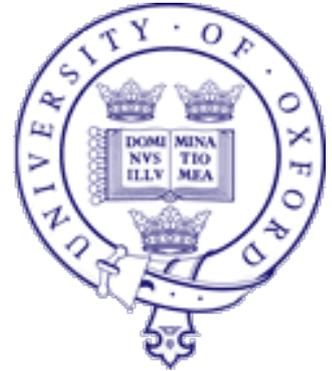


Representability problems for coarse-grained potentials

Ard Louis



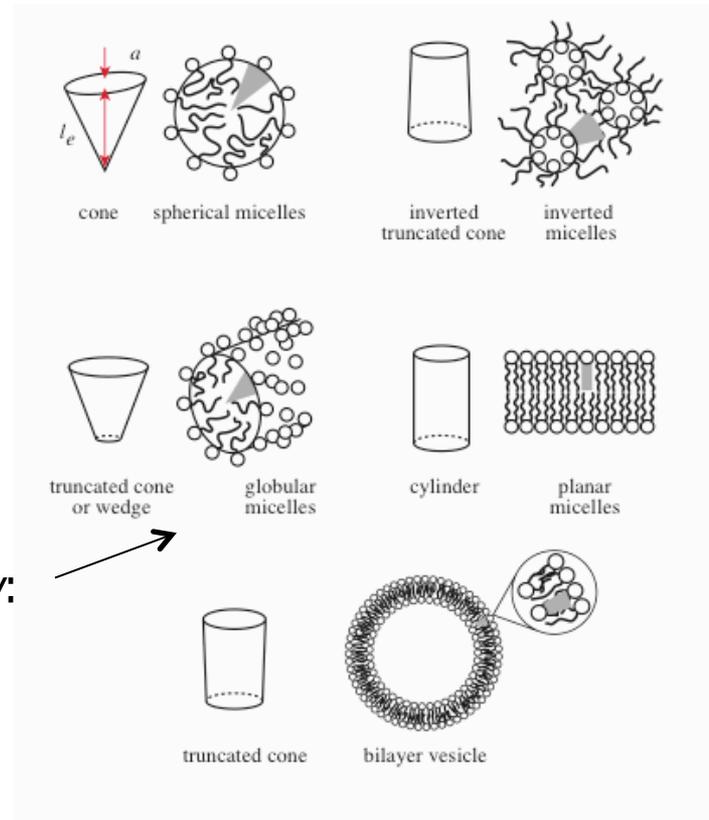
Emergent physics/chemistry and coarse-graining



"More is Different".

P.W. Anderson,

Science 177, 393-396 (1972).



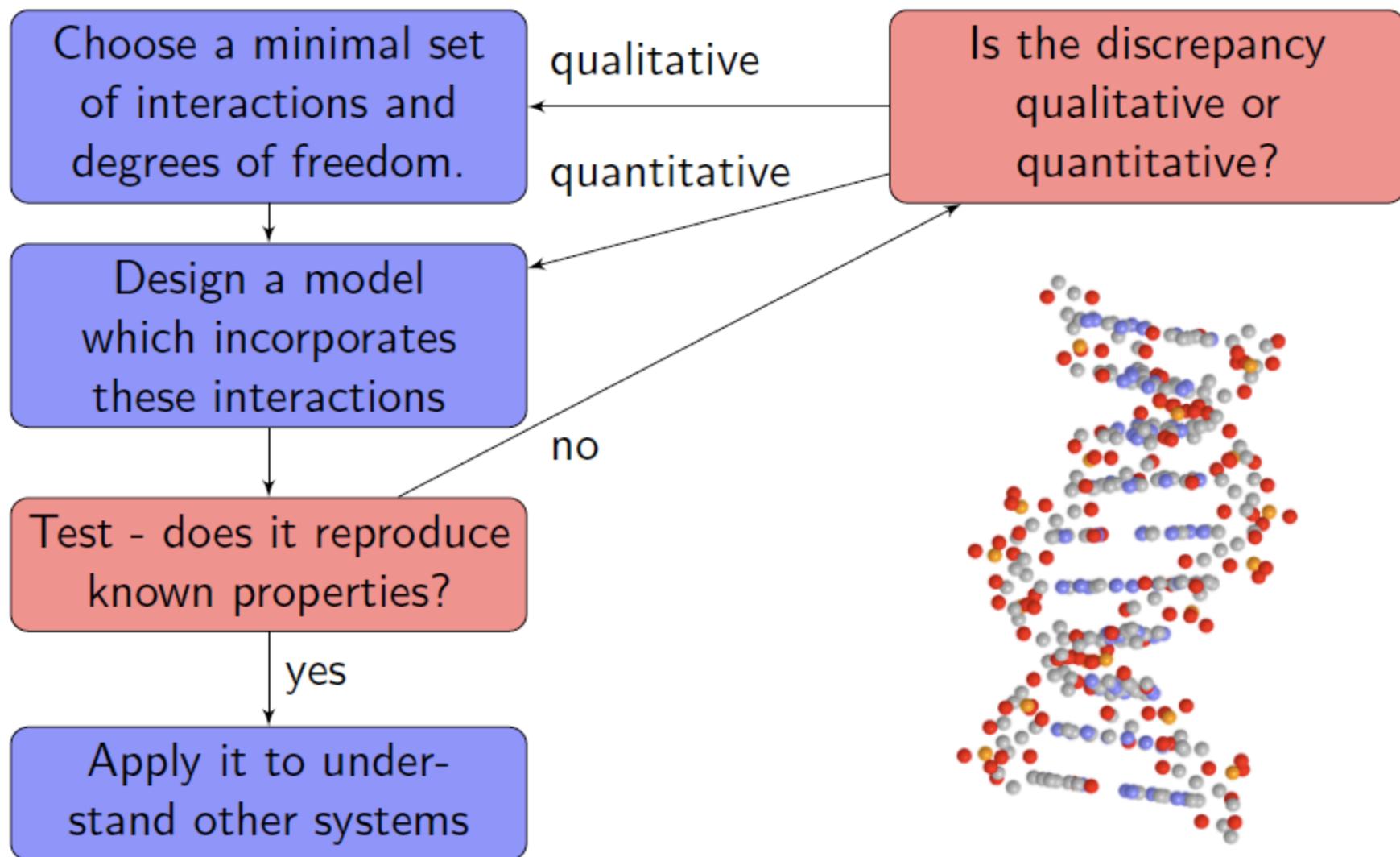
- Micellar phases emerge from geometry:
- SAW and polymer scaling
- Hard-spheres and structure/crystallization

How to model DNA self-assembly?

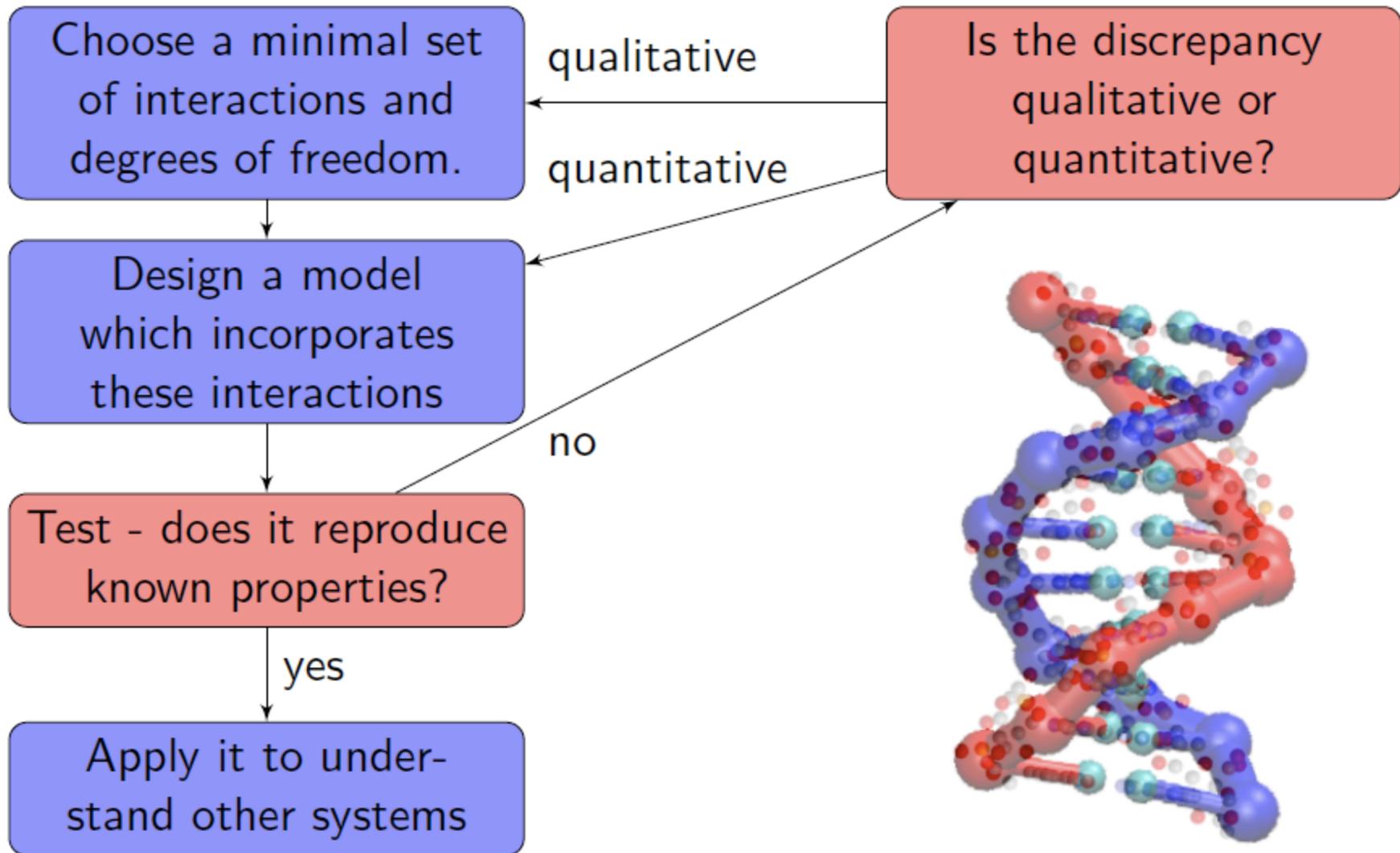
- Atomistic models
 - orders of magnitude too slow
- Bottom-up coarse-graining
 - Representability problems
- We use **top-down coarse-graining** instead
 - Self-assembly primarily determined by:
chain-like molecule with **specific binding**



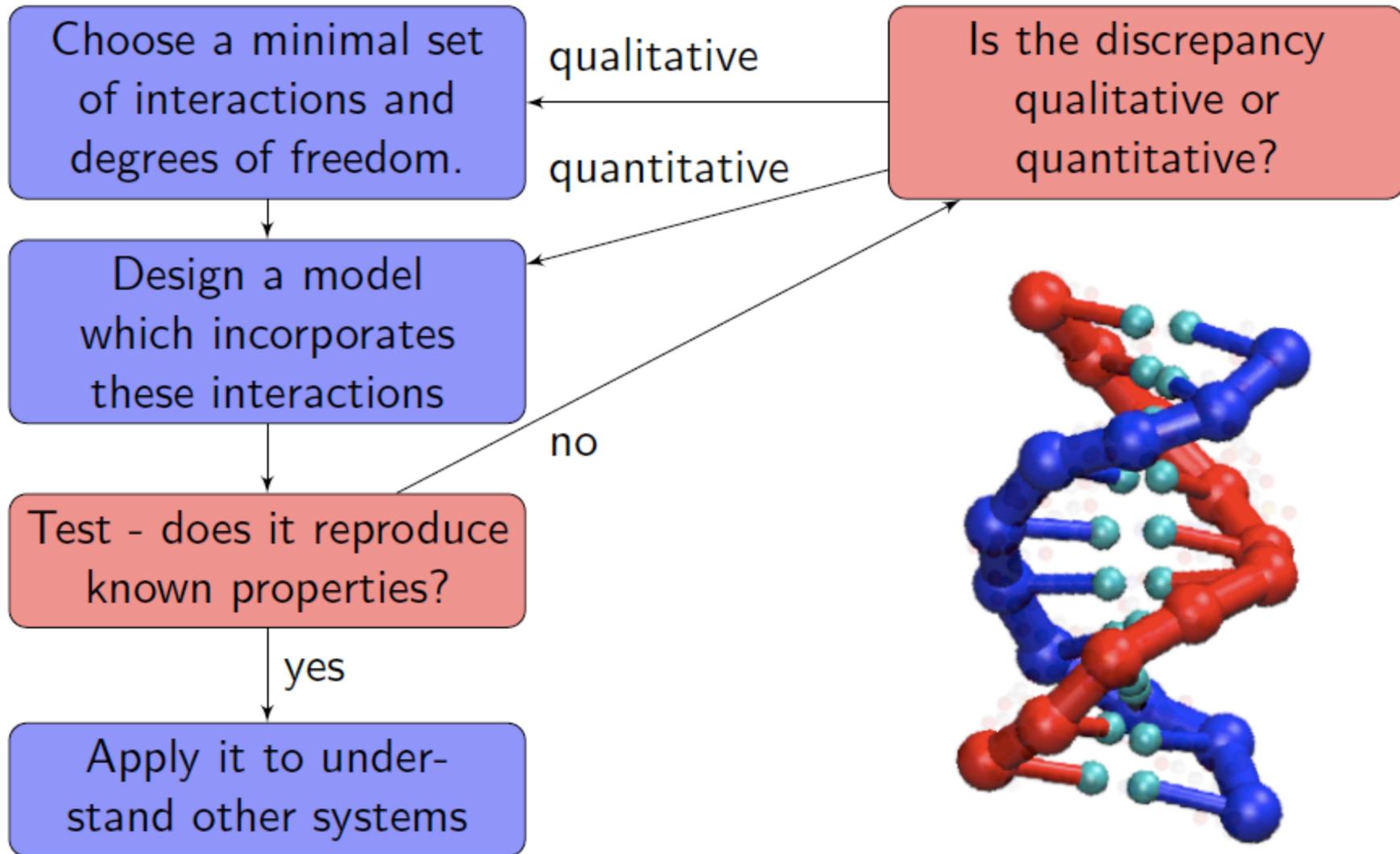
Creating a Minimal Model - a Top-Down Approach



