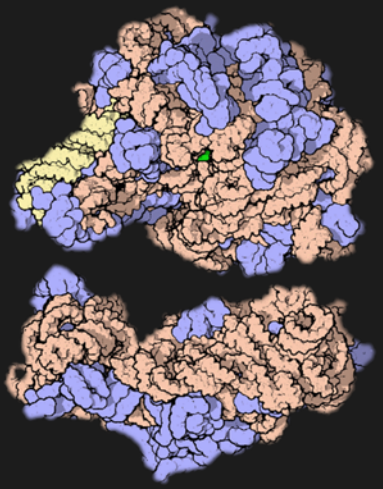
Physics of Biology, Geneva 2013

Yonatan Savir



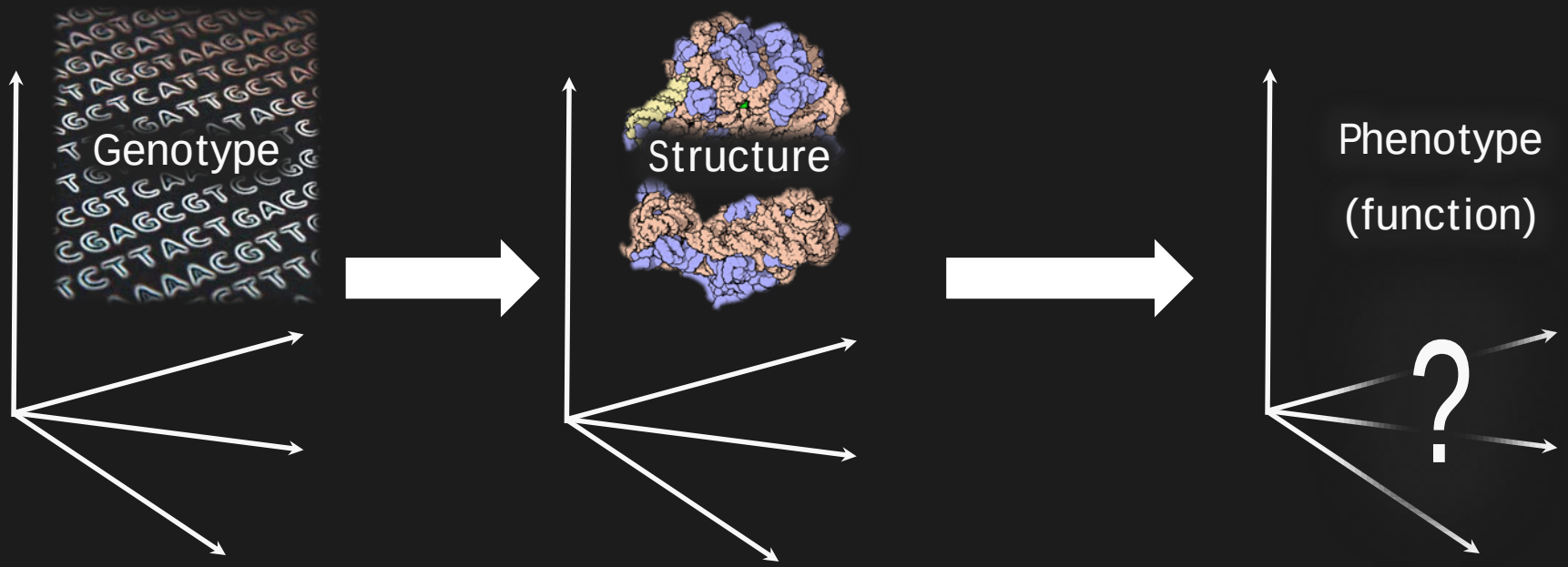
Ribosome

Function

?

Please ask questions

What is the dimension of phenotype space?



Dimension = # parameters ~ 10^4

What are the relevant degrees of freedom?

Flatland, where ribosomes evolve

Dimensional reduction

Molecular recognition problem

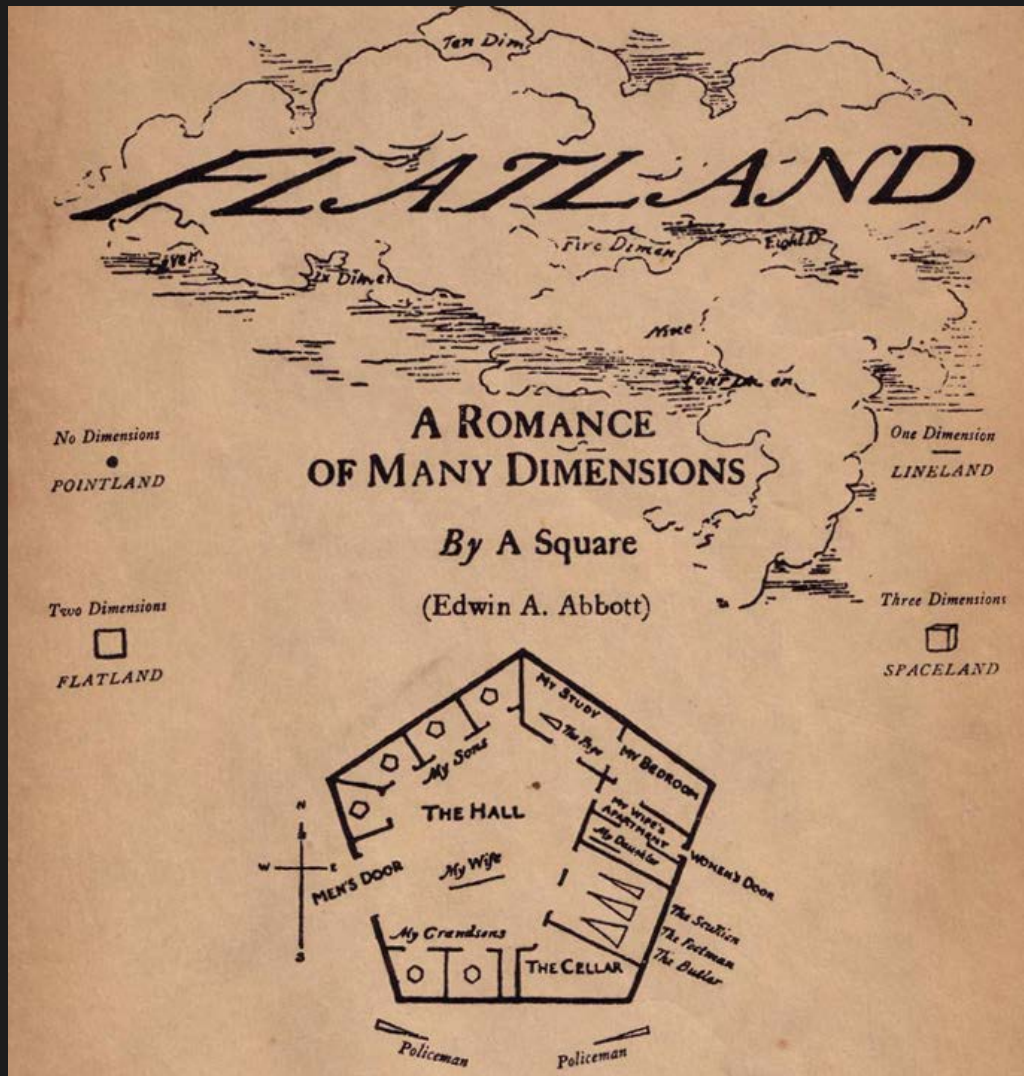
- "Fitness"?
- Conformational changes?

Generic principles

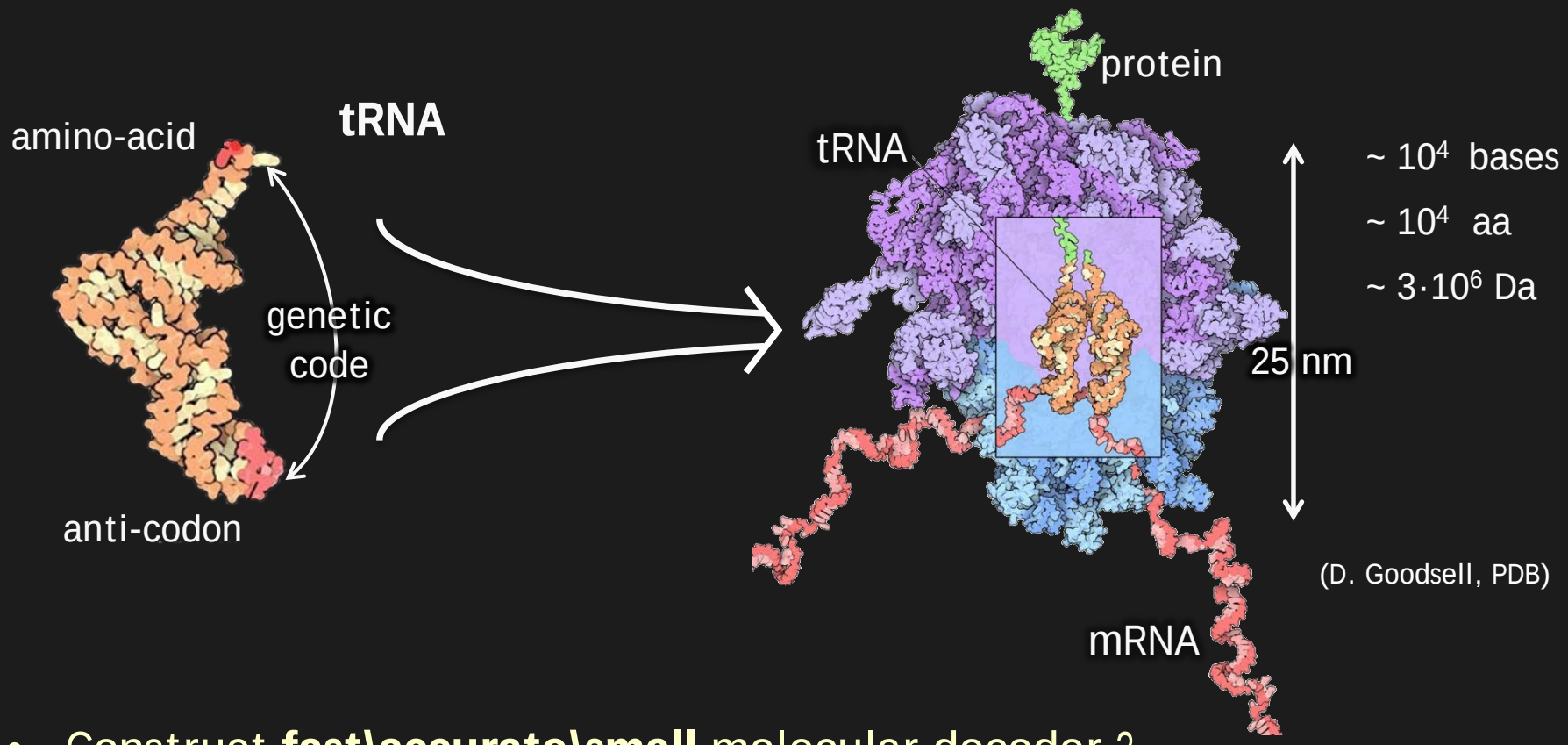
- Conformational tuning (induced fit).
- worst case scenario.

