ERGODIC PROBLEMS FOR REAL COMPLEX SYSTEMS IN CHEMICAL PHYSICS

TAMIKI KOMATSUZAKI[1,2], AKINORI BABA[1,2], SHINNOSUKE KAWAI[1], MIKITO TODA[3], JOHN E. STRAUB[4], and R. STEPHEN BERRY[5]

[1]*Molecule & Life Nonlinear Sciences Laboratory, Research Institute for Electronic Science, Hokkaido University, Kita 20 Nishi 10, Kita-ku, Sapporo 001-0020, Japan*
[2]*Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), Kawaguchi, Saitama 332-0012, Japan*
[3]*Department of Physics, Faculty of Science, Nara Women's University, Kitauoyahigashimachi, Nara 630-8506, Japan*
[4]*Department of Chemistry, Boston University, 590 Commonwealth Avenue, SCI 503, Boston, MA 02215, USA*
[5]*Department of Chemistry, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA*

CONTENTS

Advancing Theory for Kinetics and Dynamics of Complex, Many-Dimensional Systems: Clusters and Proteins, Advances in Chemical Physics, Volume 145, Edited by Tamiki Komatsuzaki, R. Stephen Berry, and David M. Leitner.

I. INTRODUCTION

How many variables or parameters are required to reveal the process of the evolution or the changes of the states in complex systems such as complex chemical networks or proteins? Consider a system of N degrees of freedom interacting with the surrounding environment of M degrees of freedom. If M is zero, the system is regarded as being isolated and usually described in a microcanonical ensemble of constant energy E. On the contrary, the case of M being infinity corresponds to condensed phase dynamics with dissipation and fluctuation arising from the surrounding environment, which is often characterized by constant temperature T and a distribution of atomic friction coefficients.

First, let us briefly review isolated reacting systems, that is, $M = 0$. The dynamic evolution takes place in the phase space of $2N$ dimensions at a constant energy E. In principle, one should be required to use $2N - 1$ independent variables to describe the events. However, as described in Chapter 4, in the well-known statistical reaction theories such as transition state theories, the rate of reaction can be formulated in terms of a substantially smaller number of parameters. For the case of condensed phase systems, the relative ratio of the number of parameters required to describe the reaction rate per the actual dimension of the system is far smaller in condensed phase than for the isolated system. For instance, Kramers theory characterizes the effects exerted by the environment through the temperature T, potential of mean force, random force, and friction. The rate of reaction can be again formulated in terms of a substantially smaller number of parameters such as barrier height, friction, and temperature with a chosen "reaction coordinate."

What is the fundamental assumption that enables us to substantially reduce the actual dimension of the system to represent the rate of complex chemical reactions? The key concept is (local) ergodicity and the resulting separation of timescales; that is, the characteristic timescale to attain ergodicity just within the reactant states is significantly shorter than the timescale of the reaction from the reactant to the

product states. As a result, the rate of reaction is found to be independent of the initial condition at the reactant state under constant temperature or energy. This is crucially relevant to the fundamental question of how complex, chemical and biological systems evolve or change their states in time.

A. Ergodicity

The best known definition of ergodicity often used in statistical mechanics is the property that the time average of a characteristic of an ergodic system is indistinguishable from the ensemble average for the distribution over all accessible points in the system's phase space. More precisely, the time average of an arbitrary function f, which is complex-valued in general, defined in a space (mathematically a smooth manifold M) in the phase space of the system, is indistinguishable almost everywhere on the M from the ensemble average over all accessible points on the M. Here, the propagation in time t obeys the behavior of a dynamical system, denoted here by U_t, which maps a point on the M uniquely (one-to-one) to a point on the M while preserving the measure on the M through U_t. Expressed as an equation, it is described as

$$\int_M f(x)dP(x) = \lim_{t\to\infty} \frac{1}{t} \int_0^t f(U_{t'}x(0))dt' \tag{1}$$

where $P(x)$ denotes a measure defined on the M that can be normalized (i.e., probability measure $\int_M dP(x) = 1$), x are continuous variables defined on the M, and $U_{t'}x(0)$ is the time propagation of $x(0)$ from 0 to t, corresponding to $x(t')$.

Let us exemplify this concept in terms of several systems: A two-dimensional dynamical system well studied in the context of ergodicity may be the "stadium billiard," almost a circle but with the two semicircles separated by parallel straight lines. Almost all trajectories in this enclosure pass through the entire interior of the enclosure; the exceptions are trajectories perpendicular to the straight sections and the one trajectory that passes between the centers of the two hemicircles.

The introduction of probability measure plays an essential role to establish the concept of ergodicity. It is because the integration in terms of the probability measure singles out such events of measure zero whenever the probability measure exists. For Hamiltonian systems, the probability measure is given by the phase space volume suitably normalized because the measure-preserving condition is guaranteed by the Liouville theorem. We discuss the case in which there is no such a measure to be normalized in Section II.

The second example is a two-dimensional torus where the ratio of the two frequencies ω_1 and ω_2 is irrational, that is, $\omega_1/\omega_2 \neq n/m$ (n, m: arbitrary positive integers). If the system satisfies this irrational condition, no trajectory can be closed on the torus (one calls such motions quasiregular) and every trajectory densely covers all the surface of the torus. Hence, we recognize the motion as being ergodic

on the torus. It should be noted that the question of ergodicity depends on which space M one considers. In the case of a torus defined by two invariants of motion, there exists no other independent invariant of motion on the torus to decompose the M into disjoint sets. This implies that the invariants of motion on *the* torus are regarded as global (and trivial) invariants of motion on that space because no invariant of motion exists to divide that space into disjoint sets. Note, however, that if one considers the M as the whole phase space of constant energy, the invariants of action surely prevent the system from wandering through accessible phase space almost everywhere. The system does not behave as ergodic in the whole phase space of constant energy.

The third example is a one-dimensional harmonic oscillator. One can regard this as a system composed of a particle with a finite angular velocity confined to a circle where the particle moves along a diameter of the circle, bouncing elastically each time it reaches the circular boundary. This integrable system also satisfies the condition of ergodicity; that is, irrespective of the initial condition on the circle, the system can cover all the accessible points and hence the time average of any function defined on the circle is equivalent to the ensemble average. Note that this is different from the case of two-dimensional torus whose ratio of the two frequencies is rational. In the latter case, depending on the initial condition on the torus, trajectories cover different regimes on the torus because of the difference in phase.

What type of property must dynamical systems possess in order to be ergodic? It has been proved that at least for a time evolution U_t defined on the phase space X that is one-to-one and preserves the probability measure P (e.g., Hamiltonian systems), if either $P(M) = 1$ (more in general $P(M \backslash X) = 0$) or $P(M) = 0$ holds for any subset M in X with $M = U_{(-t)}(M)$ (namely, if a set M is invariant under U_t in X and such an invariant set exists solely as global or empty), the system cannot have (nontrivial) invariants of motion to decompose the M into disjoint sets but can have only (trivial) global invariant of motion almost everywhere through the M. (Here, $M \backslash X$ denotes the subset of X that contains all the elements that do not belong to M.) Followed by several theorems, the resultant dynamics is known to satisfy the ergodic condition by Eq. (1) on the M [1]. The most important consequence here is that ergodicity implies neither the existence of chaos nor the loss of correlation in dynamics. (The more detailed discussions are given in Section II.)

The notation $U_{(-t)}(M)$ above should be understood in general as an inverse image of U_t, that is,

$$U_{(-t)}(M) \equiv U_t^{-1} M = \{x | U_t x \in M\} \tag{2}$$

If the time evolution of the system is invertible, that is, $U_{(-t)}x$ can be defined uniquely for each $x \in X$, it coincides with the backward propagation of M:

$$U_{(-t)}(M) = \{U_{(-t)}x | x \in M\} \tag{3}$$

However, if, for example, two different points x and y are mapped into the same point $z = U_t x = U_t y$, the image $U_{(-t)}z$ of the point z cannot be well defined while the inverse image of $\{z\}$ can be defined as a set $U_{(-t)}\{z\} = \{x, y\}$. In either case, $U_{(-t)}(M)$ is a measurable set for any measurable set M if the time evolution U_t is given by a measurable map.

B. Mixing

Most systems that interest us in the fields of chemistry and biology are more or less related to the property called "mixing," which results from the existence of chaos. The explanation of mixing is rather simple: suppose coffee in a cup; add milk to the coffee in a ratio of, for example, 70% coffee and 30% milk by volume. After stirring the (mixed) solution many times enough to mix them up, whenever one takes any arbitrary fraction of the solution, the ratio of the coffee and the milk one will find is 70% versus 30%. This situation is represented in general as follows: Suppose two arbitrary subsets A and B in the phase space X and the inverse image $U_{(-t)}B$ of the subset B. The mixing condition is formulated mathematically as

$$\lim_{t\to\infty} P(A \cap U_{(-t)}B) = P(A)P(B) \tag{4}$$

This equation means that the probability that an arbitrary point x in A will end up in B after t iterations (t is considered to be an effectively infinite number) [the left-hand side of Eq. (5)] is just the same as that of finding the B in the whole X and independent of the position of A and B in X [the right-hand side of Eq. (5)]:

$$\frac{\lim_{t\to\infty} P(A \cap U_{(-t)}B)}{P(A)} = P(B) \tag{5}$$

A less restrictive condition is called weak mixing, which states that the long-time average of the difference between $P(A \cap U_{(-t)}B)$ and $P(A)P(B)$ vanishes:

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \left| P(A \cap U_{(-t')}B) - P(A)P(B) \right| dt' = 0 \tag{6}$$

If a system is strong mixing, it satisfies the condition of weak mixing. The converse is not true. Intuitive interpretation of the weak mixing is that the mixing condition $P(A \cap U_{(-t)}B) \approx P(A)P(B)$ for very large t is satisfied for "most of the time" [1] since exceptional instances are wiped away by the process of averaging.

Any weak mixing transformation [satisfying Eq. (6)] directly results in ergodicity [2]. Assume that B is an invariant set with respect to the time propagation U_t, that is, $U_{(-t)}B = B$. Take A to be a complement to B so that $P(U_{(-t)}B \cap A) = 0$. We substitute this in the left-hand side of Eq. (6), and obtain

$$0 = \lim_{t\to\infty} \frac{1}{t} \int_0^t |P(A)P(B)|\, dt' = P(B)P(A) = (1 - P(A))P(A) \tag{7}$$

Thus, $P(A)$ must satisfy either $P(A) = 1$ or $P(A) = 0$. This means that weak mixing implies ergodicity (one can prove easily that strong mixing also does so). It should be noted that ergodicity does not necessarily imply mixing, or even weak mixing. As just introduced above, the typical system is an integrable Hamiltonian system such as a two-dimensional torus whose ratio of the two frequencies is irrational, or just the one-dimensional pendulum. The timescale for approaching the mixing state ("equilibrium state") or that of losing memory of the initial condition or correlation in dynamics in the phase space is approximately regarded as the inverse of Kolmogorov–Sinai entropy [3].

In the literature on chemical physics, relatively little attention has been paid to the conceptual difference between mixing and ergodicity. It may be because most systems of chemical and biological interest are expected to be inherently subject, to some extent, to both chaotic and mixing properties of nonlinear systems.

C. Multiplicity of Ergodicity in Complex Systems

Ergodicity is a property that can be verified only if one can examine both time and (phase) space averages. However, an interesting challenge arises if the system of interest has a rough, complicated potential surface. The reason is that the system may explore local regions thoroughly on short timescales yet require much longer times to escape from one, such local region, and move to another. If the potential surface has two or more relatively deep local minima that are separated by high or very narrow saddles, even if the system can, in principle, pass over those saddles, such passages can be relatively rare events, compared to the frequency of exploring all the places in the region of one of those local minima. Consequently, it is not unusual to find that a complex system can display two or more degrees of ergodicity. On a fairly short timescale, the system may exhibit only local ergodicity, but on a sufficiently long timescale, the system can explore its entire accessible space and be fully ergodic. If the landscape is sufficiently complex, there may be more than two or even more identifiable stages to the evolution of ergodicity.

An illustration of this behavior appears in small atomic clusters, particularly in the range of temperatures and pressures within which the cluster may exist as a "solid" or a "liquid," with the two phases in *dynamic* equilibrium, like two isomers. Under these conditions, one can see each phase-like form for some well-defined time interval, easily long enough for the internal vibrational modes to equilibrate, yet the system passes from one form to the other in some random fashion. If one tests for ergodicity using an ensemble and a single dynamical system on a long trajectory, one can probe for this property on a short or long timescale. If one looks on a very long timescale, one sees a single kind of behavior that involves exploring the entire accessible phase space, including the solid and liquid regions. If, on the other hand, one looks at a relatively short timescale with the probe (which we shall discuss shortly), then one sees two distinct kinds of behavior. One is ergodic

but only in the liquid region, and the other is ergodic in the solid-like region; the timescales on which one sees this kind of behavior are too brief for the system to be able to pass between solid and liquid. Presumably, the same kind of timescale separation holds for structural isomers that correspond to structures accessible to a molecule but only on a relatively long timescale.

The demonstration of this behavior appears in the *distributions* of sample values of Lyapunov exponents, the values of the exponential rates at which neighboring trajectories diverge. If these are obtained from long trajectories, then the distributions are unimodal, centered around the single most probable value. However, if the distributions are taken from shorter trajectories, then they are bimodal, with one maximum for the clusters in the liquid region and another for those in the solid region [13, 14] (see also Section III). The other illustrative example may be free energy landscape. In Section V, it is indicated that the morphological feature of the landscape depends on a timescale of observation. The longer the timescale, the more the number of detectable metastable states decreases, and the smoother the landscape implied by the observation.

D. The Ergodic Problem in Real Systems

The traditional concept of ergodic behavior is derived from mathematical analyses that, in turn, treat infinitely long pathways and arbitrarily large ensembles. Physical systems are finite and many of those of great interest now are small, and the timescales on which we may wish to observe them can be very brief indeed. Hence, it is appropriate to introduce heuristic analogues of the rigorous properties of ergodicity and chaos, based on the system in question satisfying some chosen criterion based on a finite, perhaps very long, time interval. If the system satisfies the chosen criterion, we may safely treat it as if it were truly chaotic or ergodic, within time intervals shorter than that of the criterion. (Sometimes this behavior has been called "cryptoergodicity" or "cryptochaos.")

In chemistry, ergodicity has been one common central property that one assumes in establishing several theories such as reaction rate theory, free energy landscape, and so on. One has also known many cases, for example, non-RRKM kinetics and the intramolecular vibrational energy redistribution (IVR) problem, that do not satisfy ergodicity. However, one has paid little attention to validating the concept of ergodicity in systems in which it is probably valid, and furthermore little has been done to explore new insights concerning the system's dynamics in terms of the concept. What we want to address here is what the appropriate tests or criteria should be, which enable us to use the concept of ergodicity, or, more precisely, to avoid invalidating the application of the concept of ergodicity in the problems of real systems in chemical physics. In that spirit, we are really asking for tests of cryptoergodicity, in the sense that we want to know when we can suppose

a system appears, in whatever ways are significant for our investigation, to be ergodic. We are concerned not with the rigorous mathematical property but with the observable behavior of the system. It is expected that plausible tests of this property, cryptoergodicity, depend on the size of system or the condition on our knowledge about the system (e.g., whether we can know the equation of motion of the system or we can solely monitor a physical quantity of the system).

The finiteness of timescale is even essential and rather inherent to a lot of phenomena of chemical interest. A typical illustrative example is a bimolecular reaction: Two molecules (reactants) collide with each other to form a metastable intermediate complex. After some time, the complex dissociates into a different set of two molecules (products). While the motion in the complex can be chaotic involving most of all degrees of freedom and subject to the issue of ergodicity, in the products and the reactants limits there are two separate molecules. Since the two molecules far apart cannot interact with each other, the system cannot be ergodic in all its dimensions before the formation and after the dissociation of the complex. The intermediate complex is the only form to be subject to ergodicity through the full dimension but the lifetime of the intermediate complex is finite. The finiteness and the value of lifetime in this case are determined by the system (not the problem of observation).

