HRMS. Zinc(II) or copper(II) complex formation was accomplished by treatment of the above compounds with 10 equiv of $ZnCl_2$ or $CuCl_2$ in PBS. All zinc(II) or copper(II) complexes were characterized by chemical shifts of their methylene protons in ^1H NMR analysis. The pyridyl zinc(II) complex was characterized previously,[37] and zinc(II) or copper(II) complex formation with these macrocyclic compounds has been reported elsewhere.[41,42,48,49] Detailed procedures and data are provided in the Supporting Information.

Biological assays

A CXCR4 binding assay for compounds, based on the inhibition of [^{125}I]CXCL12 binding to Jurkat cells, was performed as reported by Tanaka et al.[38] CXCR4 antagonistic activity was evaluated as described by Ichiyama et al[27], measuring inhibitory activity against Ca^{2+} mobilization induced by CXCL12 stimulation in HOS cells expressing CXCR4. Anti-HIV activity was determined by inhibitory activity against X4-HIV-1(NL4–3)-induced cytopathogenicity in MT-4 cells as reported by Tanaka et al.[38] An X4 HIV-1 infectious molecular clone (pNL4–3) was obtained from the AIDS Research and Reference Reagent Program. The virus NL4–3 was obtained from the culture supernatant of 293T cells transfected with pNL4–3.

Molecular modeling

Molecular modeling calculations were performed using Sybyl (version 7.0, Tripos). Energy minimization was performed using the Tripos force field and Gasteiger–Hückel charge parameters. The lowest energy conformation was obtained by random search methods.

Acknowledgements

T.T. and N.O. are supported by research fellowships for young scientists from the Japan Society for the Promotion of Science. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and Health and Labor Sciences Research Grants from the Japanese Ministry of Health, Labor, and Welfare. The authors thank Mr. Wu Honggui (Tokyo University of Science) for his assistance with the anti-HIV assay.

Keywords: azamacrocycles · Ca^{2+} mobilization · CXCR4 · HIV · structure–activity relationships

[1] T. Nagasawa, H. Kikutani, T. Kishimoto, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2305–2309.
[2] C. C. Bleul, M. Farzan, H. Choe, C. Parolin, I. Clark-Lewis, J. Sodroski, T. A. Springer, *Nature* **1996**, *382*, 829–833.
[3] E. Oberlin, A. Amara, F. Bachelerie, C. Bessia, J. L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J. M. Heard, I. Clark-Lewis, D. L. Legler, M. Loetscher, M. Baggiolini, B. Moser, *Nature* **1996**, *382*, 833–835.
[4] K. Tashiro, H. Tada, R. Heilker, M. Shirozu, T. Nakano, T. Honjo, *Science* **1993**, *261*, 600–603.
[5] C. C. Bleul, R. C. Fuhlbrigge, J. M. Casanovas, A. Aiuti, T. A. Springer, *J. Exp. Med.* **1996**, *184*, 1101–1109.
[6] K. Tachibana, S. Hirota, H. Iizasa, H. Yoshida, K. Kawabata, Y. Kataoka, Y. Kitamura, K. Matsushima, N. Yoshida, S. Nishikawa, T. Kishimoto, T. Nagasawa, *Nature* **1998**, *393*, 591–594.
[7] T. Nagasawa, S. Hirota, K. Tachibana, N. Takakura, S. Nishikawa, Y. Kitamura, N. Yoshida, H. Kikutani, T. Kishimoto, *Nature* **1996**, *382*, 635–638.
[8] Y. Zhu, Y. Yu, X. C. Zhang, T. Nagasawa, J. Y. Wu, Y. Rao, *Nat. Neurosci.* **2002**, *5*, 719–720.
[9] R. K. Stumm, C. Zhou, T. Ara, F. Lazarini, M. Dubois-Dalcq, T. Nagasawa, V. Hollt, S. Schulz, *J. Neurosci.* **2003**, *23*, 5123–5130.
[10] H. K. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. D. Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau, *Nature* **1996**, *381*, 661–666.
[11] Y. Feng, C. C. Broder, P. E. Kennedy, E. A. Berger, *Science* **1996**, *272*, 872–877.
[12] T. Koshiba, R. Hosotani, Y. Miyamoto, J. Ida, S. Tsuji, S. Nakajima, M. Kawaguchi, H. Kobayashi, R. Doi, T. Hori, N. Fujii, M. Imamura, *Clin. Cancer Res.* **2000**, *6*, 3530–3535.
[13] A. Müller, B. Homey, H. Soto, N. Ge, D. Catron, M. E. Buchanan, T. McClanahan, E. Murphy, W. Yuan, S. N. Wagner, J. L. Barrera, A. Mohar, E. Verastegui, A. Zlotnik, *Nature* **2001**, *410*, 50–56.
[14] H. Tamamura, A. Hori, N. Kanzaki, K. Hiramatsu, M. Mizumoto, H. Nakashima, N. Yamamoto, A. Otaka, N. Fujii, *FEBS Lett.* **2003**, *550*, 79–83.
[15] N. Tsukada, J. A. Burger, N. J. Zvaifler, T. J. Kipps, *Blood* **2002**, *99*, 1030–1037.
[16] J. Juarez, K. F. Bradstock, D. J. Gottlieb, L. J. Bendall, *Leukemia* **2003**, *17*, 1294–1300.
[17] T. Nanki, K. Hayashida, H. S. El-Gabalawy, S. Suson, K. Shi, H. J. Girschick, S. Yavuz, P. E. Lipsky, *J. Immunol.* **2000**, *165*, 6590–6598.
[18] H. Tamamura, M. Fujisawa, K. Hiramatsu, M. Mizumoto, H. Nakashima, N. Yamamoto, A. Otaka, N. Fujii, *FEBS Lett.* **2004**, *569*, 99–104.
[19] T. Murakami, T. Nakajima, Y. Koyanagi, K. Tachibana, N. Fujii, H. Tamamura, N. Toshida, M. Waki, A. Matsumoto, O. Yoshie, T. Kishimoto, N. Yamamoto, T. Nagasawa, *J. Exp. Med.* **1997**, *186*, 1389–1393.
[20] D. Schols, S. Struyf, J. Van Damme, J. A. Este, G. Henson, E. DeClarcq, *J. Exp. Med.* **1997**, *186*, 1383–1388.
[21] B. J. Doranz, K. Grovit-Ferbas, M. P. Sharron, S.-H. Mao, M. Bidwell Goetz, E. S. Daar, R. W. Doms, W. A. O'Brien, *J. Exp. Med.* **1997**, *186*, 1395–1400.
[22] G. A. Donzella, D. Schols, S. W. Lin, J. A. Este, K. A. Nagashima, *Nat. Med.* **1998**, *4*, 72–76.
[23] O. M. Z. Howard, J. J. Oppenheim, M. G. Hollingshead, J. M. Covey, J. Bigelow, J. J. McCormack, R. W. Buckheit, Jr., D. J. Clanton, J. A. Turpin, W. G. Rice, *J. Med. Chem.* **1998**, *41*, 2184–2193.
[24] H. Tamamura, Y. Xu, T. Hattori, X. Zhang, R. Arakaki, K. Kanbara, A. Omagari, A. Otaka, T. Ibuka, N. Yamamoto, H. Nakashima, N. Fujii, *Biochem. Biophys. Res. Commun.* **1998**, *253*, 877–882.
[25] H. Tamamura, A. Omagari, S. Oishi, T. Kanamoto, N. Yamamoto, S. C. Peiper, H. Nakashima, A. Otaka, N. Fujii, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2633–2637.
[26] N. Fujii, S. Oishi, K. Hiramatsu, T. Araki, S. Ueda, H. Tamamura, A. Otaka, S. Kusano, S. Terakubo, H. Nakashima, J. A. Broach, J. O. Trent, Z. Wang, S. C. Peiper, *Angew. Chem.* **2003**, *115*, 3373–3375; *Angew. Chem. Int. Ed.* **2003**, *42*, 3251–3253.
[27] K. Ichiyama, S. Yokoyama-Kumakura, Y. Tanaka, R. Tanaka, K. Hirose, K. Bannai, T. Edamatsu, M. Yanaka, Y. Niitani, N. Miyano-Kurosaki, H. Takaku, Y. Koyanagi, N. Yamamoto, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4185–4190.
[28] H. Tamamura, N. Fujii, *Curr. Drug Targets-Infectious Disorders* **2004**, *4*, 103–110.
[29] H. Tamamura, K. Hiramatsu, S. Ueda, Z. Wang, S. Kusano, S. Terakubo, J. O. Trent, S. C. Peiper, N. Yamamoto, H. Nakashima, A. Otaka, N. Fujii, *J. Med. Chem.* **2005**, *48*, 380–391.
[30] H. Tamamura, T. Araki, S. Ueda, Z. Wang, S. Oishi, A. Esaka, J. O. Trent, H. Nakashima, N. Yamamoto, S. C. Peiper, A. Otaka, N. Fujii, *J. Med. Chem.* **2005**, *48*, 3280–3289.
[31] G. C. Valks, G. McRobbie, E. A. Lewis, T. J. Hubin, T. M. Hunter, P. J. Sadler, C. Pannecouque, E. De Clercq, S. J. Archibald, *J. Med. Chem.* **2006**, *49*, 6162–6165.
[32] W. Zhan, Z. Liang, A. Zhu, S. Kurtkaya, H. Shim, J. P. Snyder, D. C. Liotta, *J. Med. Chem.* **2007**, *50*, 5655–5664.
[33] A. Khan, G. Nicholson, J. Greenman, L. Madden, G. McRobbie, C. Pannecouque, E. De Clercq, R. Ullom, D. L. Maples, R. D. Maples, J. D. Silversides, T. J. Hubin, S. J. Archibald, *J. Am. Chem. Soc.* **2009**, *131*, 3416–3417.
[34] G. J. Bridger, R. T. Skerlj, P. E. Hernandez-Abad, D. E. Bogucki, Z. Wang, Y. Zhou, S. Nan, E. M. Boehringer, T. Wilson, J. Crawford, M. Metz, S. Hatse, K. Princen, E. De Clercq, D. Schols, *J. Med. Chem.* **2010**, *53*, 1250–1260.

