2015年11月30日 バイオスーパーコンピューティング神戸2015

生体分子系における非平衡過程

神戸大学大学院システム情報学研究科 計算科学専攻計算生物学講座 田中 成典





リガンドドッキングやタンパク質フォールディングにおける長時間の構造緩和ダイナミクス
平衡状態(ボルツマン分布)にどのように接近していくか?その数学的構造を議論

 ◆ 量子−古典対応に基づき、ランジュバン・ダイ ナミクスを拡散量子モンテカルロ法で解析

 ・ポテンシャル場に応じて、系のエンタルピー、
 エントロピー、自由エネルギーがどのように時
 間変化するか?

Structure-Based Drug Design (SBDD)

タンパク質とリガンド(薬剤候補)の立体構造に基づく薬剤設計

近年の技術的進歩

- コンピュータの計算能力の急速な進化
- タンパク質の立体構造の解明
 - X線立体構造解析やNMRといった測定技術の進歩
 - インフラの整備(SPring-8、京)



タンパク質ーリガンド系の分子シミュレー ションによる定量的な設計が可能となった



計算機シミュレーションを用いたSBDDのメリット

- タンパク質とリガンドの相互作用(自由)エネルギーを計算できる
- 立体的な結合構造が利用あるいは予測できるため、詳細な分子間の相互作用を解析できる



ドッキング・シミュレーションの例(Uehara et al., 2015)



Fig. 4 Molecular structures of the binding pocket of neprilysin and BIR. The blue colored pose is the crystal structure, the green colored pose is a representative structure of cluster-1, and the red colored pose is a representative structure of cluster-2. Artificial Bee Colony (ABC) アルゴリズム













分子力学・動力学シミュレーション

◆ Molecular Mechanics (MM)
 ◆ 力場(force field)を利用 → エネルギー極小状態の探
 ◆ Molecular Dynamics (MD)

- ◆ 生体分子系を構成する全ての原子に対して古典 的なNewtonの運動方程式を解く
- $F_i = m_i a_i$ $\rightarrow \mu s \sim ms の世界 \land$

リガンドドッキング・シミュレーション → 有効に結合する阻害剤の探索

「与えられたポテンシャル場での最適化問題」と位置づけることができる。



Perspective: Markov models for long-timescale biomolecular dynamics

C. R. Schwantes,^{1,a)} R. T. McGibbon,^{1,a)} and V. S. Pande^{1,2,b)}

¹Department of Chemistry, Stanford University, Stanford, California 94305, USA ²Department of Computer Science, Stanford University, Stanford, California 94305, USA; Department of Structural Biology, Stanford University, Stanford, California 94305, USA; and Biophysics Program, Stanford University, Stanford, California 94305, USA

(Received 8 July 2014; accepted 27 August 2014; published online 4 September 2014)

Molecular dynamics simulations have the potential to provide atomic-level detail and insight to important questions in chemical physics that cannot be observed in typical experiments. However, simply generating a long trajectory is insufficient, as researchers must be able to transform the data in a simulation trajectory into specific scientific insights. Although this analysis step has often been taken for granted, it deserves further attention as large-scale simulations become increasingly routine. In this perspective, we discuss the application of Markov models to the analysis of large-scale biomolecular simulations. We draw attention to recent improvements in the construction of these models as well as several important open issues. In addition, we highlight recent theoretical advances that pave the way for a new generation of models of molecular kinetics. © 2014 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4895044]

 $u_{t+\tau}(\mathbf{x}) = T(\tau) \circ u_t(\mathbf{x})$ Transfer operator for probability density

$$u_{t+n\tau}(\mathbf{x}) = 1 + \sum_{i=2}^{\infty} \exp\left(-\frac{n\tau}{t_i}\right) \langle \psi_i, u_t \rangle_{\mu} \psi_i(\mathbf{x}) \psi_i(\mathbf{x}$$

Decomposition into eigenstates

CrossMark

Variational Approach to Molecular Kinetics

Feliks Nüske,[#] Bettina G. Keller,^{*,#} Guillermo Pérez-Hernández, Antonia S. J. S. Mey, and Frank Noé*

Department for Mathematics and Computer Science, Freie Universität Berlin, 14195 Berlin, Germany

ABSTRACT: The eigenvalues and eigenvectors of the molecular dynamics propagator (or transfer operator) contain the essential information about the molecular thermodynamics and kinetics. This includes the stationary distribution, the metastable states, and state-to-state transition rates. Here, we present a variational approach for computing these dominant eigenvalues and eigenvectors. This approach is analogous to the variational approach used for computing stationary states in quantum mechanics. A corresponding method of linear variation is formulated. It is shown that the matrices needed for the linear variation method are correlation matrices that can be estimated from simple MD simulations for a given basis set. The method proposed here is thus to first define a basis set able to capture the relevant conformational transitions, then compute the respective correlation matrices, and then to compute their dominant eigenvalues and eigenvectors, thus obtaining the key ingredients of the slow kinetics.