Other molecules: RecA, Rubisco…

.

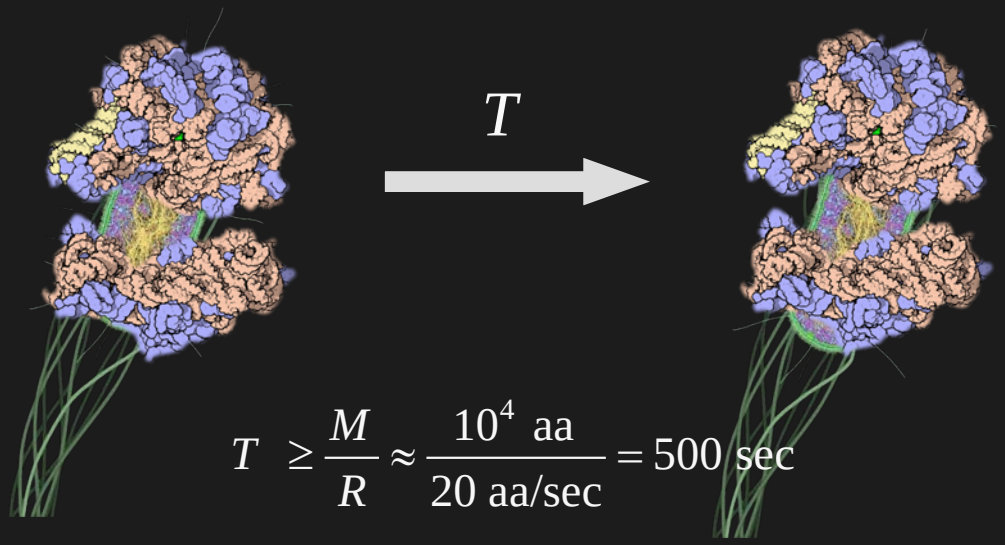


Ribosomes translate nucleic bases to amino acids



- Construct **fast\accurate\small** molecular decoder ?
- Relevant degrees of freedom ?

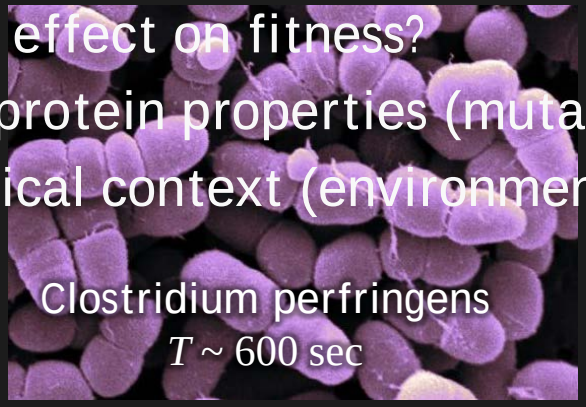
Ribosome limits self-replication rate

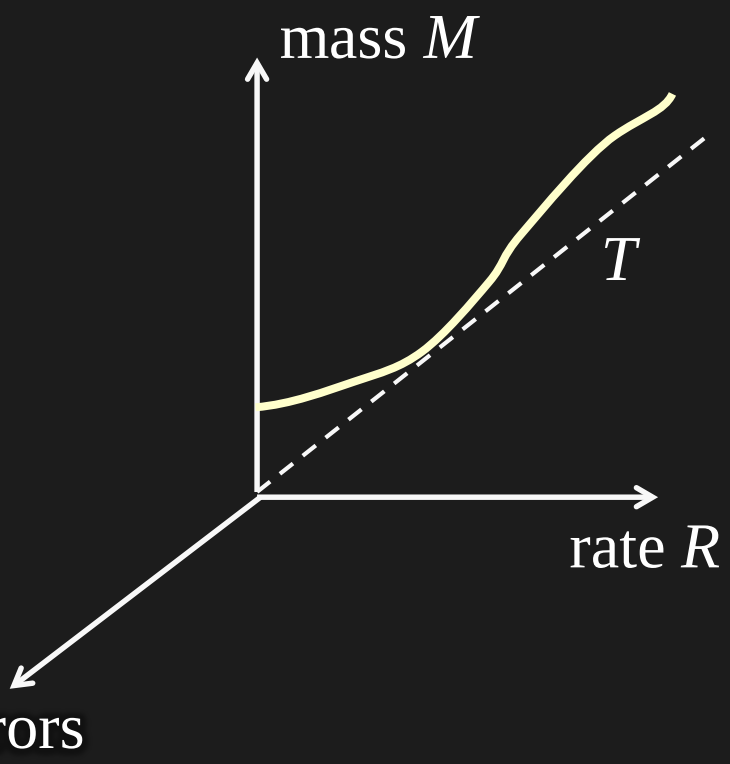


$$T \geq \frac{M}{R} \approx \frac{10^4 \text{ aa}}{20 \text{ aa/sec}} = 500 \text{ sec}$$

Accuracy effect on fitness?

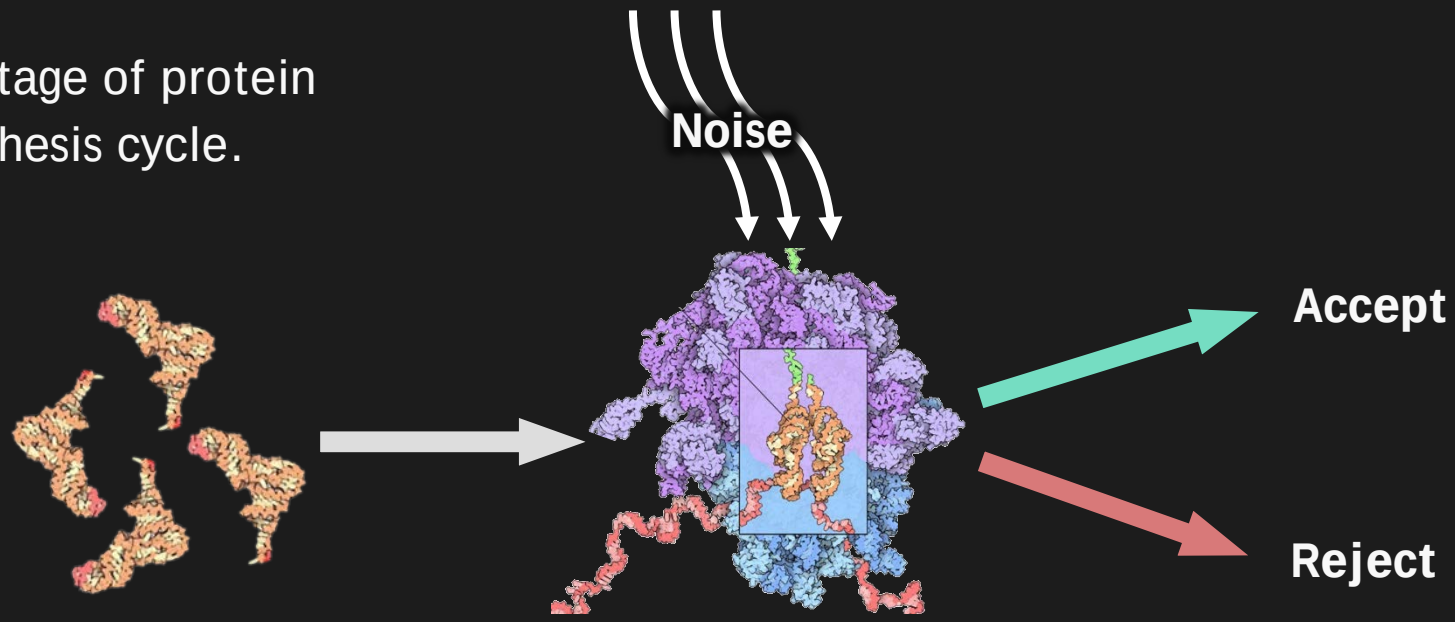
- Basic protein properties (mutations…)
- Biological context (environment…)





Decoding at the ribosome is a molecular recognition problem

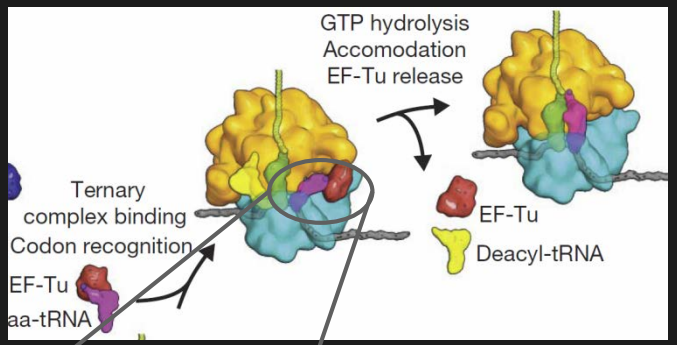
- 1st stage of protein synthesis cycle.



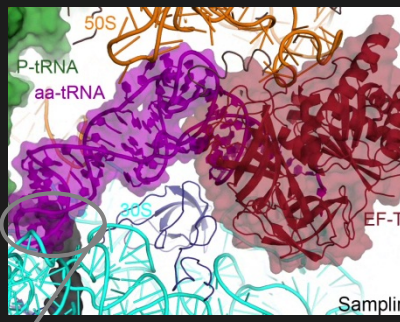
- How to evolve molecules that recognize in a **noisy** environment?
  (crowded, thermally fluctuating, weak interactions).

Ribosomal conformation changes place large energy barriers – that lack compelling explanation

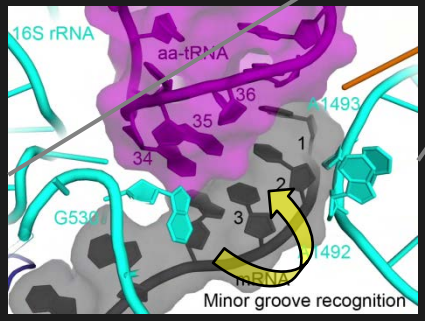
- Changes occur in all participating molecules: small\large subunits, EF-Tu and tRNA.

