Deviant effects in molecular reaction pathways

Michael S Samoilov & Adam P Arkin

In biological networks, any manifestations of behaviors substantially 'deviant' from the predictions of continuous-deterministic classical chemical kinetics (CCK) are typically ascribed to systems with complex dynamics and/or a small number of molecules. Here we show that in certain cases such restrictions are not obligatory for CCK to be largely incorrect. By systematically identifying properties that may cause significant divergences between CCK and the more accurate discrete-stochastic chemical master equation (CME) system descriptions, we comprehensively characterize potential CCK failure patterns in biological settings, including consequences of the assertion that CCK is closer to the 'mode' rather than the 'average' of stochastic reaction dynamics, as generally perceived. We demonstrate that mechanisms underlying such nonclassical effects can be very simple, are common in cellular networks and result in often unintuitive system behaviors. This highlights the importance of deviant effects in biotechnologically or biomedically relevant applications, and suggests some approaches to diagnosing them *in situ*.

Although the temporal evolution of a spatially homogeneous molecular system subject to a set of (bio)chemical reactions is often described macroscopically via the classical chemical kinetics (CCK) formalism through a set of deterministic differential equations (ODEs)[1]—otherwise referred to as 'reaction rate equations' or 'mass action kinetics'—this approach does not substantially reflect the fundamentally random and discrete nature of individual molecular interactions. The latter is well described by the chemical master equation (CME)[2–4]. Although this makes CME an intrinsically much more accurate description of biological or chemical molecular system dynamics, CCK is significantly more efficient computationally and could be viewed as derived from CME via a set of limiting approximations[5].

The conditions under which these approximations hold are often considered to be broadly valid for biological and/or simple chemical systems. However, as shown here, neither is guaranteed to be the case. Breakdowns in the limiting approximations can occur in some of the most basic processes owing to mechanisms that are particularly common in biological systems. These breakdowns then manifest themselves as various 'deviant nonclassical effects'—distinctive behaviors not accounted for by the classical molecular kinetic description—whereby the discrete and/or stochastic nature of reaction events conspires to drive the characteristic behavior of the system substantially away from that predicted by classical kinetics.

Notably, we demonstrate that although deviant effects may be stochastic in origin, this is by no means a requirement. The behavior of the system might be accurately characterized by a deterministic formalism, albeit not by CCK. In particular, although CCK dynamics is at times casually interpreted as describing the evolution of CME averages—this is only strictly accurate for purely linear (unimolecular and constant rate) reaction systems. For pathways that contain nonlinear and irreversible processes, it is more appropriate to view the deterministic CCK equations as describing—within a certain limit—the mode rather than the average behavior of the nondeterministic CME (see **Fig. 1** and Methods). This seemingly simple and technical difference is of utmost importance and may lead to unexpected qualitative differences between CCK and CME predictions. Therefore, any significant divergences between the system mode and average would then further reflect the potential deviant effects present and so can be used to diagnose possible nonclassical behaviors in practical applications.

RESULTS

Types of deviant effects

To comprehensively characterize deviant effects in molecular reaction pathways, we identify them by the corresponding approximation types (namely, by the types of limiting CME-to-CCK approximations, whose breakdowns manifest as corresponding nonclassical behaviors). **Table 1** lists the steps that appropriately yield CCK as a continuous-deterministic limit of the discrete-stochastic CME system description, thus elucidating the respective types of deviant effects.

For CCK to remain an accurate representation of the underlying system dynamics, all of these approximations have to remain valid. Any violations thereof represent fundamental failure modes of the CCK relative to the CME system description and will thus allow for deviant dynamics within the reaction network.

Although it has long been understood that the classical formulation may break down for very 'low CCK state' systems, that is, with $X \sim \sqrt{X}$ for some species[6,7], such 'discrete' effects exemplified by Type II approximations are not the only instances of potential problems with this description. Effects of Type I or III arise because of, respectively,

Howard Hughes Medical Institute, Department of Bioengineering, University of California at Berkeley, Center for Synthetic Biology, Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 977-257, Berkeley, California 94720, USA. Correspondence should be addressed to M.S.S. (mssamoilov@lbl.gov) or A.P.A. (aparkin@lbl.gov).

Published online 10 October 2006; doi:10.1038/nbt1253

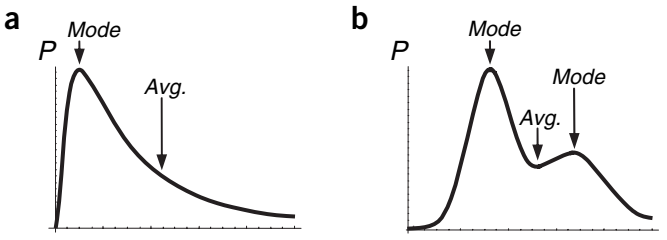

Figure 1 Characteristic CME distribution examples. Schematic representation of characteristic cases of possible instantaneous CME system state probability distributions and their effects on the fidelity of CCK description. (**a**) Because CCK is mode-based, it is most explanatory of strongly unimodal distributions—when the system is highly likely to be found in the vicinity of the probability peak. Note, the relative positions of mode and average could still be substantially different even for unimodal distributions if they are significantly skewed. (**b**) When the CME distribution is multimodal the situation is further complicated by the possibility of several distribution peaks having different probability mass. The latter could potentially drive the classical deterministic CCK description to become inaccurate or ill defined, for example, if formerly subdominant peaks grow to be more significant over time (**Fig. 3b**). Then the average may become the more representative and robust description of system behavior. (It may also be possible that neither the mode nor the average would offer an accurate representation of system dynamics, e.g., see **Fig. 5**.) Avg., average.

