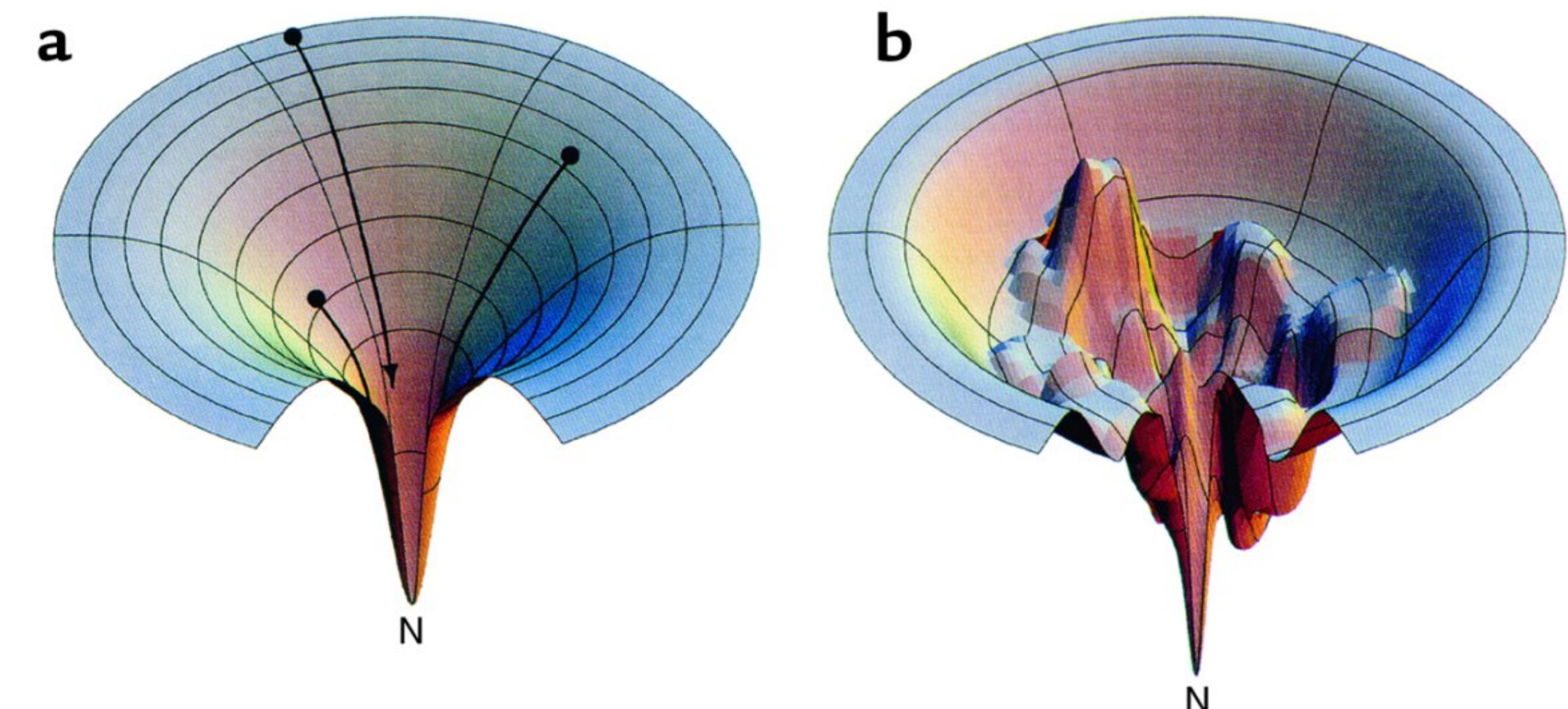
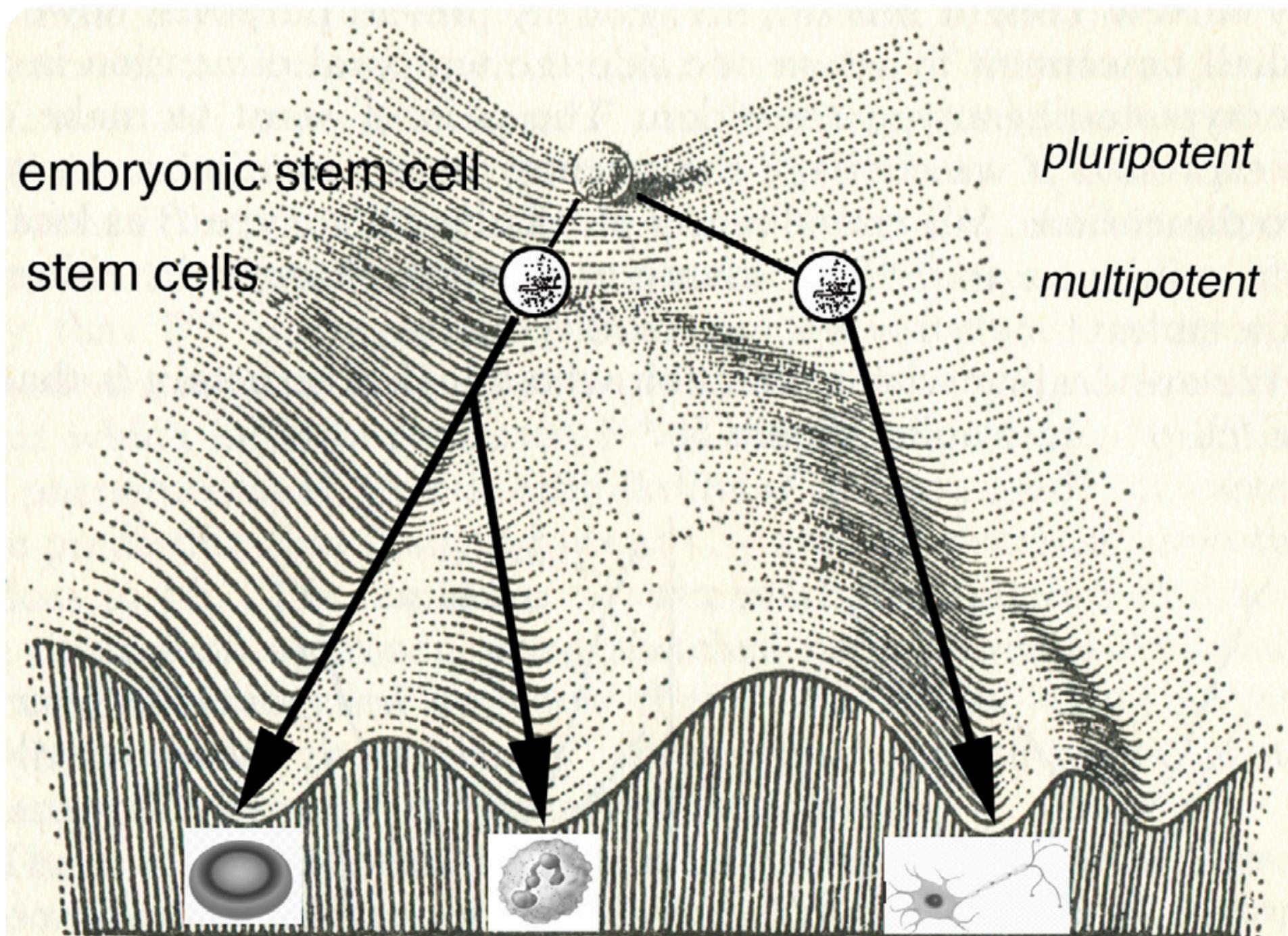
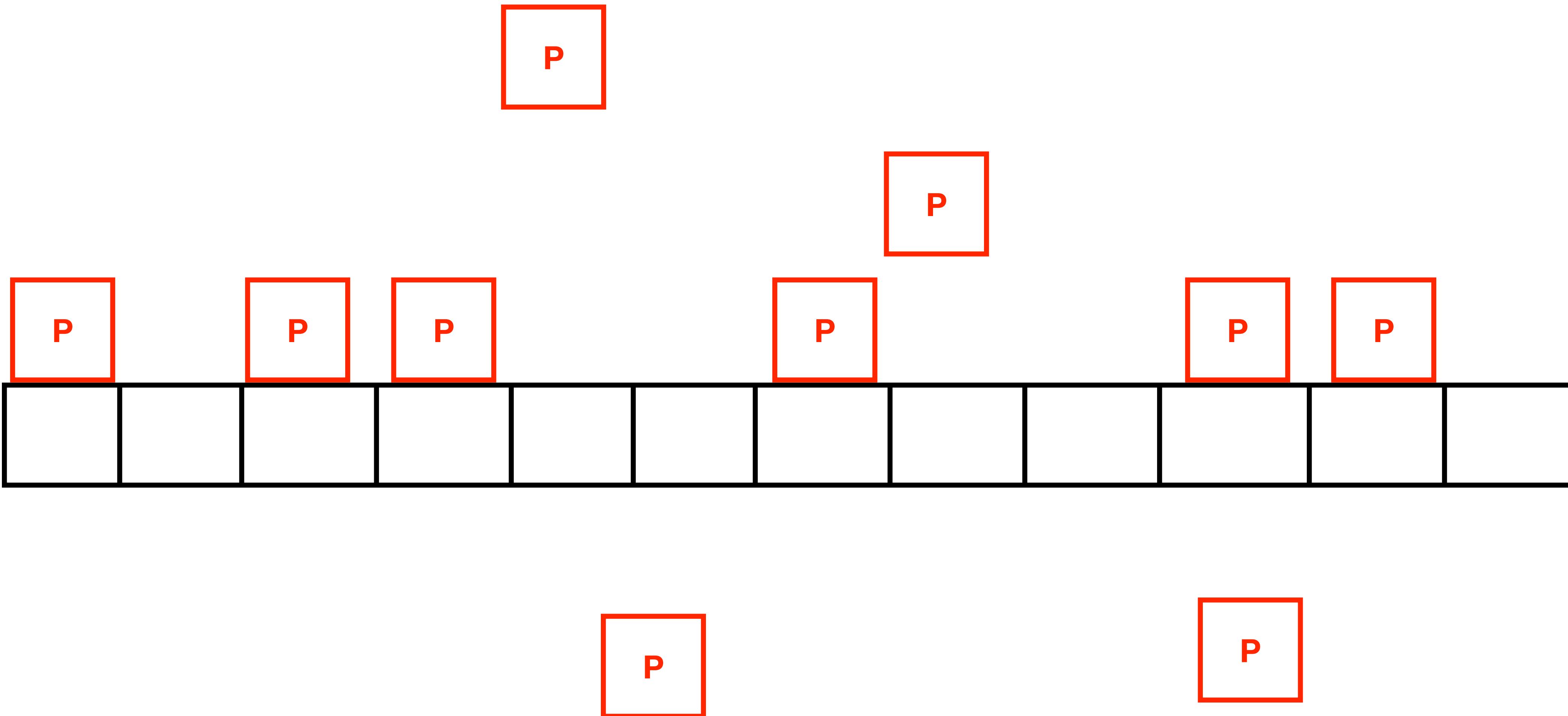