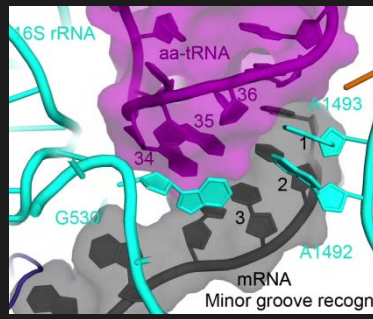


tRNA sampling



Minor groove recognition



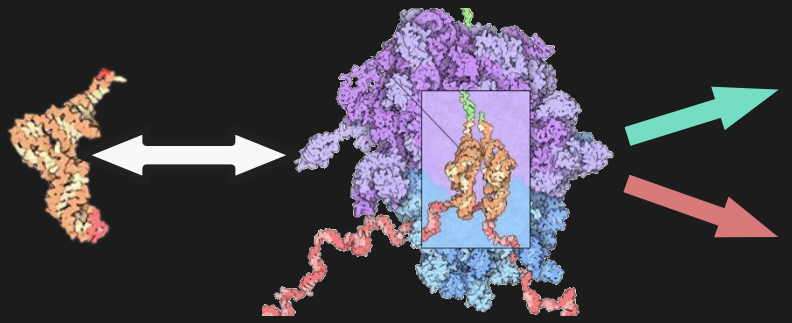


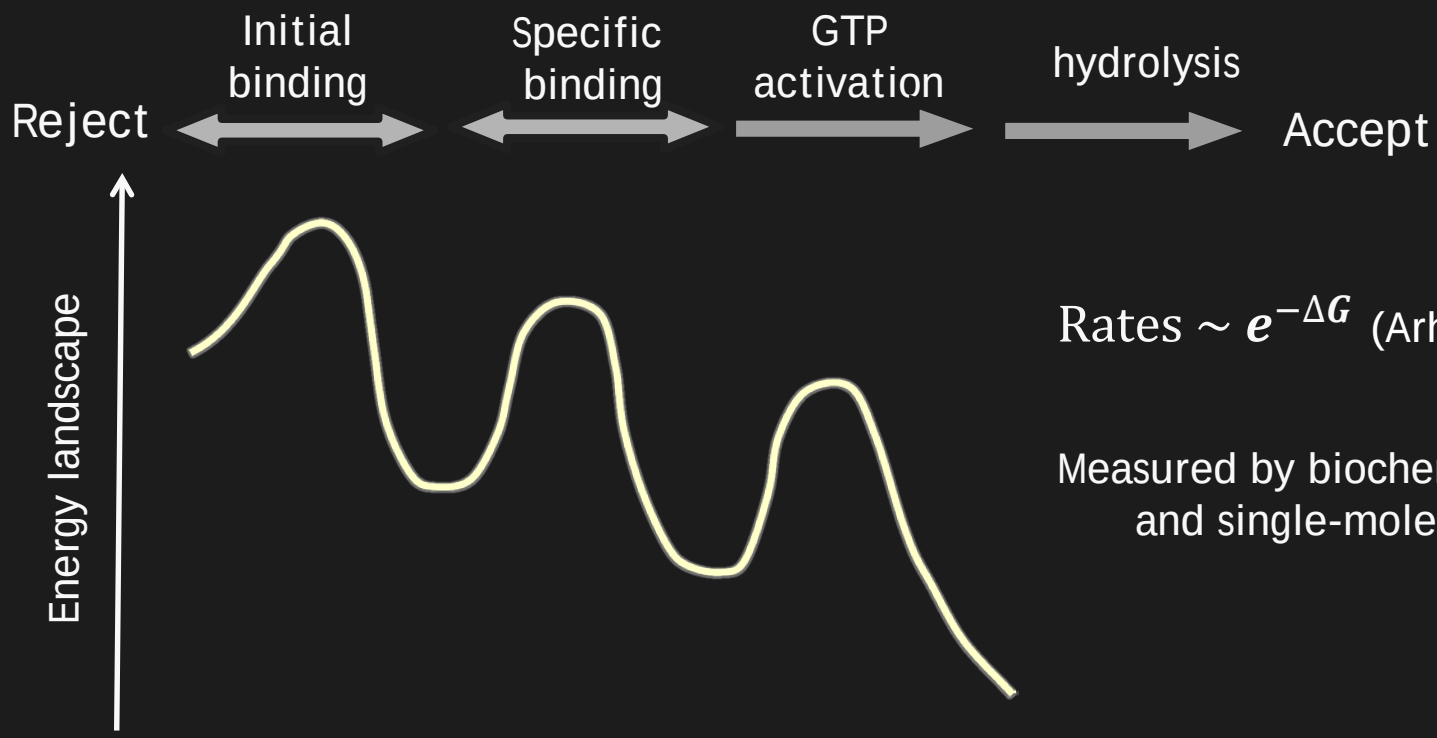
Ramakrishnan's lab

Is there a role to conformational changes in tRNA recognition?

Decoding governed by energy landscapes of tRNA recognition

- Measured kinetics: sequence of transitions.



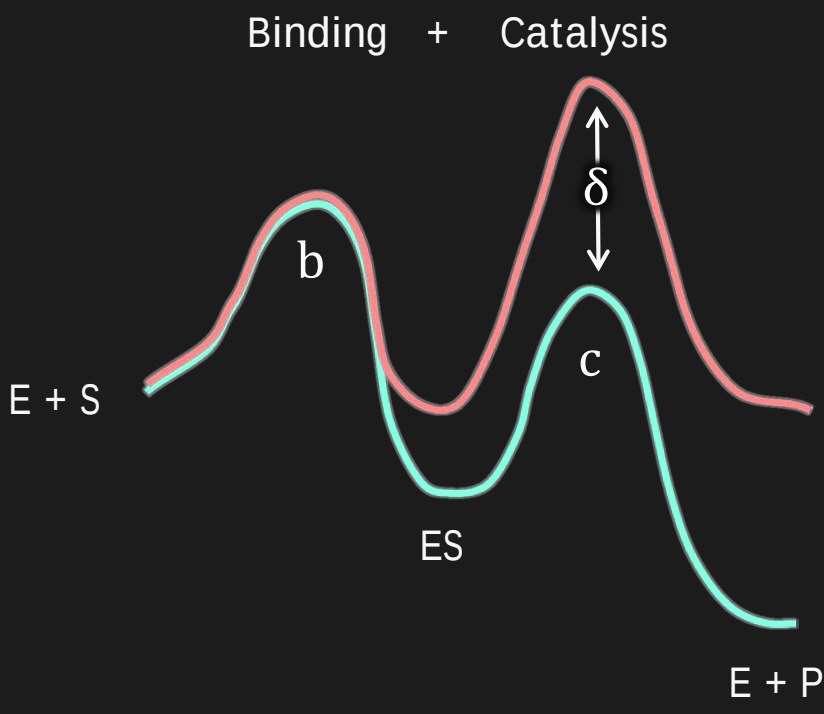


$\text{Rates} \sim e^{-\Delta G}$ (Arhenius, units of $k_B T$)

Measured by biochemistry
and single-molecule experiments.

The ribosome decoding problem is an example of recognition in the presence of competition

- Generic enzyme kinetics (MM): $E + S \leftrightarrow ES \rightarrow E + P$



$$R_B \sim \frac{1}{e^{b} + e^{c+\boldsymbol{\delta}}}$$

$$R_{\boldsymbol{G}} \sim \frac{1}{e^{\boldsymbol{b}} + e^{\boldsymbol{c}}}$$

- Analysis applies to enzymes: Balance rate R_{C} and specificity $R_{\mathrm{G}}/R_{\mathrm{B}}$?

Given the phenomenological tRNA decoding landscape:

- **Recognition performance (“fitness”) ?**
- **Effective dimension?**
- **Role of conformational changes?**

Yonatan Savir et al., *Cell* 2013

Recognition “fitness” has generic features

- “Fitness” *F* is obscure and context-dependent.:

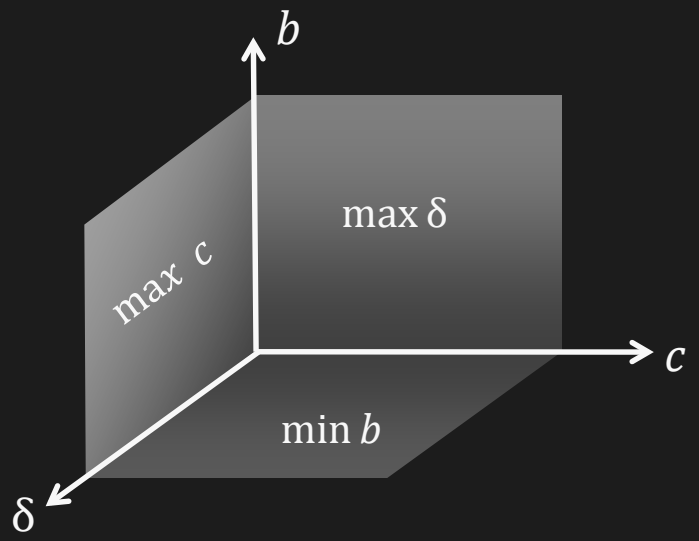
  → look for generic properties of $F(R_G(b,c), R_B(b,c+\delta))$

$$R_G \sim \frac{1}{e^b + e^c}$$

$$R_B \sim \frac{1}{e^b + e^{c+\delta}}$$

- “Biologically reasonable”: $\frac{\partial F}{\partial R_G} \geq 0,\ \frac{\partial F}{\partial R_B} \leq 0.$

- Searching for optimum in (b, δ, c) space:
  - **δ approaches biophysical limit.**
  - **Optimization is effectively 1D.**

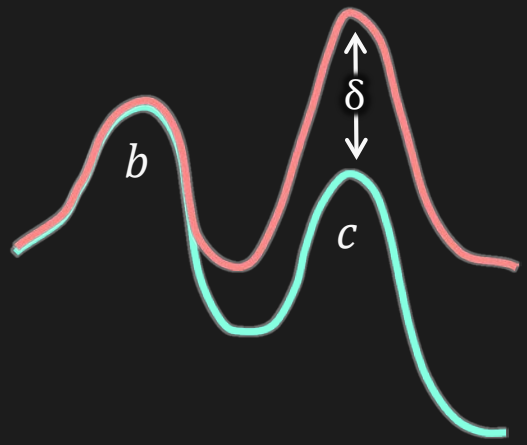


*** “Optima” are at infinity**

Adaptation in Lineland

- Example:

$$F = R_{\boldsymbol{G}} - d \cdot R_{\boldsymbol{B}} \propto \frac{1}{e^{b}+e^{c}} - \frac{d}{e^{b}+e^{c+\delta}}$$



- 1D problem: search along c.