[35] R. T. Skerlj, G. J. Bridger, A. Kaller, E. J. McEachern, J. B. Crawford, Y. Zhou, B. Atsma, J. Langille, S. Nan, D. Veale, T. Wilson, C. Harwig, S. Hatse, K. Princen, E. De Clercq, D. Schols, *J. Med. Chem.* **2010**, *53*, 3376–3388.
[36] M. Takenaga, H. Tamamura, K. Hiramatsu, N. Nakamura, Y. Yamaguchi, A. Kitagawa, S. Kawai, H. Nakashima, N. Fujii, R. Igarashi, *Biochem. Biophys. Res. Commun.* **2004**, *320*, 226–232.
[37] H. Tamamura, A. Ojida, T. Ogawa, H. Tsutsumi, H. Masuno, H. Nakashima, N. Yamamoto, I. Hamachi, N. Fujii, *J. Med. Chem.* **2006**, *49*, 3412–3415.
[38] T. Tanaka, H. Tsutsumi, W. Nomura, Y. Tanabe, N. Ohashi, A. Esaka, C. Ochiai, J. Sato, K. Itotani, T. Murakami, K. Ohba, N. Yamamoto, N. Fujii, H. Tamamura, *Org. Biomol. Chem.* **2008**, *6*, 4374–4377.
[39] G. J. Bridger, R. T. Skerlj, D. Thornton, S. Padmanabhan, S. A. Martellucci, G. W. Henson, M. J. Abrams, N. Yamamoto, K. De Vreese, R. Pauwels, E. De Clercq, *J. Med. Chem.* **1995**, *38*, 366–378.
[40] G. J. Bridger, R. T. Skerlj, S. Padmanabhan, S. A. Martellucci, G. W. Henson, M. J. Abrams, H. C. Joao, M. Witvrouw, K. De Vreese, R. Pauwels, E. De Clercq, *J. Med. Chem.* **1996**, *39*, 109–119.
[41] Y. Inouye, T. Kanamori, T. Yoshida, T. Koike, M. Shionoya, H. Fujioka, E. Kimura, *Biol. Pharm. Bull.* **1996**, *19*, 456–458.
[42] L. O. Gerlach, J. S. Jakobsen, K. P. Jensen, M. R. Rosenkilde, R. T. Skerlj, U. Ryde, G. J. Bridger, T. W. Schwartz, *Biochemistry* **2003**, *42*, 710–717.
[43] H. F. Egberink, E. De Clercq, A. L. Van Vliet, J. Balzarini, G. J. Bridger, G. Henson, M. C. Horzinek, D. Schols, *J. Virol.* **1999**, *73*, 6346–6352.
[44] M. Le Baccon, F. Chuburu, L. Toupet, H. Handel, M. Soibinet, I. Dechamps-Olivier, J.-P. Barbier, M. Aplincourt, *New J. Chem.* **2001**, *25*, 1168–1174.
[45] B. Antonioli, D. J. Bray, J. K. Clegg, K. Gloe, K. Gloe, O. Kataeva, L. F. Lindoy, J. C. McMurtrie, P. J. Steel, C. J. Sumby, M. Wenzel, *Dalton Trans.* **2006**, 4783–4794.
[46] S. P. Foxon, D. Utz, J. Astner, S. Schindler, F. Thaler, F. W. Heinemann, G. Liehr, J. Mukherjee, V. Balamurugan, D. Ghosh, R. Mukherjee, *Dalton Trans.* **2004**, 2321–2328.
[47] S. Mandal, F. Lloret, R. Mukherjee, *Inorg. Chim. Acta* **2009**, *362*, 27–37.
[48] M. Soibinet, I. De'champs-Olivier, E. Guillon, J.-P. Barbier, M. Aplincourt, F. Chuburu, M. Le Baccon, H. Handel, *Eur. J. Inorg. Chem.* **2003**, 1984–1994.
[49] R. W. Hay, M. T. Tarafder, *Transition Met. Chem.* **1990**, *15*, 490–492.

Received: December 19, 2010
Published online on February 10, 2011

高性能計算による薬剤分子設計

[1]理化学研究所　生命システム研究センター　生命モデリングコア計算分子設計研究グループ

[2]計算科学研究機構　プロセッサ研究チーム

泰地　真弘人[12]、沖本　憲明[1]

taiji@riken.jp, okimoto@gsc.riken.jp

高性能計算機の発展と計算科学

電子計算機は、1946 年に ENIAC が完成して以来ほぼ 5 年で 10 倍のペースで発展を続けており、2011 年には 10PFLOPS(10^{16} 演算/秒)の性能に到達する見込みである。さらに今後も、10 年ぐらいは同等のペースでの発展が期待される。その性能向上は主に半導体技術の発展に支えられてきた。しかし、単体プロセッサの動作速度は電力や配線遅延などの制約により限界に到達してしまい、それでも集積度の向上は継続しているため、近年では高並列処理技術なしでは性能向上が見込めなくなっている。今後は、100 万以上の大規模な並列化に取り組む必要があり、ハードウェア開発のみならずソフトウェア開発の超大規模並列化が重要な研究開発課題となっている。計算科学においても、並列化を重視した計算アルゴリズム・ソフトウェア開発が必須になりつつある。また、分子設計などでは探索のために多量の計算を行う必要がある。計算資源の増大につれ、1 万以上のシミュレーションを実行しデータ処理を行うようなことが日常的に行われるようになると、計算やエラー処理の自動化、計算結果や初期値のデータベース化などが重要になる。こうした計算の周辺技術の取り込みも今後の重要な課題である。いずれにせよ、1PFLOPS-1EFLOPS という巨大な計算資源をいかにうまく使い、科学的成果・技術的創出を達成できるかが計算科学に今後問われている課題である。次に説明する京コンピュータの利用をきっかけに、大規模計算資源を実際の成果に結びつけていくための開発が一層進むことが期待される。

京コンピュータ

京コンピュータは、現在理化学研究所と富士通株式会社で開発を進めているスーパーコンピュータである（図 1）。LINPACK ベンチマークで 10PFLOPS の性能を出すことも一つの性能目標であるが、科学的成果の創出に向けてアプリケーションの実効性能が重視されている。実際、他の世界トップクラスの計算機と比較すると、高性能計算に向けた最適化が随所に見られる。システムの完成は 2012 年度を予定しており、その年の秋から本格的な供用が開始される計画である。開発は順調に推移しており、部分的に稼働させたシステムを用いた LINPACK ベンチマークで既に 8.16PFLOPS の性能を達成し、2011 年 6 月の TOP500 リスト[1]で 1 位を飾った。その時の実行効率は 93%に達し、大規模システムでありながら高い効率を達成している。また、電力効率を競う Green 500[2]においても同時に 6 位の性能を達成した。電力効率 1, 2 位の汎用機 BlueGene/Q は別格であるが、京コンピュータはその他 GPU を用いたシステム群とほぼ互角である。GPU では実アプリケーションの実行効率は京コンピュータより低いものが多いと考えられるので、実質的には BlueGene/Q に次いで 2 位の性能であると考えられる。これらの優れた実行効率・電力効率は高いメモリ・ネットワーク帯域、水冷の採用など高性能計算(High Performance Computing, HPC)専用設計の賜物である。今後の大規模システムでは電力性能が大きな制約になることが明らかであり、京コンピュータはその面でも重要な一歩である。

京コンピュータは 8 万個以上のプロセッサを並列

につないだ大規模並列システムシステムである。さらに、1 個のプロセッサは 8 個のコアを持つので、コアの数としては 64 万以上となる。さらに、1 個のコアは 4 個の積和演算器を持つので、積和演算器の数としては 256 万個以上となり、その効率的な利用には非常に大きな並列度が必要である。

京コンピュータを特徴づける技術は、①HPC 向けの機能を持つプロセッサ“SPARC64 VIIIfx[3]”②高速で HPC 向けの機能を持つ 6 次元 Mesh-Torus ネットワーク“Tofu[4]”である。これらにより、柔軟で有効性の高い並列実行を可能としている。例えば、分子動力学計算のカーネルの実行では、50%以上の実行効率が得られている[5]。

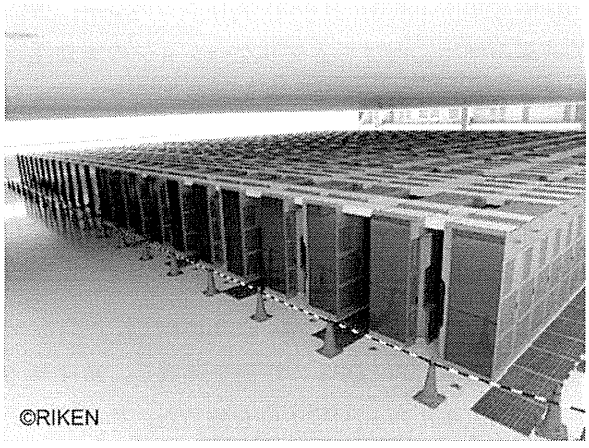


図 1. 京コンピュータの写真。(理研次世代スーパーコンピュータ開発実施本部提供)