Molecular biology events as movements in a free energy landscape of hills and valleys

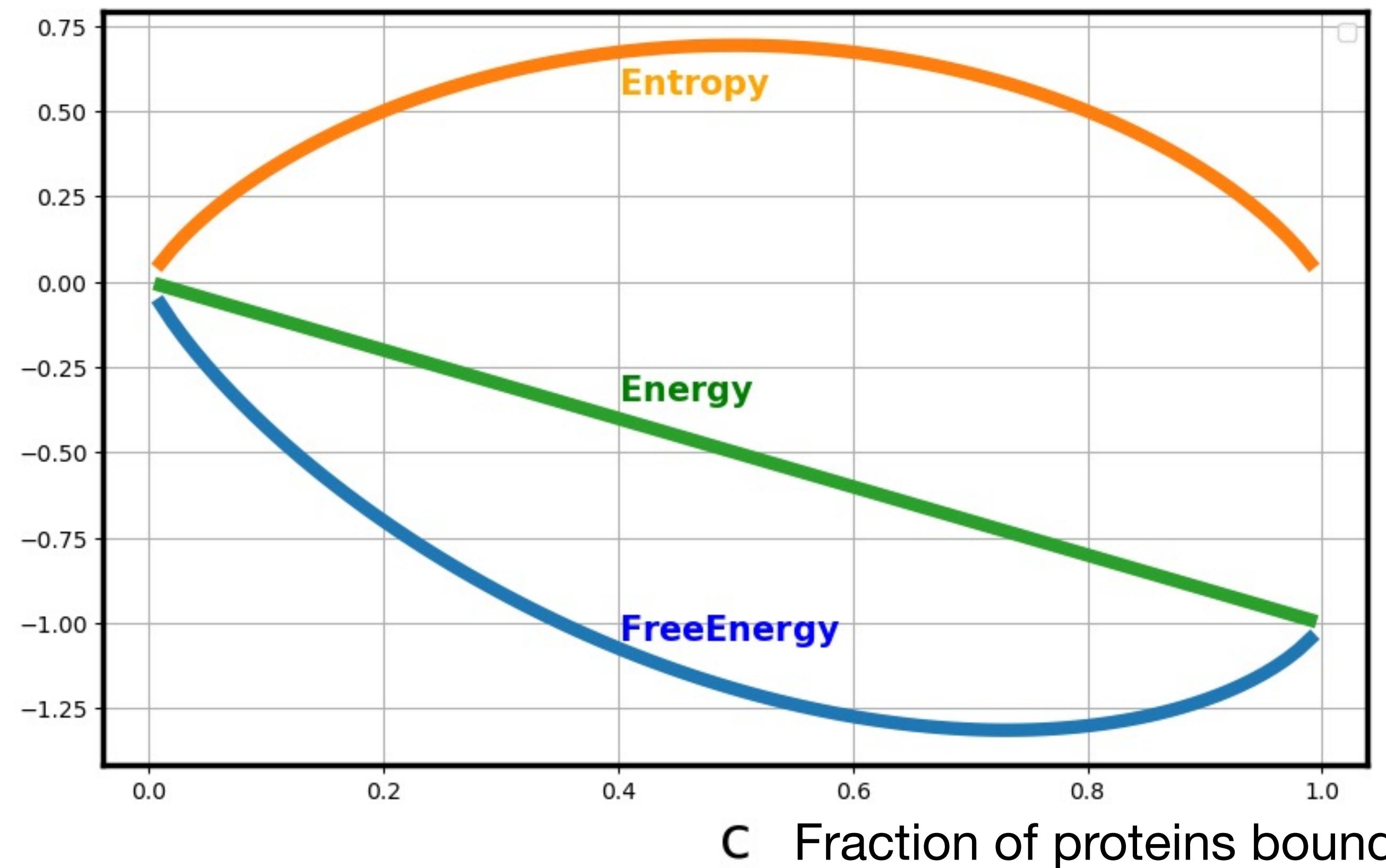
The height of these hills and depth of the valleys decide rate, stability etc in biology



Consider a bunch of proteins near DNA; can you predict what fraction of the DNA bindings sites would be covered by proteins?

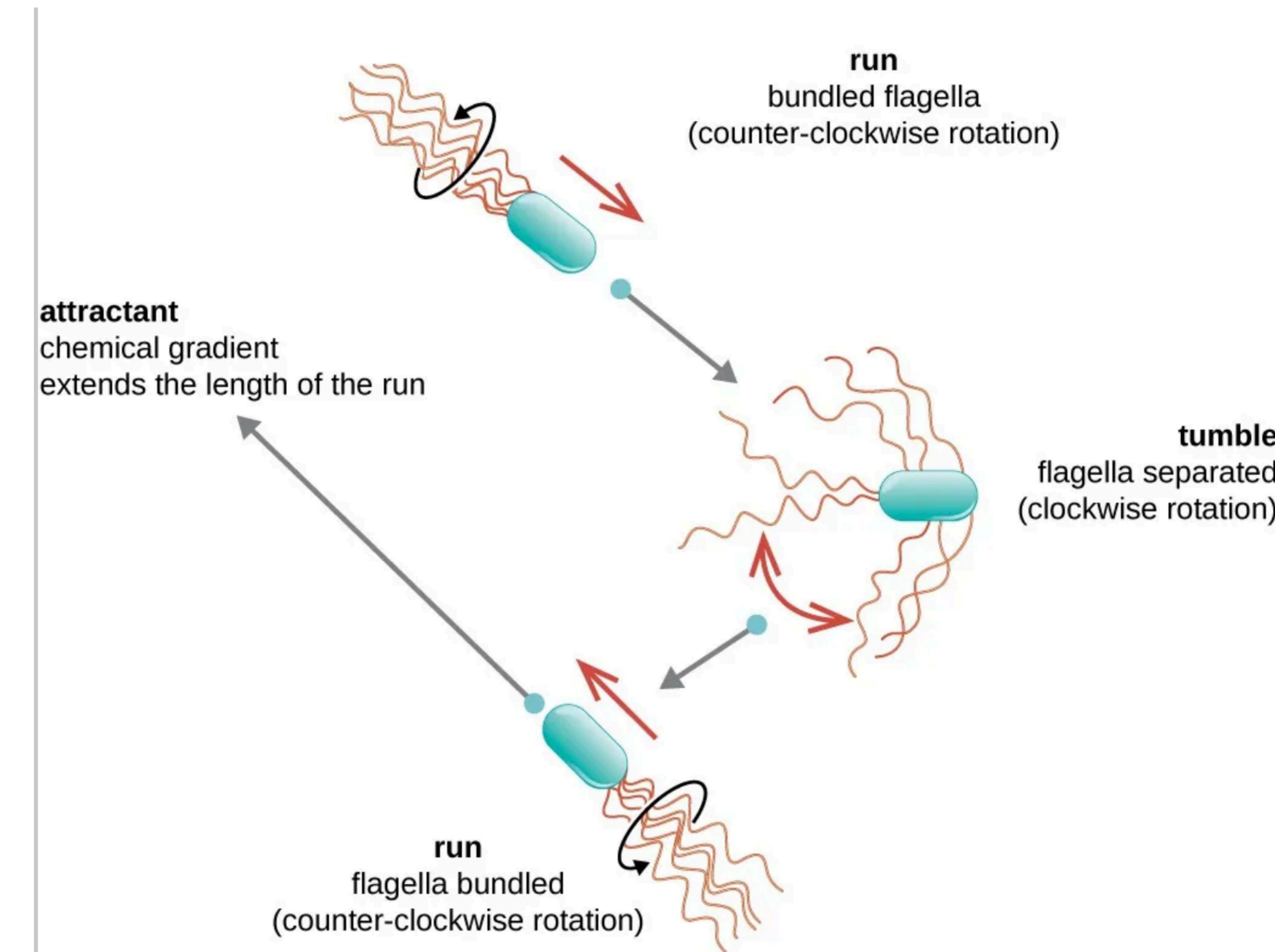


Fraction of proteins bound at equilibrium = minimum of free energy

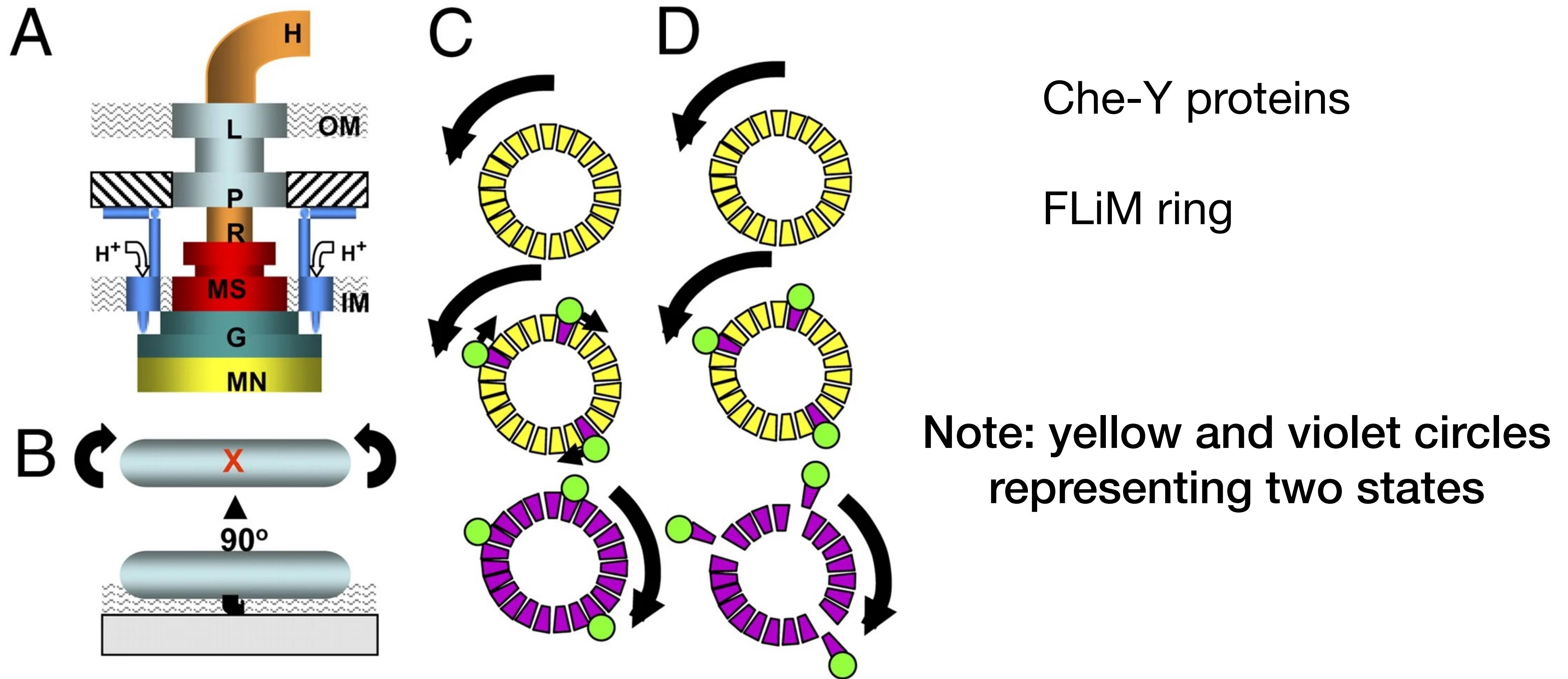


$$\text{Free Energy} = E - TS = -Nc\epsilon + Nk_B T[c \ln c + (1 - c)\ln(1 - c)]$$

