Nonequilibrium Thermodynamics of Biochemical Machines:

How many ATP's does it cost to run a functional biochemical circuit?

Yuhai Tu
IBM T. J. Watson Research Center, Yorktown Heights, NY

Collaborators:
Ganhui Lan (GWU)
Pablo Sartori (IGC, Portugal/IAS, Princeton)
Victor Sourjik (MPI, Marburg)
Yuansheng Cao (PKU/UCSD)
Chenyi Fei (PKU/Princeton)
Qiwei Yu (PKU)
Dongliang Zhang (PKU)
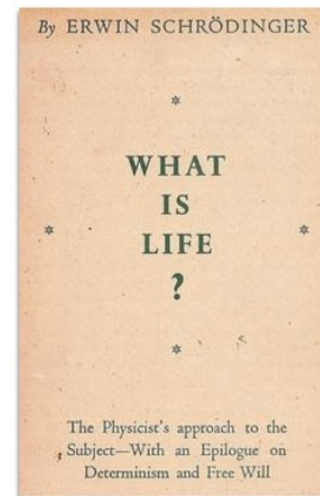Qi Ouyang (PKU)

Special Thanks:
Howard Berg (Harvard)

(The 16th Granada Seminar, Day-7, 06/17/2021)

Schrödinger's Question

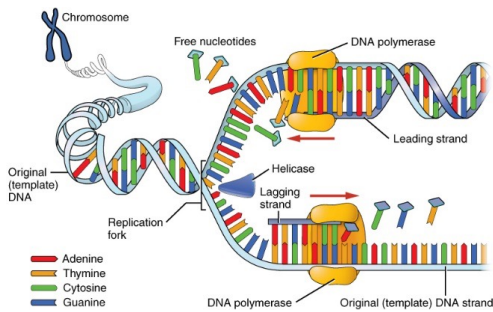**How are living (biological) systems different from non-living (physical) systems?**
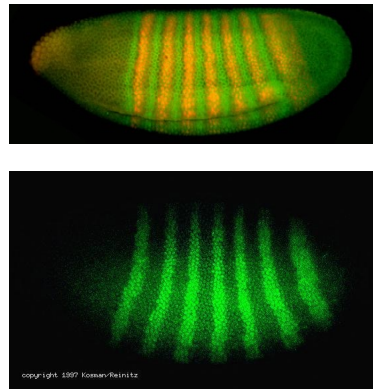

(1944)

"…life feeds on negative entropy…" – Erwin Schrodinger
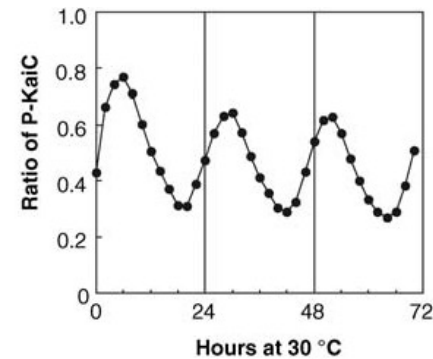
Life costs free energy

Central Question: How do living systems process information accurately with noisy components and stochastic reactions?
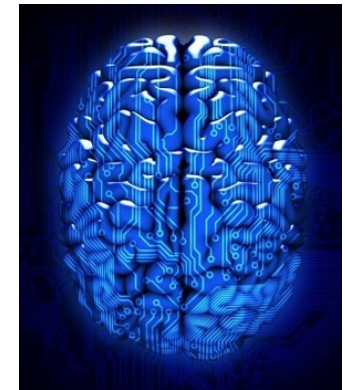

DNA Replication

Pattern Formation (drosophila (fruit fly) embryo)

Circadian rhythm (Cyanobacteria)

Learning/memory ...

- Biological systems are noisy – small # of molecules, stochastic reactions, thermal fluctuation (room temperature), noisy neuron firing patterns,….

(1) How? → Molecular Mechanism and Design Principles

(2) How much? → Energy Cost for Biological Functions

Cost of information processing: Landauer's Principle

Irreversibility and Heat Generation in the Computing Process -- IBM Journal Res. Dev., July 1961

Rolf Landauer

Landauer's Principle, motivated by Von Neumann, Rolf Landauer shows that "any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information bearing degrees of freedom of the information processing apparatus or its environment".

Each bit of lost information will cost a minimum of $kT\ln(2)$ of free energy.
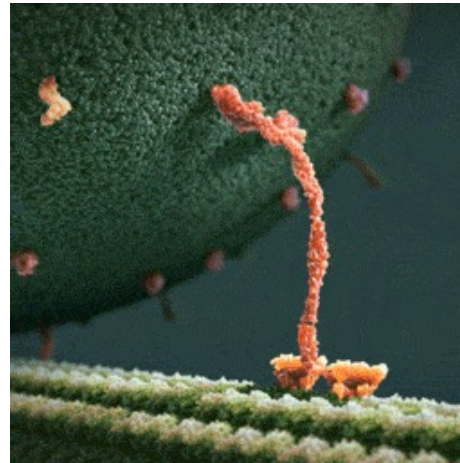
John Von Neumann

What about the energy cost of information processing in biological systems?

How is free energy used in living systems?
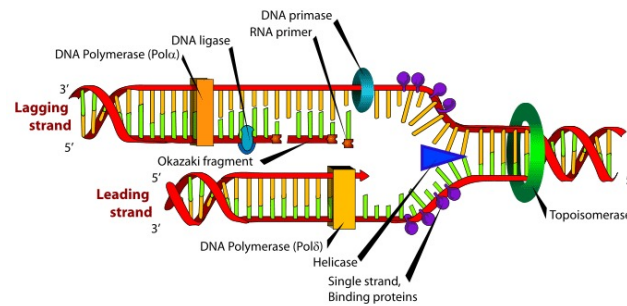
- Molecular
  Motors

Free
energy ↓

Mechanical work
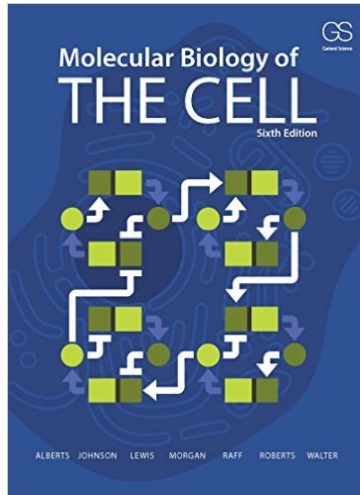

Kinesin walking
on microtubule
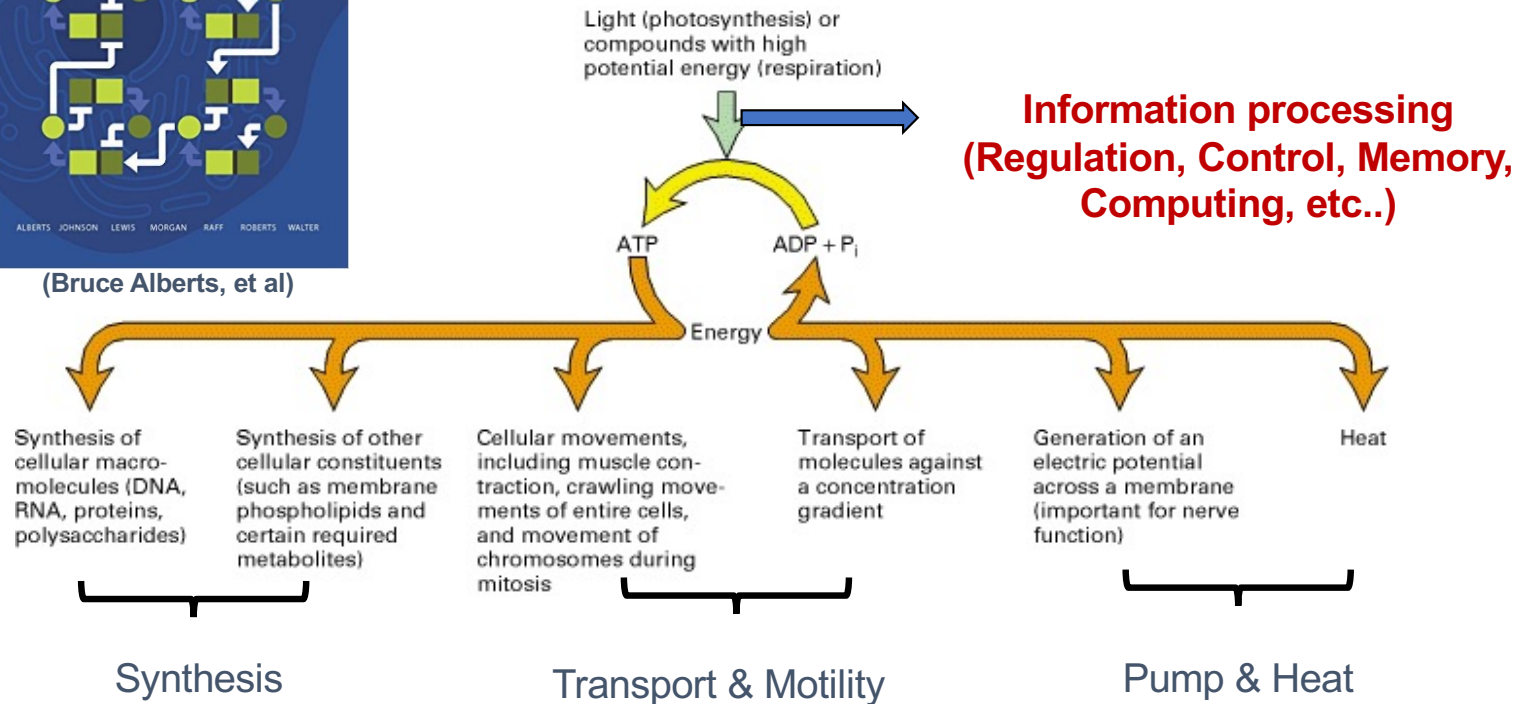
- Biological
  Synthesis

Free
energy ↓

Chemical bonds

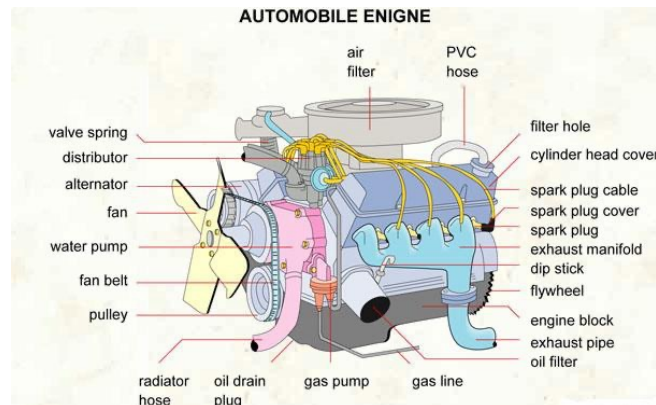# Different uses of energy in living systems

(Bruce Alberts, et al)
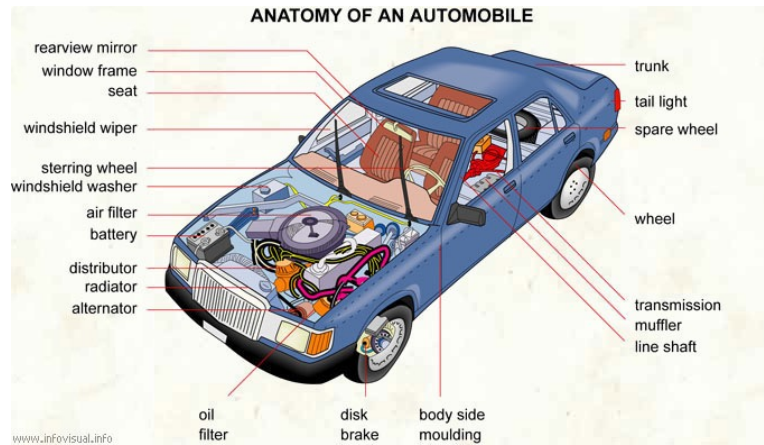Molecular Biology of THE CELL Sixth Edition
ALBERTS JOHNSON LEWIS MORGAN RAFF ROBERTS WALTER

ATP is the universal "energy currency",
which drives the **non-equilibrium processes**
of living systems.