分子シミュレーションによる薬剤スクリーニング

我々のグループの目標は、京コンピュータなどのスーパーコンピュータを活用した、大規模計算による分子設計の実現である。長期的には、酵素の設計や、複数のタンパク質が組み合わさってできる構造などの大きな機能分子の設計を目指しているが、現在はまず低分子やペプチド設計など比較的小さな分子をターゲットにした計算手法の検討を進めている。ここでは、ドッキング計算と分子シミュレーションを結合した高精度なスクリーニング手法の研究について紹介する。

タンパク質等の生体高分子と低分子の結合親和性は、阻害剤の設計において最も重要なパラメータの一つであり、これまで多くの計算手法の開発が行われてきた。特に近年では、多くのタンパク質・タンパク質−基質複合体の立体構造が実験的に求められ、Protein Data Bank[6]を初めとする公共データベースに蓄積されてきた。その結果、タンパク質の構造をベースにした薬剤設計“Structure-Based Drug Design(SBDD)”手法が盛んになった。SBDD にも、実験を中心した方法・計算を中心にした方法と様々であるが、タンパク質の構造が決まればあとは計算で合理的な設計ができれば理想的である。そのため計算による SBDD 手法の開発が古くから進められてきた。

その中でも分子ドッキングは、現在最も標準的に用いられている手法である[7][8][9]。一般的な分子ドッキング手法では、タンパク質を剛体として扱い、そのポケット部位に低分子をフレキシブルに当てはめ、親和性をスコア関数に基づき評価する。しかし、実際のタンパク質の構造は生体環境下で揺らいでおり、また基質結合に伴い大きな構造変化を起こす場合もある。従って、タンパク質によっては、分子ドッキングによるスクリーニングが十分に機能しないことも少なくない。これに対し分子動力学(MD)シミュレーションにより、タンパク質−基質の系全体に対しシミュレーションを行えば、こうしたタンパク質の柔軟性や溶媒の効果を陽に取り込むことができる。但し、MD シミュレーションはドッキングに対し遙かに計算量が多い。それ故、分子設計などの大量計算が必要なものに対する MD シミュレーションの適用は限定的であった。しかし京コンピュータなどの大規模な計算資源の登場により、シミュレーションに基づく分子設計が現実的な手法となりつつある。

我々のグループでは、分子ドッキングに基づく結合構造探索と分子シミュレーションに基づく結合自由エネルギー評価を組み合わせた手法開発を進めている[10]。分子ドッキング手法では、遺伝的アルゴリズムなどの手法を用いて低分子の配位空間を高速に探索することができる。実際に、実験的に得られている複合体構造を高い精度で予測することが可能になっている。しかし前述のように、異なる低分子間での結合親和性の強さを比較するには、分子ドッキングのスコア関数は不安定である。

一方、分子シミュレーションで低分子をタンパク質に結合させる過程を一からシミュレーションすると、シミュレーション時間がマイクローミリ秒オーダーになってしまうと考えられ、現実的でない。また実験的な結合構造は限られた基質-タンパク質にしかないので利用不可能である。そこで、分子ドッキングによって得られた複数の結合構造を基に分子シミュレーションを行い、結合親和性を評価する手法を開発した。

この手法の概念図を図2に示す。まず化合物ライブラリからケモインフォマティクス・Ligand-Based Drug Designの手法を用いて、数万-数十万程度の候補化合物を選ぶ。その後分子ドッキングを用いて、数百ー数千程度まで候補を絞り込むと同時に、複合体構造の候補を得る。この候補構造を出発点に分子動力学シミュレーションを行い、結合自由エネルギーを評価する。分子動力学シミュレーションによる結合自由エネルギー評価にはいくつかの手法があるが、ここでは比較的計算量の少ないMM/PB-SA(Molecular Mechanics/Poisson-Boltzmann

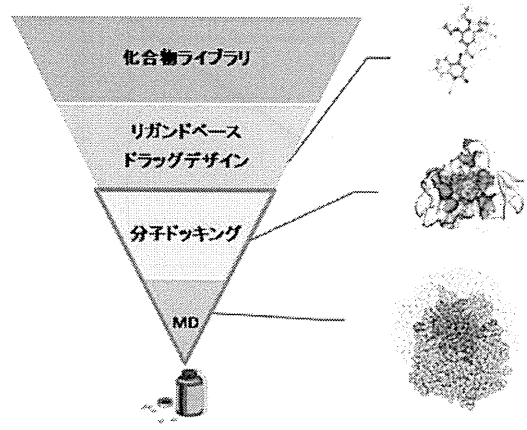


図2. 段階的なスクリーニングにより数百万の化合物ライブラリから数十ー数百の候補化合物を絞り込む過程。最後の分子ドッキングと分子動力学シミュレーションを組み合わせた薬剤スクリーニング手法が開発した部分。分子ドッキングにより結合構造を生成し、それに対し分子シミュレーションで結合親和性を高精度に求める。

SurfaceArea)法[11]で評価を行っている。ここでは、各複合体構造に対し約1ナノ秒のシミュレーションを行った。こうした大規模な計算は、後述の分子動力学シミュレーション専用計算機MDGRAPE-3の利用により可能になった。

数種類のタンパク質(Trypsin, HIV-1 protease, Acetylcholine esterase)に対しこの手法を適用した結果を図3に示す。図は濃縮率を示しており、曲線が左上に行くほど性能が高い。分子シミュレーションの適用により、より高精度な評価が安定して行えることがわかった。

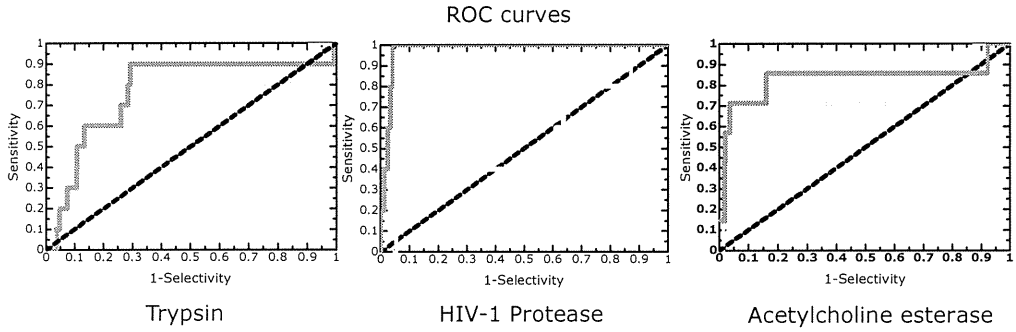


図 3. Trypsin, HIV-1 Protease, Acetylcholine esterase に対する Receiver Operating Characterisitic(ROC)曲線。赤線が本手法による結果、水色が分子ドッキングによる結果、斜め直線はランダムスクリーニングの時の理論線である。横軸が化合物の結合能の計算結果の順位、縦軸が実験的に結合することが既知の化合物がその順位までに含まれている率を示す。左上に行くほど性能が高い。

更に計算量を投入すると、より高精度な結合親和性の評価が可能なことも示されている。富士通の藤谷博士(現東大)らは、Bennett 's Acceptance Ratio に基づく自由エネルギー評価法 MP-CAFEE 法[12]を開発し、大規模計算によって結合親和性の実験値を絶対値で再現できることを示した。これには前述の MM/PB-SA 法と比して数百倍の計算時間が必要になる。また、QM/MM や FMO などの量子化学計算手法との比較についても検討を進めている[13]他、マルチドメインタンパク質の設計等も行っている[14,15]。

MDシミュレーション専用計算機

京コンピュータを用いて MD シミュレーションを行うとき、どの程度の計算ができるだろうか？分子動力学シミュレーションの計算量は、おおよそ 1 ステップ・1 粒子あたり 5×10^4 演算である。典型的な生体分子の系を 10^5 原子とすると、1 ステップの計算量は 5×10^9=5G 演算となる。今実効性能 1TFLOPS の計算機があるとすると、1 ステップにかかる実時間は 5GFLOP/1TFLOPS=5 ミリ秒となる。1 ステップのシミュレーション時間を 2 フェムト秒とすると、このときには一日に 35 ナノ秒の計算ができることになる。これは現在手が届く性能になりつつある。

しかし実効性能 1PFLOPS の計算機ではちょっと大変なことになってしまう。このとき、同規模の計算を行うと 1 ステップにかかる実時間は 5 マイクロ秒となり、一日に 34 マイクロ秒分の計算ができることになる。しかし、実際には京コンピュータやその他の汎用並列計算機では一回の通信を行うだけで 1 マイクロ秒かかってしまうので、この性能は実現不可能である。京コンピュータで実現できるのは、累積で一日数十マイクロ秒、例えば 100 ナノ秒のシミュレーションを 300 本ということになるであろう。あるいは数千万原子の系を一日 100 ナノ秒程度計算することもできるかもしれない。前述したような創薬シミュレーションでは、多くのシミュレーションを実行する必要があるのでこれでも十分である。例えば 1 個の化合物の計算に 1 ナノ秒の計算を行う場合には、1 日に 3 万化合物を扱えることになる。またより精度を向上させ、1 個あたり 1 マイクロ秒の計算を行う場合でも、1 日に 30 個の化合物を扱うことができる。ただ、タンパク質の構造緩和をさせたい時には、一本の長いシミュレーションができることが望ましい。例えば基質結合時の大きな構造変化を追いかけたいような場合である。

