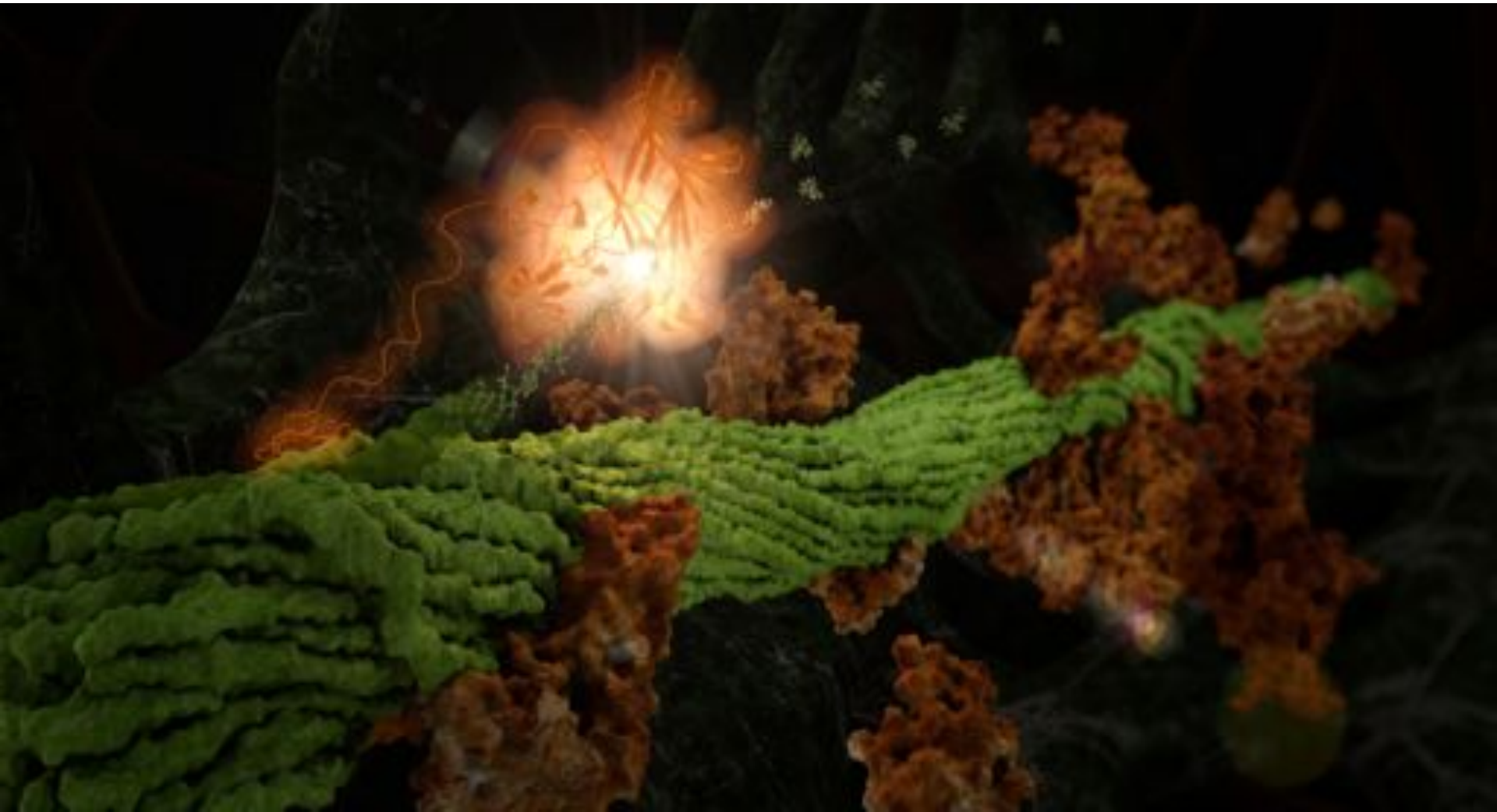




Supercomputing in Biology and Medicine



Jeremy C. Smith

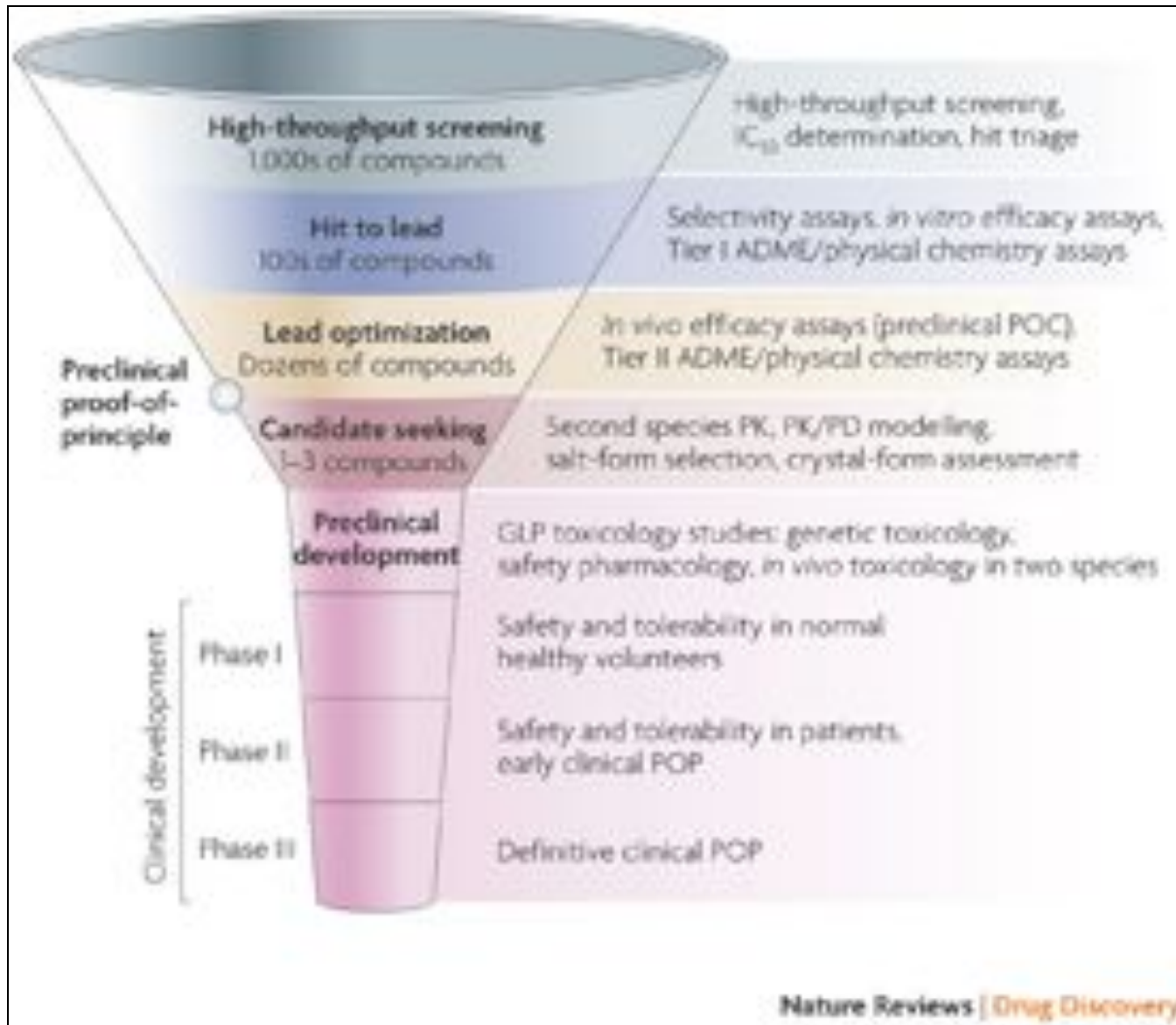


Drug Discovery

Large-Scale Molecular Dynamics Simulation

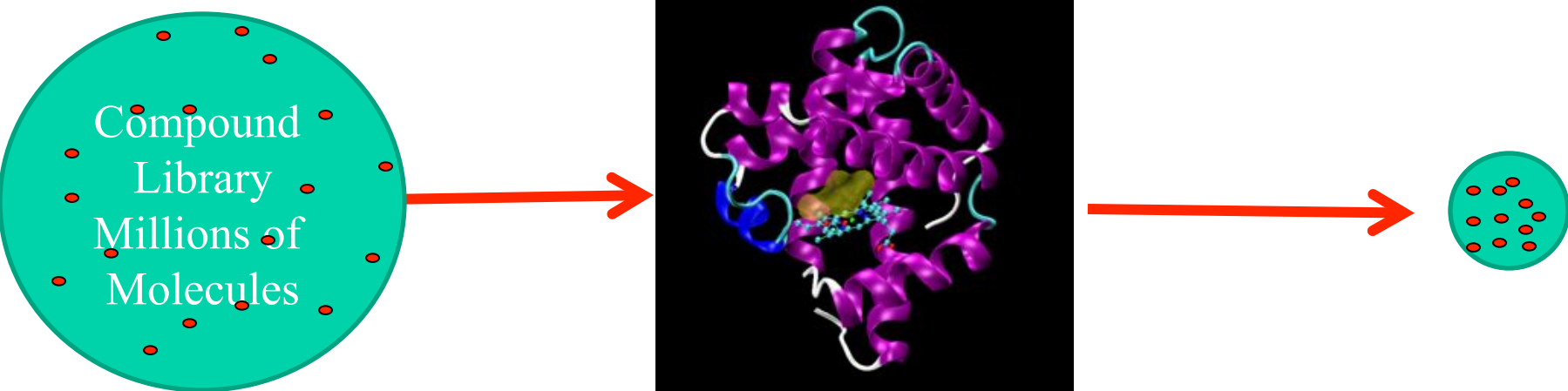


Drug Discovery





Structure-based Computational Drug Discovery



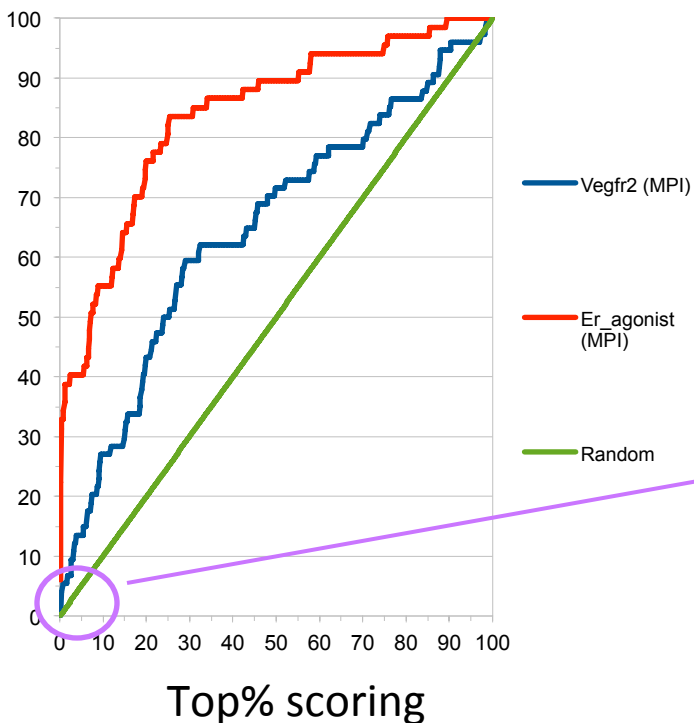
Enrich a collection of compounds with molecules most likely to bind to the drug target(s)

