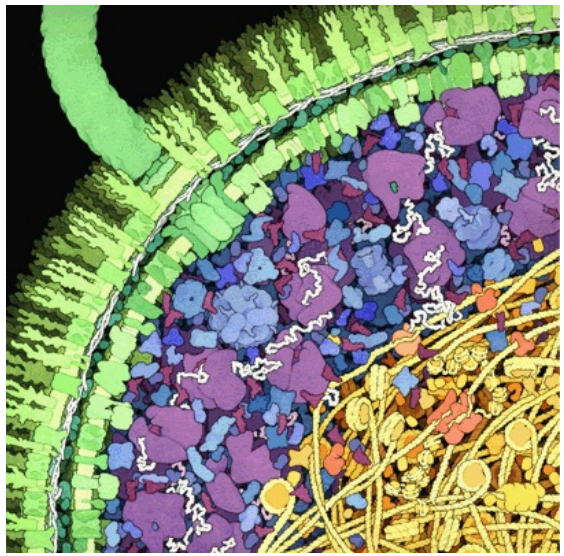
Modèles géométriques pour la prédiction des interactions macro-moléculaires

Geometric models for the prediction of macro-molecular interactions

Frederic.Cazals@inria.fr, ABS
http://team.inria.fr/abs



Inside Escherichia coli [D. Goodsell, The machinery of life]

Molecular interactions: function = structure (geometry) + dynamics

Protein complexes – physical chemistry 101

Modeling complexes: the machine learning approach

Modeling complexes: ab initio approaches

Conclusion

Outlook

Proteins and macro-molecular machines

Molecular interactions: function = structure (geometry) + dynamics

Protein complexes – physical chemistry 101

Modeling complexes: the machine learning approach

Modeling complexes: ab initio approaches

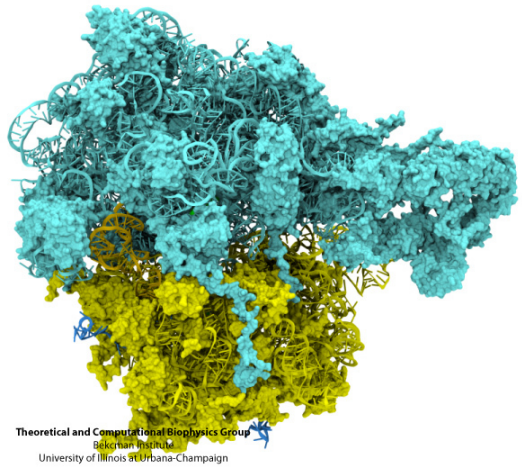
Conclusion

Outlook

The machinery of life: protein synthesis by the ribosome

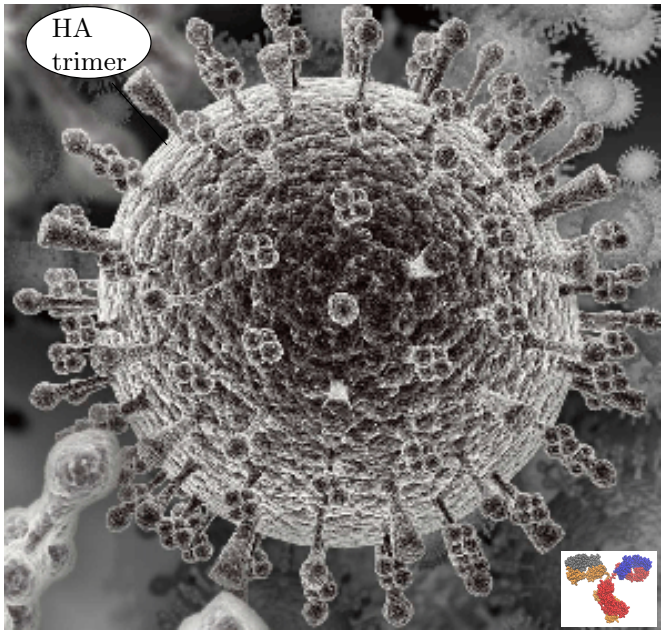
videos-science/video-ribosome-

youtube

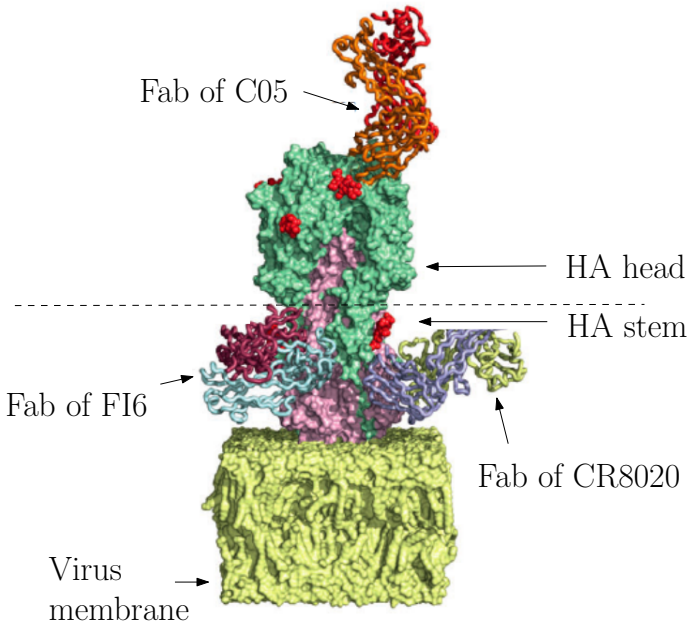


B-cell biology and antibody - antigen complexes

▷ Influenza



▷ (Broadly) neutralizing antibodies

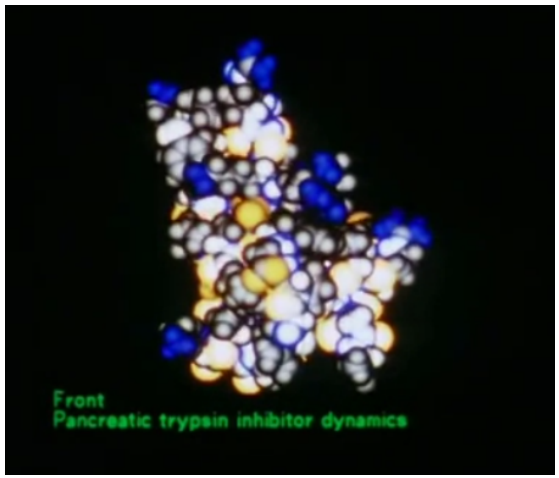


▷ Core questions on IG-Ag complexes

– Determinants of binding affinity
  relationship *affinity - avidity - virus entry inhibition*
– Role of complementarity determining regions (CDRs)
– Determinants of interaction specificity

Molecular dynamics: first simulation of a protein

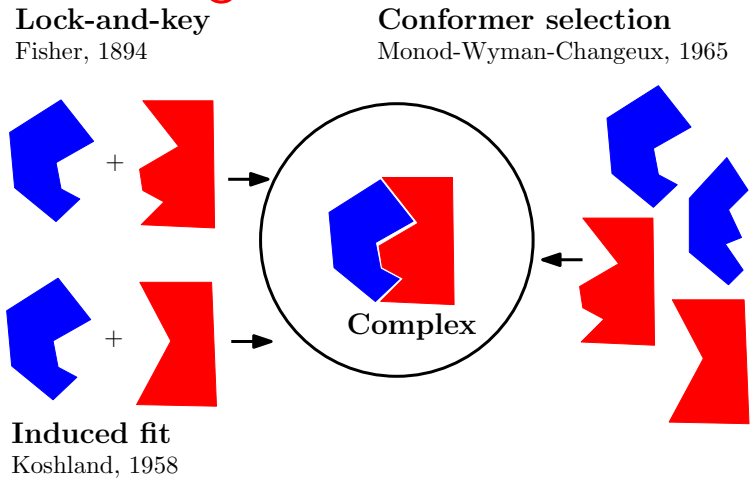
videos-science/video-michael-levitt-first-MD–simulation



About the simulation duration, quoting M. Levitt "*Cannot remember, but likely less than 100 picoseconds*"