major divergences of the classical description from, or the inability of mode-based CCK to fully account for, all the information contained within the underlying CME distribution. These do not have any explicit state/size restrictions, which may thus allow deviant behaviors to conceivably be observed in otherwise normal 'high CCK state' systems. Of course, in many systems, Type I, II and III effects may be intermixed. The goal of this work is not an investigation of such interplay, but the identification and characterization of these behaviors at the level of their fundamental genesis, including implications for the underlying system

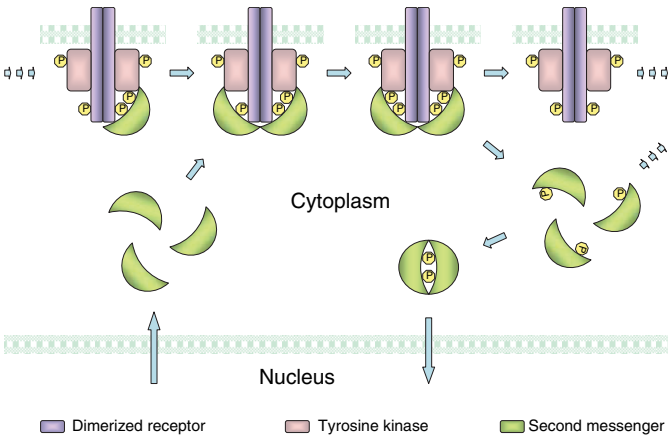

Figure 2 A signaling pathway motif. Upon dimerization, the kinase moiety of the receptor becomes activated. This allows a second messenger molecule to bind, which leads to its phosphorylation. The binding process is cooperative with the phosphorylated messenger molecules helping recruit the unphosphorylated ones to the receptor. Once activated, messenger molecules could form dimers, which are then transported into the nucleus where they may act as transcription regulators. The eventual dephosphorylation and dissociation of the dimers leads to the release of the unphosphorylated monomers back into the cytoplasm. The process continues until the pathway shuts down. (This instantiation of process (1) could be considered within the context of, for example, JAK-STAT signaling mechanism—or one of several other biomolecular pathways.)

Table 1 Classification of deviant effects

Effect	Underlying approximation
Type I	Full state probability distribution is approximated by its leading exponent, (*)
Type II	Characteristic system features occur on scales much greater than a molecule: a. Discrete molecular states are approximated by continuous concentrations, and (**) b. Combinatorial transition rates are approximated by leading power terms
Type III	The evolution of the state distribution is replaced by that of its mode, (***)

See Methods for more detail on the origins of referenced conditions.

properties and dynamics. Below we look at deliberately simple applications of this framework and demonstrate how nonclassical effects can indeed emerge in otherwise rather ordinary settings.

Practical applications

In practical biological systems applications, a number of indicators could serve to suggest the potential presence of deviant effects in a given molecular reaction pathway of interest. As discussed earlier, these may include such dynamic properties as significant deviations between the average and the mode behavior, high sensitivity to the single molecule–scale perturbations and low CCK levels of some constitutive species. Ultimately, however, the characteristic behavior of any individual biomolecular system essentially depends on the structural nature of its underlying biochemical processes *in situ*, such as the presence of bimolecular and irreversible reactions, which enable the possibility of deviant effects.

To make these notions more concrete, consider the signaling pathway depicted in **Figure 2**. This general biological motif appears in a variety of known processes, including the one derived from the proposed STAT transcription factor activation mechanism whereby the dimerized receptor, the tyrosine kinase and the second messenger correspond to the cytokine receptor, Janus kinase (JAK) and the signal transducer and activator of transcription (STAT), respectively[8–11]. Signaling is initiated by ligand binding to and subsequent dimerization of the receptor. The associated receptor kinase is then activated and the second messenger is cooperatively recruited by the phosphorylated messenger monomer to the kinase, where it is phosphorylated. Dimerization of the phosphorylated messenger (e.g., STAT-P) leads to the formation of the functional transcription factor, which is then transported into the nucleus, and dephosphorylated monomers are transported out[12]. Now, by letting X denote the phosphorylated and Y denote the unphosphorylated messenger molecules, we can systematically characterize the second messenger signaling process through its constitutive chemical reactions as an instantiation of system (1) with $n = 2$, $q = 0$, and $m = 2$:

$$X + Y \underset{k_{-C}}{\overset{k_{XY}}{\rightleftarrows}} C \xrightarrow{k_C} n\,X + q\,Z,$$
$$m\,X \underset{k_{-D}}{\overset{k_X}{\rightleftarrows}} D_m \xrightarrow{k_D} (\cdots)_1, \quad X \underset{k_c^X}{\overset{k_d^X}{\rightleftarrows}} (\cdots)_2, \qquad (1)$$
$$(n \geq 1, \quad q \geq 0, \quad m \geq 2)$$

where $(\cdots)_i$ denotes the product of the corresponding reaction. Notably, this and similar mechanisms are readily found in a wide assortment of natural settings, including, among others, the dynamic elements of branching-chain chemical reactions[1], cAMP-induced PKA control[13,14], simple models of transcriptional autoregulation ($n > 1$), Rep-family replication initiator protein regulation[15] and viral template replication,

for example, T7 phage[16,17]. It is thus remarkable that—although neither unique nor exhaustive in its ability to manifest a variety of nonclassical behaviors—the relatively simple molecular reaction process (1) turns out to encompass an array of biologically relevant exemplars that illustrate how deviant effects may arise in what appear to be rather ordinary systems. Some of these instantiations capable of characterizing individual deviant effect types are discussed next in more detail.

Type I. One simple instantiation of mechanism (1) is given in (2) where $n = 2$, $q = 0$, $m = 2$, $(\cdots)_1 = 2Y$ and $k^X_{d/c} \sim 0$ with intermediates C/D_2 in rapid-equilibrium, that is, $k_D << k_X$, k_{-D} and $k_C << k_{XY}$, k_{-C}, that is:

$$X + X \xrightarrow{k_1} Y + Y \qquad X + Y \xrightarrow{k_2} X + X \tag{2}$$

Classically, via CCK, this system is expected to have an unstable steady state at $x_1^{ss} = 0$ and a stable one at $x_2^{ss} = x_T k_2/(k_2 + 2k_1)$. However, CME analysis shows that the opposite is true. (This behavior is indeed consistent with an intuitive notion that if X stochastically runs out, it cannot be made again via (2); see **Supplementary Notes** online.) This shows that system (1) is indeed subject to strong deviant dynamics. Furthermore, it implies that any biomolecular pathway containing a reaction system similar to (1) would be capable of generating nonclassical divergences from CCK. For instance, in the case of the tyrosine kinase mechanism cited earlier, such deviant effects could manifest themselves by internally contributing to pathway shutdown rates in a manner independent of any external deactivation processes, for example, negative feedbacks, present in the system.

The full stationary CME ('nonclassical') probability distribution is a delta-function at zero state whereas its leading order exponent ('classical') approximation centers at the CCK steady state, which are separated on the system scale (**Fig. 3a**). This explicitly confirms that the nonclassical behavior here is directly attributable to Type I effects. **Figure 3b,c** also provides evidence and offers intuition as to the dynamic mechanisms of CCK description failure. Remarkably, although CCK predictions closely follow the mode of the system, as we expect, the average trajectory is not only manifestly different, but also affords the more accurate characterization of nonclassical system behavior thus representing a more useful alternative deterministic description of system dynamics deviant from CCK.

Type II. Unlike those of Type I, the nonclassical effects of Type II arise owing to the 'discrete' nature of biomolecular processes and may become significant under the more familiar low concentration (CCK state) conditions in some species or other system features sensitive to single-molecule scale. Note, however, that even though the underlying scale of these effects is typically very small, once such dynamics are present in one species, they may propagate to affect the behavior of other—including high concentration—species. For example, whereas genomic DNA and other genetic elements (e.g.,plasmids, transposons, viruses, etc.) are frequently present in single-digit copy number within or outside a cell, their molecular-scale variations can profoundly affect organism-scale behaviors, including nonclassical processes such as stochastic fate choice by temperate phage λ during cellular infection[18]. In addition, some of the best known molecular pathways, including many purely protein-mediated mechanisms, rely on similarly low-abundance species and are thus open to exhibiting deviant nonclassical behaviors. A single molecule of such agents as diphtheria toxin or enediynes can penetrate and kill a eukaryotic cell[19,20], whereas a typical *Escherichia coli* cell contains on the order of ten copies of scarce proteins (as opposed to up to 10^5 copies or more of the highly abundant ones)[21]. For instance, regulation of the lactose-sensing (uninduced) *lac* operon is controlled by only about ten molecules of *lac repressor* and even fewer molecules of β-galactosidase in wild-type *E. coli*[21].