→ much faster, much more affordable hit & lead discovery

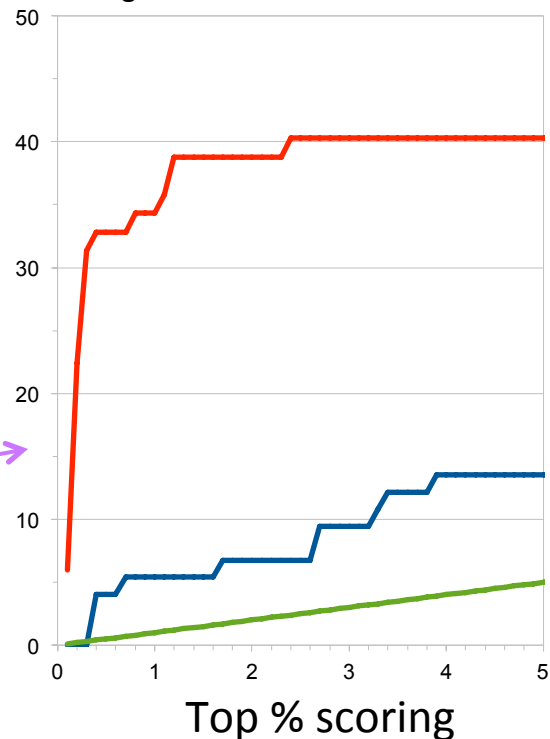
Hit Enrichment

TEST: 98,163 compound database, 2 drug targets:
ER: 67 known ligands ; VEGFr2: 74 known ligands

% of known ligands

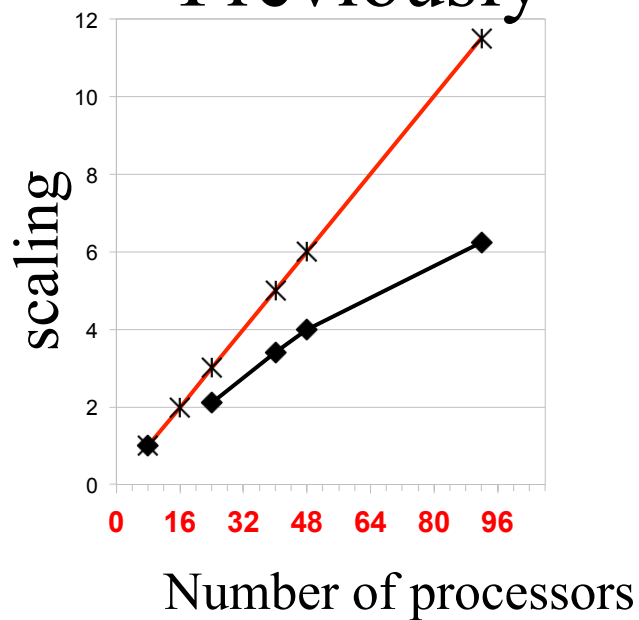


% of known ligands

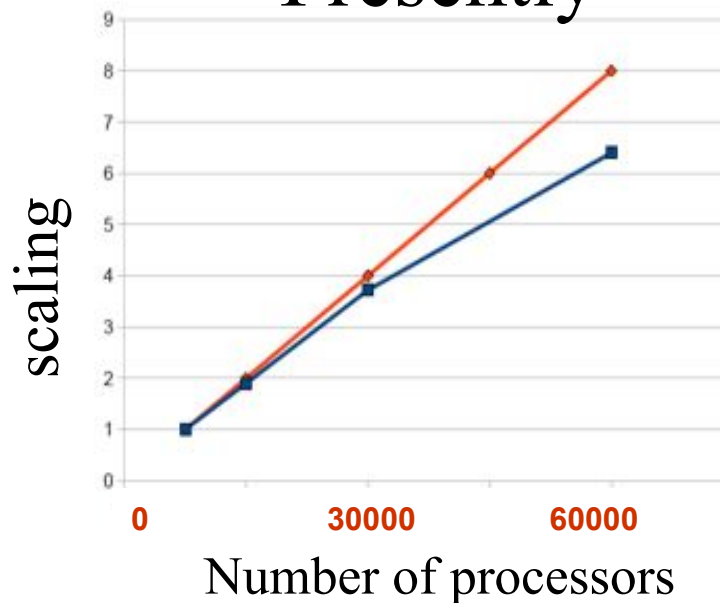


ER: Top 0.3 % of total database (295 candidates, 31 hits): Enrichment = ~104
VEGFr2: Top 0.4 % of total database (393 candidates, 4 hits): Enrichment = 10