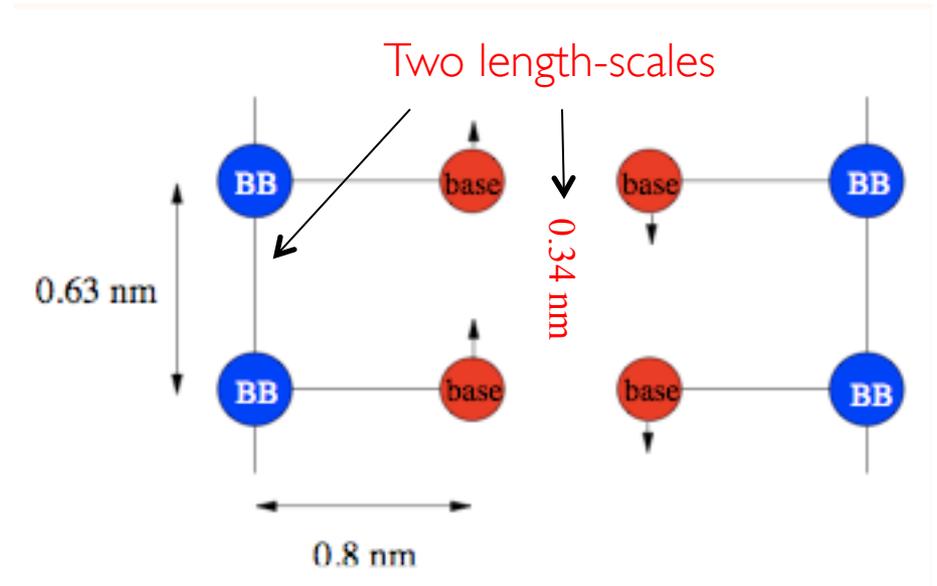
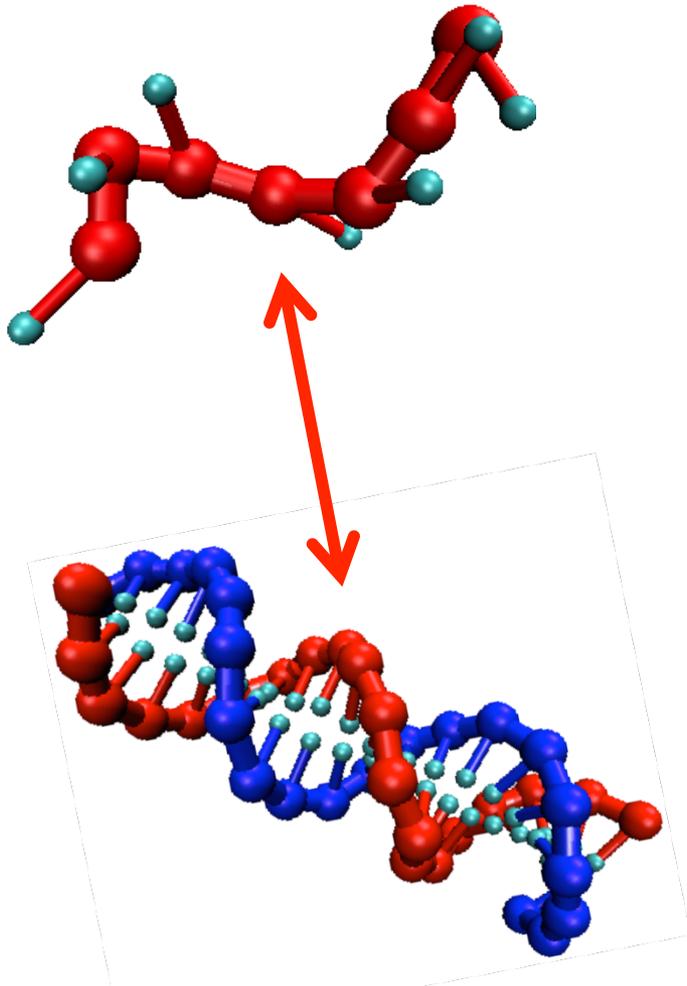
Creating a Minimal Model - a Top-Down Approach



Creating a Minimal Model - a Top-Down Approach



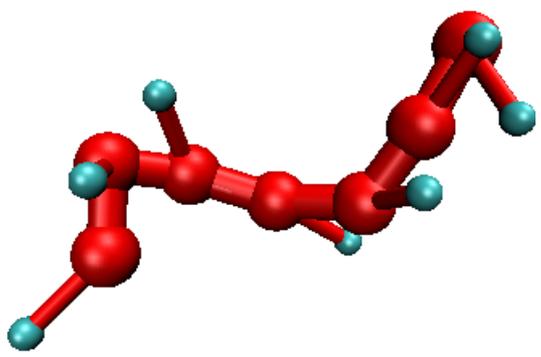
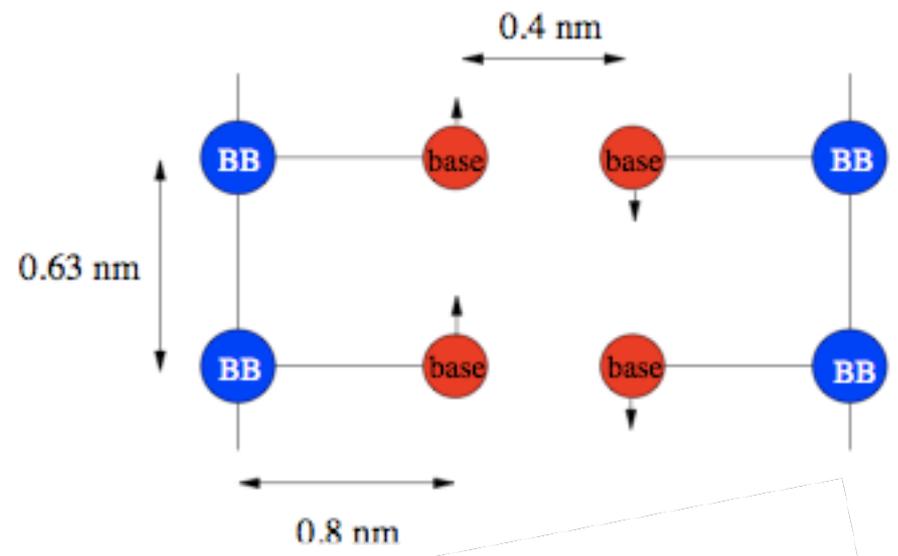
In DNA competition of 2 length-scales leads to double helix



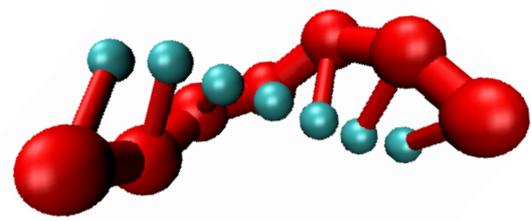
T. Ouldridge, A.A. Louis and J.P.K. Doye, Phys. Rev. Lett. **104** 178101 (2010);
J. Chem Phys. **134**, 085101 (2011)

In DNA single strands are flexible and can stack

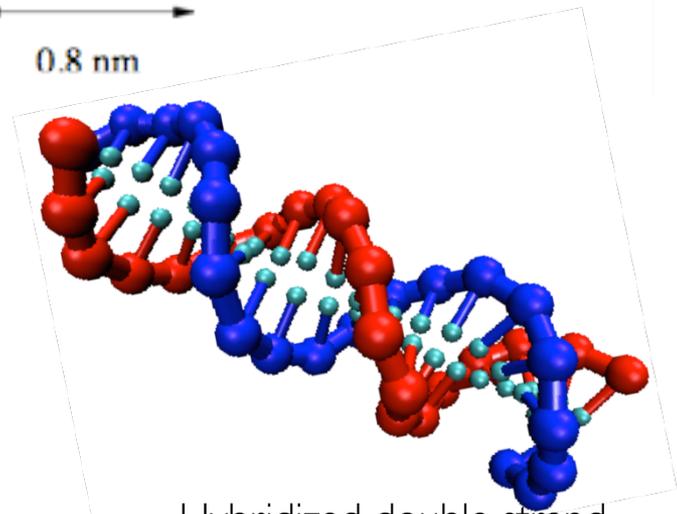
Competition between two length-scales leads to the double helix



disordered single strand



stacked single strand



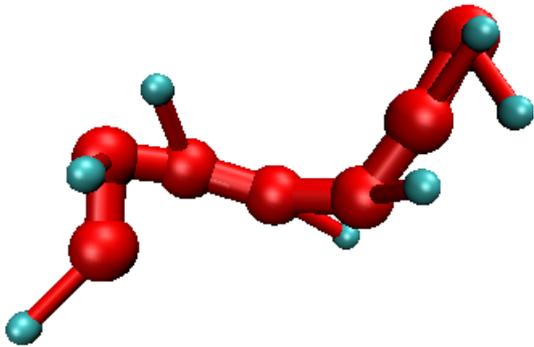
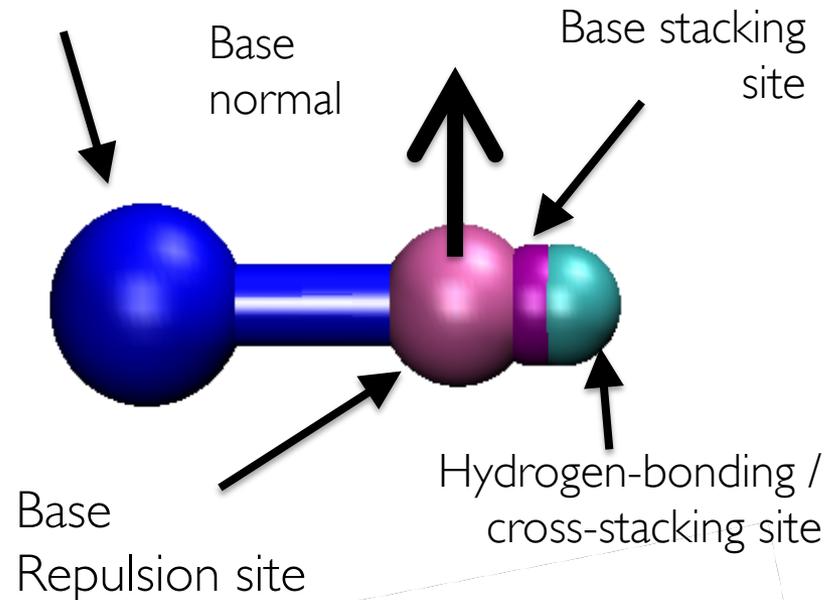
Hybridized double strand

T. Ouldridge, A.A. Louis and J.P.K. Doye, Phys. Rev. Lett. 104 178101 (2010);
J. Chem Phys. 134, 085101 (2011)

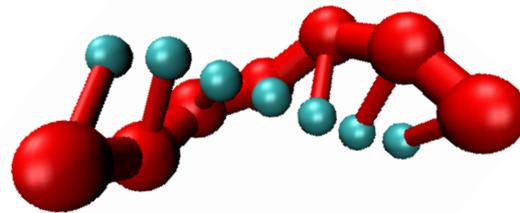
Simple coarse-grained DNA model

- Interactions

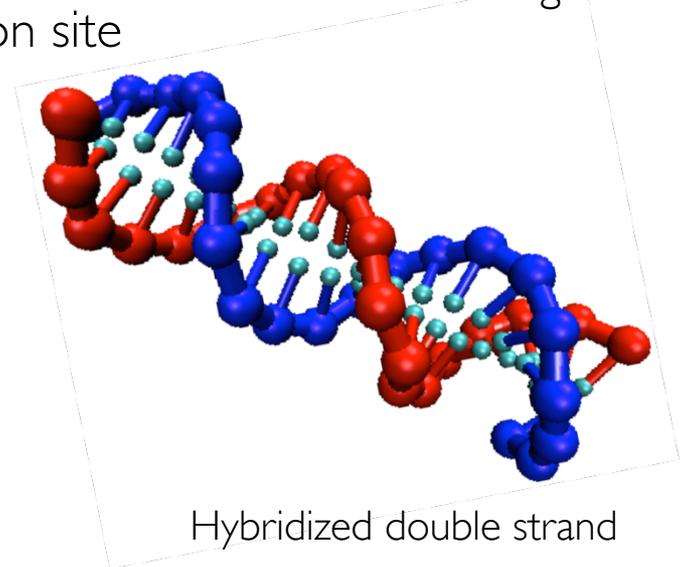
- H-bond between complementary bases
- Stacking between bases
- Backbone: FENE spring
- Helicity emerges naturally
 - Propellor twist emerges naturally
- But no minor/major groove