我々のグループでは、MD シミュレーションを加速するための専用計算機 MDGRAPE を開発してきた。MDGRAPE は、基本的には MD 計算で一番コストが高い非結合力(クーロン力・分子間力)の計算を、専用のパイプラインを用いて加速する装置である。2006 年に完成した MDGRAPE-3（図 4）は、名目ピーク性能 1PFLOPS を達成した世界最初の計算機である[16]。この計算機を用いて、実効性能 185TFLOPS

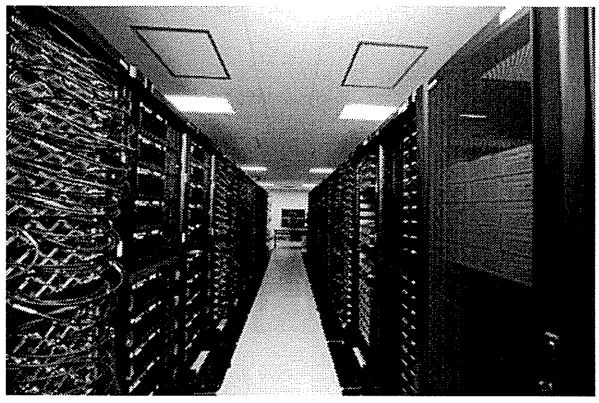


図 4. 分子動力学シミュレーション専用計算機 MDGRAPE-3 の写真。専用 LSI 約 4,800 個からなり、理論ピーク性能 1PFLOPS の性能を持つ。

の計算を達成し、2006 年の Gordon Bell 賞(Peak Performance, Honorable Mention)を受賞している[17]。しかし、MDGRAPE-3 では、力の計算の加速装置をホスト計算機(PC クラスタ)に接続する方式をとっていたため、通信性能は PC クラスタの性能に律速されていた。そのため、多量の計算を並列に実行するか、あるいは大規模系をシミュレーションするのには向いていたが、小規模系の長時間計算には向いていなかった。前述の分子シミュレーションに基づく創薬スクリーニングは、多量の並列計算の例であり、MDGRAPE-3 を用いて可能になった計算結果である。

近年、米国 D. E. Shaw Research のグループが、MD シミュレーション専用計算機 Anton を開発し[18]、ミリ秒スケールの MD シミュレーションを実現した[19]。Anton は、MDGRAPE のような力の計算パイプラインに、汎用プロセッサ・ネットワークを統合したシステム LSI を用いており、通信が非常に高速である。2 万 3 千原子の系に対し 1 ステップを 24 マイクロ秒という非常に短い時間で計算することができる。重要なのは、生命科学で扱いたいような比較的小さな系に対しては、Anton では京コンピュータでもできない計算ができるということである。

我々もこれに対し、新世代の分子シミュレーション専用計算機MDGRAPE-4の開発を進めているところである。MDGRAPE-4 に向けて、高度に専用化された力の計算パイプラインと、メモリ・汎用プロセッサ・高速なネットワークを統合したシステム LSI を開発している。MDGRAPE-4 は、換算ピーク性能で 3PFLOPS の性能を 4 ラックで実現し、10 万原子系の系を一日に 10 マイクロ秒以上実行することを目標に開発を進めている。

生命科学における高性能計算の今後

次世代シーケンサや一分子計測技術などに代表される生命定量化技術の進展に伴い、生命科学における精密なモデル化・シミュレーションの重要性が増大しつつあり、計算科学によって生命科学全体に大きな変化をもたらすことができる可能性もあるだろう。生命科学には、エクサ(10^{18})フロップススケールの計算資源を十分に使い切るだけの計算需要があるが、それを単体で使い切れるかというと難しい。前述のように、実行効率に優れた京コンピュータでも例えば長時間の MD シミュレーションを行うには適しておらず、この延長上にあるシステムではなおさらである。また多くの生命科学応用では、小ー中並列規模の計算を大量に行いたいような需要が多い。高性能計算機の「性能」がその演算性能だけでは語れなくなった現在においては、需要の実態に合わせた開発が必須になっている。これらの事情は生命科学だけに留まらないと考えられ、今後の計算機開発の重要な課題である。

謝辞

MD シミュレーションを用いたスクリーニングについては、二木紀行博士（現 Fixstars)、末永敦博士(現産総研)との共同研究である。また MDGRAPE-3 の開発は、成見哲教授(現電通大)、大野洋介博士(理研)との共同研究である。ここに感謝したい。

参考文献

[1] http://www.top500.org/

[2] http://www.green500.org/

[3] Maruyama, T. et al., *IEEE Micro***30**(2), 30-40 (2010).

[4] Ajima, Y. et al., IEEE Computer **42**(11), 36-40 (2009).
[5] 大野洋介他、私信
[6] http://www.pdb.org/
[7] Ewing, T. et al., *J. Comput. Aided Mol. Des.***15**, 411 (2001)
[8] Friesner, R. A. et al., *J. Med. Chem.***47**, 1739 (2004)
[9] Jones, G. et al., *J. Mol. Biol.***267**, 727 (1997).
[10] Okimoto, N. et al.,*PLoS Comput Biol*, **5**, e1000528 (2009)
[11] Srinivasan, J. et al., *J. Am. Chem. Soc.***120**(37), 9401-9409 (1998).
[12] Fujitani, H. et al., Phys. Rev. E**79**(2), 021914 (2009).
[13] Watanabe, H. et al., *CBI Journal***10**, 32-45 (2010).
[14] Usui, K. et al., *Protein Sci.***18**(5), 960-969 (2009).
[15] Ito, F. et al.,*Biomaterials***31**(1), 58-66 (2010).
[16] Taiji, M. et al, *Proc. Supercomputing 2003*, in CD-ROM (2003).
[17] Narumi, T. et al, *Proc. Supercomputing 2006*, in USB memory (2006).
[18] Shaw, D. E. et al, *Comm. ACM***51**(7), 91-97 (2008).
[19] Klepeis, J. L. et al., *Curr. Opin. Sturuct. Biol.***19**(2), 120-127 (2009).

たいじ　まこと MAKOTO TAIJI
理化学研究所　生命システム研究センター　生命モデリングコア　コア長
兼　計算科学研究機構　プロセッサ研究チーム　チームリーダー

バックグラウンドは物理ですが、生命科学分野での計算の加速をライフワークとして取り組んでいます。

おきもと　のりあき NORIAKI OKIMOTO
理化学研究所　生命システム研究センター　生命モデリングコア　上級研究員

バックグラウンドは薬学で、分子シミュレーションによってタンパク質と薬剤間の相互作用の原理を理解することを目指して研究しております。

連絡先　〒650-0047 神戸市中央区港島南町 7-1-26
電話　　078-940-5648（泰地）/078-940-5627（沖本）

THE JOURNAL OF PHYSICAL CHEMISTRY

ARTICLE

pubs.acs.org/JPCB

Free-Energy Landscapes of Protein Domain Movements upon Ligand Binding

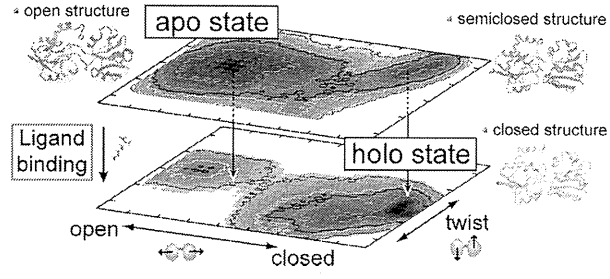
Hiroko X. Kondo,[†,‡,§] Noriaki Okimoto,[*,†,§] Gentaro Morimoto,[†,§] and Makoto Taiji[†,‡,§]

[†]High-Performance Molecular Simulation Team, Computational Systems Biology Research Group, RIKEN Advanced Science Institute, 61-1 Ono-cho, Tsurumi, Yokohama, Japan 230-0046

[‡]Department of Computational Biology, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Japan 277-8561

Supporting Information

ABSTRACT: The conformation and functions of proteins are closely linked, and many proteins undergo conformational changes upon ligand binding. The X-ray crystallographic studies have revealed conformational differences in proteins between the liganded and unliganded states. Currently, the conformational transitions that originate in the ligand binding are explained on the basis of two representative models, the induced-fit and preexisting equilibrium dynamics models. However, the actual dynamics of the proteins remain ambiguous. Though these two models are the extreme ones, it is important to understand the difference between these two, particularly in structural biology and medicinal chemistry studies. Here, we clarified the difference in the mechanisms responsible for the conformational changes induced in two proteins upon ligand binding by examining computationally determined free-energy profiles of the apo- and holoproteins. The lysine/arginine/ornithine-binding protein and maltose/maltodextrin-binding protein were chosen as the target proteins, and the energy profiles were generated by a molecular simulation approach. Our results revealed that fluctuations in the apo state and protein−ligand interactions both play important roles in conformational transition, and the mechanism is highly influenced by the fluctuations in the apo state, which are unique to a particular structure.



■ INTRODUCTION

Under physiological conditions, proteins fluctuate around their native state and often undergo conformational changes upon interaction with ligands. To examine functional conformational changes in proteins, it is crucial to understand thermal fluctuations and the induction of conformational change upon ligand binding. In general, two representative models are used to describe conformational changes in proteins, the induced-fit model[1] and the preexisting equilibrium dynamics model[2,3] (Figure 1). In the induced-fit model, an apoprotein fluctuates around its open conformation and is structurally altered in such a manner that it changes to the closed conformation upon ligand binding (1−2−4 in Figure 1). In contrast, the preexisting equilibrium dynamics model assumes that an apoprotein fluctuates between the open and closed conformations and is stabilized in the latter conformation by thermodynamic equilibrium after ligand binding (1−3−4 in Figure 1). Recently, the functional conformational change of the protein has been studied by structural,[4,5] kinetic,[6] and computational[7,8] approaches, and these support that the mechanism of the conformational changes consists of both models, induced-fit and preexisting equilibrium dynamics models,[9] not either of them. The focal point in this subject is the contribution ratio and sequence for combination of the two models. Molecular dynamics (MD) simulation is so helpful for understanding the functional conformational change in depth because this is a powerful method for studying the structure and dynamics of proteins at the atomic level.