Light (photosynthesis) or compounds with high potential energy (respiration)

**Information processing
(Regulation, Control, Memory,
Computing, etc..)**

ATP   ADP + P_i

Energy

Synthesis of cellular macro-molecules (DNA, RNA, proteins, polysaccharides)

Synthesis of other cellular constituents (such as membrane phospholipids and certain required metabolites)

Cellular movements, including muscle contraction, crawling movements of entire cells, and movement of chromosomes during mitosis

Transport of molecules against a concentration gradient

Generation of an electric potential across a membrane (important for nerve function)

Heat

Synthesis   Transport & Motility   Pump & Heat

Engineered systems

Designed to consume energy to carry out desired work

Mechanical Motion

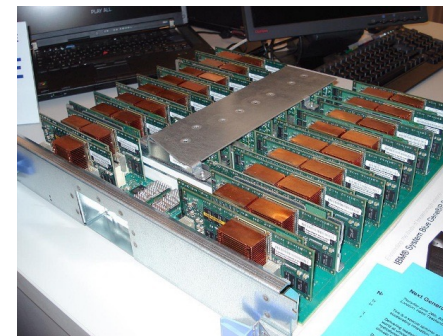ANATOMY OF AN AUTOMOBILE

AUTOMOBILE ENIGNE

Computing

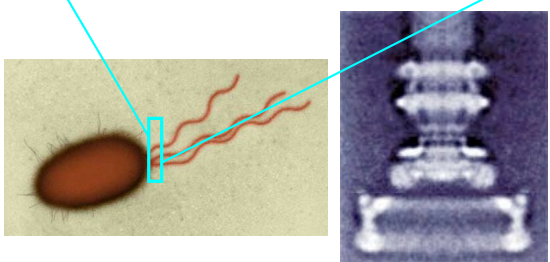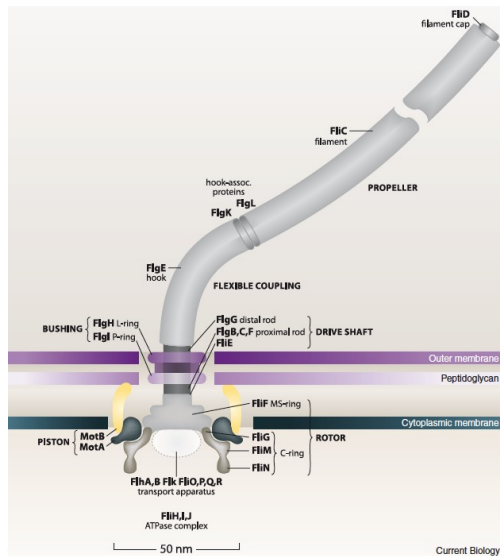**IBM BlueGene Supercomputer**

BlueGene Chips

5% of all energy consumption in the U.S. goes just to running computers --10/2018
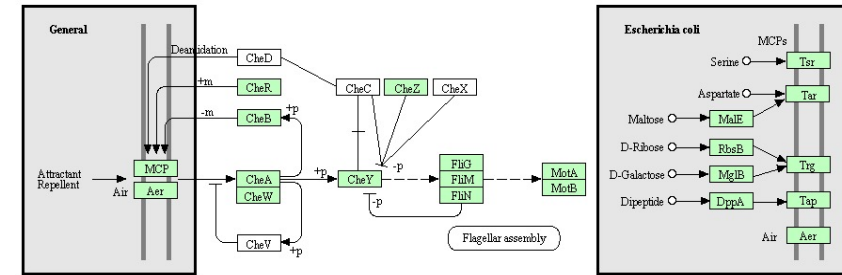
Biological machines

Evolved to consume energy to carry out desired functions
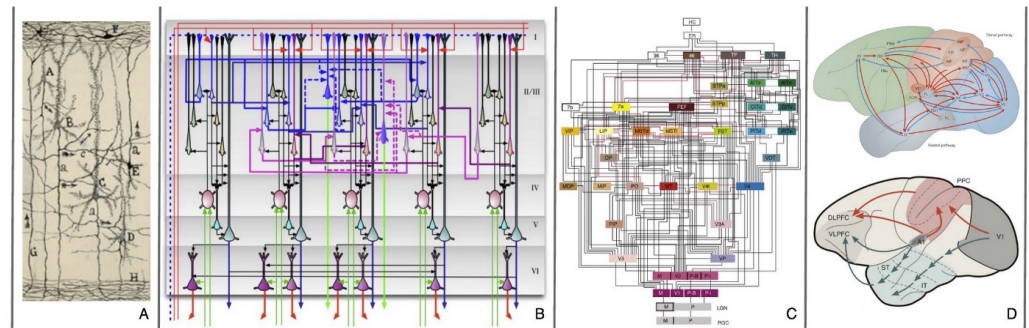
Nano-BioMotor
(Mechanical machine)

Information Processor
(Computing machine)

- **Biochemical network**

- **Neural network**

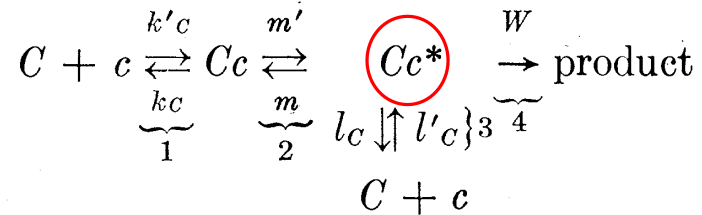**What's the energy cost of running these biochemical circuits?**

An earlier example: the kinetic proof-reading mechanism

Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity

(protein synthesis/DNA replication/amino-acid recognition)

Proc. Nat. Acad. Sci. USA
Vol. 71, No. 10, pp. 4135–4139, October 1974

$$C + c \underset{kc}{\overset{k'c}{\rightleftarrows}} Cc \underset{m}{\overset{m'}{\rightleftarrows}} \boxed{Cc*} \overset{W}{\rightarrow} \text{product}$$

$$\underbrace{}_{1} \quad \underbrace{}_{2} \quad l_C \Updownarrow l'_C \}_3 \overbrace{}^{4}$$

$$C + c$$

$p \rightarrow p^2$

$10^{-3} \rightarrow 10^{-6}$

Jacque Ninio & John Hopfield

DISSIPATION-ERROR TRADEOFF IN PROOFREADING

BioSystems, 11 (1979) 85—91

Thermodynamics of kinetic proofreading
free energy dissipation (cost)

Charlie Bennett

The Onsager Reciprocal Relationships, Detailed Balance, and the Cycle Rule

RECIPROCAL RELATIONS IN IRREVERSIBLE PROCESSES. I.

By Lars Onsager

$$J_1 = L_{11}X_1 + L_{12}X_2 + L_{13}X_3$$
$$J_2 = L_{21}X_1 + L_{22}X_2 + L_{23}X_3 \qquad (2.3)$$
$$J_3 = L_{31}X_1 + L_{32}X_2 + L_{33}X_3$$

The Onsager reciprocal relations
$$L_{12} = L_{21} \; ; \; L_{13} = L_{31} \; ; \; L_{23} = L_{32}. \qquad (2.4)$$

(Lars Onsager)

3. Analogy with Chemical Reactions

We shall compare (2.3) with the equations for a chemical monomolecular triangle reaction. Suppose that a certain substance may exist in a homogene-
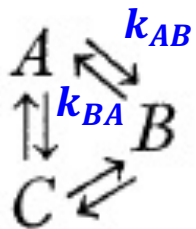
Detailed balance

The cycle rule

$$k_{AC}k_{CB}k_{BA} = k_{AB}k_{BC}k_{CA}.$$

k_{AB}

k_{BA}

$$A \rightleftarrows B \rightleftarrows C \qquad (3.1)$$

$$k_{BA}\bar{n}_A = k_{AB}\bar{n}_B$$
$$k_{CB}\bar{n}_B = k_{BC}\bar{n}_C \qquad (3.5)$$
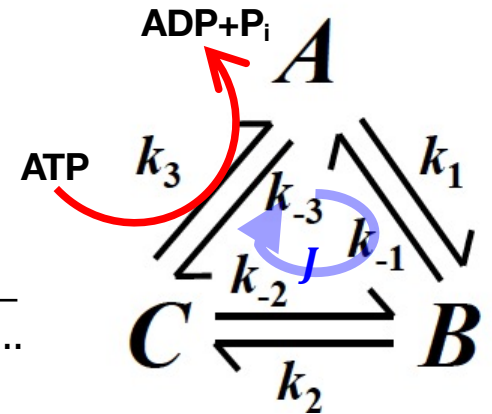$$k_{AC}\bar{n}_C = k_{CA}\bar{n}_A.$$

Biochemical systems are far from equilibrium

The cycle rule (or DB)
is broken:

$$k_1 k_2 k_3 \neq k_{-1} k_{-2} k_{-3}$$

Persistent current J
in steady state:

$$J = k_2 P_B - k_{-2} P_C = \frac{k_1 k_2 k_3 - k_{-1} k_{-2} k_{-3}}{...sum\ of\ 9\ (>0)\ terms\ ...}$$
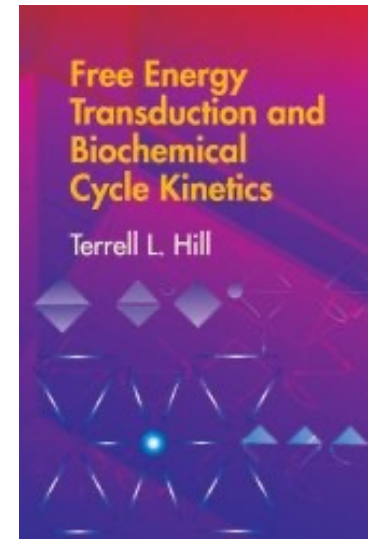
Thermodynamic force:

$$\Delta\mu = k_B T \ln \frac{k_1 k_2 k_3}{k_{-1} k_{-2} k_{-3}}$$

Free energy dissipation rate:

$$\dot{W} = J \times \Delta\mu \geq 0$$

Continuous energy dissipation (power consumption) is needed
to maintain a non-equilibrium steady state