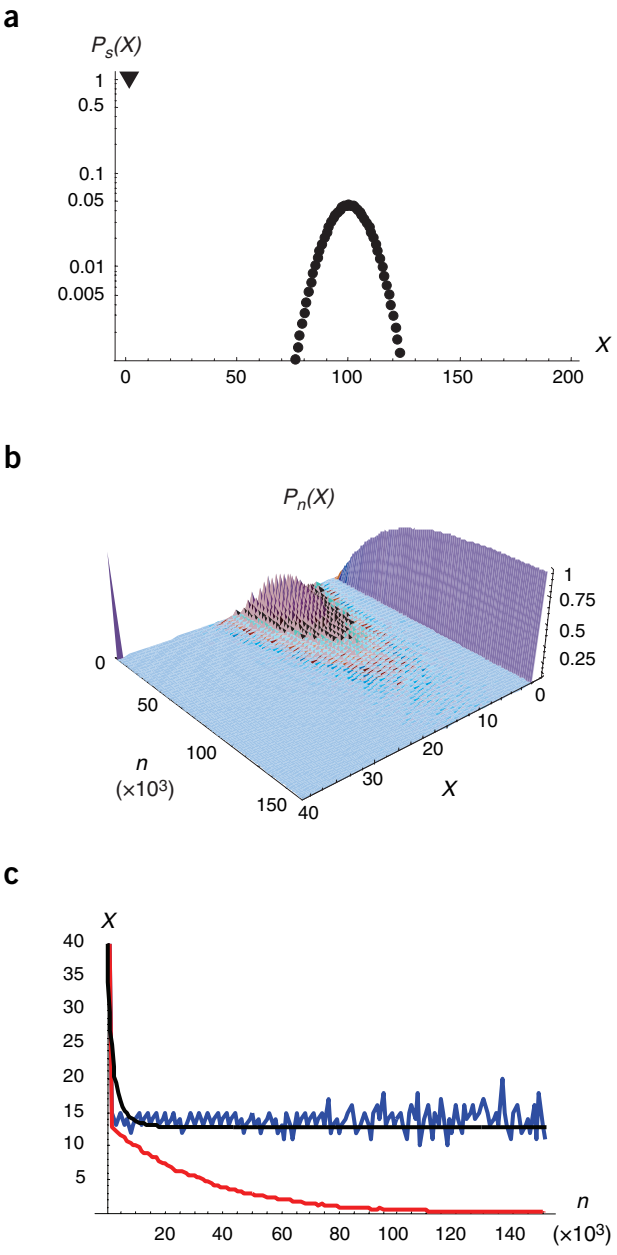

Figure 3 Type I system evolution. (**a**) Log plot of classical [•] and nonclassical [▼] stationary probability distributions for system (2) with $c_1 = 2$, $c_2 = 2$ and $X_T = 200$. (**b**) Temporal evolution of the exact probability distribution for system (2) with $c_1 = 2$, $c_2 = 0.5$ and $X(0) = 40$, $Y(0) = 20$ (the initial delta function is in the left-most corner) computed over 1,000 trajectories numerically simulated via the Gillespie Algorithm[5,49,50] (time is given in terms of reaction number, n). (**c**) Plot of the upper CME mode (blue), CME average (red) and CCK (black) trajectories for this system. Note that the molecular (discrete-stochastic) system quickly approaches the classical steady state by closely following the standard CCK dynamics, and then fluctuates around it in a way consistent with the classical stationary distribution. Yet, once the system stochastically hits a zero state—it remains there indefinitely, which causes the bulk of probability distribution mass to temporally switch from the initially prevalent peak at the classical solution, X_2^{ss}, to the ultimately dominant nonclassical one, X_1^{ss} (**Fig. 3b**). This leaves the CCK trajectory 'stuck' in the transient high-concentration peak, thus ultimately invalidating classical description.

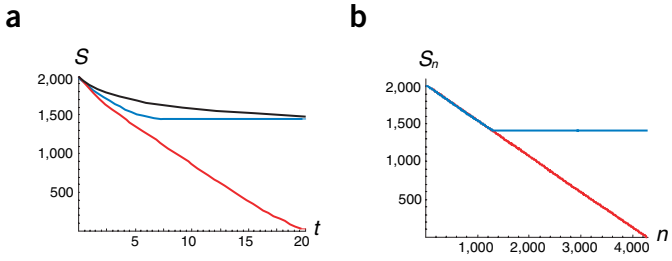

Figure 4 Type II system evolution. (**a**) Average CME trajectories in the same system for 100 simulations (red and blue) versus CCK results (black). (**b**) Results of the exact numerical simulations for system (3). The plot shows the evolution of substrate abundance versus reaction number with $c_1 = 50$, $c_2 = 5$, $c_3 = 100$, $c_d = 4{\cdot}10^5$ and $S_0 = 2{,}000$, $C_0 = P_0 = 0$, where blue trajectory corresponds to $E_0 = 100$ and red to $E_0 = 99$.

Furthermore, as noted earlier, the Type II discrete effects may not even require the low molecular copy numbers *per se* to determine the characteristic behavior of the system. For instance, consider a basic Michaelis-Menten mechanism augmented by a binary enzyme inactivation process:

$$S + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} C \xrightarrow{k_3} P + E$$
$$E + E \xrightarrow{k_d} (\cdots). \qquad (3)$$

This instantiation of system (1)—with $X = E$, $Y = S$, $Z = P$, $n = 1$, $q = 1$, $m = 2$ and $k^X_{d/c} \sim 0$ with the intermediate D_2 in rapid-equilibrium—would classically predict a monotonic decay of substrate that continues while the enzyme concentration approaches zero (**Fig. 4a**, black curve).

