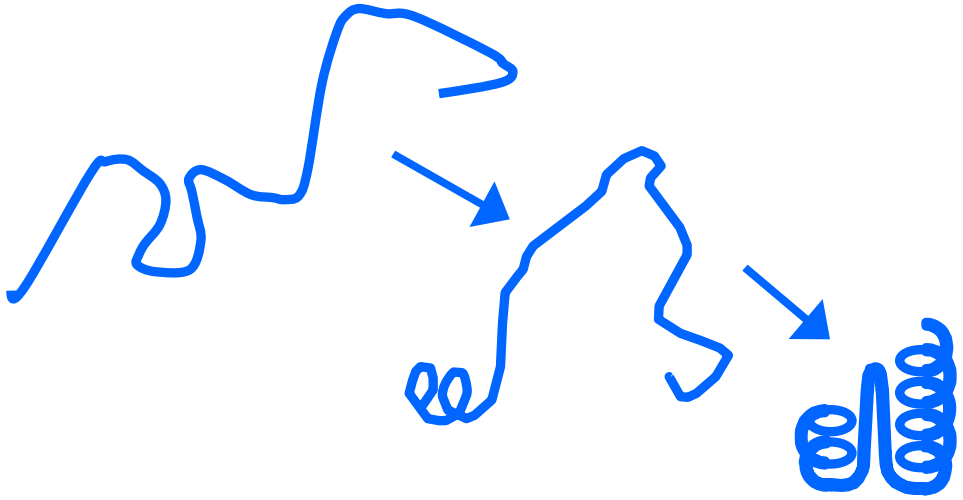
Protein Folding

How does a protein do this ?



If we could understand folding we could

- predict structure
- design proteins that fold better (more stable)
- identify essential residues for folding (not suitable for mutagenesis)

Folding should be easy

From biochemistry Übungen (protein folding – easy)

- take lysozyme / ribonuclease...
- put in 8M urea (unfolds)
- remove urea (refolds)

Conclusion ?

- the protein sequence is all you need to fold a protein
- is this true ? Not always

Alternative (logical reasoning)

- protein folding should be impossible…

Folding should be hard / Levinthal's paradox

Each amino acid has 2 or 3 or n conformations

- for a protein of m residues, it should visit n^m

If it spends 10^{-15} s at each conformation ?

- time to find one conformation for n=3 and 100 residues
  - $3^{100} \times 10^{-15}$ s $\approx 1.6 \times 10^{25}$ years

Consequence

- proteins cannot be exploring space randomly
- historic idea of "folding pathway"

Questions

- Kinetic versus thermodynamic
- What order do events happen in ? (collapse vs secondary struct)
- Is unfolding the same as folding ?
- Is folding in a test tube the same as nature ?

Kinetic versus thermodynamic minima

Anfinsen story..

- proteins can be unfolded and refolded alone
- all the information is in the sequence
- native conformations are the (free) energy minimum
- thermodynamic belief

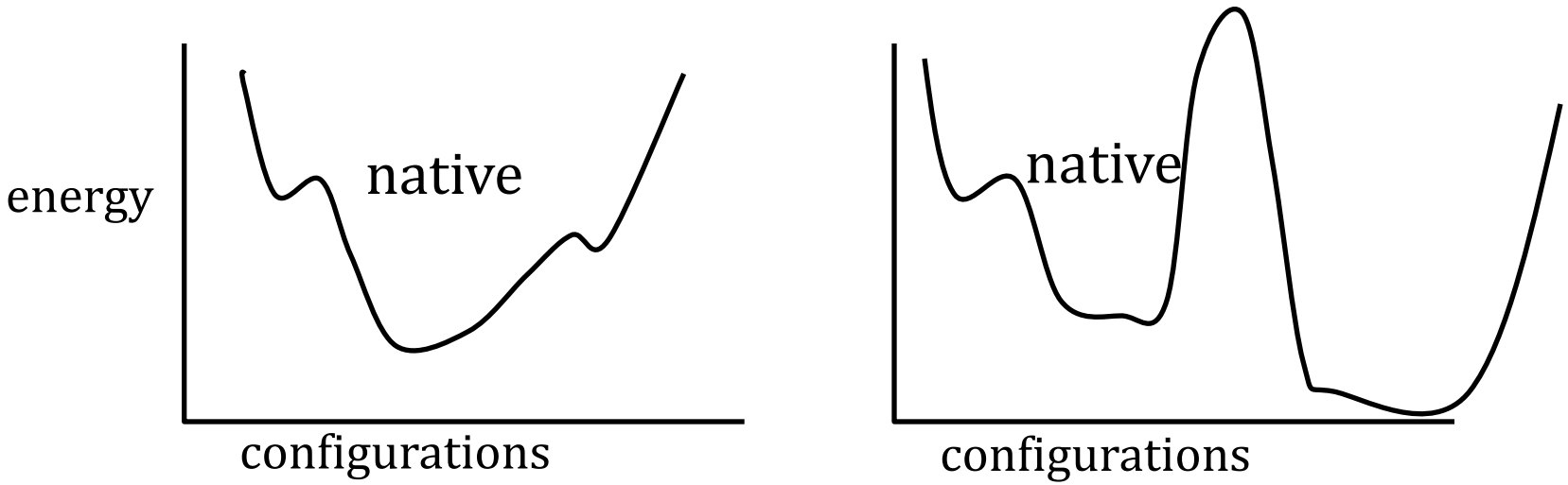
More modern

- many many proteins cannot be refolded in the lab
- consequence .. maybe they need something else
- maybe they are not always in free energy minimum

Kinetically trapped proteins..

Kinetic versus thermodynamic minima

- If proteins fold spontaneously and remain folded, they are thermodynamically determined
- If you leave a protein long enough and it unfolds, it was not in an energy minimum



thermodynamic

kinetically trapped

Kinetic versus thermodynamic minima

Thermodynamic

- protein structure prediction
  - just a matter of modelling the real world

Kinetically trapped

- we cannot predict structure from sequence just by energies

Kinetic versus thermodynamic minima

Can we see which is the case ?

- leave a protein for 10 minutes
  - see if it finds another state
- leave it for 10 years ?
- depends on barriers