dx.doi.org/10.1021/ct40091561J. Chem. Theory Comput. 2014, 10, 1739-1752

 $|\rho_{t+\tau}(x)\rangle = \mathcal{P}(\tau)|\rho_t(x)\rangle$ Operator that propagates the probability density ρ

Analogy to variational principle in quantum chemistry (LCAO-MO)

$$=\sum_{i=1}^{n} a_i |\varphi_i\rangle$$
 Expansion by basis functions

$$C_{ij} = \langle \varphi_i | \mathcal{P}(\tau) | \varphi_j \rangle_{\pi^{-1}} \qquad S_{ij} = \langle \varphi_i | \varphi_j \rangle_{\pi^{-1}}$$

n

 $|f\rangle$

Eigenvalue problem (cf., tICA, Relaxation mode analysis)

 $Ca = \lambda Sa$



Overdamped Langevin equation

$$\dot{\boldsymbol{r}} = -\frac{D}{k_B T} \nabla U(\boldsymbol{r}) + \boldsymbol{\eta}(t)$$

U(*r*) : potential field *D*: diffusion coefficient

$$\langle \eta_i(t)\eta_j(t')\rangle = 2D\delta_{ij}\delta(t-t')$$

Fokker-Planck-Smoluchowski方程式

$$\frac{\partial}{\partial t}P(\boldsymbol{r},t) = D\nabla \cdot \left[\nabla P(\boldsymbol{r},t) + \frac{1}{k_B T}\nabla U(\boldsymbol{r})P(\boldsymbol{r},t)\right]$$

 $P(\mathbf{r},t) \to Ce^{-U(\mathbf{r})/k_BT}$ 平衡状態:ボルツマン分布

(P. Faccioli *et al.*, Phys. Rev. Lett. 97 (2006) 108101.) (Villin headpiece)

Mapping to quantum mechanics

 $P(\mathbf{r},t) = e^{-U(\mathbf{r})/2k_BT}\psi(\mathbf{r},t)$ Classical-quantum correspondence

$$-\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = \hat{H}_{\text{eff}}\psi(\boldsymbol{r},t)$$

Imaginary-time Schrodinger equation

$$\hat{H}_{\mathrm{eff}} = -D\Delta + V_{\mathrm{eff}}(\boldsymbol{r})$$
 Diffusion + Branching

$$V_{\text{eff}}(\boldsymbol{r}) = \frac{D}{(2k_BT)^2} \left\{ \left[\nabla U(\boldsymbol{r}) \right]^2 - 2k_BT\Delta U(\boldsymbol{r}) \right\}$$

 $\hat{H}_{\text{eff}}\phi(\boldsymbol{r}) = E\phi(\boldsymbol{r})$ Time-independent Schrodinger eq.

$$\phi_0(\mathbf{r}) = C e^{-U(\mathbf{r})/2k_BT}$$
 Ground state with $E_0 = 0$

Path integral formulation

$$\widehat{G}(\boldsymbol{r}_{f}, t_{f} | \boldsymbol{r}_{i}, t_{i}) = \langle \boldsymbol{r}_{f} | e^{-\widehat{H}_{\text{eff}}(t_{f} - t_{i})} | \boldsymbol{r}_{i} \rangle = \int_{\boldsymbol{r}_{i}}^{\boldsymbol{r}_{f}} \mathcal{D}\left[\boldsymbol{r}(\tau)\right] e^{-S_{\text{eff}}[\boldsymbol{r}]}$$

$$S_{\text{eff}}\left[\boldsymbol{r}\right] = \int_{t_i}^{t_f} d\tau \left\{ \frac{\left[\dot{\boldsymbol{r}}(\tau)\right]^2}{4D} + V_{\text{eff}}\left[\boldsymbol{r}(\tau)\right] \right\} \quad \text{Effective action}$$

$$\tilde{P}(\boldsymbol{r}_f, t_f | \boldsymbol{r}_i, t_i) = e^{-\left[U(\boldsymbol{r}_f) - U(\boldsymbol{r}_i)\right]/2k_B T} G(\boldsymbol{r}_f, t_f | \boldsymbol{r}_i, t_i)$$