Nonclassically the behavior is different. If the number of enzyme molecules is even, the reaction proceeds until the enzyme is exhausted, at which point the process shuts down (**Fig. 4b**, blue curve). Conversely, if the number is odd, the bimolecular reaction is unable to remove the last molecule from the system, which then causes the process to continue on indefinitely (or until the substrate runs out, **Fig. 4b**, red curve). This difference provides a simple example of a nonclassical Type II effect driven by variations on the single-molecule scale, that is, odd or even parity state of the enzyme, but manifested in the high-abundance species such as substrate. Remarkably, the average trajectory still accurately reflects this system's behavior regardless of initial conditions, with a combination of bimolecular and irreversible reactions again necessary to generate the characteristic nonclassical kinetics (**Fig. 4a**).

Finally, such deviant effects may indeed be considered 'discrete' rather than 'stochastic' in the conventional sense, as here the characteristic behavior of the system (that is, whether to stop or continue on indefinitely) is essentially deterministic—controlled by the parity of the initial state and the discrete character of individual reactions regardless of their stochastic nature or the total number of molecules present (see **Supplementary Notes** for additional detail.)

Type III. These deviant effects may manifest themselves when the characteristic behavior of a system can be significantly affected purely by its underlying ability to sample states away from the CCK trajectory (which, as discussed earlier, closely follows the mode of the system state distribution; see also **Fig. 1**). This could be contrasted with the influence of discreteness in the underlying molecular states, Type II, or the approximations to the probability distribution itself, Type I, effects.

A basic enzymatic futile cycle—a ubiquitous regulatory motif in a wide variety of biological molecular pathways—can, among others, be subject to Type III effects (**Fig. 5**). Namely, the futile cycle always exhibits monostable substrate/product levels if stationary behavior of the enzymes is described using CCK (distribution modes). However, if it is examined using nonclassical analysis (whether Langevin or CME), this mechanism could be observed to display stochastic oscillations in product/substrate levels—doing so without any additional feedback/forward mechanisms and manifestly due to the enzyme activity being stochastically governed by a (unimodal) probability distribution, rather than fixed at the mode[22].

DISCUSSION

The ongoing development of systems and computational biology as well as advancements in high-throughput empirical and quantitative techniques have led to the proliferation of modeling tools targeted at biotechnological and biomedical applications. This increased reliance on *in silico* methods in many practical endeavors has made the questions of their relative accuracy as well as computational expediency increasingly more relevant. These issues are particularly important when choosing between the CCK methods of analysis and those based on the CME description of molecular processes. A needless choice of CME could lead to a computationally intractable implementation, whereas an inappropriate choice of CCK may result in the appearance of unaccounted for 'deviant' effects and corresponding misprediction or misinterpretation

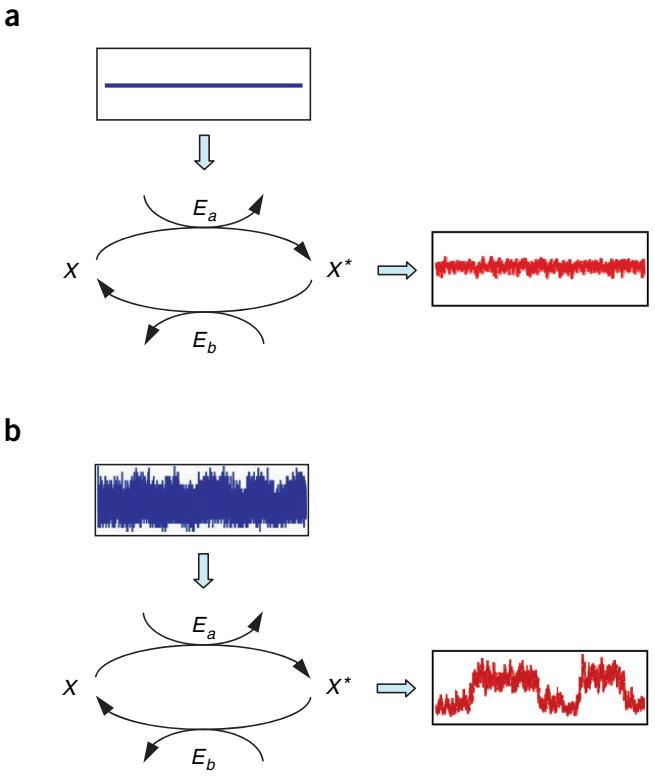

Figure 5 Type III deviant behavior. An enzymatic futile cycle is a mechanism capable of manifesting a Type III effect. (**a**) If the forward (or the reverse) enzyme is considered deterministically using its stationary mode or CCK, the dynamic behavior of the product (or substrate) is always monostable. (**b**) However, if in the same mechanism the forward enzyme behavior is considered via a full (unimodal) probability distribution, rather than just its mode, the dynamics of the product can undergo a noise-induced bifurcation with the system becoming bistable, that is, exhibiting stochastic switching/oscillations. (See ref. 22 for additional detail.)

of system behavior. Thus, a practical question arises: given its computational simplicity and global recognition, under what conditions can CCK be safely used in biological applications, and when does the possibility of deviant effects require the more accurate but more demanding CME analysis?