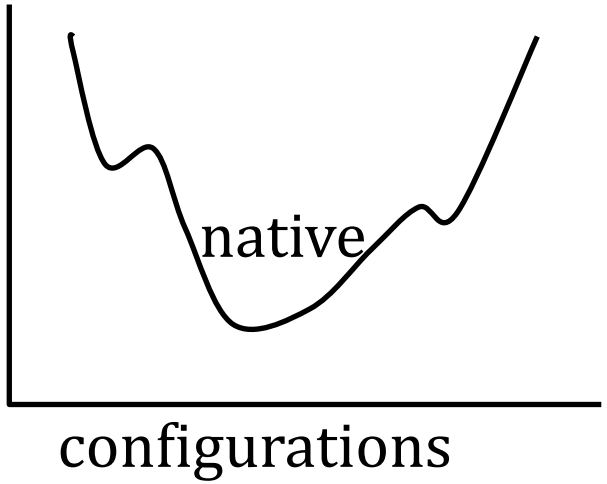
Empirically

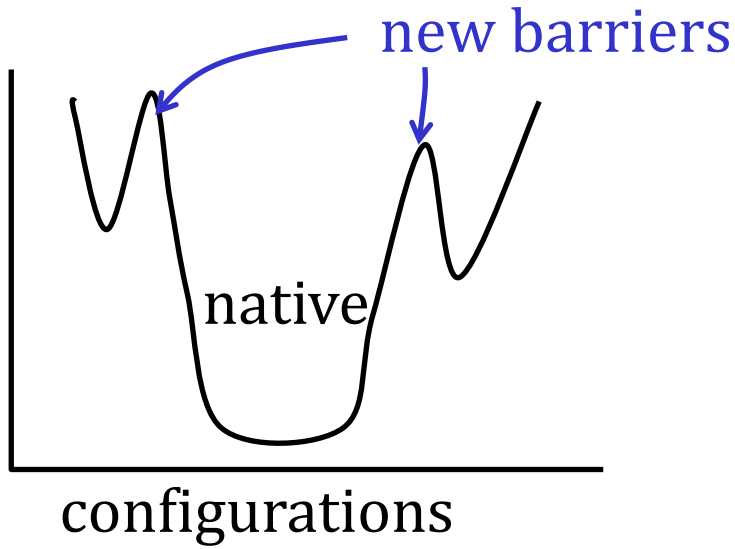
- some evidence of kinetic trapping
- some proteins do have other states
  - β-fibrils, Alzheimers, mad cow disease

Evolution / design consequences

Imagine I can predict structure and stability

- I design a better / more stable protein





- my new protein may be more stable
- it may never fold
- evolutionary implications
  - protein sequences may evolve for folding (+ structure and function)

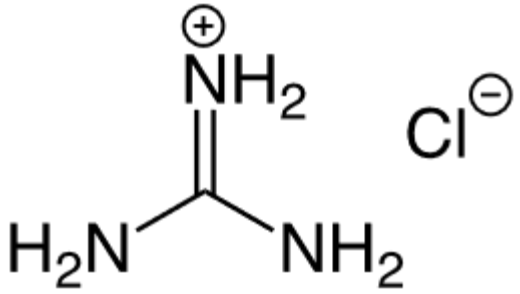
Experiments

Timescales

- maybe 10-6 s for folding
- maybe orders of magnitude slower (sometimes faster)

Experimental approaches

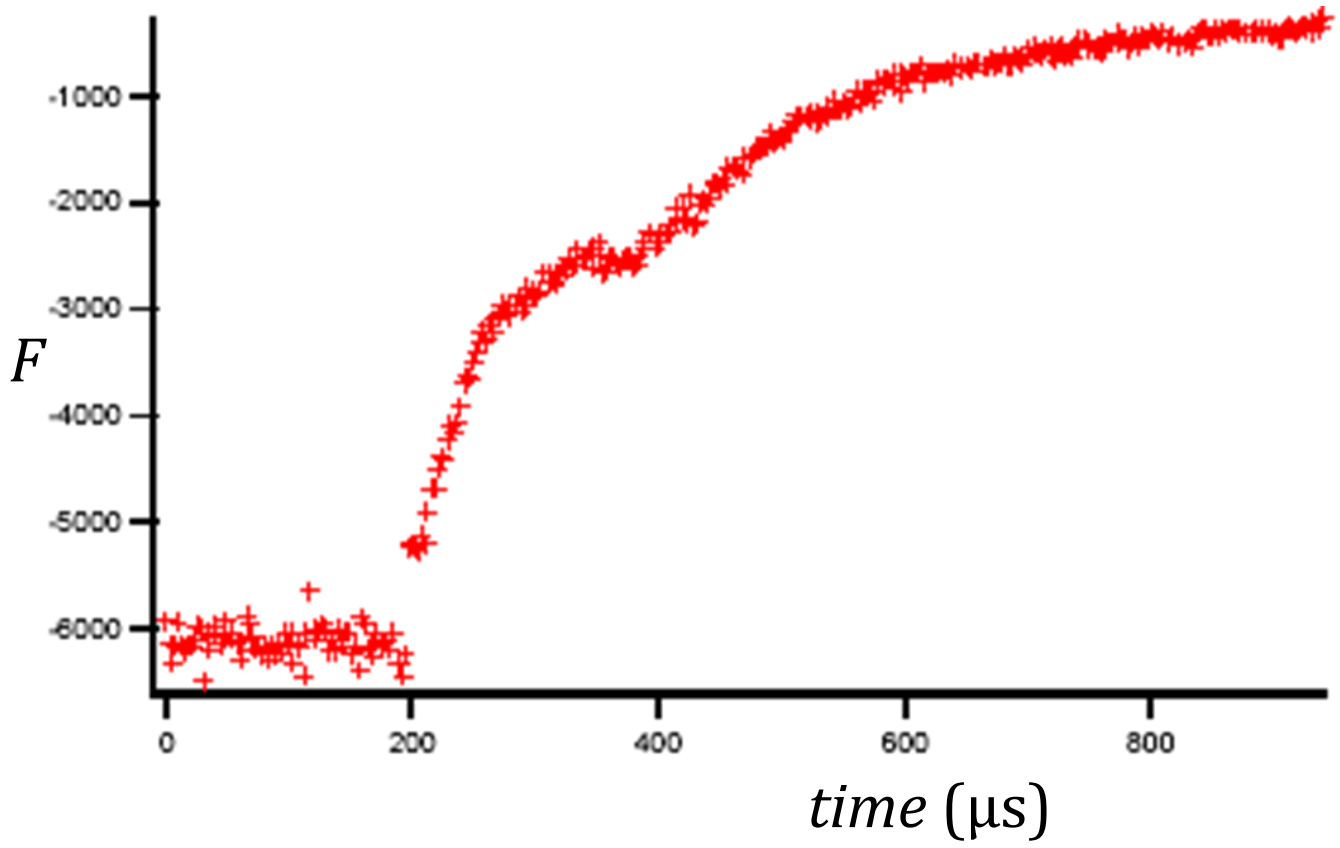
- force protein to unfold
  - chemistry (guanidinium HCL, 8M urea)
  - temperature (heat, cold)
- change conditions and watch
  - try to measure very fast
  - try to change timescale
  - try to measure unfolding



Experiments are difficult

Very difficult to measure on the µs / ms time scale

- temperature jump
- stop - flow
- fluorescence
- NMR
- circular dichroism (CD)



No experiment gives good
structural detail on this time scale

picture from Dimitriadis, G, Drysdale, A, Radford, S and Smith, A, "Investigating early events in protein folding by laser induced temperature jump"

How good are experiments ?

Technical difficulties – obvious

How relevant are experiments ?

- Imagine the perfect experiment
  - you unfold a protein (heat, salt, ..)
  - go back to native conditions and watch

  - does this tell you about protein folding ?

