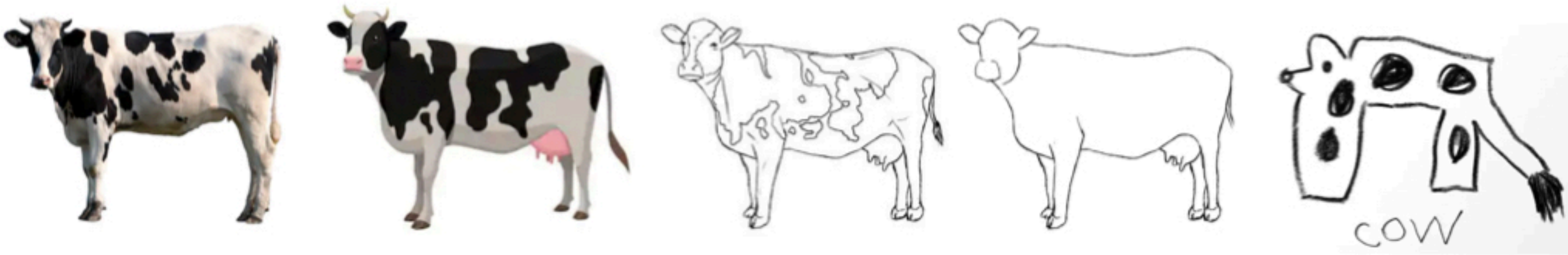
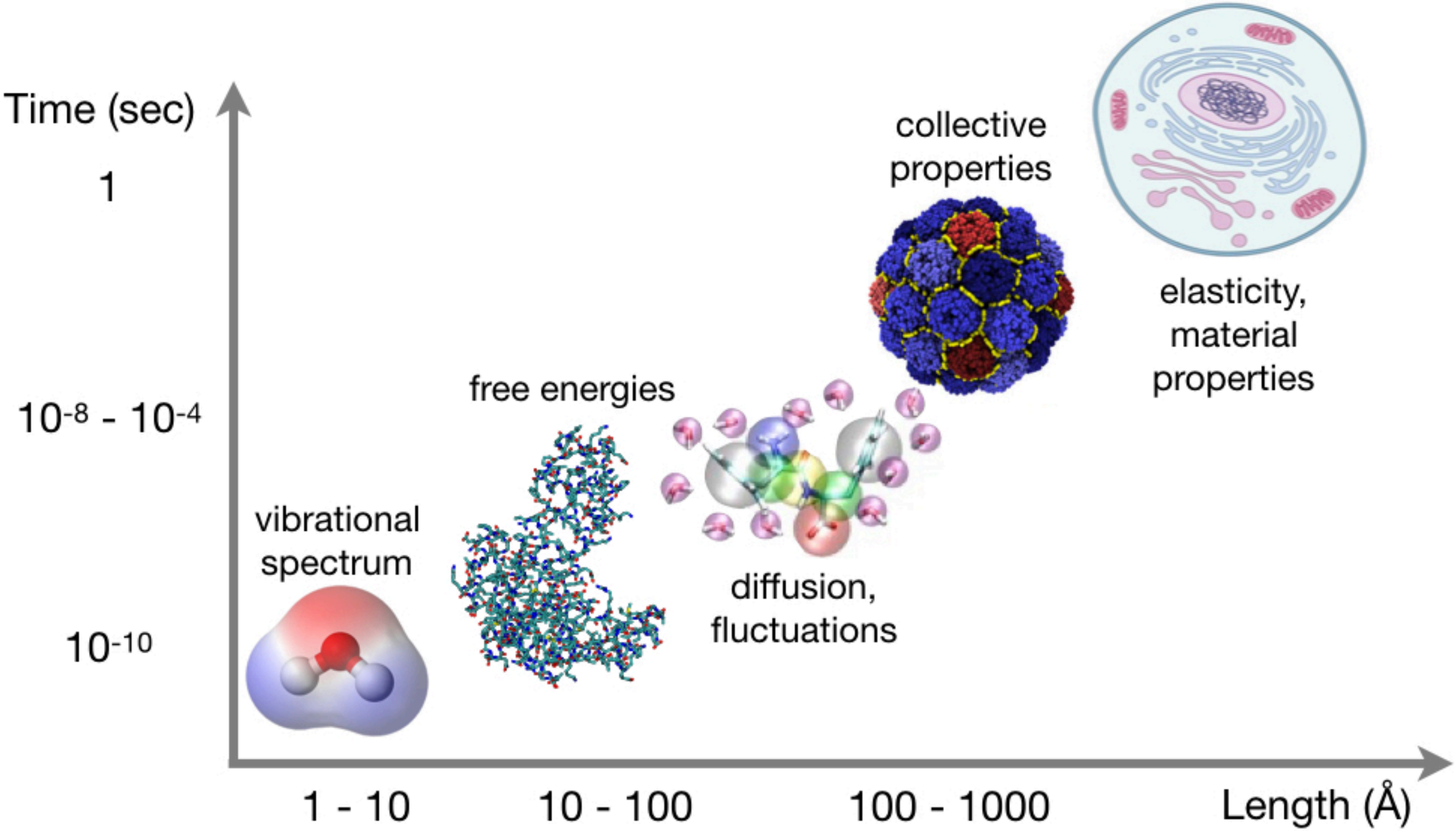
Exploring the crossroad between information theory, statistical mechanics, and biology