ADP+P$_i$

ATP

Terrell Hill

Free Energy
Transduction and
Biochemical
Cycle Kinetics

Terrell L. Hill

How living systems use energy to fight the 2nd law of thermodynamics

to create biological order or maintain biological functions

I. Ultra-sensitive biochemical switch
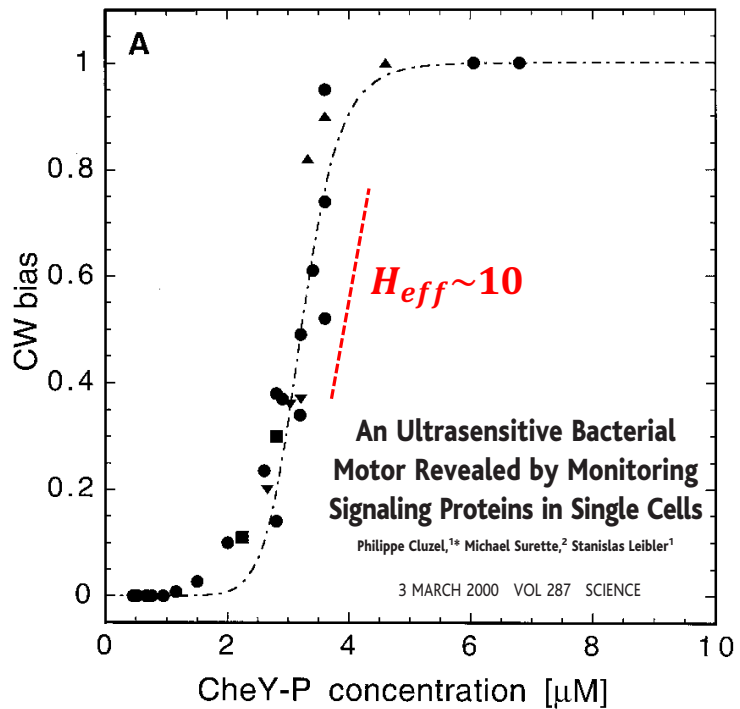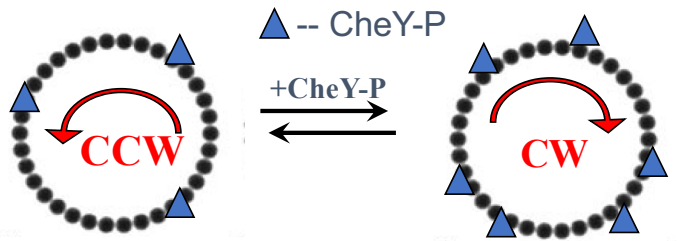
II. Sensory adaptation

III. Biochemical oscillations

IV. Synchronization

V. Pattern formation

.

.

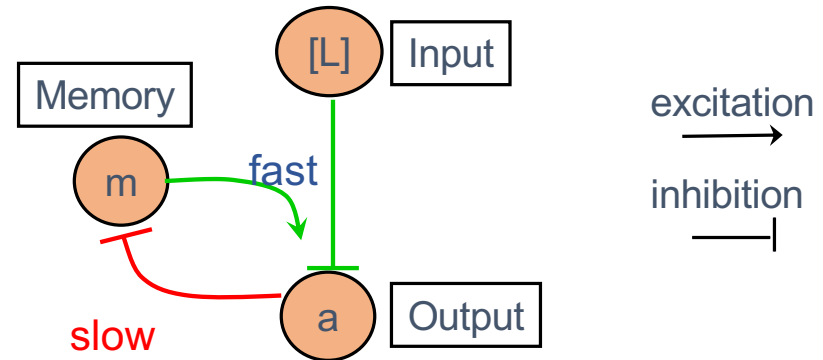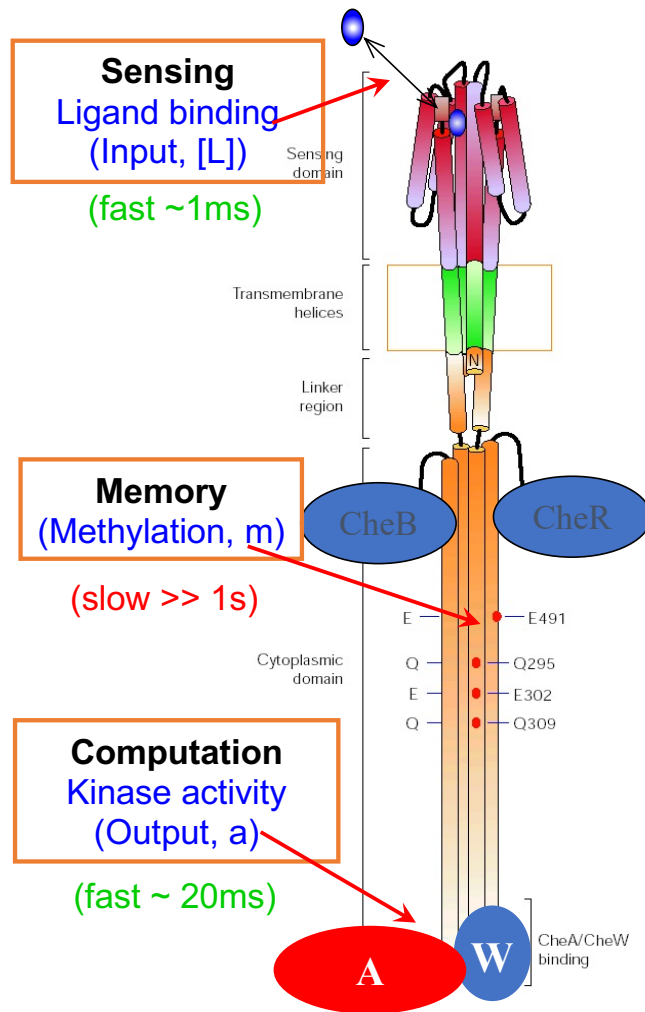# I. The energy cost of ultra-sensitivity in the bacterial flagellar motor switch

▲ -- CheY-P

+CheY-P

CCW ⇌ CW

A

CW bias

$H_{eff} \sim 10$

An Ultrasensitive Bacterial Motor Revealed by Monitoring Signaling Proteins in Single Cells

Philippe Cluzel,[1]* Michael Surette,[2] Stanislas Leibler[1]

3 MARCH 2000   VOL 287   SCIENCE

CheY-P concentration [μM]

The mechanism is non-equilibrium (DB broken)
and a higher sensitivity costs more energy

H_{eff}

$\Delta W/k_B T$

● — ● Δn=16
■ — ■ Δn=10
▲ — ▲ Δn=4
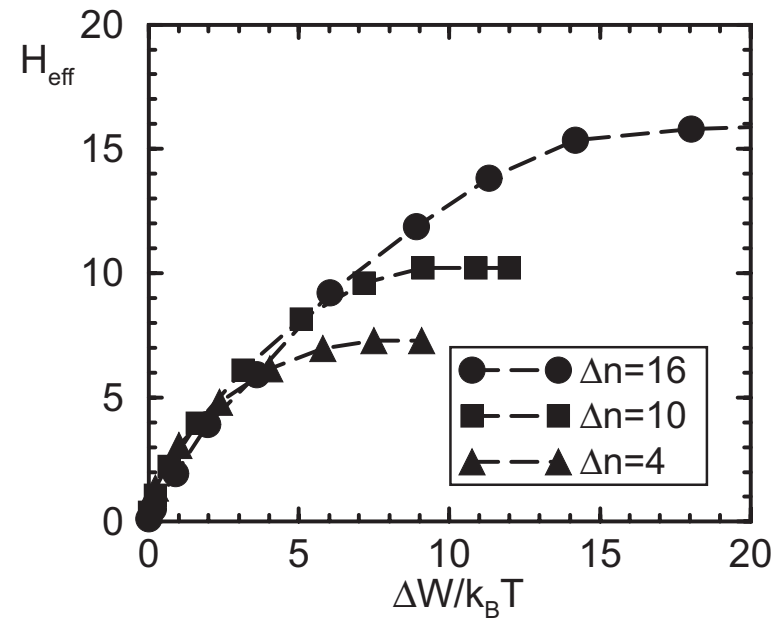
The nonequilibrium mechanism for ultrasensity in a biological switch: Sensing by Maxwell's demons

Yuhai Tu*   PNAS | August 19, 2008 | vol. 105 | no. 33 | 11737–11741

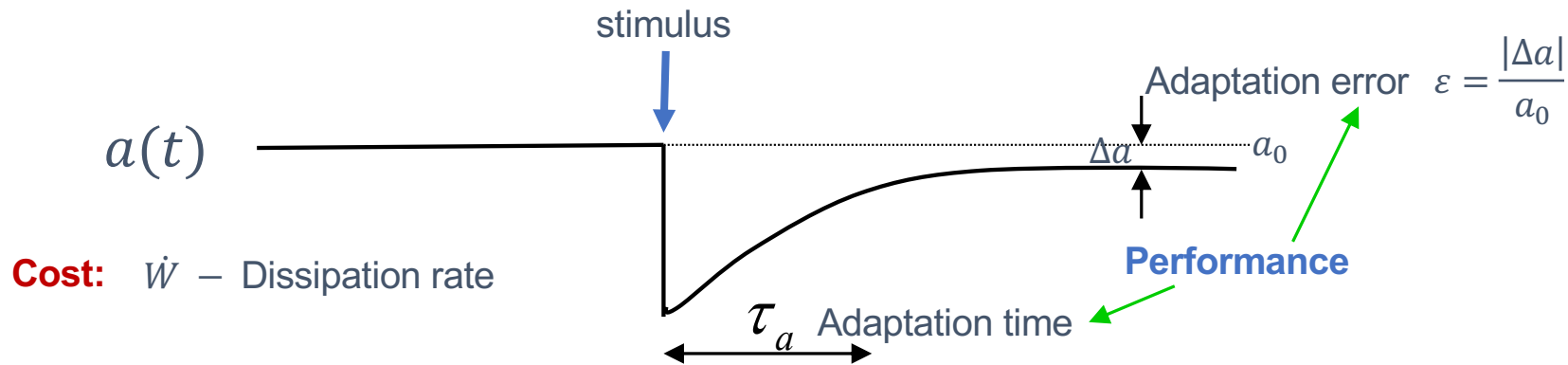# II. The molecular mechanism for accurate adaptation in bacterial chemotaxis