In this work we offer a comprehensive rigorous framework for characterizing the potential divergences between CCK and CME system descriptions, with three key approximations being the sources of basic deviant effects. By further asserting CCK as the mode and not the average of CME solution, we are able to suggest practical approaches for diagnosing the presence of nonclassical system behaviors. It is not good enough to simply use the rule of thumb that slow reactions with low CCK states are the sole reasons to consider using CME, as deviant effects could arise in high CCK states as well. Finally, we also show that certain characteristics of biological networks, such as widespread coupling of nonlinear and irreversible reaction mechanisms, make them particularly amenable to nonclassical effects. These characteristics are common throughout many biotechnologically and biomedically relevant systems, such as transcription/translation and Michaelian enzyme processes, because the nonlinearities they introduce are frequently required to effect the complex dynamics and robust regulation observed[13,23–26].

Taking these results together, we have suggested a practical approach for identifying deviant effects in biological molecular systems by explicitly considering some of the basic reaction mechanisms that happen to be well represented within biochemical networks *in situ*. Importantly, as local dynamics can then propagate through biomolecular networks—for example, via signal transduction pathways[26–31], potentially expressing themselves in species that are otherwise directly involved only in strictly classical kinetics—the presence of just one such mechanism within a larger biological process may lead to a variety of system-wide nonclassical behavior modes[22,32–40]. Given such potentially global implications of locally deviant pathway dynamics, we hope that this work may offer biologists involved with kinetic analysis or molecular modeling additional tools and intuition to help decide whether their work requires the use of discrete and stochastic chemical master equation–based methods or whether the simpler classical chemical kinetics framework is sufficient.

METHODS

The dynamics of general biomolecular and chemical reactions are classically described by the CCK formalism, whereby the temporal relaxation of a molecular system from an initial state $\vec{X}_0$ to a steady state (or curve) $\vec{X}_{ss}$ (referred to as its "trajectory")—under appropriate conditions and subject to a set of transformations

$$\nu^s_{1j} X_1 + \nu^s_{2j} X_2 + \nu^s_{3j} X_3 + \ldots \xrightarrow{k_j} \nu^f_{1j} X_1 + \nu^f_{2j} X_2 + \nu^f_{3j} X_3 + \ldots \quad (4)$$

that convert a set of $\{\nu^s_{ij}\}$ molecules into $\{\nu^f_{ij}\}$ molecules of $\{X_i\}$ with a macroscopic reaction rate constant k_j—is modeled by the differential equation system of the form:[1]

$$\frac{dx_i}{dt} = \sum_j (\nu^f_{ij} - \nu^s_{ij})\, k_j \prod_k x_k^{\nu^s_{kj}} \quad (5)$$

where $x_i \sim X_i/\Omega$ is the species concentration in volume Ω.

However, the CCK formalism, equation 5, expressly disregards the ultimately discrete nature of system states and stochastic nature of transitions that govern individual reaction events. A more comprehensive and accurate description ought to include these properties in its representation of the system dynamics. Such nondeterministic description is provided by the chemical master equation (CME)[1–3],

$$\frac{\partial P(\vec{X};t)}{\partial t} = \sum_j \left[W(\vec{X} - \vec{r}_j; \vec{r}_j) P(\vec{X} - \vec{r}_j; t) - W(\vec{X}; \vec{r}_j) P(\vec{X}; t) \right] \quad (6)$$

where $\vec{X}$ is a vector of the number of molecules for each chemical species, $\vec{r}_j$ is the change in $\vec{X}$ due to the occurrence of the j^{th} reaction, $P(\vec{X}; t)$ is the state probability distribution at time t and the $W(\vec{X}; \vec{r}_j)$'s are instantaneous transition probability rates for the reaction $\vec{X} \to \vec{X} + \vec{r}_j$. In the case of chemical processes:

$$W(\vec{X}; \vec{r}_j) = c_j \prod_k \frac{X_k!}{(X_k - \nu^s_{kj})!\nu^s_{kj}!} \quad (7)$$

where $r_{ij} = \nu^f_{ij} - \nu^s_{ij}$ and c_j's correspond to the individual microscopic reaction event rates in process (4). That is, these expressions indeed capture both the stochastic and discrete aspects of the underlying processes at the molecular level.