disordered single strand

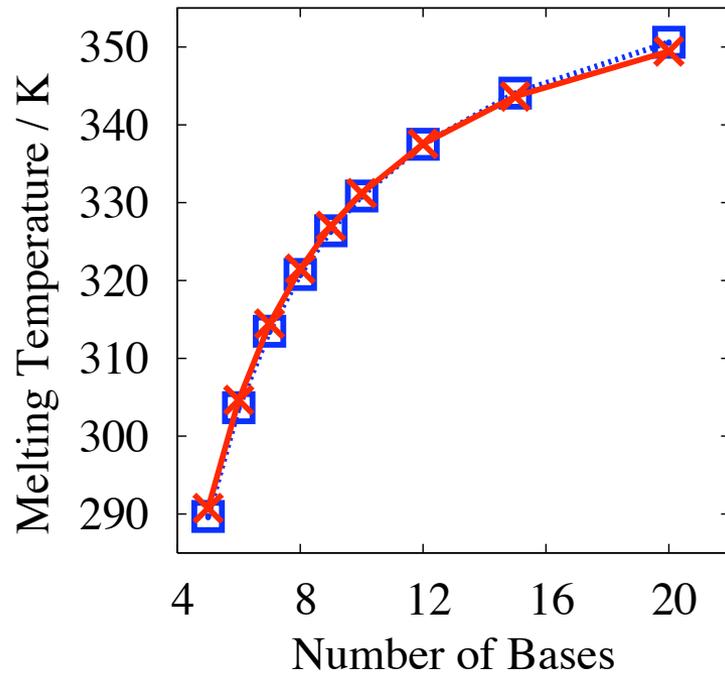


stacked single strand



Hybridized double strand

Duplex formation & length dependence

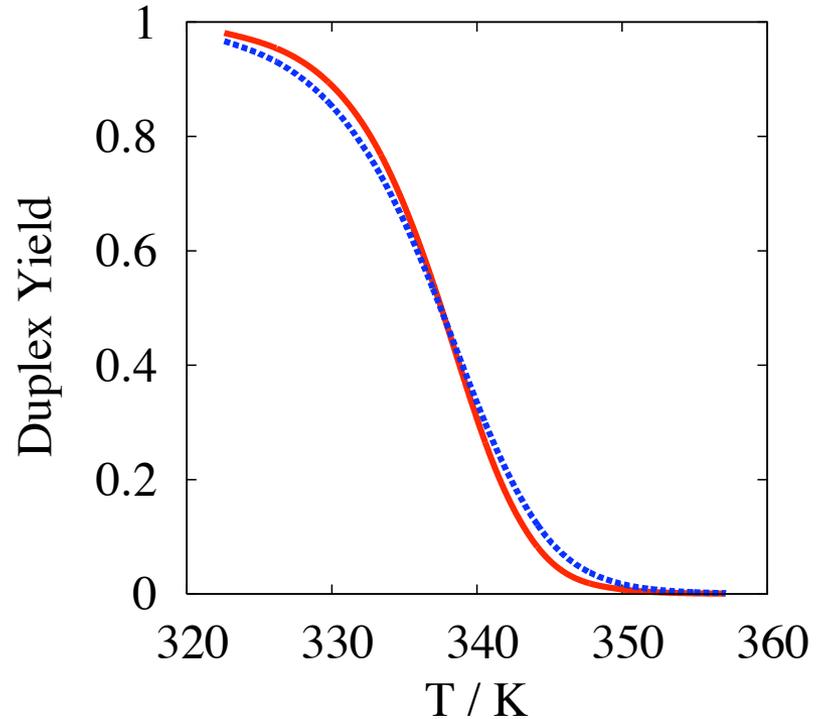


Good agreement of T_m with L is a measure of the cooperativity of the transition – influenced by the single strand cooperativity

Duplex formation & transition widths

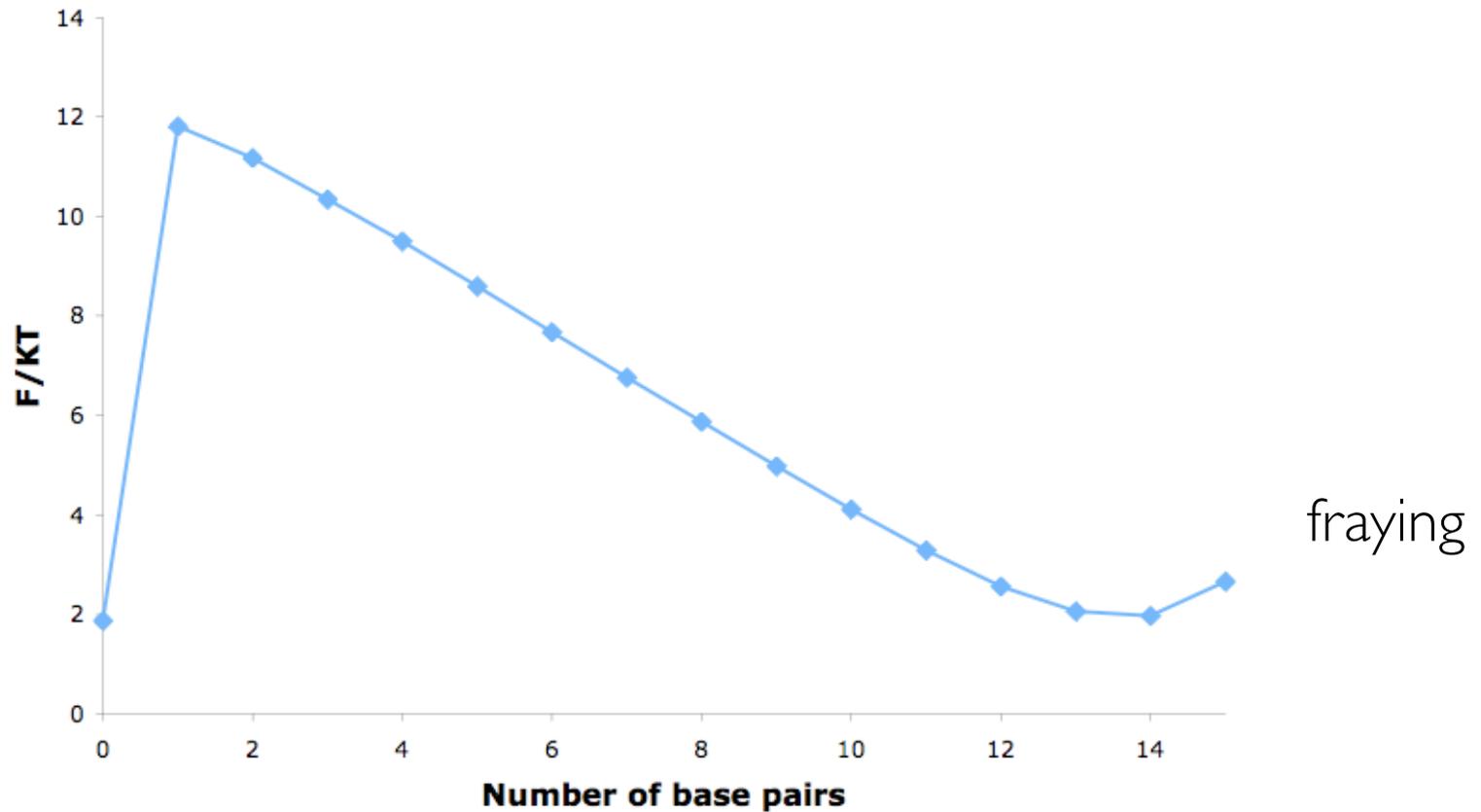
$$\Delta T_m \sim \frac{k_B T_m^2}{\Delta H}$$

$$\frac{dT_m}{d[c]} = -\frac{k_B T_m^2}{[c]\Delta H} \sim \frac{\Delta T}{[c]}$$



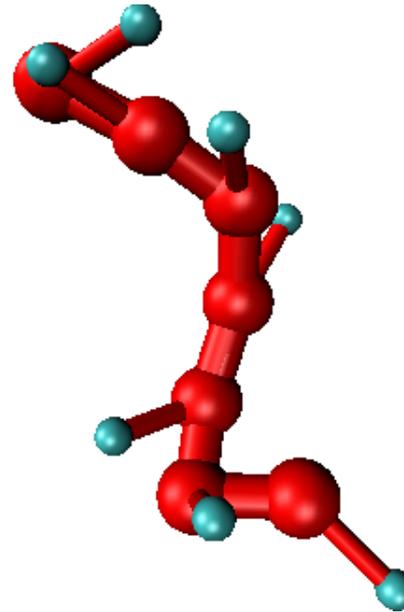
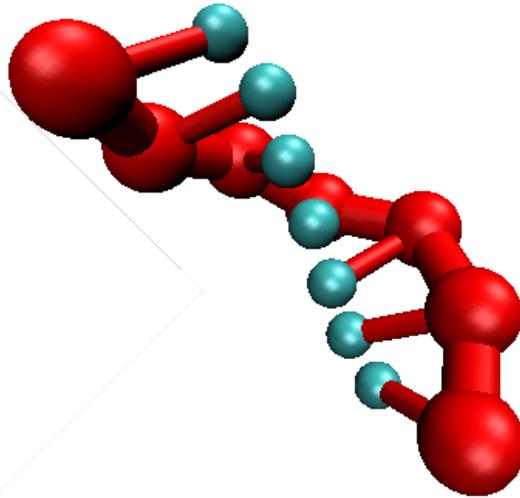
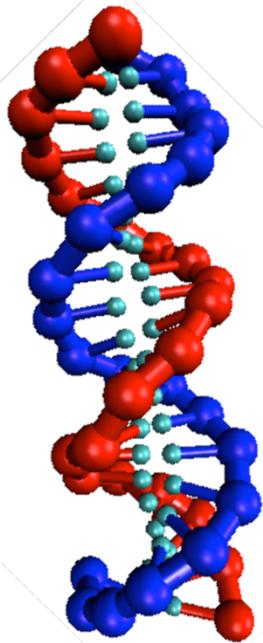
The width of the transition is related to how well you can predict the concentration dependence of the melting temperatures

Free-energy profile for duplex formation



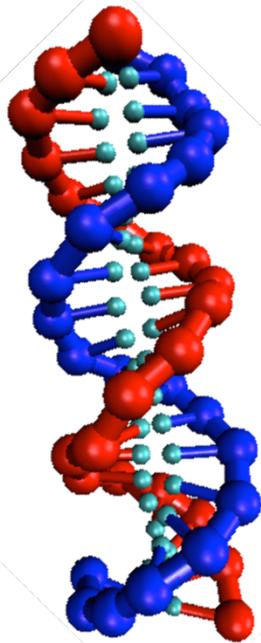
formation of a 15mer duplex

Mechanical properties



Duplex \sim 125 bp
Unstacked single strand \sim 2-4 bases
Fully stacked single strand \sim 64 bases
Twist persistence length of duplex \sim $3.74^\circ/\text{bp}$

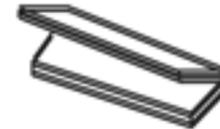
Mechanical properties – many subtleties we can't get



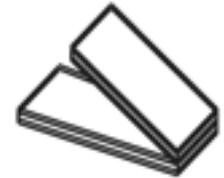
twist



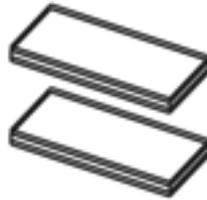
roll



tilt



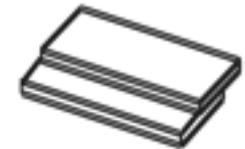
rise



slide

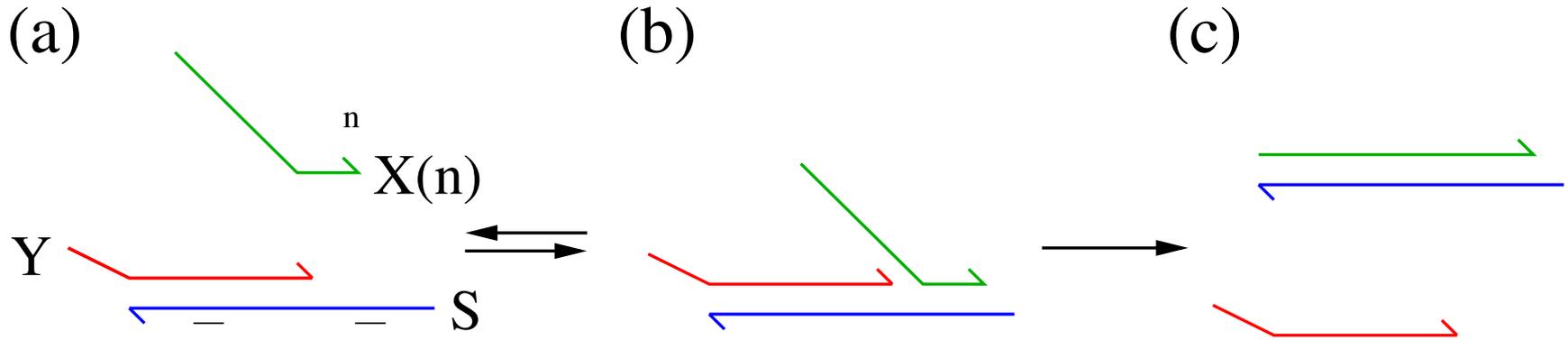


shift

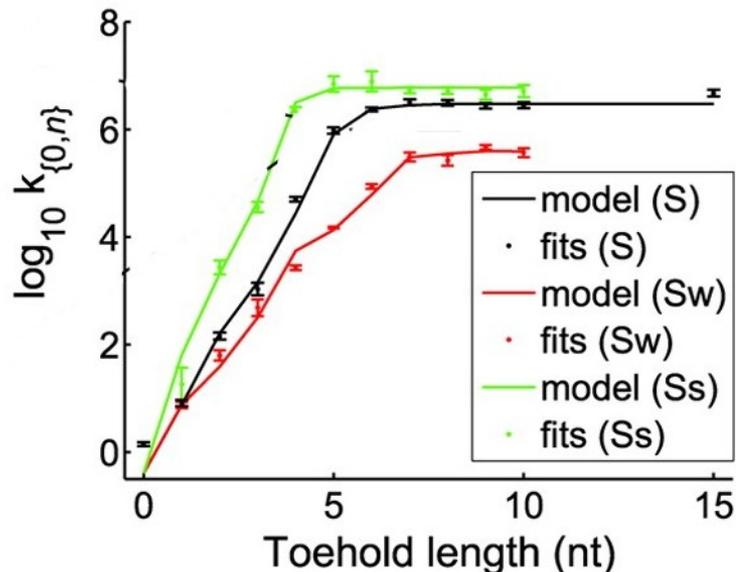


dsDNA undertwists upon initial stretching
Sequence dependent elastic properties are very very subtle – need a much better representation of excluded volume etc.....