Here, we analyzed these conformational changes in proteins from the viewpoint of the free-energy landscape. This approach was used because the difference between the two models is essentially due to differences in the free-energy profiles or the potential of mean force (PMF) obtained from MD simulations in the apo and holo states. Two proteins were selected as the model proteins, the lysine/arginine/ornithine-binding protein (LAOBP) (Figure 2a) and the maltose/maltodextrin-binding protein (MBP) (Figure 2b). These proteins belong to the superfamily of bacterial periplasmic binding proteins, and both of them contain two domains that sandwich the ligands. They mediate the passage of nutrient molecules from the periplasm into the cytoplasm of gram-negative bacteria[10] and are also involved in chemotaxis.[11] Some experimental studies have suggested that these proteins differ in their conformational transition mechanisms; LAOBP is characterized by the preexisting equilibrium dynamics[12] and MBP by the induced-fit model.[13−15] Structural

Received: December 15, 2010
Revised: April 18, 2011
Published: May 24, 2011

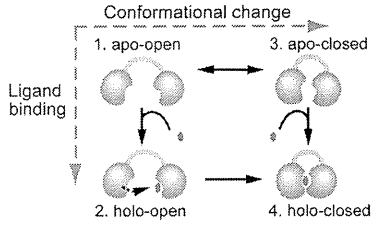


Figure 1. Two representative models for protein conformational change, the induced-fit and preexisting equilibrium dynamics models. The horizontal and vertical axes represent conformational change and ligand binding, respectively. The pathways 1→2→3 and 1→3→4 represent the induced-fit and preexisting equilibrium dynamics models, respectively.

analyses by NMR imaging and fluorescent probe-based sensing has suggested that most apo-MBP exist in the open conformation,[14,15] and the free-energy increases in proportion to the degree of domain closure.[13] A strong correlation was established between the degree of domain closure and the stability of the protein structure. These support the mechanism of conformational transition in MBP based on the induced-fit model. However, there is no direct evidence to suggest that the conformational transition mechanism in LAOBP follows the preexisting equilibrium dynamics model, except for the structural data of the apo- and holoproteins.

We determined the energy profiles of the two proteins by using the umbrella sampling technique[16] with MD simulations. Our results revealed clear differences between the free-energy landscapes of the two proteins and suggested that fluctuations in the apoproteins and protein−ligand interactions both play important roles in conformational transition. Here, we report detailed models for the conformational transition in proteins, which have been interpolated from the energy landscape.

■ MATERIALS AND METHODS

Procedure of This Study. First, we performed standard MD simulations to analyze the dynamics of LAOBP (Figure 2a) and MBP (Figure 2b) in the following four conformations, the unliganded state initiated with the open and closed conformations (unliganded-open and unliganded-closed, respectively) and the liganded state initiated with the open and closed conformations (liganded-open and liganded-closed, respectively). Next, a principal component analysis (PCA)[17,18] was applied to the trajectories obtained from the above simulations to determine the reaction coordinate for the umbrella sampling. Eigenvectors with large eigenvalues can represent the large-scale and essential conformational movements of the protein. Finally, we performed umbrella sampling simulations for both proteins in the apo and holo states and determined their free-energy profiles.

Preparation of Initial Structures for Simulations. The X-ray structures of the open-apo form (PDB entry: 2LAO[12] for LAOBP and 1OMP[19] for MBP) and closed-holo form (PDB entry: 1LST[12] for LAOBP and 1ANF[20] for MBP) were used as the initial structures of unliganded-open and liganded-closed, respectively. The initial structures of liganded-open and unliganded-closed were constructed by binding the ligand to the initial structure of unliganded-open and removing the ligand from the initial structure of liganded-closed. Lysine was used as the ligand for LAOBP, and maltose was used for MBP. The initial

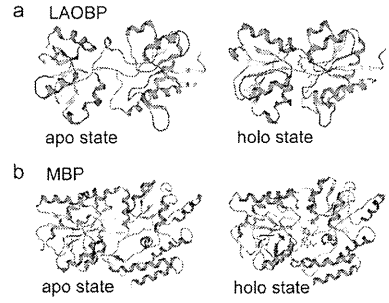


Figure 2. Crystal structures of the model proteins in the apo and holo states. (a) Front view of the crystal structures of lysine/arginine/ornithine-binding protein (LAOBP) in the presence and absence of the ligand (right and left, respectively). (b) Front view of the X-ray structures of maltose/maltodextrin-binding protein (MBP) in the presence and absence of the ligand (right and left, respectively). The residues from 170 to 175 (loop region) and those from 95 to 101 are colored blue. These are the important residues for domain movement for MBP, and the details are explained in Figure 10.

structure of liganded-open MBP was designed using the molecular docking technique with the Genetic Optimization for Ligand Docking (GOLD) program (version 3.1);[21,22] on the other hand, that of LAOBP was built by molecular modeling based on the position of the ligand in the holo crystal structure.

MD Simulation Methods. All MD simulations were performed using a modified AMBER program (version 8.0)[23] designed for MD machines on personal computers (host computers) equipped with a special-purpose computer for MD simulations, MDGRAPE-3.[24−26] All of the initial structures were solvated in a rectangular box containing TIP3P water molecules[27] under periodic boundary conditions. The minimum distance between a protein atom and the nearest image atom was 15 Å. We adopted force field ff03[28] for the protein. The SHAKE algorithm[29] was applied to bonds involving hydrogen atoms, considering an integration time step of 1.0 fs. The computations of non-bonded forces and energies of Coulomb and van der Waals interactions for MD simulations were accelerated by MDGRAPE-3, and the other calculations were performed by the host computer. In this study, Coulomb interactions were assessed by applying the particle mesh Ewald (PME) method. The real-space part for PME was calculated by the MDGRAPE-3, while the other parts for PME was calculated by the host computer. To optimize the balance of these calculation times, a cutoff distance of 14 Å for Coulomb and van der Waals interactions was used.

After energy minimization for the each system, the system was gradually heated to 310 K during the first 50 ps period for heating, and an additional 20 ns simulation was performed for data collection. The temperature and pressure were kept constant at 310 K and 1 atm, respectively, by the method of Berendsen's algorithm with a coupling time of 1.0 ps.[31] In the standard MD simulation process, to investigate the essential domain movements of LAOBP and MBP, we performed multiple simulations for each initial system using different initial atomic velocities. Hereafter, we chose the typical trajectories for the respective systems.

Furthermore, additional MD simulations for LAOBP were performed. From the results of the standard MD simulations and umbrella sampling of LAOBP, we found that the key conformation appeared in the apo state, which is called the semiclosed conformation (see Results section). To investigate the structural

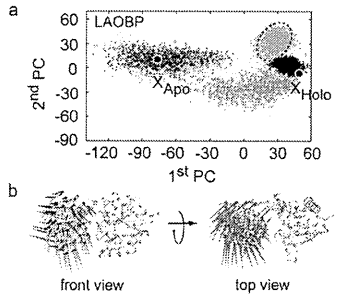


Figure 3. Fluctuation and conformational transition of LAOBP in PCA space. (a) Projection of the MD trajectories of LAOBP on the first and second eigenvectors obtained by PCA. X_{Apo} and X_{Holo} represent the open (apo) and closed (holo) crystal structures, respectively. The light-blue and blue points indicate the two kinds of unliganded-open trajectories, and the black points represent the liganded-closed trajectory. The unliganded-closed and liganded-open trajectories are shown by yellow and green points, respectively. We designated the region around the circle drawn with a broken line as the semiclosed conformation. (b) The fluctuation modes of LAOBP are indicated by the arrows along the protein backbone. The structure on the left side is seen from the same direction as that in Figure 2a. The red arrow represents the first PC, and the yellow arrow denotes the second PC. The front view of the protein structure is shown on the left, and the top view is shown on the right.

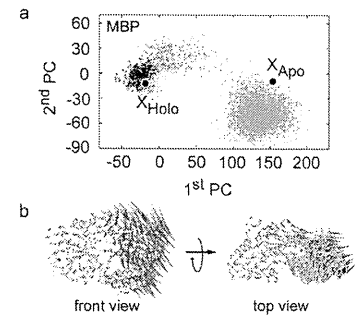


Figure 4. Fluctuation and conformational transition of MBP in PCA space. (a) Projection of the MD trajectories of MBP on the first two eigenvectors obtained by PCA. The light-blue points indicate the unliganded-open trajectory, the green points indicate the liganded-open trajectory, yellow points indicate the unliganded-closed trajectory, and the black points indicate the liganded-closed trajectory. (b) Fluctuation modes of MBP. The structure on the left side is seen from the same direction as that in Figure 2b. The red arrow represents the first PC, and the yellow arrow denotes the second PC. The views of the front and top of the structure are shown on the left and right, respectively.

properties, multiple simulations of the liganded and unliganded-semiclosed conformation were performed. At this time, the liganded-semiclosed conformation was built by referring to the position of the ligand in the holo crystal structure. The conditions of the MD simulations are same as the above simulations.