Sensing
Ligand binding
(Input, [L])

(fast ~1ms)

Sensing domain

Transmembrane helices

Linker region

Memory
(Methylation, m)

(slow >> 1s)

Cytoplasmic domain

CheB   CheR

E — E491
Q — Q295
E — E302
Q — Q309

Computation
Kinase activity
(Output, a)

(fast ~ 20ms)

A   **W**   CheA/CheW binding

Memory

m

[L]   Input

fast

slow

a   Output

excitation

inhibition

$$a = G([L], m) = \frac{1}{1 + e^{-\Delta E(m, [L])}}$$

$$\Delta E(m, [L]) \approx N\left[\alpha(m - m_0) - \ln\left(1 + \frac{[L]}{K}\right)\right]$$

$$\frac{dm}{dt} = F(a) \approx \frac{v_R(1 - a)}{(1 - a) + k_R} - \frac{v_B a}{a + k_B}$$

(YT, T. Shimizu, H. Berg, PNAS 2008)
(YT, Ann. Rev. Biophys. 2012)

The energy-speed-accuracy (cost-performance) tradeoff in sensory adaptation

stimulus

$a(t)$

Adaptation error $\quad \varepsilon = \dfrac{|\Delta a|}{a_0}$

Δa $\quad a_0$

Performance

Cost: $\dot{W}$ $-$ Dissipation rate

τ_a Adaptation time

Detailed balance is broken in the underlying feedback control mechanism for adaptation

(G. Lan, P. Sartori, S. Neumann, V. Sourjik, YT, Nature Physics, 2012)

Some calculations ......

The energy-speed-accuracy (ESA) relationship

$$\frac{\dot{W}}{k_B T} = c(a_0) \times \tau_a^{-1} \times \ln(\varepsilon^{-1})$$

(up to finite size correction....)

Entropy production rate

O(1) constant

Adaptation speed   **Adaptation accuracy**

III. Free energy is used to suppress phase diffusion in biochemical oscillations

Repressilator   Brusselator   Glycolysis
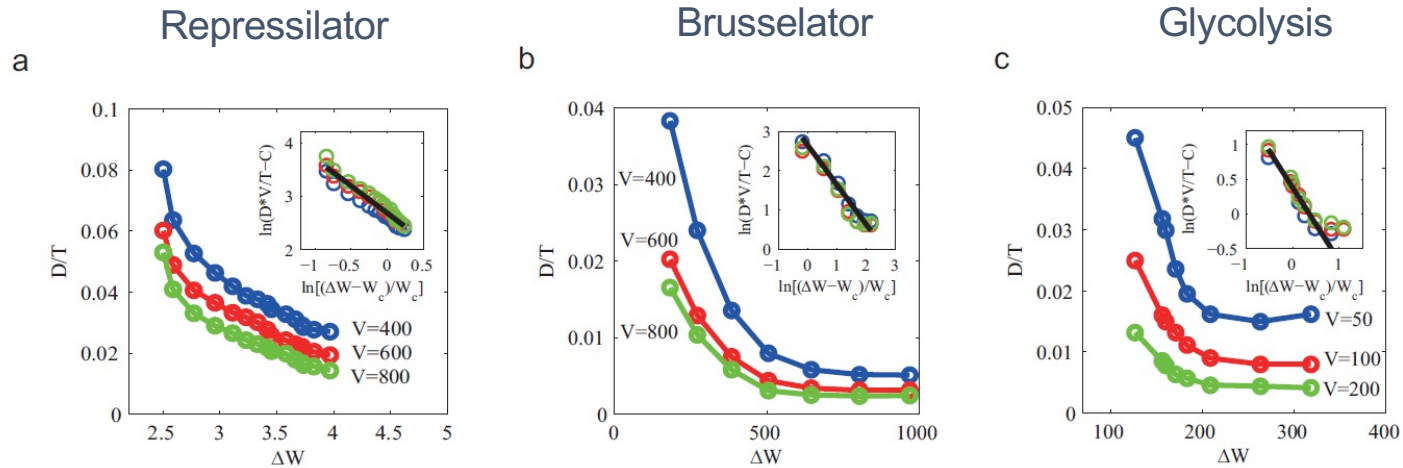
Phase diffusion constant

$$V \times \frac{D}{T} = C + \frac{W_0}{\Delta W - W_c}$$

Finite onset energy

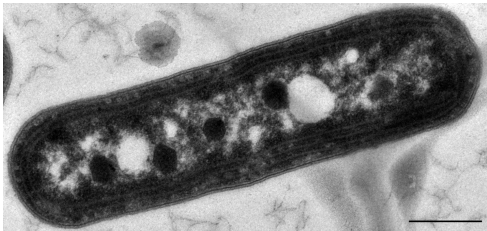$(\Delta W > W_c)$

V: volume
T: period
C, W_0: constants

Free energy dissipated per period

Valid for all three generic biochemical circuits for oscillations
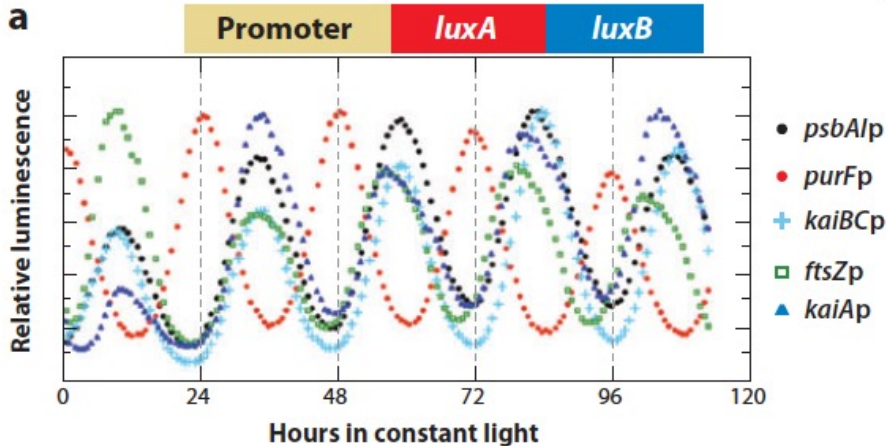
(Y. Cao, H. Wang, Q. Ouyang, Y. Tu, Nat. Phys. 2015)

IV. Synchronization of circadian clocks: the Kai system in cyanobacteria

Cyanobacteria is the simplest organism that exhibits circadian rhythm (24 hrs)
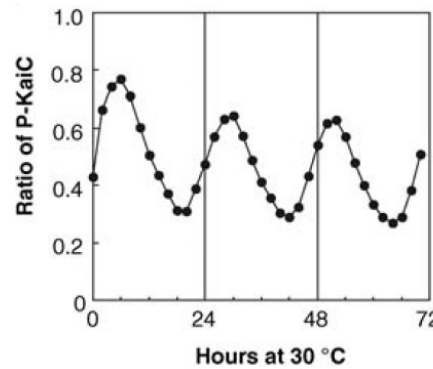

Synechococcus elongatus

robust 24-hr oscillation in vivo

(Johnson et al, *Ann. Rev. Biophys.* 2011)
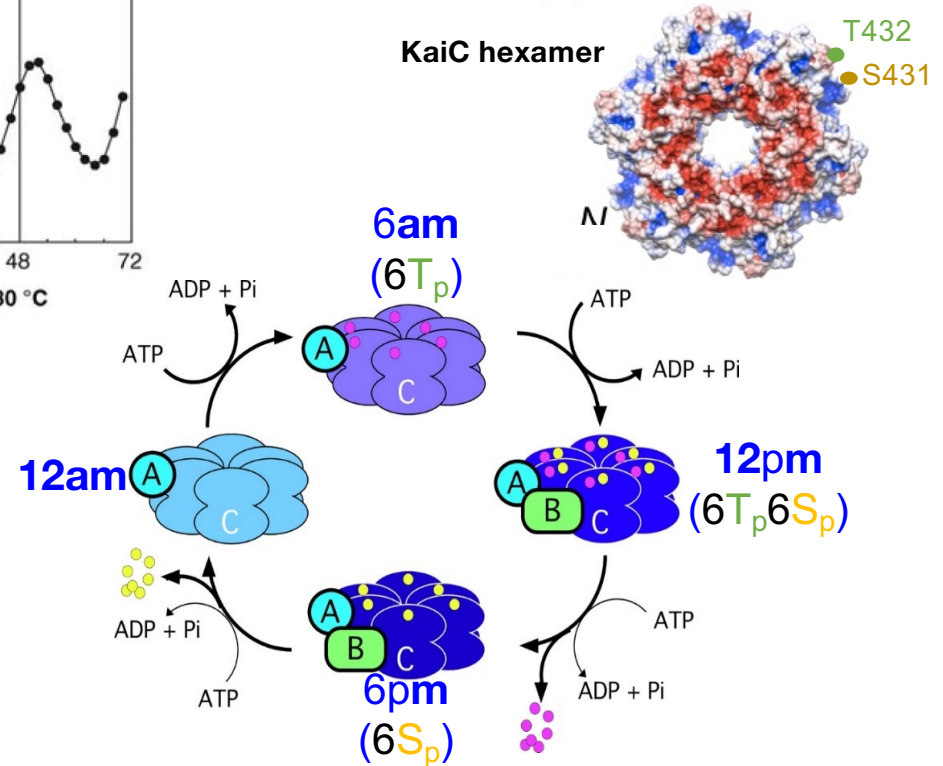
A breakthrough!