Another example is a system that exhibits transitions among multiple well regions. The degree of chaos and ergodicity can be different for different wells. They are subject to the competition between the strength of chaos and the residence time of each well, and depend on the extent that the system can attain ergodicity (or rather cryptoergodicity) in that well. This is essential for heterogeneity to emerge in establishing cryptoergodicity (we will discuss this aspect in more detail for proteins in Section IV). The more important question to be addressed is what we actually learn from the concept of ergodicity about complexity of systems such as the question of what the system actually feels under a thermally fluctuating environment.

Here, it must be noted that the introduction of the term "cryptoergodicity" is not only due to the limitation of observation but also inherent to the problems themselves whenever they invoke the change in their states. We will also come back to this issue from the viewpoint of open phase space in Section II.B. This chapter addresses the ergodic problems relevant to real complex systems from small-body systems such as atomic clusters to proteins. Here, we start with an overview, the historical background of the concept of ergodicity, and the implication of the concept in the sense of statistical mechanics in Section II. Then in Section III, we show how the *local* Lyapunov exponent distribution can unveil the ergodic property of inert gas clusters. This system may be regarded as representative of small-body systems, in contrast to systems with complex internal constraints, for example, proteins, but cluster dynamics is rich enough to start to discuss because clusters exhibit phase transition-like behavior even with small, finite number of degrees of

freedom. However, when the system of interest becomes much larger than those, although one can still compute the local Lyapunov exponent distribution, the device of the local Lyapunov exponent distribution becomes almost impossible to use. In Sections IV, we turn to the so-called ergodic measure developed for elucidating the rate of self-averaging of physical observables and characterizing the timescale of quasithermalization, and show the existence of heterogeneous multiple timescales to attain ergodicity, depending on the moiety of a protein. In Section V, we review our recent studies on the other measure to evaluate attainability and multiplicity of ergodicity in complex protein systems when one cannot access the underlying equation of motion of the system but just a time series of certain physical variables of the system such as interdye distance. We present our recent progress in deepening our understanding of the free energy landscape at single-molecule level.

II. ORIGIN OF STATISTICAL REACTION THEORY REVISITED

The most fundamental assumption of the *statistical* reaction theory is the separation of timescales; that is, the characteristic timescale for establishing equilibrium in the potential well is *assumed* to be much shorter than that for the reaction to take place. The chemical reaction proceeds while local equilibrium is maintained in the potential well. This makes it possible to apply the methods of the equilibrium statistical physics to chemical reactions. However, the recent development of theoretical and experimental studies on reaction processes reveals the necessity of going beyond the conventional statistical reaction theory [21, 22].

We consider the foundation and limitations of the statistical physics, especially its relevance for understanding reaction processes involving biomolecules. In the context of reactions, the following two features become crucial. First, reaction processes take place in open phase space regions in the sense that trajectories flow into and out of them, while the phase space is closed in the conventional statistical physics. Second, the system is inhomogeneous for reaction processes involving biomolecules, while it consists of identical particles in the traditional statistical physics. We will explain why these two features present serious issues concerning the foundation of statistical reaction theory.

In this section, we start our discussion with a brief review of the *traditional* ideas on the dynamical origin of the statistical physics. Then, we go on to argue why the above two features of the reaction processes necessitate serious reconsideration on the foundation of statistical physics. Finally, we discuss recent development of the dynamical theory concerning the statistical physics such as Sinai–Ruelle–Bowen (SRB) measure and infinite ergodic theory, and present possibility of these new ideas in the study of reaction processes.

A. Traditional Ideas of the Dynamical Origin of Statistical Physics

In the study on the mechanism of approaching equilibrium, Boltzmann introduced the model, now called the Boltzmann equation [24], using the one-particle distribution $P(p, q)$ defined on the phase space (p, q), where p and q are the momentum and the coordinate of the one particle, respectively. Under the assumption of *molecular chaos*, that is, the motions of molecules are supposed to be completely uncorrelated, he showed H-theorem, that is, the quantity $H \equiv \int P(p, q) \log P(p, q) dp\, dq$ monotonically decreases in time, indicating irreversible approach to equilibrium.

His derivation of the H-theorem met the objection from Loschmidt, who asserted that the H-theorem contradicts the time-reversal symmetry of Newton's equation of motion. In order to defend his derivation of H-theorem, Boltzmann introduced the *ergodic hypothesis* implying that H-theorem is relevant for a dominant part of the phase space.

Their argument triggered the development of the theory of *ergodicity*, which is now well established in the sense of mathematics. Here, we give a brief explanation of the theory of *ergodicity*. The following discussion is *not* limited to the Hamiltonian systems, that is, the subjects of the *traditional* studies of the statistical physics. It is also applicable to dissipative systems since dissipative systems can have invariant measures, which are *not* the phase space volume. Thus, the argument can be applied to reactions involving biomolecules surrounded by an environment, in addition to unimolecular reactions of isolated systems.

We follow the traditional argument for the foundation of statistical physics. Several good references exist both for mathematicians [35, 36] and for nonmathematicians [26, 27]. In statistical physics, the idea of ergodicity plays the role that corresponds to that of *the law of large numbers* in the probability theory [31]. In traditional statistical physics, observed values of the physical quantity are generally assumed to be equivalent to time averages over the infinite time interval. In order to apply the equilibrium statistical methods, these time averages should be independent of initial conditions.

In order to justify the above idea of ergodicity in statistical physics from the standpoint of the dynamical system theory, the first thing to ask is whether time averages over infinite time interval exist or not. To approach this, we state *Birkhoff's individual ergodicity theorem*. The theorem guarantees existence of time averages over the infinite time interval for physical quantities of a certain class.

1. *Birkhoff's Individual Ergodicity Theorem*

Suppose that the time evolution U_t with the time t is defined on the phase space X such that U_t preserves the probability measure P defined on X, that is, for any subset A of X, $P(A) = P(U_{(-t)}(A))$ holds. Let us consider a physical quantity $f(x)$

defined for $x \in X$, where $f(x)$ belongs to the set $L^1(P)$ of functions that satisfy the following condition:

$$\int_X |f(x)|\, dP(x) < \infty \tag{8}$$

This condition requires that the integral of the quantity $f(x)$ in the region $X_+ \equiv \{x \in X | f(x) > 0\}$ and that in the region $X_- \equiv \{x \in X | f(x) < 0\}$ converge, respectively.

Let $x(t) \equiv U_t x$ denote the trajectory with an initial condition $x(0) = x \in X$. Then the following quantity exists:

$$\hat{f}(x) \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t f(x(t'))dt' \tag{9}$$

for initial conditions almost everywhere concerning the probability measure P. Moreover, the function $\hat{f}(x)$ is invariant under the time evolution U_t, that is, $\hat{f}(x) = \hat{f}(U_t x)$, and

$$\int_X \hat{f}(x)dP(x) = \int_X f(x)dP(x) \tag{10}$$

holds.

According to the theorem, the time average $\hat{f}(x)$ of the quantity $f(x)$ exists for the trajectory with the initial condition x. Moreover, the invariance of the function $\hat{f}(x)$ means that the time average $\hat{f}(x)$ is constant for any initial conditions over each *individual* trajectory, hence the theorem is called the individual ergodic theorem.

However, the time average $\hat{f}(x)$ can take different values for different trajectories. Therefore, an additional requirement is needed to guarantee that time averages do not depend on initial conditions. This is the requirement of ergodicity. (Some references call it *metrical transitivity*, see, for example, Ref. 27.)

2. *Requirement of Ergodicity*

Suppose that the time evolution U_t is defined on the phase space X such that U_t preserves the probability P defined on X. The evolution U_t is called *ergodic* if either $P(A) = 0$ or $P(X\backslash A) = 0$ holds for any subset A of X with the property $A = U_{(-t)}(A)$, that is, A is invariant under U_t. We denote $X\backslash A$ the complement of A, that is, the subset of X that contains all the elements not belonging to A.

For a time evolution U_t that satisfies the requirement of ergodicity, Birkhoff's individual ergodicity theorem indicates that the time average of the physical quantity

$f \in L^1(P)$ equals its ensemble average for almost every initial condition $x \in X$, that is,

$$\hat{f}(x) = \int_X f(x')dP(x') \tag{11}$$

almost everywhere on X. This can be proved as follows. Denote $A(a) \equiv \left\{x \in X | \hat{f}(x) \geq a\right\}$ for an arbitrary value a. Both $A(a)$ and $X \backslash A(a)$ are invariant subsets because of Birkhoff's individual ergodicity theorem. Then, either $P(A(a)) = 0$ or $P(X \backslash A(a)) = 0$ holds based on the *requirement of ergodicity*. Thus, for an arbitrary value a, either $P(A(a)) = 1$ or $P(A(a)) = 0$ holds. This means that $P(A(a))$ is discontinuous at some value $\hat{a}$, indicating that $\hat{f}(x)$ takes the constant value $\hat{a}$ almost everywhere. Moreover, this constant equals the ensemble average $\int_X f(x)dP(x)$.

In the traditional argument of statistical physics, we consider the time evolution U_t under the Hamiltonian H. The measure-preserving condition is guaranteed by the Liouville theorem, that is, the probability measure P is given by the phase space volume suitably normalized, as long as U_t is defined on a certain *compact* subset X of the phase space. Assuming the requirement of ergodicity, Birkhoff's individual ergodicity theorem indicates that the time averages of the physical quantities exist and do not depend on initial conditions, that is, the idea of ergodicity in the sense of statistical physics is justified.

The following is a historical comment [27]. In the original idea, ergodicity meant that *every* point in the phase space was visited by a trajectory. However, it is impossible for a one-dimensional trajectory to cover the whole phase space of multiple dimensionality. Something one-dimensional cannot occupy *all* the points in a space of higher dimension. Therefore, the concept of ergodicity must be relaxed. Now, ergodicity is understood to mean that a trajectory covers the phase space densely, that is, it comes arbitrarily close to every point in the phase space.

In Birkhoff's individual ergodicity theorem, the condition $f \in L^1(P)$ for the physical quantity $f(x)$ is crucial. When physical quantities do not belong to this set, we can have a different situation. Also note that the situation differs completely for the cases with unnormalizable measures. These issues will be discussed later in Section II.C.

In the mathematical formulation of ergodicity, the time averages are defined over the infinite time interval. For physical situations, however, the time averages must be taken over finite time intervals. We are thus led to the question, "To what extent is ergodicity attained in the physical sense?" This issue will be discussed in the next section.

In physical problems, the correlation $\langle f(0)f(t)\rangle \equiv \int_X f(x)f(U_t x)dP(x)$ is also of interest for the physical quantity $f(x)$ in the set $L^2(P)$ of functions that satisfy

the following condition:

$$\int_X |f(x)|^2 \, dP(x) < \infty \tag{12}$$

Suppose that the correlation decays exponentially with the characteristic timescale t_c. Then, the two values $f(x)$ and $f(U_t x)$ of the physical quantity $f \in L^2(P)$ can be considered as independent as long as the time difference t is larger than t_c. This enables us to obtain the central limit theorem for the physical quantity $f \in L^2(P)$ [31].

The above argument leads us to another important property of the dynamical systems, that is, *mixing*. We call the time evolution U_t mixing, when the cross-correlations $\langle f(0)g(t)\rangle \equiv \int_X f(x)g(U_t x)dP(x)$ decay to zero for any physical quantities $f, g \in L^2(P)$ with their ensemble averages equal to zero. We have presented the definition of mixing in terms of measure theory in Section I.B. To see the equivalence of this definition with the measure-based definition of Eq. (4), put $f = \chi_A$ and $g = \chi_B$, the characteristic functions of the sets A and B:

$$\chi_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases} \tag{13}$$

Then, the correlation $\langle \chi_A(0)\chi_B(t)\rangle \equiv \int_X \chi_A(x)\chi_B(U_t x)dP(x)$ is equal to the probability $P(A \cap U_{(-t)}B)$, since the integrand $\chi_A(x)\chi_B(U_t x)$ equals 1 when both $x \in A$ and $U_t x \in B$ hold, that is, $x \in A \cap U_{(-t)}B$, otherwise $\chi_A(x)\chi_B(U_t x) = 0$. Subtracting the averages $\langle \chi_i\rangle \equiv \int_X \chi_i(x)dP(x) = P(i) \quad (i = A, B)$, respectively, we obtain $\langle(\chi_A(0) - \langle\chi_A\rangle)(\chi_B(t) - \langle\chi_B\rangle)\rangle = \langle\chi_A(0)\chi_B(t)\rangle - \langle\chi_A\rangle\langle\chi_B\rangle$ that approaches zero as the time t goes to infinity, indicating that $\langle\chi_A(0)\chi_B(t)\rangle$ approaches $\langle\chi_A\rangle\langle\chi_B\rangle$. Thus, we obtain $P(A \cap U_{(-t)}B)$ that goes to $P(A)P(B)$ as t goes to infinity.

B. Issues on Openness and/or Inhomogeneity

Here, we consider the foundation of the statistical reaction theory especially for those reactions involving biomolecules. The following two features become important: an *openness* and/or *inhomogeneity*.

The first issue is statistical properties within *open* phase space regions. In the traditional idea, the phase space is supposed to be compact, that is, closed and of finite volume. Moreover, trajectories do not flow into the phase space region and never leave it, thereby staying there for infinite time from the past to the future. The Liouville theorem guarantees that the measure-preserving property holds for the phase space volume, that is, the Lebesgue measure, and the probability measure is normalizable. In this sense, the phase space is *closed* in traditional statistical physics. On the other hand, in reaction processes, trajectories flow into and out of

the phase space region or regions that correspond to the potential well or wells. In this sense, the phase space is open in the chemical reactions.

In the chemical reactions, trajectories stay within the phase space region of a well only for a *finite* time interval. After entering the phase space region and staying there for some time, trajectories leave the region by going over a saddle and enter a new region, leading to chemical change. Thus, ergodicity in the statistical reaction theory concerns the question of the extent statistical statements are valid within finite time intervals. In the traditional theory of reactions, it is supposed that the trajectory visits almost everywhere in the phase space region in the well. If ergodicity in this local sense is satisfied, reaction processes become statistical and independent of specific initial conditions.

A closely related question was presented recently as a criticism of the traditional understanding of ergodicity [26]. In the traditional understanding, it is supposed that the trajectory visits the phase space region densely. However, Gallavotti pointed out that for systems of many degrees of freedom, it takes too long in the physical sense for the trajectory to cover the *whole* phase space densely. In other words, for *macroscopic* systems, the traditional understanding of ergodicity is irrelevant as the foundation of statistical physics.

Both the above arguments concern the necessity of introducing a criterion and a characteristic timescale so that we can estimate if ergodicity holds effectively in the physical sense. Such a criterion was proposed by Thirumalai and Straub [55, 56] called *the ergodic measure*. The quantity concerns fluctuation of time averages over *finite* timescales. If the fluctuation behaves consistently with the asymptotic behavior predicted by the law of large numbers, we can conclude that the statistical limit is effectively attained in the *physical* sense within finite timescales.

Note that, in introducing such criteria, we do *not* need to require that each trajectory covers densely the whole phase space. Rather, we need to estimate whether the asymptotic limit in the sense of the law of large numbers is attained or not. The reason why we focus our attention on this point is the following. In the traditional discussion of ergodicity, we treat *homogeneous* systems consisting of large numbers of *identical* particles. In these systems, a trajectory does *not* need to cover the whole phase space densely to exhibit statistical properties predicted based on ergodicity. It only suffices to cover a *representative* region of the phase space. Because of the permutation symmetry in systems consisting identical particles, time averages over such a representative region can be almost the same as the time average over the whole phase space. Moreover, such a representative region can be much smaller than the whole phase space. The characteristic timescale for ergodicity to hold in the physical sense can be much shorter than the timescale to cover densely the whole phase space.

The above argument leads us to the second issue that is, ergodicity for *inhomogeneous* systems. For biomolecules such as proteins, the above argument on a representative region is not readily applicable since these molecules tend to be

heterogeneous in their amino acid sequences. Moreover, in reaction processes involving biomolecules, we consider statistical aspects not necessarily in the *macroscopic* scale but in *mesoscopic* scales. For example, Thirumalai and Straub have shown, using the ergodic measure, that the degree of attaining ergodicity differs depending on the parts of the protein [56]. Their results indicate the possibility that some parts of the protein still remain out of thermal equilibrium while other parts recover equilibrium. Thus, ergodicity in *parts* of the biomolecule is of interest as a possible tool to see nonequilibrium within a *single* molecule. Such nonequilibrium situations can play an important role in the functional behavior of biomolecules as we point out later in Section II.D.