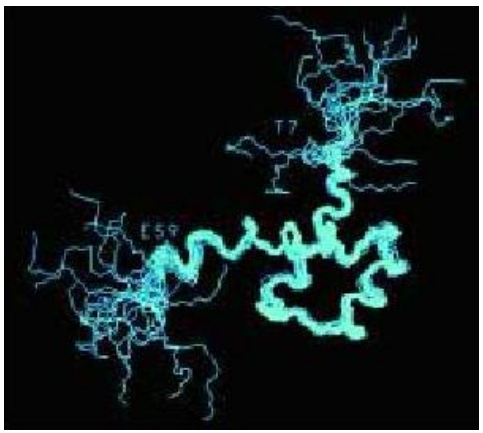
Protein interactions: docking, affinity, specificity

▷ Docking models:



- Lock-and-key: Fisher, 1894
- Induced fit: Koshland, 1958
- Conformer selection, Monod-Wyman-Changeux, 1965

▷ Flexibility matters



▷ Key ingredients:

– Geometry:
complementarity,conformations, flexibility

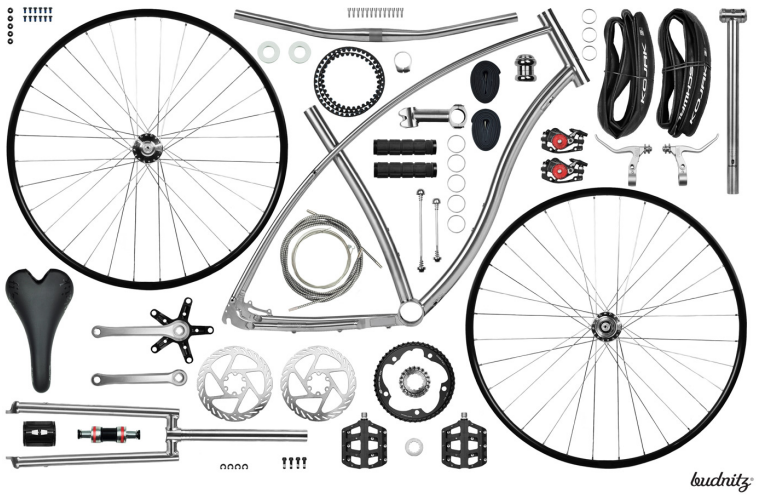
– Physics: enthalpy, entropy

▷ Major challenges (cf CAPRI):

geometry: large conformational changes

physics: entropy based affinity control

The lock and key metaphor is misleading: function is often about dynamics



$$\text{Information(spare parts)} < \text{Information(static bicycle)} \ll \text{Information(moving bicycle)}$$

Proteins and macro-molecular machines

Molecular interactions: function = structure (geometry) + dynamics

Protein complexes – physical chemistry 101

Modeling complexes: the machine learning approach

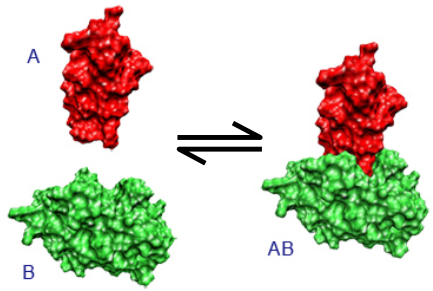
Modeling complexes: ab initio approaches

Conclusion

Outlook

Binding affinity: dissociation free energy

- Protein complexes rock back and forth



- Dissociation constant and dissociation free energy:

$$K_d = [A][B]/[AB]$$
$$\Delta G_d = -RT \ln K_d/c^\circ = \Delta H - T\Delta S.$$

- Binding affinities (thermodynamics):
  - random complex: $K_d \sim 10^{-6}$
  - high: $K_d \sim 10^{-9}$
  - very high: $K_d \sim 10^{-12}$
  - extreme: $K_d \sim 10^{-15}$

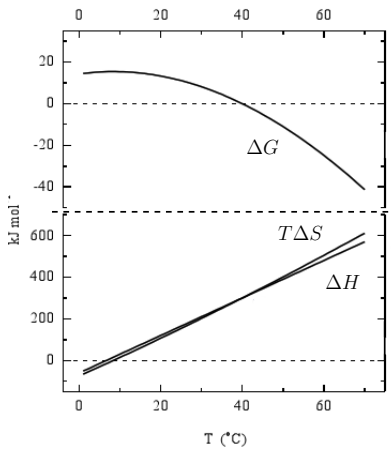
- Time scales (kinetics):
  - short-lived complexes: $10^{-6}s$ (e.g. enzyme-substrate)
  - stable complexes: $10^{3}s$ (e.g. antibody-antigen)
  - permanent complexes: $10^{6}s$ (aggregates)

Binding affinity: thermodynamics

- Dissociation constant k_D for $C \leftrightharpoons A + B$:

$$K_d = \frac{[A][B]}{[C]}; \Delta G_d = -RT \ln K_d/c^\circ = \Delta H - T\Delta S. \quad (1)$$

- The enthalpy - entropy compensation:
  - enhanced packing of interface atoms due to attractive forces: $\Delta H < 0$
  - higher packing, restricted atomic motions: $T\Delta S < 0$
- Marginal stability of proteins and complexes:



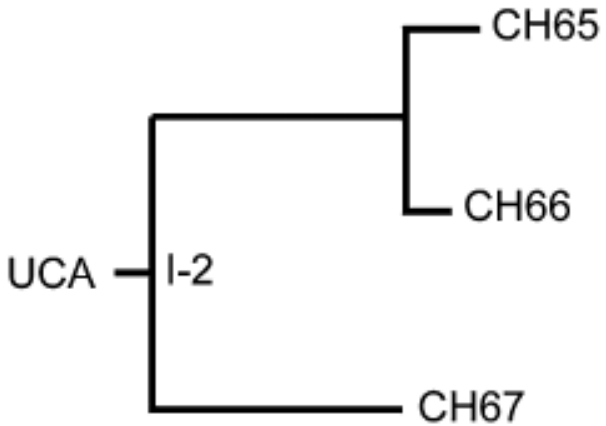
  - Large ΔH and $T\Delta S$ compensate
  - Crossing of curves difficult to predict
  - Marginals stability is key to regulation

Pict. courtesy of Alan Cooper (Thermodynamics of unfolding)

The immune response: affinity maturation

Rigidification of CDR loops limits the entropic penalty upon binding

▷ Antibodies: lineage

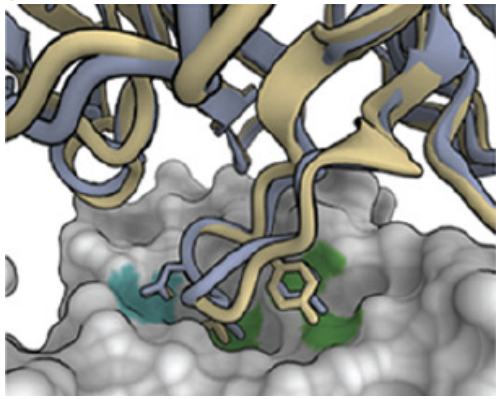


▷ Binding affinities: K_d analysis by SPR

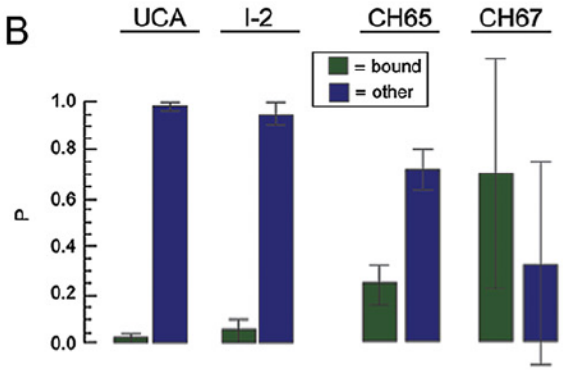
Fab	$K_d(\mu M)$
UCA	118 ± 14
I-2	142 ± 15
CH65	$0.49 \pm .10$
CH67	0.36 ± 0.04

CH65 $\sim$ CH67; wrt UCA:
$\Rightarrow \sim$ 200-fold improvement

▷ But UCA and CH65 have similar binding modes!!!