Strand displacement reaction



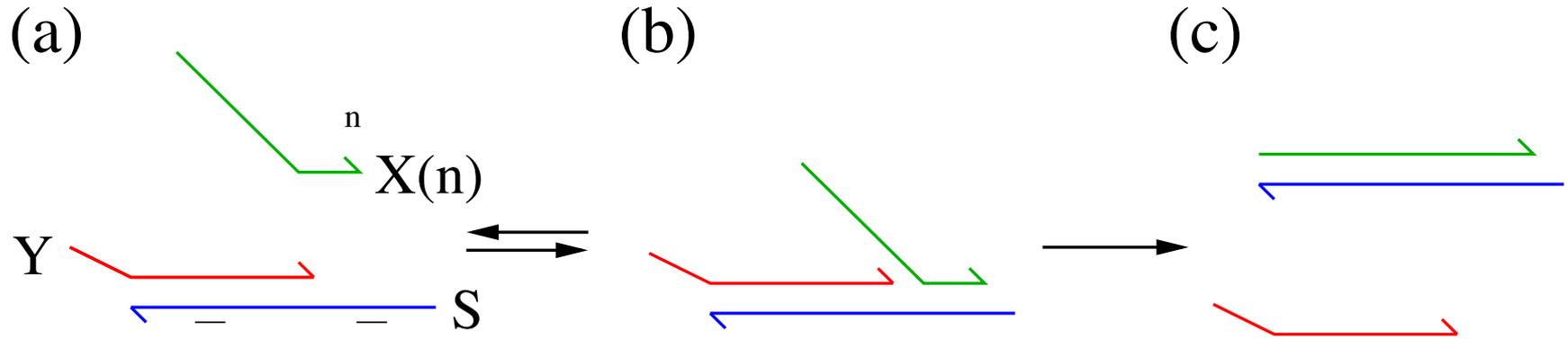
Displacement (or strand-exchange) reactions are key to dynamical behaviour of DNA nanodevices, and DNA computation



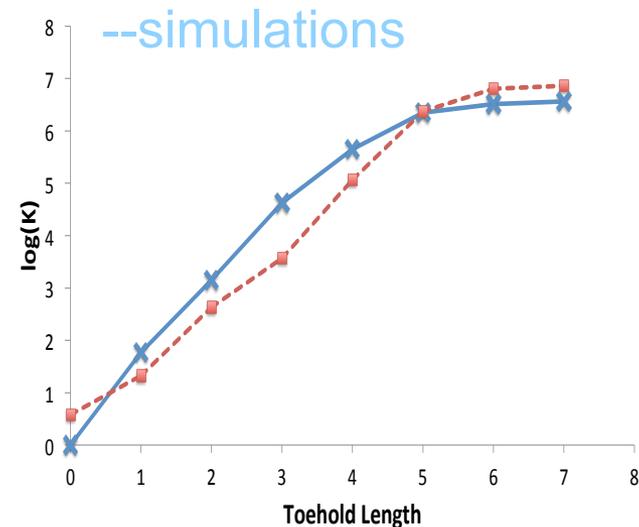
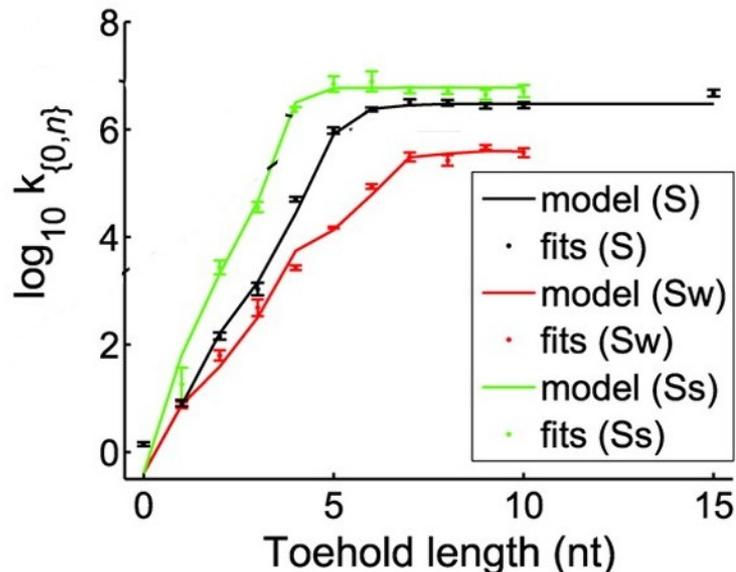
Rates change by orders of magnitude as toehold length increases, before leveling out at $nt \sim 5$

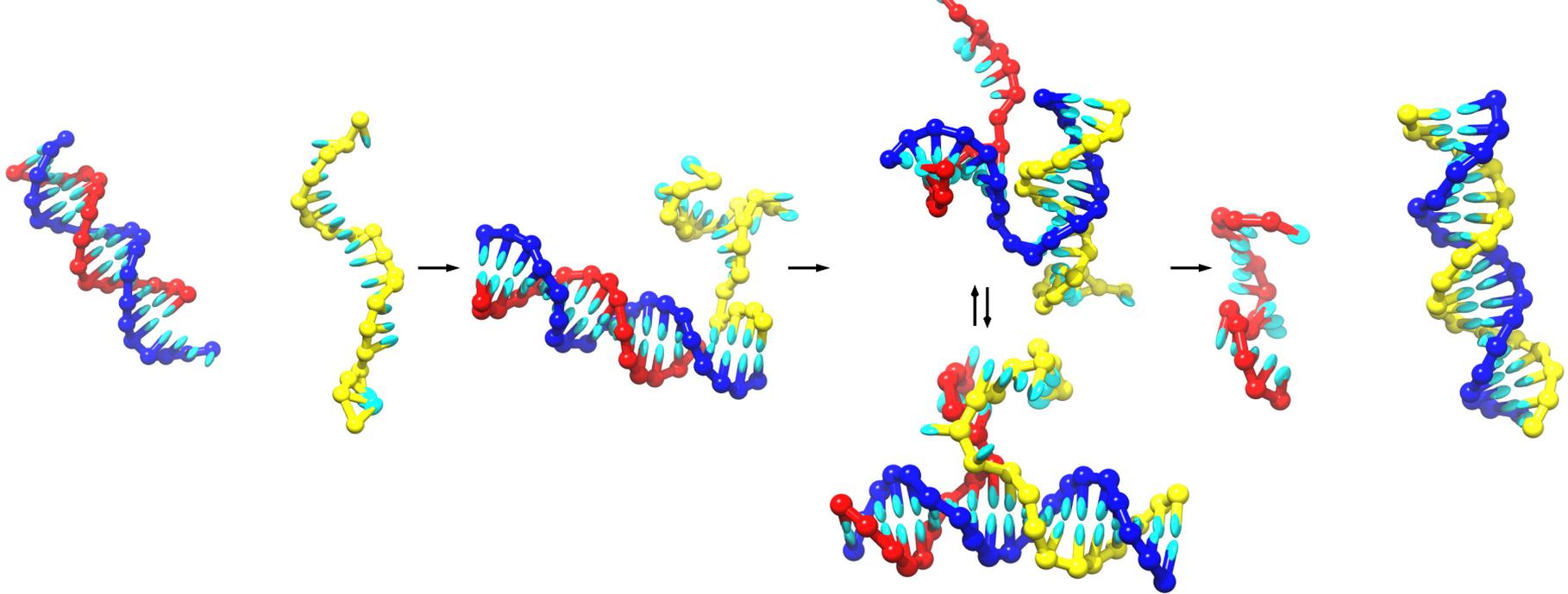
David Yu Zhang and Erik Winfree, JACS, **131**, 17305 (2009)

Strand displacement reaction

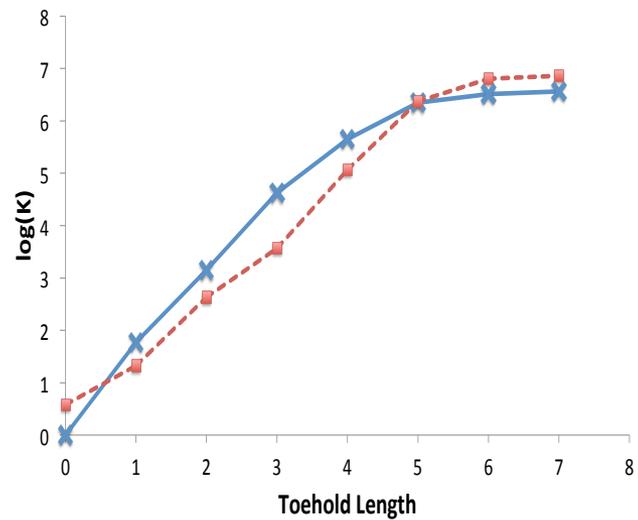
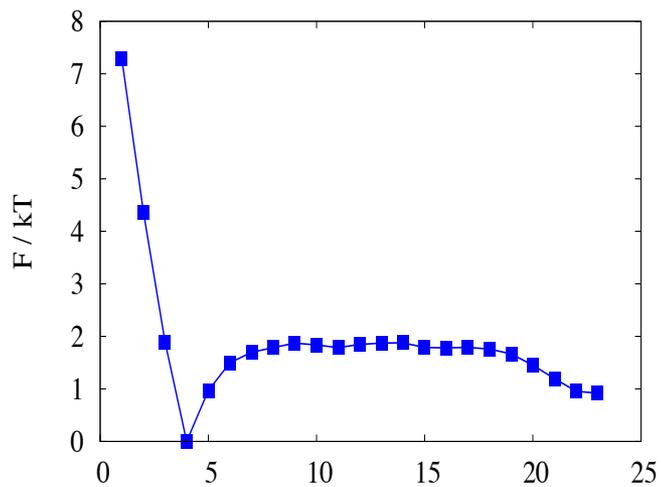


Displacement (or strand-exchange) reactions are key to dynamical behaviour of DNA nanodevices, and DNA computation

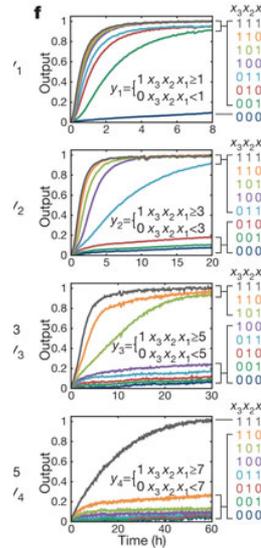
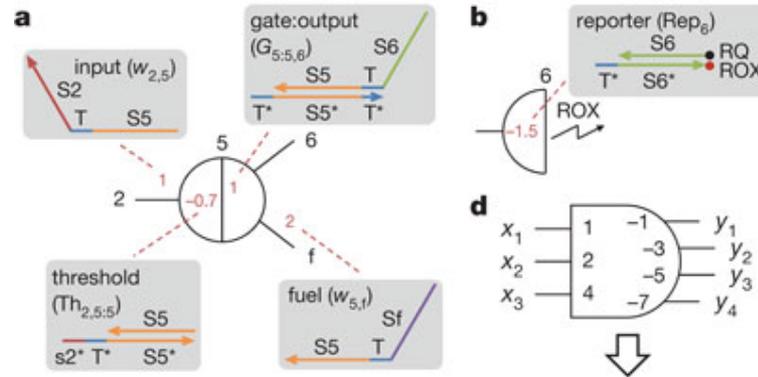
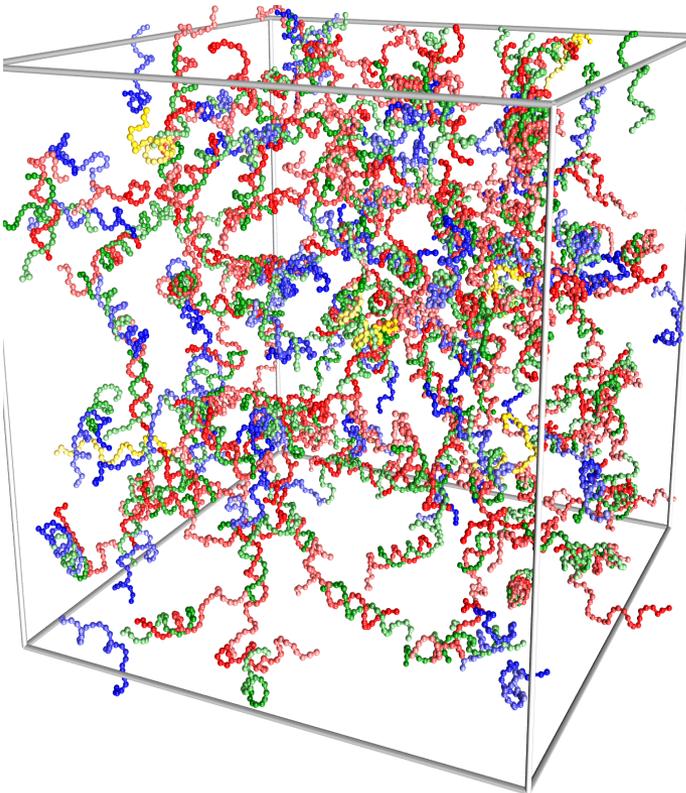