These arguments show that the ergodicity problem in the physical sense becomes even more important as we pay attention to biomolecules in mesoscopic scales. Then, openness and/or inhomogeneity become two key issues.

C. New Developments in Dynamical System Theory

Recently, new developments in the dynamical system theory offers some clues to investigate the issues related to ergodicity discussed in the previous sections. Here, we address two recent results, that is, the Sinai–Ruelle–Bowen (SRB) measure and an extension of the Birkhoff's individual ergodicity theorem.

First, we discuss the SRB and related measures [29–32, 34, 50]. In the traditional understanding of statistical physics, it is supposed that the phase space volume (exactly speaking, the Lebesgue measure) is the *only* relevant measure for statistical physics. However, in chaotic scattering processes, for example, fractal exists in the scattering events, which is singular with respect to the Lebesgue measure. In chaotic dissipative systems, a consideration of fractals becomes important due to the presence of strange attractors. These phenomena lead us to ask what the relevant physical measure is, in the sense that it corresponds to observation in experiments and numerical simulation.

The SRB measure is the measure that is smooth along the unstable invariant manifold, while it is singular along the stable invariant manifold. For compact uniformly hyperbolic systems, it is proved that the SRB measure exists [49]. Its existence can be intuitively understood as follows. Suppose a typical distribution of initial conditions on the phase space in the sense that its Lebesgue measure is positive. Through the time evolution, the distribution is stretched repeatedly along the unstable manifold. Under these processes, nonuniformity of the distribution becomes less and less pronounced leading eventually to a smooth distribution. Along the stable manifold, to the contrary, folding processes make nonuniformity of the distribution more and more steep, eventually giving rise to a singular distribution.

Suppose that we have arbitrary initial distributions that is typical in the sense that its Lebesgue measure is positive. The distribution approaches the SRB measures

under the time evolution, that is, the SRB measure is the natural invariant measure, the measure that a typical distribution of initial conditions approaches under the time evolution. Moreover, it is conjectured that the SRB measure is structurally stable, that is, it is not sensitive to random noise or a change of the parameters of the system. In this sense, it is considered as the physical measure, that is, the measure based on time averages obtained by physical observation [31, 50]. The SRB measure is expected to give a clue to understand nonequilibrium phenomena such as turbulence [53].

The theory of SRB measure has also revived the argument between Boltzmann and Loschmidt, leading to the fluctuation theorem. The fluctuation theorem states that universal behavior exists in the ratio between the probability of increasing entropy and that of *decreasing* entropy [26, 33]. Thus, the theory of SRB measure opens a new research area in nonequilibrium physics from the viewpoint of the dynamical systems.

The above discussion leads us to extend further the SRB measure to even wider situations. We should note that, in the requirement of ergodicity, whether ergodicity holds or not *depends* on which measure you use. Moreover, the phase space volume is *not* necessarily an appropriate invariant measure in chaotic scattering and systems with dissipation, as we have explained. Thus, we need to think of the question which measure we should use. The clue to answer this question is given by the existence of variational principles. The SRB measure can be characterized by the variational principle [29, 30, 34, 50]. This corresponds to the fact that equilibrium distributions are characterized through a variational principle as attaining the maximum of entropy or the minimum of the free energy. In this sense, the SRB measure enables us to extend the concepts of equilibrium distribution to nonequilibrium situations. Based on this similarity, a measure that can be characterized by the variational principle in general is called the Gibbs measure. The variational principle is formulated using Lagrange multipliers. The canonical distribution in equilibrium statistical physics is obtained by the variational principle under the constraint that the energy is given. We think further of the variational principle where the values of any physical quantities (not necessarily energy) are given by observation. This generalization introduces a new concept of measures, that is, the Gibbs measures.

For open hyperbolic systems, Gaspard and Dorfman [52] introduced a measure that is characterized by the variational principle, that is, the Gibbs measure. This measure is concentrated on the saddles of the chaotic scattering, that is, the repeller in the phase space. Given an arbitrary typical distribution of initial conditions, the closer those trajectories approach the repellers, the longer they remain in the scattering region. In the asymptotic limit of an *infinite* timescale, the invariant measure is thus defined on the repellers in the phase space. The measure has a finite value only for scattered trajectories since only the scattered trajectories are counted. Chaotic scattering introduces the singular measure that is concentrated

on the repellers. In this sense, the variational principle here means that observation of scattering trajectories uniquely singles out the relevant physical measure, which is *not* the Lebesgue measure. They show that the measure plays an important role in quantifying statistical properties of stationary events for open systems such as scattering and reaction processes where fractal structure becomes manifest in the invariant distributions.

An interesting question arises if we can extend further the concept of the Gibbs measure to normally hyperbolic invariant manifolds (NHIMs). The NHIMs are manifolds where hyperbolicity on the normal directions is stronger than that on the tangential directions. Thus, the definition of NHIMs corresponds to extensions of repellers to multidimensional dynamical systems. Therefore, normal directions to the NHIM play the role of the reaction coordinate. On the other hand, the tangential directions to the NHIMs consist of vibrational modes, which can be coupled with each other and be chaotic, as long as their hyperbolicity is weaker than hyperbolicity along the normal directions. Any typical distribution of initial conditions in the initial state will approach the NHIM located near the saddle as these trajectories leave the well leading to the reaction. The nearer they approach the NHIM, the longer they take to leave the well. Thus, we can construct the measure on the NHIM similarly to that on the repellers.

Second, we discuss an extension of the Birkhoff's individual ergodicity theorem [28, 38–44, 47]. and its relation to nonstationary processes in reactions [45]. Recently, the Birkhoff's individual ergodic theorem has been extended in the following two directions: (i) those cases in which the physical quantity $f(x)$ does not belong to $L^1(P)$ with normalizable probability measures P and (ii) those cases in which the invariant measure is not normalizable, that is, the cases that can be treated by what is known as the infinite ergodic theory [28].

For these cases, the concept of time averages is extended, and a new formulation of the law of large numbers is introduced. Then, an interesting new feature is that the asymptotic limit of time averages itself exhibits random fluctuation. Moreover, its distribution reveals a certain universal behavior. For example, Aizawa and his group have shown these universal characteristics for a class of one-dimensional maps and certain billiard systems [39–44]. The existence of universal fluctuation suggests that the statistical reaction theory can be extended to those reactions in which the traditional concept of ergodicity does not hold. Such cases can include the reaction processes in the mixed phase space where the reaction rate constant does *not* exist because of the fractional behavior such as power law in the distribution of the residence times, anomalous diffusion, and $1/f$ spectra. See the chapter 3 and Ref. [62–64].

The infinite ergodic theory can be important in those phenomena where extreme events play a crucial role in reaction processes. These days, extreme events in natural and social science receive an intense attention [25, 37] since these extreme events play a decisive role in phenomena such as earthquakes and great depressions,

although they are rare. In particular, when the probability of extreme events is larger than that predicted by the Gaussian distribution, the predictions based on the Gaussian can lead to catastrophic disasters in the society.

In considering extreme events in reactions involving biomolecules, existence of the gap is important between the characteristic timescale for reactions (as rapid as picoseconds for ligand binding or local conformational change) and that for biological functions (as slow as milliseconds to seconds or hours for protein folding, signaling, or transport). This wide gap in the characteristic timescales implies that even extremely rare events in terms of microscopic reactions can be considered frequent in the timescales of biological functions. This phenomenon is similar to the geological events, in which earthquakes are rare events in the characteristic timescale of individual human being while they are frequent on the timescale of geological events. Inspired by such a similarity, the term "protein quake" was coined for describing behavior of the protein [51]. These authors also noticed a hierarchical structure of substates and an associated distribution of bottlenecks that give rise to "broken ergodicity" and nonergodic behavior of the protein on a given finite dynamical timescale. Existence of common features in the protein and geological events suggests that the study of extreme events from the viewpoint of the infinite ergodic theory can lead to finding new universal aspects in nonequilibrium phenomena.

In order to analyze reaction processes from the viewpoint of extreme events and their universality, we need to extend the study of Aizawa's group to multidimensional dynamical systems. For example, Shojiguchi et al. have shown that nonstationary and power law behavior exists in systems where resonance overlap in the Arnold web is nonuniform and sparse in the well [62–64]. There is a possibility that the asymptotic distributions of physical quantities in such nonstationary systems exhibit the universal distribution.

D. Biomolecules as Maxwell's Demon

In order for biomolecules to play a role in information processing, they must be under nonequilibrium conditions as the celebrated argument of Maxwell's demon indicates [23, 57, 58]. Maxwell's demon is a tiny existence of a molecular size, which can differentiate molecules, one from another, on the basis of a property such as energy. Maxwell showed that its existence would lead to violation of the second law of thermodynamics [23]. Now the commonly accepted view is that the fluctuation of equilibrium conditions invalidates the original argument of Maxwell [57, 58]. However, there is still a possibility that nonequilibrium conditions enable the demon to work its task of differentiating molecules, that is, a kind of information processing [57, 59]. In particular, the demon is studied based on the fractional behavior of dynamical systems although their studies are limited to systems of two degrees of freedom [60, 61].

The question then arises how nonequilibrium conditions are maintained at the molecular level, and whether a dynamical mechanism exists that contributes to maintain nonequilibrium conditions. In order to investigate these questions, the theory of reactions should go beyond the traditional concept of ergodicity. This study will reveal an intrinsic dynamical mechanism of biomolecules so that the molecule is capable of exhibiting the ability to process information.

III. ERGODICITY IN ISOMERIZATION OF SMALL CLUSTERS

Small clusters of atoms have emerged as very useful tools to help us understand how ergodic and chaotic behavior enter in the kinetics and dynamics, not only of their own motions but also of much more complex systems. This is partly because analyzing the behavior of a system of 3, 4, ...,10, ..., even to 50 or 100 particles is now a reasonable task with modern computing tools and partly because the complexity of the multidimensional configurational and phase spaces in which the particles move grows extremely rapidly with the dimensionality of the space, that is, with the number of degrees of freedom of the multiparticle system.

Some of the aspects of ergodicity that have emerged from the study of clusters are as follows: the importance of the differences in behavior in different local regions of the multidimensional potential surface, the utility of local probes such as local Lyapunov exponents, and the time evolution of ergodicity, from local to global character. We can learn how to identify and characterize the specific directions in phase space that are responsible for the magnitude and direction of the local Lyapunov exponents, the components that are the primary local propagators of ergodicity. The Lyapunov exponents, particularly their local analogues (which we simply call "local Lyapunov exponents," based on finite trajectories of some desired length), reveal the directions and extent to which a trajectory tends to carry a system away from its locality and hence the extent to which a trajectory moves to explore some different region of configuration and phase space. We remind the reader that Lyapunov exponents are the measures of how neighboring trajectories diverge or converge locally from one another, and that for Hamiltonian (conservative) systems, these appear in positive and negative pairs. The traditional concept of Lyapunov exponent is based on the average behavior over the full, accessible phase space.

We begin this discussion with a short review of how we learn the different kinds of behavior in different regions of the potential surface. The first indication of this came from the observation that the positive Lyapunov exponents of the three-particle triangular Lennard–Jones cluster, LJ_3, and the sum of those exponents, the Kolmogorov entropy, increase with the energy of the system, up to the range in which the system can just pass over the energy saddle of the linear configuration. In that energy range, the system behaves in a more ordered fashion than at slightly

lower energies [4]. Another measure studied in that investigation was the effective Hausdorff dimension, the dimension of the space in which the three atoms move on a timescale consistent with observations, for example, nanoseconds, but brief compared to the time for mode coupling in nearly harmonic molecules, for example, milliseconds [5–7]. Simulations at low energies, corresponding to about 2–10 K, show Hausdorff dimensions of 3.1–3.5, as one would expect from the three normal modes of vibration of a triangular molecule such as LJ_3 when its motions are essentially harmonic. The deviation from precisely 3 is a measure of the degree of mode coupling at those energies. However, at an energy corresponding to 18.2 K, the Hausdorff dimension is 5.9; the maximum possible is the number of degrees of freedom in phase space that are not individually conserved, which, for n is $6n - 10$ or 8. Hence, the Hausdorff dimension tells us that this three-body system is already quite nonrigid at this energy, although it doesn't quite have full freedom in its phase space. Likewise, the Kolmogorov entropy (K-entropy) or sum of Lyapunov exponents increases steadily at an accelerating rate from energies corresponding to about 1 K up to a maximum at an energy equivalent to 28 K, *drops* to a local minimum around 30 K, and then increases again. The drop occurs just at the energy that allows passage over the saddle at the linear configuration of the molecule [8].

In the fully chaotic liquid range, the n positive Lyapunov exponents λ_n increase according to a power law $\lambda_n = \alpha n^{\beta}$. The slope α increases rapidly with increasing temperature or energy; the exponent β is essentially unity at all energies or temperatures [8]. This analysis also examined the way the K-entropy depends on the range of interaction between atoms; this range can be varied systematically if, as used in this work, one represents the interaction between pairs of atoms with a Morse potential, $V(r) = \epsilon \exp[-2\rho(r - r_0)] - 2\exp[-\rho(r - r_0)]$. A value of ρ of 3 corresponds to the longest range of pairwise potential known between two atoms in a diatomic molecule; a value of about 7 corresponds, likewise, to the shortest range exhibited by pairs of atoms in diatomic molecules. Short-range interactions give rise to very rough energy landscapes with extensive parts of the topography at high energies; long-range interactions give rise to smoother landscapes with deep, well-defined minima [9]. The study by Hinde et al. [8] showed that the K-entropy of three-particle clusters with Morse interactions between particles has an energy dependence that clearly distinguishes the systems with very long-range interactions from others with shorter ranges of interaction. Those with ρ of 3 have K-entropies that rise monotonically with energy and flatten at high energies; those with ρ of 5 or more have maxima in their K-entropies, as they move to regions of high potential energy and hence lower kinetic energy on their potential surfaces, as Fig. 1 shows.

Other, closely related systems have revealed similar behavior. Linear triatomic clusters have larger maximum Lyapunov exponents than triangular clusters at the lowest energies at which the linear form can exist; but at higher energies, the

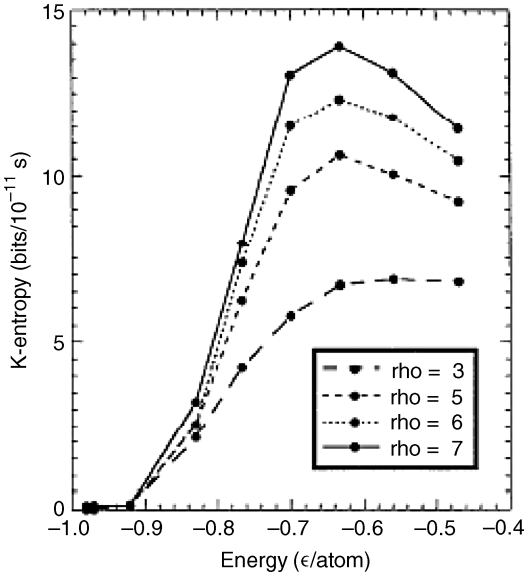

Figure 1. K-entropy for three-body systems with Morse potentials of various ranges. The shortest range here is that with ρ of 7. The span from 3 to 7 is approximately that of the known diatomic molecules, when they are represented by Morse potential interactions. (Reproduced with permission from Ref. 8. Copyright 1992, by the American Institute of Physics.)

two are very similar [10]. If some of the energy of the system is in rotational motion, then the system tends to be less chaotic, as indicated by a lower maximum Lyapunov exponent than for the case of pure vibration. However, varying the energy in rotation can reveal periodic transitions between regular and chaotic motion [11]. This point was explored in more detail to reveal that the volume of phase space occupied by regular trajectories is a nonmonotonic function of the angular momentum and depends on the coupling between kinetic and potential energy [12].