Previously



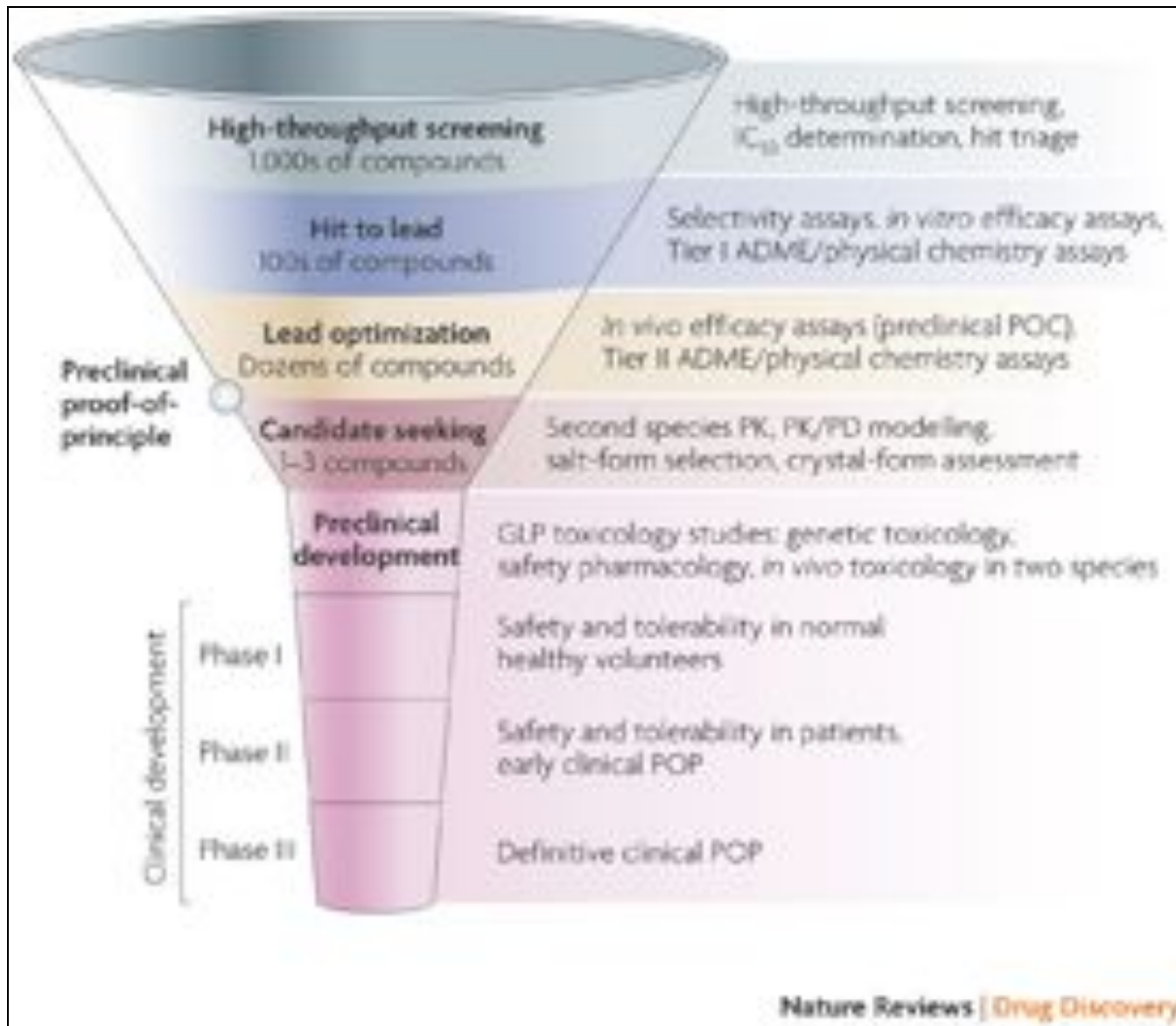
Presently



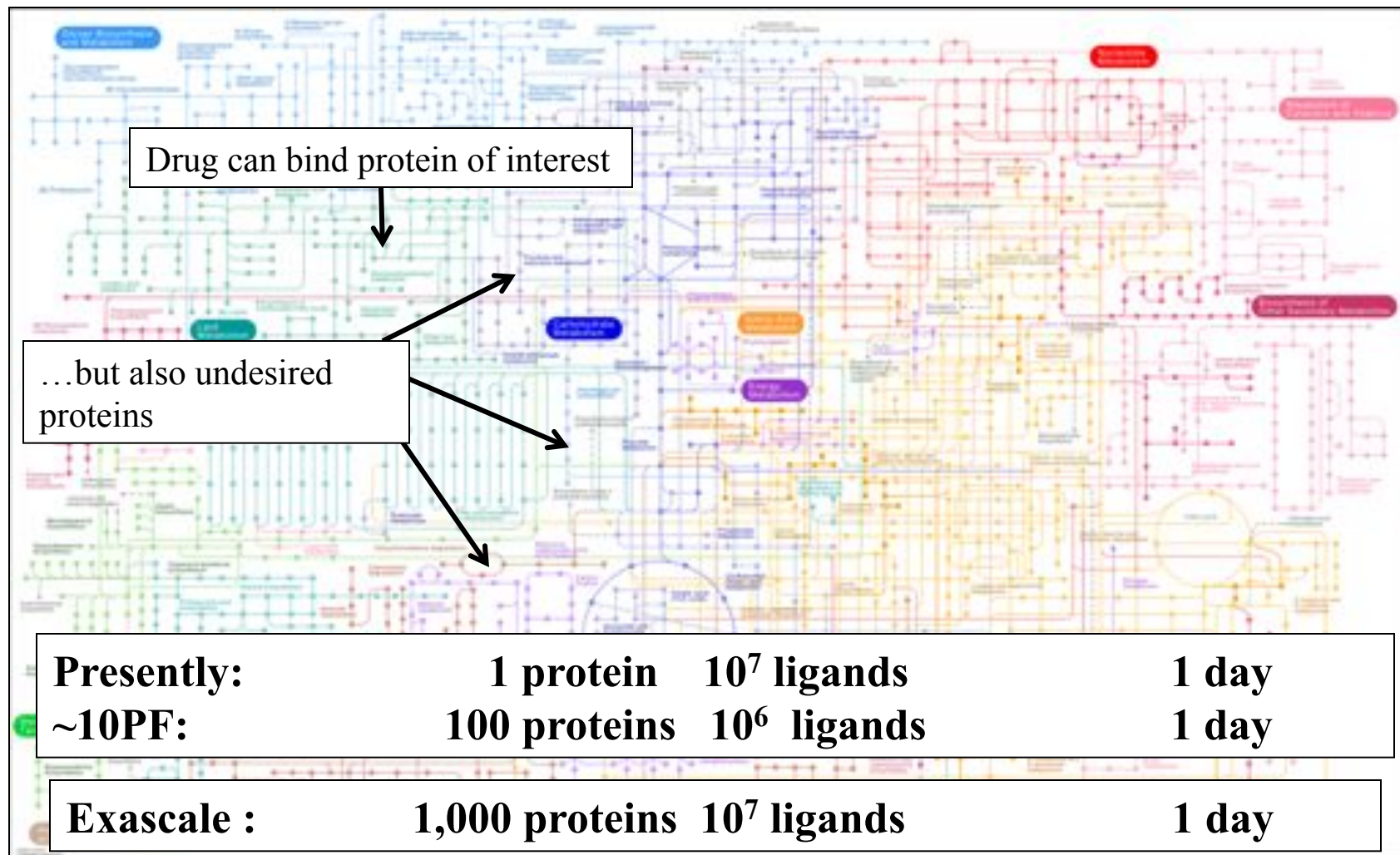
Previously :	1 protein	10^4 ligands	~days
Presently :	1 protein	10^7 ligands	1 day