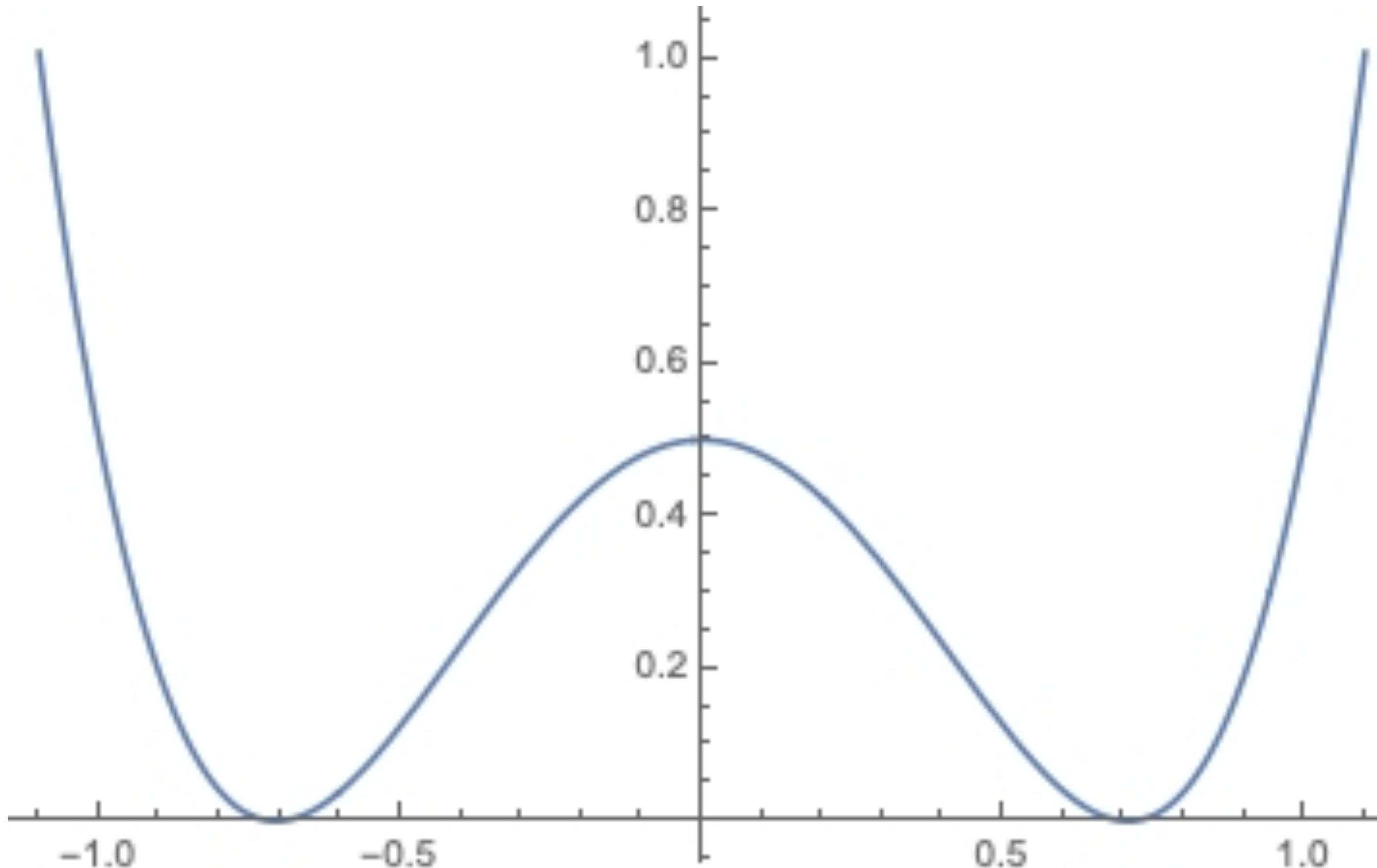
Bacteria rotate its flagella clockwise and counter-clockwise to change direction



To achieve rotation, a set of proteins bound switch between two states

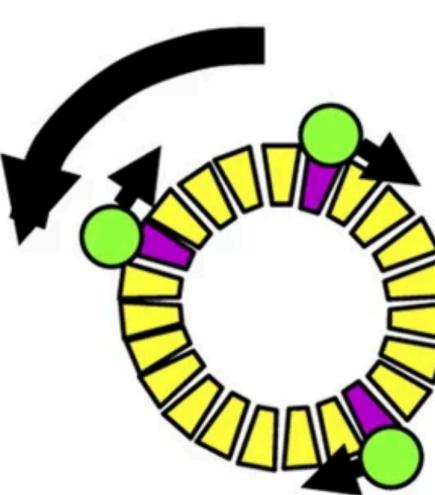
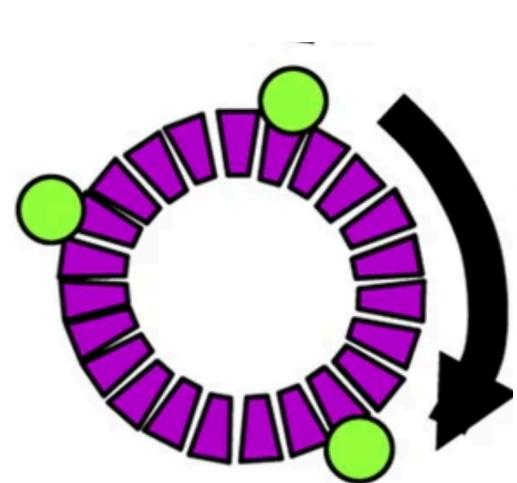


The underlying free energy



$$F(m) = Am^4 - Bm^2$$

A and B are two positive numbers representing property of protein interactions



Just like mass and gravitational interactions form real hills and valleys we see, particles and chemical interactions can form a landscape with hills and valleys!

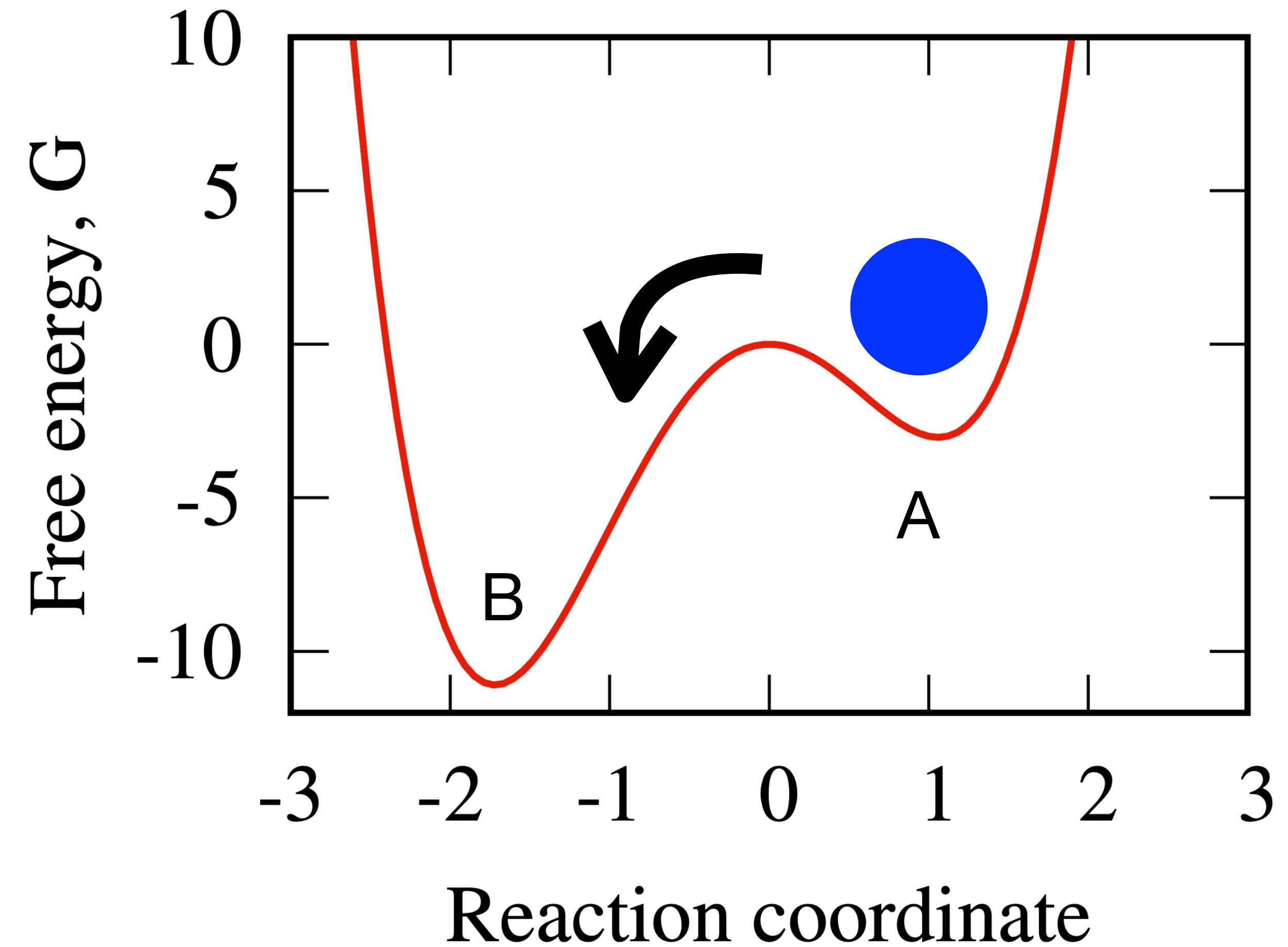
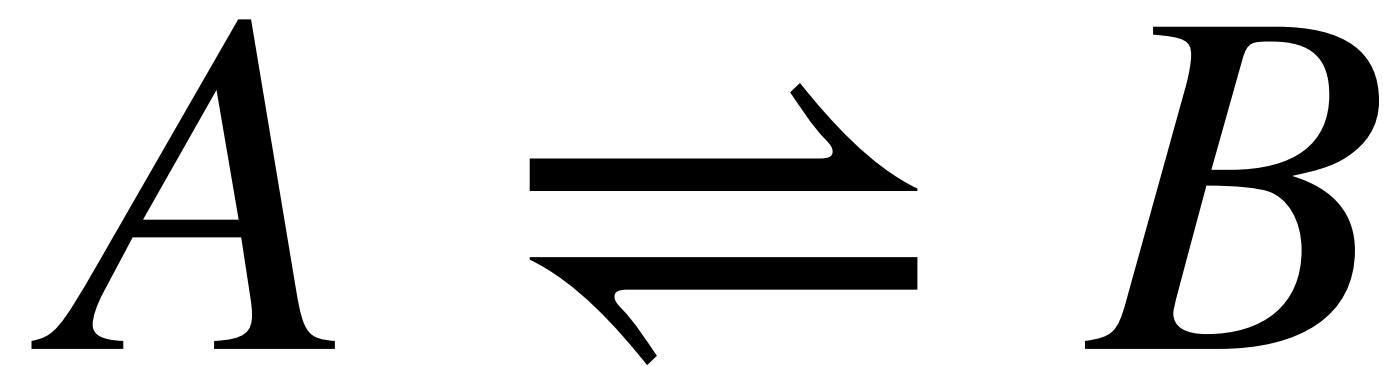
Biomolecules “see” this landscape — hills and valleys — as they move/react