Nature versus experimental folding

Nature

- protein is synthesised on ribosome
- 150 mM salt, pH relatively neutral, 300 K

Experiment

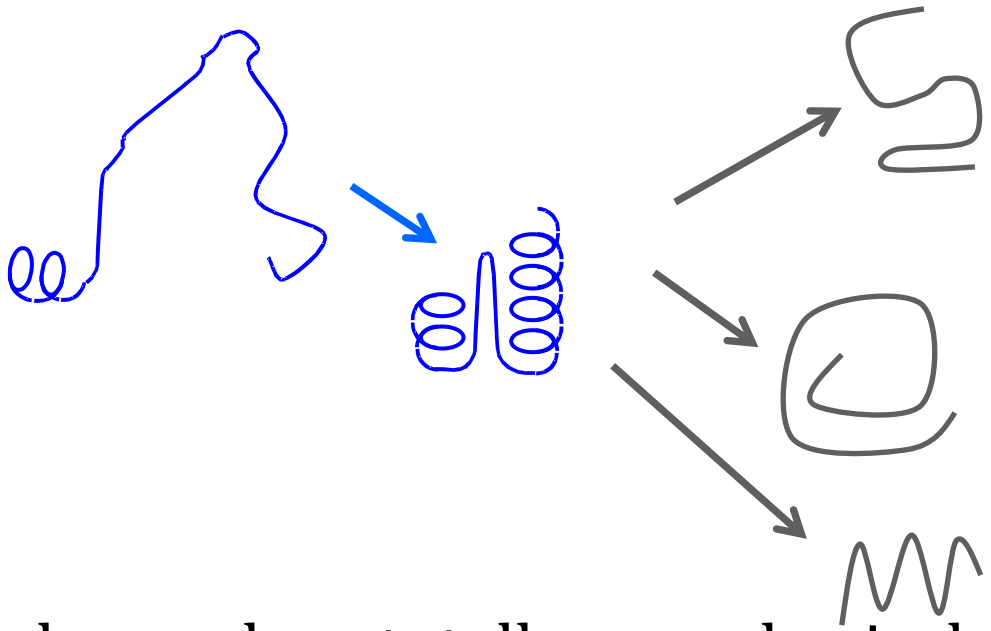
- high salt, temperature, ..

Difference

- under exotic conditions, protein visits unusual conformations

Simulations

- Can one simulate folding ? Not really
- Unfolding ?
  - start with native protein and heat it until it unfolds
  - is this the unfolding pathway ? Is it the opposite of folding ?



nature

unfolding

- Going backwards may be a totally non-physical, irrelevant pathway

Forgetting atomistic detail

- What are questions we can ask ?
- What can we guess without any calculations ?
- Questions
  - is there such a thing as a folding pathway ?
  - how should we look at folding ?
    - secondary structure forms first and is rearranged
    - hydrophobic residues come together and then secondary structure forms ?
    - a few important contacts are formed, then structure forms

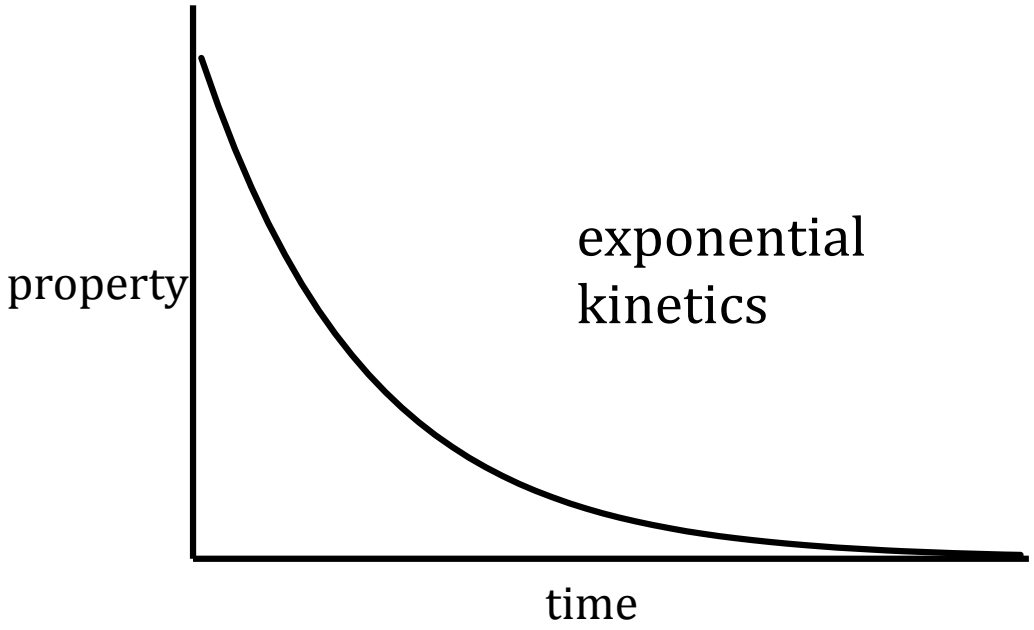
Kinetics in general

What have we seen so far ?

- most properties have something like exponential decay
- property $= ae^{-\alpha t}$
- rate of change proportional to quantity present

A model should explain at least this

- later by ensemble view..
  where are barriers ?



What do we know ?

Possibilities

Proteins form secondary structure first

- helices and sheets then arrange themselves

OR

Hydrophobic collapse

- hydrophobic residues find each other
- backbone rearranges and secondary structure is fixed

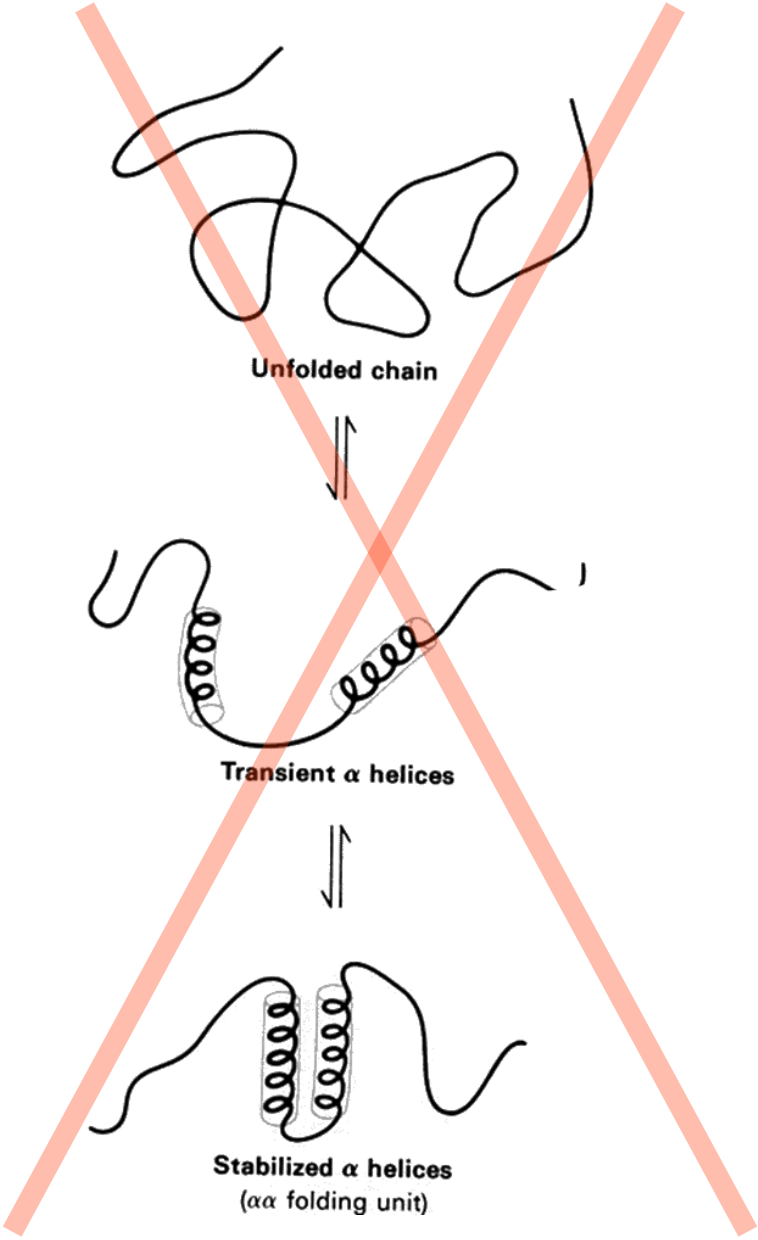
Sidechain or backbone driven

Old textbook

- local secondary structure forms, then reorganises
- secondary structure depends largely on backbone