(measured: $\Delta = c - b$)

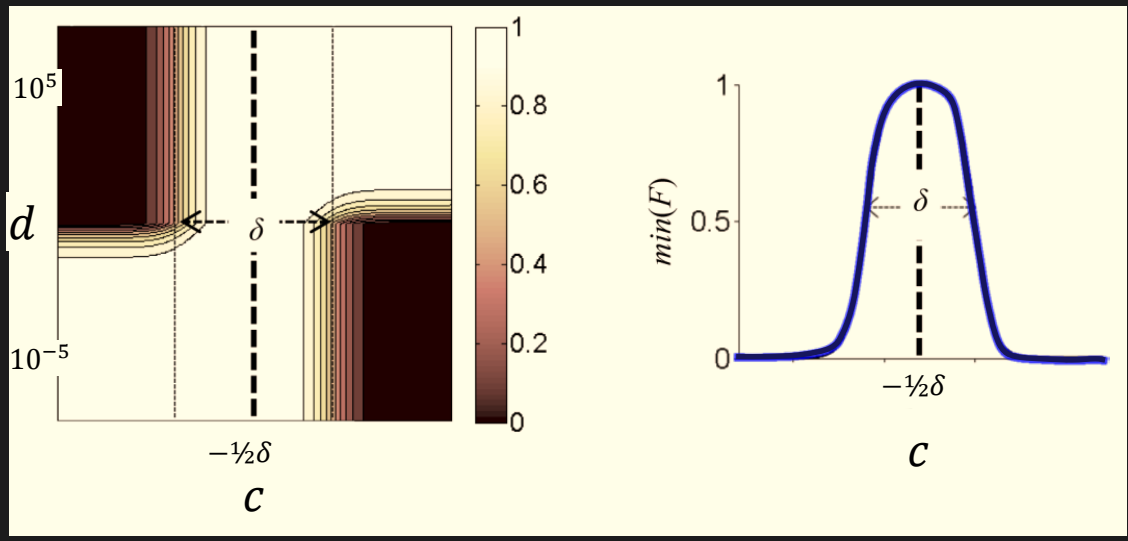
Has the ribosome adapted ?

Adaptation in Lineland

- Tuning parameter, e.g.

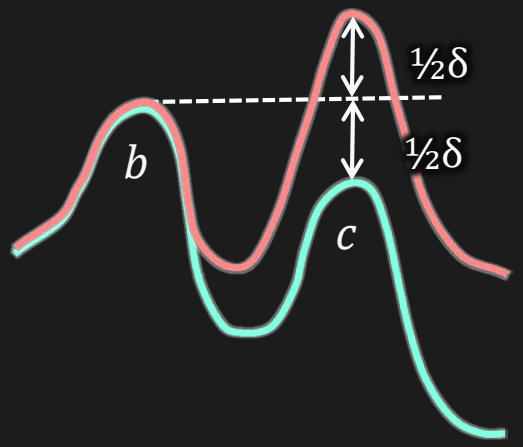
$$F = R_{\boldsymbol{G}} - d \cdot R_{\boldsymbol{B}}.$$

"Worst case scenario": $\max(\min(F))$.



- Max-Min solution is **symmetric:**

$$\Delta = c - b \approx -½\delta$$

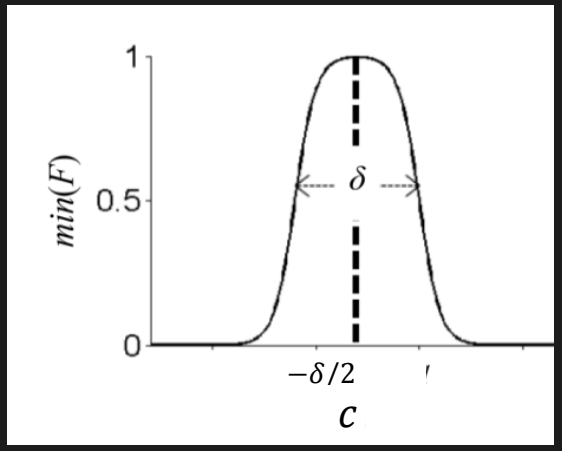


Solution is valid for wide range of fitness functions

- Ribosome optimal in wide region:

$$5\cdot 10^{-6} = e^{-\delta} \le d \le e^{\delta} = 2\cdot 10^{5}.$$

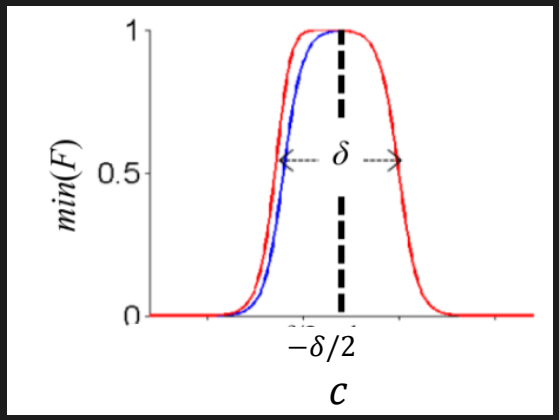
$F = R_G - d \cdot R_B$



- **General feature**: any fitness function $F(R_G, R_B)$ *exhibits maximum as long as both rates are "relevant".*

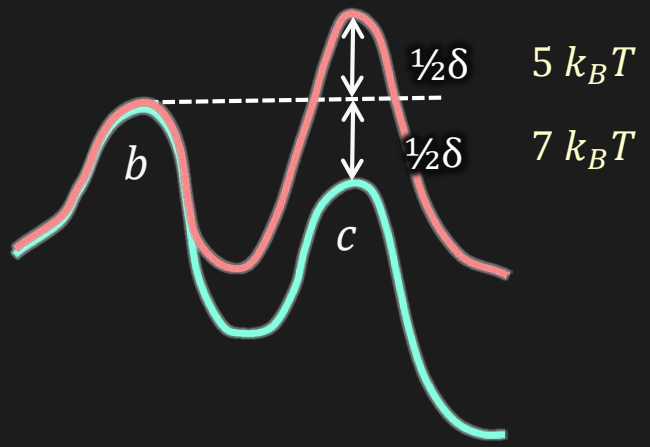
$$e^{-\delta} < \left| \frac{\partial F}{\partial R_B} \Big/ \frac{\partial F}{\partial R_B} \right| < e^{\delta}$$

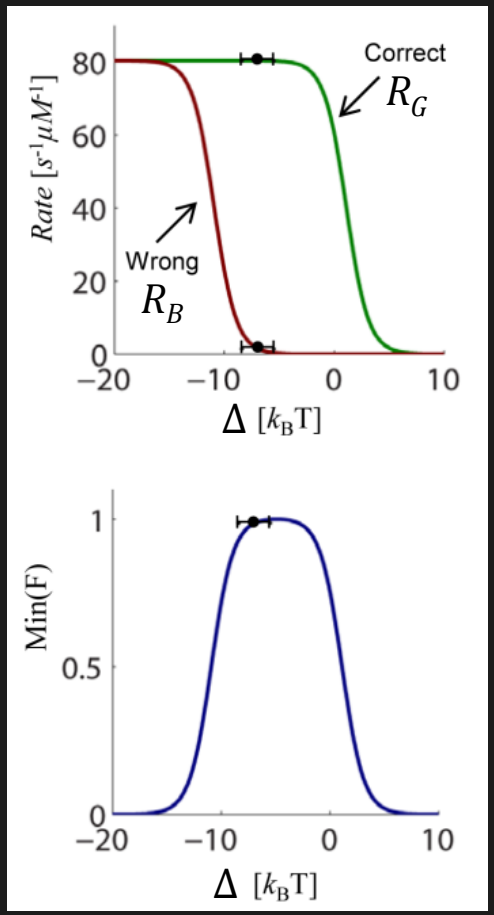
$R_G - d \cdot R_B{}^2$ $R_C - d \cdot (R_B/R_G)$



Ribosome energy barriers appear adapted

- **Max-Min predicts symmetry** $\Delta \approx -\frac{1}{2}\delta$.
- Measurements: $\Delta \approx -7\ k_BT,\ \delta \approx 12 k_BT$.

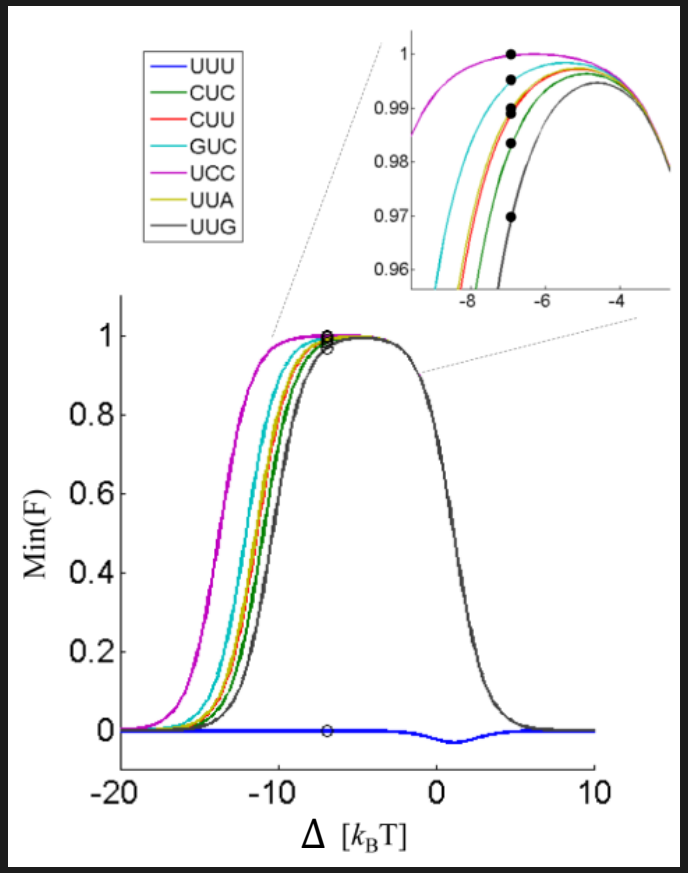




The ribosome shows adaptation

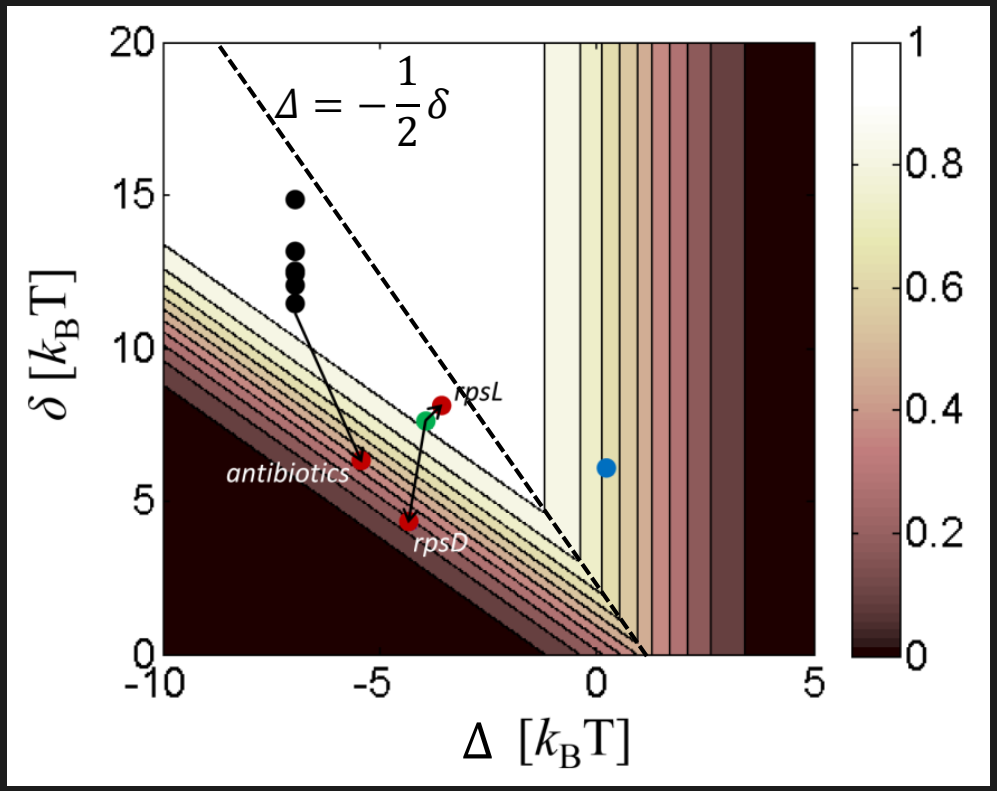
(according to Max-Min prediction).