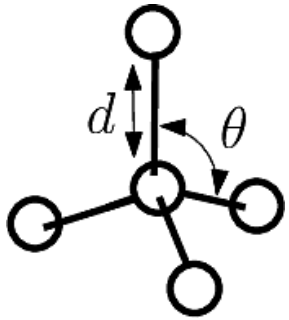
▷ Solution: time spent bound conformations – long MD simulations



▷Ref: Harisson et al; PNAS 110, 2013

Force fields: the potential energy of a (bio-)molecular system

- The $3n-6$ degrees of freedom of a molecule:



  - types for atoms (element, bonds)
  - covalent: bond lengths, angles
  - non covalent: pairwise distances
  - solvent model

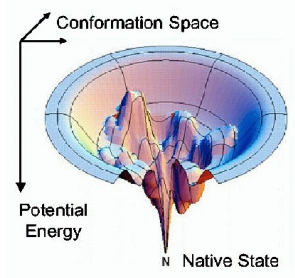
- Potential energy:

$$U_{\text{total}} = E_{\text{bond}} + E_{\text{angle}} + (E_{\text{proper}} + E_{\text{improper}}) + (E_{\text{vdw}} + E_{\text{electro}}) \qquad (2)$$

E_{bond}: bonds
E_{angle}: covalent angles
E_{proper}: proper dihedrals
E_{improper}: improper dihedrals
E_{vdw}: van der Walls
E_{electro}: electrostatics

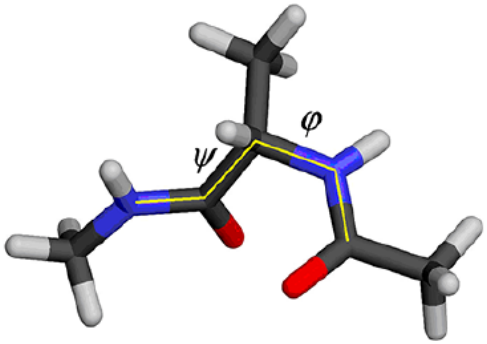
- Examples:

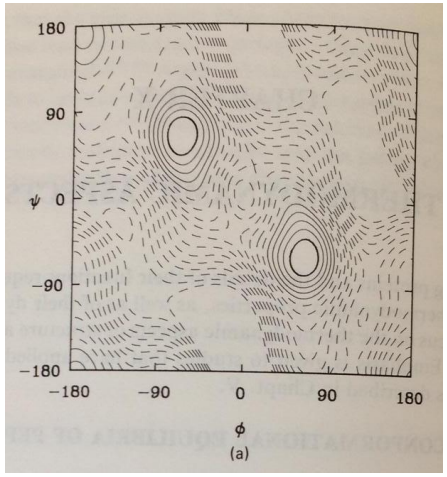
  - AMBER: $S_u = (73, 133, 112, 3, 14, 758)$
    1093 unique parameters
  - CHARMM: $S_u = (85, 152, 209, 13, 33, 1)$
    493 unique parameters
  - MARTINI: $S_u = (16, 4, 0, 2, 21, 3)$
    46 unique parameters

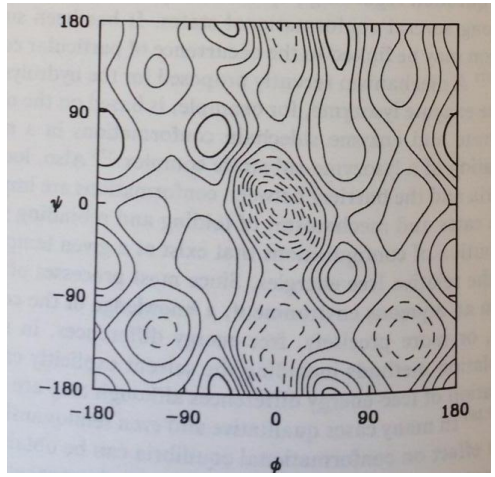


Potential energy landscapes: illustration

- Potential energy map: vacuum versus solvated

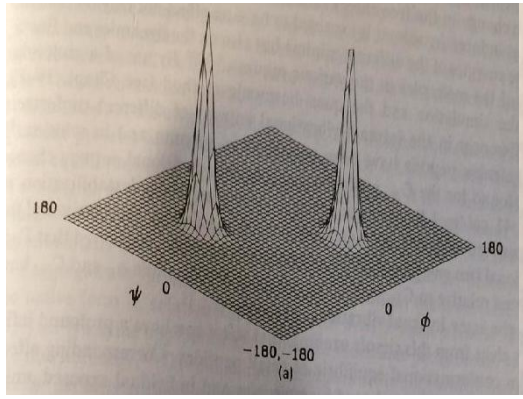


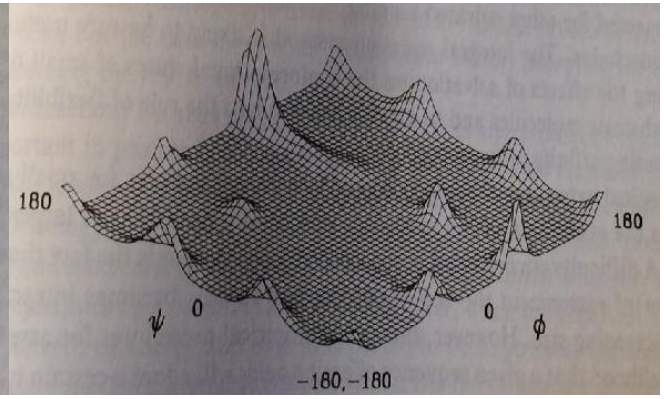




- Corresponding Boltzmann-weighted probability maps:

– Solvent stabilizes many more conformers–hydrogen bonding.
– Dramatic incidence of the PES and FES.

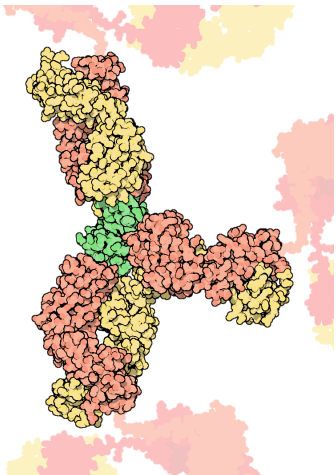




- Ref: Petitt, Karplus, Chem. Phys. Lett., 121, 1985

Binding affinity: direct calculation

▷ A standard antibody-antigen complex:



▷ Model without solvent:

- FAB of antibody $\sim$ 3000 atoms
- Antigen (lysozyme) $\sim$ 1000 atoms
- One conformation: 1 point in $\mathbb{R}^{3\times 4000}$

▷ ΔG_d as a multidimensional integral:

$$\Delta G = -\frac{1}{\beta} \ln \left(\frac{1}{8\pi^2} \frac{C_A C_B}{C_{AB}} \frac{\int e^{-\beta U(r_{AB})} dr_{AB}}{\int \left(e^{-\beta U(r_A)} dr_A\right) \left(\int e^{-\beta U(r_B)} dr_B\right)} \right) \tag{3}$$

▷Ref: Woo ad Roux, PNAS 102 (19), 2005

Free energy, density of states, and volume calculations

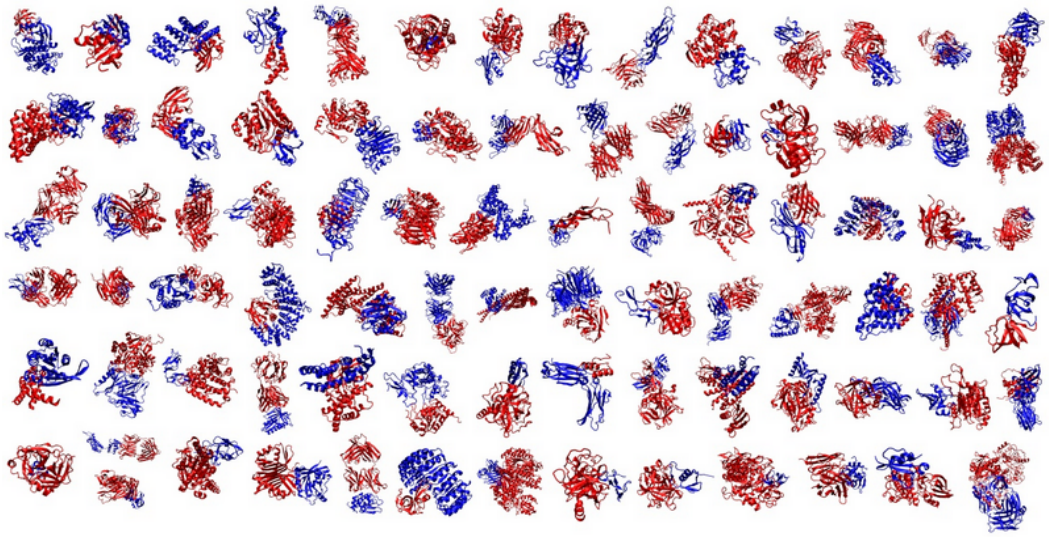
- ▷ Density of states
- ▷ Volume of polytopes: hardness