**Free energy landscape is very much like
the potential energy landscape, except
that both energy and entropy play a role!**

Biomolecules “see” this landscape – hills and valleys – as they move/react

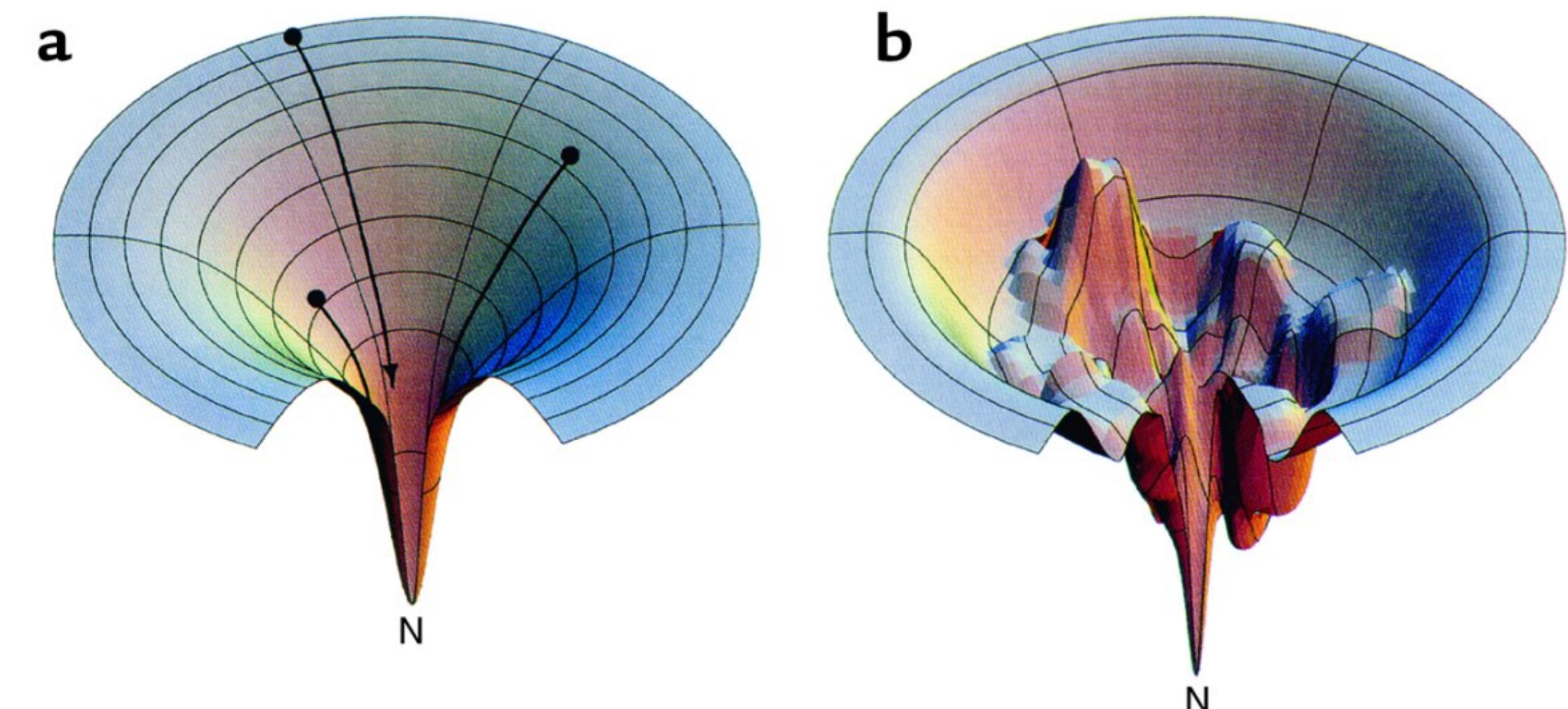
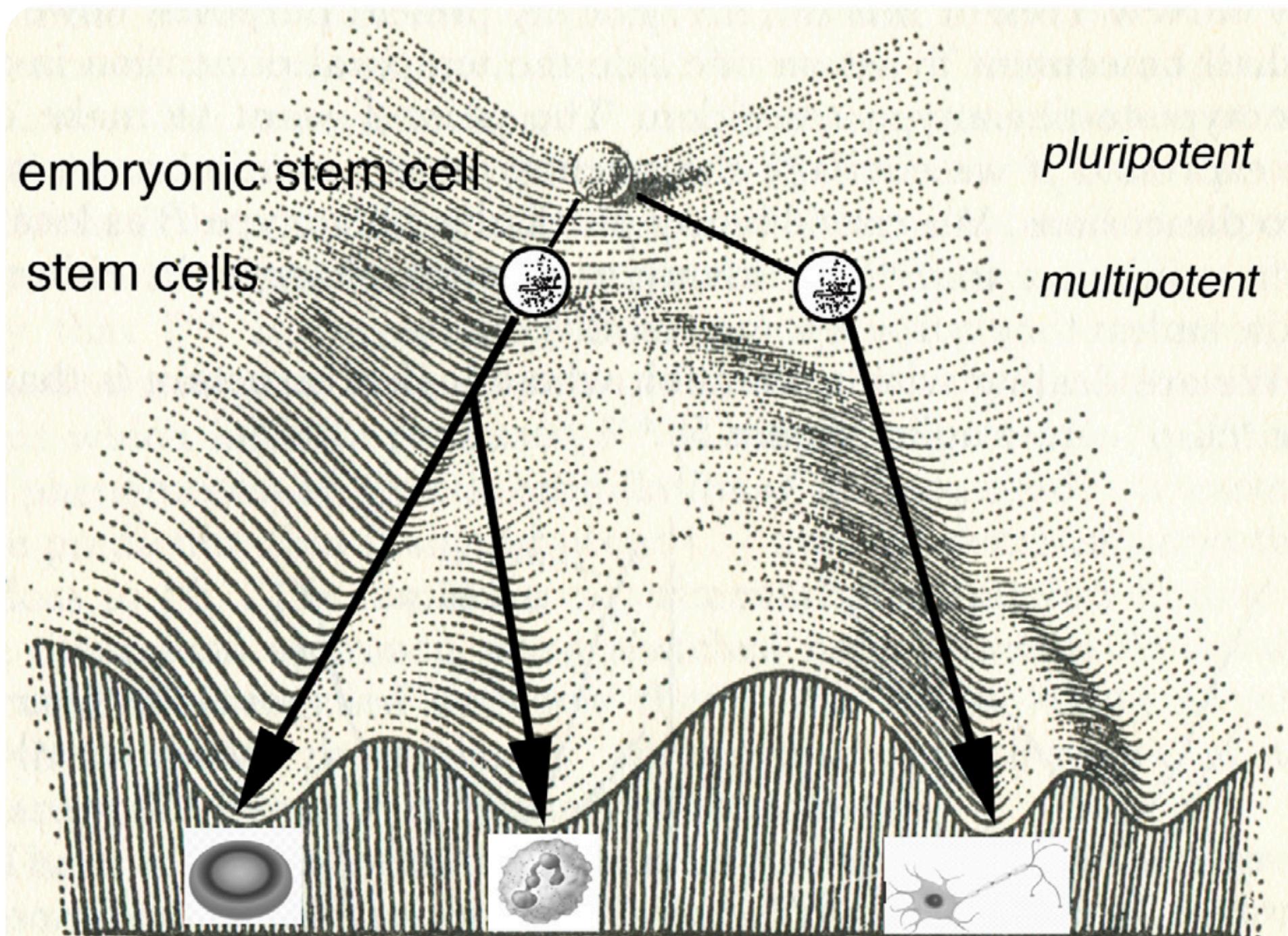
All biological events (chemical reactions) are happening in such free energy landscapes

Chemical reaction can be thought of as moving in free energy landscape

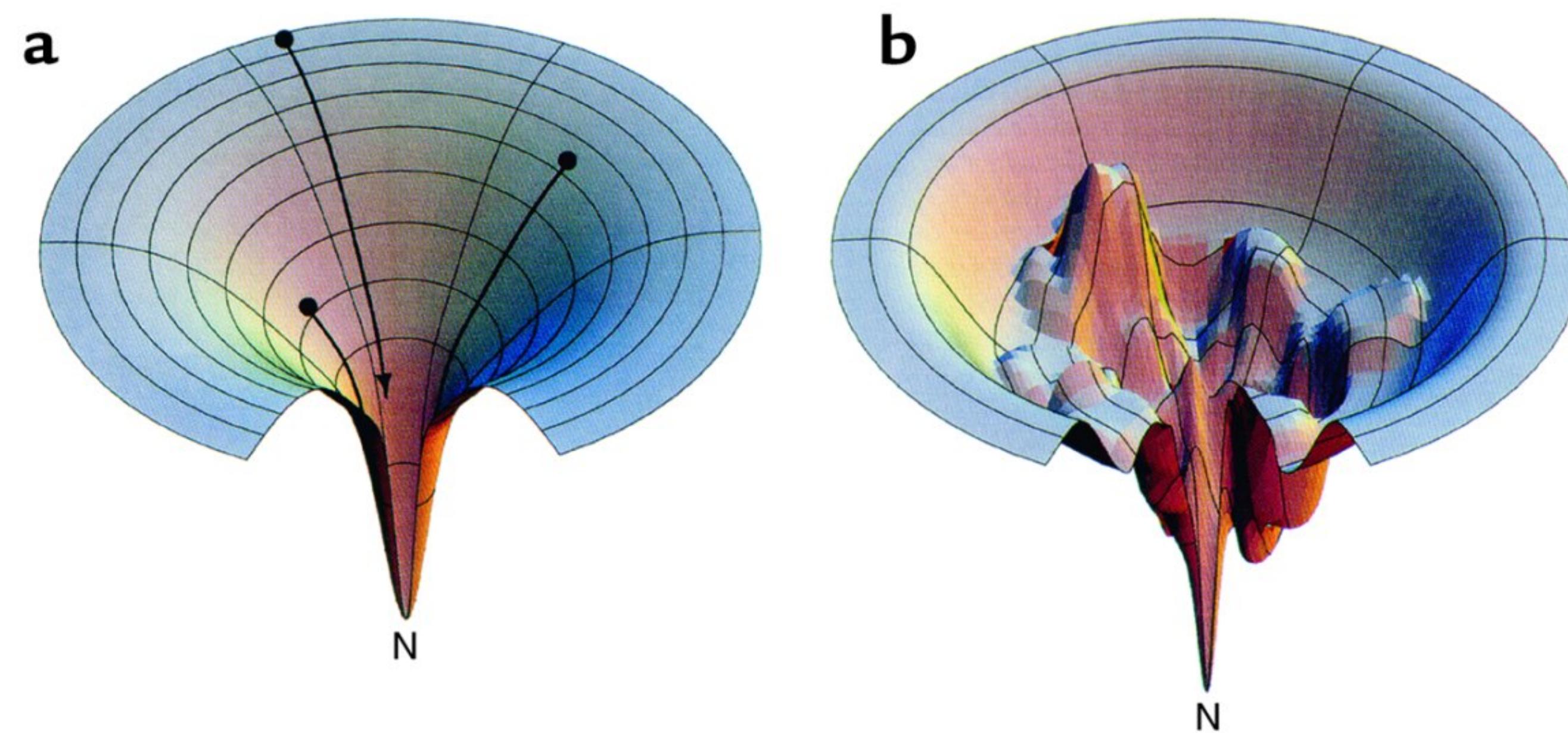


Molecular biology events as movements in a free energy landscape of hills and valleys