Alternative

- sidechains are very important



Stryer, L. Biochemistry, 1981, WH Freeman, "Biochemistry", page 36

Sidechains might be important

- backbone view does not predict collapse of protein
  why would helices and sheets find each other ?
- α-helix and β-strand propensity is weak
  - isolated peptides are not stable
  - β-strands often depend on long-range H bonds
- helix / strand formation depends on environment / solvent and is not known in open structure
- fold is largely predictable / characterised by pattern of HP (sidechains)
- proteins are most sensitive to mutations in core
  (they are important for stability)

Types of pathway

From classical chemistry we would like a path

unfolded → A→ B→C→ folded (native)

Slightly more complicated – detours and sideproducts (X) allowed

Basic idea

- molecules may get sidetracked, but
  - every molecule sees A, B, C...
- where does it come from ?
  - Levinthal's answer to paradox
    - there must be a preferred pathway
- old view / microstate pathway approach

Consequence of simple pathway

Two state kinetics

- A → B or B → C might be part of transition barrier
- pathway with detours explains multi-state kinetics (if necessary)

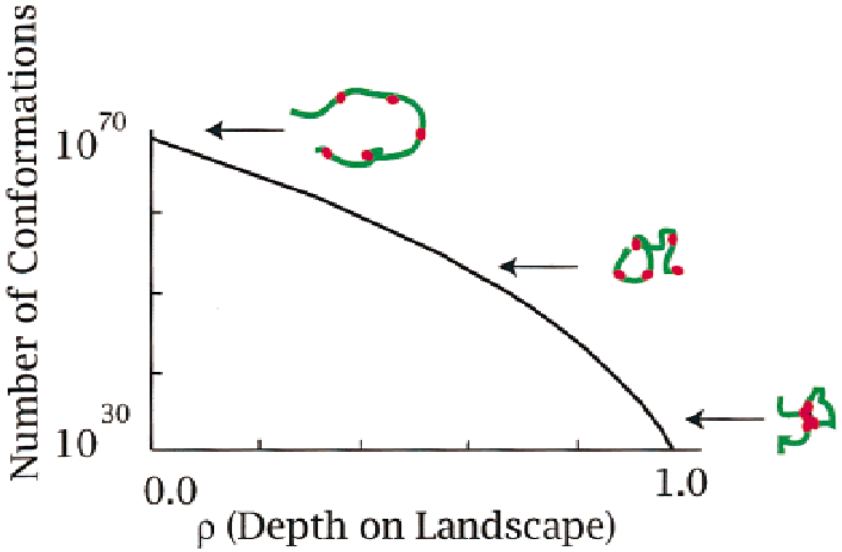
Does it sound intuitively reasonable ?

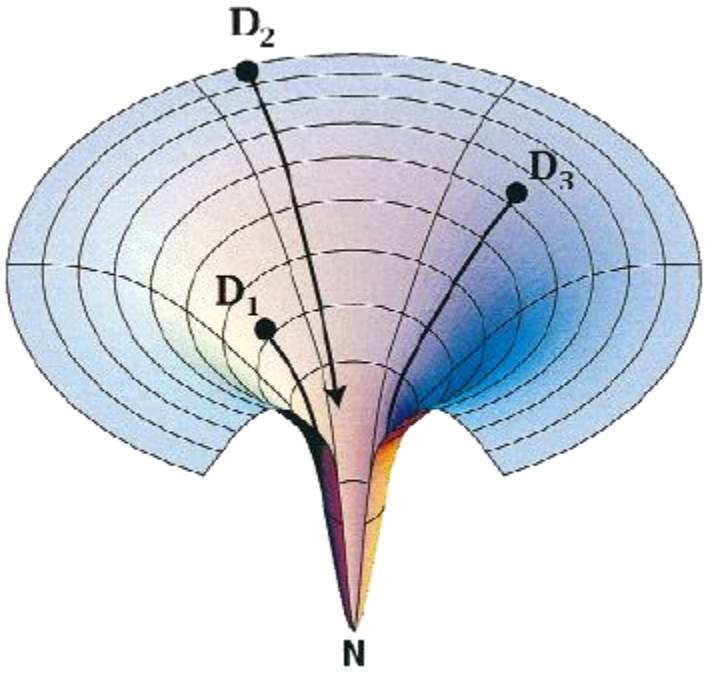
- what if a mutation perturbs A or B or C ?
  - whole pathway might break
  - maybe OK (this is why some mutants do not fold)

Do you need conventional pathway to explain barriers ?

Ensemble view

- conformation space is huge
- will a protein be able to find a neat path through it ?
- should we even look for paths ?
- consider a multitude of paths...
- is this merely a cute picture ?
- first implication...
  - entropy vs energy





Dill, K.A., Protein Sci., 8, 1166-1180, 1999
Polymer principles and protein folding

Ensemble view

As a protein folds

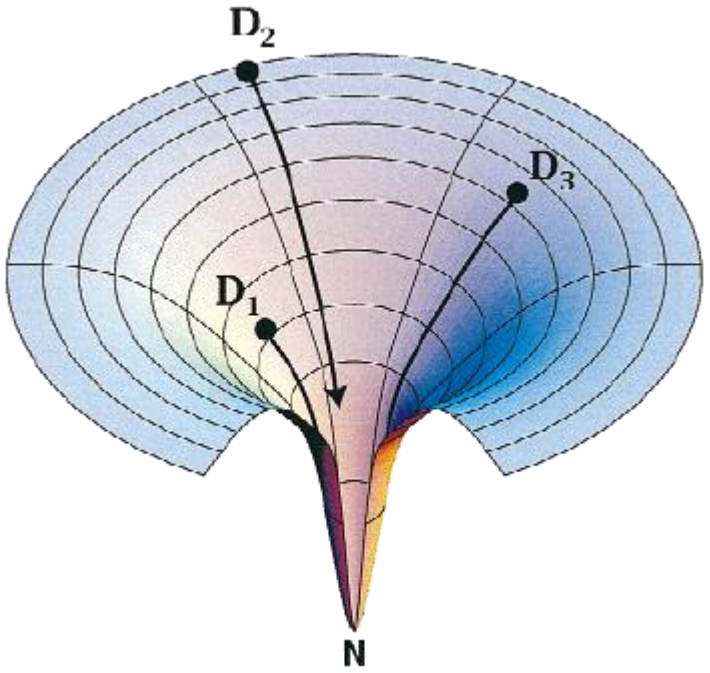
- potential energy goes down (happy)
- number of possible states goes down
- entropy goes down (unhappy)

Interpretation in chemical terms

- do we have a off-pathway intermediate ?

Do we have a reaction coordinate ?