KaiA, KaiB, KaiC, +ATP

Reconstitution of Circadian Oscillation of Cyanobacterial KaiC Phosphorylation in Vitro
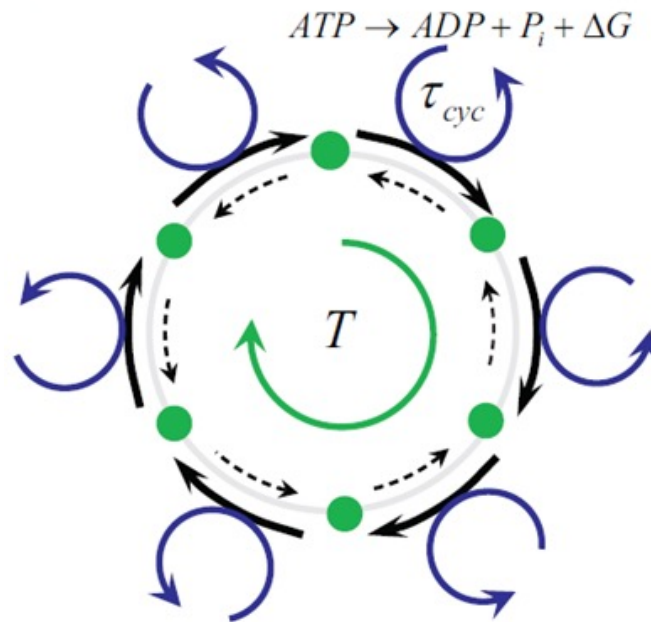
(Nakajima, …, T. Kondo, Science, 2005)

(Takao Kondō)

KaiC hexamer

T432
S431

6am
$(6T_p)$

12pm
$(6T_p6S_p)$

12am

6pm
$(6S_p)$

ADP + Pi
ATP
ATP
ADP + Pi
ADP + Pi
ATP
ATP
ADP + Pi

A simple model of a single biochemical clock: The Poisson clock model

Energy is used to suppress backward (reverse) reactions – reversal of time

$$ATP \rightarrow ADP + P_i + \Delta G$$

τ_{cyc}

T

(Y. Cao, H. Wang, Q. Ouyang, Y. Tu, Nat. Phys. 2015)

An Interesting Puzzle:

KaiC monomer

U

dP   P

S   T

dP   P

ST
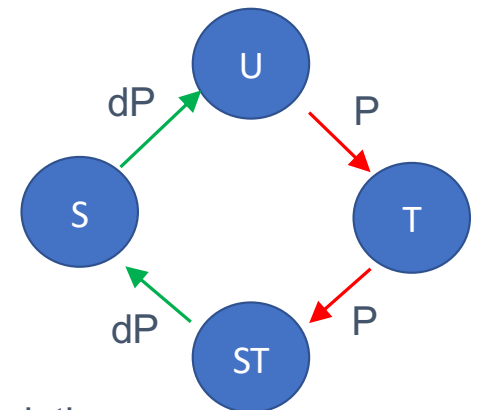
P –phosphorylation
dP- dephosphorylation

Only 2 ATP are needed for the P-dP cycle.

But, ~16 ATP are hydrolyzed per KaiC per day
(Terauchi et al, PNAS 2007)

What are the other 14 ATP's used for?

Coupled Molecular Clocks and the Cost of Coupling
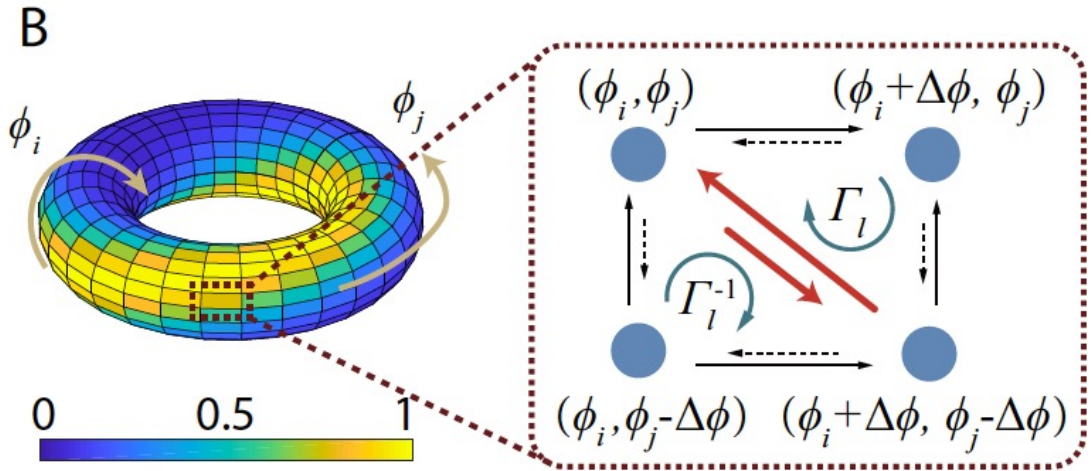
One finite-state Poisson clock
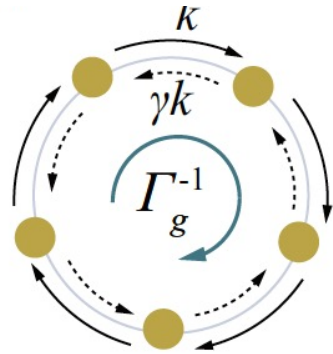
global clock cycle

$$\Gamma_g \equiv \prod_n k_n^- / \prod_n k_n^+ = \gamma^N \neq 1,$$

$$\Gamma_l = e^{-\Delta E_{ij}} \times \gamma^{-1} \times \gamma = e^{-\Delta E_{ij}} \neq 1,$$

Coupled clocks -- exchange interactions

B

scaled $P(\phi_i, \phi_j)$

local exchange cycles

$$k_{ex}((\phi_i + \Delta\phi, \phi_j - \Delta\phi) \to (\phi_i, \phi_j)) = \frac{\Omega}{m} \exp(-\Delta E_{ij}/2),$$

$$k_{ex}((\phi_i, \phi_j) \to (\phi_i + \Delta\phi, \phi_j - \Delta\phi)) = \frac{\Omega}{m} \exp(\Delta E_{ij}/2),$$

$$E(\phi_i - \phi_j) = -\frac{E_0}{2}\cos(\phi_i - \phi_j)$$

An analytical solution for the many-oscillator phase distribution (the mean-field limit)

The phase distribution function of m interacting oscillators $P(\phi_1, \phi_2, ..., \phi_m, t)$ satisfies **the Fokker-Planck equation**:

$$\frac{\partial P}{\partial t} = k \sum_i \frac{\partial}{\partial \phi_i} \left(-e_g + \frac{\partial}{\partial \phi_i} \right) P + \frac{\Omega}{m} \sum_{i<j} \frac{\partial}{\partial \varphi_{ij}} \left(2E'(\varphi_{ij}) + \frac{\partial}{\partial \varphi_{ij}} \right) P,$$

The processive speed: $v = k e_g = -k \ln \Gamma_g / 2\pi$ Interaction "energy" $E(\varphi_{ij})$ with $\varphi_{ij} = \phi_i - \phi_j$

Exact steady-state solution: $P_s(\vec{\phi}) = Z^{-1} \exp(-\beta E_t(\vec{\phi})),$ $E_t = \frac{2}{m} \sum_{(i<j)} E(\phi_i - \phi_j)$
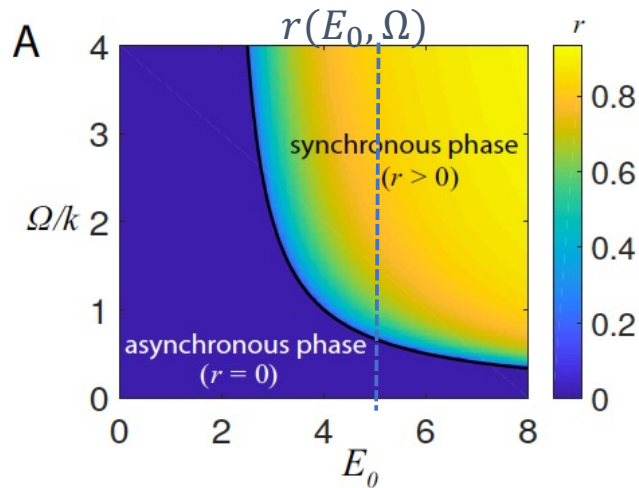
An effective temperature: $T_{eff} = \beta^{-1} = 1 + k/\Omega$ decreases with exchange frequency Ω

The nonequilibrium phase transition and energy cost of synchronization

Synchronization order parameter $r \in [0,1]$

$$re^{i\psi} \equiv \frac{1}{m}\sum_{j=1}^{m} e^{i\phi_j}, \qquad \psi = vt$$

A

$r(E_0,\Omega)$

synchronous phase $(r > 0)$

asynchronous phase $(r = 0)$

Critical line: $\dfrac{\Omega}{\Omega+k}E_0 = 2$

$r > 0$ when $\dfrac{\Omega}{\Omega+k}E_0 > 2;$

$r = 0$ when $\dfrac{\Omega}{\Omega+k}E_0 \leq 2$

Energy dissipation rate

$$\dot{W} = \frac{1}{m}\int \Big[\underbrace{\sum_i \frac{J_i^2}{kP_s}}_{\text{procession energy (single clock)}} + \underbrace{\sum_{i<j}\frac{J_{ij}^2}{\frac{\Omega}{m}P_s}}_{\text{coupling energy}}\Big]d\vec{\phi},$$

$$W = \dot{W}\times T = W_0 + W_{ex}$$

$$W_0 = 2\pi e_g = -\ln\Gamma_g$$
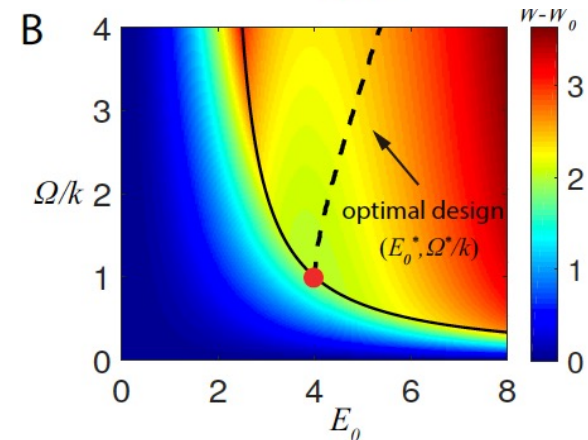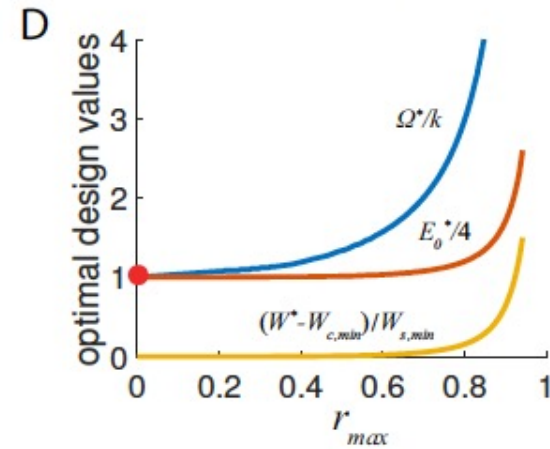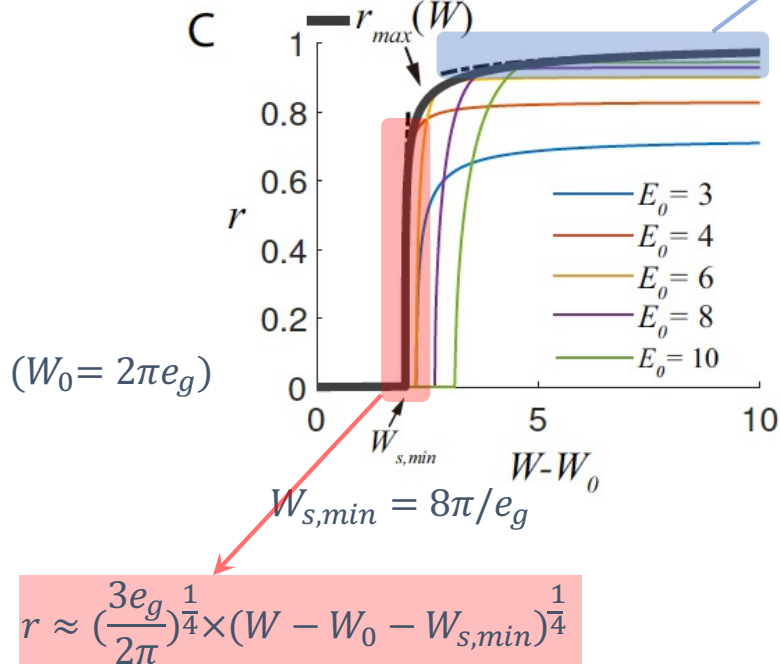
$W_{ex}(E_0,\Omega)$
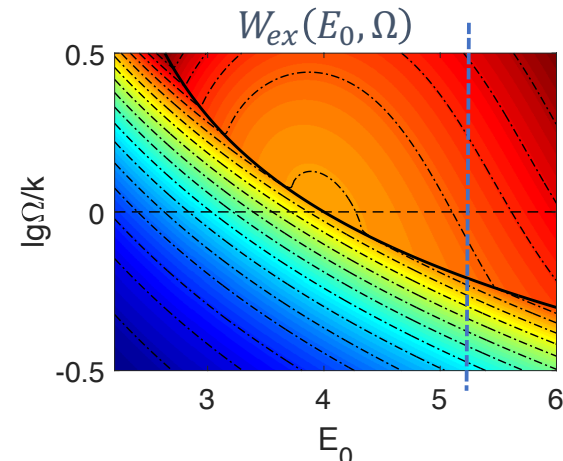
The optimal design for maximal synchronization with a fixed energy budget