Coaxial stacking is important; displacement slowed by several effects



DNA computing OR gate



Neural network computation with DNA strand displacement cascades

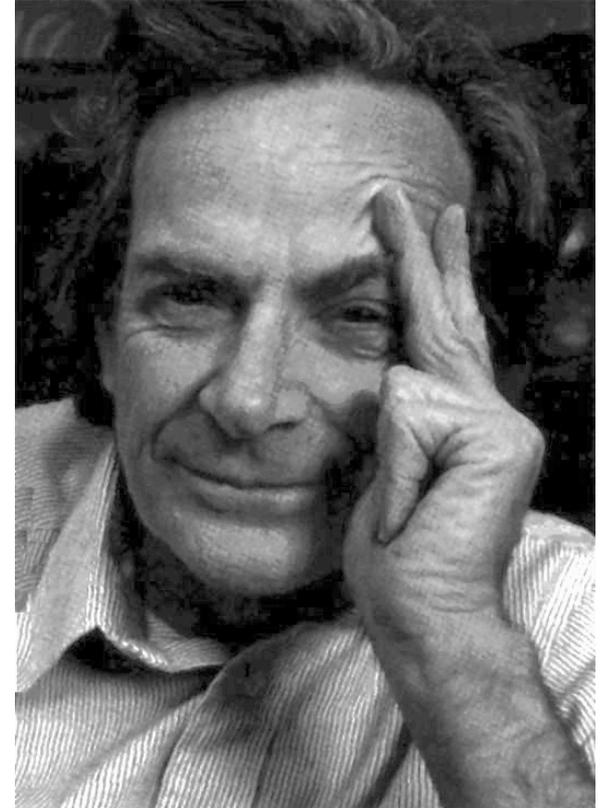
LL Qian *et al. Nature* 475, 368-372 (2011)

Simulation on GPU equiv ~ 10 min real time

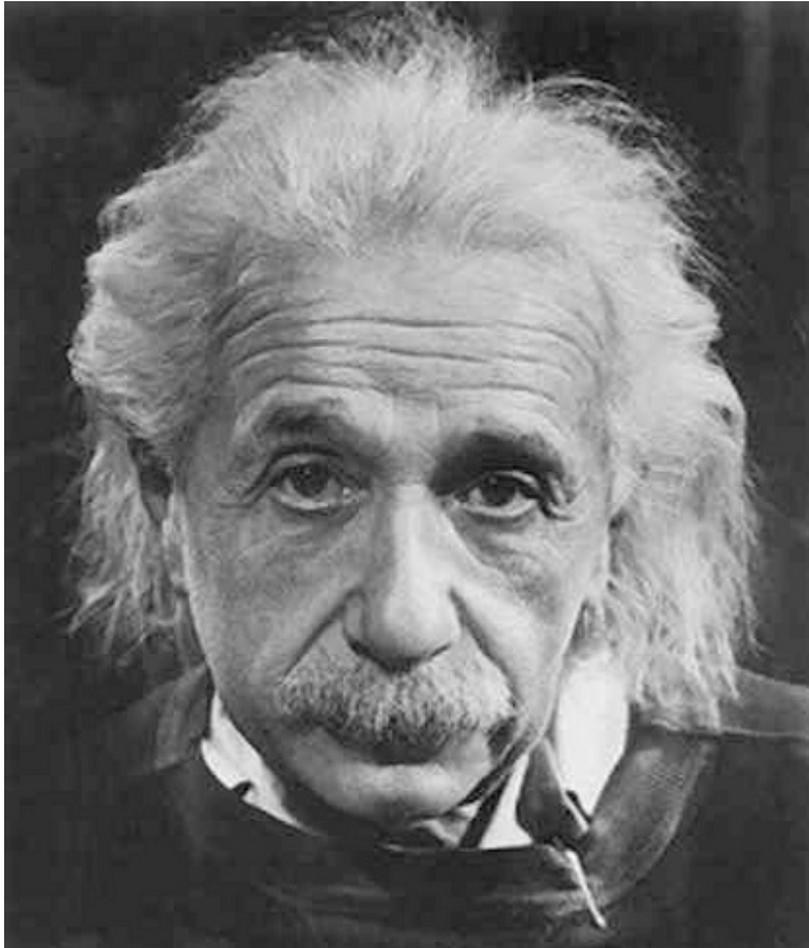
Just another pretty movie?

The first principle is that you must not fool yourself--and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that.

-- R.P. Feynman,
"Cargo Cult Science" (1974)



Einstein's Razor



Make things as simple as possible, but no simpler.

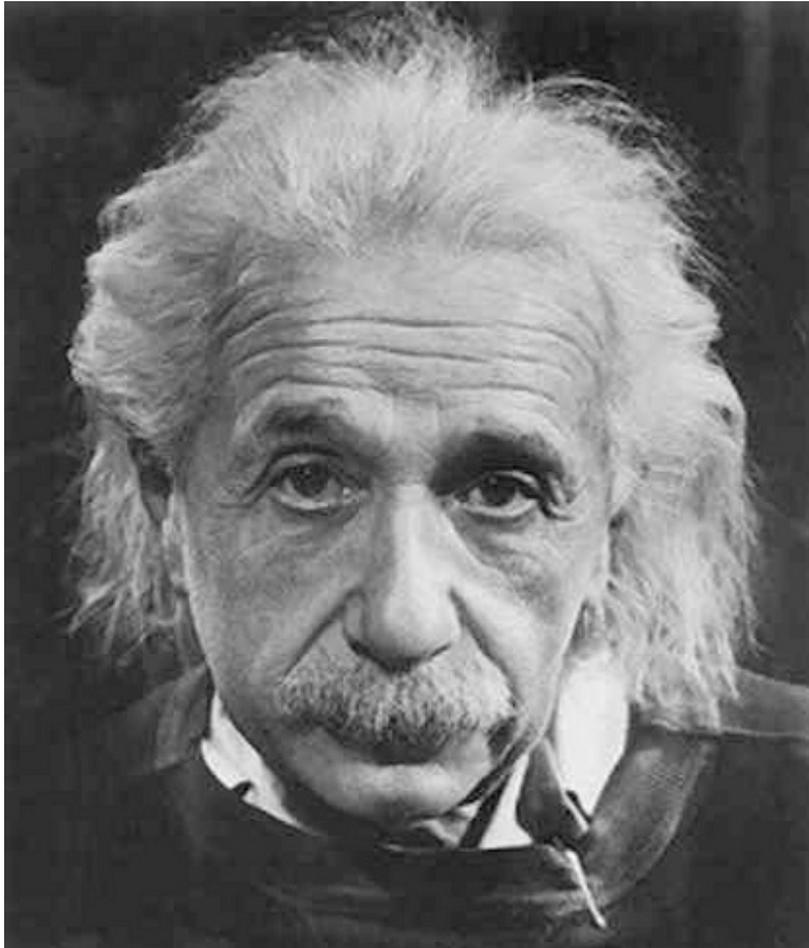
What he really said:

It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.

The Herbert Spencer Lecture,
delivered at Oxford (10 June 1933)

Survival of the fittest

Einstein's Razor & Coarse-grainer's nightmare



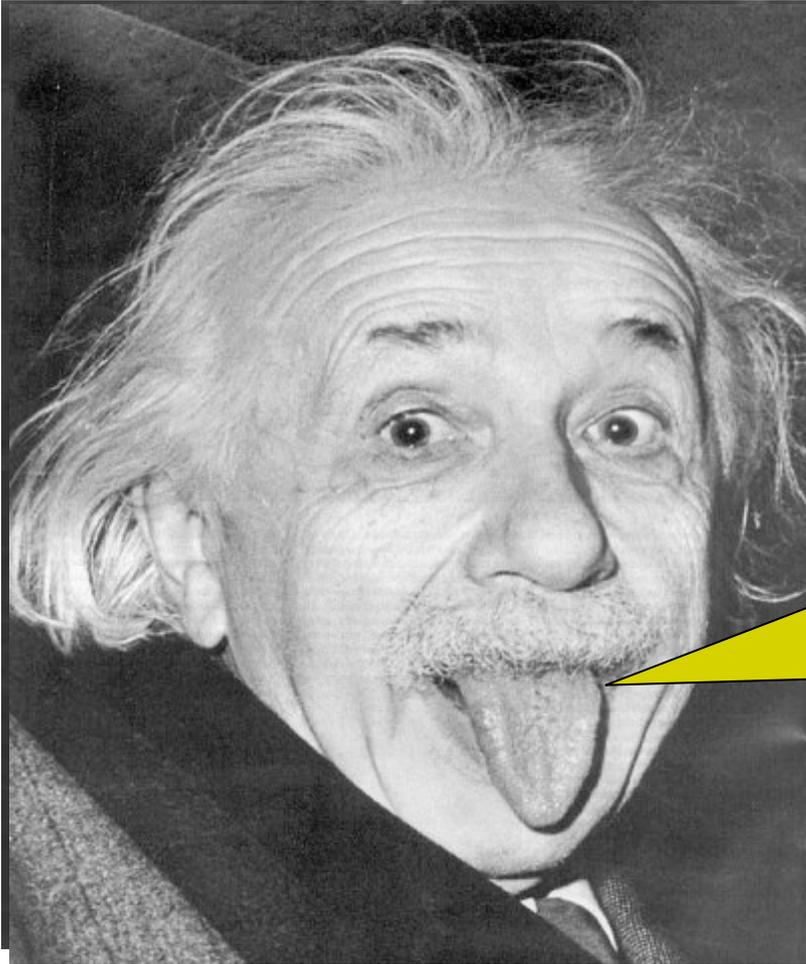
Make things as simple as possible, but no simpler.

You work really hard to derive a tractable coarse-grained system, only to find



Survival of the fittest

Einstein's Razor & Coarse-grainer's nightmare:



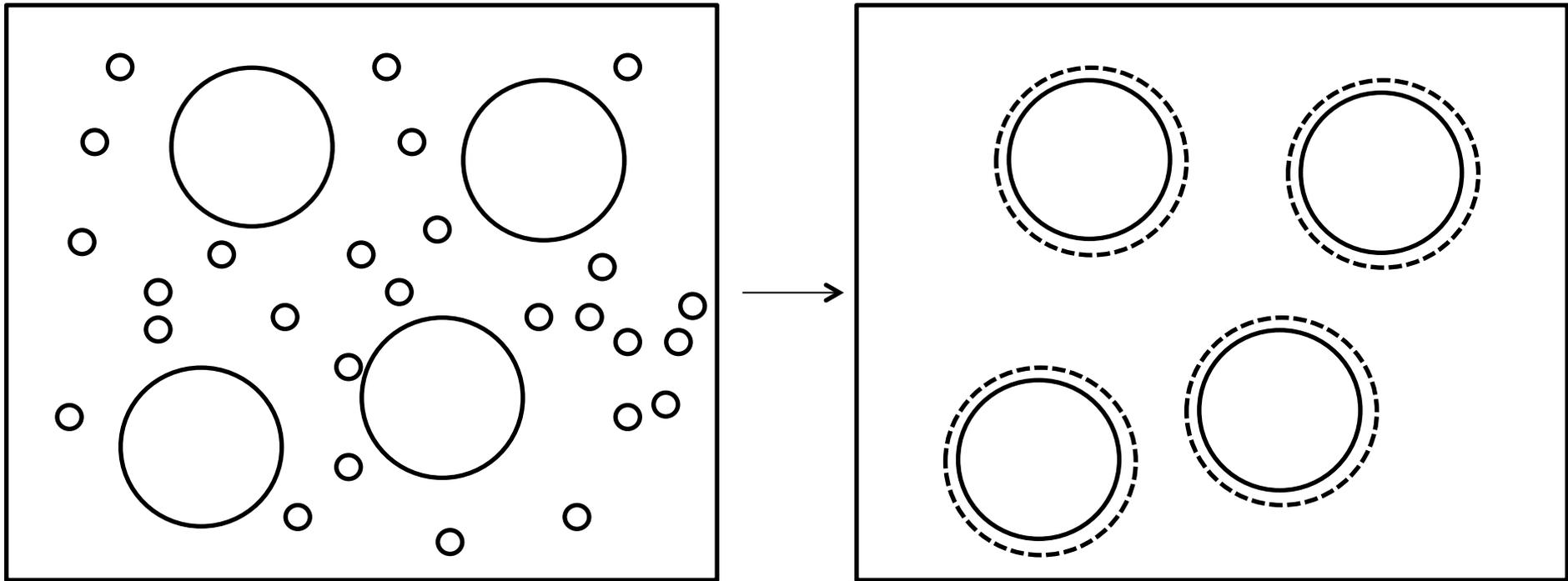
Make things as simple as possible, but no simpler.

Wrong:
too simple!!!

No free lunch theorems

- Effective-potentials for depletion systems
- Representability and potentials
 - A) Two and three-body potentials
 - B) Isotropic model for water
- Dynamics?
- DNA?

Case-study 1: Depletion interactions



Coarse-grain by equating partition functions (or all distribution functions of CG variables)

Only maps to an **effective Hamiltonian system** with effective potentials in grand-canonical or semi-grand ensembles.