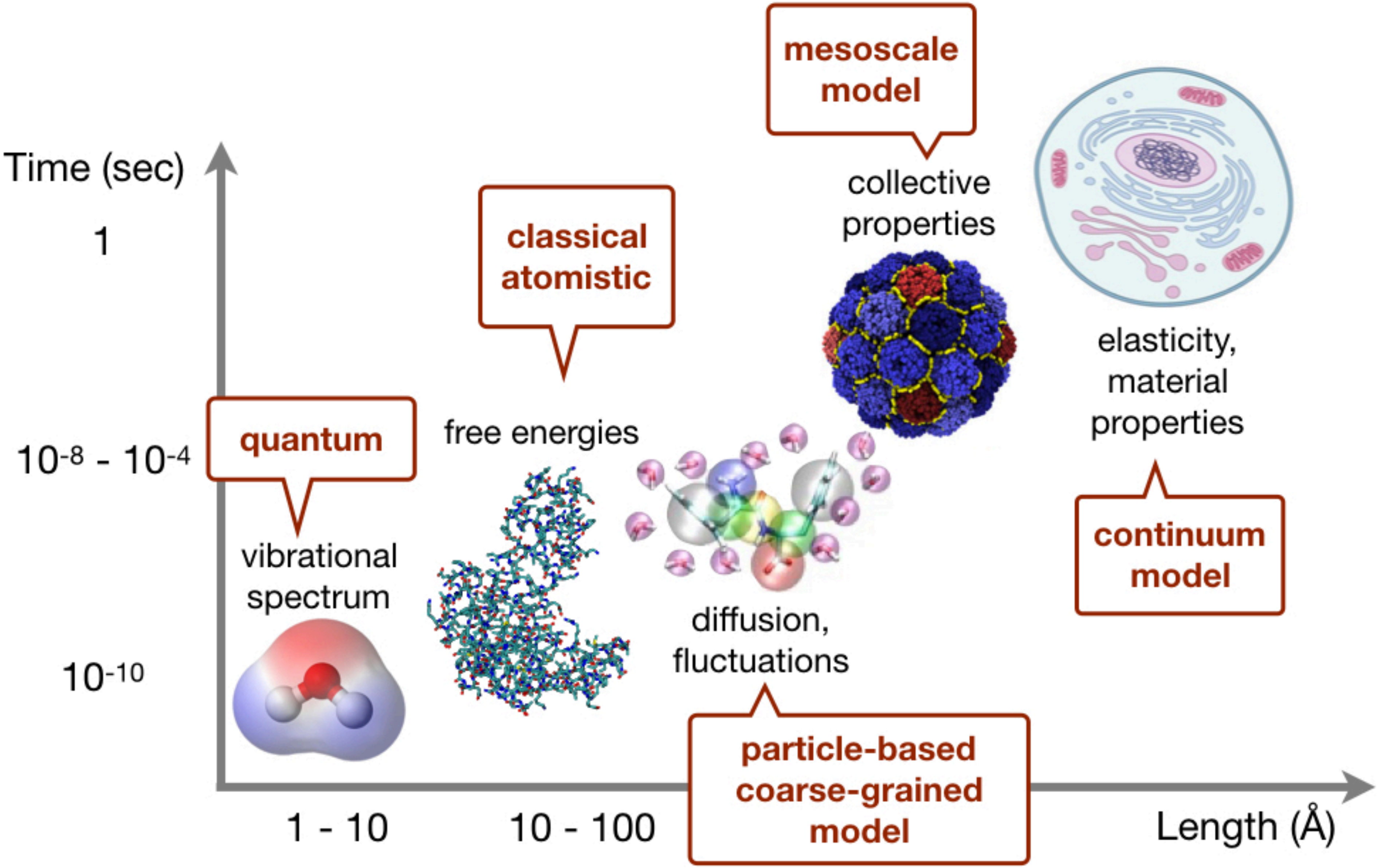
Raffaello Potestio
University of Trento
Physics Department

