーションの対象となる空間を patch と呼ばれる単位に分割し、計算を行う各プロセッサに分配する。各プロセッサは分配された patch に含まれる原子の座標や速度のデータを扱う。この patch の大きさはカットオフの距離を考慮し、隣接

する patch のみで化学結合、van der Waals 力や静電相互作用力の直接計算部ができるように設定される。このように空間的な局在性を考慮することによって大規模な系に対しても計算量やプロセス間の通信を低減させることが可能である。

相互作用力の計算は力の種類ごとに compute object を定義して行う。化学結合による力を計算する compute object では、計算に必要となる原子が属する patch に基づきその力が計算されるプロセスを 1 つ決定する。これにより複数 patch、複数プロセスにまたがる原子を要する計算を効率的に行い、データの通信量も低減させることが可能である。

化学結合によらない van der Waals 力や静電相互作用力の計算が分子動力学法の実行時間の 80~90%程度を占める。NAMD ではこれらの計算に対して GPU を利用することで計算を高速化させている。NAMD の実装では複数の CPU プロセスが 1 つの GPU を共有し、そのプロセスが持つ空間内の原子について非結合力の計算を行う。このように非結合力の計算のみに GPU を適用し、複数の CPU プロセスで GPU を共有させることで、計算機の構成に対して柔軟でスケーラビリティの高い実装を行っている。

### 4. 大規模分子動力学シミュレーションの GPGPU における性能測定

本研究では TSUBAME2.0 のノードを 700 台、GPU を 2100 台まで使用して、NAMD のベンチマークデータを測定した。対象とした紅色光合成細菌の系はタンパク質、脂質二重膜、溶媒からなる総原子数 20,305,122 個の極めて巨大な系である。使用した NAMD、Charm++のバージョンはそれぞれ 2.8、6.3.2 である。コンパイルには gcc version 4.3.4、cuda toolkit 4.0 を用いている。測定は 4 ノードから 700 ノードまでノード数を増加させ、ノード数の増加に対する NAMD のスケーリングの性能を調べた。また、各ノード数について GPU をノードあたり 0 台から 3 台まで使用するデータの測定も行った。

図2に、結果は ns/day、すなわち 1 日あたりに進むシミュレーションの秒数で示されている。通常、ノード数や GPU の台数が増加すると、計算量に対してデータ転送や同期のコストが増加することからパフォーマンスが伸びないことが多い。対して NAMD はノード数、GPU 数双方の増加に対して 700 ノード 2100 GPU までパフォーマンスの向上をみることができた。スケーリング性能だけではなく GPU の利用によるパフォーマンスの向上も目を見張る。NAMD では GPU を 1 台利用するだけでも、CPU のみの場合に比べ 3 倍程度のパフォーマンスの向上がみられる。