The height of these hills and depth of the valleys decide rate, stability etc in biology

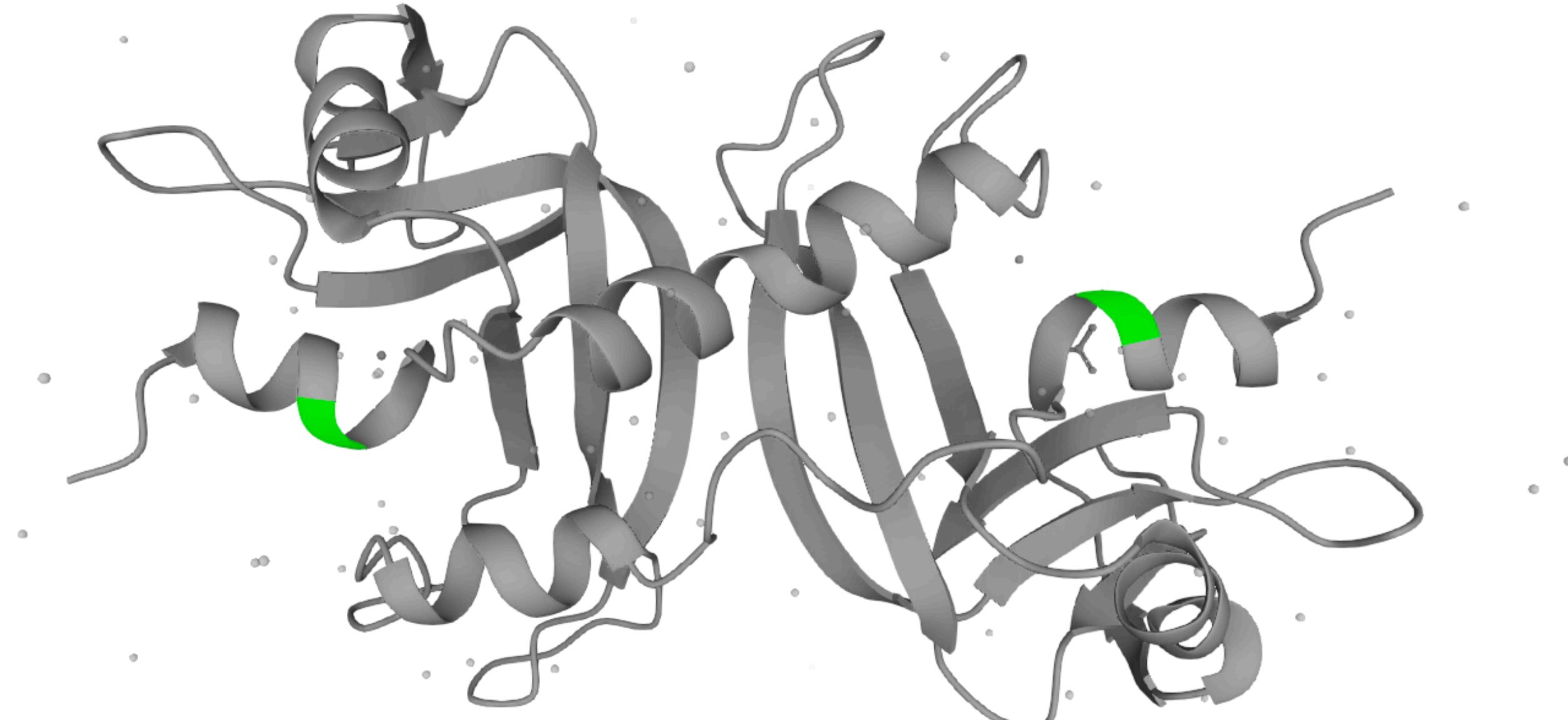


Protein folding

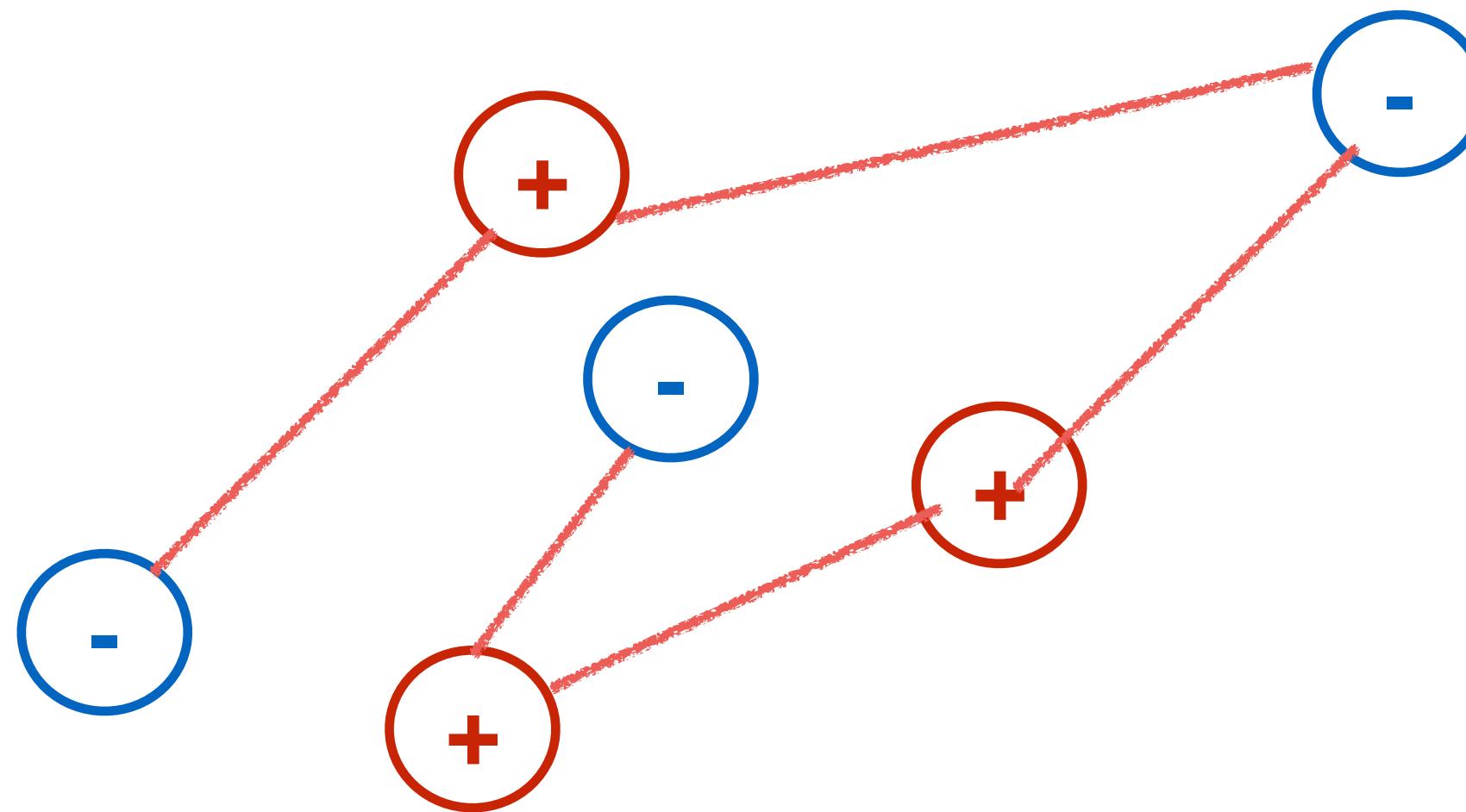


One-dimensional Amino Acid sequence decides 3D configuration of proteins

MALKSLVLLSLLVLVLLRVQPSLGKETAAAKFERQHMDSSSTAASSS
NYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAVCSQKNVACKNGQ
TNCYQSYSTMSITDCRETGSSKYPNCAYKTTQANKHIIVACEGNPYVP
VHF DASV



Interaction between amino acids
lead to a minimum free energy
state



But this leads to a paradox!

(Levinthal's paradox)

(Cyrus Levinthal 1969)

Levinthal's paradox

For a protein, infinitely large number of different configurations are possible!

How can a protein fold to the “right” (native) configuration within a finite time?

Levinthal's paradox

There are infinitely large number of different configurations possible!

If it explores all configurations to know the minimum free energy, it should take infinitely long time to find the “right” shape!