Gran Sasso Science Institute
June 24, 2020

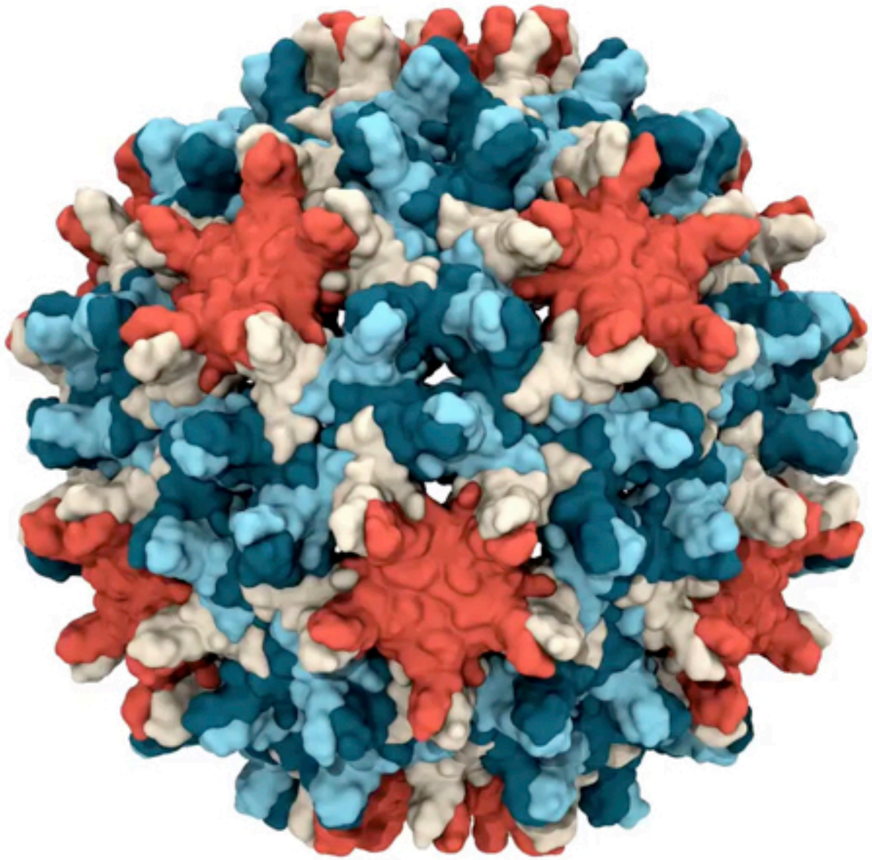
The multiscale challenge in biological systems