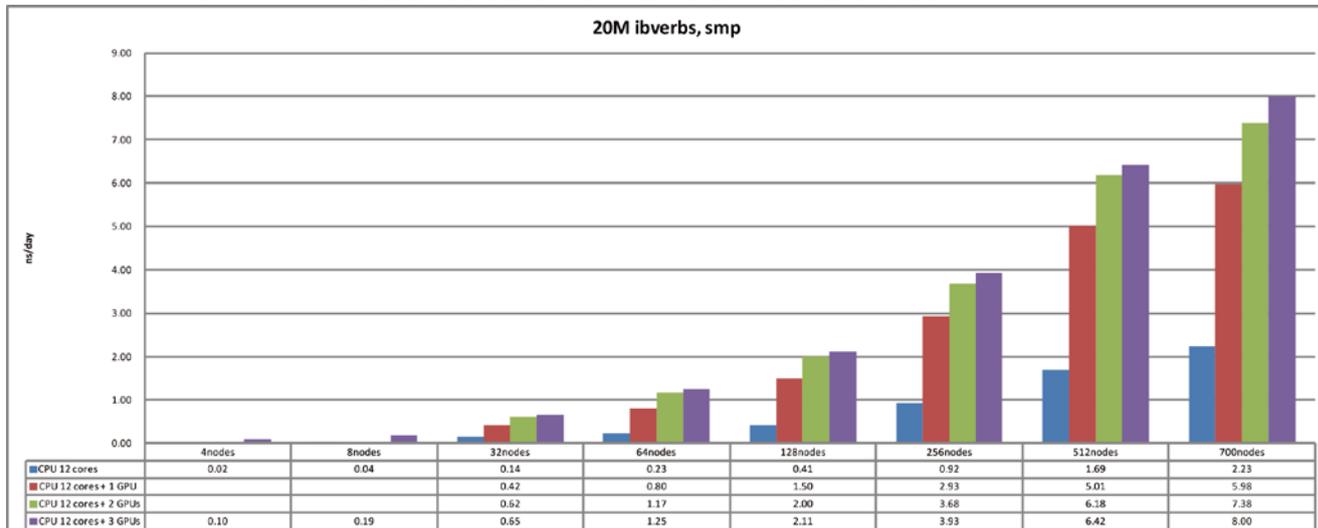


図 2: NAMD のパフォーマンスベンチマーク

5. 大規模分子動力学シミュレーションによる光捕集タンパク質の自己組織化過程の解析

本研究では Illinois 大学の Klaus Schulten らの研究グループが作成した紅色光合成細菌の分子モデルに対して、TSUBAME2.0 の 700 ノード 2100GPU を用いて分子動力学シミュレーションを行った。データの I/O などコストがかかりベンチマークと比較するとパフォーマンスは低下するものの、グランドチャレンジ期間中に合計 7 ナノ秒のシミュレーションを実行することに成功した。

図 3 に 7 ナノ秒間の系のポテンシャルエネルギーの変化を示す。ポテンシャルエネルギーは、系が不安定な場合やシミュレーションがうまくいっていない場合に平均値が大きく変動する。今回の結果では概ね一定の値を保っており安定したシミュレーションを実行できていたと判断することができる。

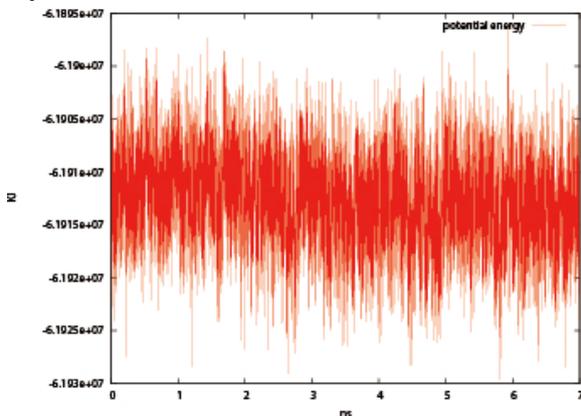


図 3: ポテンシャルエネルギーの変化

次に光捕集タンパク質の自己組織化過程についての解析を行った。紅色光合成細菌の光合成の反応過程では反応中心 RC に対して光捕集タンパクの自己組織化が起きることが報告されている。本研究では反応中心を持つ LH1 複合体と反応中心を

持たない LH2 複合体の重心距離の変化に着目し解析を行った。以降の議論では図 4 に示すように各 LH1 複合体、LH2 複合体にアルファベット 1 文字の名前をつけて区別する。A~Z、a~j が LH2 複合体に割り振られ、k~q が LH1 複合体に割り振られている。

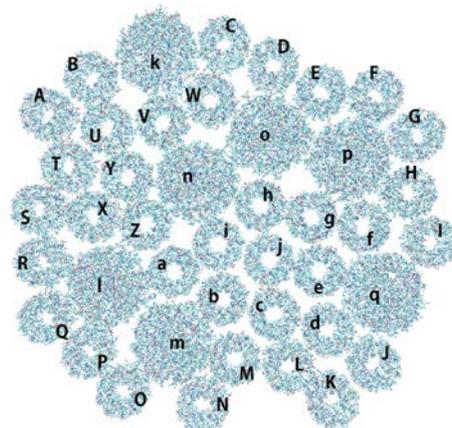


図 4: 各タンパク質のラベル

図 5 は各 LH1 複合体から、近傍の LH2 複合体への距離の時間変化をプロットした図である。この図をよると、例えば LH1 複合体 l と LH2 複合体 a や LH2 複合体 Q、LH1 複合体 p と LH2 複合体 H、LH1 複合体 n と LH2 複合体 V などは時間経過とともに重心間距離がゆるやかに縮んでいる傾向がみとれる。特に LH1 複合体 n と LH2 複合体 V の距離が縮んでいるのに対し、LH1 複合体 k と LH2 複合体 V の距離はほぼ一定であり、LH1 複合体 n と LH2 複合体 i の距離は増加傾向にあることから、LH1 複合体 n が引き寄せられていると推察される。