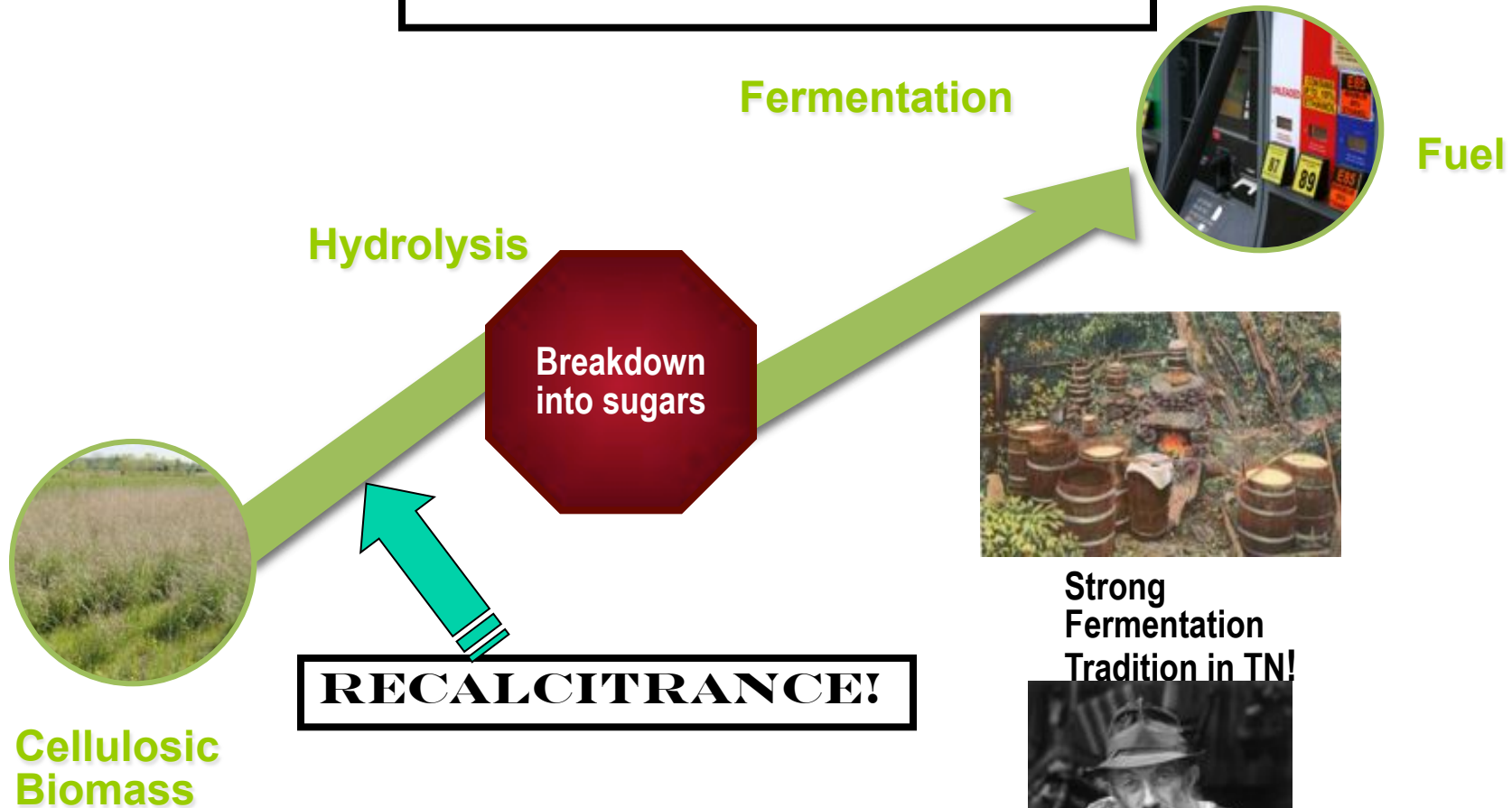
Drug Discovery



10-100 Petaflops: Drugome Exploration, Drug Toxicity, Repurposing



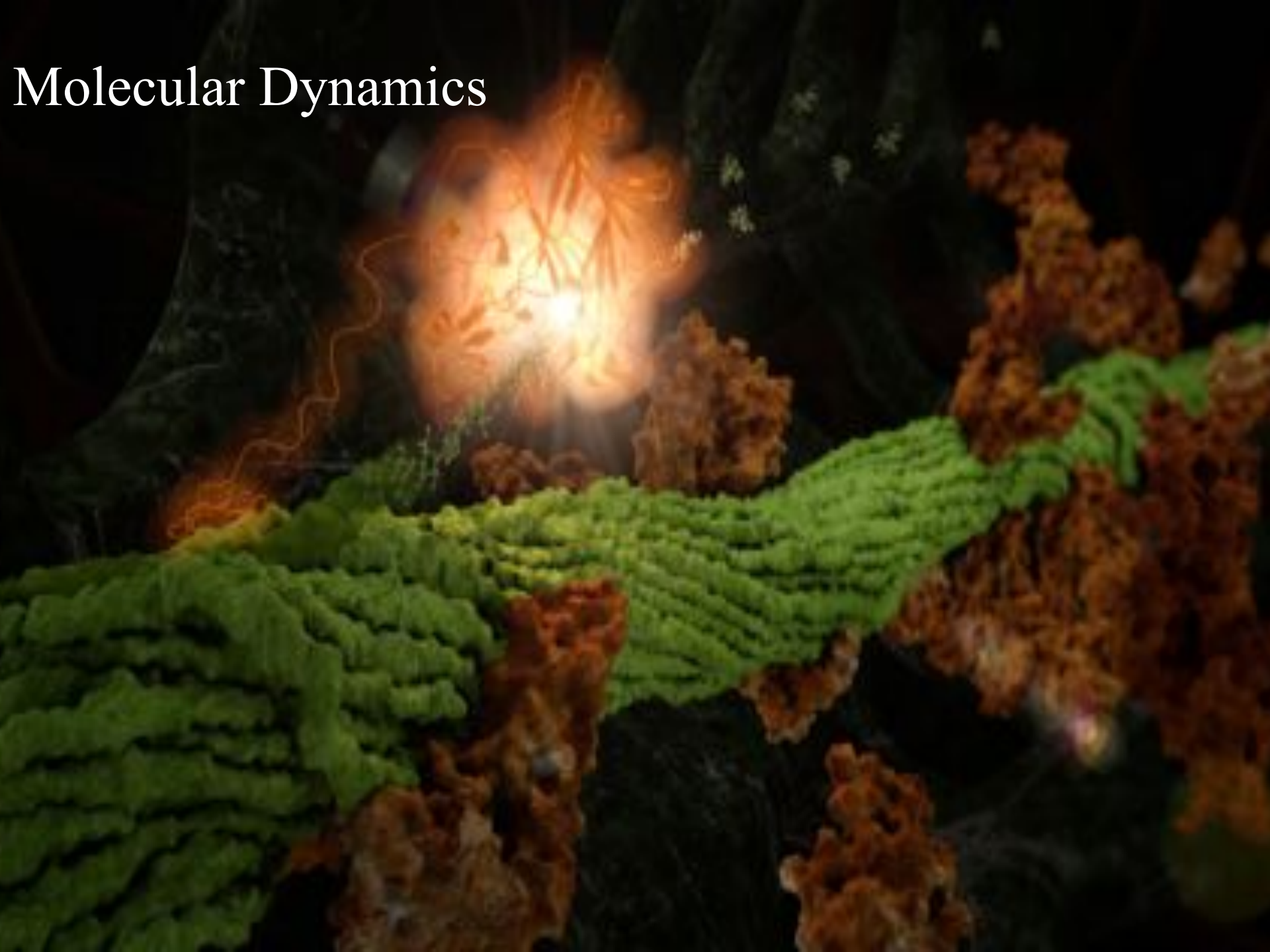
Cellulosic Ethanol



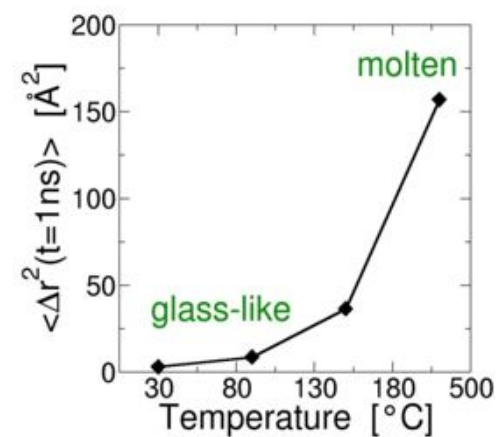
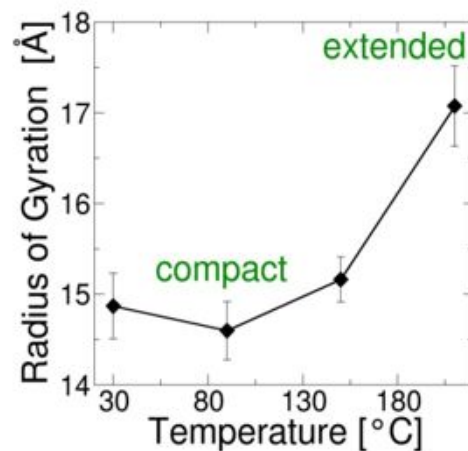
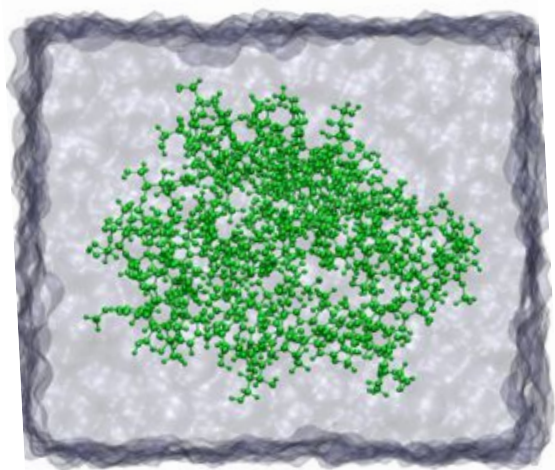
Strong
Fermentation
Tradition in TN!



Molecular Dynamics

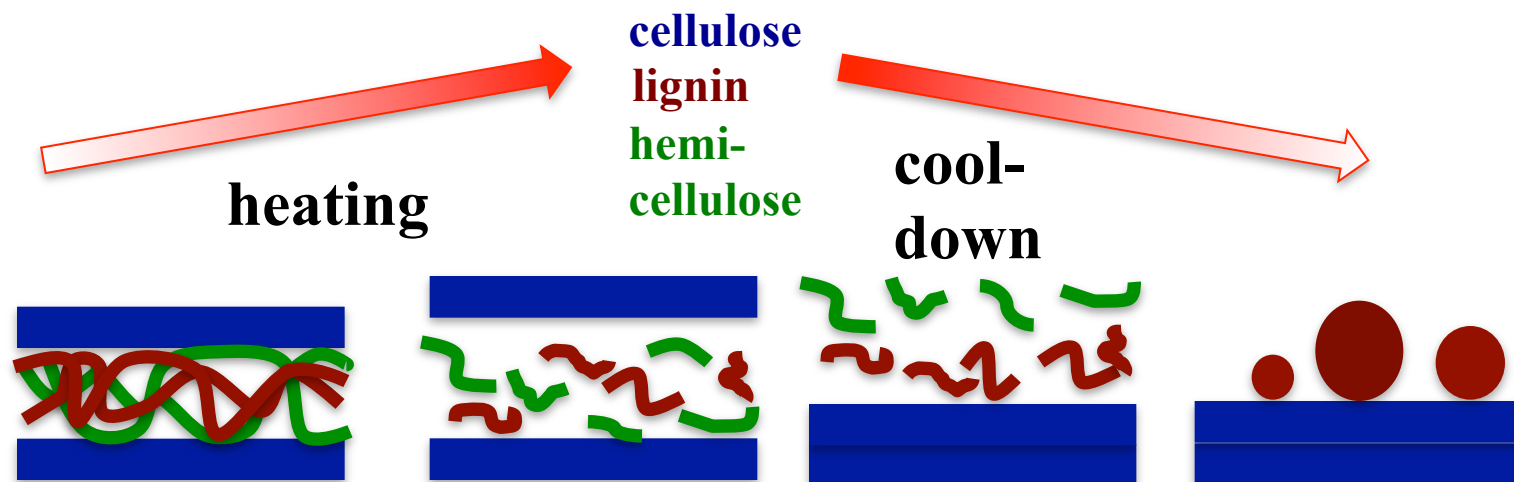


DOE INCITE: MD Simulation of Softwood Lignin

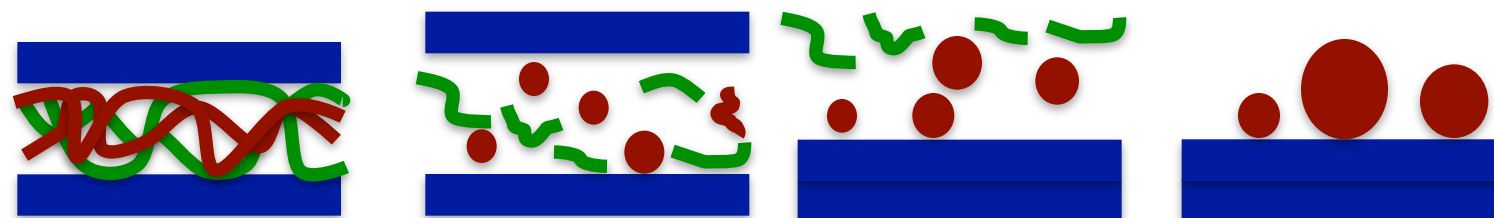


Petridis et al JACS 133 20277 (2011)

Lignin Aggregation During Heating Phase of Dilute-acid Pretreatment



aggregation occurs during cool-down

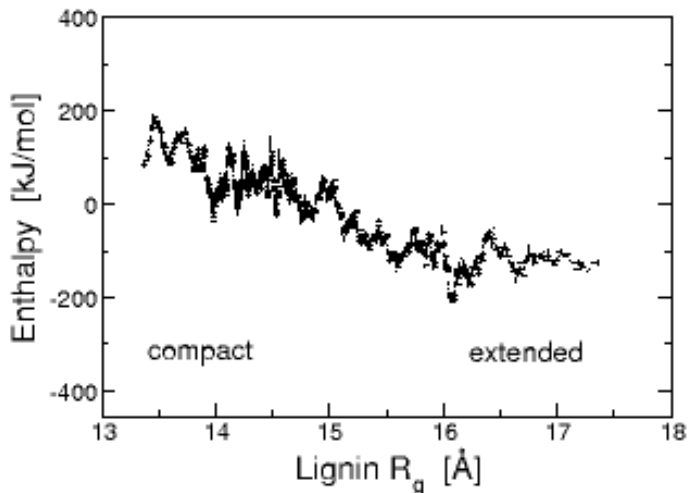


aggregation occurs during heating

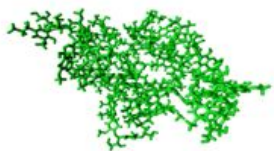
Why does Lignin Collapse at Room Temperature?