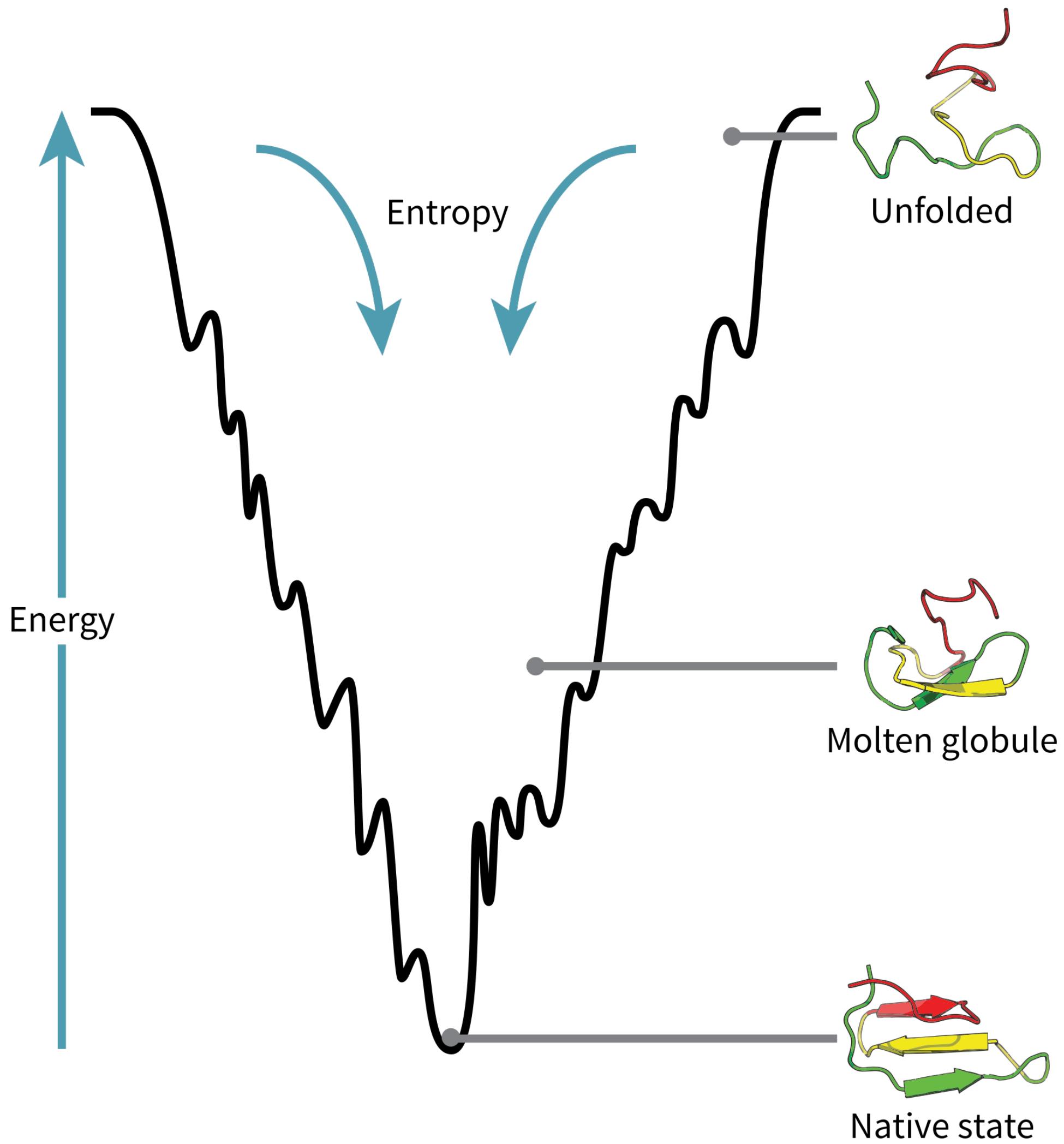
Levinthal's paradox

There are infinitely large number of different configurations possible!

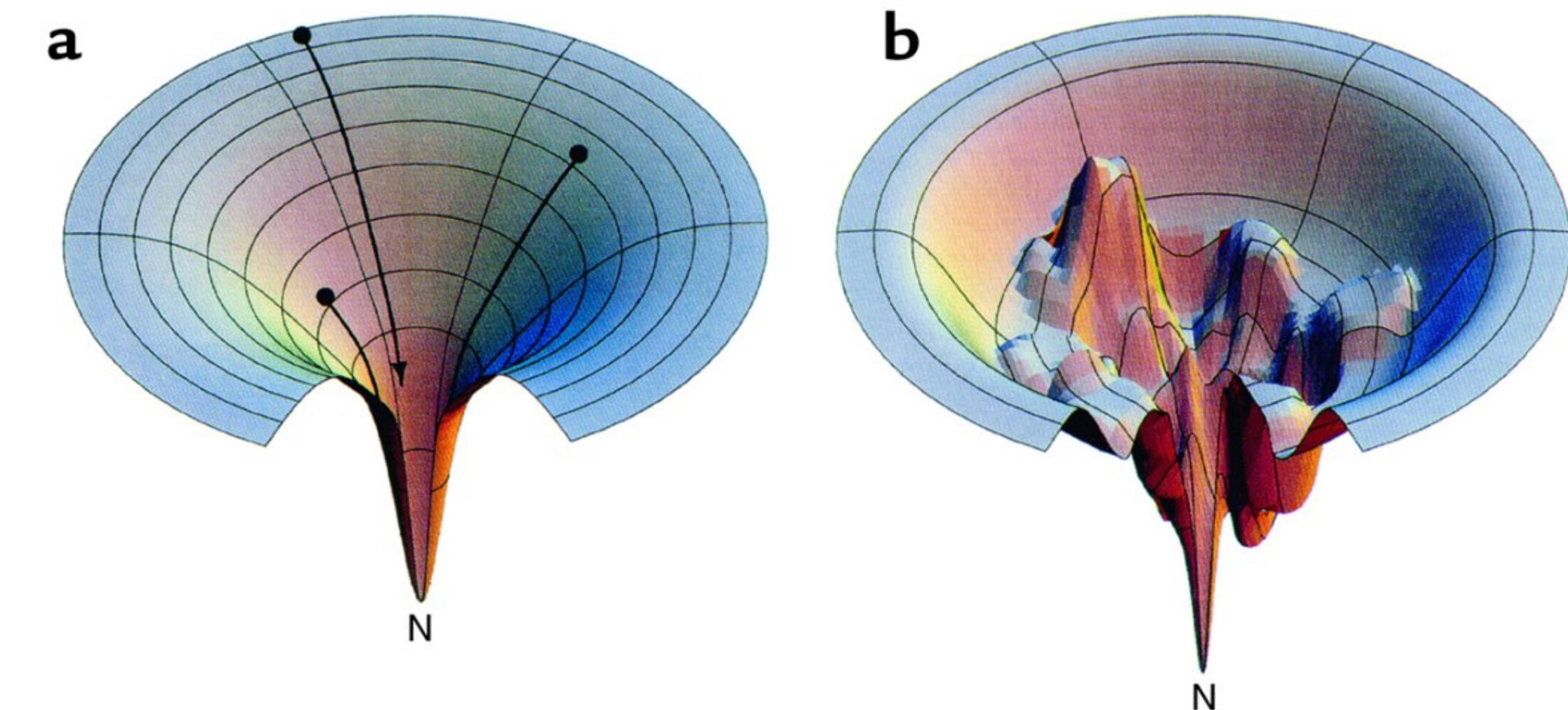
If it explores all configurations to know the minimum free energy, it should take infinitely long time to find the “right” shape!

BUT, they fold in a fraction of a second!

How could proteins fold quickly?



**Protein folding happens in a
“funnel-like” free energy
landscape**



Optional challenge: doable for an IITB UG student*

**Can you compute free energy of an
“ideal” protein chain?**

You will immediately see signatures of the funnel-like landscape!

Ideal chain: a polymer chain with no interaction; no energy (like the ideal gas)!

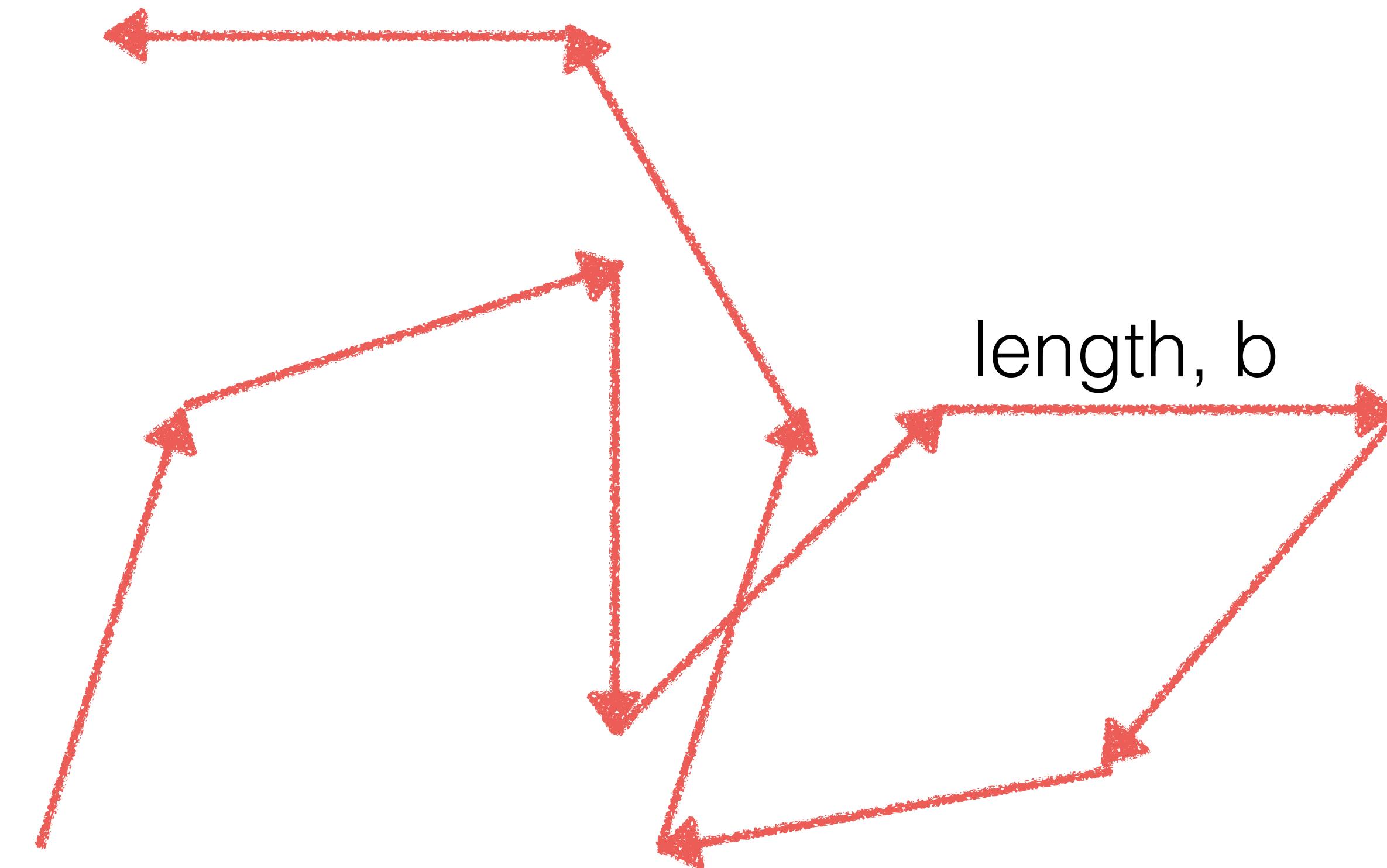
Do you know programming? It's a small code!
One can also derive with a pen and paper!

Ideal protein chain : Simplest protein chain you can imagine

N randomly oriented “monomers” with no interaction; no energy

They can take
any orientation;

They can
overlap!