The second way that atomic clusters have opened an approach to the study of ergodicity and chaos has been in the area of finding timescales for the establishment of ergodic behavior in ever larger regions of the energy landscape [13, 14]. The probe to find the range of exploration in this approach is the *distribution* of effective Lyapunov exponents for brief, moderate, and long time intervals, but always just of *finite-time-based* Lyapunov exponents, not extrapolated to infinite time as one would determine traditional Lyapunov exponents. Clusters are particularly useful for this because, given their small sizes, they can exhibit dynamic coexistence of different phase-like forms in equilibrium over ranges of temperature and pressure, whether solid and liquid or different solid forms. Typically, under such conditions

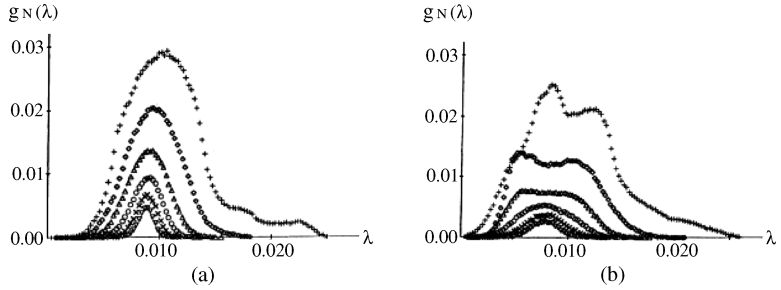

Figure 2. Distributions of sample values of Lyapunov exponents for three-atom clusters ("Ar_3" with Lennard-Jones potentials), taken from finite-path samples, at two temperatures. (**a**) 28.44 K; (**b**) 30.65 K; at 28.44 K, the system is below the linear saddle; at 30.65 K, it can pass over the saddle. The lowest distributions are based on 8192 time steps, and the successively higher sets are based on half the number of steps of the distribution below, so the second lowest are based on 4086 steps, and the highest on the shortest number, only 256 time steps. (Reproduced with permission from Ref. 14. Copyright 1993 by the American Physical Society.)

of coexistence, the residence time in one phase-like form is long relative to the time of vibrational periods or of thermal equilibration of the vibrational degrees of freedom [15, 16].

Figure 2 shows two sets of distributions of the sample values of the largest Lyapunov exponent for Ar_3 from molecular dynamics simulations at two temperatures, 28.44 and 30.65 K. The lowest "curves" are based on 8192 time steps of 10^{-14} s; the next higher, on 4086 steps, and so on, to the highest, which is based on only 256 time steps. The crucial point is that for short times, and a suitable temperature, even the argon trimer shows a bimodal distribution of Lyapunov exponents. This is more vivid with Ar_7, for which Fig. 3 shows the distributions of sample values of the largest Lyapunov exponents for short trajectories, of only 256 steps, as functions of both the value of the exponent and the kinetic energy at which that value occurred. The essential point of these figures is the passage from a narrow, unimodal distribution at low energies, through a region of bimodal distribution, to a high-energy region where the distribution is again unimodal but broad. Figures 2 and 3 show how, for brief times, systems explore only local regions; but for longer times, they visit their entire accessible phase space. Moreover, with probes such as the one used here, we can determine the timescale for passage from localized behavior to global. Global studies reveal some of the characteristics of larger clusters, notably their phase behavior, but the information in the distributions of local Lyapunov exponents gives additional insight into coexistence of phases even for clusters of over a thousand atoms [17].

One recognizes intuitively, and Hessian matrices demonstrate, that different directions of motion play different roles in the multidimensional configuration space of a several- or many-particle system. One very recent development has

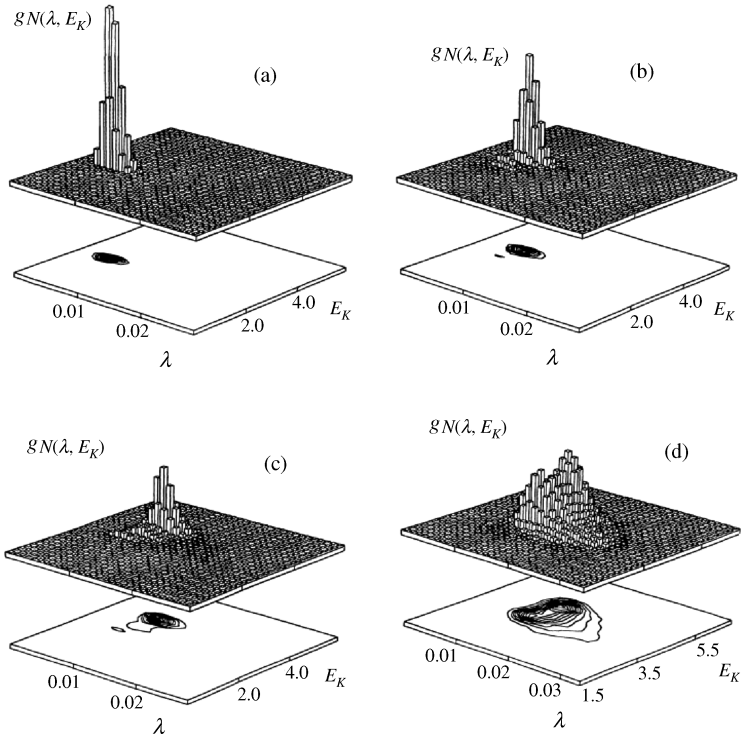

Figure 3. Distributions of sample values of finite sample Lyapunov exponents for a seven-particle cluster with Lennard–Jones interactions ("Ar_7"). The distributions are based on 256 time steps; they are expressed as functions of kinetic energy E_k and λ, the local value of the Lyapunov exponent. The distributions correspond to total energies of -0.355, -0.341, -0.328 and -0.300×10^{-13} erg. The units of λ are bits per 10^{-14} s and of E_k are 10^{-15} erg. (Reproduced with permission from Ref. 14. Copyright 1993 by the American Physical Society.)

explored this issue, with the goal of identifying the coordinates that play the most important roles in carrying a system from one local region to another [18]. This study uses a simple, Lennard–Jones cluster of three atoms as its model, in order to explore the distributions and participation ratio spectra of both traditional and local Lyapunov exponents. With even this very simple system, one can see that ergodicity develops on different timescales for different regions of phase space. Naturally, the regions most susceptible to unstable trajectories are those nearest to saddles. This particular study uses Gram–Schmidt vectors rather than the actual Lyapunov vectors, but the former are very close approximations to the latter, especially for very local investigations. This three-body system is a convenient device to begin to explore the kinds of information that one can extract from traditional and local Lyapunov exponents and the distributions of the latter. This is, in some

ways, a consequence of the fact that this system, with nine degrees of freedom and seven constants of motion (three components of momentum and angular momentum, and energy), has only two pairs of nonzero Lyapunov exponents, which, of course, come in matching positive and negative pairs. We refer to the larger as λ_1 and its most negative counterpart as λ_{18}, and the other two as λ_2 and λ_{17}. The investigation evaluated not only the Lyapunov exponents themselves but also the inverse participation ratios [19, 20], which measure the number of degrees of freedom that participate in the direction associated with each Lyapunov exponent. The distributions of the local Lyapunov exponents narrow steadily, as the length or duration of the trajectory extends. The distributions are quite narrow for trajectories of 2000 or more time steps, but very broad for only 100 or 200 time steps. Some bimodality of the sort observed by Amitrano and Berry was also seen in this work. This behavior is clear in the distributions in Figs. 4 and 5, for the larger and smaller positive, finite-interval Lyapunov exponents and the corresponding inverse participation ratios. Low values of the latter indicate many of the modes are participating in the direction corresponding to that Lyapunov exponent. One can

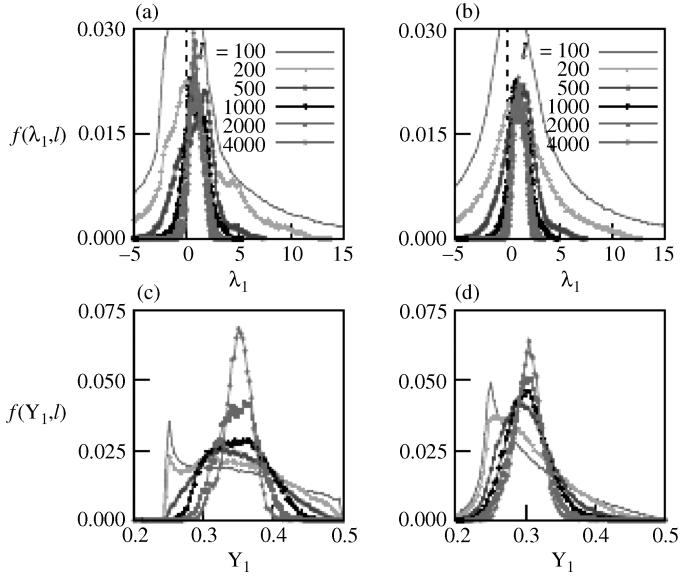

Figure 4. Distributions of the larger Lyapunov exponent λ_1 for ranges of sample time intervals l, in (**a**) and (**b**), and of the corresponding participation ratios Y_1. The participation ratio is a measure of the number of degrees of freedom that contribute to the direction of motion of each Lyapunov eigenvector. For (**a**) and (**c**), an amount of energy $E = -1.58\epsilon$ was put initially into the symmetric stretching mode, and for (**b**) and (**d**), the same energy was put initially into the asymmetric bending mode. The shortest interval sampled was 100 time steps, indicated by the thin curve without any dot, the lowest in (**a**) and (**b**), and the longest, 4000 time steps, is the most peaked in all four panels. (Reproduced with permission from Ref. 18.)

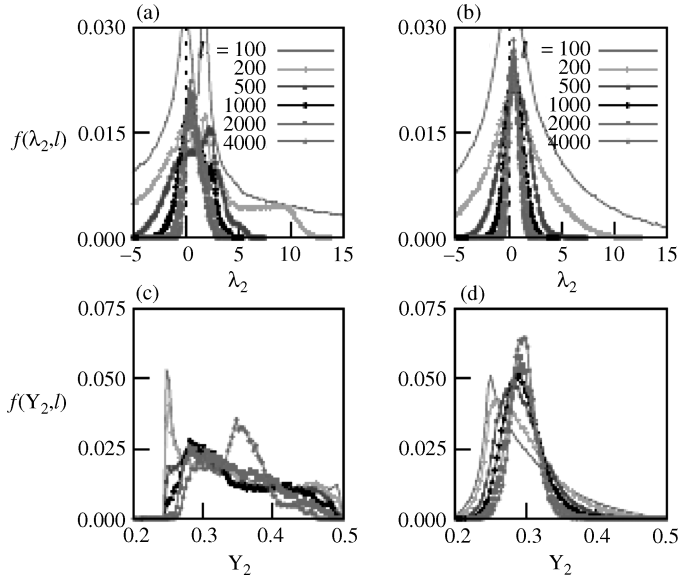

Figure 5. Distributions of the smaller Lyapunov exponent λ_2 for the same ranges of sample time intervals as in Fig. 4. All the notations are the same as in that figure. The most significant difference here is the bimodality of the two lowest curves in (**a**), corresponding to the system being in either of the two regions for such short intervals. (Reproduced with permission from Ref. 18.)

also see that the asymmetric bending mode of this triangular system plays an earlier role in inducing chaotic behavior than does the symmetric bending mode, in the sense that the asymmetric mode couples with the symmetric stretch at lower energies than does the symmetric bend. A result that emerges from these calculations is a coupling, perhaps surprising, of the excited *symmetric stretching mode* and *symmetric bending mode* with the asymmetric bending mode. Symmetry strictly forbids this, but tiny round-off errors in computation are sufficient to create small perturbations that break the symmetry and enable the coupling of asymmetric and symmetric modes.

Hence we can recognize the utility of local Lyapunov exponents as devices to help elucidate local dynamics, beyond the global features revealed by the traditional Lyapunov exponents.

IV. EXPLORING HOW PROTEINS WANDER IN STATE SPACE USING THE ERGODIC MEASURE AND ITS APPLICATION

The "complexity" of the energy landscape of proteins is responsible for the rich behavior observed in the dynamics of proteins [65–67]. The rugged energy surface arises from the presence of many energy scales in proteins due to the intrinsically

heterogeneous nature of the systems [68]. The equilibrium and dynamical properties of proteins are thought to be determined by a temperature-independent multidimensional potential hypersurface consisting of many minima (conformational substates), maxima, and saddle points. That general view of solids and liquids has a long history, dating back to Eyring. However, the ambitious project to provide a more quantitative assessment of the character of the underlying hypersurface using computational simulations has established in the "inherent structure" theory of Stillinger and Weber [69] and the "conformational substates" view of Frauenfelder [70]. In this picture, the distribution of energies for the minima, the volume of the basins, and the distribution of barrier heights separating these substates determine the thermodynamics and dynamics of the system. This point has been confirmed by the disorder seen in X-ray crystallographic studies and in the wide distribution of timescales for protein motion seen in the ligand photodissociation/rebinding experiments of Frauenfelder and coworkers on heme proteins [71].

Beginning with the pioneering study of Czerminski and Elber [68], computational studies have provided an increasingly quantitative description of the distribution of minimum energy conformations, the rate of exploration of these conformations, and relation to observable properties such as free energies and relaxation for small peptides [72, 73], model proteins [74, 75], and atomistic models of larger peptides [76] and proteins [77].

Recently, there has also been a focus on the application of sophisticated measures of phase space structures, typically restricted in applications to small molecules of relatively few degrees of freedom, to larger molecules and peptides. A focus of particular interest is the identification of local modes in proteins that may couple selectively to a few specific protein modes but relatively weakly to the larger density of states of the surrounding protein and solvent. Leitner and coworkers have pioneered the application of a number of methods, originally developed for the study of energy transfer in solids, to vibrational energy and heat flow in proteins [78] (see Chapter 3). Those methods have been applied and extended by Straub and coworkers to identify mode-specific energy transfer pathways of amide I vibrations in small peptide-like molecules [79, 80], globular proteins [81], and porphyrin and heme groups [82, 83] (see Chapter 1).

While applications to the study of energy flow in proteins have focused on dynamics in a constant temperature ensemble, there have also been significant experimental [84] and theoretical studies [85] focused on Hamiltonian (constant energy) flow in peptide-like molecules and small peptides. Significant developments enabling the experimental and theoretical study of biomolecules in the gas phase coupled with dramatic enhancements in computational power have led to the application of sophisticated methods for the study of phase space structures, previously restricted to the study of a few degrees of freedom systems and small molecules, to biomolecules [86]. A beautiful example can be found in the work of

Farantos who applied methods for computation of periodic orbits to examine the phase space structure of the alanine dipeptide [87]. An extension of that approach has recently been applied to interpret vibrational spectra in proteins [88]. These applications demonstrate the significant potential for the future study of the phase space structure of biomolecular systems.

A. The Kinetic Energy Metric as a Probe of Equipartitioning and Quasiequilibrium

One approach to exploring the nature of the rugged energy landscape and the rate at which observable properties are sampled is through measuring the convergence of averages over dynamics trajectories using replica molecular dynamics – the generalized ergodic measure originally introduced by Thirumalai and Mountain [89–91] and applied to a wide range of systems including proteins [92, 93]. This technique to examine the rate of sampling kinetic energy and atomic force has been shown to be a useful analytical tool for investigating timescales for energy equipartitioning and conformational space sampling [92, 94]. Interestingly, similar measures have been developed in other fields with issues of broken ergodicity where a state of quasiequilibrium is established other than the canonical thermal distributions, including self-gravitating systems [95].

In this section, we review the theory of the ergodic measure applied to estimate the rate of self-averaging of physical observables and characterize the dynamics in phase space using replica molecular dynamics. To provide insight into the behavior of the ergodic measure, the kinetic energy metric is evaluated analytically for the Langevin model and the force metric is evaluated analytically for a system of normal modes. In each case, the rate of convergence is shown to provide a measure of fundamental properties of the system dynamics on an underlying energy landscape.

Suppose we have an observable F that can be written as a function of time $F_i(t)$ for the ith atom of a system of N atoms, such as the kinetic energy $F_i(t) = m_i v_i^2(t)/2$. Writing the time average of $F_i(t)$ as $f_i(t)$ and the average of $f_i(t)$ over all N atoms of the system as $\bar{f}(t)$, we define the mean square difference of the individual $f_i(t)$'s from the average $\bar{f}(t)$ as

$$\Omega(t) = \frac{1}{N} \sum_{i=1}^{N} [f_i(t) - \bar{f}(t)]^2 \tag{14}$$

This is known as the fluctuation metric [94]. It can be shown that for an ergodic system after a short time the function $\Omega(t)$ decays to zero as $1/t$ as $\Omega(0)/\Omega(t) \simeq Dt$ (see Fig. 6) [90]. The power law decay of $\Omega(t)$ to zero at long times implies that the system is "self-averaging" and the slope is proportional to a diffusion constant for the exploration of the range of values (space) accessible to the variable $F(t)$. This is a necessary, but not sufficient, condition for the system dynamics to be ergodic.