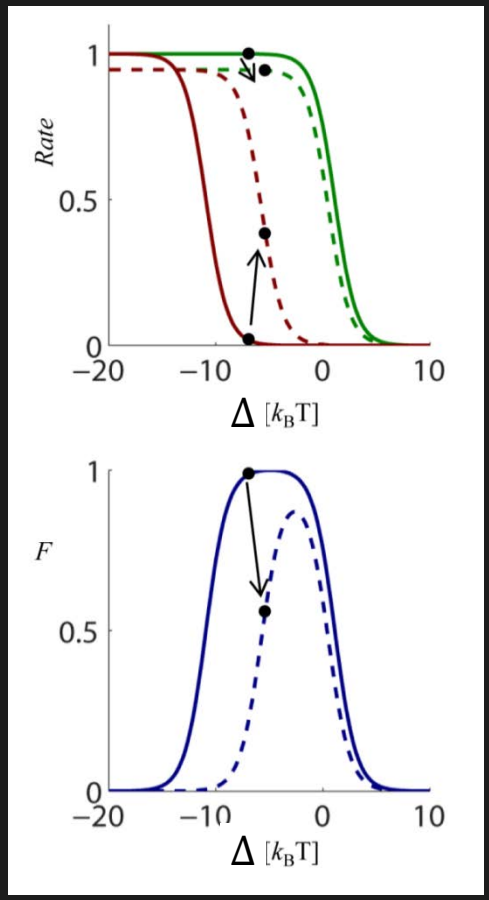
Decoding is adapted for all measured tRNAs



- Except for UUC: $\{\mathrm{UUC}, \mathrm{UUU}\} \rightarrow$ phenylalanine.

Predicted adaption regime of other ribosomes





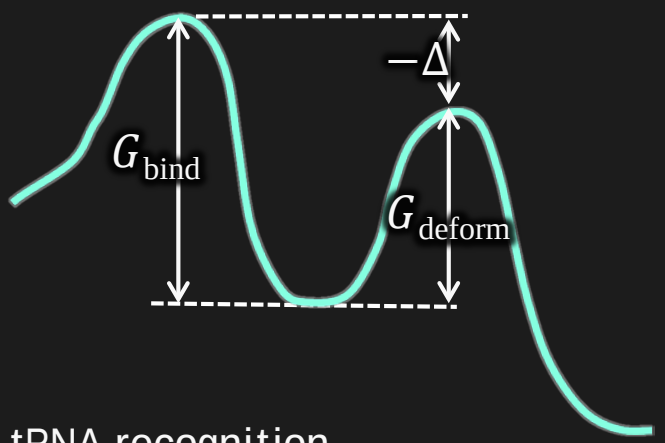
- Adaptation regime in the space of all possible landscapes, $-\delta \leq \Delta \leq 0$.
- Mutations\antibiotics push away from adaptation.

What is the role of conformational changes?

- Energy barrier includes binding and deformation energies:

$$\Delta = G_{\text{deform}} - G_{\text{bind}} = -\tfrac{1}{2}\delta \approx 6\ \mathrm{k_B T}.$$

- For any $G_{\text{bind}} \geq 6\ \mathrm{k_B T},\ G_{\text{deform}} \geq 0.$



- **Conformational proofreading:** deformation improves tRNA recognition.

Energy barrier that discerns the right target from competitors.

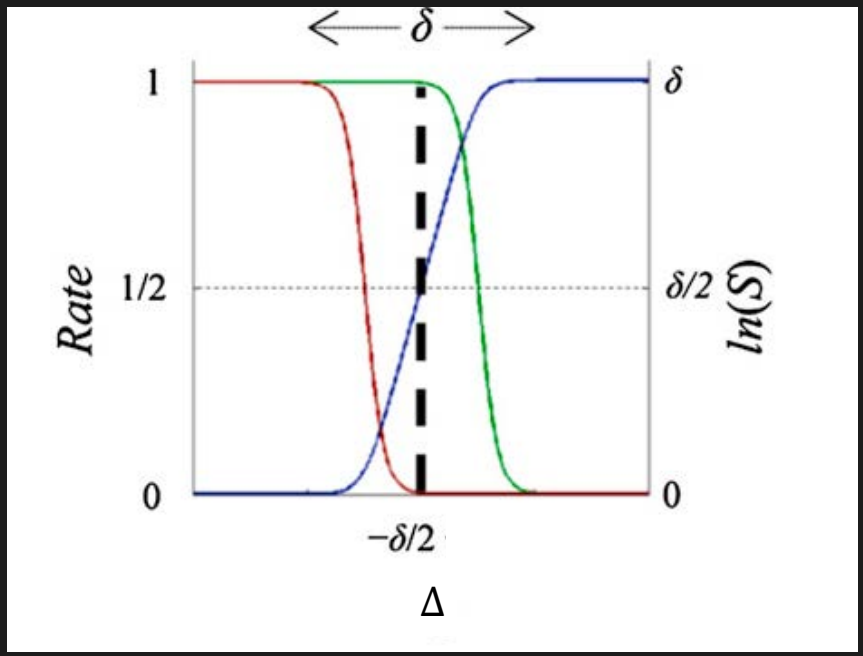
Enzymatic specificity is only square of the maximal one

- Interplay of rate R_G and specificity, $S = R_G/R_B$.
- Maximal specificity: $S \leq S_{max} = e^{\delta}$.

$$S^* = S_{max}^{1/2} = e^{\delta/2}$$

- Balance of rate and specificity.
- Hint for kinetic proofreading:

$$S^* \rightarrow (S^*)^2 = S_{max}$$



Yonatan Savir et al., *Cell* 2013

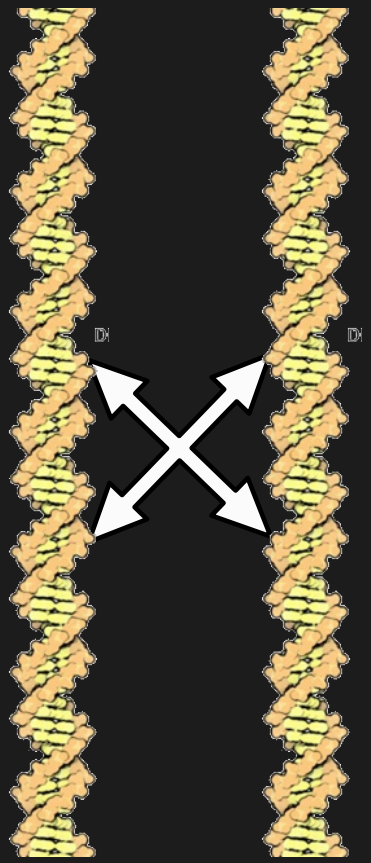
- Are these Generic properties?
- **What about Enzymes?**

Recombination machinery recognizes homologous DNA

- Exchange between two *homologous* DNAs.
- Essential for genome integrity and diversity.

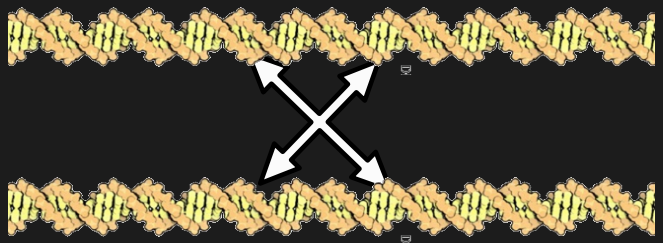
- Task: Detect homologous DNA target among many incorrect lookalikes.

- DNA *stretches* during recombination: large deformation energy barrier.

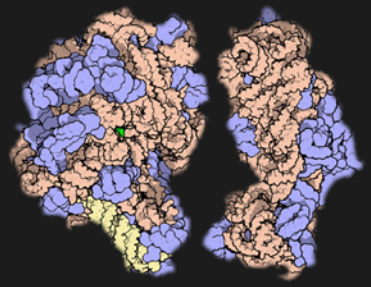


Yonatan Savir et al., *Mol Cell* 2010

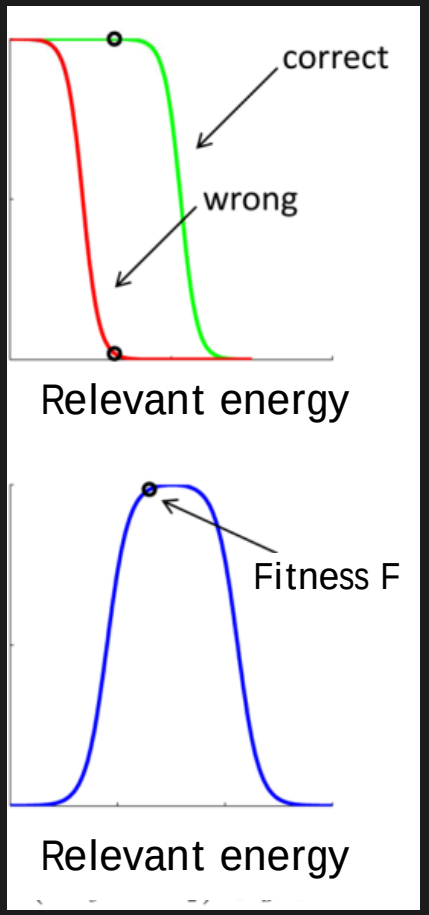
Energy barriers for optimal recognition may be a general design principle of recognition systems with competition



Recombination optimizes extension energy of dsDNA.