In a **pure canonical ensemble**, one cannot decompose the potential into pairwise, three-body etc.. terms... see e.g. AAL, *J. Phys.: Condens. Matter* **14**, 9187 (2002)

AO model treated here :M. Dijkstra, JM Brader and R Evans, *J. Phys.: Condens. Matt.* **11**, 10079 (1999)

Case-study 1: Depletion interactions: Lessons?

1. Coarse-graining by equating partition functions or distribution functions is really a re-summation -- formally this could be done by tracing out any variable.
2. Be careful which ensemble you work in if you want to *map* to an effective potential or Hamiltonian system. [Canonical is suspect].
3. Phase-diagrams and thermodynamics can be worked out, but be careful to include zero-body and one-body terms – remember McMillan-Mayer.
4. Dynamics must be treated with care.

More, see: AAL, J. Phys.: Condens. Matter **14**, 9187 (2002)

Case-study 1b: Debye-Huckel potentials

Charge-neutrality means Canonical ensemble is natural & is effectively a one-component system

$$\beta v_{DH}(r; \rho) = \frac{Z^2}{r} \exp[-\kappa(\rho)r]. \quad \kappa(\rho) = \sqrt{4\pi Z\rho}$$

$$Z_c = \frac{\beta P}{\rho} = \int_0^\rho \frac{\partial \beta P(\rho')}{\partial \rho'} \frac{d\rho'}{\rho} = \int_0^\rho [1 - \rho' \hat{c}(k=0; \rho')] \frac{d\rho'}{\rho}$$

In RPA approximation $c(r) = -\beta v(r)$ $\beta \hat{v}(k=0; \rho) = \frac{4\pi Z^2}{\kappa^2} = \frac{Z}{\rho}$,

$$Z_c^{RPA} = 1 + Z,$$

$$Z_{vir} = \left(1 + \frac{1}{2}Z\right) - \frac{2}{3}\beta\pi\rho \int r^2 \left[h(r) \left(r \frac{\partial v_{DH}(r)}{\partial r} \right) - g(r) 3\rho \frac{\partial v_{DH}(r)}{\partial \rho} \right].$$

These terms reduce Z_{vir} further

No free lunch theorems

- Effective-potentials for depletion systems

- Representability and potentials
 - A) Two and three-body potentials
 - B) Isotropic model for water

- Dynamics?

Case study 2: Representability problems for pair potentials

Coarse-graining throws away information (**No free lunch**)

Question: when does this matter?

Intuitions for effective potentials $v^{\text{eff}}(r)$:

representability problems: You can't simultaneously represent all the properties of the underlying system **at one state point** with **one** coarse-grained potential. AAL, J. Phys.: Condens. Matter **14**, 9187 (2002)

Contrast with **transferability**: different $v^{\text{eff}}(r)$ at **different state points**

Exhibit A:

Original system 

Coarse-grained system

2 and 3 body potentials

pair potential $v^{eff}(r)$ only

$$H = K + \sum_{i < j} w^{(2)}(r_{ij}) + \sum_{i < j < k} w^{(3)}(\mathbf{r}_{ij}, \mathbf{r}_{ik}, \mathbf{r}_{ki}), \quad \longrightarrow \quad v^{eff}(r) = w^{(2)}(r) + \delta v(r)$$

Energy route:

$$U(N, V, T) = \frac{1}{2} \rho^2 \int d\mathbf{r}_1 d\mathbf{r}_2 g(r_{12}) w^{(2)}(r_{12}) + \frac{1}{6} \rho^3 \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 g^{(3)}(\mathbf{r}_{12}, \mathbf{r}_{23}, \mathbf{r}_{31}) \times w^{(3)}(\mathbf{r}_{12}, \mathbf{r}_{23}, \mathbf{r}_{31}),$$

$$U(N, V, T) = \frac{1}{2} \rho^2 \int d\mathbf{r}_1 d\mathbf{r}_2 g(r_{12}) v_U^{eff}(r_{12}).$$

$$\delta v_U(r) = \frac{1}{3} \rho \int d\mathbf{r}_3 \frac{g^{(3)}(\mathbf{r}_{12}, \mathbf{r}_{23}, \mathbf{r}_{31})}{g(r_{12})} w^{(3)}(\mathbf{r}_{12}, \mathbf{r}_{23}, \mathbf{r}_{31})$$

$g(r)$ from full Hamiltonian

Structure route:

$$g(r) \leftrightarrow v^{eff}(r)$$

Same $g(r)$ in coarse-grained system

$$v^{eff}(r) = w^{(2)}(r) + \delta v_g(r) \text{ (unique)}$$

Comparison

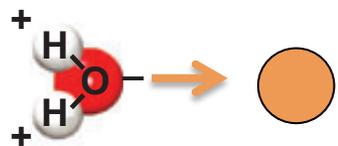
$$\frac{\delta v_U(r)}{\delta v_g(r)} = \frac{1}{3} + O[(w^{(3)})^2; \rho^2].$$

← (low density weak $w^{(3)}$ limit)

Representability problems: one potential can't simultaneously represent multiple properties of the system AAL, J. Phys.: Condens. Matter 14, 9187 (2002); Faraday

Exhibit
B:

H₂O with spherically symmetric pair potentials?



$v_g(r) \leftrightarrow g_{OO}(r)$ - “bottom up” from Tip4pEw water

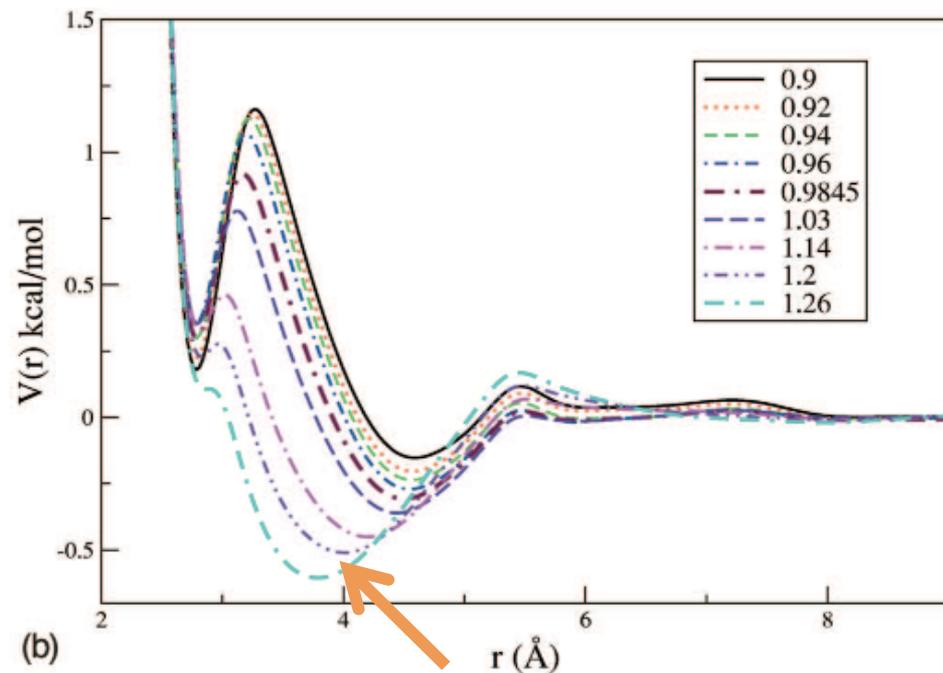
The potential is **unique**

R.L. Henderson Phys. Lett. **49A**, 197 (1974)

J.T. Chayes and L. Chayes, J. Stat. Phys. **36**, 471 (1984)

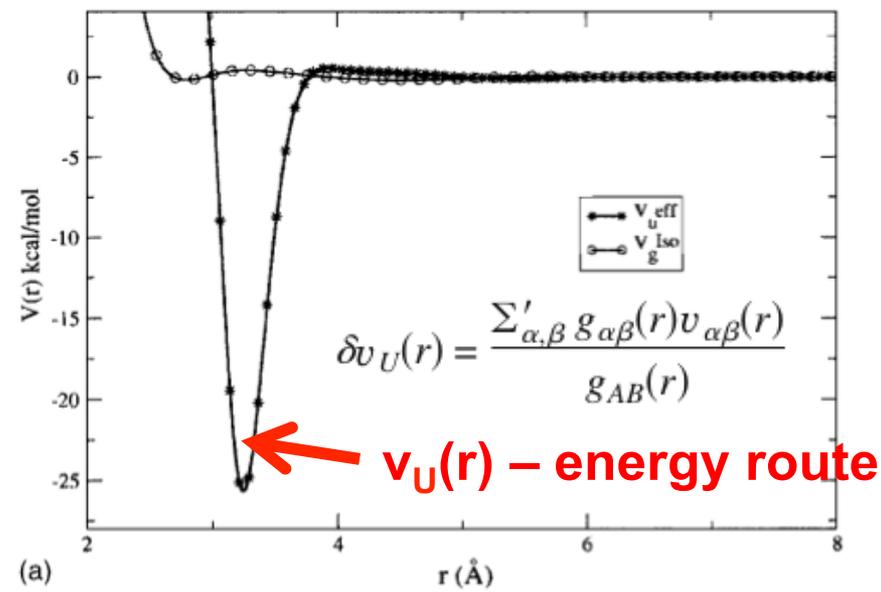
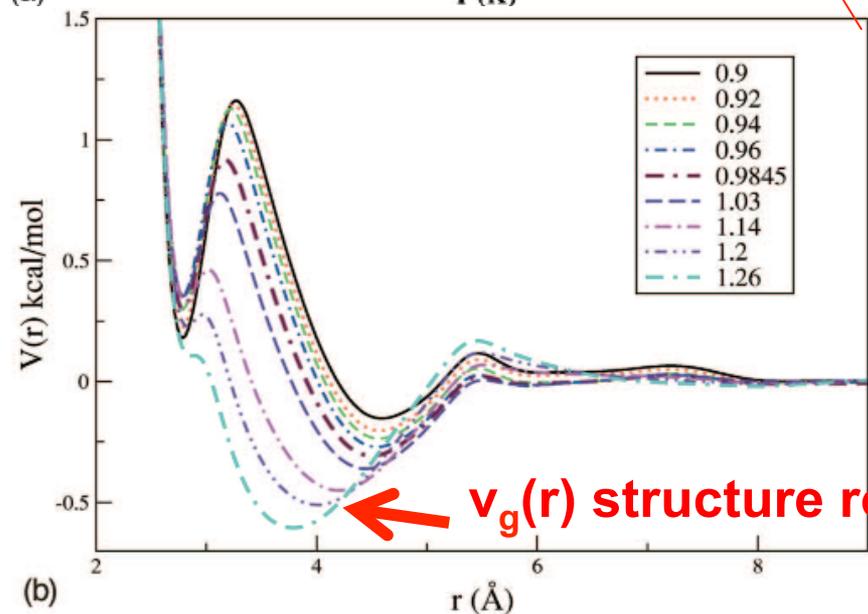
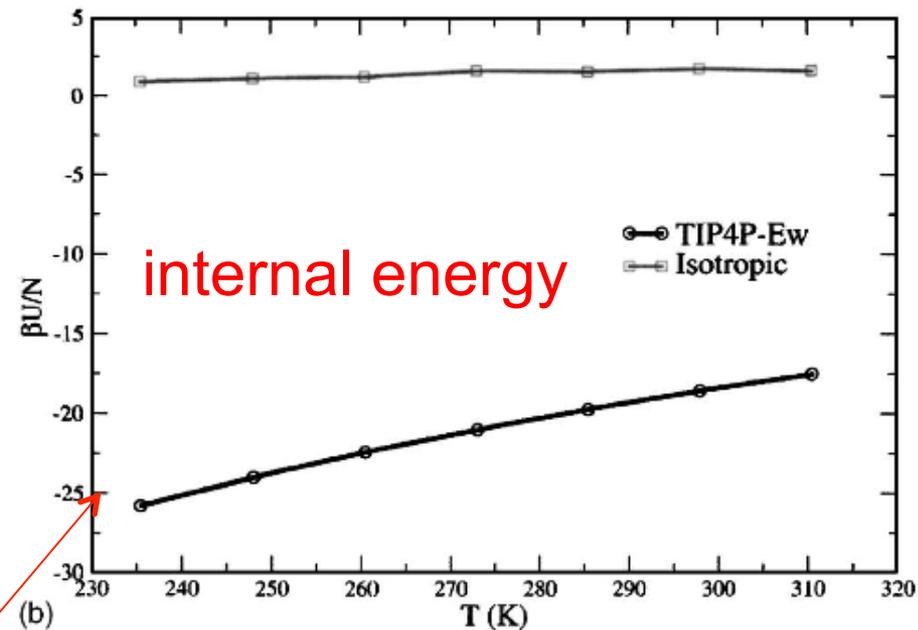
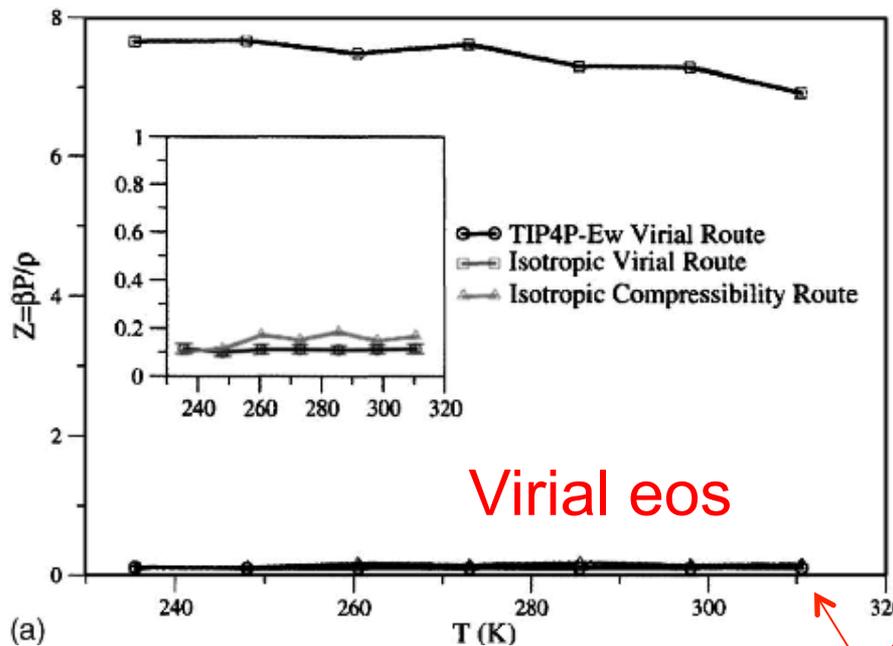
Thermodynamics through compressibility route