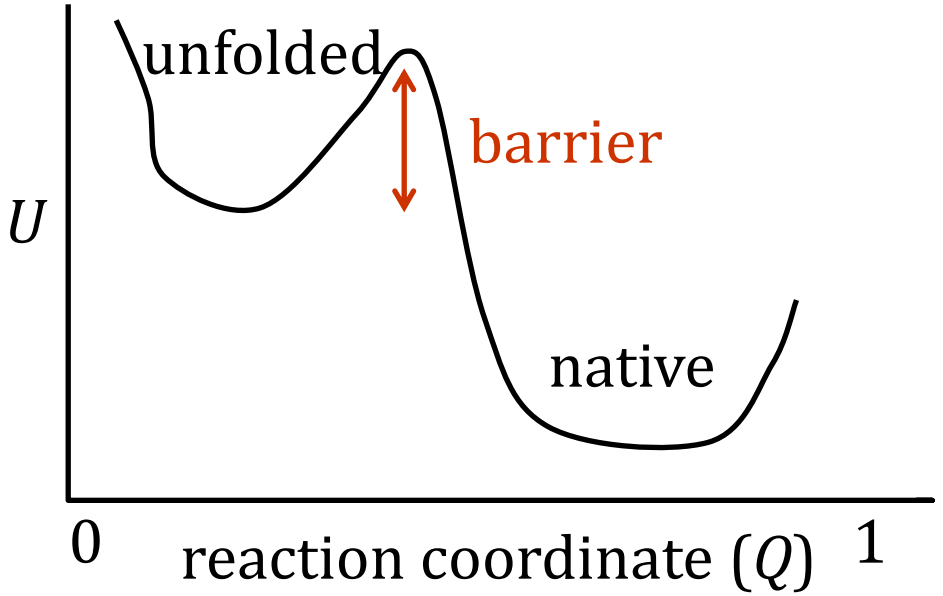
- not a classic one
- can we invent one ? yes
  - if two atoms are in contact in the final structure
    - native contact
  - Q = fraction of contacts which are correct



Reaction path

Is this like a chemical reaction ?

- no
- many molecules have same Q, but different conformations



We want at least two state kinetics

- where does barrier come from ?

Entropic barriers

Nature cares about free energies

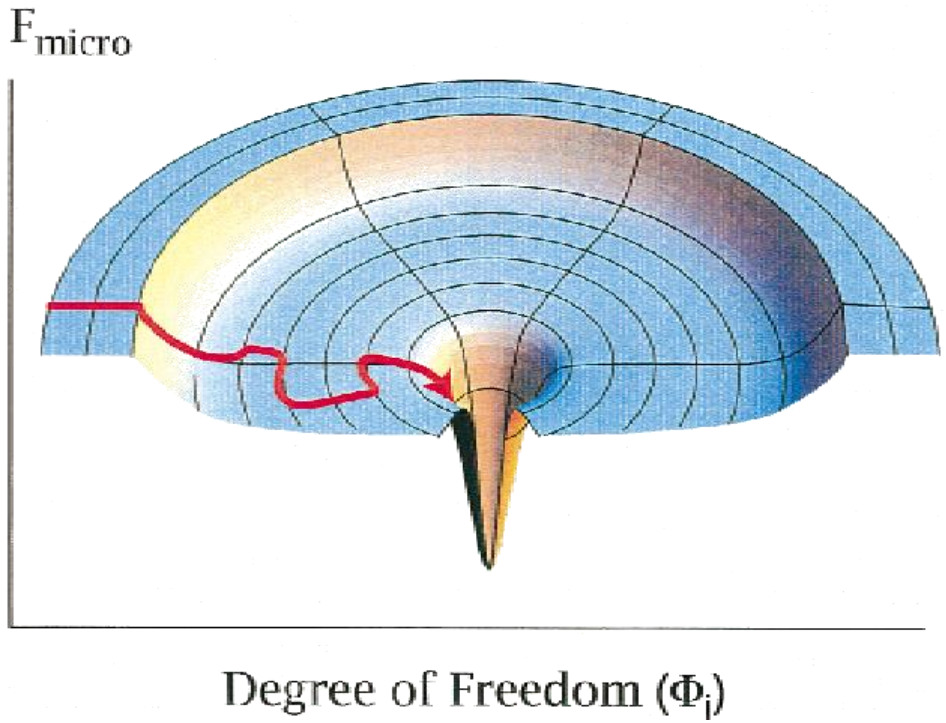
$$\Delta G = U - T\Delta S$$

If a molecule walks around

- it takes a long time
- looks the same as an energy barrier

- Are these pictures useful ?
- Do they agree with calculation ?

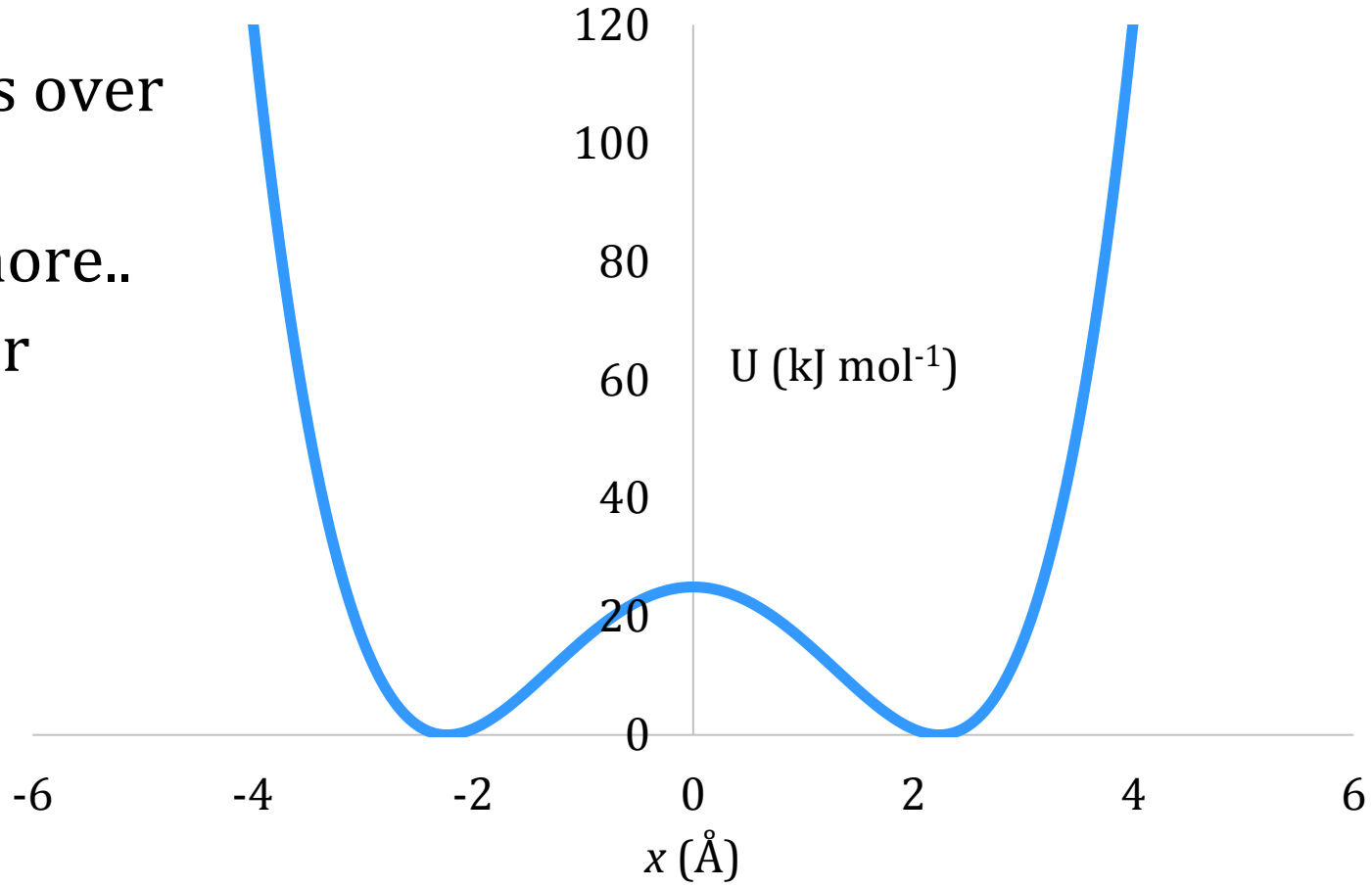
- have we seen an entropic barrier before ?
  - an entropic valley ?