▷Ref: Dyer, Freeze, Kannan, J. ACM 38(1), 1991

▷Ref: Lovász, Vempala, J. Comput. Syst. Sci., 71(2), 2006

Protein interactions: the structure affinity benchmark

http://bmm.cancerresearchuk.org/~bmmadmin/Affinity/



▷ Dissociation constant vs affinity

$$\Delta G_d = -RT \ln K_d/c^\circ$$

▷ NB: in general, bound partners only do not suffice to get accurate predictions

- ▷ 144 protein complexes
  - 17 IG - Ag complexes
- ▷ Binding affinity known: ITC, SPR
  - caveat: order of magnitude matter (pH, ion strength, . . .)
- ▷ Three crystal structures known: bound complex + 2 unbound partners

▷Ref: Kastritis et al; Protein Science (20), 2011

Estimating Kd: two routes

- Learning: regression
  - Databases of crystal structures + affinity measurements
  - Regression models involving relevant variables
- From first principles
  - Atomic models of the partners
  - A force field and a thermodynamic sampling algorithm

Proteins and macro-molecular machines

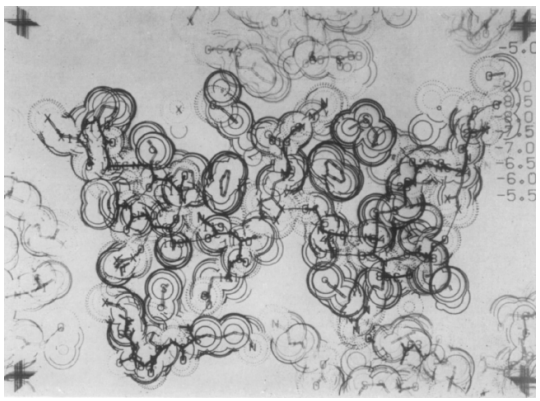
Binding affinity estimation as a regression problem

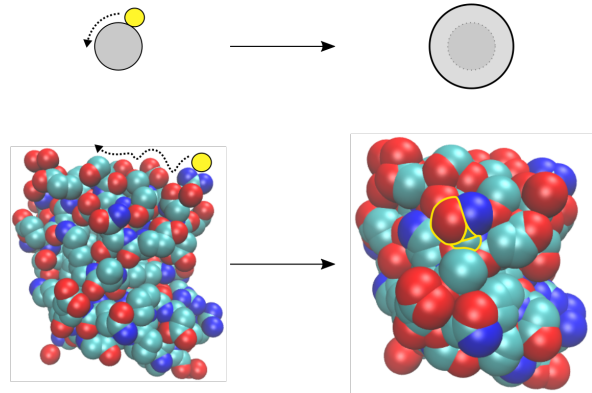
- Regression:
  - Regression: predicting the value of a continuous (dependent) variable from the values of other (independent) variables.
  - ΔG is the dependent variable
  - Many types of regressors: least squares, regularized least squares, k nearest neighbours, regression trees, multivariate adaptive splines, ...
- Adequate variables: two classes of methods
  - Large collections of parameters coding distances, biochemical properties (H-bonds, properties of a.a.), conservation of a.a., etc.

    NB: requires a close monitoring to avoid overfitting.
  - A small number of them: more precise encoding of enthalpy and entropy related quantities.
- Overfitting and sparsity
  - Variable selection and regularization via the LASSO
  - Sparse model enumeration + cross validation

Solvent Accessible Models: the birth

"*The successful elucidation of the structure of a protein by single-crystal diffraction procedures provides a list of atomic co-ordinates whose reliability will vary in different parts of the molecule.*"
"*The topology of the surface of a protein is intimately related to its function; parts of the surface are directly involved in interactions with other molecules; the solvent- protein interface is almost certainly related to the structure of the native molecule; and the chemical reactivity of the various functional groups will depend on their relation to this interface.*"





▷Ref: Lee and Richards, JMB, 3 (55), 1971

▷Ref: M.L. Connolly, J. Appl. Crystallography, 1983

▷Ref: Akkiraju and Edelsbrunner, Discrete Appl. Math., 1996

Solvent Accessible Models: the rise

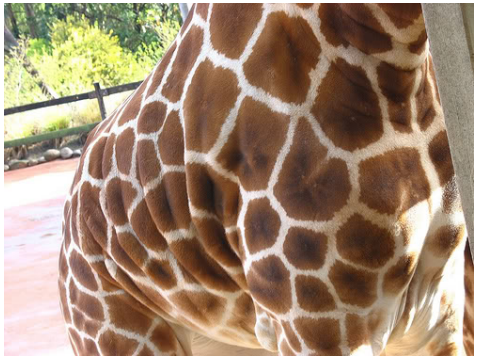
From Chotia, *Structural invariants in protein folding*:
"*An analysis of 15 protein structures indicates: First, the loss of accessible surface area by monomeric proteins on folding–proportional to hydrophobic energy–is a simple function of molecular weight; second, the proportion of polar groups forming intramolecular hydrogen bonds is constant; and third, protein interiors are closely packed, each residue occupying the same volume as it does in crystals of amino acids.*"

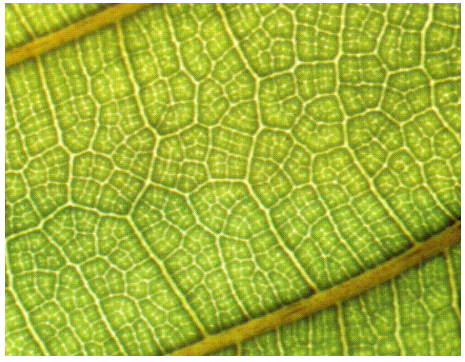
From Janin, *Principles of protein–protein recognition*:
"*The formation of the protein–protein interface by the insulin dimer, the trypsin-PTI complex and the* $\alpha\beta$ *oxyhaemoghbin dimer removes 1,130–1,720* ^{2} *of accessible surface from contact with water. The residues forming the interface are close packed: each occupies the same volume as it does in crystals of amino acids. These results indicate that hydrophobicity is the major factor stabilising protein–protein association, while complementarity plays a selective role in deciding which proteins may associate.*"

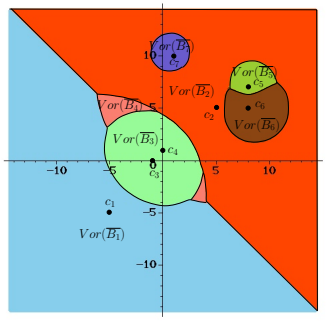
▷`Ref: Chothia, Nature 254, 1975`

▷`Ref: Janin, Nature 256, 1975`

Voronoi diagrams in Biology, Geology, Engineering

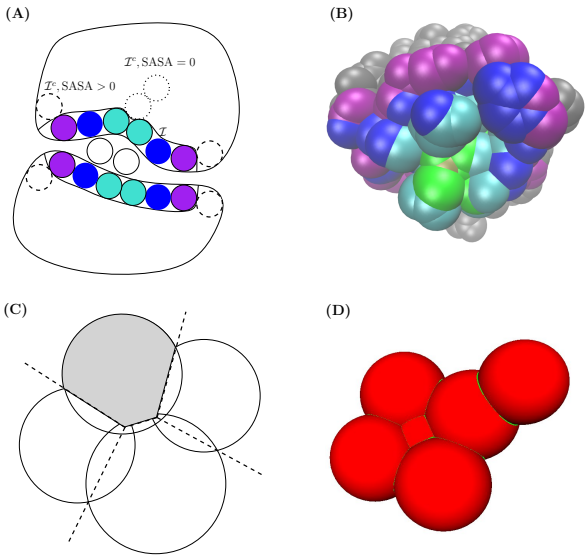






Our parameters: overview

▷ Our variables: proxys for enthalpy and (vibrational) entropy variations upon binding, the latter based on packing properties



