Internally-funded improbability: a thermodynamic account of agency

Benjamin Gaskin

University of Sydney

An outline of the argument—

- **Living systems are dissipative structures**
  - Dissipative structures are not living systems
  - Living systems use energy for work indirectly
  - Organisational closure is insufficient for agency
  - An epistemic cut is necessary for agency

- **Agency is internally-funded improbability**
  - Maxwell's demon is a minimal agent
  - Landauer's principle sets a cost for agency
  - Agency can be quantified by KL divergence
  - Agency can be formalised as a lower bound

The etymologies of patiency and agency

- The reconstructed PIE root of agency is ***h_2eǵ-: “to drive”**

- The Latin root of patiency is **patior**: **“suffer, experience, wait”**
  - The PIE root is disputed, perhaps *peh$_1$- (“to hurt”) or *pet- (“to fly, fall”)


Captured by SD 1045's onboard camera during
Category 4 Hurricane Sam, Sept. 30 2021
SAILDRONE
NOAA

How do living systems differ?

Living systems are dissipative structures

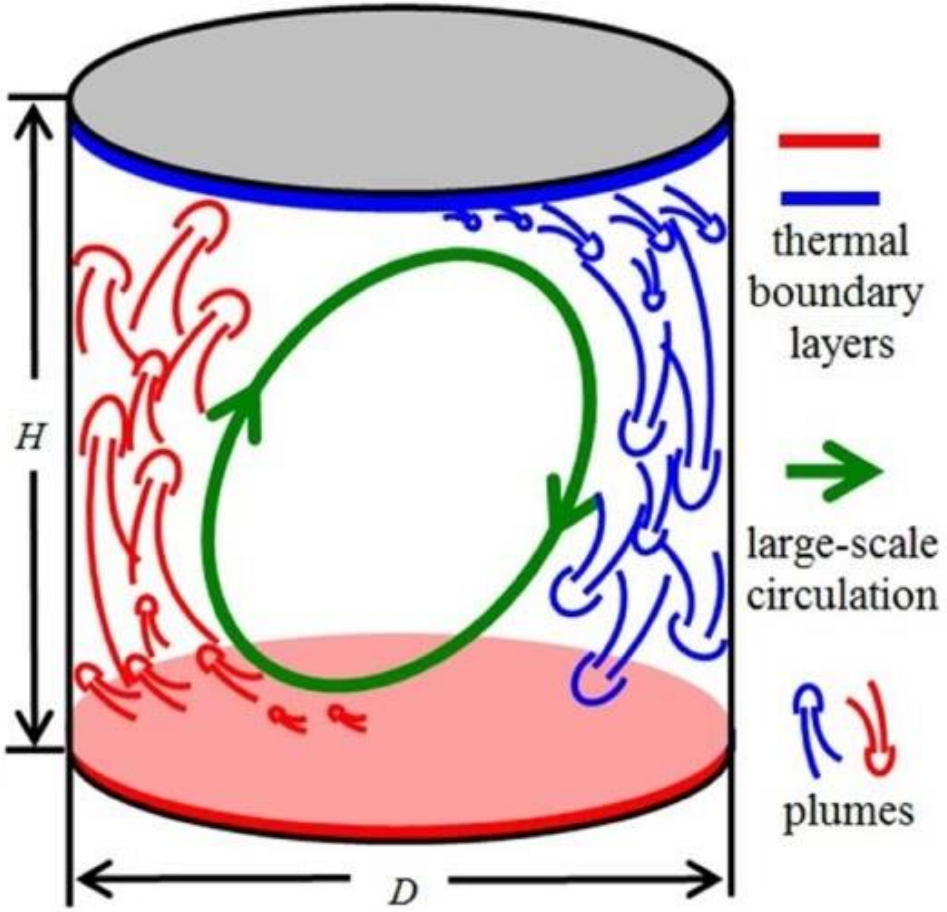
Hurricanes are dissipative structures

Dissipative structures are not living systems

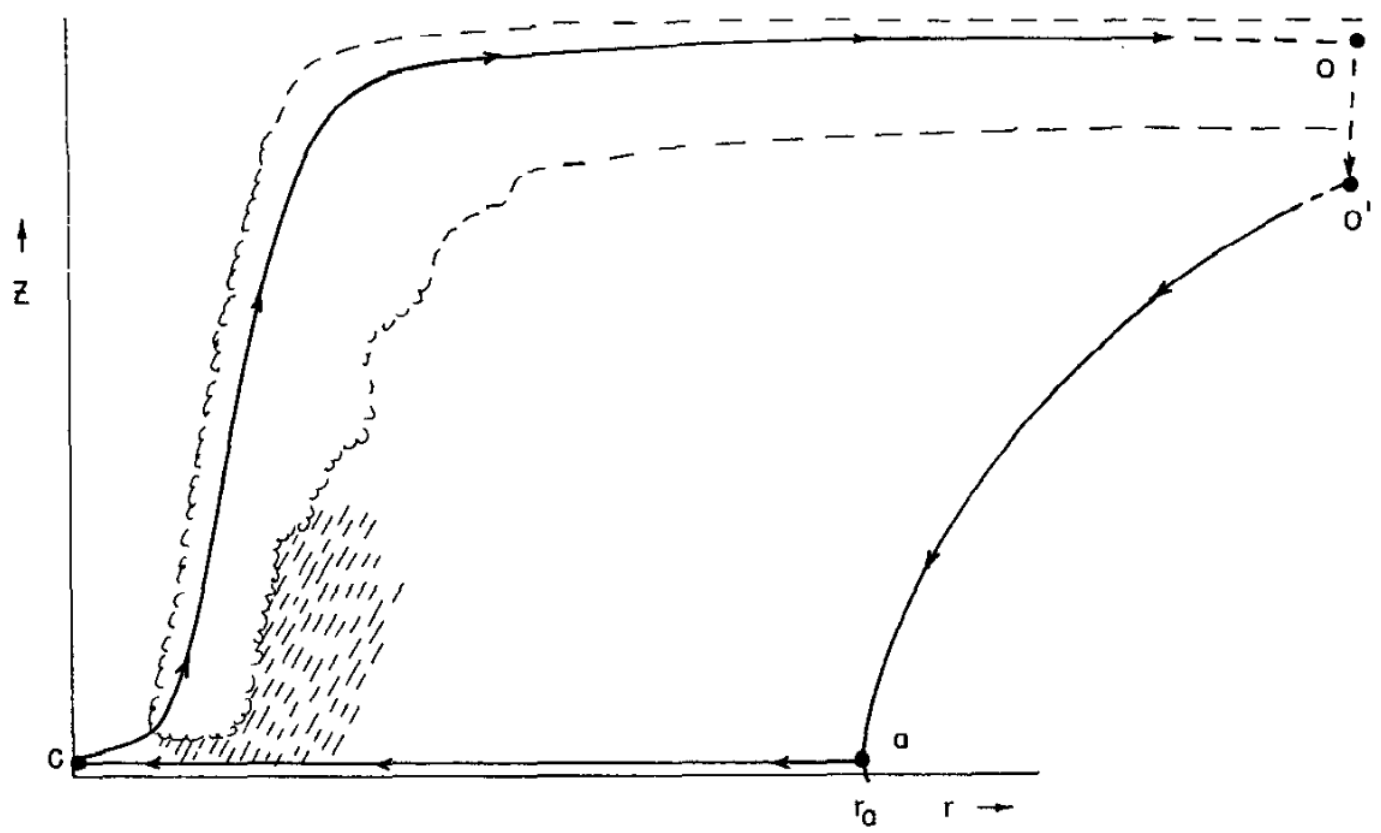
Living systems use energy for work indirectly

Living systems are dissipative structures

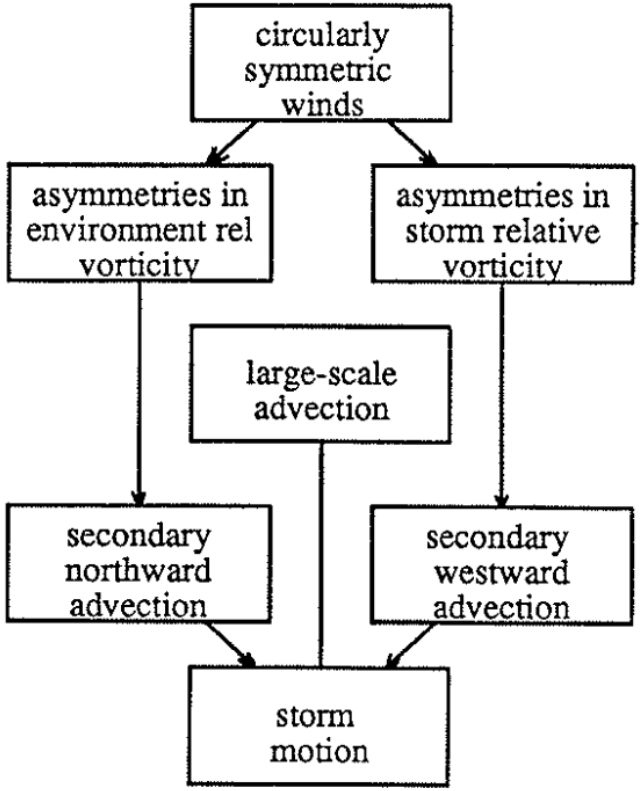
- Prigogine & Lefever (1973):
  - “One has **structures which are created by the continuous flow of energy and matter** from the outside world. Their maintenance requires a critical distance from equilibrium, i.e. a minimum level of dissipation. For all these reasons we have called them **dissipative structures**.”