:分布関数 P(r,t) のpropagator

この非平衡過程をシミュレートできないか? → 拡散量子モンテカルロ(DMC)法



ボルツマン分布 → 自由エネルギー F₀

$$F_{0} - F(t)$$

$$= \int d\mathbf{r}U(\mathbf{r}) [P_{0}(\mathbf{r}) - P(\mathbf{r}, t)]$$

$$+ k_{B}T \left[\int d\mathbf{r}P_{0}(\mathbf{r}) \ln P_{0}(\mathbf{r}) - \int d\mathbf{r}P(\mathbf{r}, t) \ln P(\mathbf{r}, t) \right]$$

$$= \int d\mathbf{r}U(\mathbf{r}) [P_{0}(\mathbf{r}) - P(\mathbf{r}, t)]$$

$$+ k_{B}T \left\{ \int d\mathbf{r} [P_{0}(\mathbf{r}) - P(\mathbf{r}, t)] \ln P_{0}(\mathbf{r}) + \int d\mathbf{r}P(\mathbf{r}, t) \ln \frac{P_{0}(\mathbf{r})}{P(\mathbf{r}, t)} \right\}$$

$$= \int d\mathbf{r}U(\mathbf{r}) [P_{0}(\mathbf{r}) - P(\mathbf{r}, t)]$$

$$+ k_{B}T \left\{ \int d\mathbf{r} [P_{0}(\mathbf{r}) - P(\mathbf{r}, t)] \left[\ln C - \frac{U(\mathbf{r})}{k_{B}T} \right] + \int d\mathbf{r}P(\mathbf{r}, t) \ln \frac{P_{0}(\mathbf{r})}{P(\mathbf{r}, t)} \right\}$$

$$= k_{B}T \int d\mathbf{r}P(\mathbf{r}, t) \ln \frac{P_{0}(\mathbf{r})}{P(\mathbf{r}, t)}$$

$$F_{0} は最小自由エネルギーを与える$$

$$\ln x \le x - 1 \text{ for } x > 0 \qquad (等号は x = 1) \text{ Jb}$$

$$F_{0} - F(t) \le k_{B}T \int d\mathbf{r} P(\mathbf{r}, t) \left[\frac{P_{0}(\mathbf{r})}{P(\mathbf{r}, t)} - 1 \right] = 0$$

$$\overrightarrow{F}(t) \ge F_{0} \qquad (等号は P(\mathbf{r}, t) = P_{0}(\mathbf{r}))$$

結局、 拡散モンテカルロ法により、波動関数の時間発展を シミュレートすれば、その長時間極限として基底状態が 得られ、量子-古典マッピングにより、それが自由エネ ルギー最小の分布(ボルツマン分布)を与える。









 $k_{\rm B}T = 1.0, D = 1.0, \Delta t = 10^{-4}, M = 20,000$ walkers

x^(m)(0) = 2.0 (m = 1,2,…, M に対する波動関数の時間変化



1次元調和ポテンシャル(続き)



 $k_{\rm B} {\rm T}$ = 0.1, D = 1.0, Δt = 10⁻⁴ M = 20,000 walkers

 $x^{(m)}(0) = 2.0 \ (m = 1, 2, \dots, M)$ に対する波動関数の時間変化





エントロピー減少

シミュレーションの例2:1次元二重井戸ポテンシャル

 $U(x) = x^4 - 2x^2$



 $k_{\rm B}T = 1.0, D = 1.0, \Delta t = 10^{-4}, M = 20,000$ walkers

 $x^{(m)}(0) = -1.0 \ (m = 1, 2, \dots, M)$ に対する波動関数の時間変化







$$U(x) = x^{2} + e^{-2x^{2}} - 3e^{-2(x+1)^{2}} - 15e^{-20(x-1)^{2}}$$

 $k_{\rm B}T = 1.0, D = 1.0, \Delta t = 10^{-3}, M = 20,000$ walkers



U(x), V_{eff}(x), Φ₀(x)の形状



1次元非対称二重井戸ポテンシャル(続き)



エントロピーは最初増加、しばらく 安定した後、分布が x > 0 に移行 するのに伴い、減少する。

分布の x > 0 への移動は エンタルピーの減少が駆動 \rightarrow ordered state



遷移過程で、ポテンシャル エネルギー(エンタルピー) は大きくゆらぎながら、平均 としては減少







2次元非対称二重井戸ポテンシャル(続き)



Summary

- ◆ 等温古典確率系緩和ダイナミクス ⇔ 量子系基 底状態探索
- Langevin eq., Fokker-Planck eq. ⇔
 Schrodinger eq., Feynman path integral
 ⇔ Diffusion Monte Carlo
- ◆ U, S, F の時間的変動の追跡
- Entropy dominance vs. Enthalpy
 - ◆ 急峻なポテンシャルの谷間に落ちて行くとき、エントロピーの減少が見られる(explicitly illustrated)。
- ◆ 多自由度系への拡張も容易
- ◆ 有限温度のoptimizationの一般的枠組