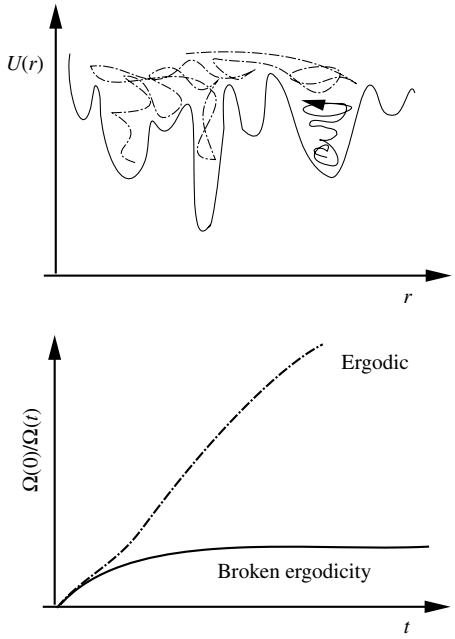

Figure 6. Two trajectories depicted on the background of a rugged energy landscape (*top*) and the corresponding reciprocal ergodic measure $\Omega(0)/\Omega(t)$ for an ergodic system and a system demonstrating "broken ergodicity" (*bottom*).

The slope of $\Omega(t)$ is proportional to the generalized diffusion constant D for the observable F that can be written $D\Omega(0) = l^2/\tau$ where $\Omega(0)$ is the mean square fluctuation of the property F and τ is the timescale for taking a "step" of generalized mean square length $l^2 = \Omega(0)$ in sampling the fluctuations of the property F in phase space. In this way, the ergodic measure may be used to explore the rate of exploration of phase space in complex systems characterized by a rugged free energy landscape.

Imagine that phase space is divided in two regions A and B by an impassable barrier. Given enough time any trajectory will explore all of the allowed phase space. For a set of trajectories started in region A, $\Omega(t)$ will decay to zero, and the property $F(t)$ will appear to be self-averaging. However, unless we have started one of our trajectories in region B we cannot know that the partition exists and the system in not ergodic. Therefore, the decay of $\Omega(t)$ to zero is a necessary but not sufficient condition for ergodicity.

The ergodic measure is readily calculable while alternative measures of ergodicity (or stochasticity) such as Lyapunov exponents [96] are considerably more involved and not as obviously relevant to the convergence of thermodynamic

properties as the ergodic measure. While there are strong connections between the convergence of the ergodic measure and the rate of spectral entropy production [89], the ergodic measure has been shown to provide significantly greater insight into the underlying protein dynamics. Moreover, the ergodic measure is readily computed for systems of arbitrarily large dimension, a great advantage in studies of protein dynamics.

It is possible to derive the diffusion constant for the kinetic energy metric by assuming that the velocity is a Gaussian random variable. The kinetic energy (or local temperature) metric $\Omega_{KE}(t)$ can be expressed in terms of the fluctuations of the kinetic energy $\delta f_i(t) = (m v_i^2(t) - 3k_B T)/2$ as

$$\Omega_{\mathrm{KE}}(t) = \frac{1}{t^2} \int_0^t ds_1 \int_0^t ds_2 \, C_i(s_1 - s_2) \tag{15}$$

where in the limit of large N we identify $C_i(t)$ as the equilibrium time correlation function of the fluctuations in the kinetic energy about its equilibrium average value.

Berne and Harp noted that the velocity may be modeled as a Gaussian random variable if the information entropy corresponding to the probability of having the velocity at time t and the velocity at time 0 is maximized [97]. Through that approximation the autocorrelation function for any higher moments of the velocity may be calculated in terms of the normalized velocity autocorrelation function $\psi_i(t)$ for the ith atom. The autocorrelation function for the fluctuation of the kinetic energy about its equilibrium average value may then be written $C_i(t) = (3/2)(k_B T)^2 \psi_i^2(t)$ and the diffusion constant for the kinetic energy metric is

$$D_{\mathrm{KE}} = \frac{1}{2}\left[\frac{1}{N}\sum_i \int_0^\infty dt \psi_i^2(t)\right]^{-1} \tag{16}$$

providing a means to determine the slope of $\Omega_{\mathrm{KE}}(0)/\,\Omega_{\mathrm{KE}}(t) \simeq D_{\mathrm{KE}} t$ for a particular model of the system dynamics characterizing $\psi_i(t)$.

B. The Kinetic Energy Metric as a Probe of Internal Friction

The Langevin description of the motion of atoms in proteins is often used to interpret kinetic experiments. The dynamics captured by the Langevin model is the foundation of modern reaction rate theory. In particular, the Langevin model in combination with a normal mode description of the protein has been used to interpret inelastic neutron scattering data for proteins [98]. Starting from the normal mode description and assuming the friction tensor to be diagonal, the velocity autocorrelation function for each of the $3N$ normal coordinates is of the form $\psi_i(t) = \exp(-\beta_i t/2)[\cos(a_i t) + (\beta_i/2a_i)\sin(a_i t)]$ with $a_i^2 = \omega_i^2 - \beta_i^2/4$ and $\omega_i^2 = \kappa_i/m_i$, where ω_i, κ_i, m_i, and β_i are the normal mode frequency, the effective

harmonic force constant and effective mass, and friction of the ith normal mode of the system, respectively.

Attempts to determine the friction acting on atoms in a protein, whose motion is bounded, have relied on fits to approximate forms of the correlation functions for the position and velocity [99, 100]. For the kinetic energy metric, using Eq. (16) we find that the generalized diffusion constant assumes the form

$$D_{\mathrm{KE}} = \frac{1}{2}\left[\frac{1}{N}\sum_i \frac{1}{2\beta_i}\right]^{-1} = \left\langle \frac{1}{\beta} \right\rangle^{-1} \tag{17}$$

when $\kappa_i \geq 0 \ \forall \ i$. This is a useful result. It shows that through a straightforward determination of the diffusion constant for the kinetic energy metric, the friction acting on the motion of a particle may be characterized. Note that the value of D_{KE} is dominated by the *smallest* value of the friction. The asymptotic convergence of the ergodic measure provides information on those degrees of freedom that most slowly approach equilibrium. This reflects the heterogeneous nature of phase space that has been associated with interesting properties in chemical kinetics, particularly in the low friction energy diffusion regime [101].

As has been shown by Baba and Komatsuzaki in their recent work focusing on the interpretation of single-molecule spectroscopy [122], a detailed understanding of the timescale for the establishment of local equilibrium between states is essential to the interpretation of single-molecule dynamics in terms of dynamics on a free energy landscape. Presumably, the rate of attaining local equilibrium, observed in that work, is intimately connected to the distribution of rates for establishing local self-averaging, discussed here in the context of the ergodic measure. Their definition of local equilibrium states (LES) and its use in the decomposition and construction of a free energy landscape is explored in Section V.

Average values of the relaxation rate cannot be used to characterize the rates of relaxation for specific protein modes. The method presented here provides a straightforward means of estimating the friction for all atoms of the protein in a way that allows for a global analysis of the dominant pathways for kinetic energy relaxation. More detailed normal mode-based approaches have been developed to provide insight into the mode-specific nature of the energy transfer pathways defining the system's exploration of phase space [78].

C. The Force Metric as a Probe of the Curvature in the Energy Landscape

The fluctuation metric $\Omega_{\mathrm{KE}}(t)$ provides a measure of the timescale for self-averaging a given observable, such as the kinetic energy, over a single trajectory. An alternative is to calculate averages over independent trajectories and measure the rate at which these independent averages converge to the same value as they must

for an ergodic system. This idea has been applied to assess the rate of convergence for the various contributions to the atomic force (bonds and angles, dihedrals, and nonbonded van der Waals and Coulombic) in peptides and proteins [94].

To explore this idea, we define the force metric and evaluate it for the special case of a harmonic system. The time average of the force on the ith atom, $\vec{f}_i^a(t)$, is defined as for the fluctuation metric where a indicates that the average is calculated over the ath trajectory. Given two trajectories a and b, with independent starting configurations, we define the difference between the averages calculated over each trajectory as the metric

$$d_F(t) = \frac{1}{N}\sum_{j=1}^{N} |\vec{f}_j^a(t) - \vec{f}_j^b(t)|^2 \tag{18}$$

In an ergodic system, at long times self-averaging is achieved and the metric decays to zero as $d_F(t)/d_F(0) \simeq 1/D_F t$, where D_F means the corresponding diffusion constant. Conversely, if $d_F(t)$ does not decay as a power law we infer that there trajectories a and b must sample distinct free energy basins.

When the system dynamics consists of small excursions about a well-defined average structure, it may be a good approximation to model the system dynamics using a quenched normal mode approximation. The potential is expanded in a Taylor series of the $3N$ coupled coordinates about a mechanically stable equilibrium position where $\overleftrightarrow{\kappa}$ is the $3N \times 3N$ dimensional force constant matrix or matrix of second derivatives about the equilibrium position. The normal mode transformation diagonalizes $\overleftrightarrow{\kappa}$ by a unitary transformation defined by the matrix $\overleftrightarrow{a}$ resulting in $3N$ normal mode coordinates and corresponding frequencies ω_i. The force metric may then be written as

$$d_F(t) = \frac{4k_B T}{NF^2}\sum_{i}^{3N}\langle m\rangle_i[1 - \cos(\omega_i t)] \tag{19}$$

where we have defined the "average mass of the ith normal mode" as $\langle m\rangle_i = \sum_j^N m_j|a_{ji}|^2$ using the normalization condition that $\sum_j^N |a_{ji}|^2 = 1$. In this context, F is a vector observable of interest, such as the force on a given atom or internal coordinate, and F^2 is its mean-square value.

For proteins the average mass is fairly independent of the mode number and we can approximate $\langle m\rangle_i = \bar{M}$, where $\bar{M}$ is the average atomic mass for the peptide or protein atoms [93]. The initial value of the force metric can be expressed in terms of the second moment of the vibrational density of states $d_F(0) = 6\bar{M}k_B T\langle\omega^2\rangle$, where $\langle\omega^2\rangle = (1/3N)\sum_i^{3N}\omega_i^2$, leading to a remarkably simple form for the asymptotic limit of the force metric:

$$d_F(0)/d_F(t) \simeq \frac{1}{2}t^2\langle\omega^2\rangle \tag{20}$$

Application to the Debye model leads to the result $\langle\omega^2\rangle = 3\omega_D^2/5$ providing a means of calculating the approximate Debye frequency for the system from the curvature of the total force metric.

For the underdamped normal modes, inertial motion dominates and the metric shows superconvergence as $1/t^2$. However, in the case that the modes are coupled through nonlinear forces, phase space may be divided by vague tori [102] (see also the generalization to multidimensional systems in Section III in Chapter 4) or as an Arnold web [103] (see also Section II in Chapter 3), slowing convergence (see Fig. 6). In such cases, the ergodic measure may be used to assess the extent of broken ergodicity, as a convenient alternative to the more demanding computation of Lyapunov exponents [89].

The use of the force metric to assess the average curvature of the underlying free energy landscape has a variety of potential applications. For example, Stillinger and LaViolette applied the inherent structure theory to examine the validity of the Lindemann melting criterion that solids become unstable when the mean-square fluctuations of the atoms approach 10% of the lattice spacing, for simple solids [104]. That idea was extended to examine the nature of the folding transition in proteins [105] leading to the interesting conclusion that in the native state the interior of proteins may be considered "solid-like" while the surface behavior is more "liquid-like." Deeper insight into the nature of those phases could presumably be gained through computation of the distribution of Lyapunov exponents, as demonstrated by Berry and coworkers in their studies of the dynamics of atomic clusters and discussed earlier in this chapter. More recently, a connection has been established between the curvature of the underlying free energy landscape and the stability of proteins [106]. In that work, the mean square gradient of the potential, related to the underlying curvature of the free energy landscape, was shown to be intimately connected to the statistical temperature, $T(E) = (\partial S(E)/\partial E)^{-1}$, and the depth of the free energy basin containing the native state of the protein.

D. Extensions of the Ergodic Measure to Internal Energy Self-Averaging

Using the fluctuation metric, we have shown that the kinetic energy is equipartitioned on a timescale of picoseconds. Evidence suggests that longer time relaxation associated with conformational transitions in the peptide is best explored using the metric of the nonbonded (Coulombic and van der Waals) potential energy. The equations for the energy metric are found by substituting the scalar energy $e_i^\alpha(t)$ for the ith atom in the αth trajectory into expressions where the corresponding force vector leads to

$$d_E(t) = \frac{1}{N}\sum_{i=1}^{N}[e_i^a(t) - e_i^b(t)]^2 \tag{21}$$

In a rugged energy landscape, the nonbonded energy metric shows rapid initial convergence followed by a slow, long-time decay. The initial convergence

is significantly greater at higher temperatures (see Fig. 6). In many proteins, the plateau in the reciprocal metric is reached within 3 ps for $T < 240$ K [93]. This indicates that at lower temperatures the peptide motion is confined to fluctuations about a single local free energy basin without significant conformational transitions on a 75 ps timescale. At 300 K, there is a significant region of linear convergence followed by a plateau beyond 15 ps. This behavior clearly indicates the presence of a wide distribution of timescales for the protein motion. Moreover, within the short timescale several free energy basins are sampled as indicated by the change in the slope of $d_E(t)$. The longer time relaxation is related to infrequent barrier crossing, largely in the form of dihedral angle transitions (discussed below) and the diffusive motion of subdomains of the protein that may shift in relative orientation [107].

E. Probing the Heterogeneity of Energy Flow Pathways in Proteins

In the study of heme protein dynamics, the concept of rapid intramolecular vibrational relaxation within the heme has been a significant focus of experimental and theoretical study for decades [108, 109]. In the protein carboxymyoglobin, the heme group and its host protein share only one covalent bond—that between the proximal histidine and the iron atom. The heme is otherwise kept in place by roughly 90 van der Waals contacts with surrounding protein atoms, much like a molecule in solution. The extent of this isolation can be brought to light using a modified form of the ergodic measure defined by an average over all atoms of the reference system that may, for example, be the heme alone, the protein alone, or the protein and solvent bath. The results are shown in Fig. 7 and summarized in Table I [110].

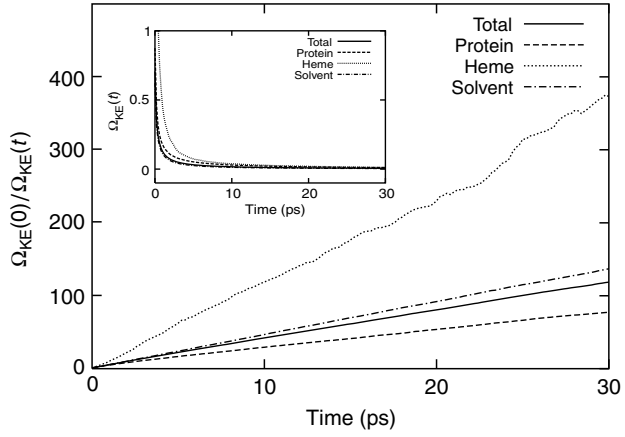

Figure 7. Convergence of the reciprocal ergodic kinetic energy metric $\Omega_{KE}(0)/\Omega_{KE}(t)$ computed from the molecular dynamics trajectory of the protein carboxymyoglobin in aqueous solution at 300 K for subsets of atoms in the protein, solvent, and heme.

TABLE 1
Summary of the Inhomogeneous Character of the Average Static Friction Computed for Specific Elements of the Solvated Heme Protein System in Carboxymyoglobin at 300 K

Region	γ_0 (ps^{-1})
System	
Total	2.12
Protein	1.87
Heme	6.76
Solvent	7.49
Backbone	
Oxygens	9.56
Nitrogens	9.45
α-Carbons	6.92
Carbonyl carbons	3.78
Hydrogens	0.91
Residues	
Charged	2.39
Aliphatic	1.69
Aromatic	1.09

Energy redistribution within the heme itself occurs on a faster timescale than the thermalization of the heme with its environment. This indicates that the rate-limiting step in heme relaxation following photolysis is the "doorway" to energy transfer between the heme and its surroundings. The location and effectiveness of a given IVR pathway can depend sensitively on the nonlinear dynamics of the multimode system and the details of coupling between local modes that determine the character of dynamics in phase space.