Dill, K.A., Protein Sci., 8, 1166-1180, 1999, Polymer principles and protein folding

entropy – story from the Übung

- we have an energy surface
- system prefers low energy regions
  - no surprise
- heat the system and it jumps over barrier
- heat the system more and more..
  - it did not cross the barrier more often
  - why ?

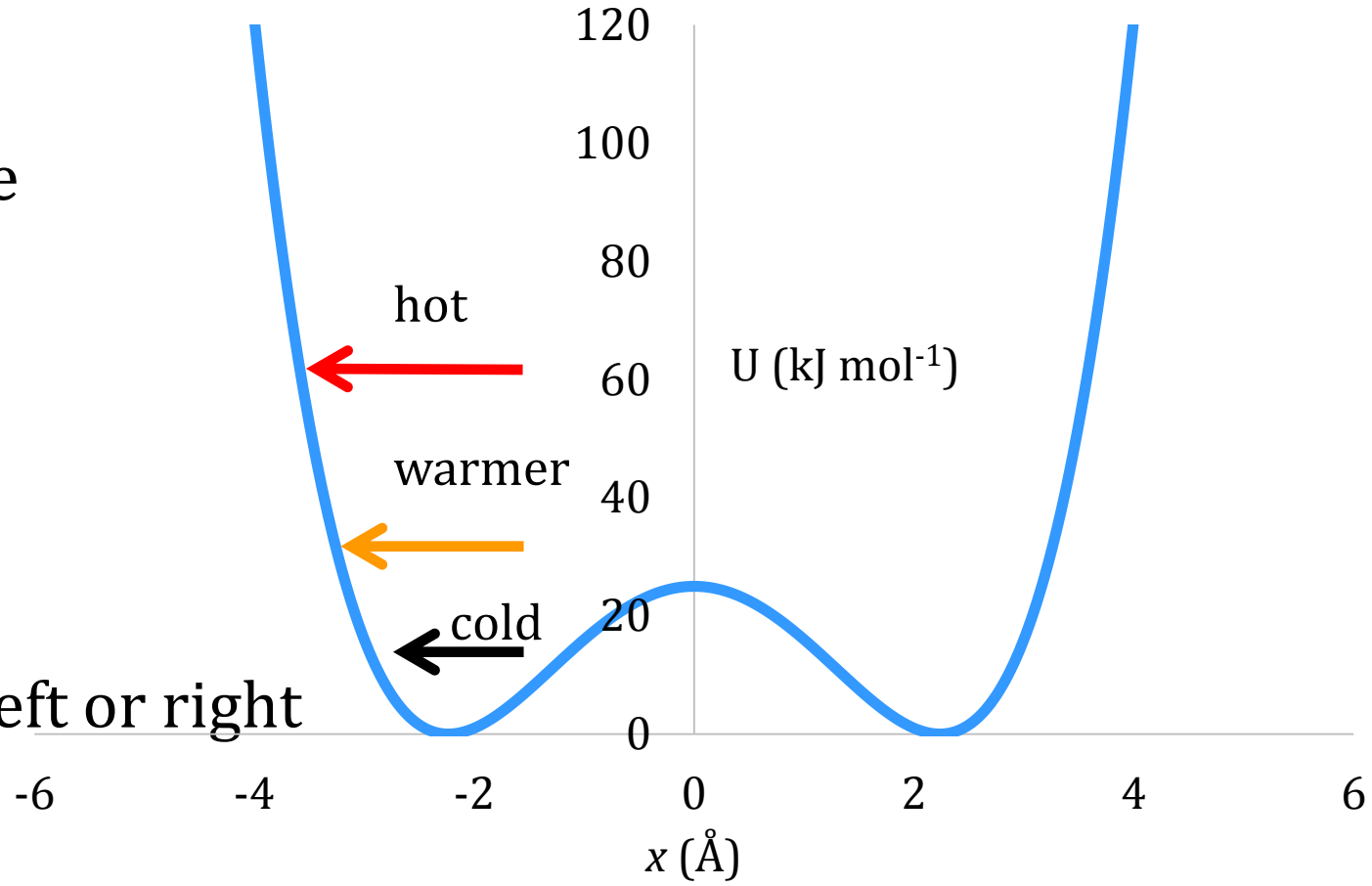


entropy – story from the Übung

- from potential energy
  - system does not want move far from centre

- as we heat the system
  - there is more space to the far left or far right

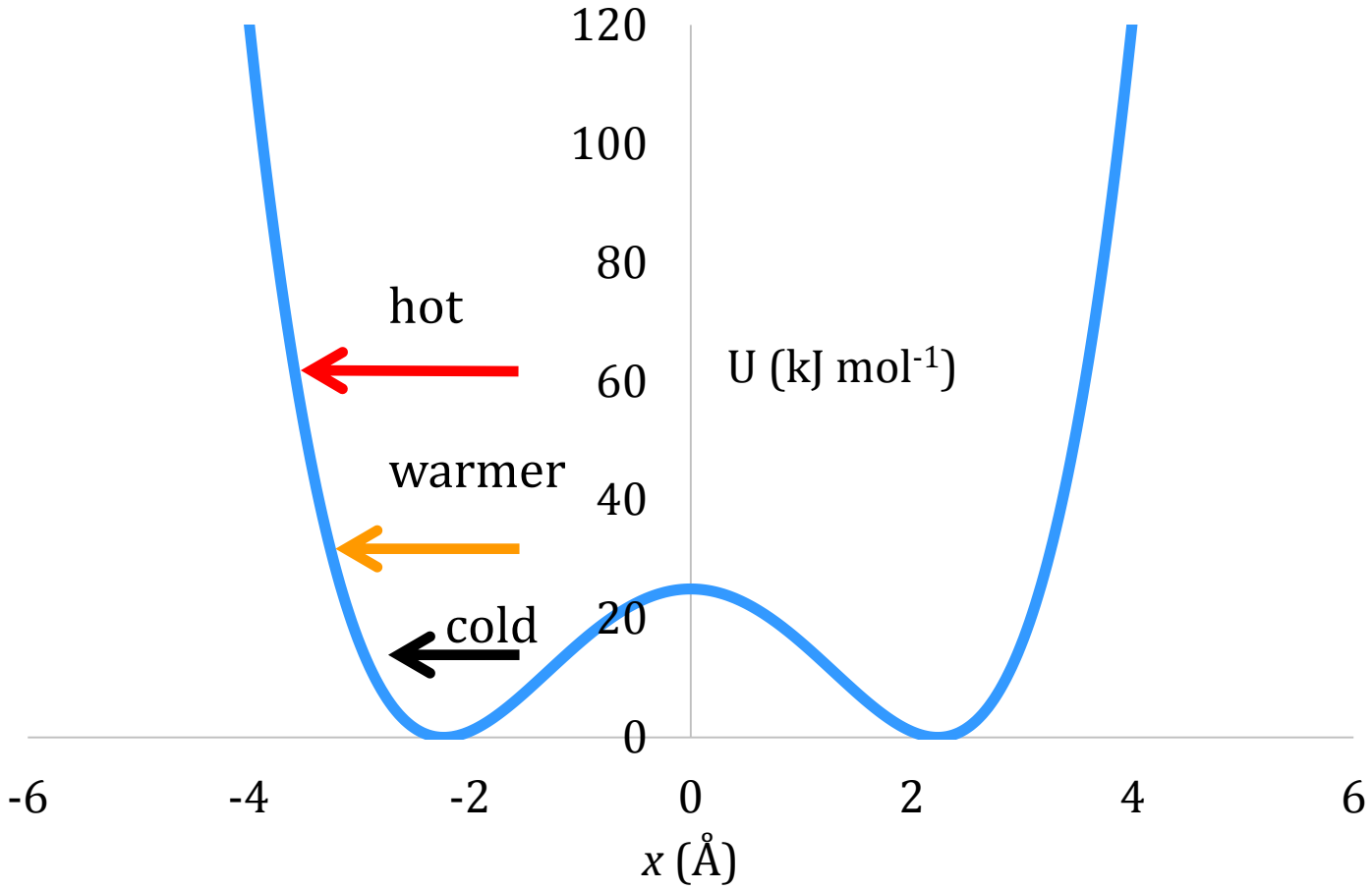
  - how much space ?
    - potentiall infinite
- when it is hot
  - the system can be far to left or right