Ribosome optimizes energy barriers of decoding



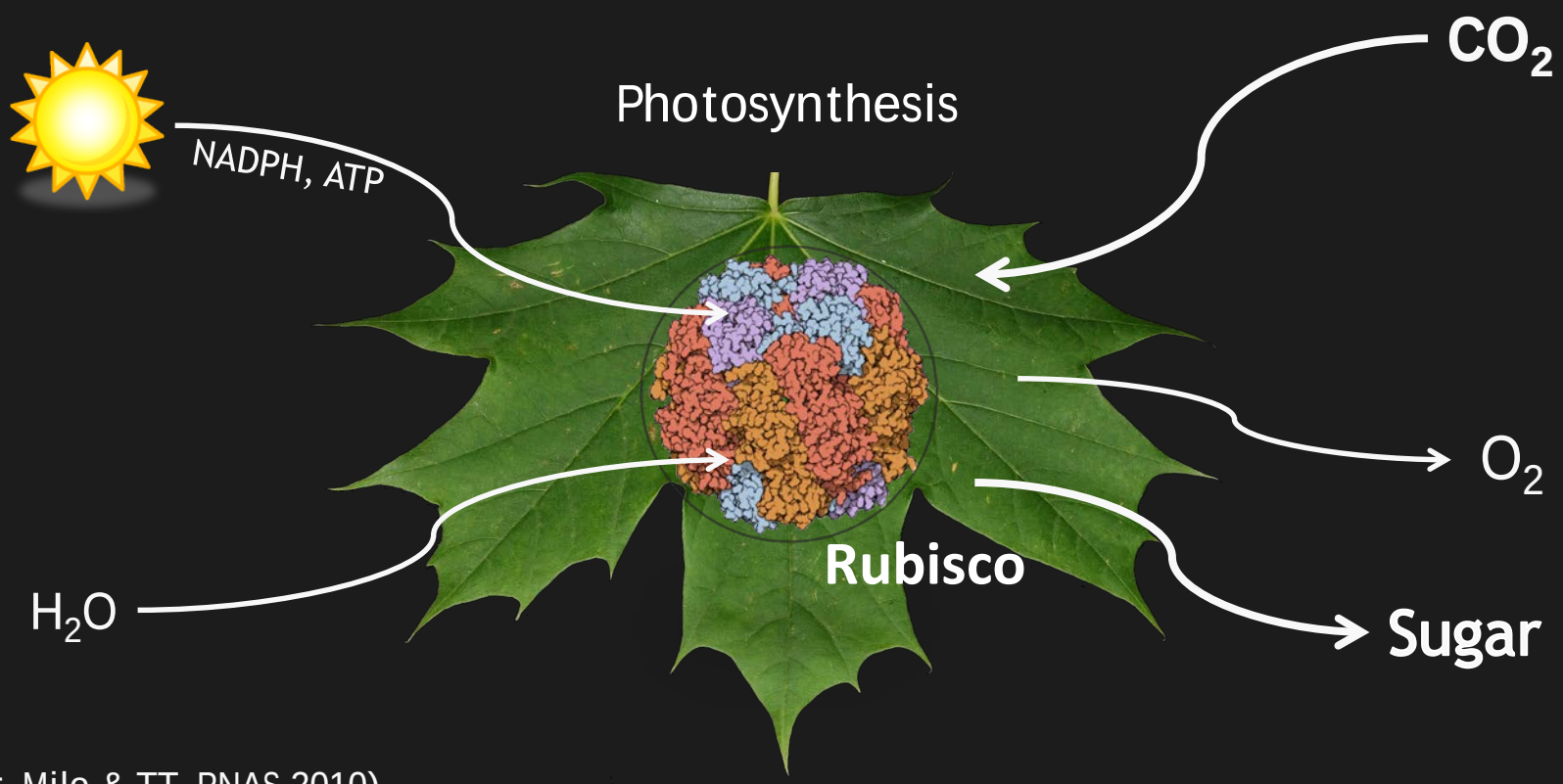
Cees Dekker et al., *Mol Cell*, 2012

- **Conformational proofreading:** mechanism follows from optimization of information transfer function.
- May explain **induced fit** (Koshland 1958).
  Why molecules deform upon binding to target?

- **Applies to any enzymatic kinetics in the presence of competition…**

Rubisco captures atmospheric CO_2 to make sugar

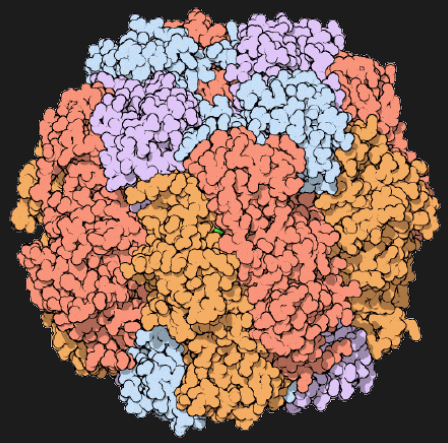
- Photosynthesis fixes carbon into organic forms.
- Rubisco captures CO_2.
- Complex of 8 large + 8 small subunits (540 kDa).



(Savir, Noor, Milo & TT, PNAS 2010)

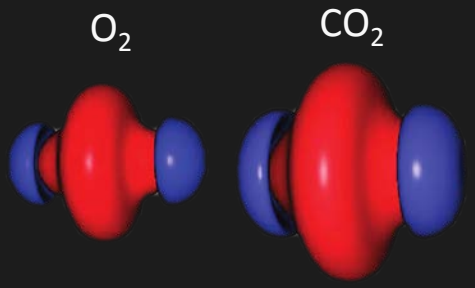
Rubisco is thought to be a slow and confused enzyme

- Most abundant protein on Earth.
- Catalyzes most carbon fixation.



- **Very slow** catalysis (~ 3-10 CO_2 /sec).
- **Low specificity:** confuses O=C=O and O=O.

- Can Rubisco be improved? – little success so far.



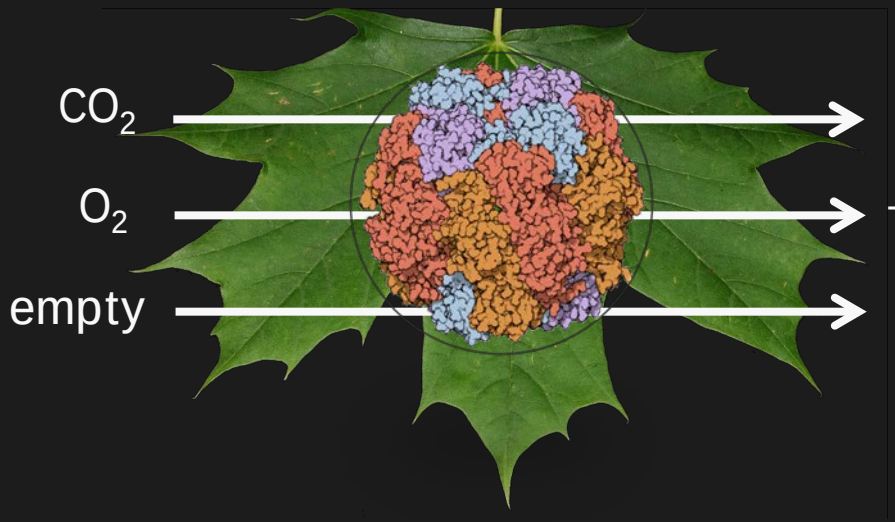
- Already optimized by evolution?
- Inefficiency due to biochemical constraints?

- Is Rubisco **optimal**, **constrained** or both?

Rubisco needs to recognize CO_2 above a high background of O_2

- Four parameters:2 affinities to CO_2 and O_2 : $\{K_C \; K_O\}$
  2 catalysis rates : $\{\nu_C \quad \nu_O\}$

Atmosphere:
20% O_2
0.04% CO_2



- Overall fitness = net photosynthesis rate (carbon fixed)

$$F = \sum_b \nu_b \times P_b = \nu_C P_C - \tfrac{1}{2}\nu_O P_O + 0 \cdot P_{\text{empty}}$$

- Optimal Rubisco?