$$r_{max}(W) = \max(r(E_0, \Omega|W))$$

$$(E_0^*, \Omega^*) = \underset{\Omega, E_0}{\operatorname{argmax}} \, r(E_0, \Omega|W)$$

$$r \approx 1 - \frac{\pi}{e_g(W - W_0)}$$

$$(W_0 = 2\pi e_g)$$

$$W_{s,min} = 8\pi/e_g$$

$$r \approx \left(\frac{3e_g}{2\pi}\right)^{\frac{1}{4}} \times (W - W_0 - W_{s,min})^{\frac{1}{4}}$$

How do individual KaiC hexamers synchronize with each other?

KaiC hexamers exchange monomers

(Kageyama et al, …, T. **Kondo, Mol. Cell, 2006)**
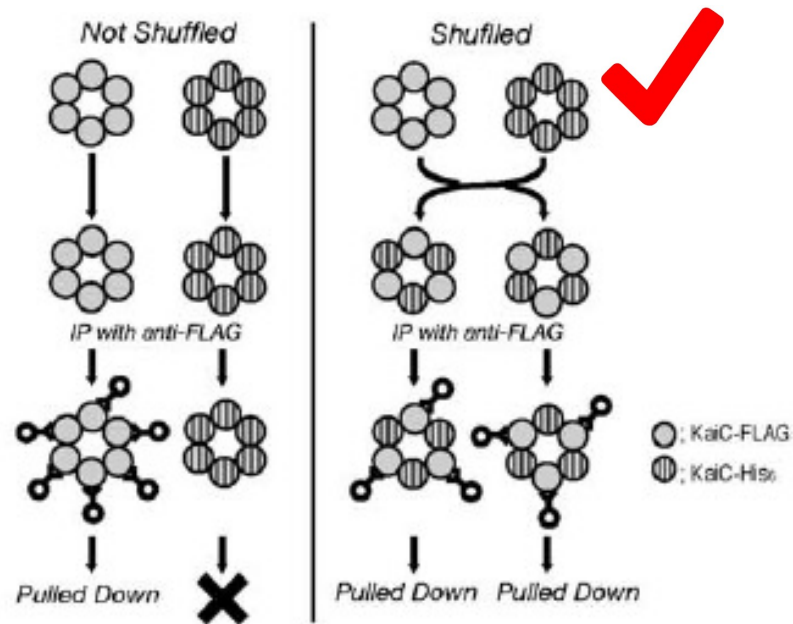
Monomer shuffling can lead to synchronization

**E. Emberly & N. Wingreen, PRL 2006**
T. Mori et al, PLoS Biol. 2007
….

A model of synchronization based on monomer-shuffling in the Kai system

The Puzzle: Only 2 ATP are needed for the P-dP cycle , 16 ATP are hydrolyzed per KaiC per day
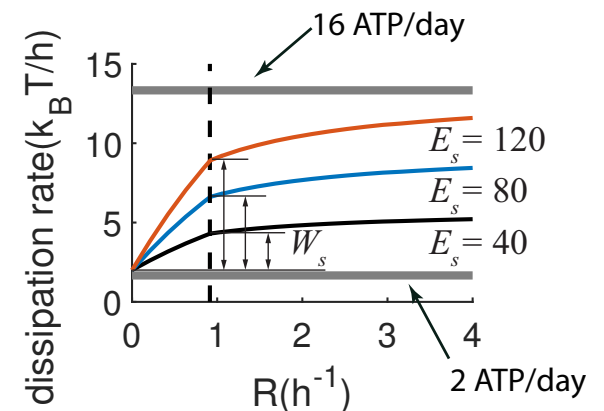(Terauchi et al, PNAS 2007)

Our model

(1) Dynamics of individual KaiC hexamer -- PdP cycle

(2) Exchange coupling between two KaiC hexamers

Monomer-shuffling between same conformation, P(◯) or dP(▢)
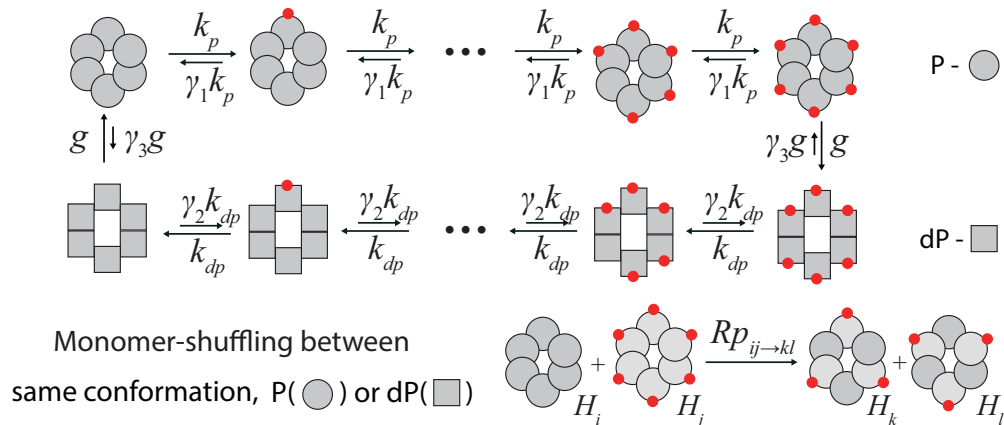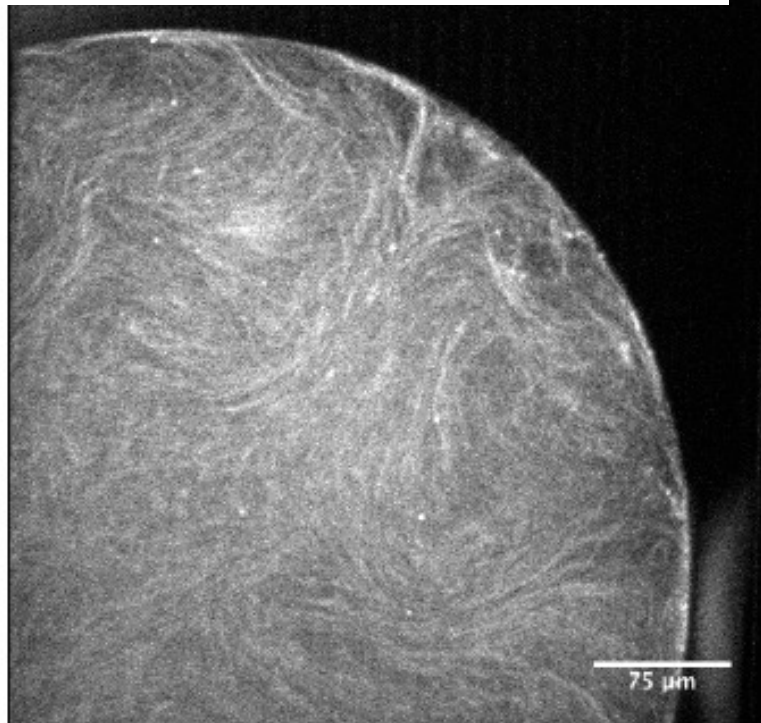
Varying
R– exchange (shuffling) rate
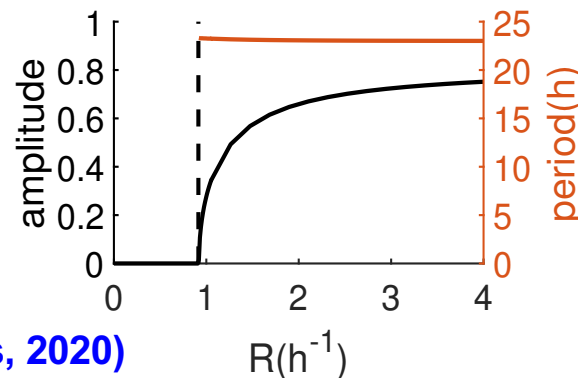
(DL Zhang et. al. Nature Physics, 2020)

The inverse scaling of the dissipation rate

The efficiency puzzle

Active Gel: Kinesin –Microtubule mixture

75 μm

How much energy is dissipated from flow?