entropy – story from the Übung

regions far left or far right seem attractive

- because they are potentiall large (at high temperature)

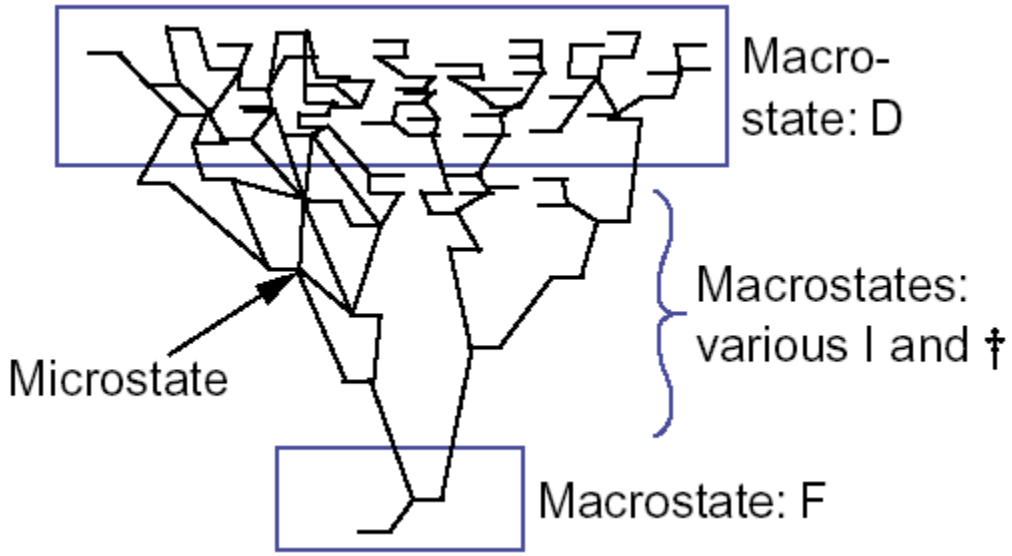


Interpretation of landscape

Does this disagree with conventional pathway ?

$A \rightarrow B \rightarrow C$?

- mostly at early stages
- there is a multiplicity of "A"
- when near native, there are relatively few conformations, so there may be something more like a pathway

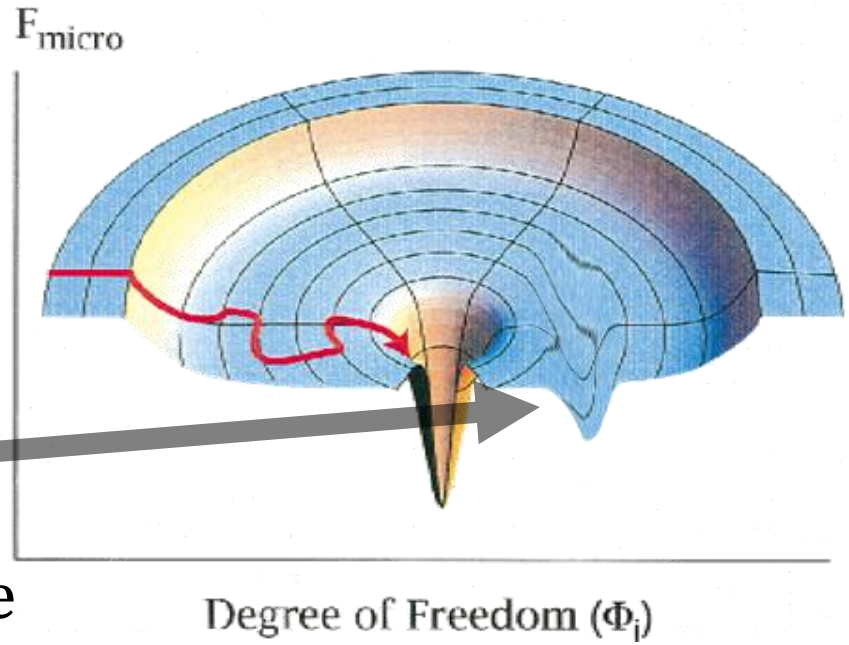


Gruebele, M, Curr. Opin. Struct. Biol. 12, 161-168, 2002, Protein folding: the free energy surface

Plausible ? Agreement with experiment ?

Difficult ..

- Experiments measure an average over all molecules
  - these are the same in different models
  - should we expect well defined intermediates ?
    - not really
    - what if one sees them ?
      - say they are valleys on the energy landscape