Deterministic CCK limit of CME. Although a number of methods can be employed in analytically relating the master equation to a certain deterministic limit of some underlying system dynamics[41–44], our goal of practically categorizing the nature of the approximations required to obtain the CCK limit of the CME representation may be accomplished using the approach of Kubo *et al.*[45–48] There, it was shown that a general CME solution with deterministic initial conditions can be represented as:

$$P(\vec{x}, t) = C \exp\left(-\Omega\left[s(\vec{x}, t) + O(1/\Omega)\right]\right) \quad (*)$$

where $s(\vec{x},t)$ is the leading order coefficient in the effective volume, Ω, expansion of $P(\vec{x},t)$. If we consider a continuous extension of (*) with respect to $\vec{x}$ in the large system limit, then substituting (*) into equation 6 and expanding in powers of Ω—while recalling that for natural molecular systems only $\nu \le 2$ reactions are relevant—yields to $O(\Omega^0)$,

$$\partial_t s(\vec{x}, t) = -\sum_{j=1}^{R} w_j(\vec{x}) \left(e^{\vec{r}_j \cdot \vec{\nabla} s(\vec{x}, t)} - 1 \right) \quad (**)$$

with $W(\vec{X}; \vec{r}_j) = \Omega w_j(\vec{x}) + O(\Omega^0)$ where $w_j(\vec{x}) = k_j \prod_k x_k^{\nu^s_{kj}}$. This is, in fact, a Hamilton-Jacobi equation whose most probable trajectory is:

$$\frac{d\vec{x}}{dt} = \sum_{j=1}^{R} \vec{r}_j w_j(\vec{x}) e^{\vec{r}_j \cdot \vec{\nabla} s(\vec{x}, t)}; \qquad \frac{d\vec{\nabla} s(\vec{x}, t)}{dt} = \sum_{j=1}^{R} \vec{\nabla} w_j(\vec{x}) \left(1 - e^{\vec{r}_j \cdot \vec{\nabla} s(\vec{x}, t)} \right) \quad (***)$$

Notice that if the trajectory starts in the $\vec{\nabla} s(\vec{x};t) = 0$ state—that is at the distribution mode, as is the case for the deterministic initial conditions—it will continue in that state all along the trajectory:

$$\frac{d\vec{x}}{dt} = \sum_{j=1}^{R} \vec{r}_j w_j(\vec{x}) \quad (***)$$

which is indeed the CCK evolution equation 5.

Deterministic mode versus average trajectory. From the above, it is clear that the CCK equations, given in (5) more closely describe the mode trajectory in the appropriate limit of the underlying CME probability distribution rather than its average. Alternatively, the actual average behavior of the system can be written down directly from equation 6 as a different set of deterministic ODEs, namely:

$$\frac{d}{dt} \langle \vec{x} \rangle = \sum_{j=1}^{R} \vec{r}_j \langle w_j(x) \rangle \quad (8)$$

Although the two trajectory types, the average and the mode, become the same for linear systems, they will generally be described by different equations if bimolecular or higher order reactions are present. (High-order reactions yield $w_j(x)$ terms of order x^n in (***), with the corresponding term in equation 8 proportional to $\langle x^n \rangle$. However, for $n > 1$: $\langle x^n \rangle \neq \langle x \rangle^n$, in general, and therefore comparing (***) and equation 8 shows that $x(t) \neq \langle x \rangle(t)$ likewise.) That is, although the correspondence between the mode and average system behavior may be close for certain conditions, in pathways where nonlinear reactions are present this congruity is by no means guaranteed. (For instance, the average of some strongly skewed or bimodal distribution will not be representative of its mode behavior (**Fig. 1**).) In fact, this work provides some explicit examples of chemical reaction mechanism where such mode (CCK) and average (CME) system behaviors diverge rather dramatically.