Overall, application of the ergodic measure to explore protein dynamics has deepened our appreciation for the inhomogeneous character of the distribution of rates of phase space sampling that correlates with the details of the protein structure.

V. EXTRACTING THE LOCAL EQUILIBRIUM STATE (LES) AND FREE ENERGY LANDSCAPE FROM SINGLE-MOLECULE TIME SERIES

In this section, we review our recent studies on the extraction of LES, in which the system exhibits ergodicity in the metastable state, from single-molecule time series such as interdye distance in single-molecule measurements. We also revisit the concept of the free energy landscape of proteins and discuss what kinds of energy landscape proteins actually experience during the course of their time evolution.

Recent experimental developments in single-molecule spectroscopy have provided us with several new insights into not only the distribution of the observable but also the dynamical information at single-molecule level [111–115]. Fluorescence resonance energy transfer (FRET) experiments monitor the fluorescence intensity from donor (D)/acceptor (A) molecules attached in a single protein. The observed quantities are expected to trace the time evolution and the distribution of the D–A distance at the single-molecule level. For example, some experimental studies have indicated the existence of abnormal diffusion depending on the timescale at which one observes the dynamical events [114], heterogeneous pathways for protein folding in adenylate kinase [113], and different timescales of relaxation in an intermediate state and in the unfolded state of iso-1-cytochrome *c* [116]. Yang et al. [114] showed in single-molecule time series of flavin reductase with a bound flavin that abnormal diffusion emerges for a timescale less than 10^{-1} s in the fluorescence lifetime fluctuation while it turns to normal Brownian diffusion for longer timescales than 10^{-1} s. Rhoades et al. [113] observed a broad distribution in the timescale of the folding transitions and an importance of non-Markovian conformational dynamics especially for slow transitions (>1 s) by trapping adenylate kinase within surface-tethered lipid vesicles. Kinoshita et al. [116] found using a new single-molecule detection technique employing a capillary flow system that iso-1-cytochrome *c* (known as having a collapsed intermediate state) exhibits relatively slower conformational dynamics in the unfolded state, compared to that in the intermediate state. As argued by Talaga et al. [117], for large conformation changes of proteins, there exist some spurious interactions caused by immobilization from the surface linkage when using direct surface-linking techniques. It should be noted that the above experimental systems [113, 116] were specially designed so as to be free from the artifacts with the surface linkage technique and to detect large conformational change for longer time duration compared to the confocal microscopy experiment.

However, the most fundamental question of *what type of energy landscape single molecules actually see during the course of such dynamical evolution* remains unresolved. To address this question, an essential goal must be the development of a device or means to extract the relevant information concerning the local equilibrium states and their network from scalar single-molecule time series.

A. Extracting LES from Single-Molecule Time Series

There exist several problems in the single-molecule measurements [118–121] in addition to the problem discussed by Talaga et al. [117]. One of the most cumbersome obstacles is the so-called "degeneracy problem" due to the dimensionality of the observable: even when the system resides in different physical states, the value of the observable (scalar time series) is not necessarily different and may be degenerate due to the finite resolution of the observation and the nature of

the projection of the underlying multidimensional information onto the chosen observable.

Baba and Komatsuzaki have recently developed a new method for extracting LES from a given scalar time series as free as possible from the degeneracy problem and constructing an effective free energy landscape [122]. In short, the crux is to evaluate states not by the value of the time series at each time but by the short-time distributions in the neighborhood of each time with a time window τ. The short-time distributions reflect not only the value at each time but also the higher order moments, that is, variance, skewness, kurtosis, and so on in the vicinity of *the* time. Thus, the short-time distribution can differentiate the states that are degenerate in the value itself as much as possible under the limitation of the available information (i.e., solely the scalar time series). The secondary crux is to present the criteria of assigning if an obtained state can be regarded as LES or not within the timescale τ by checking the timescale separation between the τ and the escape time from the state candidate. This procedure naturally revisits the concept of the free energy landscape and provides us with a fundamental question of what type of energy landscape the system actually follows in a chosen timescale. In the following, we present their procedure briefly.

Figure 8 shows their scheme to construct a set of state candidates from time series $s(t)$. Suppose that $s(t)$ is recorded with an equal interval from t_1 to t_n. First, they extract "short segments" in a time window $(t_m - \tau/2, t_m + \tau/2]$ in the vicinity of t_m and construct the short-time probability density function $g_m^{(\tau)}(s)$ ($m = k+1, k+2, \ldots, n-k$, where k is larger than the size corresponding to $\tau/2$) (see Fig. 8a). Second, they quantify the degree of proximity of two probability density functions by using Kantorovich metric [123] defined by

$$d_K(p_i \| p_j) = \int_{-\infty}^{\infty} ds \left| \int_{-\infty}^{s} ds' \left(p_i(s') - p_j(s') \right) \right| \tag{22}$$

where $p_i(s)$ and $p_j(s)$ are two arbitrary probability density functions with respect to s. It was found [122] that the d_K is much more appropriate than the most commonly used measures, for example, Kullback–Leibler divergence (relative entropy) [124] and Hellinger distance [125], in differentiating the distance between two probability density functions. This is due to the fact that Kullback–Leibler divergence and Hellinger distance give rise to a single value (the former is ∞ and the latter is $\sqrt{2}$) when $p_i(s)$ and $p_j(s)$ have no overlap in the variable s. Figure 8b illustrates the metric relationship (regarding d_K) among $g_m^{(\tau)}(s)$ by projecting onto a two-dimensional plane so as to maintain the metric relationship among them [126] (note that in the actual procedure they do not need any projection of the full dimension into such a lower dimension). Each node corresponds to each $g_m^{(\tau)}(s)$ at a different time t_m. Third, Baba and Komatsuzaki partition the set of $g_m^{(\tau)}(s)$ into a union of "clusters" on the full-dimensional metric space as illustrated by clusters

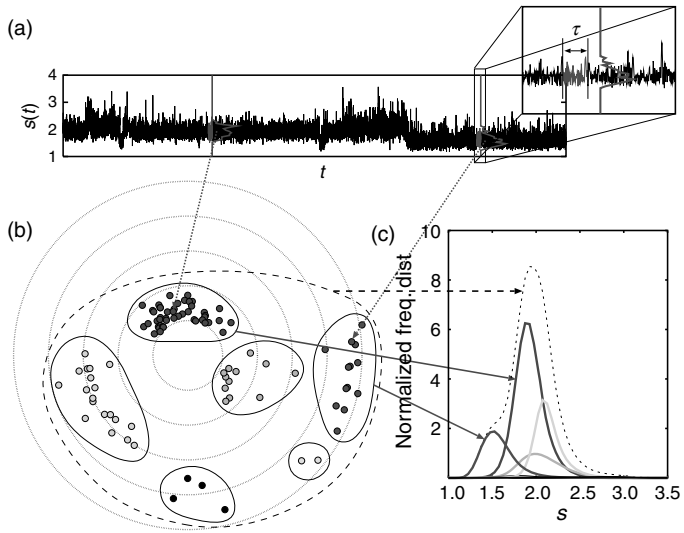

Figure 8. (**a**) A time series $s(t)$. For every mth time step t_m, the short-time probability density function $g_m^{(\tau)}(s)$ ($\int g_m^{(\tau)}(s)ds = 1$) is evaluated for a time window $(t_m - \tau/2, t_m + \tau/2]$. (**b**) A two-dimensional projection of a set of $g_m^{(\tau)}(s)$ so as to maintain the "metric" relationship among the $g_m^{(\tau)}(s)$ [determined by Eq. (22)]. Each node or circle corresponds to each $g_m^{(\tau)}(s)$ at different t_m (for the visual clarity, not all but every other 10,000 sampled points of $g_m^{(\tau)}(s)$ are plotted from the time series in (**a**)). The set covered by the dotted line indicates the full set of $g_m^{(\tau)}(s)$ corresponding to the full $s(t)$. Different subsets (covered by solid lines) of different nodes correspond to the different state "candidate" where the composite $g_m^{(\tau)}(s)$ are classified as the same group on this metric space in the full dimension. (**c**) The corresponding frequency distributions of the four major state "candidates" with respect to s. (Reproduced with permission from Ref. 122. Copyright 2007 by the National Academy of Sciences.)

surrounded by solid lines in Figure 8b. Each cluster composed of $g_m^{(\tau)}(s)$, all of whose shapes are almost the same, naturally provides a *candidate* of state. Figure 8c shows the corresponding frequency distributions of the four major state candidates. Note here that this procedure does not need to assume any shape of the distribution function associated with the underlying states (although one has often assumed it as Gaussian in the conventional analysis with fixing the total number of states).

The most important step is to incorporate the concept of local equilibrium states into the "candidates of states." They considered the following criteria in the timescale for assigning the candidate of state as an LES:

$$\tau_{\mathrm{eq}}(i) < \tau < \tau_{\mathrm{esc}}(i) \tag{23}$$

where $\tau_{\mathrm{eq}}(i)$ denotes the characteristic timescale of the system to be locally equilibrated within the state i and $\tau_{\mathrm{esc}}(i)$ that of the system to escape from the state i

to the other. That is, they considered a "candidate of state" (= a grouped subset of $g_m^{(\tau)}(s)$) should satisfy the condition, if it can be regarded as an LES, that a timescale for which the system is locally equilibrated inside the "state candidate i" is shorter than the timescale of observation τ and the τ is longer than a timescale at which the system escapes from it.

One can easily check the latter condition $\tau < \tau_{\mathrm{esc}}(i)$ by identifying which "candidate of state" the system resides in, enters to, and leaves from along the time series by checking to which cluster the short-time distribution at each time belongs. However, the equilibration timescale of the state i, $\tau_{\mathrm{eq}}(i)$, in the former condition is inaccessible by using $g_m^{(\tau)}(s)$. Theoretically, the state classified as an LES should possess a unique local distribution of the observable whenever the system revisits the same state along the course of time evolution. It is because, by definition, the system cannot escape any LES before locally equilibrated, resulting in a certain unique distribution function of any physical quantity. Therefore, their method can implicitly take into account the former condition in the procedure of grouping $g_m^{(\tau)}(s)$ into a set of approximately unique distribution functions. Their method, hence, assigns a state candidate as an LES if $\tau < \tau_{\mathrm{esc}}(i)$ is satisfied, otherwise as a non-LES at the given time window τ.

B. Revisiting the Concept of Free Energy Landscape

In order to validate the existence of barriers or "transition states" on the free energy landscape, the other condition is required, that is, (local) detailed balance. One can assess the local detailed balance as follows: The above procedure enables us to evaluate residential probability P_i of the ith LES, that is, how frequently the system (re)visits the ith LES along the time series. In addition, one can evaluate the rate constants $k_{i\to j}$ (and $k_{j\to i}$) from the $i(j)$th LES to the $j(i)$th LES, that is, the averaged, inverse of life time for which the system resides in the $i(j)$th LES before leaving to the $j(i)$th LES. The local detailed balance condition is represented by

$$k_{i\to j} P_i \simeq k_{j\to i} P_j \tag{24}$$

Suppose that we can utilize canonical transition state theory (TST) of the reaction rate, that is,

$$k_{i\to j} = k_B T/h \exp(-\Delta F^{\ddagger}_{i\to j}/k_B T) \tag{25}$$

where $\Delta F^{\ddagger}_{i\to j}$, k_B, h, and T denote the free energy barrier from the ith LES to the jth LES, Boltzmann constant, Planck constant, and absolute temperature, respectively. Then one can evaluate the free energy barrier $F^{\ddagger}$ on the free energy landscape that links the free energy minima F_i and F_j of ith LES and jth LES (see Fig. 9):

$$F_i = -k_B T \ln P_i \tag{26}$$

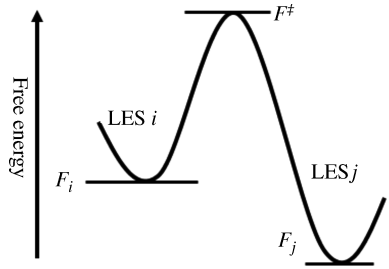

Figure 9. Free energy surface.

and

$$F^{\ddagger} = F_i - k_B T \ln\left(\frac{h}{k_B T} k_{i\to j}\right) = F_j - k_B T \ln\left(\frac{h}{k_B T} k_{j\to i}\right) \tag{27}$$

where F_i and $F^{\ddagger}$, respectively, denote the relative free energy of the ith LES and the relative free energy barrier linking the ith and jth LES. Note here that unless the local detailed balance is satisfied, the second equality in Eq. (27) does not hold. This implies that one can neither identify the relative free energy of the barrier $F^{\ddagger}$ nor construct the landscape of free energy unless the detailed balance holds. (Kramers theory [127] and Grote–Hynes theory [128] tell us that the canonical TST provides an upper bound of the rate constant. The free energy barrier derived from Eq. (27) can be affected by the existence of viscosity from the environment [129, 130]. An appropriate correction may be required for better estimation of the free energy barrier [112].)

Let us revisit the concept of free energy landscape. The free energy landscape $F(\mathbf{Q})$ is usually defined as a function of m-dimensional progress variables $\mathbf{Q}$:

$$Z(\mathbf{Q}) = \int\int\ldots\int d\mathbf{q}\, d\mathbf{p}\, \delta^m(\mathbf{Q}(\mathbf{q}) - \mathbf{Q}) \exp\left(-\frac{E(\mathbf{p},\mathbf{q})}{k_B T}\right) \tag{28}$$

$$F(\mathbf{Q}) = -k_B T \log Z(\mathbf{Q}) \tag{29}$$

where $E(\mathbf{p},\mathbf{q})$ denote the total energy of the system (e.g., protein) as a function of its momenta $\mathbf{p}$ and coordinates $\mathbf{q}$ coupled with a heat bath of temperature T. δ^m, $\mathbf{Q}(\mathbf{q})$, and $Z(\mathbf{Q})$ denote multidimensional Dirac's delta function defined by $\delta^m(\mathbf{z}) = \delta(z_1)\delta(z_2)\cdots\delta(z_m)$, progress variables by which free energy landscape is depicted (usually a certain function of $\mathbf{q}$), and the partition function with respect to $\mathbf{Q}$, respectively. As for the definition, one can always compute this landscape as a function of the chosen $\mathbf{Q}$ by assuming that all the degrees of freedom distribute according to the Boltzmann distribution except a set of the "frozen" $\mathbf{Q}$. Quite often one has argued the *dynamics* in $\mathbf{Q}$ on the landscape of $F(\mathbf{Q})$. The question to be

addressed here is this: on what type of energy landscape a single protein actually feels along the evolution in the **Q** space? In the other words, what should be the hidden assumption to validate the picture of dynamics on free energy landscape $F(\mathbf{Q})$?

The key condition is that the system must attain ergodicity to yield the Boltzmann distribution within the space complement to the **Q** space in the sense of statistical mechanics (see Section II.). That is, in general, the characteristic timescale with respect to the **Q** motions is much longer than those of the other degrees of freedom so that the system can move about the complementary space at each "constant" **Q**. Furthermore, such timescale separation between **Q** and the other degrees of freedom should maintain through the whole regime where $F(\mathbf{Q})$ is constructed. Some of the other degrees of freedom complement to **Q** may become slow compared to the timescale in **Q** depending on the regime of the whole space. However, such a possibility of changing the relationship in timescale between each variable **Q** and the rest should not be invoked in the context of the free energy landscape.

One may notice the essential difference between the free energy profile in Fig. 9 composed of the LES and the free energy barrier and free energy landscape defined by Eqs. (28) and (29). The former is defined by LESs and free energy barrier based on the local ergodicity and local detailed balance that should be considered in the space or coordinates to describe the landscape (e.g., some **Q**). That is, the former definition refers to some requisite conditions in the **Q** dynamics. On the other hand, the latter definition of free energy landscape [Eqs. (28) and (29)] never requires any *a priori* condition in the **Q** dynamics themselves except the timescale separation. The "free energy landscape" defined by Eqs. (28) and (29) is more appropriately referred to as "the potential of mean force landscape" because it invokes no constraint on the **Q** dynamics.