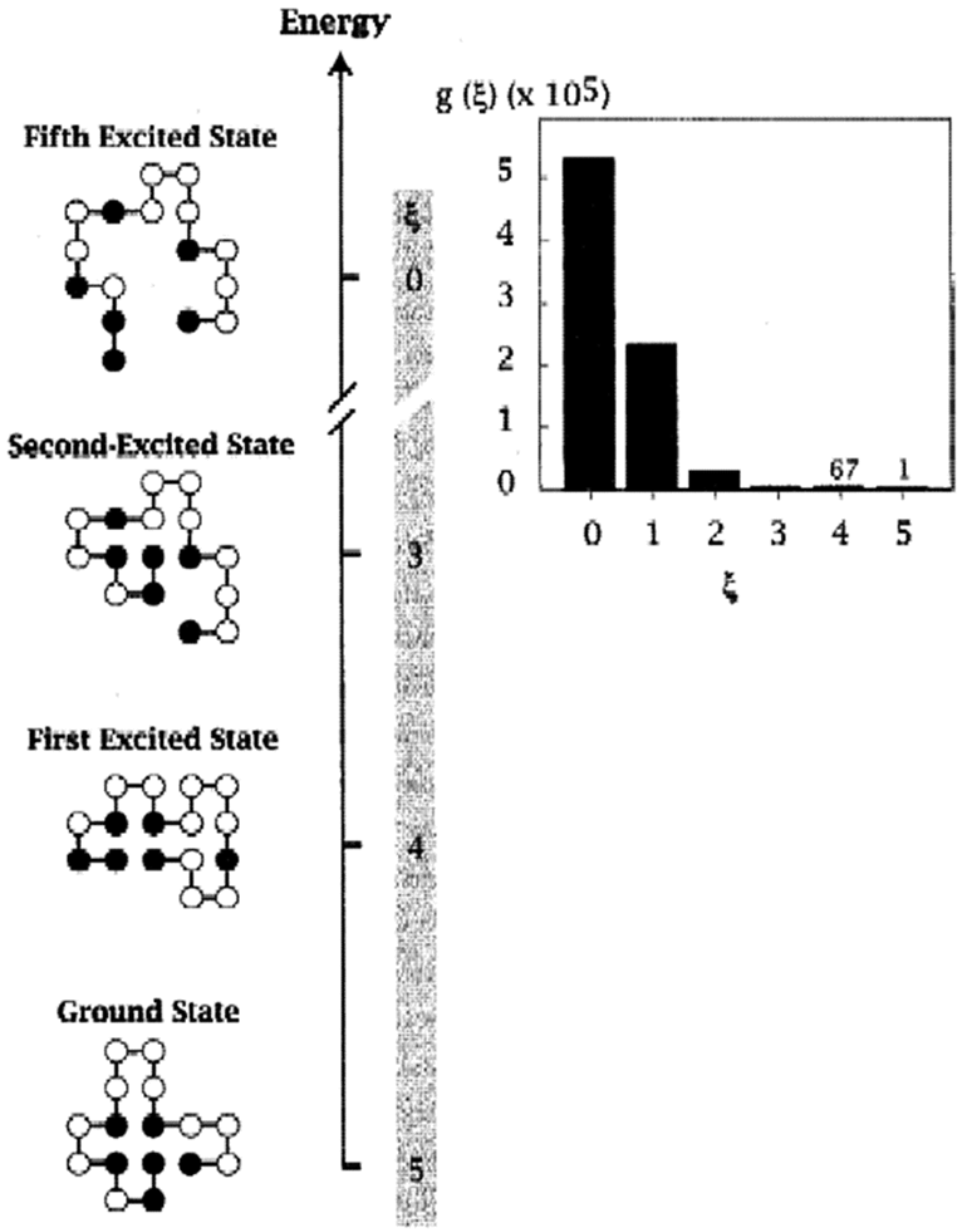
- Hard to find testable predictions

Simulations ?

- Cannot simulate a protein in detail
- Can simulate a discretised, simple protein
  - one point per amino acid
  - coordinates site on grid /lattice / gitter points

Question

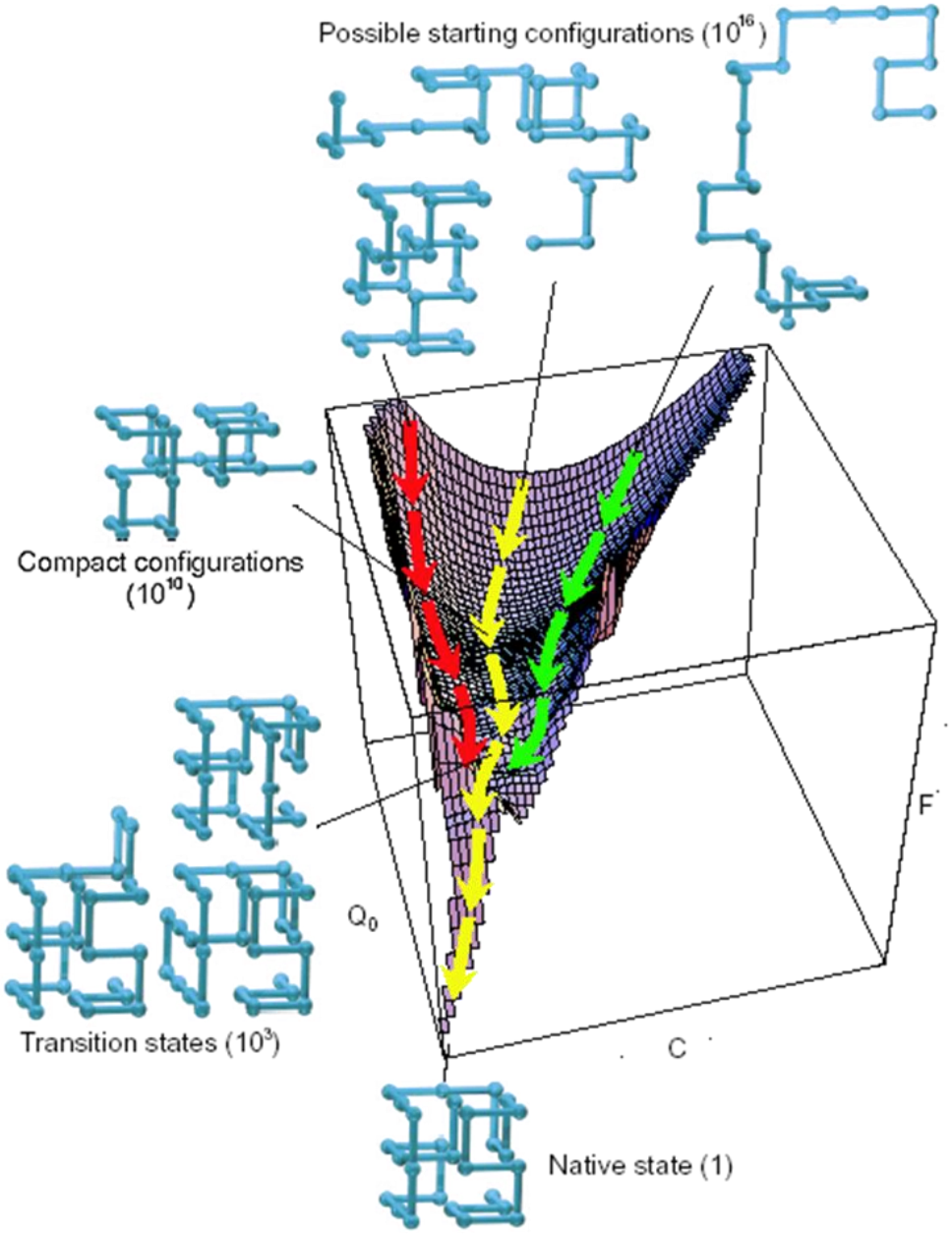
- lowest possible energy E_{min}
- how many conformations have
  - $E = E_{mi}$? $E = E_{mi}$ + a bit ? …



A larger calculation

27 residue

- simple lattice model
- estimations by sampling
  - not exhaustive
  - Q_0 correct contacts
  - C total contacts
  - F free energy



Summary

- Experiment vs. theory
  - experiment usually gives us averages
  - most calculations look at details
- Very different views on folding may be hard to distinguish
- Folding may be guided by sidechains (not hierarchical)
- Early folding may be best modelled by very crude models (so space can be sampled)
- Any useful model should predict exponential kinetics (or more complicated)