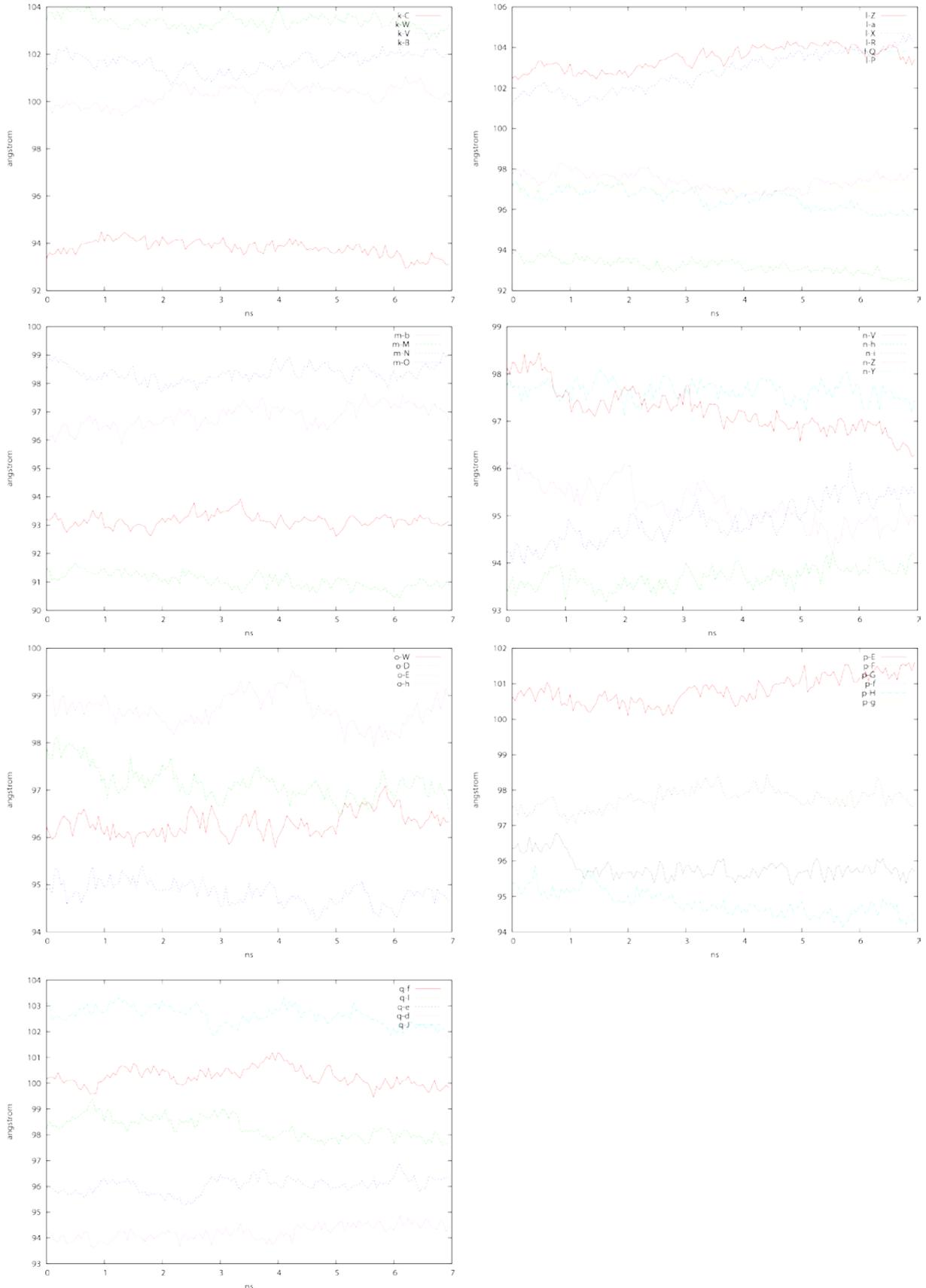


図 5: 各 LH1 と近傍の LH2 間の距離の推移

## 6. まとめ

本研究では、紅色光合成細菌の光捕集複合体と反応中心の大規模な分子モデルに対して TSUBAME2.0 の 700 ノードを用いて分子動力学シミュレーションを行うことで、GPGPU の性能評価