- **Enthalpy**
- $\Delta H \approx +200$ kJ/mol **Unfavorable**

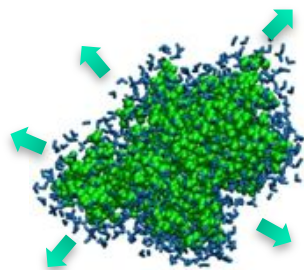
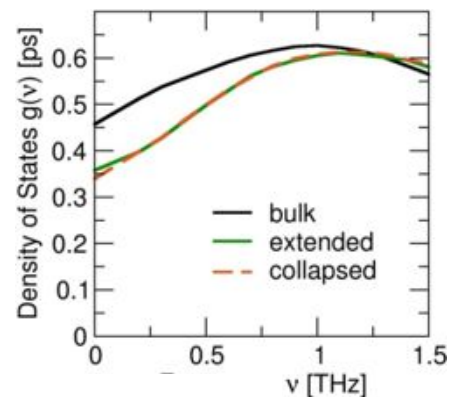


- **Lignin configurational entropy**
- $-T\Delta S_{\text{conf}} \approx +10$ kJ/mol **Unfavorable**



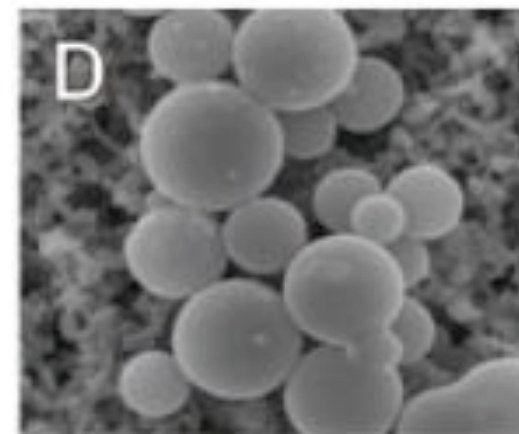
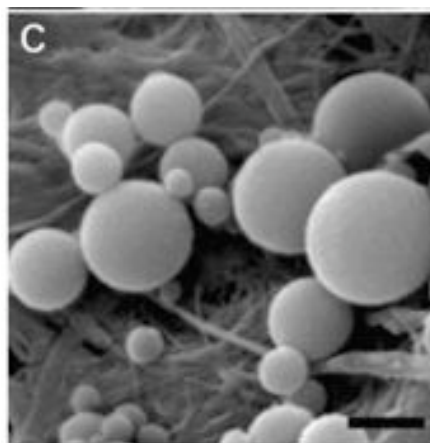
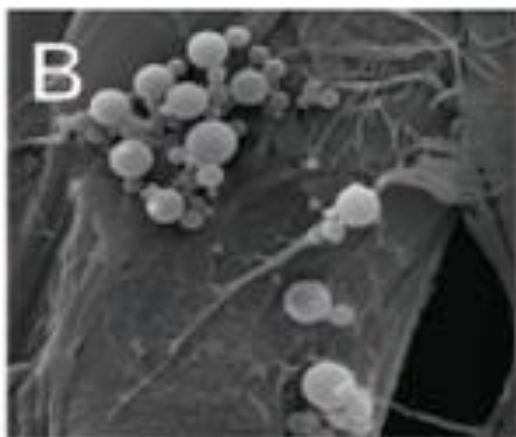
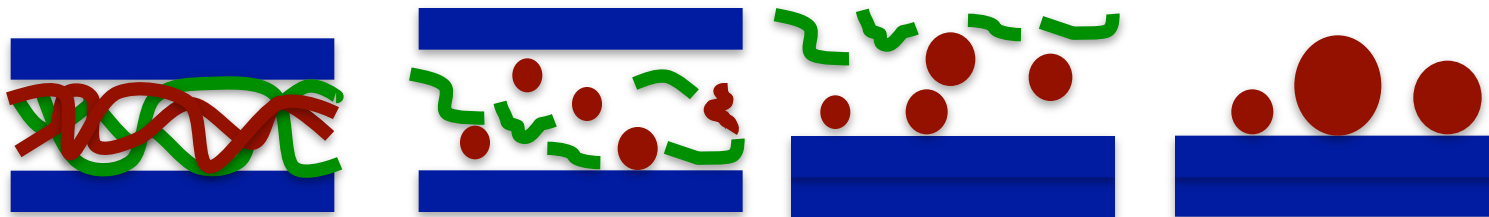
Petridis et al JACS 133 20277 (2011)

- Hydration water translational & rotational entropy
- $-T\Delta S_{\text{t+r}} \approx -100$ kJ/mol **Favorable**

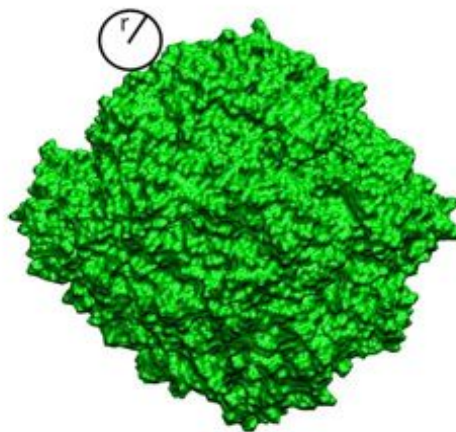


- Hydration water compressibility
- $-T\Delta S_{\text{fluc}} \approx -300$ kJ/mol **Favorable**

**Collapse Driven by Removal of
Entropically Unfavorable Water
Molecules from Lignin Surface to Bulk**



Are Lignin Aggregates Spheres?

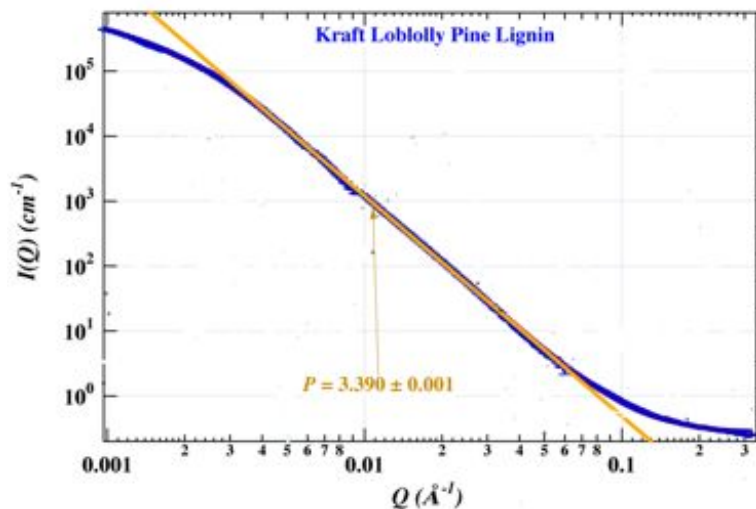


Molecular
Dynamics

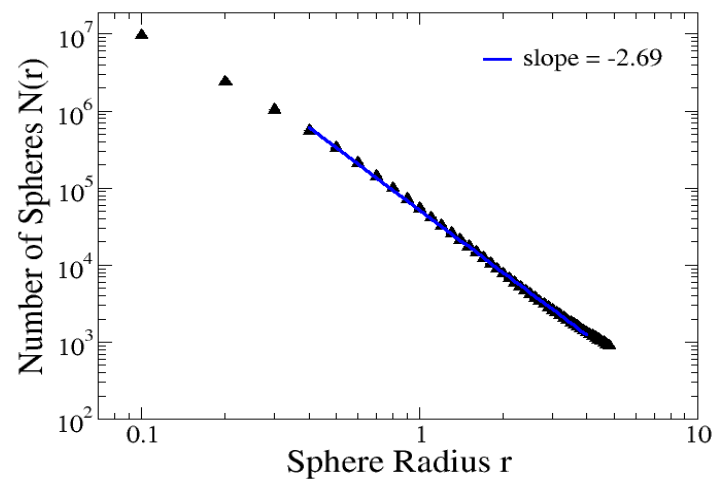
$$N(r) = r^{-d_s}$$

Small-Angle
Neutron Scattering

$$S(q) \propto Q^{d_s - 6}$$

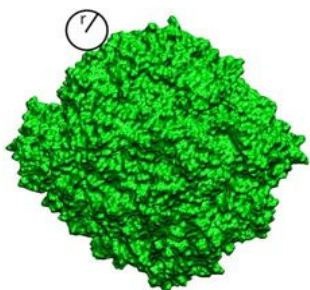


$$d_s = 2.62 \pm 0.02$$

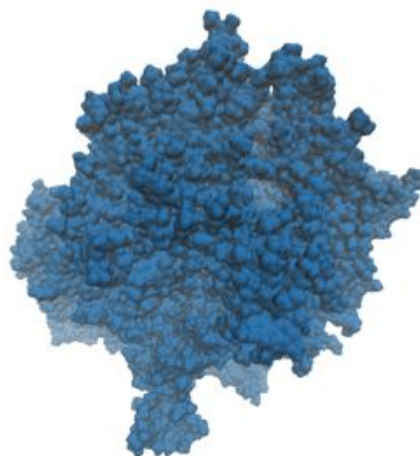


$$d_s = 2.65 \pm 0.01$$

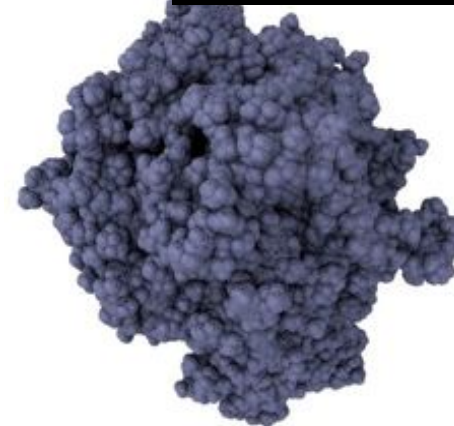
Surface Fractals over Three Orders of Magnitude