Hurricanes are dissipative structures



Potential Intensity (PI) Theory



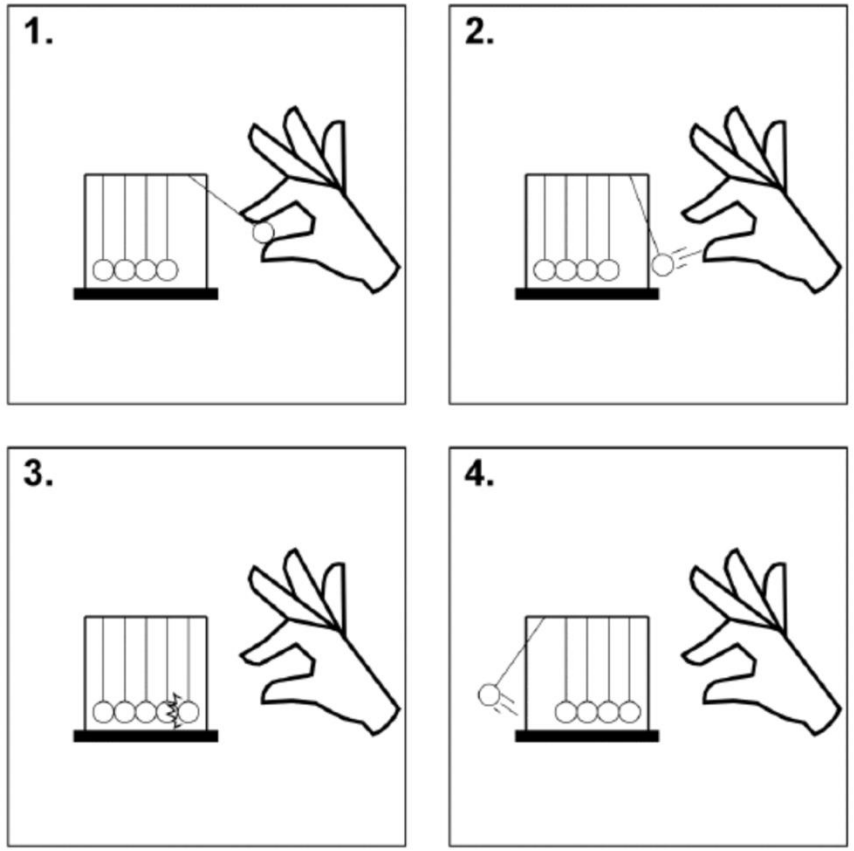
Beta and Advection Model (BAM)

Dissipative structures are not living systems

- **Bauer's three laws of motion of living matter**
  1. **Spontaneous activity**
     - Changes occur **unprompted by any external action**
  2. **Active response**
     - Reactions are **underdetermined by external actions**
     - **Scale and location** need not correspond to the action
  3. **Stable non-equilibrium**
     - Energy is **invested in a structural opposition to equilibrium**
     - Converted into **internal source of work**: 'free structural energy'

Living systems use energy for work indirectly

- Life is characterised by the indirect use of energy: **first to build internal structures, then these fund work**
  - Watterson likens this to Newton's cradle
  - Bauer calls this **'free structural energy'**
- Not merely structural, **also dynamics**
  - This is **static, internal work** (e.g., NESS)
- This structural energy provides the **material of outward, external work**
  - We will argue that it **also funds agency**



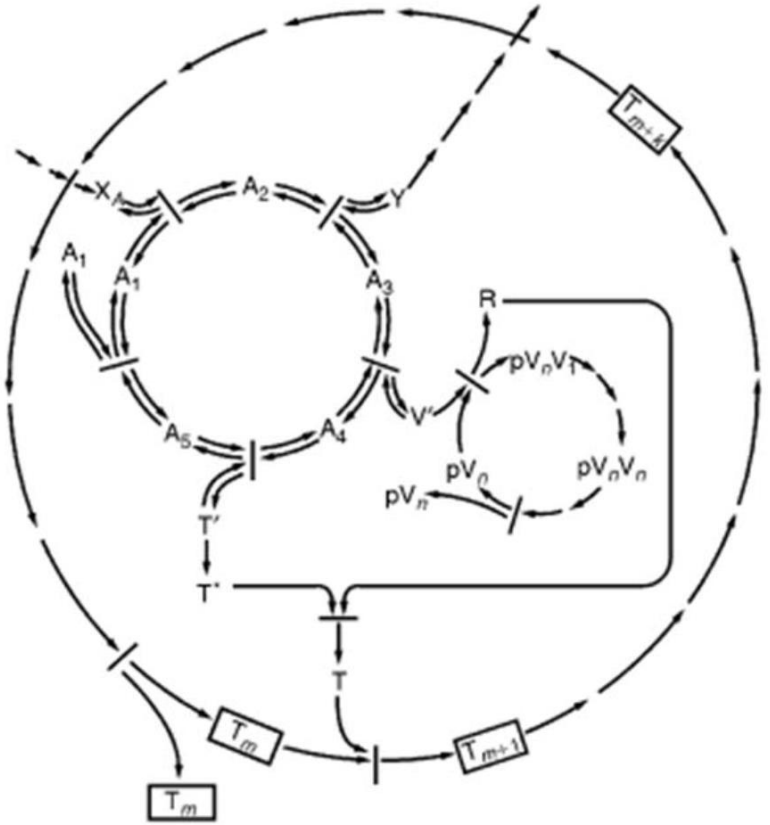
The evolution of agency

Organisational closure is insufficient for agency

An epistemic cut is necessary for agency

Organisational closure is insufficient for agency

- Gánti's (1975) **chemoton** model
  - **Three coupled subsystems**
    - **Metabolic**
    - **Membrane**
    - **Template**
  - **Autocatalytic**
  - **Stoichiometric**
- Mossio & Moreno (2010): "adaptive control was presumably absent … stability against perturbations in time would rather rely on internal redundancies and feedback loops."
  - Control relationships need separate timescales
  - This requires the dynamic decoupling of systems