Flow velocity, $v \sim 1\mu m/sec$

Flow correlation length, $L \sim 100\ \mu m$

Viscosity, $\eta \sim 1\ mPa\ s$

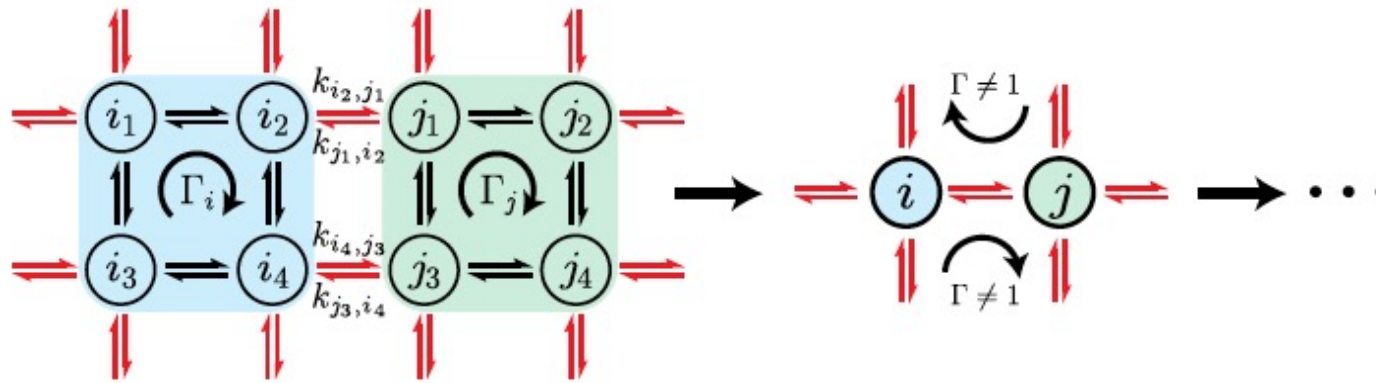Sample size, $V_s \sim 1\ \mu L$

$$P = \eta\, v^2\, V_s\, /\, L^2 \sim 10^{-7}\ nW$$

Calorimetry measurement $\Rightarrow \sim 10 nW$

$\Rightarrow$ **Efficiency $\sim 10^{-8}$!**

-- PJ Foster*, J Bae*, B Lemma, J Zheng, Z Dogic, DJ Needleman, JJ Vlassak, "Dissipation and energy propagation across scales in an active cytoskeletal material", *in preparation*

Coarse graining of a reaction network in state space

Coarse-graining of the states

$$P_i^{ss} = \sum_{\alpha=1}^{r} P_{i_\alpha}^{ss}, \qquad P_j^{ss} = \sum_{\alpha=1}^{r} P_{j_\alpha}^{ss}.$$

Coarse-graining of the rates

$$k_{i,j} = \frac{J_{i,j}}{P_i^{ss}} = \frac{1}{P_i^{ss}} \sum_{(\alpha,\beta)} J_{i_\alpha, j_\beta} = \frac{\sum_{(\alpha,\beta)} k_{i_\alpha, j_\beta} P_{i_\alpha}^{ss}}{\sum_{\alpha=1}^{r} P_{i_\alpha}^{ss}}.$$

The inverse scaling of the dissipation rate

Entropy production rate

$$\dot{W} = \sum_{i<j}(J_{i,j} - J_{j,i})\ln\frac{J_{i,j}}{J_{j,i}},$$

The random flux model ($k_{i,j} = i.i.d.P(k)$)

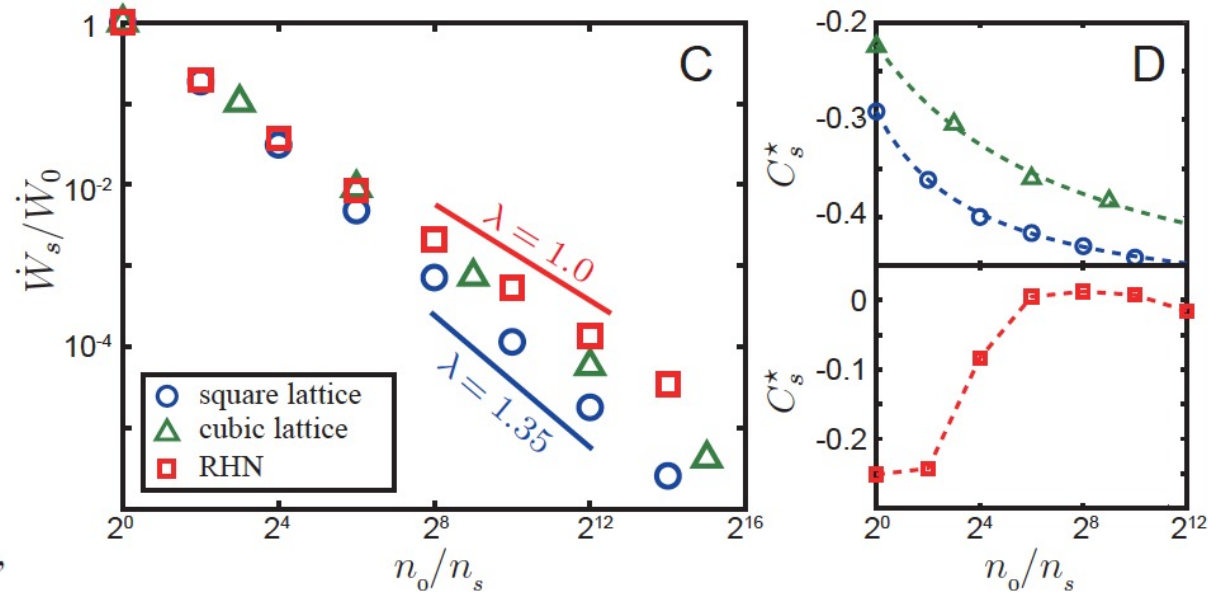$$\frac{\dot{W}(n_s)}{\dot{W}(n_0)} = \left(\frac{n_0}{n_s}\right)^{-\lambda},$$

$$\lambda = d_L - \log_r(1 + C^*),$$

$$d_L \equiv \frac{\ln[L(n_s)/L(n_0)]}{\ln(n_s/n_0)}.$$

$$C^* = \frac{\langle A_{i_\alpha,j_\beta}(A_{i,j} - A_{i_\alpha,j_\beta})\rangle_{i_\alpha,j_\beta}}{\sqrt{\langle A^2_{i_\alpha,j_\beta}\rangle_{i_\alpha,j_\beta}\langle(A_{i,j} - A_{i_\alpha,j_\beta})^2\rangle_{i_\alpha,j_\beta}}},$$

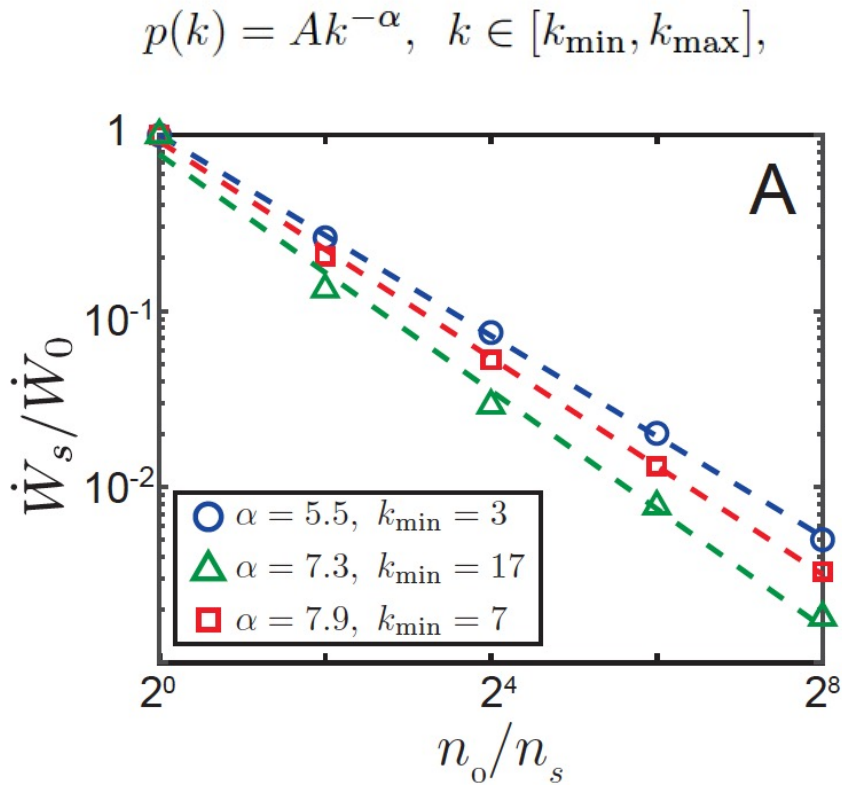$$(A_{i,j} = J_{i,j} - J_{j,i})$$

n_0 − number of states in the microscopic system
n_s − number of states coarse − grained s times
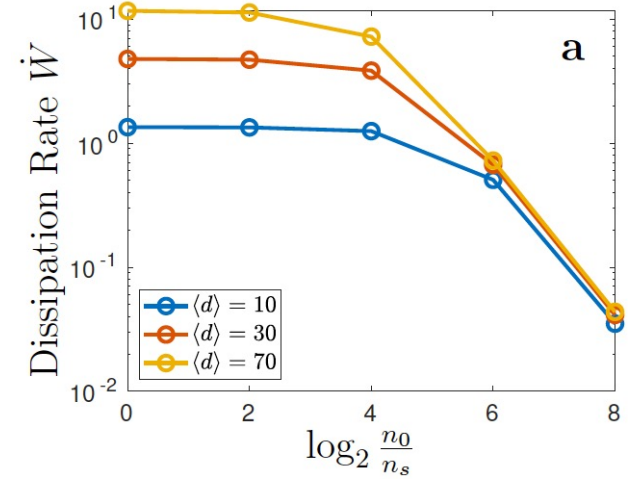L − number of reactions (links) in the network

The inverse scaling law (ISL) requires self-similarity of the reaction networks
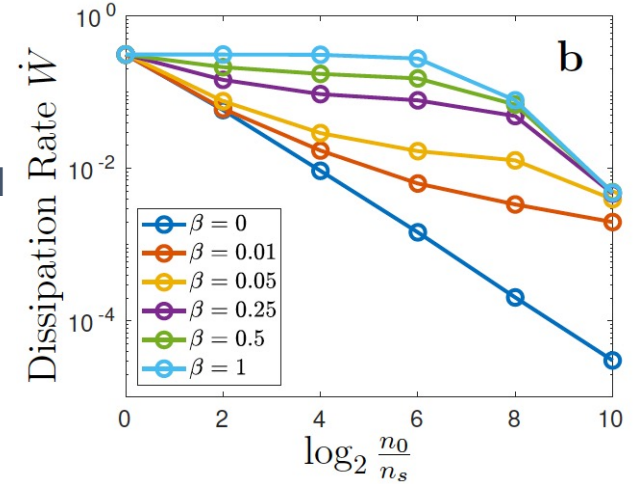
No ISL in networks w/o self-similarity