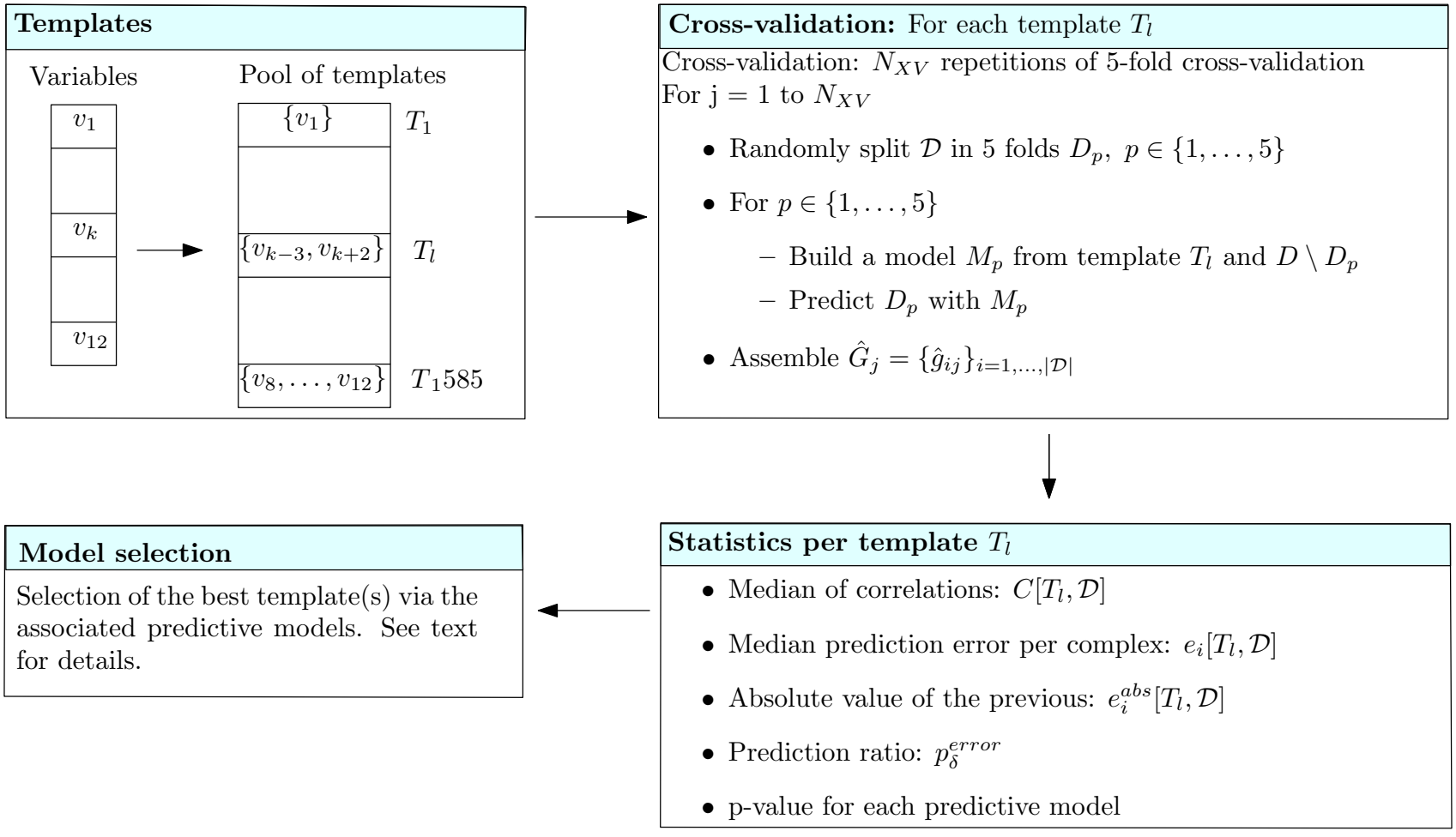
▷ Or particular interest

- IVW-IPL: inverse volume-weighted internal path length
- $\text{NIS}^{charged}$: fraction of charged residues on the non-interacting surface (NIS)

- **(A)** Binding patch and labeling of interface atoms The non interface atoms ($\mathcal{I}^c$) are split into those which retain solvent accessibility (SASA > 0, dashed balls), and those which do not (SASA $= 0$, dotted balls)

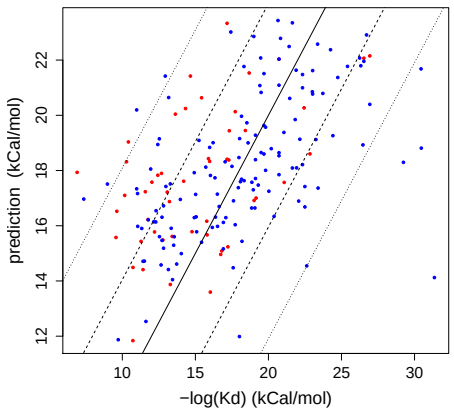
  NB: **Buried Surface Area** or **BSA**: area of colored spherical caps
- **(B)** Shelling order of an atom: smallest number of atoms traveled to reach an exposed non interface atom, i.e. an atom belonging to $\mathcal{I}^c$ and with SASA > 0 (in grey)
- **(C,D)** Atomic packing: via Voronoi volumes

Statistical methodology

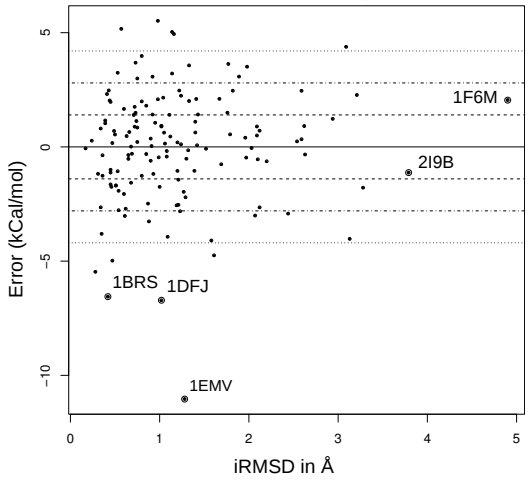


Results on the structure affinity benchmark

▷ Predictions vs measurements



▷ Hardness vs flexibility



▷ State-of-the-art binding affinity estimates on the SAB:

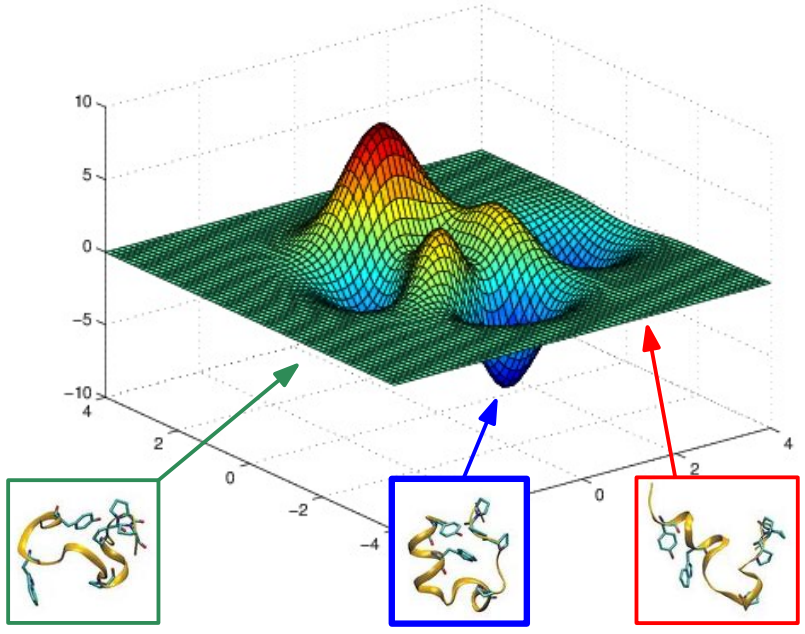
– Whole SAB: K_d within one and two *OOM* in 48% and 79% of cases
  high resolution (2.5Å): K_d within one and two *OOM* in 62% and 89%
– Absence of correlation between prediction hardness and protein flexibility

▷ References:

– 1 OOM (order of magnitude) $\Leftrightarrow$ 1.4 kcal/mol
– kT per molecule, or RT per mole at room temperature: 0.6 kcal/mol
– ΔG_d, exp. errors $\sim$ 0.3 kcal/mol