C. Extracted LES of a Minimalistic Protein Model at Different Temperatures

Suppose that we have a set of single-molecule time series of interdye distance $d(t)$ that could be converted from the FRET intensity as a function of temperature such as that in Fig. 10. How can one retrieve the underlying local equilibrium states and their free energy landscape from such limited information?

Baba and Komatsuzaki illustrated the potential of their method by using the scalar time series of the end-to-end distance of an off-lattice 3-color, 46-bead model protein [131, 132] at several temperatures [122]. This model (called the BLN model) is composed of hydrophobic (B), hydrophilic (L), and neutral (N) beads. The global potential energy minimum for the sequence, $B_9N_3(LB)_4N_3B_9N_3(LB)_5L$, folds into a β-barrel structure with four strands (see Fig. 11). Two peaks are seen in the heat capacity: one corresponds to the collapse temperature ($\sim 0.65\epsilon$ [134]) at which the BLN model transitions from the

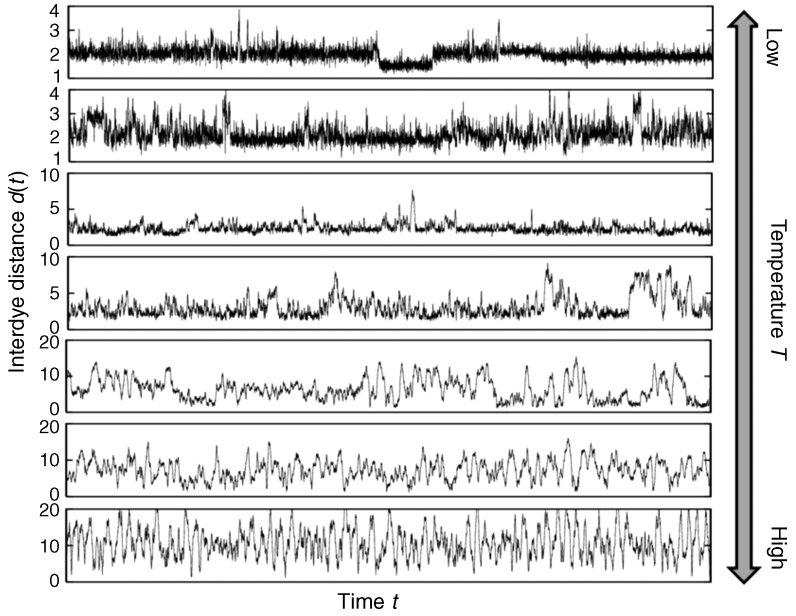

Figure 10. An "interdye distance" $d(t)$ of a protein at different temperatures. The $d(t)$ plotted here is the end-to-end distance of an off-lattice 3-color, 46-bead model protein [131, 132]. Temperature T, from the top to the bottom column, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and $2.0\epsilon/k_B$ (ϵ is the unit of energy of the model protein and k_B the Boltzmann constant). The isothermal MD simulation was performed by Berendsen's thermostat [133]. After a long simulation for equilibration, the value of d was recorded every 100 steps along the 13 million-step trajectory. Here the MD step, Δt, corresponds to $\sim$1/180 of the timescale of one vibration of the bond. The coupling constant γ with the thermostat was chosen as $\sim 1/(200\Delta t)$ so that the canonical distribution can be quickly attained during the course of MD simulation [122].

extended to the compact, collapsed states, and the other to the folding temperature (0.27 [135]-0.35ϵ [134]) where it folds into the global potential energy minimum [136, 137].

Figure 12 shows the result of the analysis of time series $d(t)$, that is, the normalized frequency distributions $f(d)$ of the assigned LES and non-LES with a chosen time window. The $f(d)$ is defined by, in the absence of any broadening effects of signal in the measurement,

$$f(d) = \sum_{m \in \text{cluster } i} \delta(d - d(t_m)) / \sum_{m \in \text{all clusters}} \int_{-\infty}^{\infty} ds \delta(s - d(t_m)) \qquad (30)$$

The larger the $f(d)$, the more the system resides in the LES/non-LES. Here the time window τ was set to be $10{,}000\Delta t$, which corresponds to $\sim$55 oscillations of the bond vibration and 50 times longer than the timescale of the coupling $\sim 1/\gamma$.

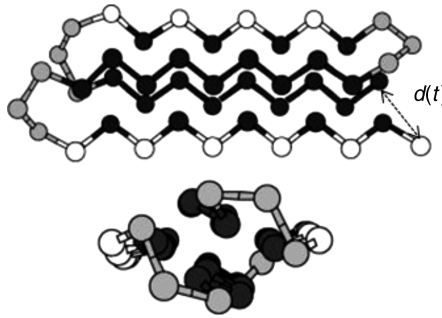

Figure 11. The global minimum structure of BLN protein model. The black, white, and gray color beads correspond to hydrophobic (B), hydrophilic (L), and neutral (N) beads, respectively. The potential energy function is described by $V = (K_r/2)\sum_i (r_i - r_0^i)^2 + (K_\theta/2)\sum_i (\theta_i - \theta_0^i)^2 + \sum_i [A(1+\cos\Phi_i) + B(1+\cos 3\Phi_i)] + 4\epsilon \sum_{i<j-3} S_1[(\sigma/R_{ij})^{12} - S_2(\sigma/R_{ij})^6]$, where $S_1 = S_2 = 1$ for BB (attractive) interactions, $S_1 = 2/3$ and $S_2 = -1$ for LL and LB (repulsive) interactions, and $S_1 = 1$ and $S_2 = 0$ for all the other pairs involving N, expressing only excluded volume interactions. $K_r = 231.2\epsilon\sigma^{-2}$ and $K_\theta = 20\epsilon/\text{rad}^2$, with the equilibrium bond length $r_0^i = \sigma$ and the equilibrium bond angle $\theta_0^i = 1.8326$ rad.

At 0.3 and 0.4 ϵ/k_B, four and three large LESs are identified. As T increases to $0.5\epsilon/k_B$, one can see three large LESs into which some of the LES/non-LES at $0.4\epsilon/k_B$ are unified. The three large LES observed at $0.5\epsilon/k_B$ are interpreted as unified into one large LES at $0.6\epsilon/k_B$. This implies that the system quite often traverses back and forth the three unified LES at $0.6\epsilon/k_B$, crossing over the barriers that link the three LES observed at $0.5\epsilon/k_B$ within the chosen τ. Note also that some "delocalized" distributions start to emerge (with low probabilities) besides this large unified state at $0.6\epsilon/k_B$. At $0.7\epsilon/k_B$, although the "compact" large LES ceases to exist, delocalized distributions become more significant with higher probabilities. Note that from 0.6 to $0.7\epsilon/k_B$, the "center" of the LES migrates from the short to the long end-to-end distance regions, which corresponds to the transition from the collapsed state to the unfolded state. This manifests the existence of T_c between 0.6 and $0.7\epsilon/k_B$, which is consistent with the previous assignment in terms of heat capacity $\sim 0.65\epsilon/k_B$ [134]. At $0.8\epsilon/k_B$ above T_c, two distributions are classified as LES while the other distributions violate Eq. (22) in the τ.

The most striking messages of this illustrative example are twofold: First, at single-molecule level with a finite timescale, LES and non-LES can generally coexist as a function of temperature and timescale of the observation. Second, one can see that none of the distributions is classified as LES at $0.7\epsilon/k_B$ within the chosen τ. This indicates that, on the timescale τ, in neither the compact states nor more delocalized, denatured states, the system can be well equilibrated (i.e., the residential times in them are shorter than the τ) although either at higher or lower

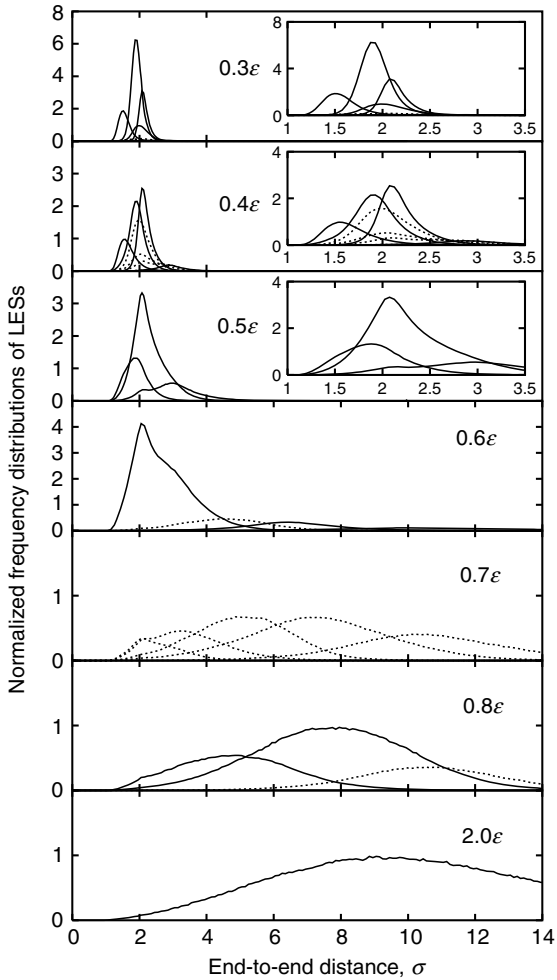

Figure 12. $f(d)$ of LES/non-LES constructed from the $d(t)$ at different temperatures, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and $2.0\epsilon/k_B$. In the insets of 0.3–0.5ϵ, the magnified graphs in the horizontal axis are also given. The unit of vertical axis is 10^{-2}. The $f(d)$ indicated by the dotted lines are assigned as non-LESs. (Reproduced with permission from Ref. 122. Copyright 2007 by the National Academy of Sciences.)

temperatures the system can more likely equilibrate at least in some of the states at the same chosen timescale τ. We believe that this property of the attainability of local ergodicity depending on the "temperature" is generic irrespective of the system when the system has several potential basins [138].

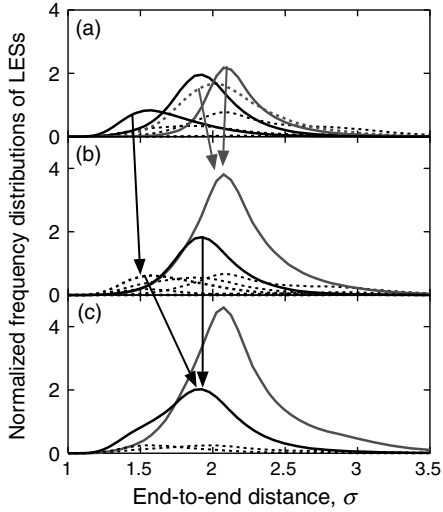

Figure 13. The dependence of LES/non-LES on the time window τ at $T = 0.4\epsilon$. The number of the sample points to evaluate the local distributions (n_S) corresponds to (**a**) 400, (**b**) 500, and (**c**) 2000, respectively. The unit of vertical axis is the same as in Fig. 12. The solid and dashed lines indicate the LES and non-LES distributions, respectively. The gray arrows indicate the merger of a non-LES into an LES as $n_S = 400 \rightarrow 500$. The black arrows indicate the change from LES to non-LES as $n_S = 400 \rightarrow 500$ and the merger of a non-LES into an LES as $n_S = 500 \rightarrow 2000$. (Reproduced with permission from Ref. 122. Copyright 2007 by the National Academy of Sciences.)

It was indicated in Fig. 13 that the topography of the landscape depends on the timescale of observation; two metastable states are unified as one if the timescale of observation is longer than the escape timescale for which the system can visit mutually these two states [122, 138]. The different time windows naturally lead to different coarse-grained LESs and different coarse-grained free energy landscapes, if they exist, the system should find at the different timescales.

This corresponds to a manifestation of transition from a set of two "ergodic" states (i.e., locally ergodic) to a single "ergodic" state, as observed in small atomic clusters in Section III, that is, a bimodal to unimodal peak in the local Lyapunov exponent distribution for a "phase transition" in a small finite system.

D. Outlook

In this section, we have reviewed a new method to extract effective free energy landscapes from single-molecule time series. If and only if the local equilibrium and the local detailed balance are satisfied in a chosen timescale of observation τ, one can construct the effective free energy landscape for the regions where the system wanders frequently. In the procedure, the equilibration time for which

the system can be locally equilibrated inside the chosen state(s) was not assessed directly but rather implicitly taken into account by checking invariance of the probability density function of the observable. One should note the emphasis of "local" in the phrase "local detailed balance." At the single-molecule level, in principle, detailed balance does not hold globally unless the length of time goes almost to infinity and the ergodicity in the sense of statistical physics is known to be satisfied in advance. However, one can still expect the existence of the detailed balance in a certain local regime with a certain timescale τ where free energy basins and barriers are well defined.

VI. FUTURE PERSPECTIVES

In this chapter, we have addressed the fundamental background and practical approaches for addressing ergodic problems in complex systems in chemical physics. Real complex systems have provided us with challenging subjects in ergodicity: multiplicity of ergodicity, heterogeneity of timescales to attain ergodicity for each moiety of the system, the openness of the phase space, and so on. Multiplicity of ergodicity is closely related to the issue of the time-dependent nature of the free energy landscape whose morphological features depend on the timescale of observation. The openness of the phase space demands that the observer incorporate the *finiteness of the timescale* of the observation when considering the ergodic properties of a system. In exploring the ergodic character of real systems, one must carefully choose criteria or measures of ergodicity and self-averaging. In practice, it is impossible to assure that the system satisfies the mathematical condition of ergodicity. Local Lyapunov exponent distributions provide insight into the distinction between chaotic and less chaotic regions, dimensionality of the exploration of the system through the full phase space, and a timescale of merging two or more ergodic regions. When one can compute the local Lyapunov exponent distribution, this measure provides substantial insight into the detailed system dynamics. However, when the size of the system becomes large, say, more than a hundred degrees of freedom or so, it is impossible to compute such an elaborate distribution. For those cases, the ergodic measure provides a practical measure of the rate of self-averaging that may be readily applied to systems of any dimension if one can know the equation of motion. Furthermore, the more apparent distinction between small-body systems and larger many-body systems such as proteins is the existence of multiple time- and space scales. As presented in Section III of Chapter 3, the divergence of nearby trajectories at infinitesimally small deviation in the phase space is expected to become almost irrelevant to protein dynamics in the large. For many-body systems such as proteins, many outstanding problems to be addressed now are rooted in the existence of a strongly inhomogeneous distribution of timescales that must be sampled in order to attain ergodic sampling. Moreover,

the exact timescale will depend on the moiety of the system, as shown in Section IV, and its relationship with protein function in the space of nonequilibrium degrees of freedom. When one cannot know the equation of motion or cannot compute dynamical information because of the limitation of computational power, a possible measure to address the concept of ergodicity should be based on the limited dynamical information obtained by experiments such as single-molecule time series. The method of LES analysis developed by Baba and Komatsuzaki can be a strong tool to retrieve the ergodic property from single-molecule measurements. Recently, their approach was applied to investigate which interdye distances (which pair of amino acid residues) are suitable to retrieve the underlying free energy landscape of protein G [139].

The much more challenging question is this: How is the concept of ergodicity/nonergodicity relevant to biological functions in more complex systems such as signal transduction in a cell or protein–protein interaction and other recognition/signaling problems? To look into this challenging problem, it is crucial to establish a device or means to quantify the robustness of the concept in terms of time series by linking with single-molecule biophysics. The argument of Maxwell's demon in Section II and the existence of heterogeneous distribution of timescales to achieve ergodicity at each moiety and hierarchy in Sections III and IV may suggest us that the system has not to "feel" full ergodicity or full equilibrium at timescales to characterize or be relevant to biological functions even though most degrees of freedom are more likely expected to be thermalized. That study provides a glimpse of the insights that may be achieved through the application of methods for probing ergodicity to complex biological systems.

Acknowledgments

T.K. and S.K. thank Dr. Hiroshi Teramoto for his fruitful discussions and comments. J.E.S. gratefully acknowledges the generous support of the National Science Foundation (CHE-0750309) and Boston University's Center for Computational Science. T.K. acknowledges financial support from JSPS, JST/CREST, Grant-in-Aid for Research on Priority Areas "Systems Genomics" and "Real Molecular Theory," MEXT. M.T. acknowledges financial support from JSPS, Grant-in-Aid for Research on Priority Area "Real Molecular Theory," MEXT, and Nara Women's University Intramural Grant for Project Research. S.K. acknowledges the support by research fellowships of the Japan Society for the Promotion of Science for Young Scientists. RSB wishes to acknowledge the hospitality of the Aspen Center for Physics, where much of his contribution was developed.