An epistemic cut is necessary for agency

- Gánti's template subsystem provides a **seed of agency**
  - **Initially random**, templates meaningful only in length and quantity
  - Sequences are energy degenerate, **not constrained**
  - Through secondary structure, these can have **functional effects**
- This functional effect corresponds to **Pattee's epistemic cut**
  - The meaning of symbols must be seen in terms of **complementarity**
  - Symbols **constrain dynamics according to their selection history**

A thermodynamic account of agency

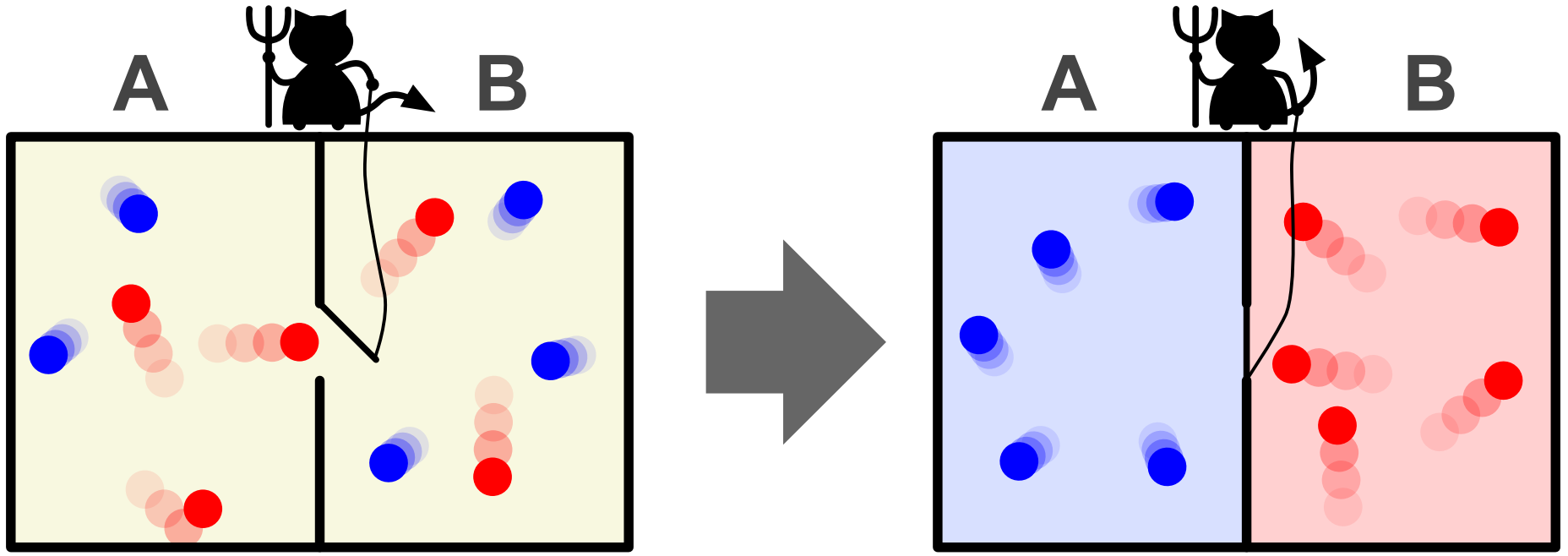
Maxwell’s demon is a minimal agent

Landauer’s principle sets a cost for agency

Agency can be quantified by KL divergence

Agency can be formalised a a lower bound

Maxwell's demon is a minimal agent