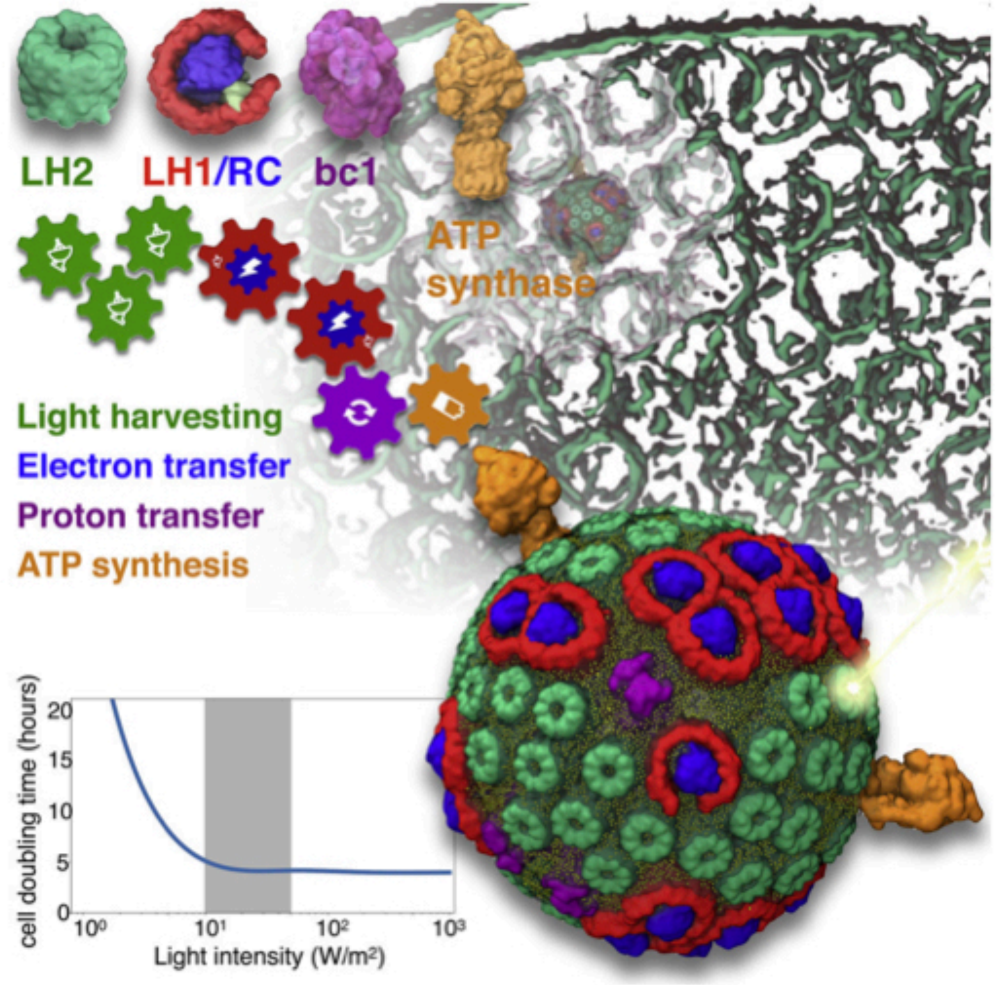


The multiscale challenge in biological systems



The current frontier of large-scale MD simulation



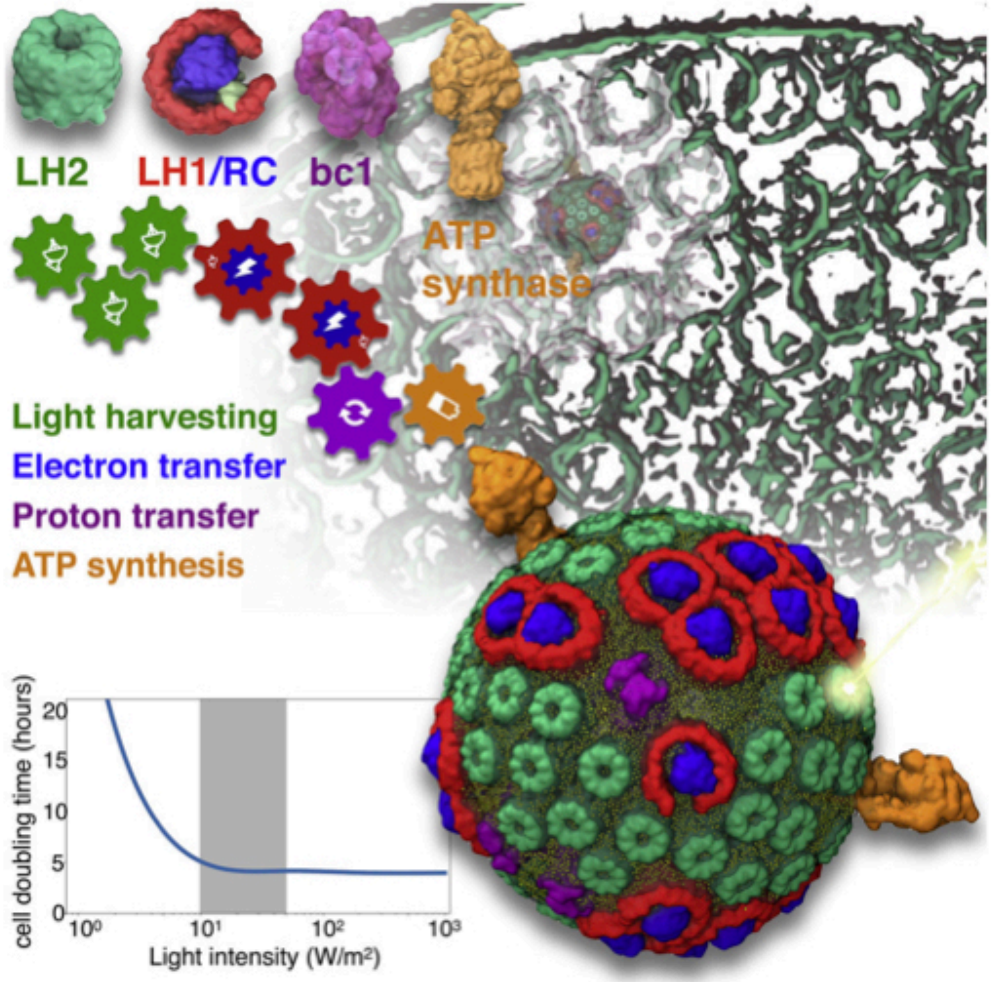


Hadden *et al.*, eLife 2018

Singharoy *et al.*, Cell 2019

The current frontier of large-scale MD simulation

GPU clusters for HPC



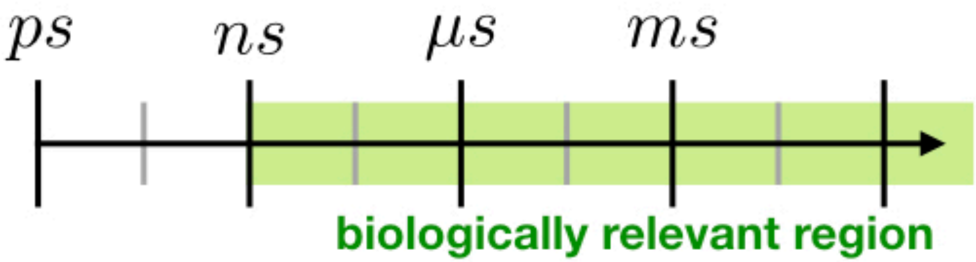