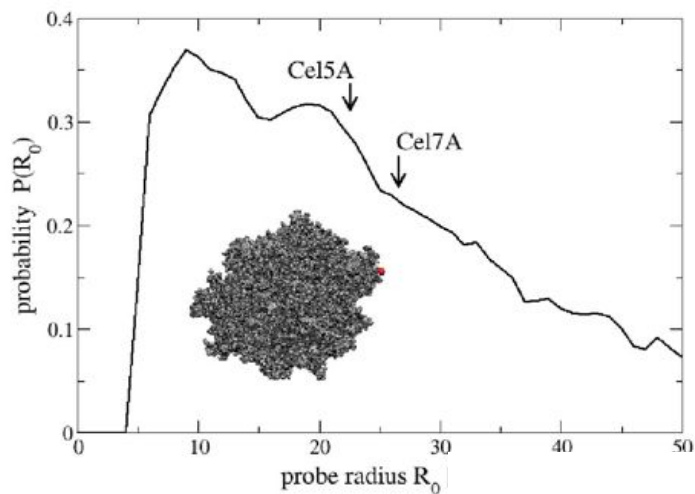
$$R_g = 4.2 \text{ \AA}$$



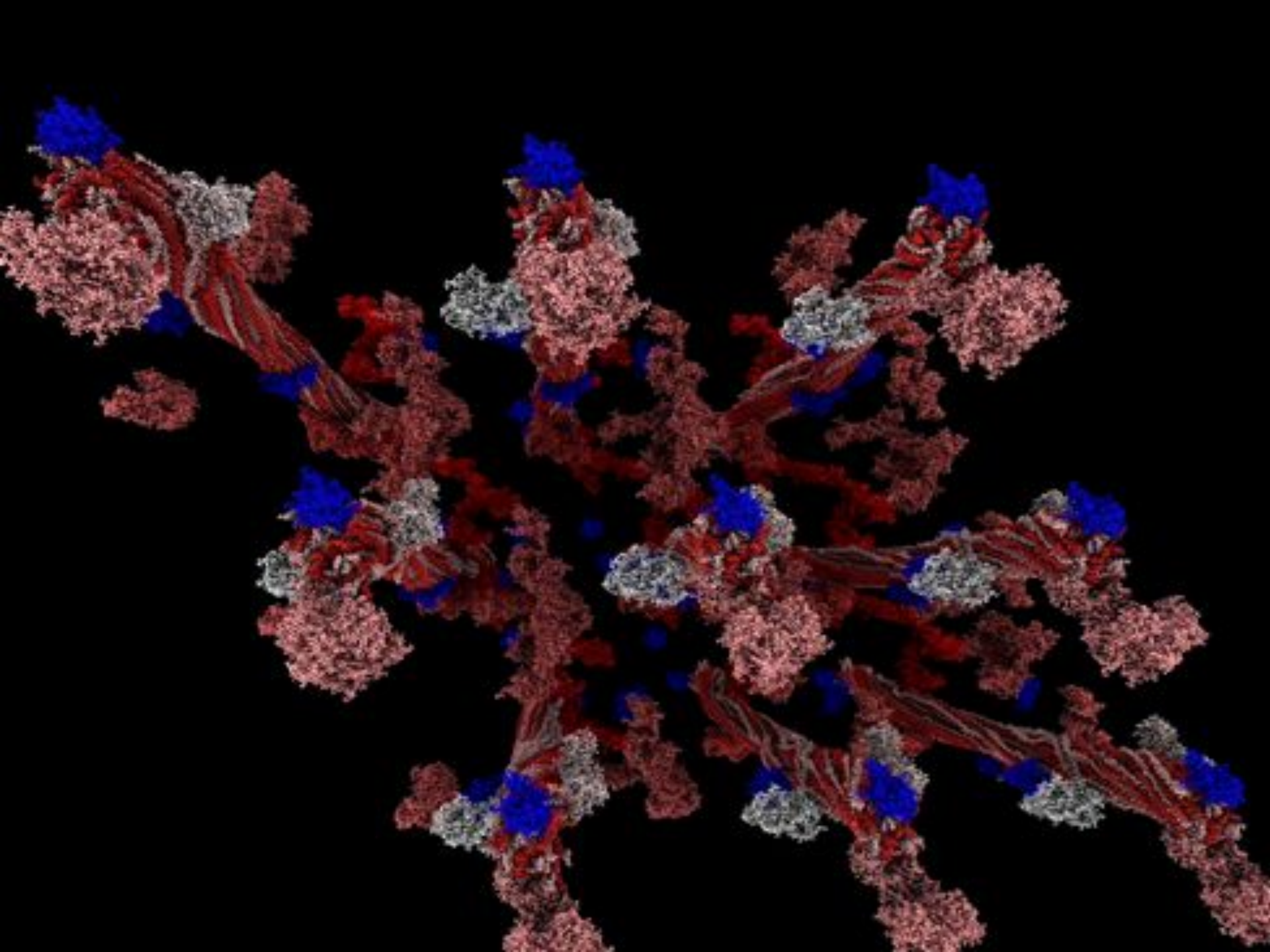
$$R_g = 42 \text{ \AA}$$



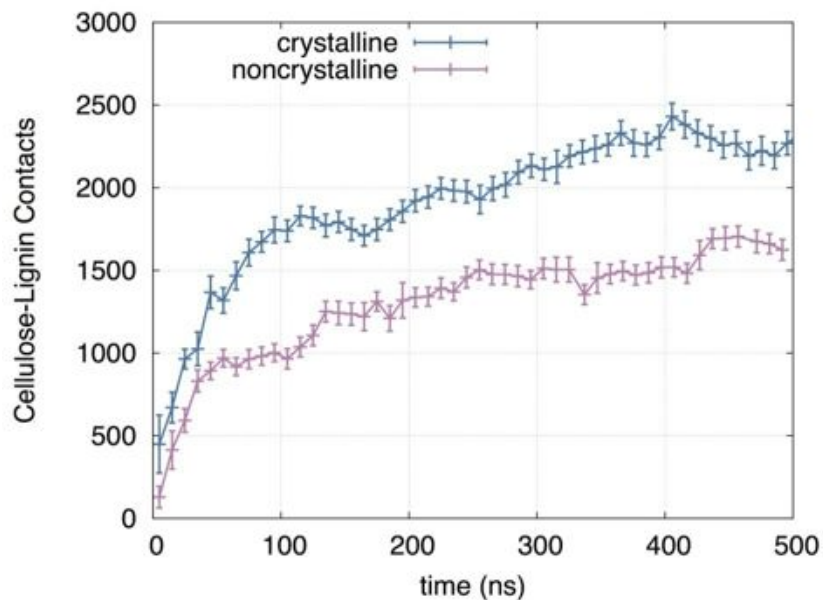
$$R_g = 420 \text{ \AA}$$



Enzyme:lignin
interaction
distribution



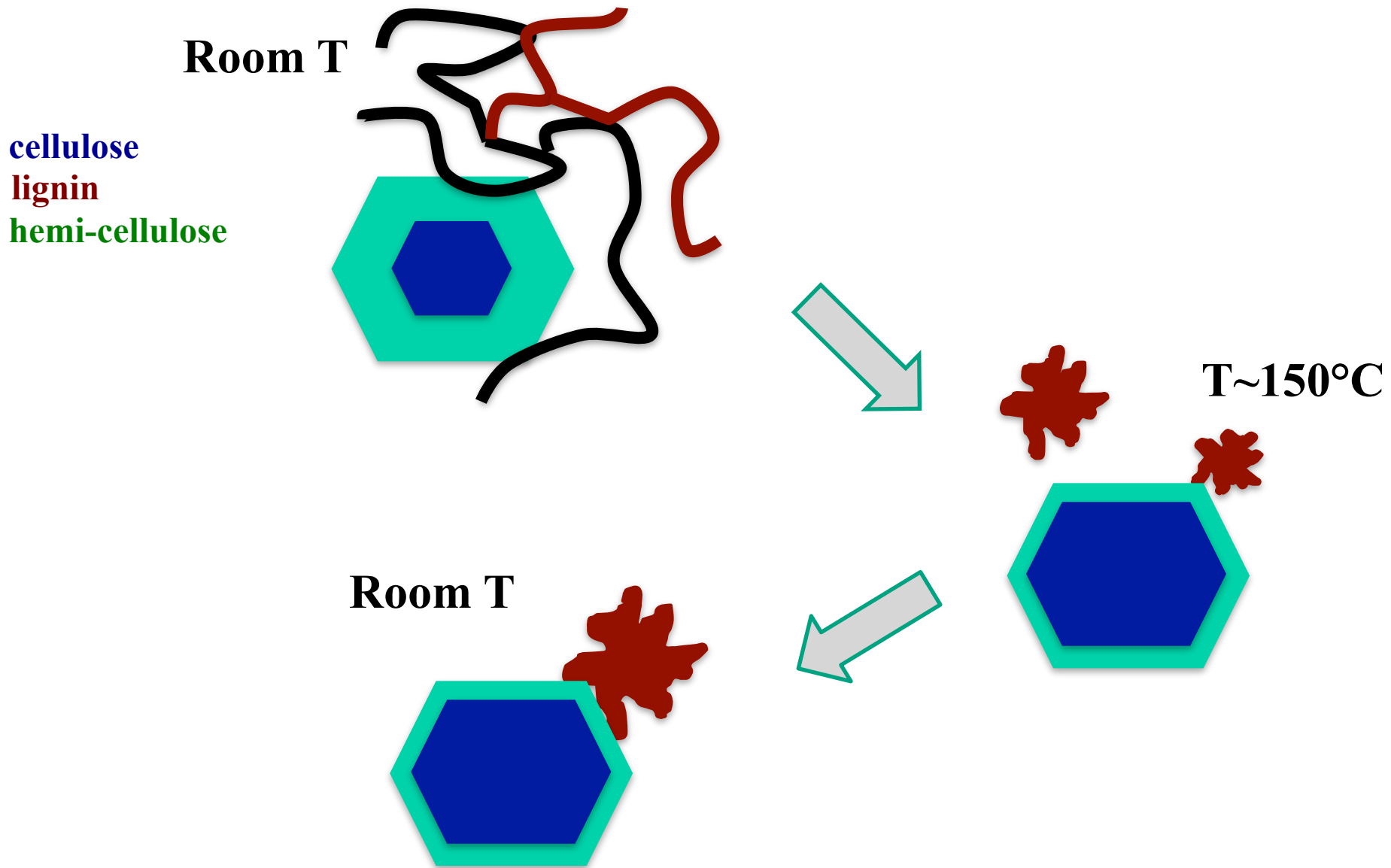




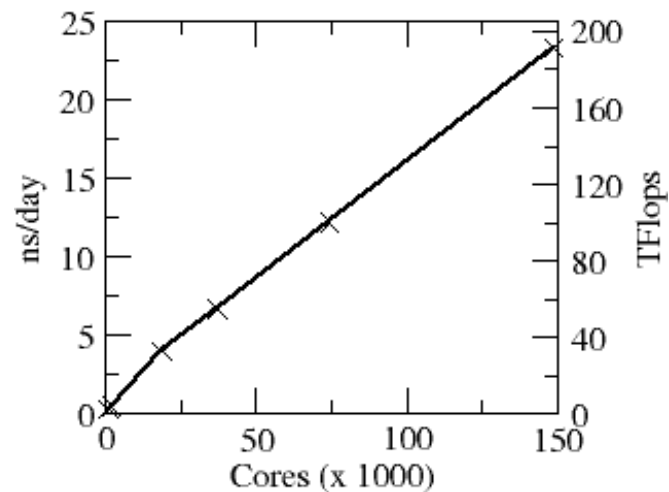
Interface	Interaction Energy Density (kJ/mol/nm ²)
lignin: crystalline cellulose	-49±2
lignin: non-crystalline cellulose	-50±2
water : crystalline cellulose	-94±2
water : non-crystalline cellulose	-107±2

Solvent-Driven Preferential Association of Lignin with Crystalline Cellulose