Landauer's principle sets a cost for agency

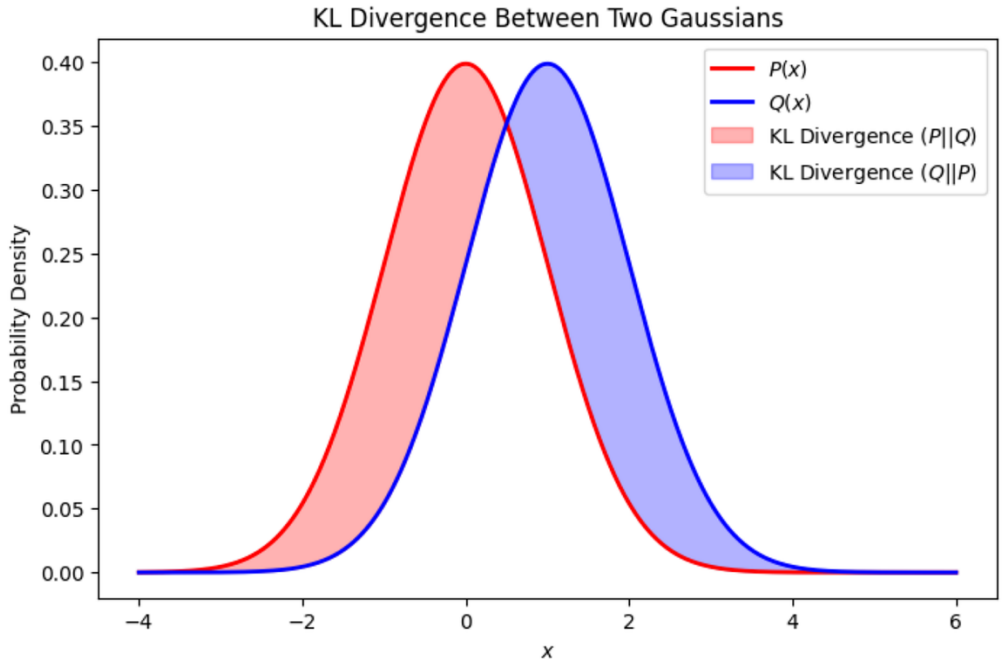
$$E_{\min} = k_{\mathrm{B}} T \ln 2$$

- Maxwell's demon is consistent with the second law of thermodynamics because of **the cost of information**
  - Minimal cost of erasing one bit
  - Boltzmann's constant (~1.38×10^{-23} J/K)
  - Absolute temperature in kelvins
  - Entropy of a binary choice

Agency can be quantified by KL divergence

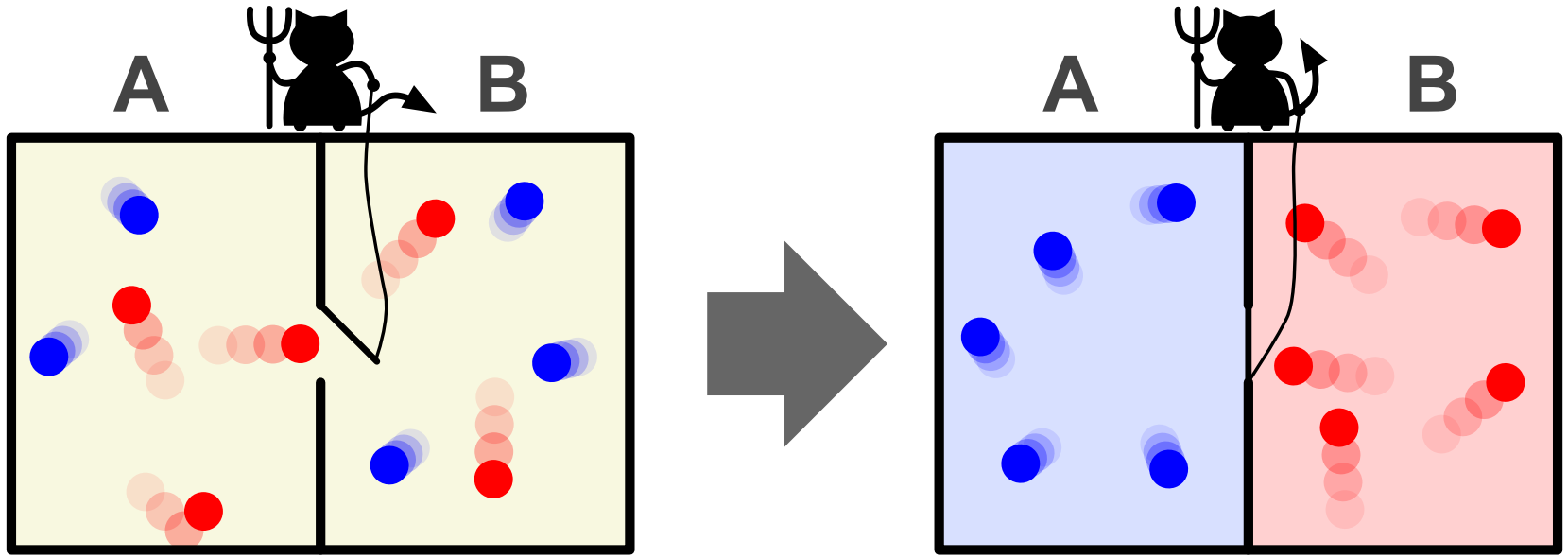
$$D_{KL}(P|Q) = \int P(x)\, ln \frac{P(x)}{Q(x)}\, dx$$

- Kullback-Leibler (KL) divergence measures the **distance (in bits) between two distributions**
  - Reference distribution (P)
  - Comparison distribution (Q)
  - Pointwise information distance
  - Weighted local divergence
  - Expectation under P