Singharoy *et al.*, Cell 2019

The current frontier of large-scale MD simulation

GPU clusters for HPC

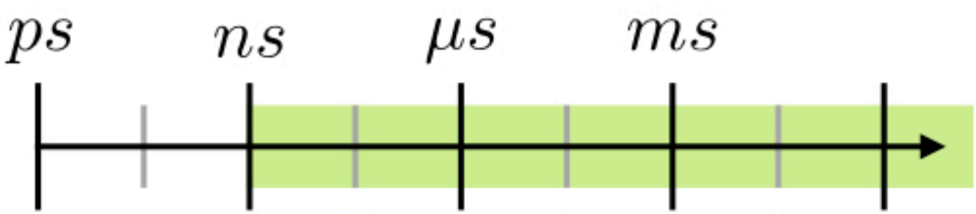
ANTON supercomputer, D.E. Shaw research

The current frontier of large-scale MD simulation



ANTON supercomputer, D.E. Shaw research

The current frontier of large-scale MD simulation

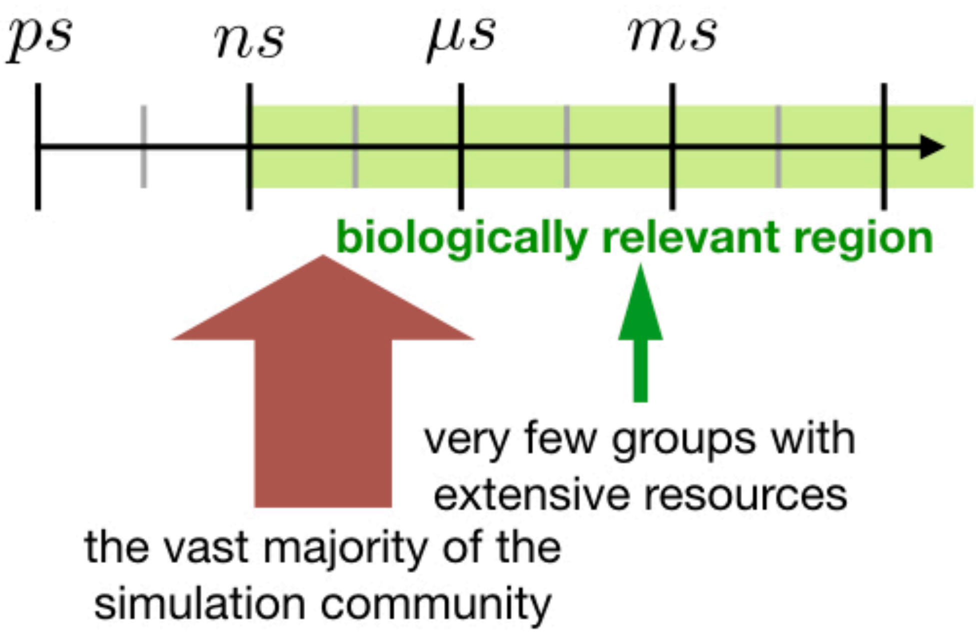