Umbrella Sampling. Umbrella sampling is a method used to generate the free-energy profile along specified reaction coordinates. After removal of the rotational and translational motions, trajectories of four kinds (unliganded-open, liganded-open, unliganded-closed, and liganded-closed) were subjected to PCA using a positional covariance matrix of Cα atoms. On the basis of the examination of the contribution ratio of each eigenvalue obtained from PCA (see Supporting Information), we selected the eigenvectors with large contribution ratios, namely, the first and second PCs for LAOBP and the first PC for MBP, and used them as the reaction coordinates for the umbrella sampling simulation. The unliganded-closed trajectory was not used in the analysis of MBP because the experimental studies showed that the liganded-closed MBP interacts with another membrane protein, followed by the opening of the domains and release of the ligand, which suggests the necessity of an external force for the transition from the closed to the open conformation;[32] our simulation supports the experimental studies because the conformational changes were hardly observed in the MD trajectory.

In each sampling simulation, we performed energy minimizations on the initial structures with constraint of the concerned Cα atoms and a 50 ps heating simulation followed by an equilibration run with the biasing potential. We then extracted the biased distribution of the first (and second) PC values. The biasing potential (V_i) is given by

$$V_i = \frac{1}{2}k\sum_j (\Phi - \Phi_{ij})^2 \qquad \Phi = \mathbf{r}\cdot\mathbf{X}_j \tag{1}$$

where k is the force constant, X_j is the jth eigenvector, and Φ_{ij} is the reference value of the jth PC value for ith simulation. The other parameters were identical to those used for the free MD simulations.

For the umbrella sampling simulations of LAOBP, the Cα atoms in all of the residues were restrained with force constants of 0.2−0.8 kcal/(mol·Å^2) for the first and second PC values. For MBP, the Cα atoms in residues 6−365 were restrained only for the first PC values in the same way as the simulation of LAOBP. The force constants and length of simulations were determined such that the distribution of PC values at each reference value sufficiently overlapped with that of the adjacent reference value. In this study, we performed more than 900 ps MD simulations to equilibrate the systems at the respective reference values, and the last 500 ps trajectories for the respective simulations were collected for PMF calculation. The total numbers of the umbrella sampling simulations were 892 for LAOBP and 69 for MBP.

The biased distributions obtained by the umbrella sampling simulations were rendered unbiased by using the weighted histogram analysis method (WHAM).[33] According to the WHAM equation, the unbiased distribution $p(\Phi)$ is given as the sum of the unbiased probability distribution obtained from the ith simulation $p_i^{\text{unbiased}}(\Phi)$

$$p(\Phi) = \frac{\sum_i n_i(\Phi)}{\sum_i N_i \exp\{(f_i - V_i(\Phi))/k_BT\}} \tag{2}$$

$$f_i = -k_BT \ln\{\int d\Phi\, p_i^{\text{unbiased}}(\Phi) \exp(-V_i(\Phi)/k_BT)\} \tag{3}$$

where $n_i(\Phi)$ is the number of states at Φ in the ith simulation and N_i represents the total number of time slices in the ith simulation.

The PMF $\Delta E(\Phi)$ is calculated from the following equation using the unbiased distribution function $p(\Phi)$ obtained from WHAM

$$\Delta E(\Phi) = E(\Phi) - E(\Phi_0) = -k_BT \ln \frac{p(\Phi)}{p(\Phi_0)} \tag{4}$$

where Φ_0 is the reference first or second PC value.

dx.doi.org/10.1021/jp111902t |J. Phys. Chem. B 2011, 115, 7629–7636

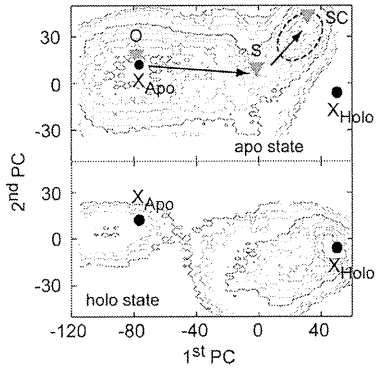


Figure 5. Free-energy landscapes of LAOBP. Free-energy landscape of the apo (upper) and holo (lower) states. As the first PC value increases, the degree of domain closure increases. X_{Apo} and X_{Holo} represent the open (apo) and closed (holo) crystal structures, respectively, and the region around the circle drawn with a broken line represents the semiclosed conformation. Contour lines are drawn at every 2 kcal/mol. O, S, and SC indicate the open, saddle, and semiclosed conformations, respectively.

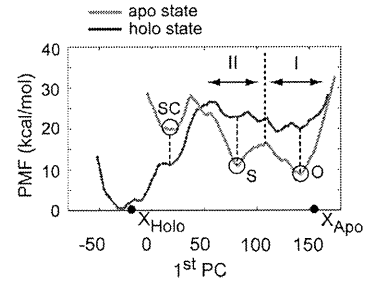


Figure 6. Free-energy profiles of MBP. The free-energy profiles of MBP in the apo (red line) and holo (blue line) states. In the profiles, the energy difference between the lowest PMF value in the apo state and that of the holo state correspond to the free-energy difference (8.4 kcal/mol) derived from the experimental data. As the first PC value decreases, the degree of domain closure increases, which is reverse to the case of LAOBP. The apo and holo crystal structures are designated as X_{Apo} and X_{Holo}, respectively. The open circles with O, S, and SC represent the first (open), second (semiopen), and third (semiclosed) energy minima, respectively. In the holo state, the difference in the free energies of the global energy minimum and the second local minimum was 19.0 kcal/mol, and the energy barrier from the open to closed conformations was 8.0 kcal/mol. For explanation of the structural transitions upon ligand binding, conformations in the apo state were classified by the regions I and II.

■ RESULTS

Conformational Fluctuations of Apo- and Holoproteins. Prior to detailed analysis by the free-energy approach, we performed preparatory MD simulations of LAOBP and MBP to check the conformational dynamics of apo- and holoproteins and to investigate the effect of ligand binding. Then, we applied PCA to the coordinates of the Cα atoms to determine the reaction coordinates of conformational change. PCA can extract significant fluctuation modes of the proteins and helps in classifying the conformations clearly.[18] The coordinates were sampled from the four kinds of trajectories (unliganded-open, liganded-open, unliganded-closed, and liganded-closed) obtained from these MD simulations. In the PCA, we used five trajectories for LAOBP, two kinds of the unliganded-open trajectories, the liganded-closed, the unliganded-closed, and the liganded-open trajectories. For MBP, three trajectories, the unliganded-open, liganded-closed, and liganded-open trajectories, were used. The unliganded-closed trajectory was not used in the analysis of MBP (see the Material and Methods section).

Figures 3a and 4a show the projection of the MD trajectories of LAOBP and MBP, respectively, onto the first and second eigenvectors. Each vector clearly shows that the fluctuation mode of the first PC is the open−closed motion and that of the second PC is the twisting motion in both proteins (Figures 3b and 4b). In LAOBP, the contribution ratios of the first and second principal components were 69 and 18%, respectively. The projection indicated that the MD trajectories spread widely, covering the experimental open and closed conformations. The liganded-closed trajectory clustered around the experimental closed conformation and the liganded-open trajectory gradually moved from the open conformation to the closed one. The unliganded-open trajectories widened along the direction of the first PC, and one of these almost reached the closed conformation but not the experimental one (Figure 3a, light-blue points). The conformations in this cluster were termed "semiclosed". This result indicates that the transition from the open to closed conformation that occurs in the presence of the ligand is different, especially in the second PC direction, namely, in the twisting motion of the domains from the transition that occurs in the absence of the ligand.

The contribution ratios of the first two principal components of MBP were 72 and 9%, respectively. The liganded-closed and unliganded-open trajectories clustered around the experimental closed and open conformations, respectively, although the unliganded-open trajectories spread along the direction of the second PC, suggesting that the two domains of the protein in the unliganded-open simulation are slightly twisted in comparison with the X-ray open structure (Figure 4a). This conformational difference from the crystallographic study would be caused by the solvent because the NMR study showed apoproteins that contain slightly twisted fluctuations (1.3 ± 0.6° for closure and 2.4 ± 1.9° for twist) in the solvent.[14] The liganded-open trajectory moved from the open conformation to the closed one along the direction of the first PC. The hydrogen bond network between the protein and ligand had already formed at an early stage of the liganded-open simulation. Therefore, it is believed that such interactions induce conformational transition.

Free-Energy Profiles for Domain Movements in LAOBP. We generated the free-energy profiles of the apo and holo states of LAOBP using the umbrella sampling method. In this study, we chose the first and second eigenvectors obtained from PCA as the reaction coordinates. As mentioned in the previous section, the first and second PCs represent the open−closed and twisted motions of the two domains, respectively.

Figure 5 shows the two-dimensional free-energy profile of LAOBP in the apo state on a plane defined by the two reaction coordinates. Two energy minima were found in the profile; the global minimum almost corresponded to the crystal open structure of the apoprotein, and the other minimum corresponded to the semiclosed conformation. The difference in the free energy between the global minimum and the other local minimum was estimated to be 2.0 kcal/mol, and the energy barrier from the former to the latter was approximately 9.7 kcal/mol. However, the PMF value of the closed conformation (X_{Holo} in Figure 5)

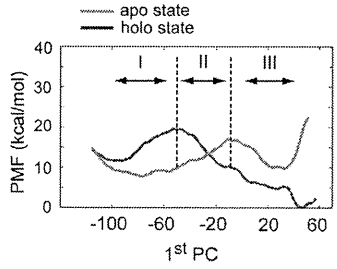


Figure 7. The energy profiles along the first PC of LAOBP in the apo and holo states. In these profiles, the energy difference between the lowest PMF value in the apo state and that of the holo state corresponds to the free-energy difference (7.6 kcal/mol) derived from the experimental data. For explanation of the structural transitions upon ligand binding, conformations in the apo state were classified by the regions I, II, and III.