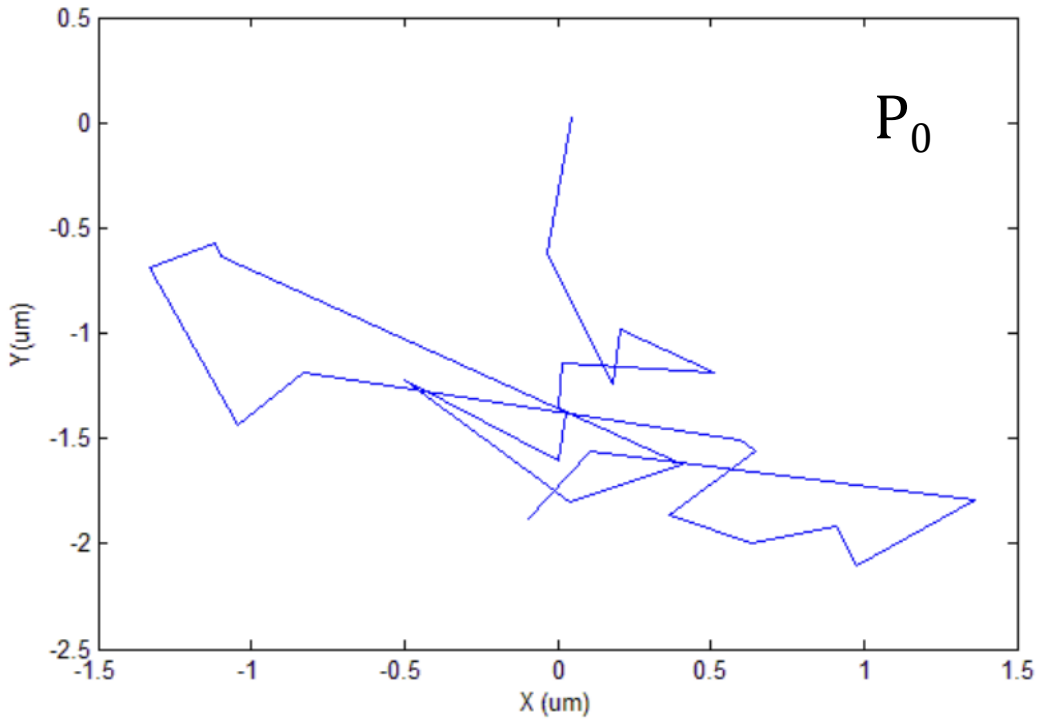
Agency can be formalised as a lower bound

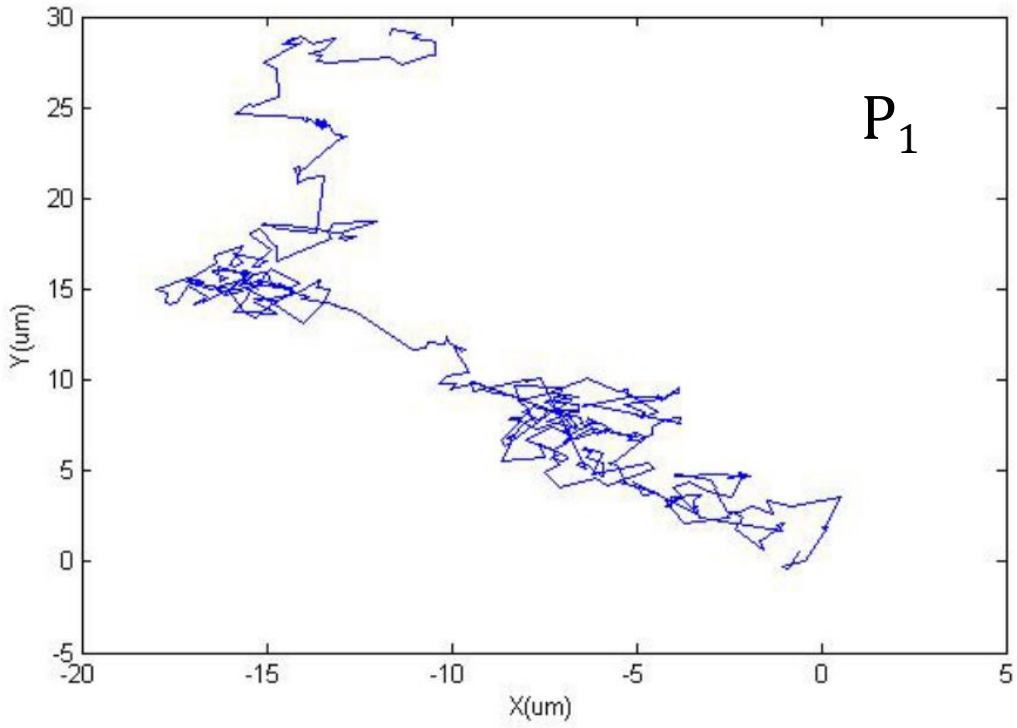
$$Q_{min} \geq k_B T \, D_{KL}(P_1 | P_0)$$