$$\rho k_B T \chi_T = 1 + 4\pi\rho \int r^2 [g_{AB}(r) - 1] dr,$$



$v_g(r)$ changes with density,
=> **Transferability problems** for $g(r)$

Representability problems are severe



Representability problems are severe

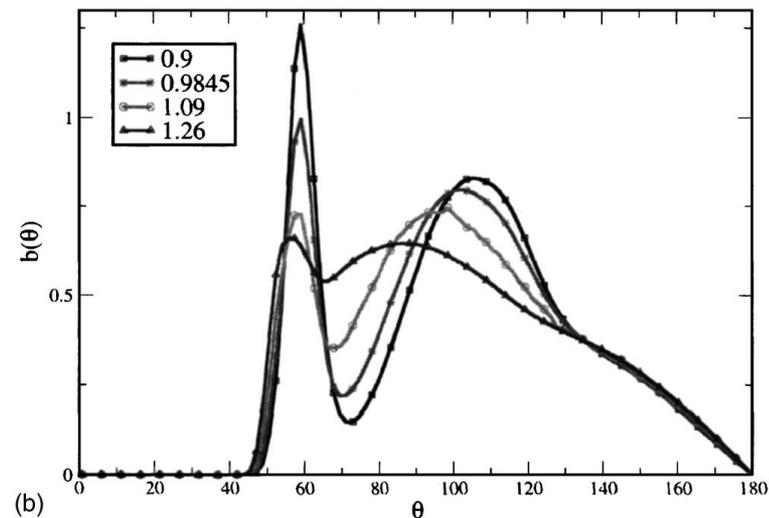
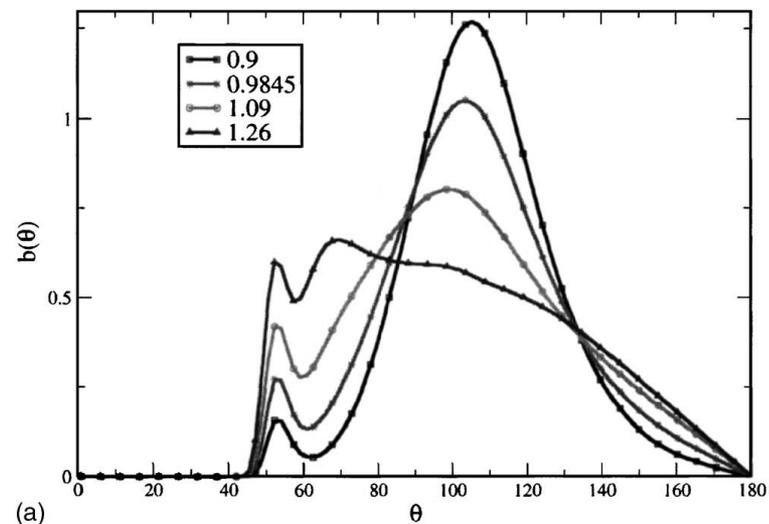
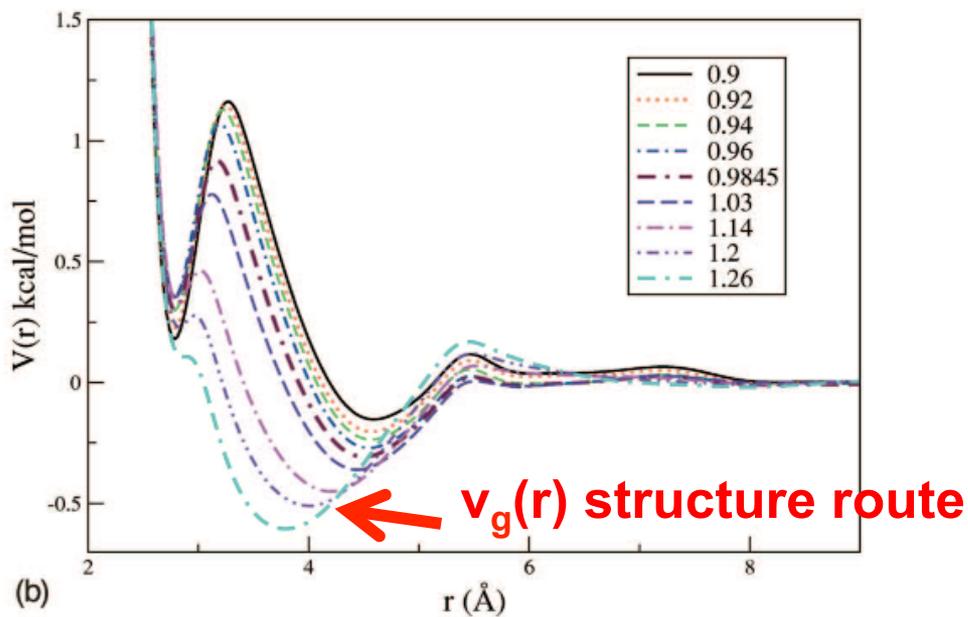
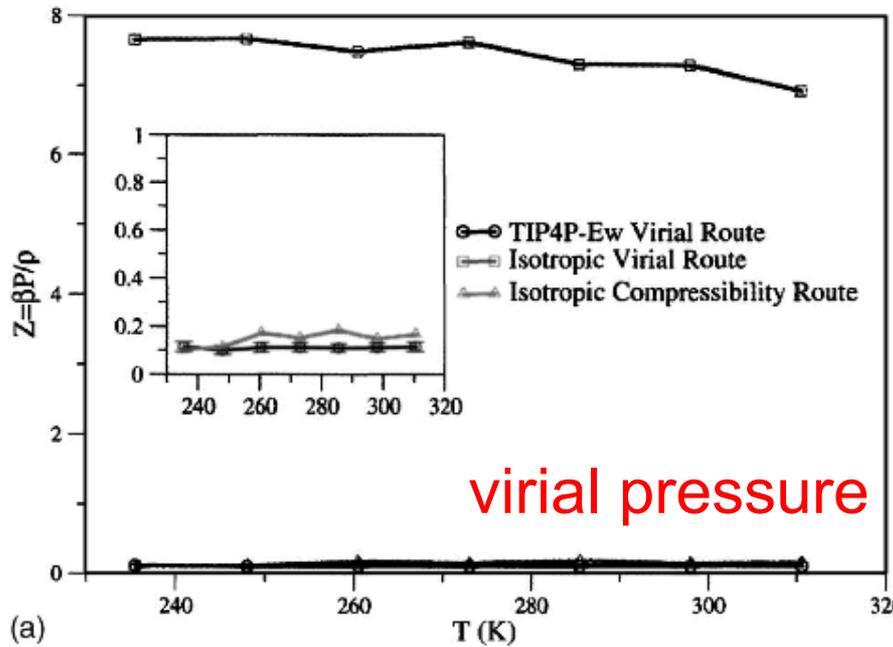


FIG. 5. Bond angle distributions at $T=235.5$ K as a function of density for (a) TIP4P-Ew and (b) isotropic potentials. R_C was chosen as 3.4 \AA .

No free-lunch theorems for representability



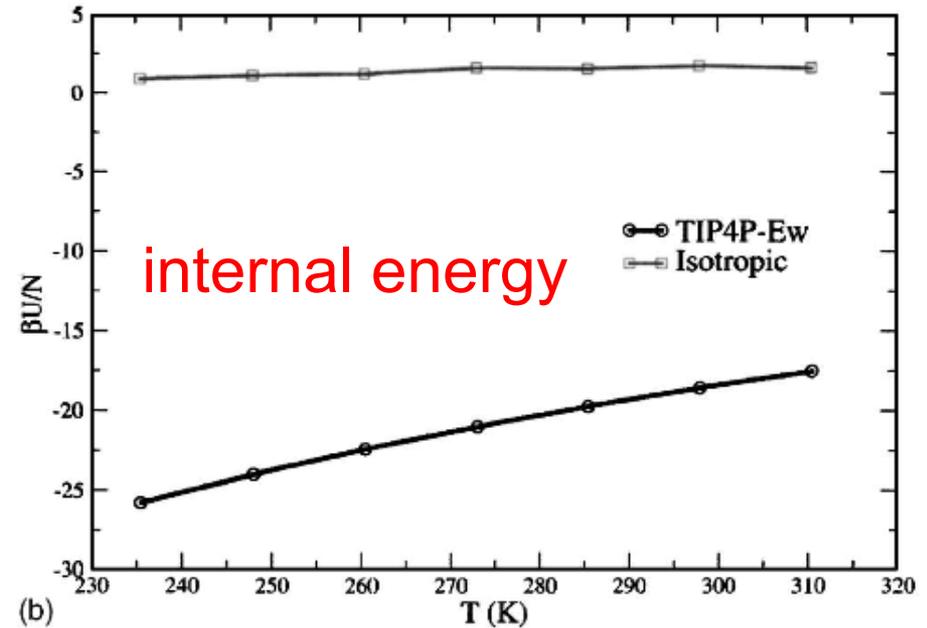
(a) Compromise by fitting to multiple properties?

$$U(N, V, T) = 2\pi \frac{N^2}{V} \int_0^\infty g_{AB}(r) v_U^{\text{eff}}(r) r^2 dr,$$

$$P = \rho k_B T - \frac{2\pi\rho^2}{3} \int r^3 dr g(r) \frac{dv(r)}{dr}$$

$$\rho k_B T \chi_T = 1 + 4\pi\rho \int r^2 [g_{AB}(r) - 1] dr,$$

See also Kremer/Mainz army on H₂O



Representability:

A single potential can't represent all properties simultaneously – All potentials are at best compromises

Fitting too tightly to one property (e.g. structure) may increase errors in another (e.g. thermodynamics)

Corrections to virial equation from density dependence?

$$Q(N, V, T) = \frac{\Lambda^{-3N}}{N!} \int d\mathbf{r}^N \exp \left\{ -\beta \sum_{i < j} v(r_{ij}; \rho) \right\}$$

Λ is the usual thermal de Broglie wavelength. The volume derivative in

$$\beta P = \left(\frac{\partial \log Q(N, V, T)}{\partial V} \right)_{N, T}$$

$$Z_{vir}^{\rho} = \frac{\beta P}{\rho} = 1 - \frac{2}{3} \beta \pi \rho \int_0^{\infty} r^2 \left\{ r \frac{\partial v(r; \rho)}{\partial r} - 3\rho \frac{\partial v(r; \rho)}{\partial \rho} \right\} g(r) dr$$

—Coarse-grained potentials do not generate a Hamiltonian system

“We record our opinion that the use of density dependent effective pair potentials can be misleading unless it is recognized that these are mathematical constructs to be used in specified equations rather than physical quantities”

J.A. Barker, D. Henderson and W.R. Smith Mol. Phys. 17, 579 (1969) – see also John Rowlinson and other forefathers

New equations with your coarse-graining scheme?

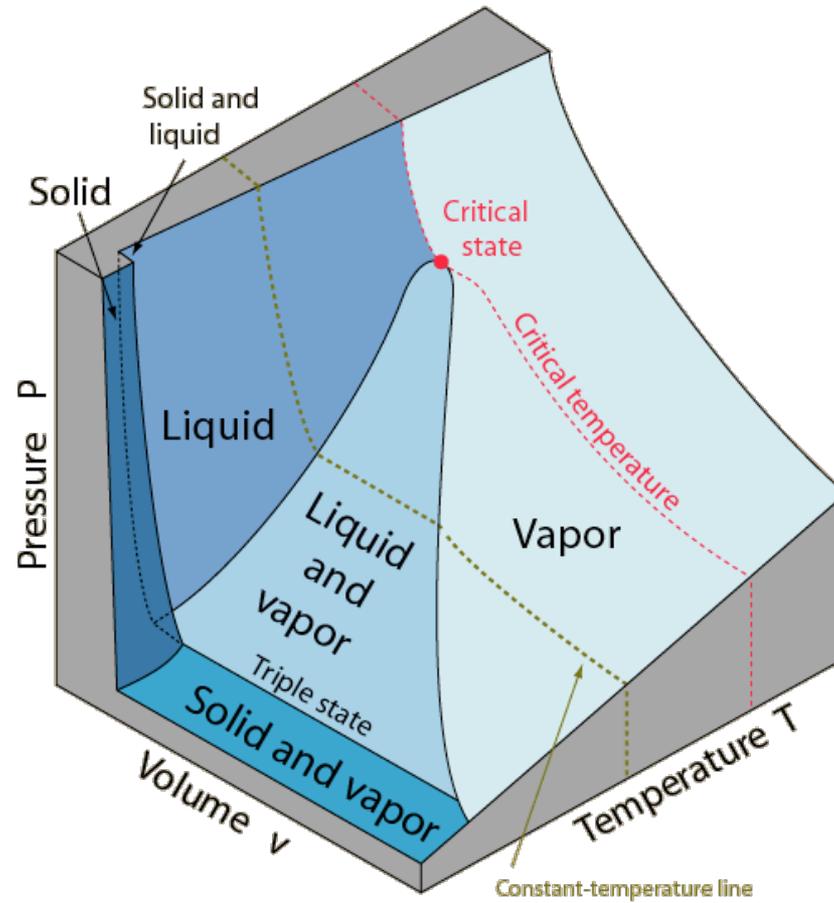
$$U(N, V, T) = 2\pi \frac{N^2}{V} \sum_{\alpha, \beta} \int v_{\alpha\beta}(r) g_{\alpha\beta}(r) r^2 dr, \quad \delta v_U(r) = \frac{\sum'_{\alpha, \beta} g_{\alpha\beta}(r) v_{\alpha\beta}(r)}{g_{AB}(r)},$$

$$U(N, V, T) = 2\pi \frac{N^2}{V} \int_0^\infty g_{AB}(r) v_U^{\text{eff}}(r) r^2 dr, \quad v_U^{\text{eff}} = v_{AB}(r) + \delta v_U(r),$$

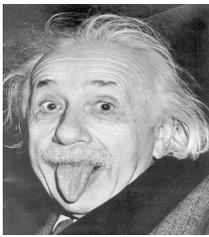
Representing properties differently in your CG scheme, e.g. take into account :

- 1) Entropy differences etc...
- 2) One-body terms?
- 3) Phase-diagrams

CG moves phase-boundaries ...