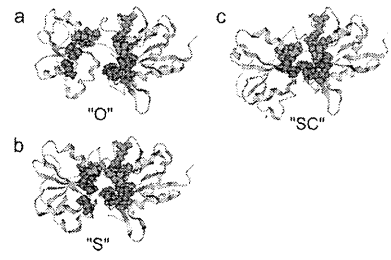


Figure 8. Protein structures in the open (a), saddle (b), and semiclosed (c) conformations with the surface residues of the pocket region. These correspond to the typical structures marked O, S, and SC in Figure 5. The surface residues are represented by blue spheres, and important residues are highlighted in red.

was quite high, exceeding 30 kcal/mol. Thus, the energy analyses clearly indicate that the apoprotein can fluctuate between the open and semiclosed conformations but cannot form the closed conformation.

Two energy minima were also identified in the free-energy profile of LAOBP in the holo state (Figure 5). The global minimum almost corresponded to the crystal closed structure of the holoprotein, and the other minimum corresponded to the experimental open structure. The difference in the free energy of the global minimum and the other local minimum was 11.4 kcal/mol, and the energy barrier from the open to closed conformation was 8.3 kcal/mol. These results revealed that most holoproteins are in the closed structure.

Free-Energy Profiles for Domain Movement in MBP. In the case of MBP, only the first PC was adopted as the reaction coordinate because the contribution ratio of the first PC (72%) was large enough to describe the conformational change, and the computational cost for the simulation of MBP was much higher than that of LAOBP due to the system size of the MD simulation.

Three energy minima were found in the free-energy profile of MBP in the apo state along the first PC reaction coordinate (Figure 6, red line). The global minimum almost corresponded to the crystal open structure. In the basin around the second lowest minimum, termed the "semiopen" conformation, the two domains came closer than those in the conformation of the global minimum. Another basin, called the "semiclosed" conformation, was found to be nearer in the first PC value to the experimental closed structure than to the experimental open one. There is a relatively large energy difference between this and the open structure (11.0 kcal/mol), and the probability of existence for the configurations in this conformation is about 10^{-8} times smaller than that of the global minima. This indicates that most of the apoprotein molecules are in equilibrium between the open and semiopen conformations but cannot change to the closed conformation.

The two basins were detected in the free-energy profile of the holoprotein (Figure 6, blue line). In the one basin, there is the global minimum that almost corresponded to the experimental closed structure of the holoprotein (Figure 6). The other one is broad and rugged and includes the open and semiopen conformations (marked O and S, respectively, in Figure 6). The most stable conformation in the holoprotein is the closed one (near X_{holo} in Figure 6), as in the case of LAOBP.

■ DISCUSSION

In the present study, we investigated the conformational and dynamic properties by analyzing MD trajectories of LAOBP and MBP and then calculated the free-energy profiles of these two proteins. The results indicated that the conformational properties are so different, especially in apo state, although both of them are strongly dominated by the open−closed and twisting motions of two domains. Furthermore, the energy profiles are closely related to these conformational properties. Here, we discuss the detailed mechanisms of the conformational transition, key structures, and ligand recognition.

Mechanisms of Conformational Transition in LAOBP. The energy profiles of apo- and holoproteins provide important information for understanding the mechanism of conformational transition in LAOBP. In this study, we calculated the binding free energy for ligand binding to the open conformation using the experimental dissociation constant for lysine. Using a binding free energy of 7.6 kcal/mol, the energy profiles of apo- and holoproteins were related along the first PC, as shown in Figure 7. On the basis of these diagrams, we propose three pathways of structural transition with ligand binding. In the first pathway, the ligand binds to the open conformation, which induces an open-to-closed conformational transition in the holo state (Figure 7, I). The second pathway is similar to the first, except that ligand binding occurs when the unliganded protein is in intermediate conformations between the open and semiclosed conformations (Figure 7, II). In this pathway, no energy barrier is required for the conformational transition. The third pathway requires ligand binding preceded by a transition from the open to the semiclosed conformation in the apo state (Figure 7, III, and see next subsection). If these three pathways are interpreted from the viewpoint of traditional models on protein conformational change, the first mechanism corresponds to the induced-fit model, and the second and third are based on a combination of the preexisting equilibrium dynamics and induced-fit models.

Semiclosed Conformation in Apo-LAOBP. Because the conformations in the semiclosed conformation have not been observed experimentally, we analyzed the roles of ligand binding and domain movement from a structural viewpoint. The free-energy profile of LAOBP suggests that the semiclosed conformation is generated from the open one via the saddle conformation. This conformational transition pathway is indicated by arrows in Figure 5, and the typical structure in each conformation (O, S, and SC) is shown in Figure 8a−c. As shown in these figures, the open conformation can allow domain movement because there is

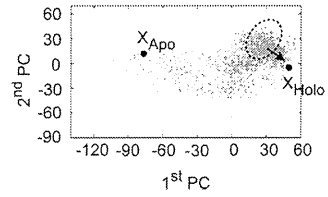


Figure 9. Projection of the additional MD trajectories of LAOBP on the first two eigenvectors obtained by PCA. The gray and magenta points indicate the trajectories of unliganded and liganded LAOBP starting from the semiclosed structures, respectively. The apo and holo crystal structures are designated as X_{Apo} and X_{Holo}, respectively. The region around the circle drawn with a broken line represents the semiclosed conformation. The arrow shows the conformational transition observed in the trajectory colored magenta.

enough space between the two domains (Figure 8a). However, in the saddle state, such a space vanishes, and the two domains are in contact with each other, indicating that the domains cannot close any longer along the direction of the first PC (Figure 8b). As shown in Figure 8c (the semiclosed conformation), the twist of the two domains and the reorientation of the side chains prevent the collision of the residues in the pocket and enable the domains to semiclose. This suggests that not only does the apoprotein adopt the fluctuation mode in the direction of the first PC, but it also moves along the direction of the second PC because the twist of the domains to escape collision is needed for domain closure. In comparison with the free-energy profile of the holoprotein, the transition pathway from the open to closed conformation in the holo state differed from the pathway from the open to the semiclosed conformation in the apo state (Figure 5). The conformational change in the holo state is dominated by the first PC rather than the second PC. This difference in the energy profiles of the apo and holo states is due to the effect of ligand binding.

In order to examine whether the semiclosed conformation changes into the open or closed conformation, we also performed additional simulations starting from the semiclosed conformation in the presence and absence of the ligand (liganded-semiclosed and unliganded-semiclosed). The results showed that conformational transition from the semiclosed to the closed conformations occurred during the MD simulation of liganded-semiclosed. In the unliganded-semiclosed simulation, conformational transition was also observed between the open and semiclosed structures (Figure 9). These results suggest that the semiclosed conformation is stable in the equilibrium and that the conformational change to the closed conformation occurs if a ligand binds to the semiclosed protein.

Mechanisms of Conformational Transition in MBP. Similarly to the case of LAOBP, the energy profiles of apo- and holoproteins of MBP were roughly connected along the first PC with the experimental binding free energy of maltose (8.4 kcal/mol)[9] (Figure 6). The energy diagram suggests two pathways for the transition from the open to the closed conformation upon ligand binding. Similar to LAOBP, the difference between these two pathways lies in the conformation that ligand binding occurs. The pathway in which the ligand binds to the protein in the open conformation (Figure 6, I) corresponds to the typical induced-fit model, whereas that in which the ligand binds to the protein in the semiopen conformation (Figure 6, II)

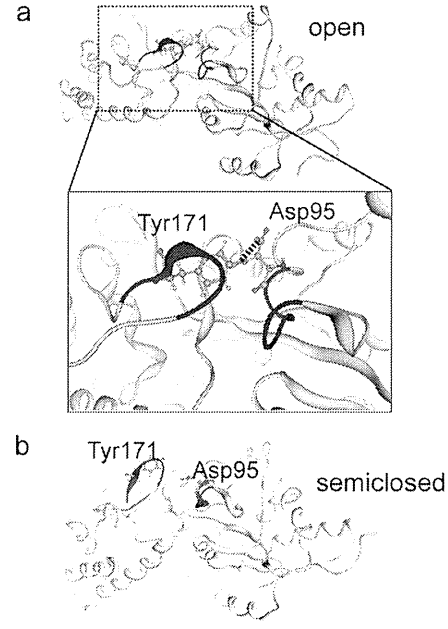


Figure 10. Comparison of the structure and location of the loop region where the salt bridge is formed in the open or semiclosed structures of MBP. (a,b) The residues from 170 to 175 (loop region) and those from 95 to 101 are colored blue. Residues Asp95 and Tyr171 are represented in the form of a ball-and-stick model. The magnification of the salt bridge region is shown on the bottom of (a). The salt bridge is formed in this region in the open conformation in (a) but is not formed in the semiclosed conformation in (b).

involves the preexisting equilibrium dynamics model within the framework of the induced-fit model.