Agency can be formalised as a lower bound

$$Q_{min} \geq k_B T \, D_{KL}(P_1 | P_0)$$

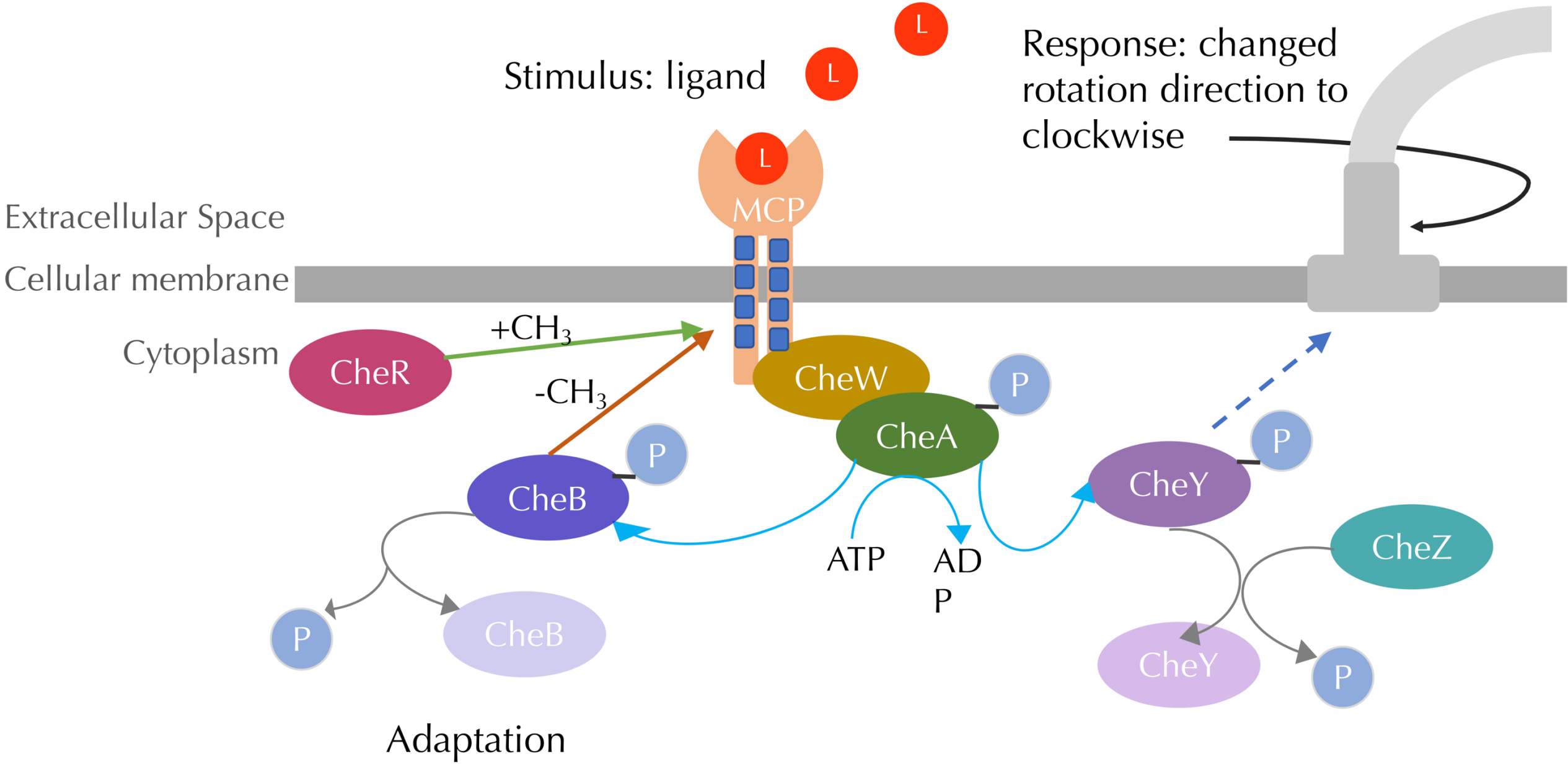


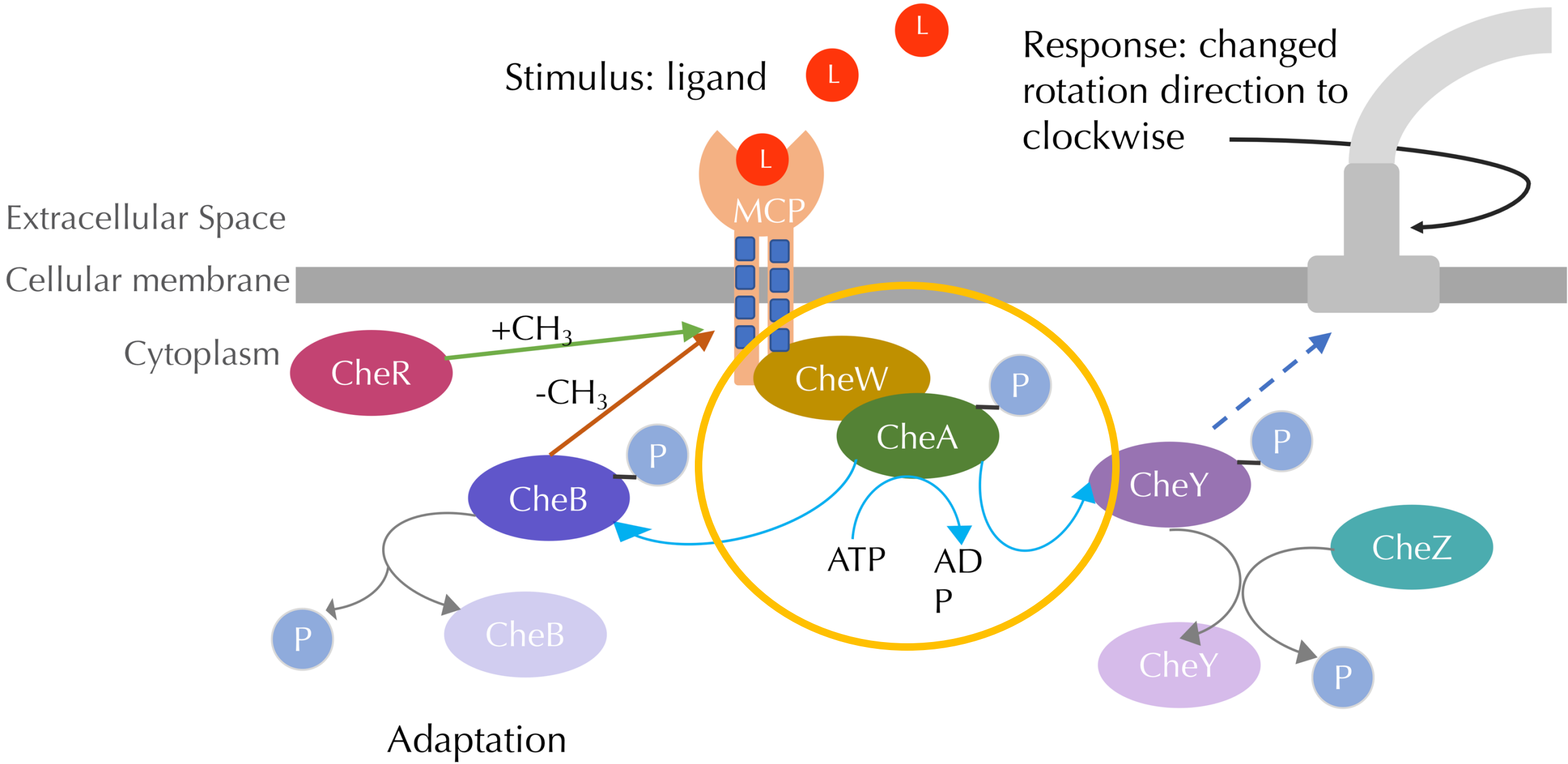


Energy in bacterial chemotaxis

Energy funds improbability throughout E. coli

Energy funds improbability through life


Stimulus: ligand
L
L
L
Response: changed rotation direction to clockwise
Extracellular Space
Cellular membrane
Cytoplasm
MCP
CheR
+CH3
-CH3
CheW
CheA
P
CheB
P
CheY
P
ATP
AD P
CheZ
P
CheB
CheY
P
Adaptation


Stimulus: ligand
L
L
L
Response: changed rotation direction to clockwise
Extracellular Space
Cellular membrane
Cytoplasm
MCP
CheR
+CH3
-CH3
CheW
P
CheA
P
CheB
P
CheY
P
ATP
AD P
CheZ
P
CheB
P
CheY
Adaptation

Energy funds improbability throughout E. coli

- Sensitivity of the flagellar switch
  - The nonequilibrium mechanism for ultrasensitivity in a biological switch (Tu et al., 2008)
- Sensitivity of the receptor complex
  - A nonequilibrium allosteric model for receptor-kinase complexes (Hathcock et al., 2023)
- Substrate selectivity in ABC transport proteins
  - Omnipresent Maxwell's demons orchestrate information management in living cells (Boël et al., 2019)
- Error correction in DNA replication
  - Thermodynamics of error correction (Sartori & Pigolotti, 2015)
- Directed movement in ribosomes
  - Trajectories of the ribosome as a Brownian nanomachine (Dashti et al., 2014)

Energy funds improbability throughout life

- **Substrate selectivity in ABC transport proteins**
  - Omnipresent Maxwell's demons orchestrate information management in living cells (Boël et al., 2019)
- **Error correction in DNA replication**
  - Thermodynamics of error correction (Sartori & Pigolotti, 2015)
- **Directed movement in ribosomes and motor proteins**
  - Trajectories of the ribosome as a Brownian nanomachine (Dashti et al., 2014)
  - Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins (Hwang & Karplus, 2019)

Conclusion

What is an internally-funded improbability?

Cf. FEP: “With him it is standing on its head”