Thoughts on representability



1. Effective potentials are not Hamiltonian potentials
 - a) Beware of using equations that treat them as such
 - E.g. Corrections to virial equation from density dependence – mostly wrong
 - But better thermodynamic equations might be derivable from coarse-graining scheme
 - b) To map to something analogical to a Hamiltonian, care must be taken with ensembles (canonical ensemble may not be ideal)
 - c) Be careful about one-body terms & missing degrees of freedom
 - d) “Energetic” and “Entropic” character can change

2. Effective potentials are compromises – there is no free lunch
 - a) It’s probably better not to fit too tightly to just one property (e.g. structure), but rather multiple properties at once.
 - b) STOP & THINK: no substitute for physical insight, e.g. symmetries etc...
 - a) Effective emergent models
 - b) Nearby phase boundaries
 - c) “Automated” coarse-graining techniques and black box methods should be treated with suspicion.

1. Representability and transferability are probably related. (State dependence)

Further thoughts on representability?

- “Emergent models” v.s. models for fine-graining
 - Different approaches?
- Fitting to thermodynamic quantities?
- Testing representability issues

There is no such thing as a free lunch!
but cost can be brought down by physical insight

No free lunch theorems

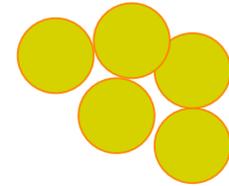
- Effective-potentials for depletion systems
- Representability and potentials
 - A) Two and three-body potentials
 - B) Isotropic model for water

• Dynamics?

More difficulties with dynamics ...



$$D \sim R^{-1}$$



$$D \sim R_g^{-1}$$

Diffusion coefficient of assemblies

Telescoping down?

Brownian Dynamics $D \sim N^{-1}$

Flattened energy land-scapes

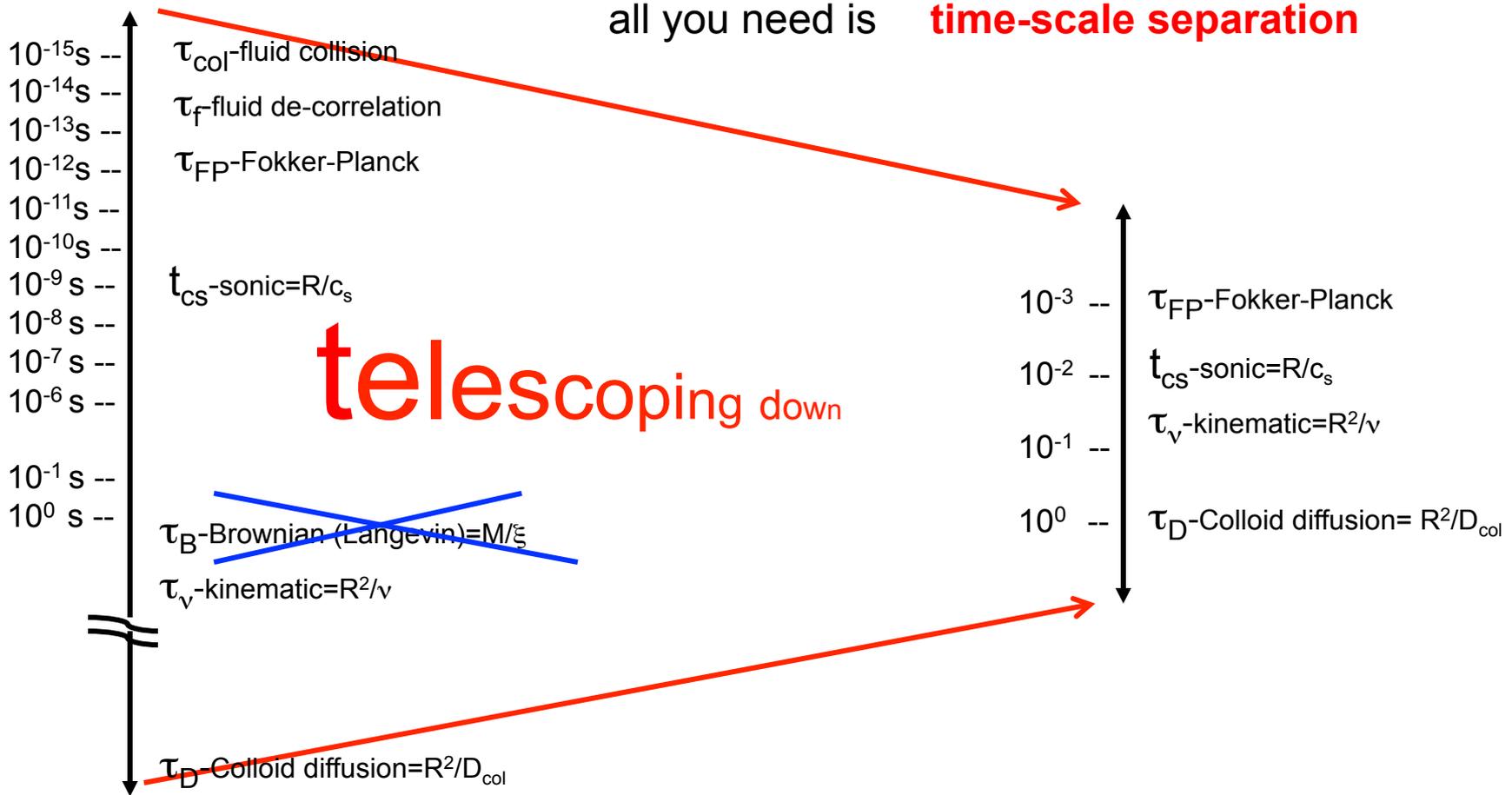


DYNAMICS? <http://arxiv.org/abs/1001.1166>;

J.T. Padding and AAL *Phys. Rev. E* **74**, 031402 (2006)

Coarse-graining: Telescoping time-scales

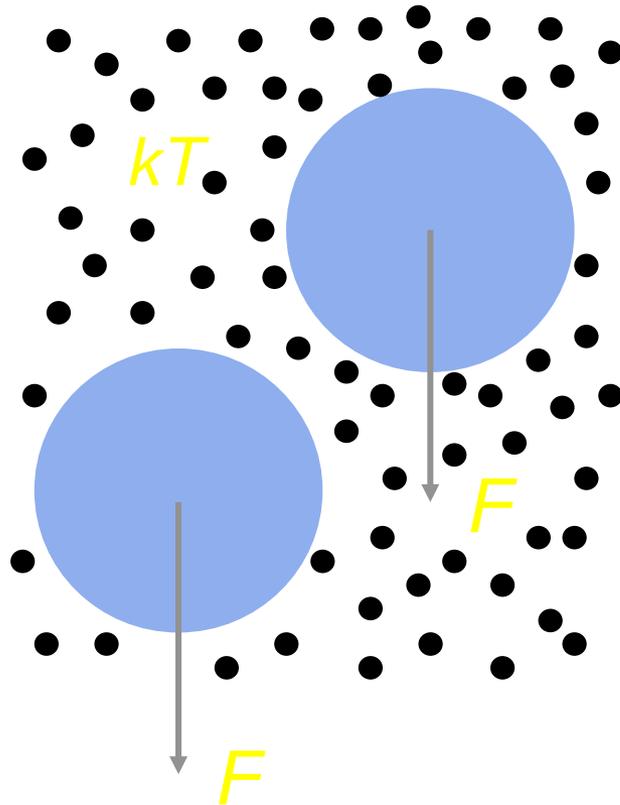
Colloid of diameter l μm



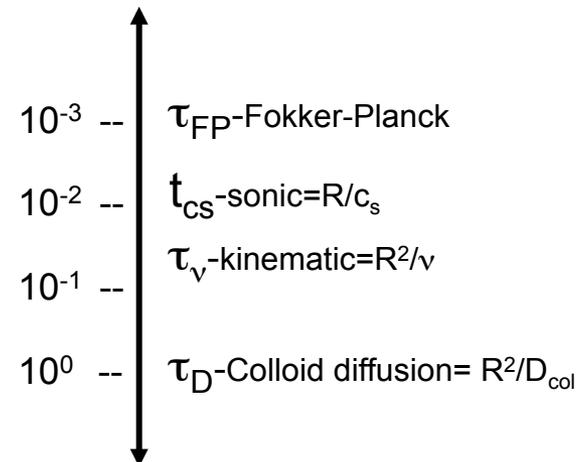
τ_{B} has no independent physical significance for colloids:
simple Langevin equation

beware of

How do I map to physical time/length scales?



1. mapping to diffusion time
2. mapping to kinematic time
3. mapping potentials



telescoping down

Coarse-graining changes energy landscapes

Extremely unlikely that all barriers are lowered to give the same time-scales – so multiple time-scales mappings in one simulation ...

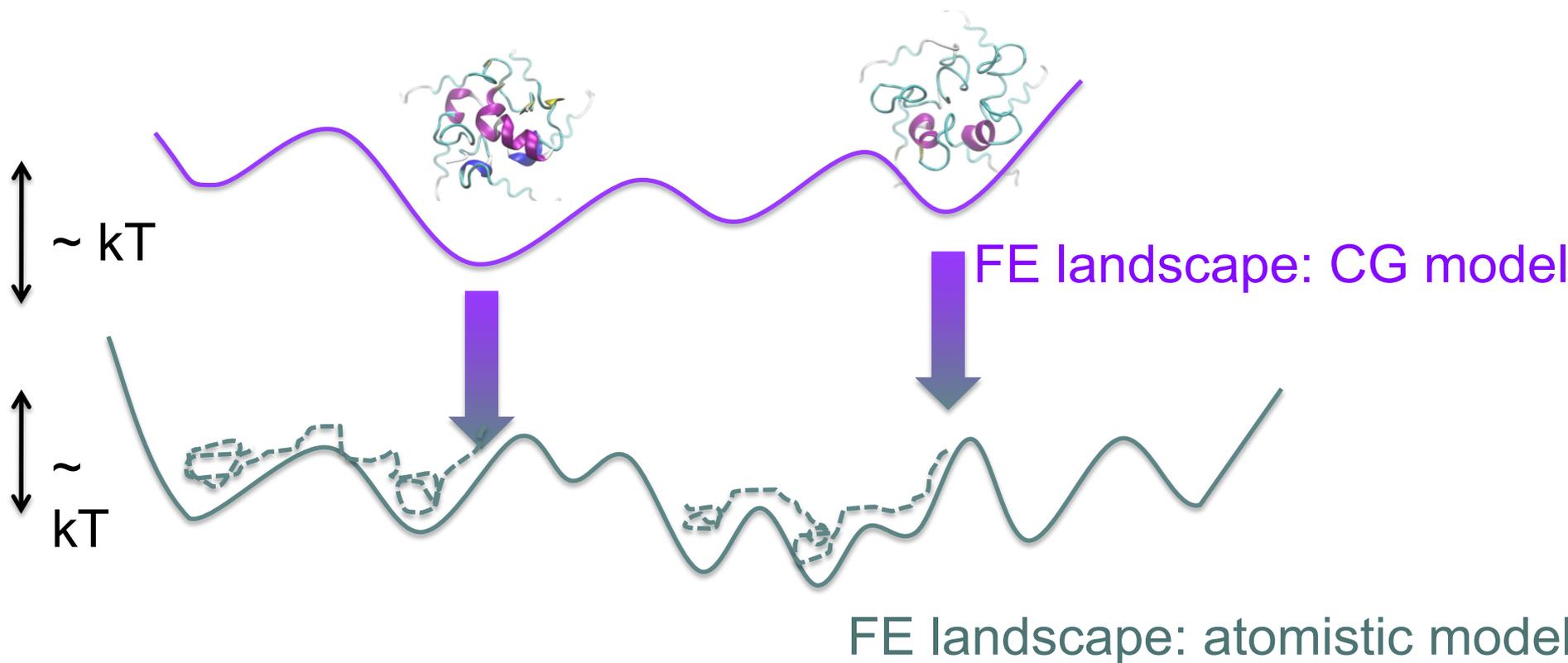


Figure from C. Peter