Analogy: ideal gas. N particles with no interactions; they can take any position;
they can overlap

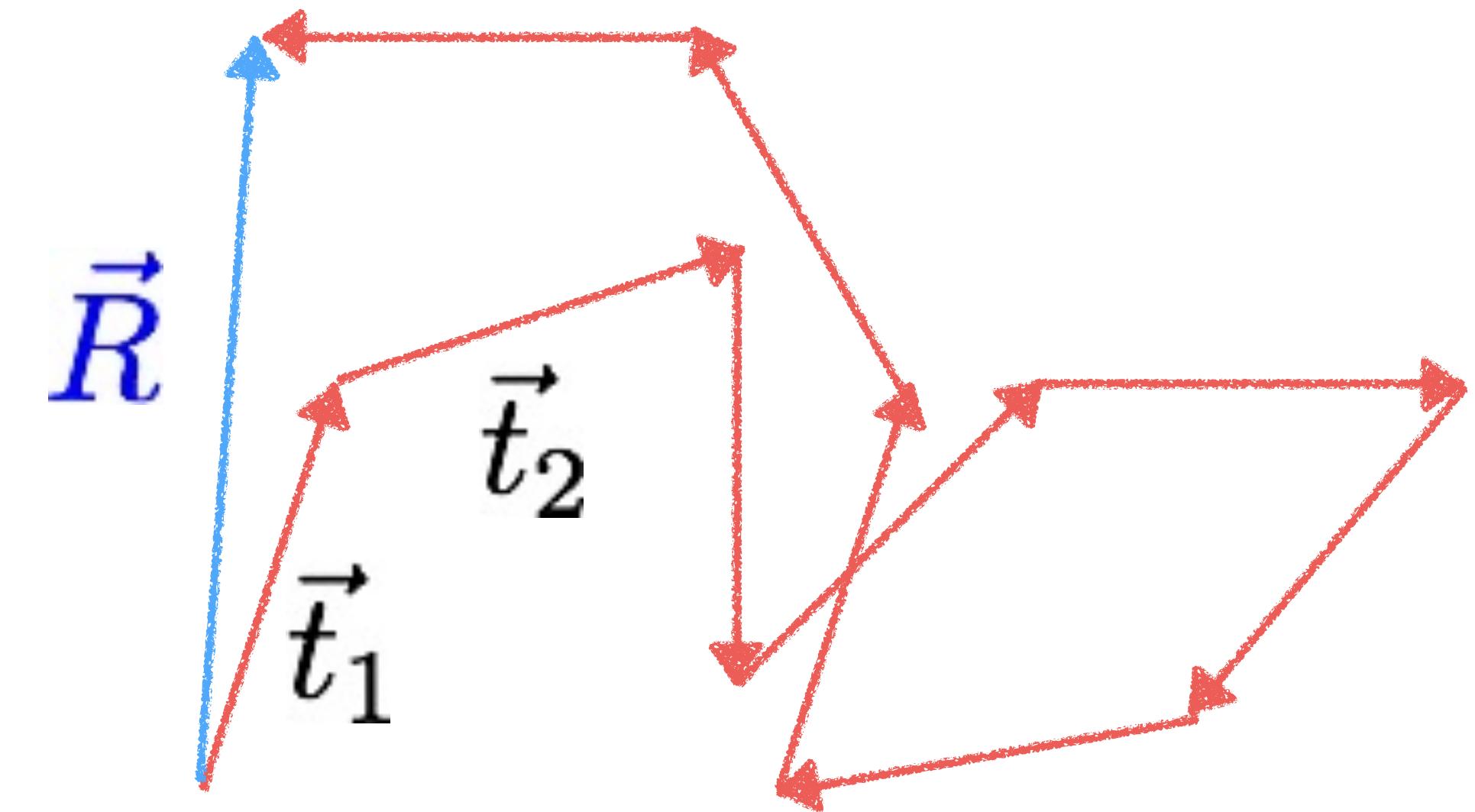
Free energy of an “ideal” protein chain!

Can you enumerate all possible configurations? Equal probability for any configuration!

Plot the distribution of end-to-end vector \vec{R}

Can you write a code?

Or derive mathematically?



(If there is time in this course, I will explain this!)

Can we make ANY quantitative
statement about such structures?

**How many of you want to predict
stock market and make money?**

Stock market is probabilistic/stochastic

If you cannot even predict stochastic (probabilistic) behaviour of an “ideal protein” chain, you have no hope of predicting stock market!

This is the simplest stochastic system you can imagine!

Article

Highly accurate protein structure prediction with AlphaFold

<https://doi.org/10.1038/s41586-021-03819-2>

Received: 11 May 2021

Accepted: 12 July 2021

Published online: 15 July 2021

Open access

 Check for updates

John Jumper^{1,4}✉, Richard Evans^{1,4}, Alexander Pritzel^{1,4}, Tim Green^{1,4}, Michael Figurnov^{1,4}, Olaf Ronneberger^{1,4}, Kathryn Tunyasuvunakool^{1,4}, Russ Bates^{1,4}, Augustin Žídek^{1,4}, Anna Potapenko^{1,4}, Alex Bridgland^{1,4}, Clemens Meyer^{1,4}, Simon A. A. Kohl^{1,4}, Andrew J. Ballard^{1,4}, Andrew Cowie^{1,4}, Bernardino Romera-Paredes^{1,4}, Stanislav Nikolov^{1,4}, Rishub Jain^{1,4}, Jonas Adler¹, Trevor Back¹, Stig Petersen¹, David Reiman¹, Ellen Clancy¹, Michal Zielinski¹, Martin Steinegger^{2,3}, Michalina Pacholska¹, Tamas Berghammer¹, Sebastian Bodenstein¹, David Silver¹, Oriol Vinyals¹, Andrew W. Senior¹, Koray Kavukcuoglu¹, Pushmeet Kohli¹ & Demis Hassabis^{1,4}✉

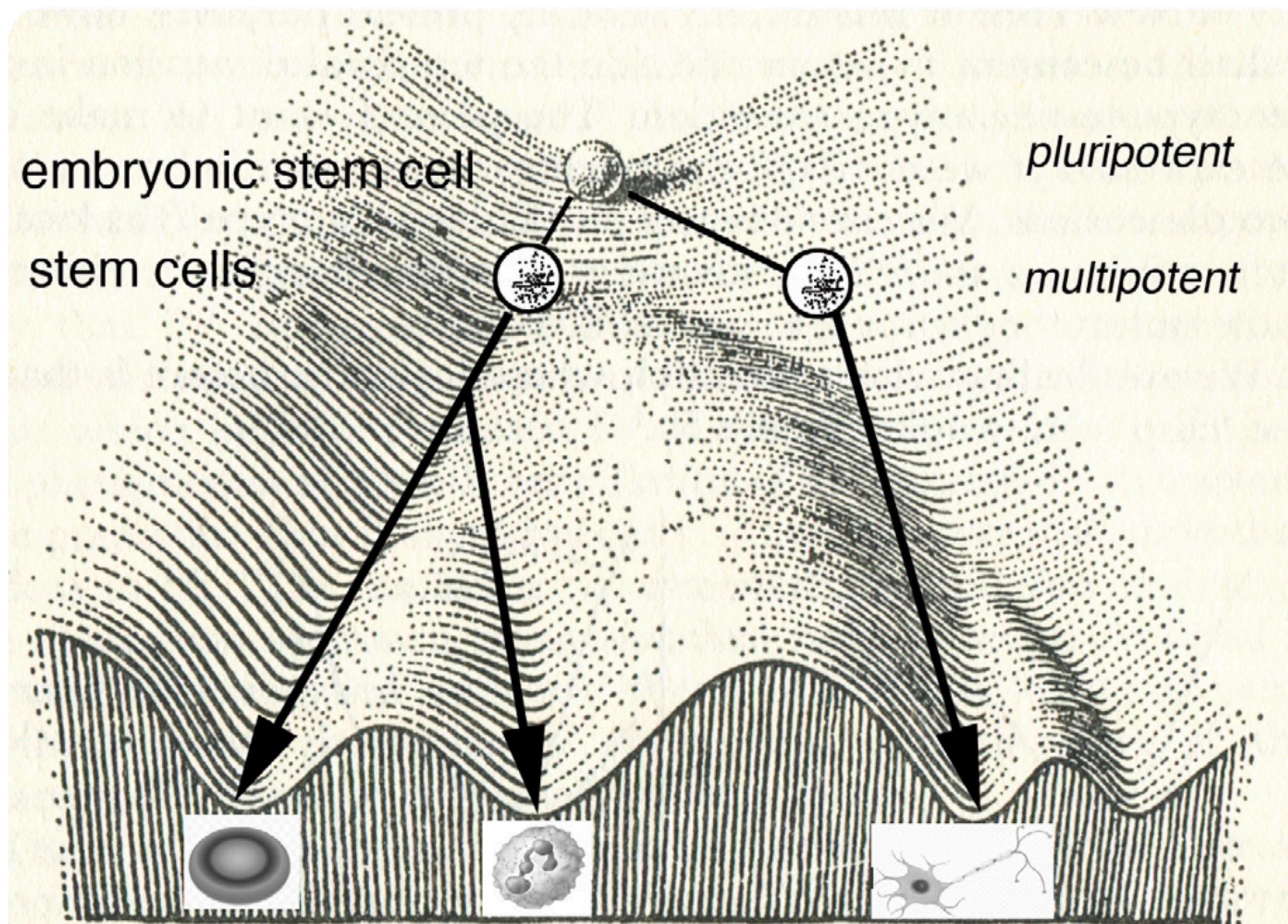
Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort^{1–4}, the structures of around 100,000 unique proteins have been determined⁵, but this represents a small fraction of the billions of known protein sequences^{6,7}. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid sequence—the structure prediction component of the ‘protein folding problem’⁸—has been an important open research problem for more than 50 years⁹. Despite recent progress^{10–14}, existing methods fall far short of atomic accuracy, especially when no

Jumper et al Nature 2021

¹DeepMind, London, UK. ²School of Biological Sciences, Seoul National University, Seoul, South Korea. ³Artificial Intelligence Institute, Seoul National University, Seoul, South Korea. ⁴These authors contributed equally: John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Demis Hassabis.