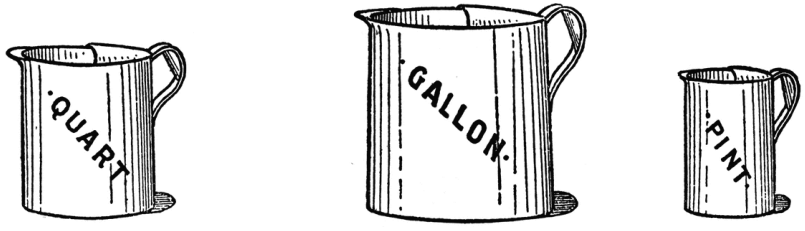
Proteins and macro-molecular machines

Emergence of function from Structure – Thermodynamics – Dynamics



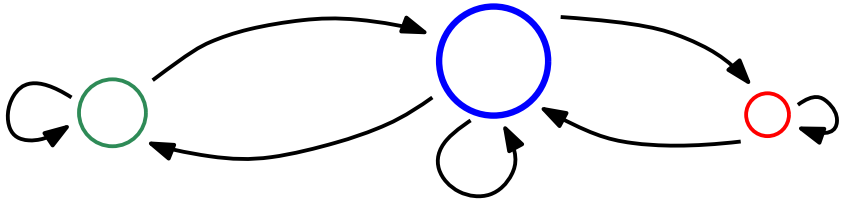
Potential Energy Landscape

- large number of local minima
- enthalpic barriers
- entropic barriers



Structure: stable conformations i.e. local minima of the PEL

Thermodynamics: meta-stable conformations i.e. ensemble of conformations easily inter-convertible into one - another.



Dynamics: transitions between meta-stable conformations e.g. Markov state model

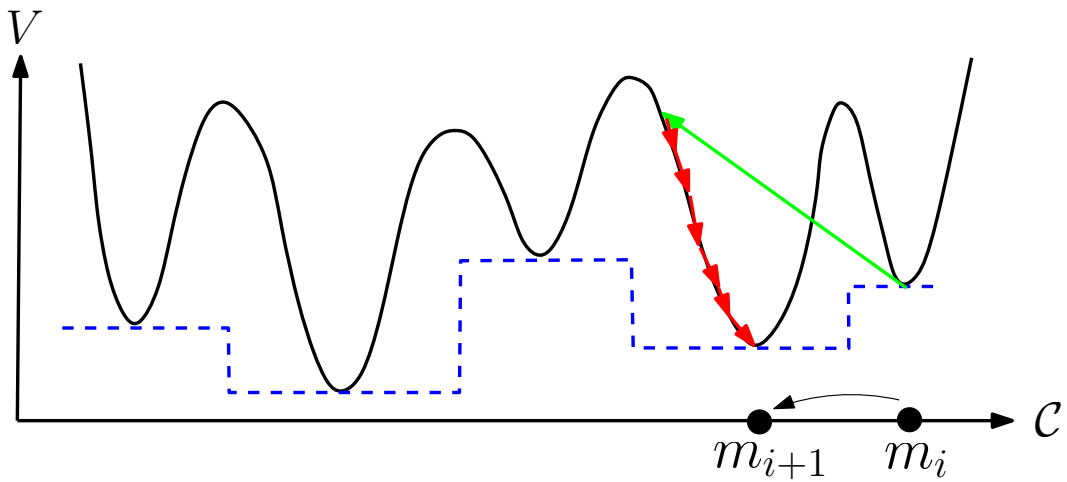
Contributions discussed

- (Structure) Sampling potential energy landscapes
- (Thermodynamics) Simplifying potential energy landscapes

Exploring Potential Energy Landscapes:

basin hopping

- ▷ Goal: enumerating low energy local minima
- ▷ Basin-hopping and the basin hopping transform
  - Random walk in the space of local minima
  - Requires a *move set* and an *acceptance test* (cf Metropolis)
    and the ability to descend the gradient (*quenching*)
    aka energy minizations
- ▷ Limitation: no built-in mechanism to avoid staying trapped

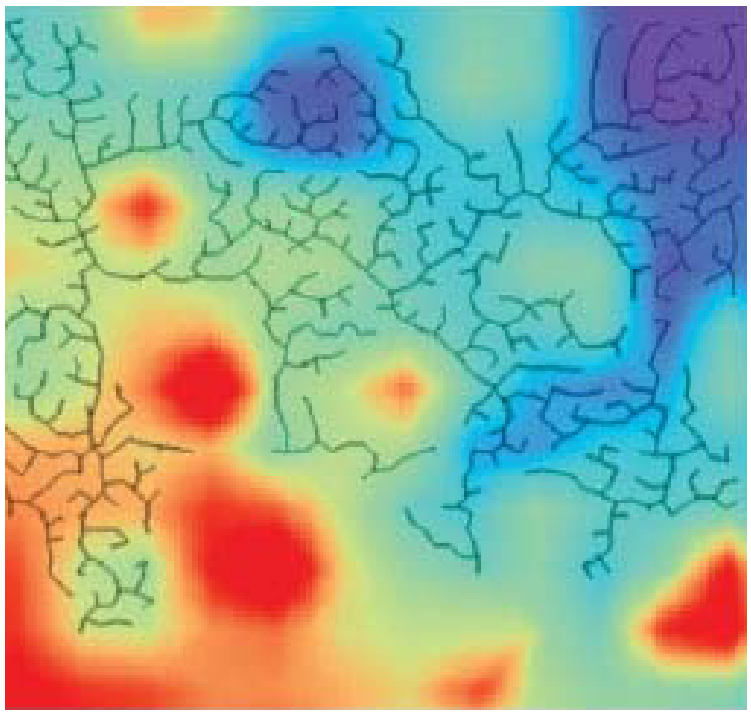


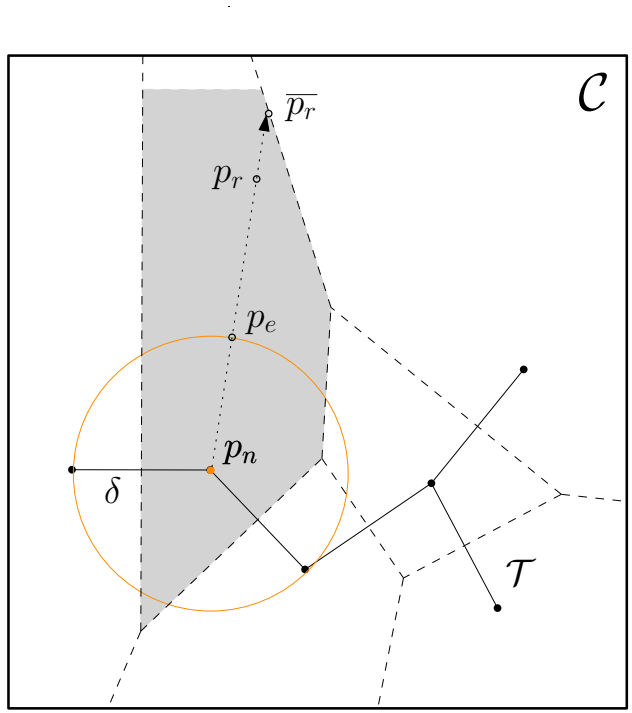
▷Ref: Li and Scheraga, PNAS, 1987

Exploring Potential Energy Landscapes:

transition based rapidly exploring random trees (T-RRT)

- ▷ Goal: sample basins and transitions
- ▷ Algorithm growing a random tree favoring yet unexplored regions
  - node to be extended selection: *Voronoi* bias
  - node extension: interpolation + Metropolis criterion (+temperature tuning)
- ▷ Limitation: oblivious to local minima



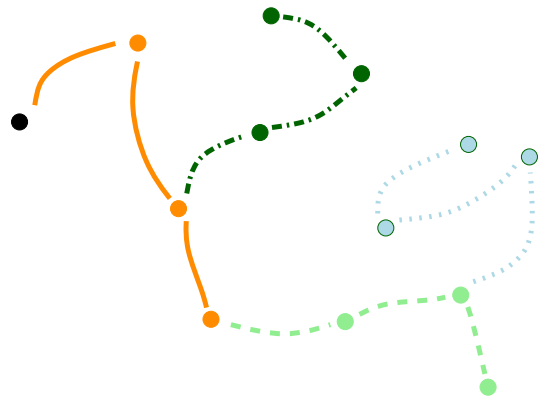


▷Ref: LaValle, Kuffner, IEEE ICRA 2000

▷Ref: Jaillet, Corcho, Pérez, Cortés, J. Comp. Chem, 2011

Exploring energy landscapes:
a generic approach yielding BH, T-RRT,...