New View of Pretreatment



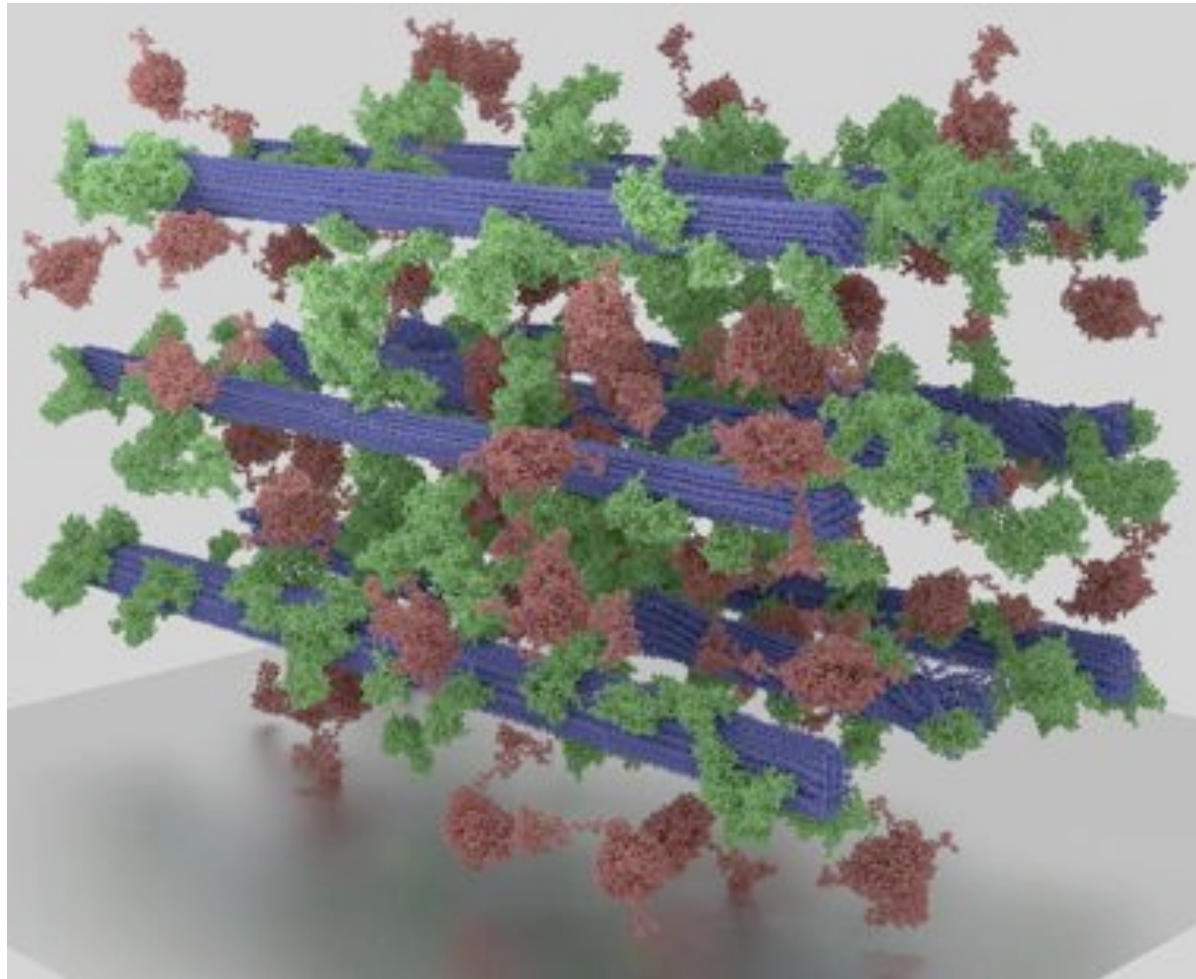
Supercomputer Scaling



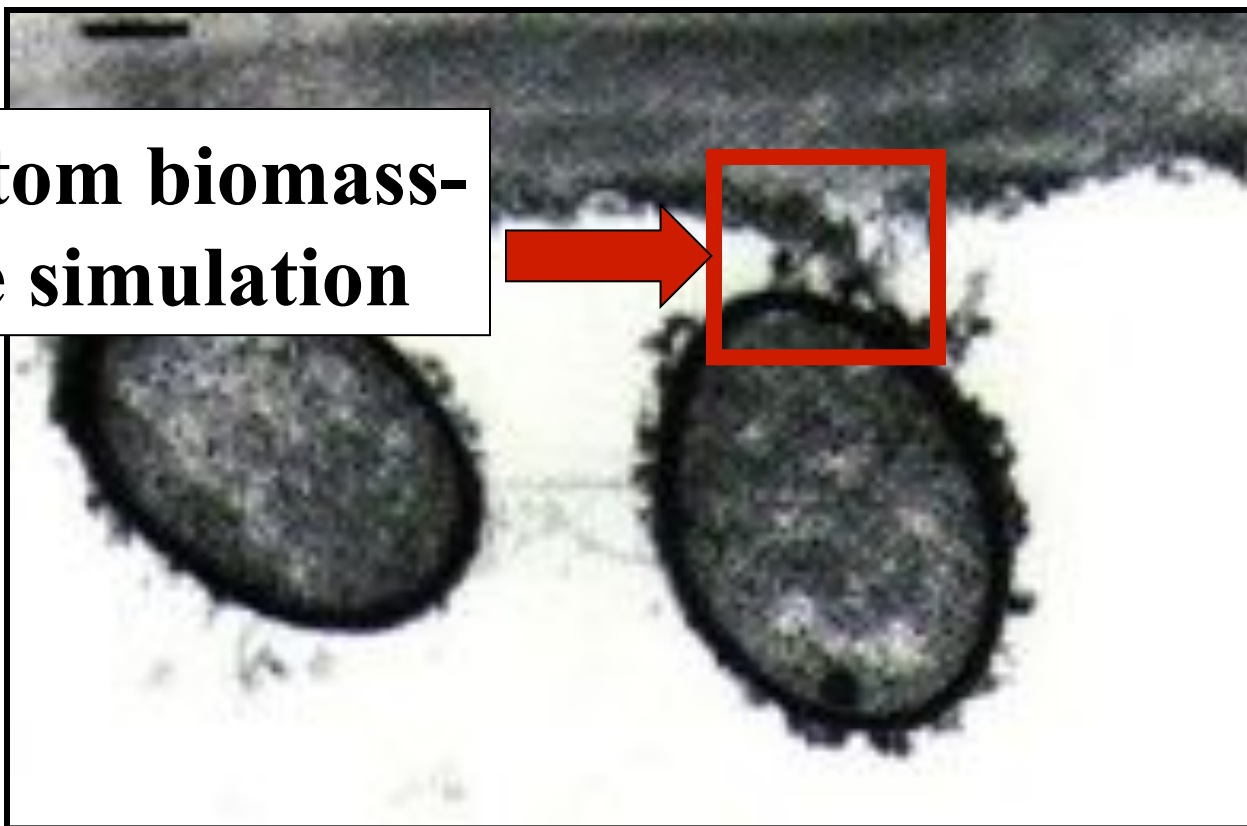
~ 100 million atoms.

- Scales to 150,000 cores

24M-atom capability-class simulation of enzyme binding to pretreated lignocellulose



300M atom biomass-
microbe simulation



Hydrolysis of biomass by *Clostridium thermocellum*

A scanning electron micrograph (SEM) of a human embryonic stem cell. The cell is a large, rounded, yellowish-brown structure with a textured surface, sitting on a network of purple and blue fibers. The background is a greenish, textured surface.

Exascale?