Agency is costly across scales

What is an internally-funded improbability?

- **Living systems convert energy via internal work into structure**
  - Bauer's free structural energy provides the energy necessary to agency
- **There is a minimal cost to manipulate physical information**
  - Landauer's principle sets the limit for Maxwell's demon as an agent
- **The effect of symbols is underdetermined by their physics**
  - Pattee's epistemic cut is the complementarity of history and dynamics

Cf. FEP: “With him it is standing on its head”

$$F = D\big(q(H)|p(H)\big) - \langle \ln p\,(\,O \mid H\,)\rangle_q$$

- F is the organism’s **surprise that the environment is in a given state**
  - This measure is **later included in G to motivate a homeostatic response**
- Our thermodynamic account of agency instead measures the **organism’s activity as divergence from a passive distribution**
  - This might be seen as the **environment’s ‘surprise’ at a given response**
- There are **no beliefs, priors, or models** in our account of agency
  - These **may emerge in agentic systems**, but they are **not the fundament**
- Our account is thus aetiologically **more plausible and parsimonious**
  - Something like the FEP may be true of living systems, **but how does it arise?**

Agency is costly across scales

- There are always **costs to diverging from a passive distribution**
  - **Chemical:** P_0 as passivity in reactions
  - **Neural:** P_0 as habit in behaviour
  - **Cultural:** P_0 as convenience in morality
- The complexity of higher levels renders this **effectively analogical**
  - We might pay **mental costs** for opposing our own habitual inertia
  - We might pay **social costs** for contradicting conventional immorality
- Such costs **may have an energetic basis**, but are more complex
  - Psychic costs may correspond to a **supervenient form of energy**
  - Interpersonally-mediated costs **relate to metabolism indirectly**

End.

bgas0204@uni.sydney.edu.au