- Input: potential energy function with million, billion, trillion of local minima
- Goal: enumerate low energy + persistent local minima
- Hybrid algorithm: alternate BH and T-RRT extensions



- Key ingredients:
  - Boost the exploration of yet-unexplored regions – Voronoi bias
  - Meaning-full management of distances – due to concentration phenomena
  - Favor spatial adaptation – local Metropolis-Hasting tests

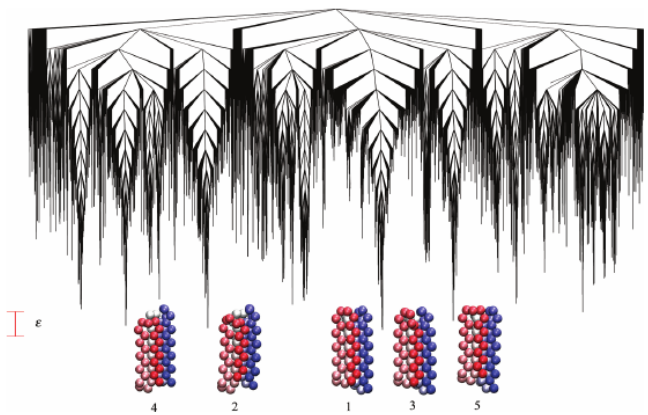
Ref: Roth, Dreyfus, Robert, Cazals; J. Comp. Chem.; 2015

Protein model BLN69: model and force field

- Description:
  - Three types of Beads: : hydrophobic(B), hydrophylic(L) and neutral(N)
  - Configuration space of intermediate dimension: 207
  - Challenging: frustrated system
  - Exhaustively studied: DB of $\sim 450k$ critical points (Industry)

$$V_{BLN} = \frac{1}{2} \cdot K_r \sum_{i=1}^{N-1} (R_{i,i+1} - R_e)^2 + \frac{1}{2} K_0 \sum_{i=1}^{N-2} (\theta_i - \theta_e)^2 + \epsilon \cdot \sum_{i=1}^{N-3} [A_i(1 + \cos\phi_i) + B_i(1 + 3\cos\phi_i)]$$
$$+4\epsilon \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} \cdot C_{ij}[(\frac{\sigma}{R_{i,j}})^{12} - D_{ij}(\frac{\sigma}{R_{i,j}})^{6}]$$

- Disconnectivity graph: describes merge events between basins

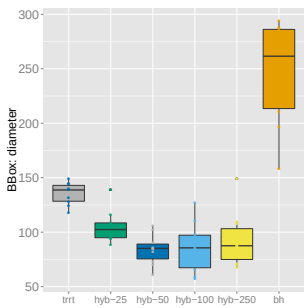


- Ref: Honeycutt, Thirumalai, PNAS, 1990
- Ref: Oakley, Wales, Johnston, J. Phys. Chem., 2011

Exploring energy landscapes: performances of Hybrid

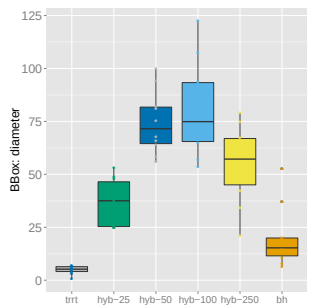
- ▷ Contributions: enhanced exploration of low lying regions of a complex landscape
- ▷ Protocol:
  - Contenders: BH, T-RRT, Hybrid for various parameter values b
  - Count and assess the local minima reported from two reference databases:
    $BLN69 - min - all$: 458,082 minima
    BLN69-min-E_{-100}: 5932 minima.

- Bounding box $\emptyset$: all mins



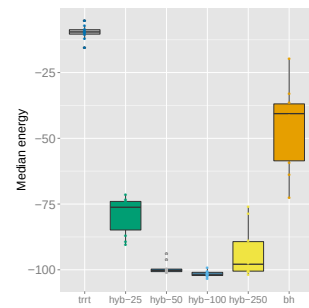
$BLN69 - min - all$

vs low lying



$BLN69 - min - E_{-100}$

- Median energies



$BLN69 - min - all$

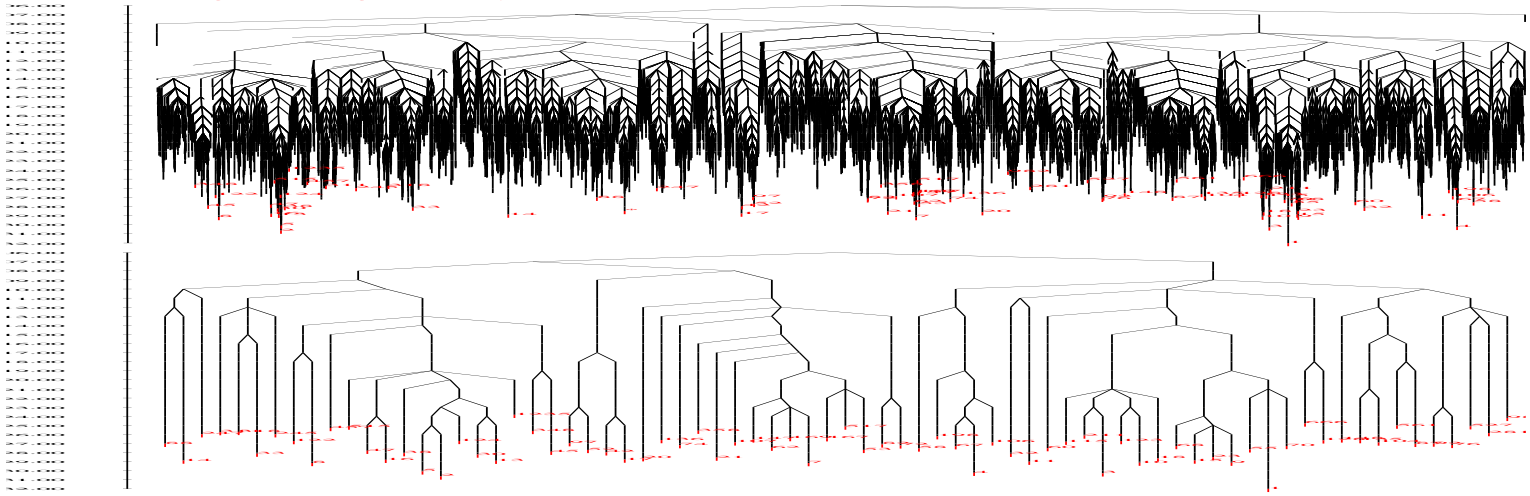
- ▷ Assessment:
  - Combines critical building blocks:
    minimization, spatial exploration boosting, nearest neighbor searches
  - Bridging the gap to thermodynamics

▷Ref: Oakley et al; J. of Physical Chemistry B; 2011

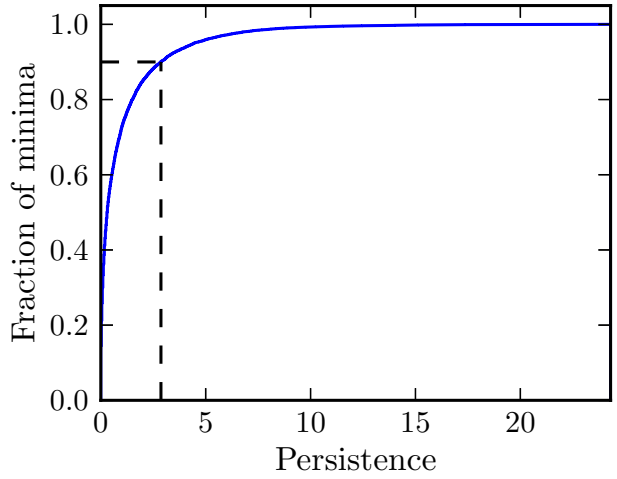
▷Ref: Roth, Dreyfus, Robert, Cazals; J. Comp. Chem.; 2015

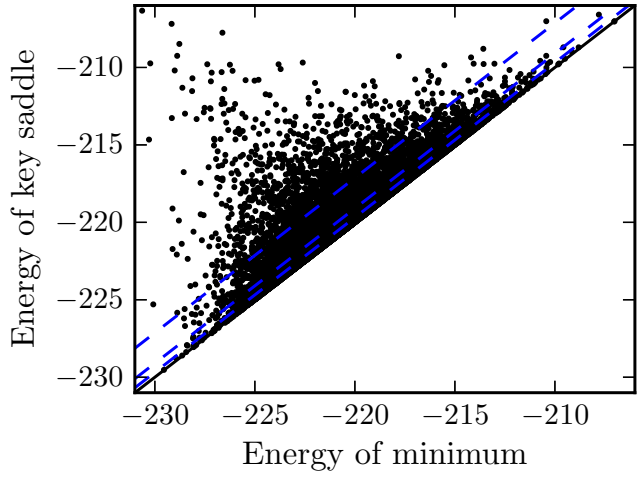
Binary Lennard-Jonnes LJ_{60}

- Coarse graining the system:



- Using the distribution of barriers' heights:





- Ref: Carr, Mazauric, Cazals, Wales; J. Chem. Phys.; 2016

Proteins and macro-molecular machines

Learning vs ab-initio approaches: different philosophy?