Note: Supplementary information is available on the Nature Biotechnology website.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Berry, R.S., Rice, S.A. & Ross, J. *Physical chemistry*, edn. 2 (Oxford University Press, New York; 2000).
2. Gillespie, D.T. A rigorous derivation of the chemical master equation. *Physica A (Amsterdam)* **188**, 404–425 (1992).
3. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam; 1992).
4. Gillespie, D.T. *Markov Processes* (Academic Press, Boston; 1992).
5. Gillespie, D.T. in *Handbook of Materials Modeling* (ed. Yip, S.) 1735–1752 (Springer, Netherlands; 2005).
6. McQuarrie, D.A., Jachimowski, C.J. & Russell, M.E. Kinetics of small systems. II. *J. Chem. Phys.* **40**, 2914–2921 (1964).
7. Zheng, Q. & Ross, J. Comparison of deterministic and stochastic kinetics for nonlinear systems. *J. Chem. Phys.* **94**, 3644–3648 (1991).
8. Behrmann, I. *et al.* A single STAT recruitment module in a chimeric cytokine receptor complex is sufficient for STAT activation. *J. Biol. Chem.* **272**, 5269–5274 (1997).
9. Levy, D.E. & Darnell, J.E., Jr. Stats: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol.* **3**, 651–662 (2002).
10. Kisseleva, T., Bhattacharya, S., Braunstein, J. & Schindler, C.W. Signaling through the JAK/STAT pathway, recent advances and future challenges. *Gene* **285**, 1–24 (2002).
11. Yamaoka, K. *et al.* The Janus kinases (Jaks). *Genome Biol.* **5**, 253 (2004).
12. McBride, K.M., McDonald, C. & Reich, N.C. Nuclear export signal located within the DNA-binding domain of the STAT1 transcription factor. *EMBO J.* **19**, 6196–6206 (2000).
13. Voet, D., Voet, J.G. & Pratt, C.W. *Fundamentals of Biochemistry* (Wiley, New York; 1999).
14. Zhang, X. *et al.* Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proc. Natl. Acad. Sci. USA* **102**, 4459–4464 (2005).
15. Giraldo, R. & Fernandez-Tresguerres, M.E. Twenty years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids. *Plasmid* **52**, 69–83 (2004).
16. Endy, D., Kong, D. & Yin, J. Intracellular kinetics of a growing virus: a genetically structured simulation for bacteriophage T7. *Biotechnol. Bioeng.* **55**, 375–389 (1997).
17. Srivastava, R., You, L., Summers, J. & Yin, J. Stochastic vs. deterministic modeling of intracellular viral kinetics. *J. Theor. Biol.* **218**, 309–321 (2002).
18. Arkin, A., Ross, J. & McAdams, H.H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648 (1998).
19. Clardy, J. & Walsh, C. Lessons from natural molecules. *Nature* **432**, 829–837 (2004).
20. Salyers, A.A. & Whitt, D.D. *Bacterial Pathogenesis: A Molecular Approach*, edn. 2 (ASM Press, Washington, D.C.; 2002).
21. Stryer, L. *Biochemistry*, edn. 4 (W.H. Freeman, New York; 1995).
22. Samoilov, M., Plyasunov, S. & Arkin, A.P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. USA* **102**, 2310–2315 (2005).
23. Stelling, J., Sauer, U., Szallasi, Z., Doyle, F.J., III & Doyle, J. Robustness of cellular functions. *Cell* **118**, 675–685 (2004).
24. Simpson, M.L., Cox, C.D., Peterson, G.D. & Sayler, G.S. Engineering in the biological substrate: Information processing in genetic circuits. *Proc. IEEE* **92**, 848–863 (2004).
25. Marhl, M., Perc, M. & Schuster, S. Selective regulation of cellular processes via protein cascades acting as band-pass filters for time-limited oscillations. *FEBS Lett.* **579**, 5461–5465 (2005).
26. Samoilov, M., Arkin, A. & Ross, J. Signal processing by simple chemical systems. *J. Phys. Chem. A* **106**, 10205–10221 (2002).
27. Ting, A.Y. & Endy, D. Signal transduction. Decoding NF-kappaB signaling. *Science* **298**, 1189–1190 (2002).
28. Xiong, W. & Ferrell, J.E. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460–465 (2003).
29. Pedraza, J.M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).
30. Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. & Elowitz, M.B. Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
31. Guido, N.J. *et al.* A bottom-up approach to gene regulation. *Nature* **439**, 856–860 (2006).
32. Davis, K.L. & Roussel, M.R. Optimal observability of sustained stochastic competitive inhibition oscillations at organellar volumes. *FEBS J.* **273**, 84–95 (2006).
33. Blake, W.J., Kærn, M., Cantor, C.R. & Collins, J.J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
34. Qian, H., Saffarian, S. & Elson, E.L. Concentration fluctuations in a mesoscopic oscillating chemical reaction system. *Proc. Natl. Acad. Sci. USA* **99**, 10376–10381 (2002).
35. Raser, J.M. & O'Shea, E.K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
36. Paulsson, J., Berg, O.G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA* **97**, 7148–7153 (2000).
37. Gillespie, D.T. & Mangel, M. Conditioned averages in chemical kinetics. *J. Chem. Phys.* **75**, 704–709 (1981).
38. Wolf, D.M., Vazirani, V.V. & Arkin, A.P. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J. Theor. Biol.* **234**, 227–253 (2005).
39. English, B.P. *et al.* Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat. Chem. Biol.* **2**, 87–94 (2006).
40. Newman, J.R.S. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846 (2006).
41. Gillespie, D.T. The chemical Langevin equation. *J. Chem. Phys.* **113**, 297–306 (2000).
42. Kurtz, T.G. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57**, 2976–2978 (1971).
43. Van Kampen, N.G. The expansion of the master equation. *Adv. Chem. Phys.* **34**, 245–309 (1976).
44. Gitterman, M. & Weiss, G.H. Some comments on approximations to the master equation. *Physica A (Amsterdam)* **170**, 503–510 (1991).
45. Kubo, R., Matsuo, K. & Kitahara, K. Fluctuation and relaxation of macrovariables. *J. Stat. Phys.* **9**, 51–96 (1973).
46. Kitahara, K. The Hamilton-Jacobi-equation approach to fluctuation phenomena. *Adv. Chem. Phys.* **29**, 85–111 (1973).
47. Dykman, M.I., Mori, E., Ross, J. & Hunt, P.M. Large fluctuations and optimal paths in chemical kinetics. *J. Chem. Phys.* **100**, 5735–5750 (1994).
48. Samoilov, M. & Ross, J. One-dimensional chemical master equation: uniqueness and analytical form of certain solutions. *J. Chem. Phys.* **102**, 7983–7987 (1995).
49. Cao, Y., Li, H. & Petzold, L. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *J. Chem. Phys.* **121**, 4059–4067 (2004).
50. Cao, Y., Petzold, L.R., Rathinam, M. & Gillespie, D.T. The numerical stability of leaping methods for stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **121**, 12169–12178 (2004).