References

1. P. Walters, *An Introduction to Ergodic Theory*, Springer, New York, 1982.
2. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Addison-Wesley, 1989.
3. A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, 2nd ed., Springer, New York, 1992.
4. T. L. Beck, D. M. Leitner, and R. S. Berry, *J. Chem. Phys.* **89**, 1681 (1988).

5. P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983).
6. P. Grassberger and I. Procaccia, *Physica D* **9**, 189 (1983).
7. P. Grassberger and I. Procaccia, *Phys. Rev. A* **28**, 2591 (1983).
8. R. J. Hinde, R. S. Berry, and D. J. Wales, *J. Chem. Phys.* **96**, 1376 (1992).
9. D. J. Wales, *Energy Landscapes*, Cambridge University Press, Cambridge, 2003.
10. E. Yurtsever and N. Elmaci, *Phys. Rev. A* **55** 538 (1997).
11. E. Yurtsever, *Phys. Rev. A* **58** 366 (1998).
12. E. D. Belega, D. N. Trubnikov, and L. L. Lohr, *Phys. Rev. A* **63** 043203 (2001).
13. C. Amitrano and R. S. Berry, *Phys, Rev. Lett.* **68**, 729 (1992).
14. C. Amitrano and R. S. Berry, *Phys. Rev. E* **47**, 3158 (1993).
15. R. S. Berry, in *Theory of Atomic and Molecular Clusters*, J. Jellinek, ed., Springer, Berlin, 1999, Chapter 1 p. 1.
16. R. S. Berry, *Compt. Rend. Phys.* **3**, 1 (2002).
17. F. Calvo, *J. Chem. Phys.* **108**, 6861 (1998).
18. J. R. Green, J. Jellinek, and R. S. Berry, *Phys. Rev. E*, **80**, 066205 (2009).
19. A. Pikovsky and A. Politi, *Nonlinearity* **11**, 1049 (1998).
20. H.-L. Yang and G. Radons, *Phys. Rev. E* **71**, 036211 (2005)
21. M. Toda, *Adv. Chem. Phys.* **123**, 153 (2002).
22. M. Toda, *Adv. Chem. Phys.* **130A**, 337 (2005).
23. J. C. Maxwell, *Theory of Heat*, Dover, reprinted in 2001.
24. L. Boltzmann, *Lectures on Gas Theory*, Dover, Reprinted in 1995.
25. E. J. Gumble, *Statistics of Extremes*, Dover, reprinted in 2004.
26. G. Gallavotti, *Statistical Mechanics*, Springer, 1999.
27. M. Toda, R. Kubo, and N. Saito, *Statistical Physics*, Vol. I, Springer, 1992.
28. J. Aaronson, *An Introduction to Infinite Ergodic Theory*, American Mathematical Society 1997.
29. C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, Cambridge University Press, Cambridge, 1993.
30. R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Lecture Notes in Mathematics* Vol. 470, Springer, 1975.
31. P. Collet and J.-P. Eckmann, *Concepts and Results in Chaotic Dynamics*, Springer, 2006.
32. J. R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, Cambridge University Press, Cambridge, 1999.
33. D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, 2nd ed, Cambridge University Press, Cambridge, 2008.
34. P. Gaspard, *Chaos, Scattering and Statistical Mechanics*, Cambridge University Press, Cambridge, 1998.
35. A. I. Khinchin, *Mathematical Foundation of Statistical Mechanics*, Dover, 1949.
36. P. Walters, *An Introduction to Ergodic Theory*, Springer, 2007.
37. S. Albeverio and V. Jentsch, and H. Kantz, eds., *Extreme Events in Nature and Society*, Springer, 2006.
38. Y. Aizawa, *Prog. Theor. Phys.*, **72**, 659 (1984).
39. T. Akimoto and Y. Aizawa, *J. Korean Phys. Soc.* **50**, 254 (2007).
40. T. Akimoto, *J. Stat. Phys.* **132**, 171 (2008).

41. T. Akimoto and Y. Aizawa, *Chaos* **20**, 033110 (2010).
42. S. Shinkai and Y. Aizawa, *Prog. Theor. Phys.* **116**, 503 (2006).
43. S. Shinkai and Y. Aizawa, *Prog. Theor. Phys.* **116**, 515 (2006).
44. T. Miyaguchi and Y. Aizawa, *Phys. Rev. E* **75**, 066201 (2007).
45. G. Margolin and E. Barkai, *J. Stat. Phys.* **122**, 137 (2006).
46. A. Rebenshtok and E. Barkai, *J. Stat. Phys.* **133**, 565 (2008).
47. N. Korabel and E. Barkai, *Phys. Rev. Lett.* **102**, 050601 (2009).
48. G. D. Birkhoff, *Proc. Natl. Acad. Sci. USA* **17**, 656 (1931).
49. R. Bowen and D. Ruelle, *Invent. Math.* **29**, 181 (1975).
50. J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.*, **57**, 617 (1985).
51. A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, R. D. Young, *Proc. Natl. Acad. Sci. USA* **82**, 5000 (1985).
52. P. Gaspard and J. R. Dorfman, *Phys. Rev. E* **52**, 3525 (1995).
53. D. Ruelle, *Prog. Theor. Phys. Suppl.* **64**, 339 (1978).
54. X. J. Wan and C. K. Hu, *Phys. Rev., E* **48**, 728 (1993).
55. J. E. Straub, A. B. Rashkin, and D. Thirumalai, *J. Am. Chem. Soc.* **116**, 2049 (1994).
56. D. E. Sagnella, J. E. Straub, and D. Thirumalai, *J. Am. Chem. Soc.* **113**, 7702 (2000).
57. L. Brillouin, *Science and Information Theory*, Dover, reprinted in 2004.
58. H. S. Leff and A. F. Rex, eds., *Maxwell's Demon 2*, IOP, Bristol, 2003.
59. M. M. Millonas, *Phys. Rev. Lett.* **74**, 10 (1995).
60. G. M. Zaslavsky and M. Edelman, *Phys. Rev. E* **56**, 5310 (1997).
61. G. M. Zaslavsky, *Hamiltonian Chaos and Fractional Dynamics*, Oxford University Press, Oxford, 2004.
62. A. Shojiguchi, C. B. Li, T. Komatsuzaki, and M. Toda, *Phys. Rev. E* **75**, 025204(R) (2007).
63. A. Shojiguchi, C. B. Li, T. Komatsuzaki, and M. Toda, *Phys. Rev. E* **76**, 056205 (2007).
64. A. Shojiguchi, C. B. Li, T. Komatsuzaki, and M. Toda, *Phys. Rev. E* **77**, 019902(E) (2007).
65. H. Frauenfelder and P. G. Wolynes, *Science* **229**, 337–345 (1985).
66. A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young, *Proc. Natl. Acad. Sci. USA* **82**, 5000–5004 (1985).
67. A. Ansari, J. Berendzen, D. Braunstein, J. B. Johnson, P. Ormos, T. S. Sauke, R. Scholl, and A. Schulte, *Biophys. Chem.* **26**, 337–355 (1987).
68. R. Czerminski and R. Elber, *J. Chem. Phys.* **92**, 5580–5601 (1990).
69. F. H. Stillinger and T. A. Weber, *Science* **225**, 983–989 (1984).
70. H. Frauenfelder, S. G. Sligar, and P.G. Wolynes, *Science* **254**, 1598–1603 (1991).
71. H. Frauenfelder, F. Parak, and R. D. Young, *Annu. Rev. Biophys. Chem.* **17**, 451–479 (1988).
72. O. M. Becker and M. Karplus, *J. Chem. Phys.* **106**, 1495–1517 (1997).
73. B. W. Church and D. Shalloway, *Proc. Natl. Acad. Sci. USA* **98**, 6098–6103 (2001).
74. B. W. Church and D. Shalloway, *J. Chem. Phys.* **111**, 6610–6616 (1999).
75. T. Komatsuzaki, K. Hoshino, Y. Matsunaga, G. J. Rylance, R. L. Johnston, and D. J. Wales, *J. Chem. Phys.* **122**, 084714 (2005).
76. B. Tarus, J. E. Straub, and D. Thirumalai, *J. Mol. Biol.* **379**, 815–829 (2008).
77. B. Tarus, J. E. Straub, and D. Thirumalai, *J. Am. Chem. Soc.* **128**, 16159–16168 (2006).

78. D. M. Leitner, *Annu. Rev. Phys. Chem.* **59**, 233–259 (2008).
79. H. Fujisaki, K. Yagi, J. E. Straub, and G. Stock, *Int. Natl. J. Quantum. Chem.* **109**, 2047–2057 (2009).
80. Y. Zhang, H. Fujisaki, and J. E. Straub, *J. Phys. Chem. A* **113**, 3051–3060 (2009).
81. H. Fujisaki and J. E. Straub, *J. Phys. Chem. B* **111**, 12017–12023 (2007).
82. Y. Zhang, H. Fujisaki, and J. E. Straub, *J. Phys. Chem. B* **111**, 3243–3250 (2007).
83. Y. Zhang, H. Fujisaki, and J. E. Straub, *J. Phys. Chem. A* **130**, 025102 (2009).
84. B. C. Dian, A. Longarte, P. R. Winter, and T. S. Zwier, *J. Chem. Phys.* **120**, 124304 (2005).
85. J. K. Agbo, D. M. Leitner, D. A. Evans, and D. J. Wales, *J. Chem. Phys.* **123**, 133–147 (2004).
86. Y. Matsunaga and T. Komatsuzaki, *AIP Conf. Proc.* **708**, 342–343 (2004).
87. S. C. Farantos, *J. Chem. Phys.* **126**, 175101 (2007).
88. V. Daskalakis, S. C. Farantos, and C. Varotsis, *J. Am. Chem. Soc.* **130**, 12385–12393 (2008).
89. D. Thirumalai and D. Thirumalai, *J. Stat. Phys.* **57**, 789–801 (1989).
90. R. D. Mountain and D. Thirumalai, *J. Phys. Chem.* **93**, 6975–6979 (1989).
91. R. D. Mountain and D. Thirumalai, *Physica A* **210**, 453–460 (1994).
92. J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **90**, 809–813 (1993).
93. J. E. Straub, A. Rashkin, and D. Thirumalai, *J. Am. Chem. Soc.* **116**, 2049–2063 (1994).
94. J. E. Straub and D. Thirumalai, *Proteins* **15**, 360–373 (1993).
95. T. Tsuchiya, N. Gouda, and T. Konishi, *Astrophys. Space Sci.* **257**, 319–341 (1998).
96. A. J. Lichetenberg and M. A. Lieberman, *Regular and Stochastic Motion*, Springer, New York, 1983.
97. G. D. Harp and B. J. Berne, *J. Chem. Phys.* **49**, 1249–1254 (1968).
98. J. C. Smith, *Q. Rev. Biophys.* **24**, 227–291 (1991).
99. J. A. McCammon, P. G. Wolynes, and M. Karplus, *Biochemistry* **18**, 927–942 (1979).
100. S. Swaminathan, T. Ichiye, W. F. van Gunsteren, and M. Karplus, *Biochemistry* **21**, 5230–5241 (1982).
101. J. E. Straub, M. Borkovec, and B. J. Berne, *J. Chem. Phys.* **86**, 4296–4297 (1987).
102. R. B. Shirts and W. P. Reinhardt, *J. Chem. Phys.* **77**, 5204–5217 (1982).
103. A. Shojiguchi, A. Baba, C. B. Li , T. Komatsuzaki, and M. Toda, *Laser Phys.***16**, 1097–1106 (2006).
104. R. A. LaViolette and F. H. Stillinger, *J. Chem. Phys.* **83**, 4079–4085 (1985).
105. Y. Zhou, D. Vitkup, and M. Karplus, *J. Mol. Biol.* **285**, 1371–1375 (1991).
106. J. Kim, T. Keyes, and J. E. Straub, *Phys. Rev. E* **79**, 030902 (2009).
107. R. Elber and M. Karplus, *Science* **235**, 318–321 (1987).
108. Y. Mizutani and T. Kitagawa, *Science* **278**, 443–446 (1997).
109. E. R. Henry, W. A. Eaton, and R. M. Hochstrasser, *Proc. Natl. Acad. Sci. USA* **83**, 8982–8986 (1986).
110. D. E. Sagnella, J. E. Straub, and D. Thirumalai, *J. Chem. Phys.* **113**, 7702–7711 (2000).
111. X. S. Xie and J. K. Trautman, *Annu. Rev. Phys. Chem.* **49**, 441–480 (1998).
112. B. Schuler, E. A. Lipman, and E. A. Eaton, *Nature* **419**, 743–747 (2002).
113. E. Rhoades, E. Gussakovsky, and G. Haran, *Proc. Natl. Acad. Sci. USA* **100**(6), 3197–3202 (2003).
114. H. Yang, G. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie, *Science* **302**, 262–266 (2003).

115. E. Barkai, Y. Jung, and R. Silbey, *Annu. Rev. Phys. Chem.* **55**, 457–507 (2004).
116. M. Kinoshita, K. Kamagata, M. Maeda, Y. Goto, T. Komatsuzaki, and S. Takahashi, *Proc. Natl. Acad. Sci. USA* **104**, 10453–10458 (2007).
117. D. S. Talaga, W. L. Lau, H. Roder, J. Tang, Y. Jia, W. F. DeGrado, and R. M. Hochstrasser, *Proc. Natl. Acad. Sci. USA* **97**, 13021–13026 (2000).
118. L. P. Watkins and H. Yang, *Biophys. J.* **86**(6), 4015–4029 (2004).
119. L. Edman and R. Rigler, *Proc. Natl. Acad. Sci. USA* **97**(15), 8266–8271 (2000).
120. J. B. Witkoskie and J. Cao, *J. Chem. Phys.* **121**(13), 6361–6372 (2004).
121. O. Flomenbom, J. Klafter, and A. Szabo, *Biophys. J.* **88**, 3780–3783 (2005).
122. A. Baba and T. Komatsuzaki, *Proc. Natl. Acad. Sci. USA* **104**(49), 19297–19302 (2007).
123. A. Vershik, *J. Math. Sci.* **133**(4), 1410–1417 (2006).
124. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 1991.
125. W. J. Krzanowski, *J. App. Stat.* **30**(7), 743–750 (2003).
126. U. Brandes, P. Kenis, J. Raab, V. Schneider, and D. Wagner, *J. Theor. Politics* **11**, 75–106 (1999).
127. H. A. Kramers, *Physica* **7**, 284–304 (1940).
128. J. T. Hynes, *Theory of Chemical Reaction Dynamics*, CRC Press, Boca Raton, FL, 1985, Chapter 1, pp. 171–234.
129. N. D. Socci, J. N. Onuchic, and P. G. Wolynes, *J. Chem. Phys.* **104**, 5860–5868 (1996).
130. D. K. Klimov and D. Thirumalai, *Phys. Rev. Lett.* **79**, 317–320 (1997).
131. J. D. Honeycutt and D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **87**, 3526–3529 (1990).
132. R. S. Berry, N. Elmaci, J. P. Rose, and B. Vekhter, *Proc. Natl. Acad. Sci. USA* **94**, 9520–9524 (1997).
133. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684–3690 (1984).
134. Z. Guo and C. L. Brooks, III, *Biopolymers* **42**, 745–757 (1997).
135. P. W. Pan, H. L. Gordon, and S. M. Rothstein, *J. Chem. Phys.* **124**(2), 024905 (2006).
136. Z. Y. Guo and D. Thirumalai, *Biopolymers* **36**, 83–102 (1995).
137. Z. Guo, C. L. Brooks, III, and E. M. Boczko, *Proc. Natl. Acad. Sci. USA* **94**, 10161–10166 (1997).
138. A. Baba and T. Komatsuzaki, *Phys. Chem. Chem. Phys.*, **13**(4), 1395–1406 (2011).
139. P. Schuetz, R. Wuttke, B. Schuler, and A. Caflisch, *J. Phys. Chem. B*, **114**(46), 15227—15235 (2010).