**ISL exists
in scale-free networks**

$$p(k) = Ak^{-\alpha}, \quad k \in [k_{\min}, k_{\max}],$$
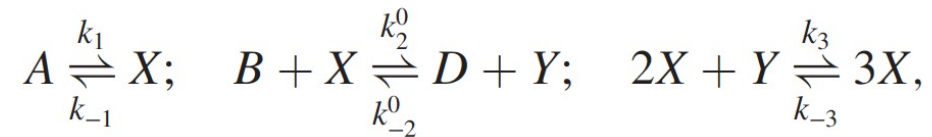
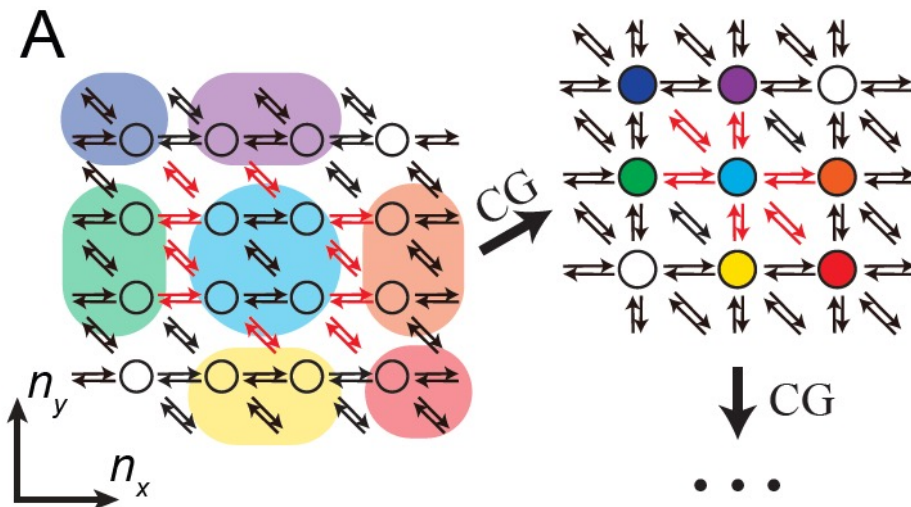# The dissipation scaling exists in realistic biochemical networks

(1) The reversible Brusselator model for biochemical oscillations

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X; \quad B + X \underset{k^0_{-2}}{\overset{k^0_2}{\rightleftharpoons}} D + Y; \quad 2X + Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 3X,$$

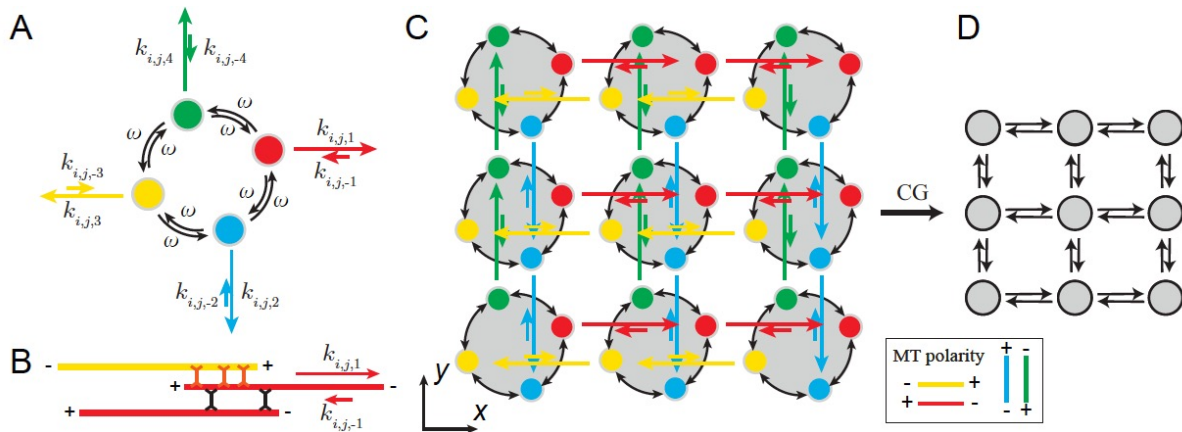**Coarse-graining (CG) procedure**
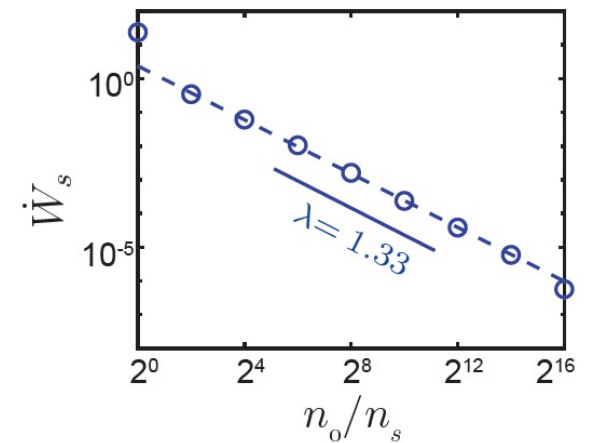
Inverse dissipation scaling
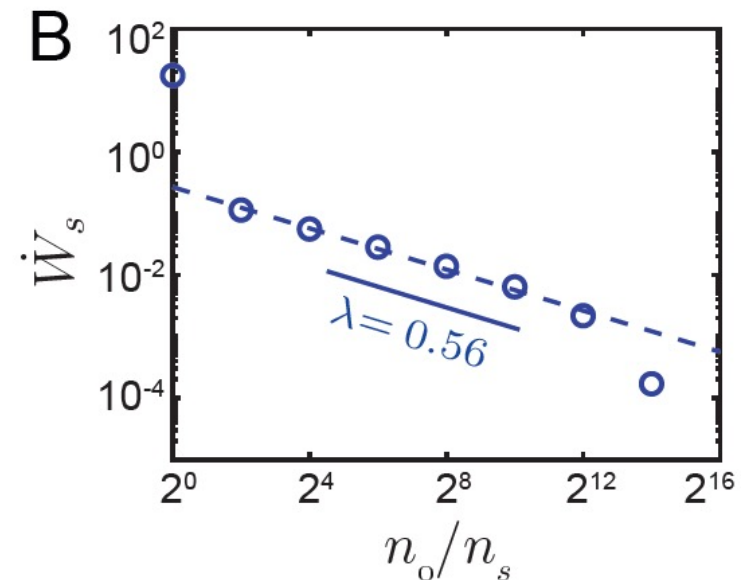
The dissipation scaling exists in realistic biochemical networks

(2) A simple model for kinesin–microtubule active flow system

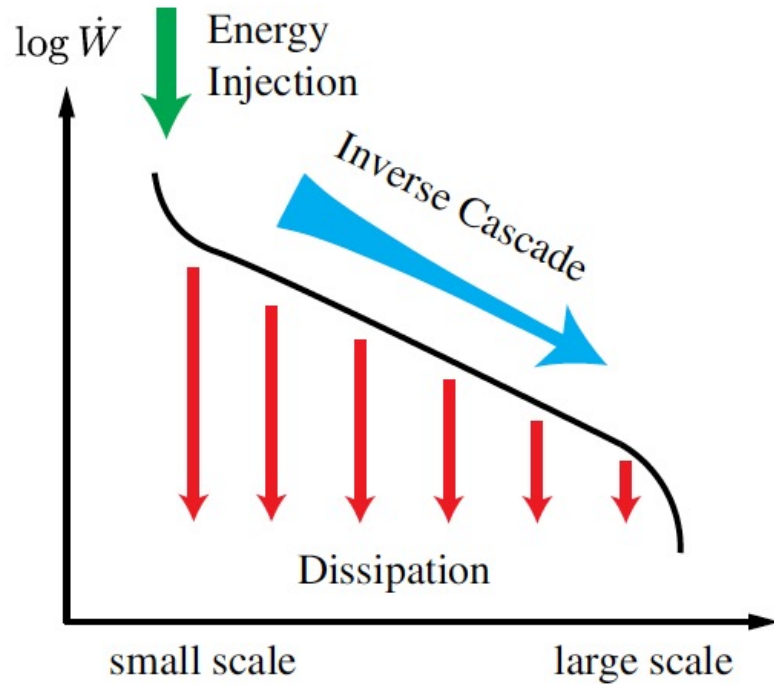**Coarse-graining (CG) procedure**

Inverse dissipation scaling

kinesin persistent run length $l_0 \sim 0.6 - 1\mu m$
active flow scale $l_f \sim 100\mu m$

$$\frac{\dot{W}_f}{\dot{W}_0} \approx \left(\frac{l_0}{l_f}\right)^{3\lambda_{3D}} \approx 10^{-7.4} - 10^{-8.2} \qquad (\lambda_{3D} \approx 1.23)$$
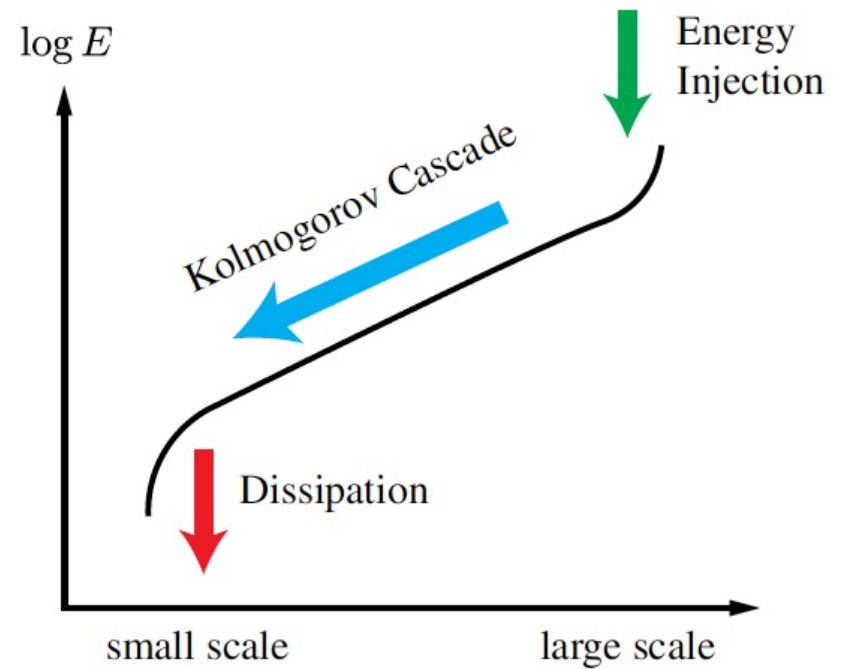
~10⁻⁸

An interesting comparison

Non-equilibrium reaction networks

Homogeneous fully-developed turbulent

(Qiwei Yu, DL Zhang, YT, PRL, 126 (8), 2021)

Understanding biochemical circuits: a nonequilibrium physics perspective

- **Biological functions are maintained by continuous energy consumption.**

  Living systems are in non-equilibrium steady state (NESS).
  The "resting energy" enhances their active functions, e.g., speed, accuracy, etc..

- **Energy costs limit the optimal performance (e.g., Accuracy, Speed).**

  Type I accuracy: mean value $\qquad$ **Type II accuracy: fluctuation**

  $$k_B \mathrm{T} \ln(\epsilon^{-1}) \cong c_0 \dot{W} \times \tau_a \qquad D_\theta = C + \frac{W_0}{\Delta W - W_c}$$

  Adaptation, error correction, **time control**, spatial and length control, self-assembly, **synchronization**, memory & learning, etc…

- **Design principles for efficient biochemical networks.**

- **Dissipation occurs in all scales – smaller scale has higher dissipation rate. Inverse scaling law exists in some self-similar extended networks. More needs to be done …**

Thank You!