- No since the development of force fields involves:
  - tuning the parameters so as to match experimental data
    using (small) organic molecules

  NB: identical methods: optimization, cross-validation, Bayesian models
  - extrapolating to bio-molecules

Table 1. Comparison of Water Model Performance at 298.15 K, 1.0 atm[a]

property	expt.	TIP3P	SPC/E	TIP4P	TIP4P-Ew	TIP4P/2005	TIP3P-FB (this work)	TIP4P-FB (this work)	iAMOEBA
ρ/g cm^{-3}	0.997	0.98	0.994	0.992	0.995	0.993	0.995	0.996	0.997
ΔH_{vap}/kcal mol^{-1}	10.52	10.05	10.43	9.90	10.58	10.93	10.71	10.80	10.94
$\alpha/10^{-4}$ K^{-1}	2.56	9.2	5.0	4.4	3.2	2.8	4.1 (1)	2.5 (1)	2.5 (1)
$\kappa_T/10^{-6}$ bar^{-1}	45.3	57.4	46.1	60	48	46	44.5 (3)	45.2 (2)	41.1 (4)
C_p/cal mol^{-1} K^{-1}	18.0	18.74	18.3	18.9	19.2	19.0	19.1 (1)	19.0 (1)	18.5 (2)
$\varepsilon(0)$	78.5	94	68	53	62	58	81.3 (9)	77.3 (4)	80.7 (11)
$D_0/10^{-5}$ cm^2 s^{-1}	2.29	6.05	2.97	4.05	2.83	2.59	2.28 (2)	2.21 (2)	2.54 (2)
η/mPa s	0.896	0.321	0.729	0.494	0.72	0.855	0.91 (2)	0.94 (3)	0.85 (2)
σ/mJ m^{-2}	71.8	52	63	59	65	69	64 (1)	70 (1)	69 (1)
TMD (°C)	+4	−91	−36	−20	+1	+5	−12	+4 (1)	+4 (1)

[a]Properties listed are density ρ, heat of vaporization ΔH_{vap}, thermal expansion coefficient α, isothermal compressibility κ_T, isobaric heat capacity C_p, static dielectric constant $\varepsilon(0)$, self-diffusion coefficient D_0, shear viscosity η, surface tension σ, and temperature of maximum density TMD. The polarizable and relatively complex iAMOEBA model (right column) is included for comparison because it was parameterized using ForceBalance and a similar data set.

Ref: Pande et al, The J. Phys. Chem. letters, 5 (11), 2014

What are we critically missing to enter the era of atomic level engineering?

- Fundamental insights into equilibrium thermodynamics require:
  - potential energy: enhanced exploration algorithms
    akin to shape / model learning
  - free energy: enhanced multicanonical sampling algorithms
    akin to high dimensional volume calculations
  - dynamics: multi-scale Markov state models

- Countless breakthroughs in terms of applications:
  - biology: understanding processes; understanding evolution
    coding sequences $\sim$ 80 millions in `UniProt/TrEMBL`
    structures: $125,000$ in the Protein Data Bank
  - medicine: immunology, cancer, neurosciences,...
  - material sciences
  - synthetic biology

▷Ref: UniProt/Trembl: http://www.ebi.ac.uk/uniprot/TrEMBLstats
▷Ref: PDB: http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html

Proteins and macro-molecular machines

Hall of fame

- More than 20 structural biology-related Nobel Prizes in 50 years:
  - J. Kendrew and M. Perutz, chemistry 1962: *for their studies of the structures of globular proteins*
  - F. Crick, J. Watson and M. Wilkins, medecine 1962: *for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material*
  - C. Anfinsen, chemistry 1972: *for his work on ribonuclease, especially concerning the connection between the amino acid sequence and the biologically active conformation*
  - K. Wutricht and J. Fenn, chemistry 2002: *for the development of methods for identification and structure analyses of biological macromolecules*
  - R. Kornberg, chemistry 2006: *for his studies of the molecular basis of eukaryotic transcription*
  - V. Ramakrishnan, T. Steitz, A. Yonath, chemistry 2009: *for studies of the structure and function of the ribosome*
  - M. Karplus, M. Levitt, A. Warshell, chemistry 2013: *for the development of multiscale models for complex chemical systems*

Methods: molecular simulation

The Nobel Prize in Chemistry 2013
Martin Karplus, Michael Levitt, Arieh Warshel

The Nobel Prize in Chemistry 2013

© Harvard University
Martin Karplus

Photo: © S. Fisch
Michael Levitt

Photo: Wikimedia Commons
Arieh Warshel

The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.

Connexions between my talk and Prof. Boissonnat's course

C1: Modéles géométriques discrets
→ Voronoi models in various guises

C2: La puissance de l'aléa
→ Randomized constructions, Monte Carlo algorithms

C3: Le calcul géométrique
→ Robust geometric predicates and constructions
The Computational Geometry Algorithms Library – code and spirit!

C4. Génération de maillages
→ The Poisson-Boltzmann equation

C5: Courbes et surfaces
→ Surface / shape reconstruction
→ Convergence of regressors

C6: Espaces de configurations
→ Conformational spaces: exploration, planning

C7. Structures de données géométriques
→ Geometric approximation theory, geometric optimization

C8: Analyse géométrique et topologique des données
→ Topological persistence, geometric/topological data analysis

References


F. Cazals, H. Kanhere, and S. Loriot. Computing the volume of union of balls: a certified algorithm. *ACM Transactions on Mathematical Software*, 2011.


F. Cazals, F. Proust, R. Bahadur, and J. Janin. Revisiting the Voronoi description of protein-protein interfaces. *Protein Science*, 15(9), 2006.


S. Marillet, M-P. Lefranc, P. Boudinot, and F. Cazals. Dissecting interfaces of antibody - antigen complexes. . . *Frontiers in immunology*, 34(8), 2017.


S. Marillet, P. Boudinot, and F. Cazals. High resolution crystal structures leverage protein binding affinity predictions. *Proteins: structure, function, and bioinformatics*, 1(84), 2015.


F. Cazals, T. Dreyfus, D. Mazauric, A. Roth, and C.H. Robert. Conformational ensembles and sampled energy landscapes: Analysis and comparison. *J. of Computational Chemistry*, 36(16), 2015.


A. Roth, T. Dreyfus, C.H. Robert, and F. Cazals. Hybridizing rapidly growing random trees . . . improved exploration of energy landscapes. *J. of Computational Chemistry*, 37(8), 2016.


J. Carr, D. Mazauric, F. Cazals, and D. J. Wales. Energy landscapes and persistent minima. *The Journal of Chemical Physics*, 144(5), 2016.


F. Cazals and D. Mazauric. Optimal transportation problems with connectivity constraints. Inria Research Report 8991, 2016.


F. Cazals and T. Dreyfus. The Structural Bioinformatics Library: modeling in biomolecular science and beyond. *Bioinformatics*, 1–8, 2016.

Acknowledgments

- Former Inria director and mentor
  - Gilles Kahn (†)
- Colleagues and students
  - Jean-Daniel Boissonnat
  - Joël Janin
  - Sylvain Pion (software engineering), Dorian Mazauric (algorithms)
  - Charles Robert (biophysics), Pierre Boudinot (immunology), Félix Rey (virology)
  - PhD students: Sébastien Loriot, Tom Dreyfus, Andrea Roth, Simon Marillet, Augustin Chevallier, Romain Tetley
- Inria ... Algorithms-Biology-Structure is already 10 years old