Experimental structural analyses have provided conformational information on the MBP apoprotein. NMR studies by Millet et al.[13] and Evenäs et al.[14] as well as an X-ray crystallographic study[19] proposed that the apoprotein adopted the open conformation. On the other hand, NMR analysis by Tang et al. suggested that the apoprotein was in an equilibrium between a major open species (95%) and a minor partially closed one (5%).[34] Although the results of these conformational studies are interpreted in slightly different ways, their conclusions are the same, that is, the open conformation is the most stable conformation for the apo state. Our results agreed with the experimental data presented above, although the partially closed structure detected by the aforementioned NMR study[34] did not agree with the semiclosed or semiopen conformations obtained in our study. The NMR structure was somewhat strained because the hinge region of the closed conformation differs from that of the holo-closed structure. Furthermore, our energy profile of the apoprotein also supports the experimental result obtained by Millet and co-workers[13] that the free energy increases in proportion to the degree of closure between the two domains. In particular, the calculated increase in the free-energy values from the open to the semiclosed conformation in the apo state shows a similar tendency in qualitative aspect as compared with the results of the experimental study in which the Ile329Trp/Ala96Trp mutant was used. Unlike LAOBP, the formation of hydrogen bonds between the protein and its ligand was observed at an early stage of conformational change in the

liganded-open MD trajectory of MBP. This suggests that the induction of conformational changes upon ligand binding is dominant for MBP.

Structural Basis of Stability in the Open Conformation in Apo-MBP. We analyzed the conformations from the MD simulations in order to clarify the key conformational factor for the stability of open and semiopen MBP. Comparison of the conformations obtained from MD simulations revealed that the salt bridges between the loop in the hinge region (Tyr171 or Gly174) and the C-terminal domain (Asp95, Asn100, or Gly101) were maintained in most of the open and semiopen conformations (Figure 10a), but these interactions were almost lost in the more closed conformation due to domain movement (Figure 10b). These conformational analysis results suggest that salt bridge interactions play a role in the interactions that occur on the hook between two domains and largely contribute to the conformational stability of the apoproteins. The fastening and unfastening of the hook almost correlates with the free-energy profile obtained in this study. This is also supported by the results of the NMR study, which show changes in the backbone stability upon ligand binding.[14] This suggests that the conformational transition from the semiopen to semiclosed accompanies the collapse of these salt bridges; therefore, this causes the high-energy barrier. Interestingly, such interactions between the two domains were not observed in LAOBP.

Relationship between the Conformational Transition Mechanisms and Ligand Recognition. It is experimentally known that LAOBP binds three types of substrate molecules (lysine, arginine, and ornithine). The size of these molecules is relatively smaller than that of the lumen of the binding cavity. On the basis of the above experimental information and our simulations, we think that the fluctuation between the open and semiclosed conformations in the apo state is essential for the formation of protein−ligand interactions.

On the other hand, MBP can bind to various maltodextrins that have numerous polar functional groups and a wide range of sizes, including nonphysiological ligands. It has been shown experimentally that MBP can distinguish physiological ligands from nonphysiological ones based on differences in their binding modes, and these nonphysiological ligands cannot induce protein activity.[35] Recently, several structural studies on the β-cyclodextrin-bound state of MBP have been reported.[36−38] An X-ray crystallographic study revealed that MBP with β-cyclodextrin, a nonphysiological ligand, adopts the open conformation,[36] and the researches on NMR imaging suggested that MBP fluctuates around the partly closed conformations.[37,38] One of the NMR conformations is very similar to the semiopen conformation observed in our study (main chain RMSD value from the experimental structure[38] (PDB entry: 2H25): 1.6 Å). The above experimental results and our computational results suggest that the difference between the structures determined by X-ray crystallography and NMR imaging is caused by the fluctuation of proteins between the open and semiopen conformations rather than by the interactions of β-cyclodextrin and MBP. Therefore, we propose that, unlike nonphysiological ligands, physiological ones can induce the conformational transition, and this leads protein activity.

■ CONCLUSIONS

On the basis of the free-energy calculations of proteins in the apo and holo states, we analyzed the mechanism underlying conformational changes in proteins upon ligand binding. The results indicated that the mechanisms in LAOBP and MBP differ in terms of the free-energy profiles and structural features. The remarkable difference between the two proteins is in the energy profiles of the apo state rather than in those of the holo state.

Moreover, structural analyses suggested that the structural difference between LAOBP and MBP strongly influences the dynamical and energetic aspects of conformational changes upon ligand binding (Figure 10), indicating that the magnitude and properties of protein fluctuation are regulated by the three-dimensional structure of the protein itself. On the basis of these results, we concluded that the conformational change upon ligand binding can be explained by a combination of the preexisting equilibrium dynamics and induced-fit models but not solely by either of the two models; the mechanism of LAOBP is mainly governed by the preexisting equilibrium dynamics model, while that of MBP is principally guided by the induced-fit model.

The present study showed the strong relationship between protein structure and fluctuations in the apo state. Moreover, the results confirmed the observation that fluctuations in the apo state are very important for conformational transition upon ligand binding. We believe that our approach is effective not only for understanding conformational change upon ligand binding but also for drug discovery. In particular, the energy profiles of apo and holo states can help to predict the conformational transition from the apo state to the holo state, which is important in the design and discovery of more diverse and novel drugs.

■ ASSOCIATED CONTENT

Supporting Information. Plots of the eigenvalues obtained from PCA. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Address: Laboratory for Computational Molecular Design, RIKEN Quantitative Biology Center, #R303, 7-1-26 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo, Japan 650-0047. Phone: +81-78-940-5627. Fax: +81-78-304-4956. E-mail: okimoto@gsc.riken.jp.

Present Addresses

§The affiliation and address have been changed as follows. Laboratory for Computational Molecular Design, RIKEN Quantitative Biology Center, #R303, 7-1-26 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047.

■ ACKNOWLEDGMENT

This study was supported by the RIKEN Advanced Science Institute (ASI) and the Protein 3000 Project and Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT). We are thankful for the computational resources of the RIKEN Integrated Cluster of Clusters (RICC). We also thank Dr. Oscar Millet and Dr. Lewis E. Kay for the gifts of the mutant structures of MBP and Dr. Futatsugi for his advice.

■ REFERENCES

(1) Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98.
(2) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88.
(3) Tsai, C. J.; Kumar, S.; Ma, B.; Nussinov, R. *Protein Sci.* **1999**, *8*, 1181.
(4) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. *Nature* **2007**, *450*, 838.
(5) Xu, Y.; Colletier, J. P.; Jiang, H.; Silman, I.; Sussman, J. L.; Weik, M. *Protein Sci.* **2008**, *17*, 601.
(6) Hammes, G. G.; Chang, Y. C.; Oas, T. G. *Proc. Natl. Acad. Sci. U. S.A.* **2009**, *106*, 13737.
(7) Okazaki, K.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11182.
(8) Miller, D. W.; Dill, K. A. *Protein Sci.* **1997**, *6*, 2166.
(9) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789.
(10) Davidson, A. L. *J. Bacteriol.* **2002**, *184*, 1225.
(11) Manson, M. D. *Adv. Microb. Physiol.* **1992**, *33*, 277.
(12) Oh, B. H.; Pandit, J.; Kang, C. H.; Nikaido, K.; Gokcen, S.; Ames, G. F.; Kim, S. H. *J. Biol. Chem.* **1993**, *268*, 11348.
(13) Millet, O.; Hudson, R. P.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12700.
(14) Evenas, J.; Tugarinov, V.; Skrynnikov, N. R.; Goto, N. K.; Muhandiram, R.; Kay, L. E. *J. Mol. Biol.* **2001**, *309*, 961.
(15) Medintz, I. L.; Deschamps, J. R. *Curr. Opin. Biotechnol.* **2006**, *17*, 17.
(16) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187.
(17) Kitao, A.; Hirata, F.; Go, N. *Chem. Phys.* **1991**, *158*, 447.
(18) Kitao, A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164.
(19) Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A. *Biochemistry* **1992**, *31*, 10657.
(20) Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E. *Structure* **1997**, *5*, 997.
(21) Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, 43.
(22) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
(23) Case, D. A.; Darden, T. A.; Cheatham, T. E. r.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman., D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER*, version 8.0; University of California: San Francisco, CA, 2004.
(24) Narumi, T.; Ohno, Y.; Okimoto, N.; Koishi, T.; Suenaga, A.; Futatsugi, N.; Yanai, R.; Himeno, R.; Fujikawa, S.; Ikei, M.; Taiji, M. A 185 Tflops simulation of amyloid-forming peptides from Yeast Prion Sup35 with the special-purpose computer System MD-GRAPE3. *Proc. Supercomputing* [CD-ROM] **2006**.
(25) Taiji, M. MDGRAPE-3 chip: a 165 Gflops application specific LSI for molecular dynamics simulations. *Proc. Hot Chips* [CD-ROM] **2004**, *16*.
(26) Taiji, M.; Narumi, T.; Ohno, Y.; Futatsugi, N.; Suenaga, A.; Takada, N.; Konagaya, A. Protein explorer: a petaflops special-purpose computer system or molecular dynamics simulations. *Proc. Supercomputing* [CD-ROM] **2003**.
(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(28) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999.
(29) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(30) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(31) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(32) Chen, J.; Lu, G.; Lin, J.; Davidson, A. L.; Quiocho, F. A. *Mol. Cell* **2003**, *12*, 651.
(33) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011.
(34) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078.
(35) Hall, J. A.; Ganesan, A. K.; Chen, J.; Nikaido, H. *J. Biol. Chem.* **1997**, *272*, 17615.
(36) Sharff, A. J.; Rodseth, L. E.; Quiocho, F. A. *Biochemistry* **1993**, *32*, 10553.
(37) Skrynnikov, N. R.; Goto, N. K.; Yang, D.; Choy, W. Y.; Tolman, J. R.; Mueller, G. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *295*, 1265.
(38) Xu, Y.; Zheng, Y.; Fan, J. S.; Yang, D. *Nat. Methods* **2006**, *3*, 931.