とクロマトフォア内のタンパク質の相互作用を解析することを目的とした。

分子動力学シミュレーションは数カ月単位で実行されることもあり、系が大きくなるほど長時間の実行時間が必要となる。GPU は比較的安価で導入が可能であり、計算速度の向上が見込めることから、

多くの研究者にとって分子動力学シミュレーションに要する時間を短縮させるための選択肢として期待されている。本研究で性能評価に用いた総原子数 20,305,122 個の系においても、CPU のみでシミュレーションした場合に対し、GPU を用いた場合、128 ノードの時に最大 5.1 倍の高速に実行することが可能であり、700 ノードにおいても 3.6 倍ほど高速に実行が可能であることを示した。また、我々が別に示したように、GPU を用いることでエネルギー消費を抑えることも可能である[11]。

膜環境中の複数のタンパク質の関与による現象は、従来の小さい分子系のシミュレーションで観測することは難しい。大規模な系を作成し、そのシミュレーションを実行できる計算機環境があつて初めて観測することができる現象である。本研究では GPU で加速された TSUBAME 2.0 の大規模な計算環境を用いることで 3 日程度の期間に 7 ナノ秒のシミュレーションを実行し、紅色光合成細菌のタンパク質の自己組織化の過程をシミュレーションすることに成功した。しかし、分子系が大きくなることにより長時間のシミュレーションが必要となる。今回得られた結果も、自己組織化の初期の過程の兆候が見られたが、まだ完全な現象の再現には及ばない。今後さらにシミュレーションを実行し、解析を加えていく必要があると考えられる。

66, pp.27-85 (2003).

[9] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry* **26**, pp.1781-1802 (2005).

[10] L. V. Kale, E. Bohm, C. L. Mendes, T. Wilmarth and G. Zheng, Programming Petascale Applications with Charm++ and AMPI. In Petascale Computing, *Algorithms and Applications*, pp. 421-441, Chapman & Hall / CRC Press, ed D Bader (2004).

[11] S. Du, T. Udagawa, T. Endo and M. Sekijima, Molecular Dynamics Simulation of a Biomolecule with High Speed, Low Power and Accuracy Using GPU-Accelerated TSUBAME2.0 Supercomputer, Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2011 (APSIPA ASC 2011), pp.1-5, 2011

#### 参考文献

[1] A. Listorti, J. Durrant and J. Barber, Artificial photosynthesis: Solar to fuel, *Nature Materials* **8**, pp. 929 - 930 (2009).

[2] M. W. Kanan and D. G. Nocera, In Situ Formation of an Oxygen-Evolving Catalyst in Neutral Water Containing Phosphate and Co<sup>2+</sup>, *Science* **321**, pp. 1072-1075 (2008).

[3] 塚谷祐介, 浅井智広, 大岡宏造, 緑色硫黄細菌の非循環型の光合成電子伝達系, *光合成研究* **20**, pp.100-106 (2010).

[4] 大友征宇, 紅色光合成細菌における光捕集タンパク質の多様性, *光合成研究* **20**, pp.109-117 (2010).

[5] M. K. Sener, J. Hsin, L. G. Trabuco, E. Villa, P. Qian, C. N. Hunter, and K. Schulten, Structural model and excitonic properties of the dimeric RC-LH1-PufX complex from Rhodospirillum rubrum, *Chemical Physics* **357**, pp.188-197, (2009).

[6] J. Hsin, J. Strümpfer, M. Sener, P. Qian, C. N. Hunter, and K. Schulten. Energy transfer dynamics in an RC-LH1-PufX tubular photosynthetic membrane, *New Journal of Physics* **12**, pp. (2010).

[7] M. Sekijima, C. Motono, S. Yamasaki, K. Kaneko, and Y. Akiyama, Molecular dynamics simulation of dimeric and monomeric forms of human prion protein: Insight into dynamics and properties, *Biophysical Journal* **85**, pp.1176-1185 (2003).

[8] J.W. Ponder and D.A. Case, Force fields for protein simulations, *Advances in Protein Chemistry*