biologically relevant region

very few groups with
extensive resources

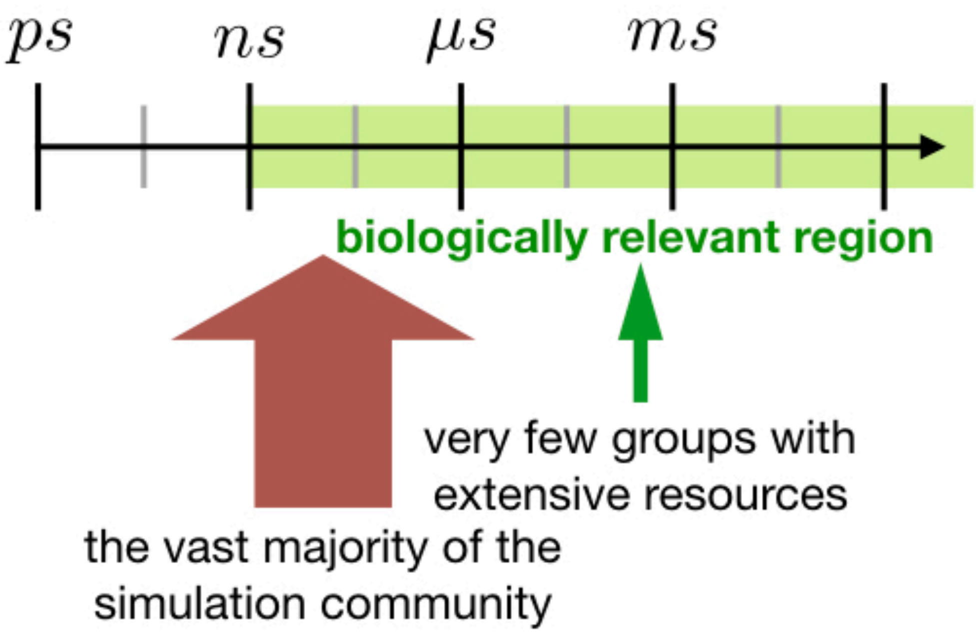
ANTON supercomputer, D.E. Shaw research

The current frontier of large-scale MD simulation



ANTON supercomputer, D.E. Shaw research

The current frontier of large-scale MD simulation



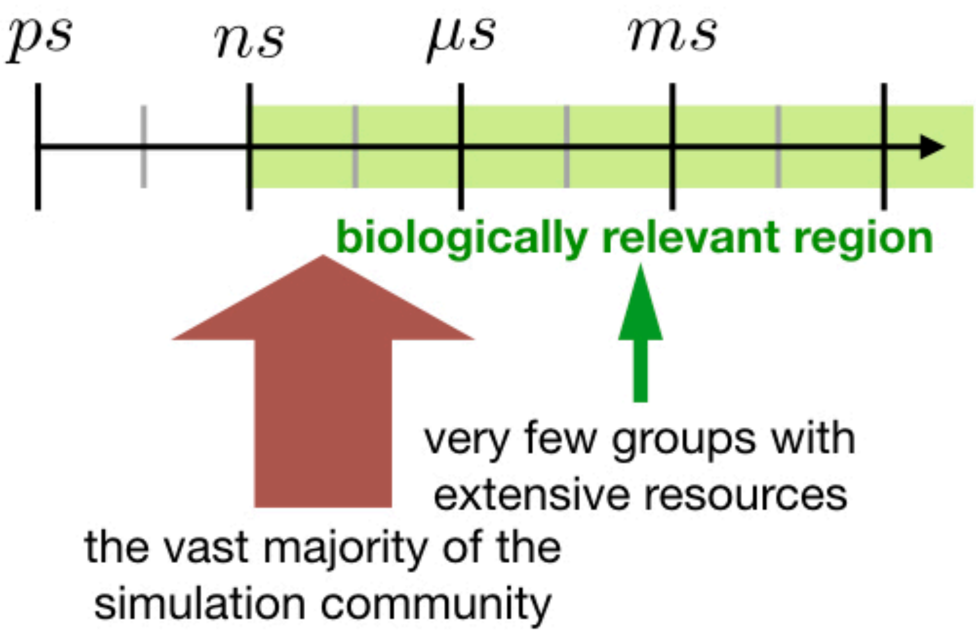
Need for massive computational resources