**10^{10} atoms
= Living Cell**

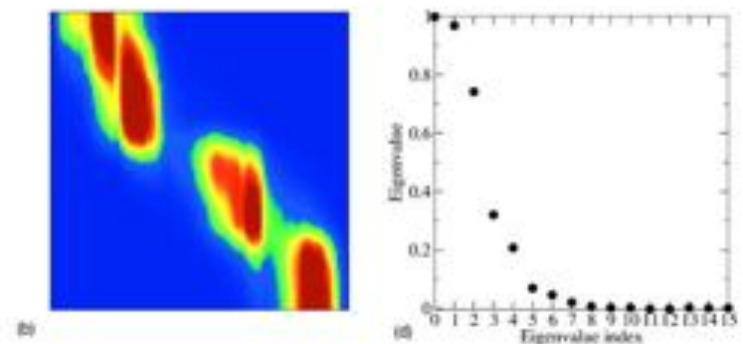
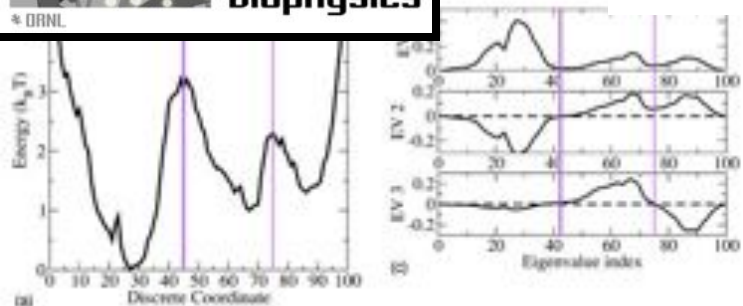
Human Embryonic Stem Cell

But...

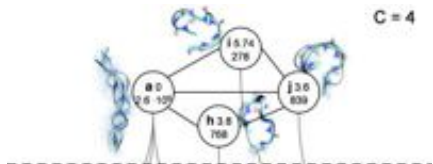
Microsecond Timescale Limitation!

New Concepts Needed....

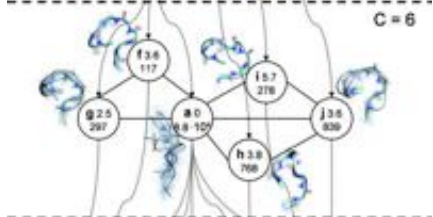
Exascale Concepts



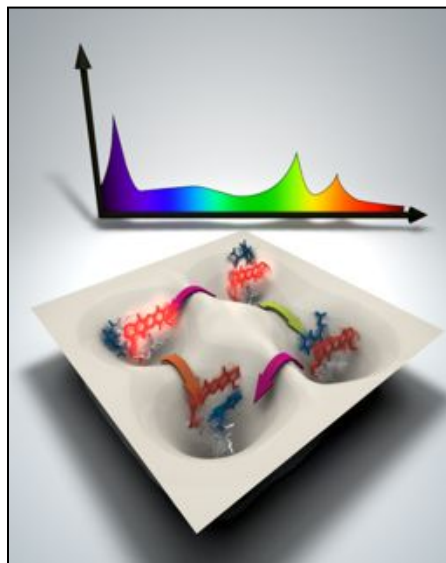
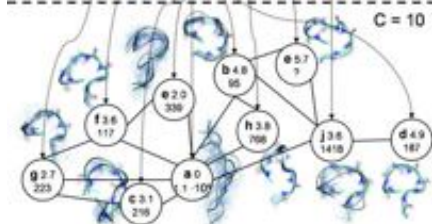
C = 4



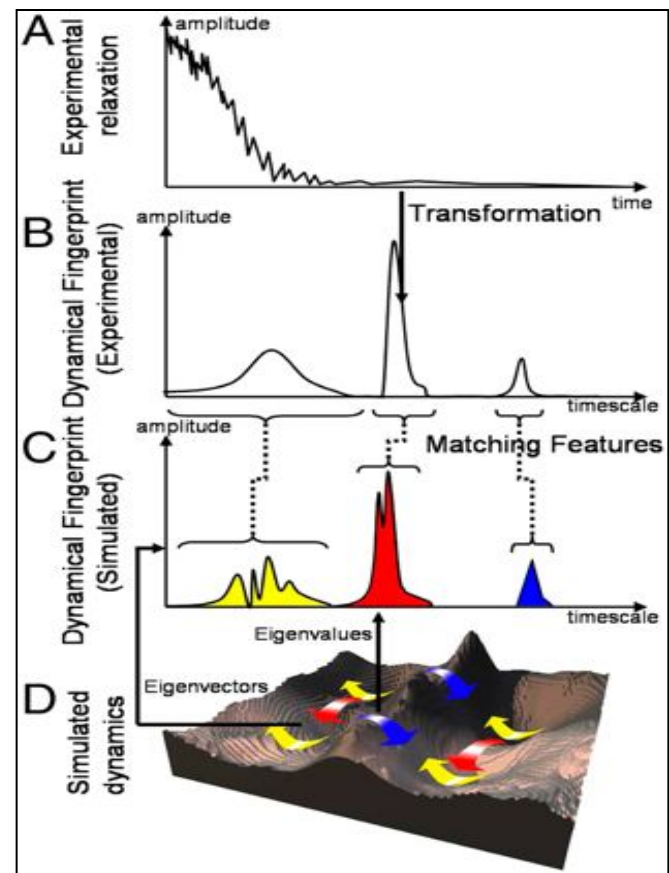
C = 6



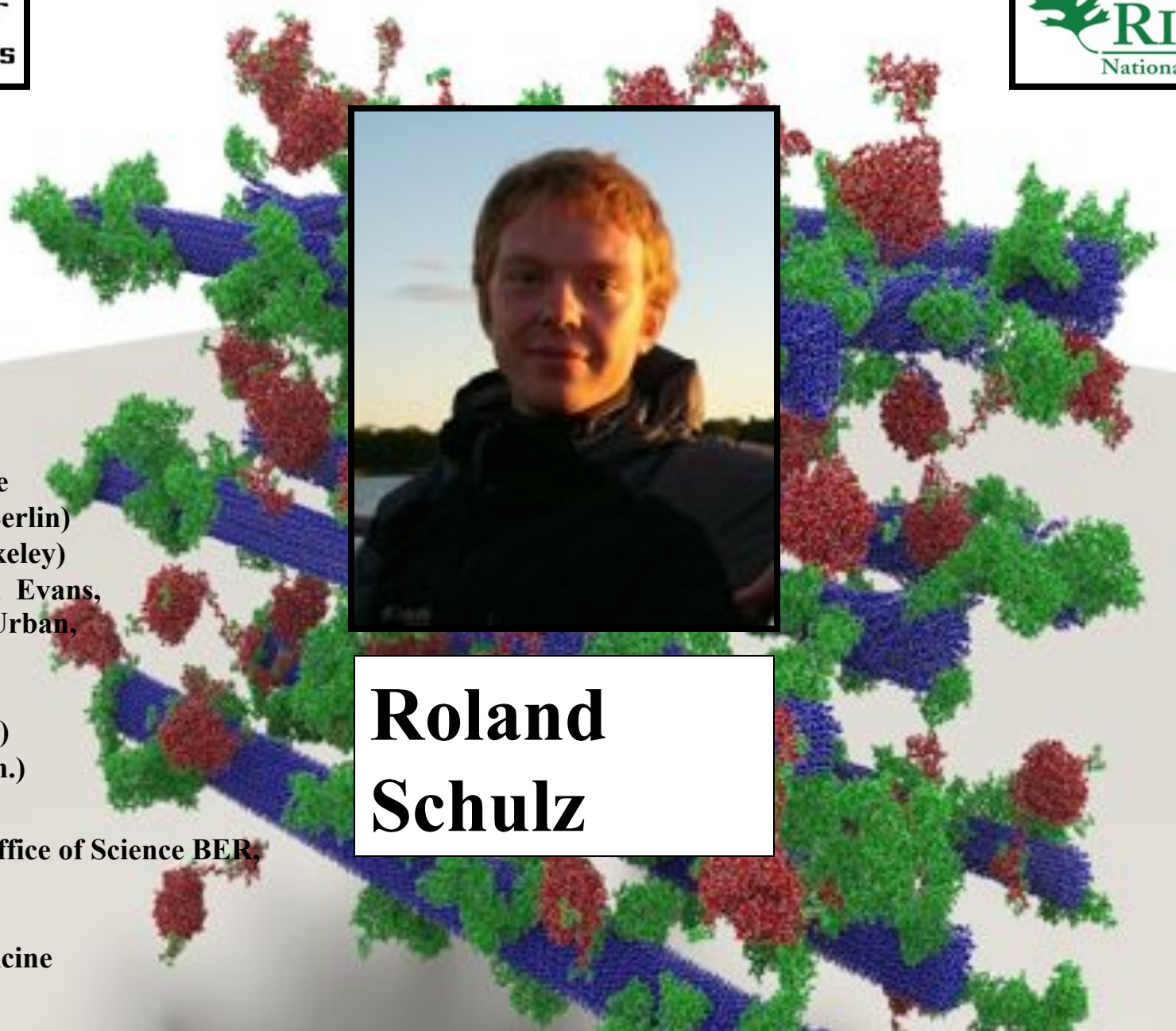
C = 10



Dynamical Fingerprints



JCP 126 840 (2007); 134 244108 (2011); PNAS 108 4822 (2011)



**Roland
Schulz**

Co-workers

- Barbara Collignon**
- Benjamin Lindner**
- Roland Schulz**
- Loukas Petridis**
- Sally Ellingson**

Collaborators

- Ilia Horenko, Frank Noe**
- Christof Schuette (FU Berlin)**
- John Chodera (UC Berkeley)**
- Venky Pingali, Barbara Evans,**
- William Heller, Volker Urban,**
- Alexei Sokolov (ORNL)**
- Marcus Foston,**
- Art Ragauskas (GaTech)**
- Jerome Baudry (U. Tenn.)**

Funding

- U.S. Dept. of Energy (Office of Science BER, ASCR)**
- NSF Biosystems Cluster**
- NIH Translational Medicine**
- DFG**
- DOE INCITE.**