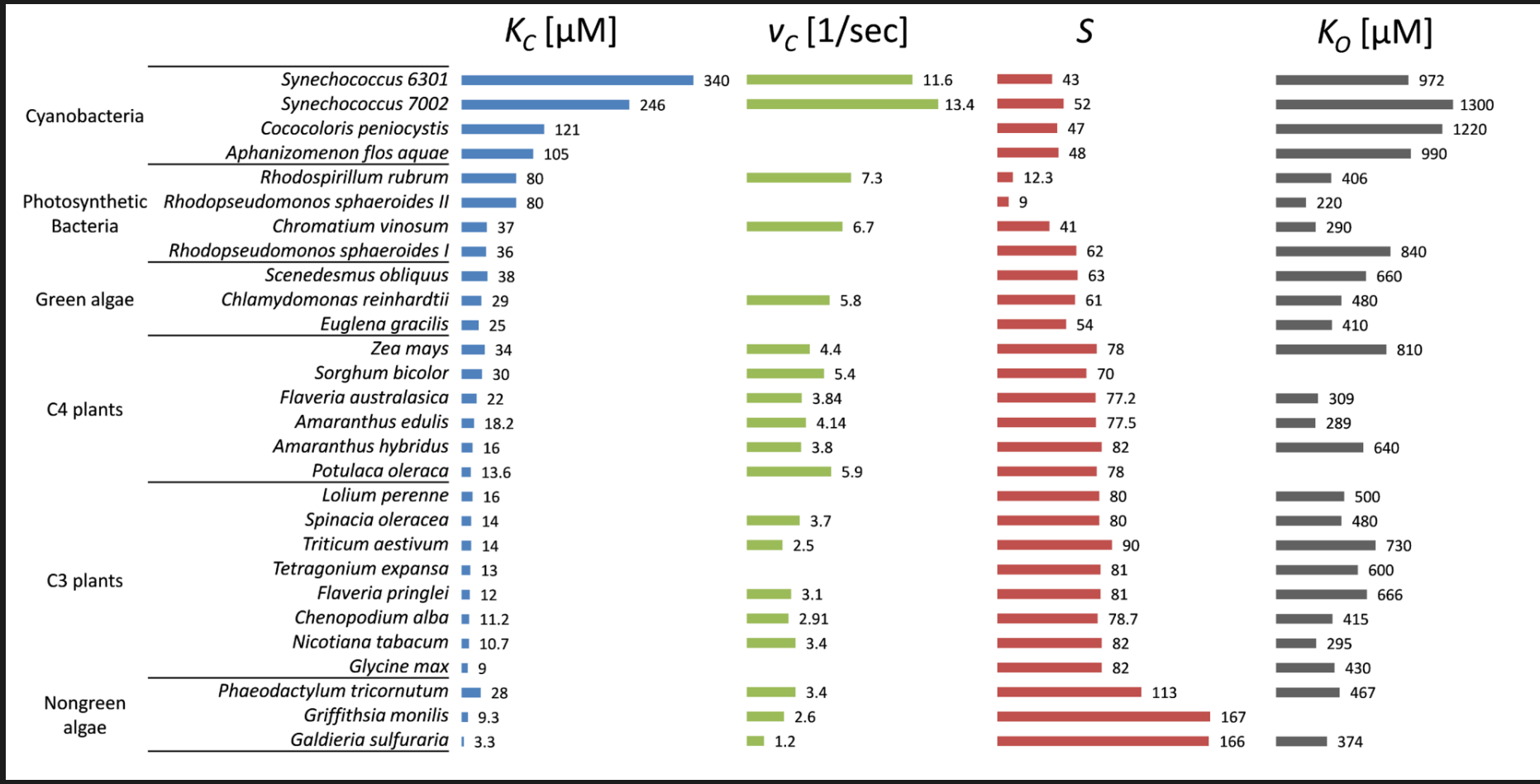
3-state partition

$$P_C = \frac{\frac{[CO_2]}{K_C}}{1 + \frac{[CO_2]}{K_C} + \frac{[O_2]}{K_O}},$$

$$P_O = \frac{\frac{[O_2]}{K_C}}{1 + \frac{[CO_2]}{K_C} + \frac{[O_2]}{K_O}}.$$

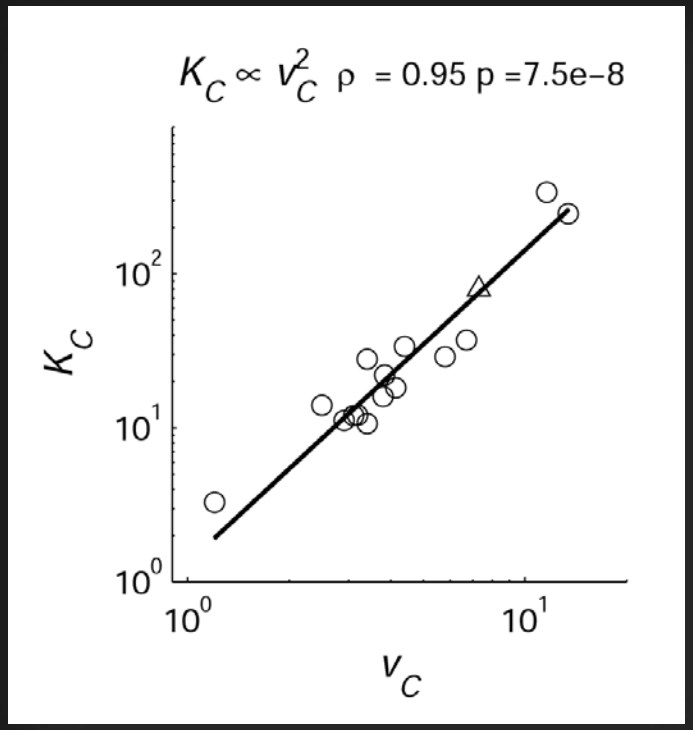
Rubisco kinetic parameters were measured in diverse organisms

- Analysis of data from 28 eukaryotes and prokaryotes:



Rubisco's parameter space is effectively 1D

- Data resides in **4D** space (S, v_C , K_C , K_O).
- But constrained to **1D power law.**
- PCA: >90% is 1D → power law correlations:



$$K_C \propto v_C^{2.0\pm0.2}$$

$$K_O \propto v_C^{0.5\pm0.1}$$

$$S \propto v_C^{-0.5\pm0.1}$$

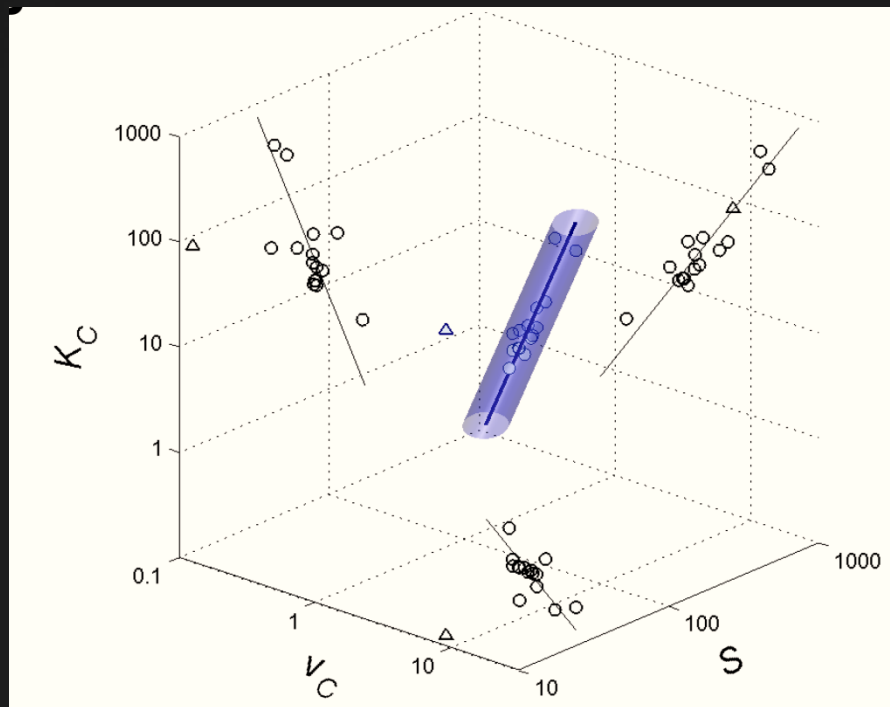
Dimensional reduction

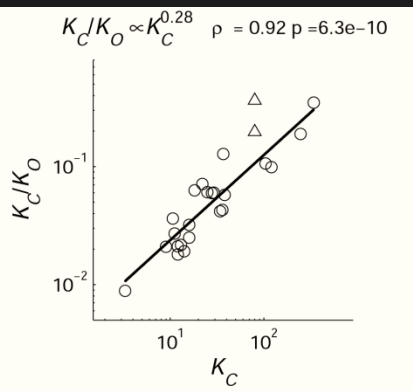
("Constrained Plasticity")

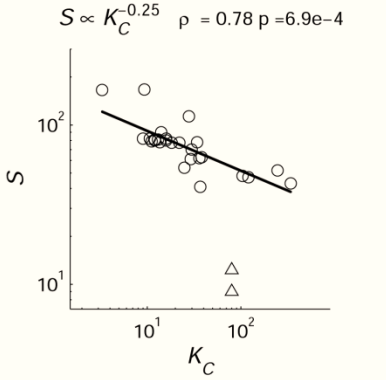
(Savir, Noor, Milo et al., PNAS 2010)

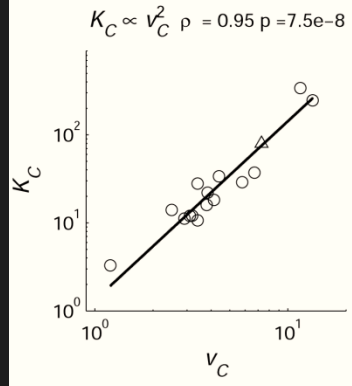
Effective 1D landscape for Rubisco evolution

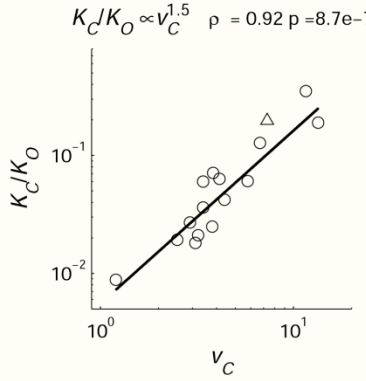
- All organisms scattered around a straight line in 4D parameter space.

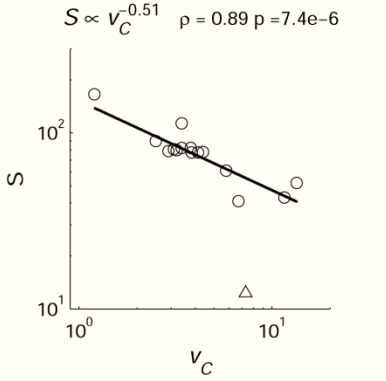






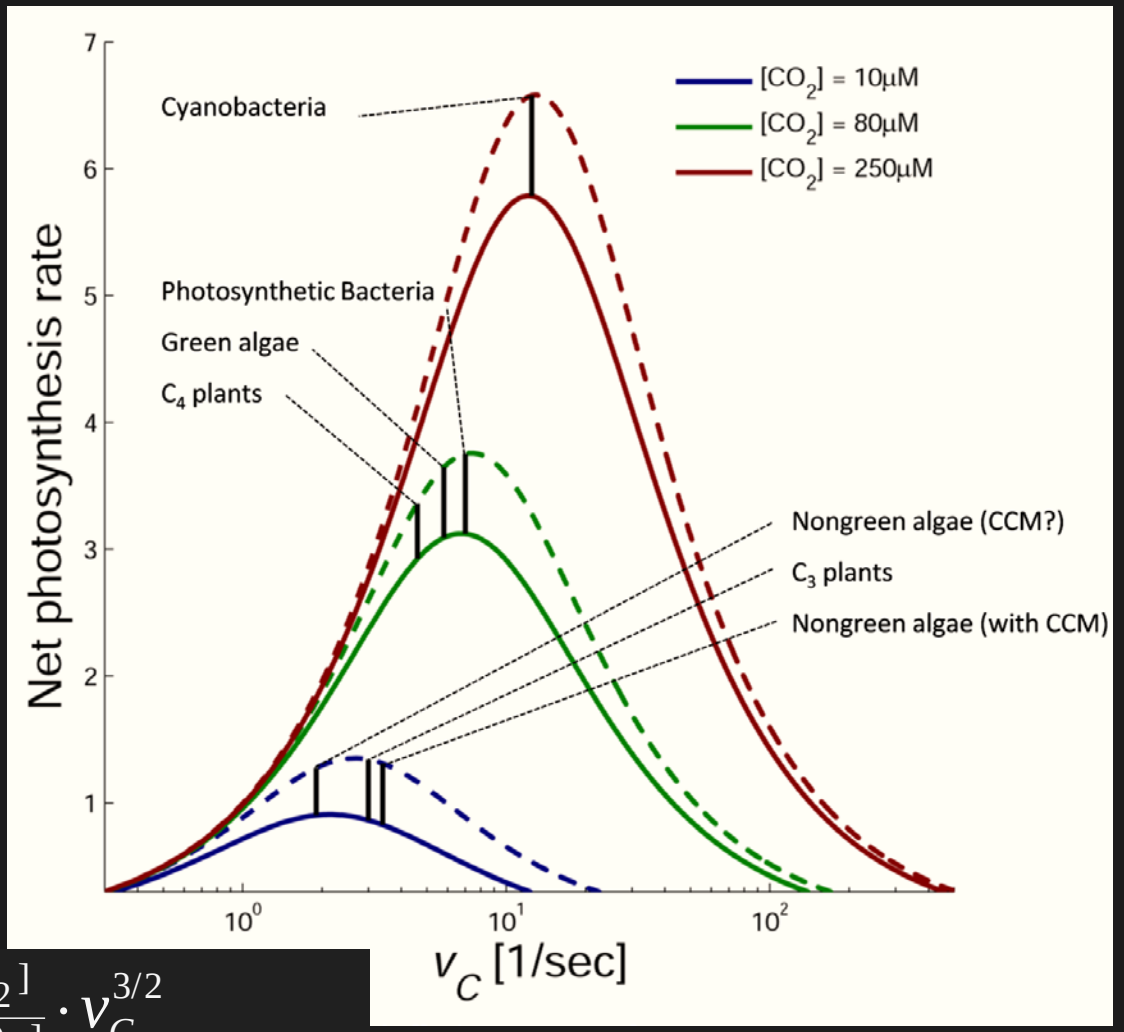






Rubisco energy barriers are nearly optimal for recognition of CO_2 in their habitats