Bound to one system/setup due to cost

Huge amount of data, difficult storage/analysis

ANTON supercomputer, D.E. Shaw research

The current frontier of large-scale MD simulation



Need for massive computational resources

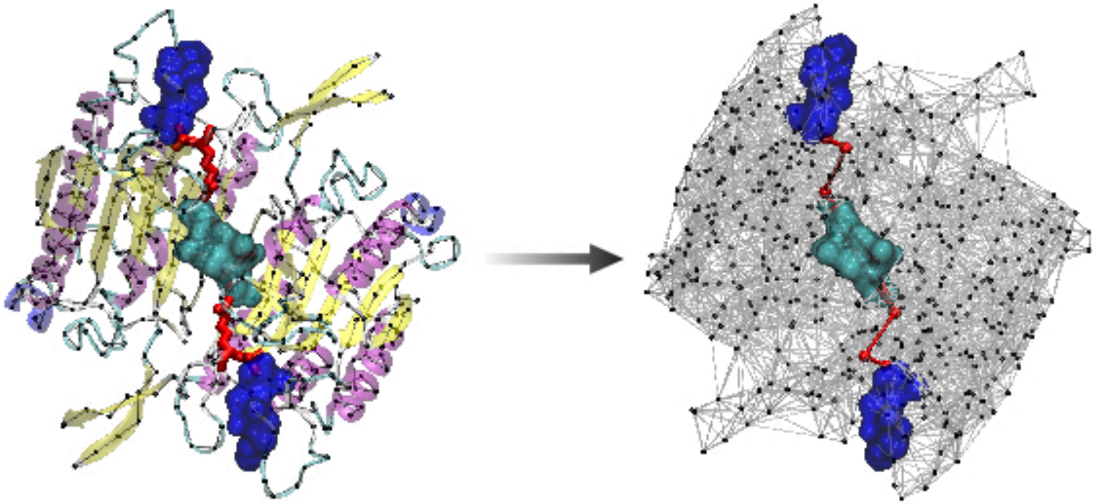
Bound to one system/setup due to cost

Huge amount of data, difficult storage/analysis

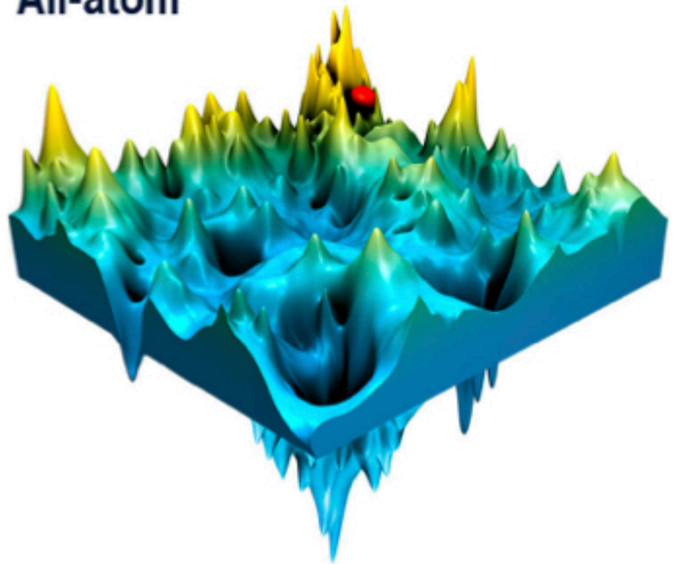
All-atom model required,
but coarse-grained information desired

ANTON supercomputer, D.E. Shaw research

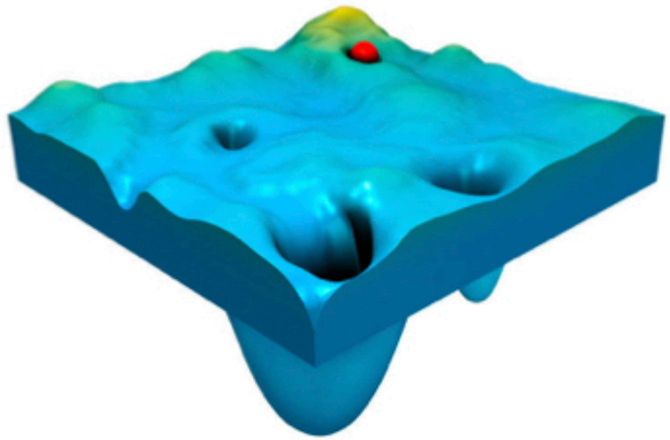
Coarse-grained models: larger, faster. Better?