✉e-mail: jumper@deepmind.com; dhcontact@deepmind.com

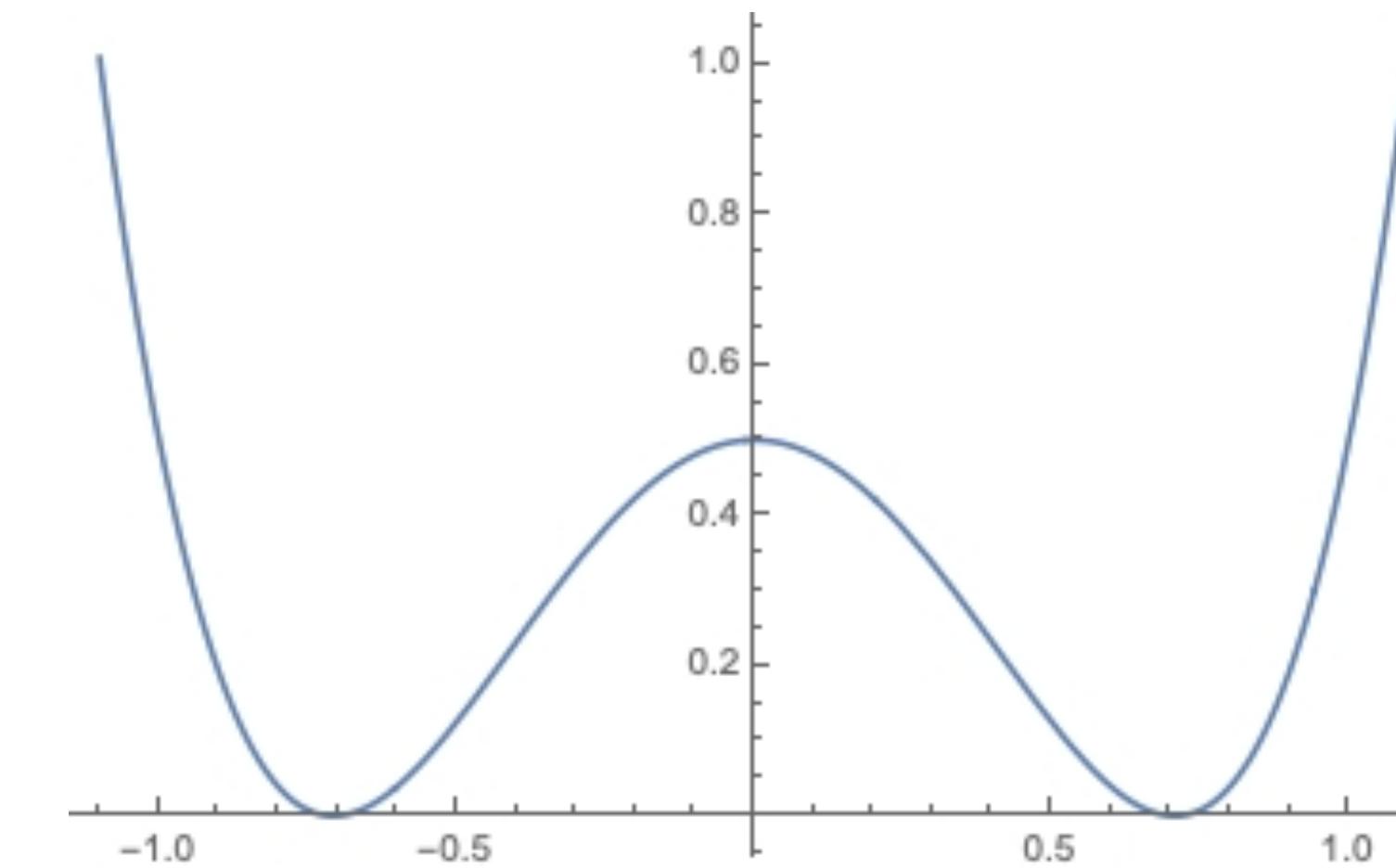
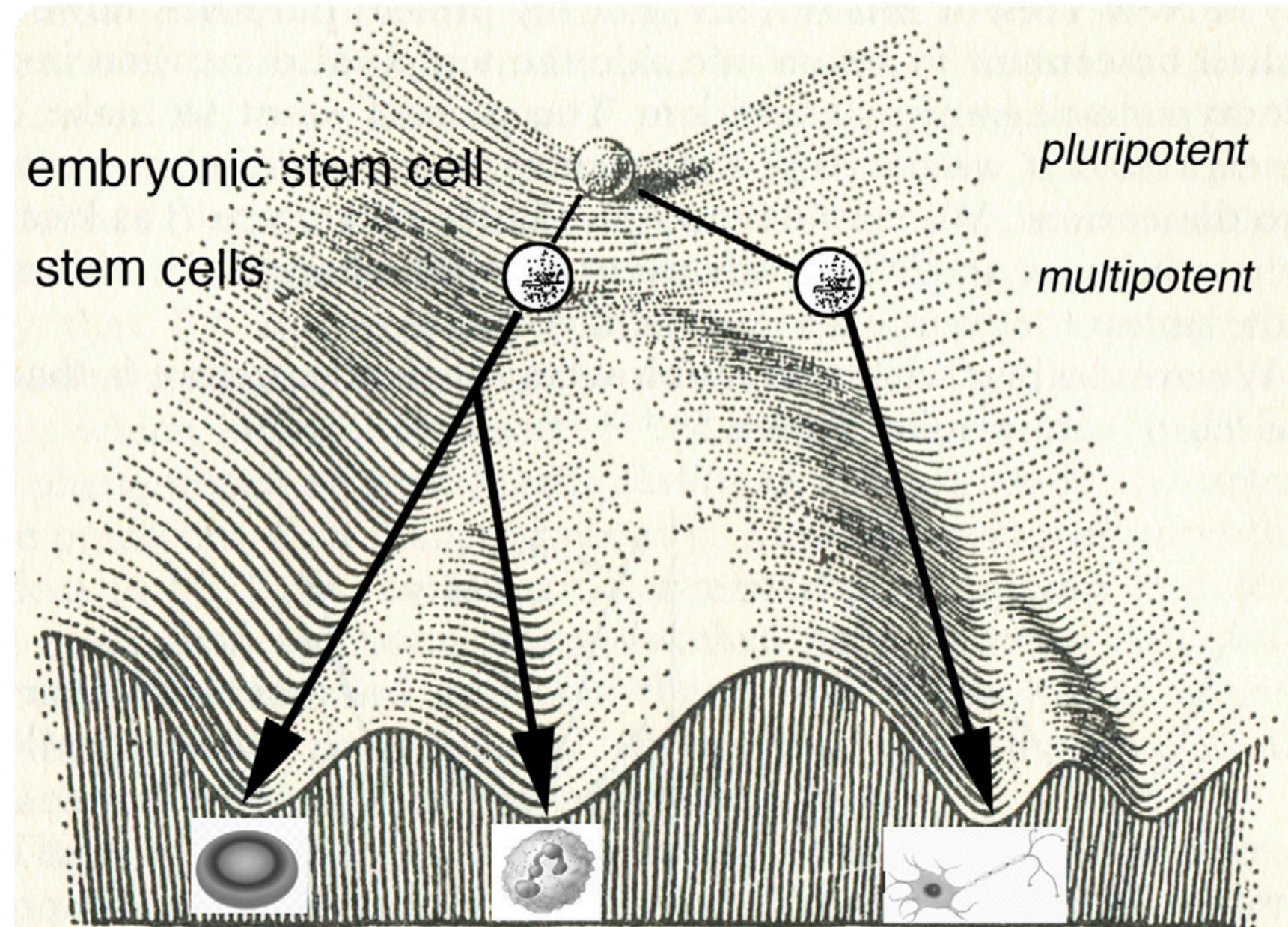
Differentiation of cells as movements in a landscape



Waddington's epigenetic landscape



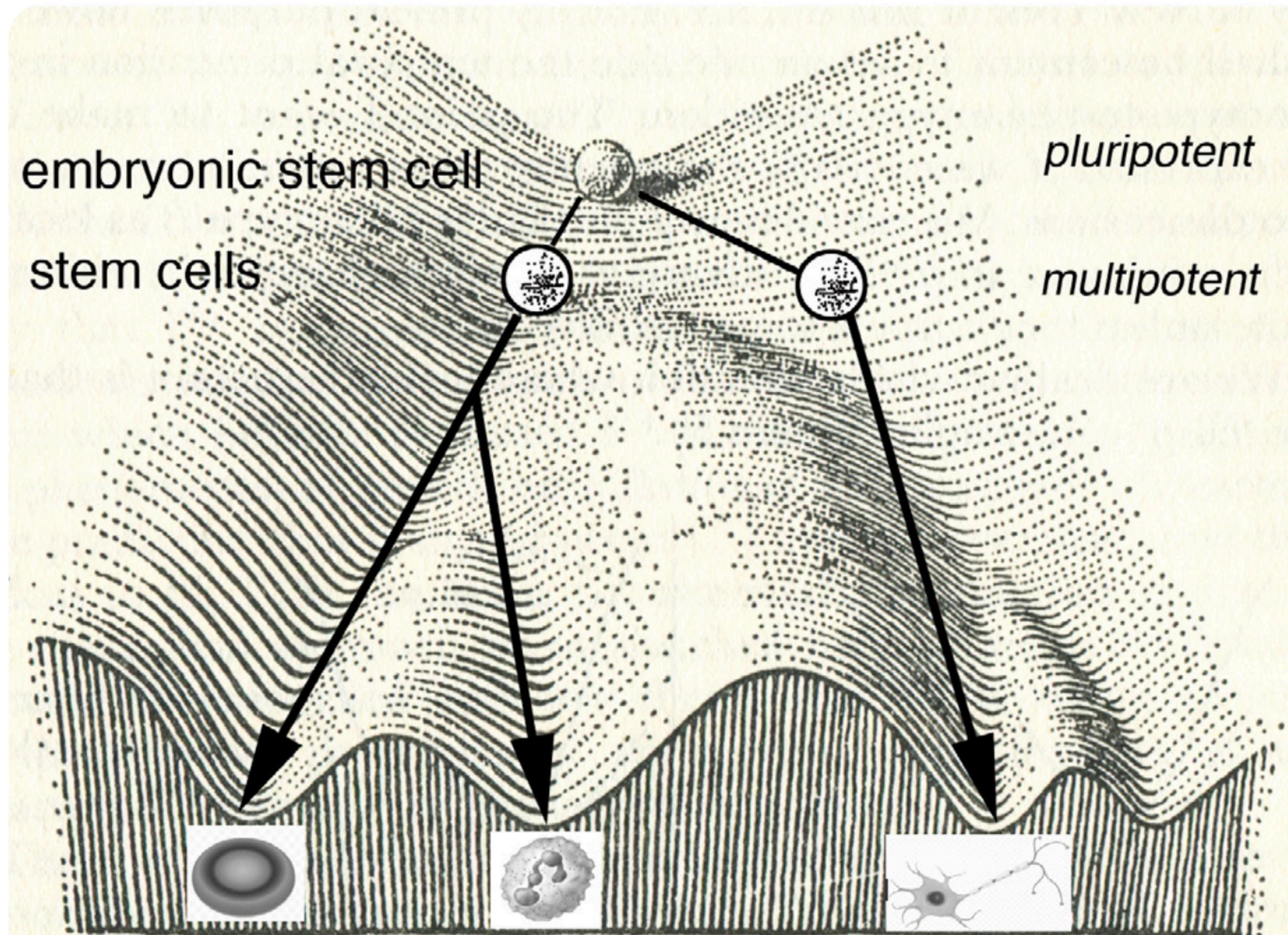
Imagine differentiation as a complicated “phase transition”



Free energy landscape for paramagnet-ferromagnet phase transition

Free energy landscape for water-ice phase transition etc are well studied

Differentiation of cells as movements in a landscape



Waddington's epigenetic landscape

Finding out precise shape of epigenetic landscape is an open unsolved research question!