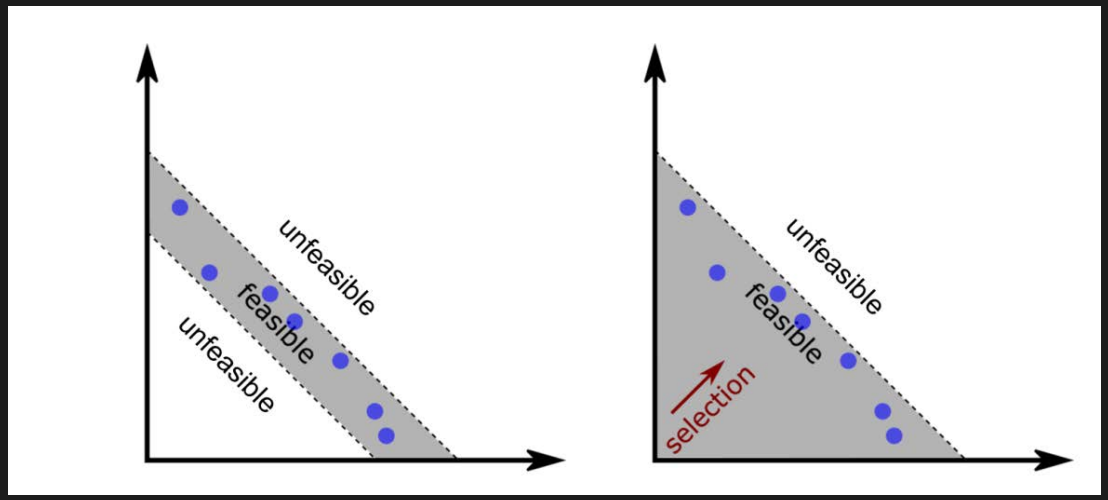
- Max(F) as function of v_C in given habitat (O_2 + CO_2).



$$F(v_C) = \frac{v_C - 3 \cdot 10^{-3} \cdot \frac{[O_2]}{[CO_2]} \cdot v_C^{3/2}}{1 + 1.3 \cdot \frac{1}{[CO_2]} \cdot v_C^2 + 5 \cdot 10^{-3} \cdot \frac{[O_2]}{[CO_2]} \cdot v_C^{3/2}},$$

What is the topology of molecular phenotype space?

- **Interplay** of constraints and evolution?

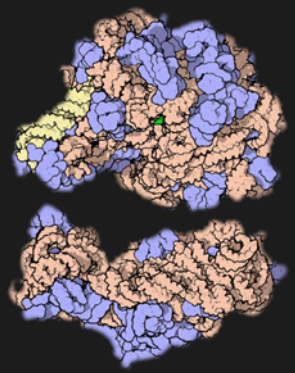


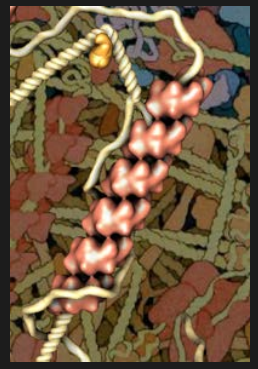
(hint: “soft” fluctuations)

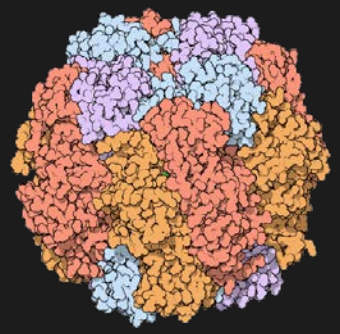
- Response to **long term climate changes**? (CO_2, Temp….)

- **Molecular phenotype spaces are low-dimensional.**
  (it that really generic?)

- **Simple thinking about function may explain some features and yield basic principles of molecular recognition systems.**





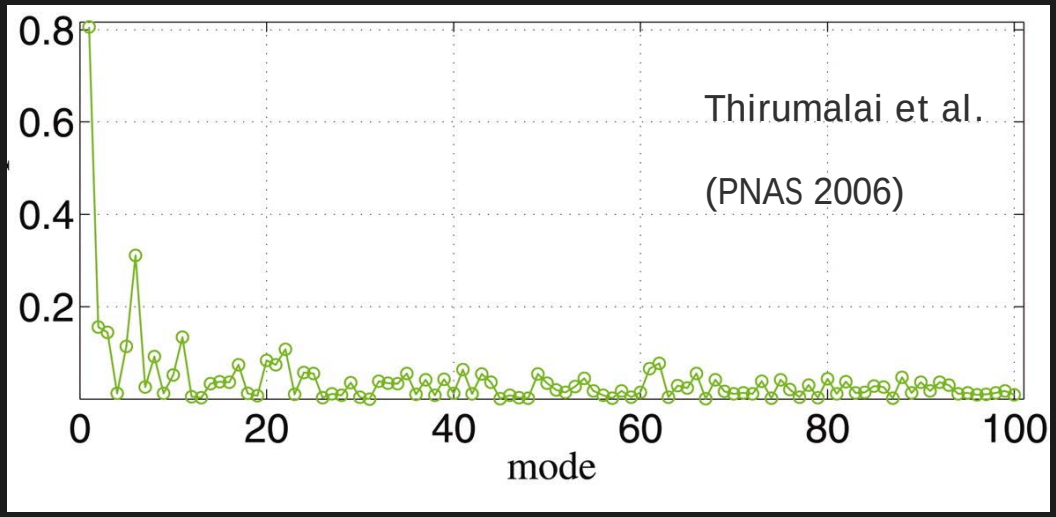


Thanks:
Yonatan Savir
Elad Noor
Ron Milo

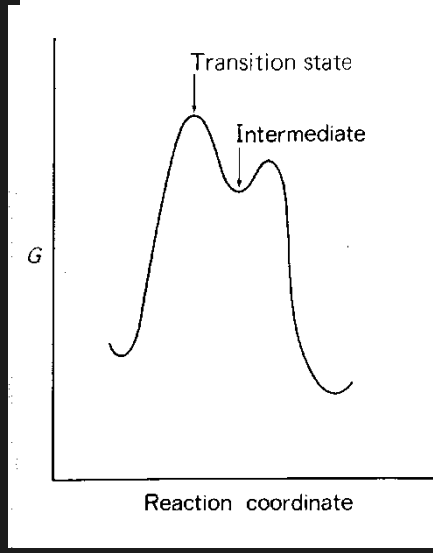
Yonatan Savir: *Cell* 2013, *Molecular Cell* 2010, *PNAS* 2010

What is the relation to basic physical properties of proteins as an evolvable state of matter?

- What are the **degrees-of-freedom** underlying dimensional reduction?
  - Hints: transition states effectively describe kinetics.
  - Low modes dominate protein function.
  - "Normal modes" (Ranganthan, Leibler, Rivoire…).



$$k_1 = \left(\frac{kT}{h}\right) \exp\left(\frac{-\Delta G^{\ddagger}}{RT}\right)$$



Transition states reduce the dimensionality of effective parameter space

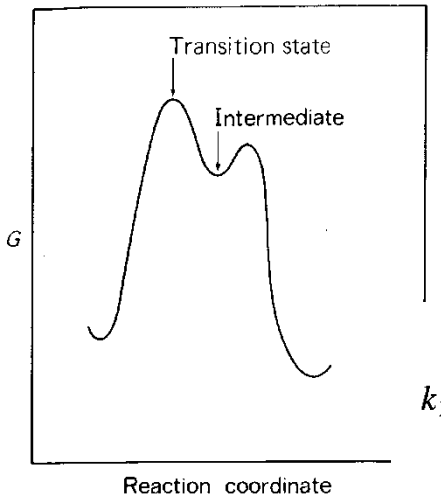


Figure 2.1 Transition states occur at the peaks of the energy profile of a reaction, and intermediates occupy the troughs.

$$k_1 = \left(\frac{kT}{h}\right) \exp\left(\frac{-\Delta G^{\ddagger}}{RT}\right)$$

FEBRUARY, 1935 JOURNAL OF CHEMICAL PHYSICS VOLUME 3

The Activated Complex in Chemical Reactions

HENRY EYRING, *Frick Chemical Laboratory, Princeton University*

(Received November 8, 1934)

leave no doubt of the proposed method of procedure in a particular case. We may write for the specific reaction rate constant for a reaction of any order

$$k_i = c(F_a/F_n)(\bar{p}/m^*) = c(F_a'/F_n)(kT/h)e^{-E_0/kT} \tag{10}$$

where F_a is simply the partition function (or Zustandsumme) for the activated state and F_n is the same quantity for the normal state. F_a' is the partition function for the activated complex for all the normal coordinates except the one in which decomposition is occurring. The partition function for this normal coordinate is included in the factor $(kT/h)e^{-E_0'/kT}$. The other quantities have been defined.

Thanks:

Yonatan Savir
Elad Noor
Ron Milo

Yonatan Savir et al.: *Cell* 2013, *Molecular Cell* 2010, *PNAS* 2010

Questions, future directions…

- Basic logic of molecular information systems from simple analysis.

- **Constrained phenotypic plasticity** in low dimensional landscapes:

  Generic phenomenon in proteins?

- What are the **degrees-of-freedom? ("modes")**
- Basic physical properties of proteins as an evolvable state of matter.

Acknowledgements: Yonatan Savir

- Recombination: Savir et al., *Mol Cell* 2010.
- Rubisco: Savir et al., *PNAS* 2010.
- Ribosome: Savir et al., *Cell* 2013.