All-atom

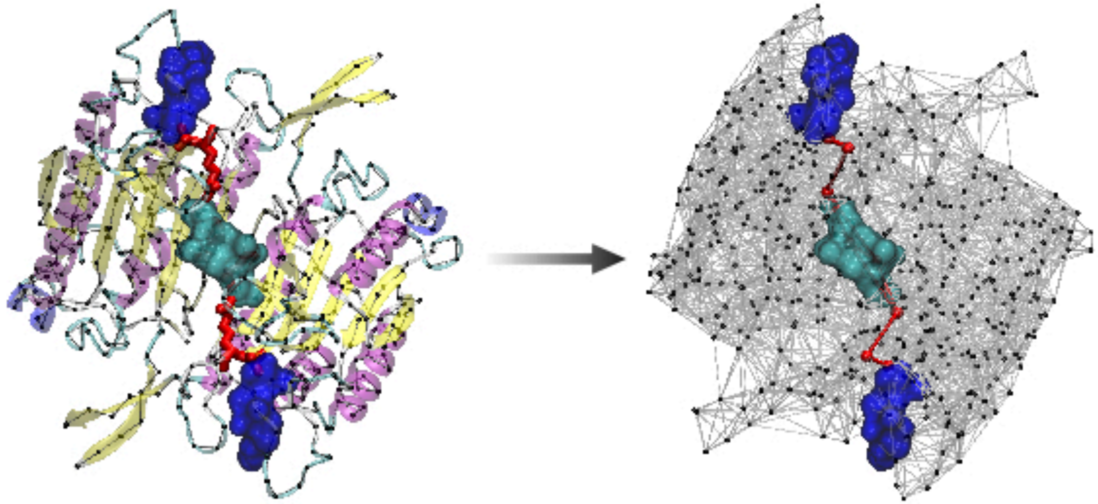


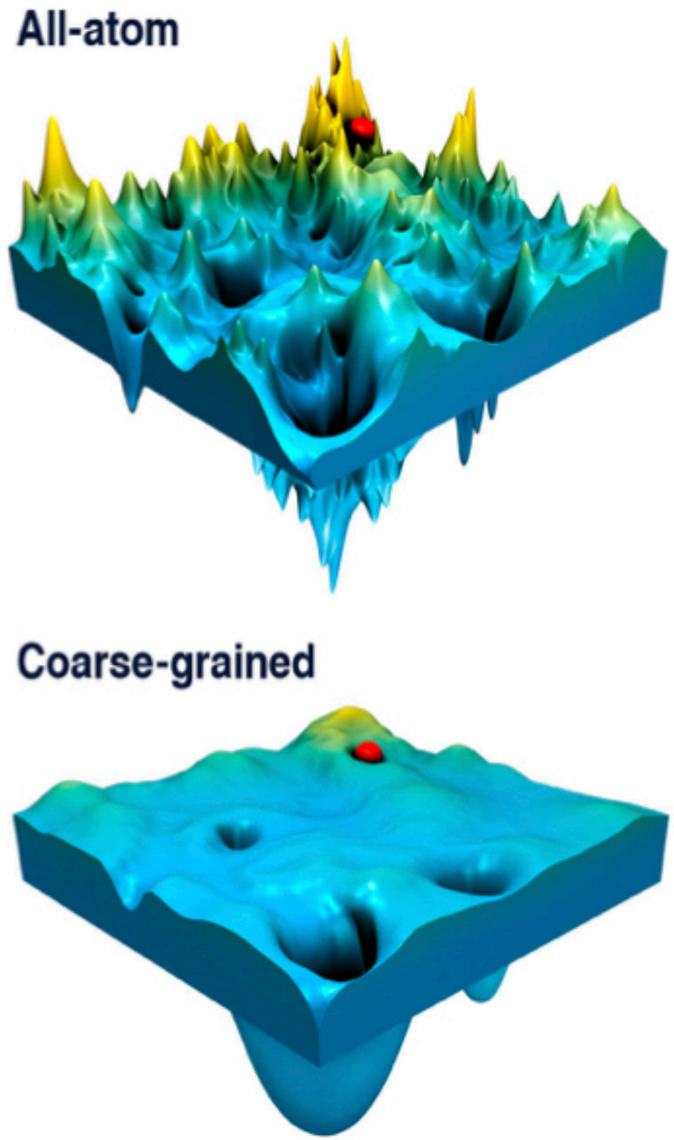
Coarse-grained



- Reduced number of DoFs
- Short-range interactions
- Faster dynamics
- Larger integration time steps
- Vast model library

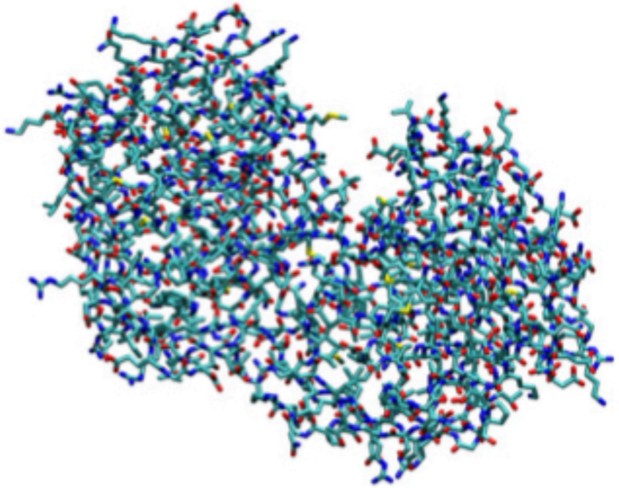
Coarse-grained models: larger, faster. Better?





Intrinsically low resolution
Only long-wavelength properties
No chemistry-dependence
No multi-scale processes

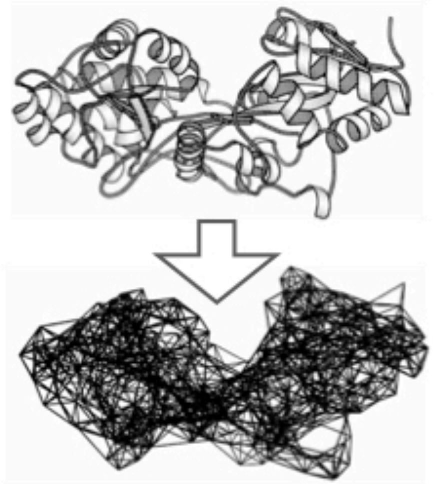
The zoo of CG protein models



all-atom

$$E_p = \frac{1}{2}\sum_{\text{bonds}} K_b(b-b_0)^2 + \frac{1}{2}\sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \frac{1}{2}\sum_{\text{dihedrals}} K_\phi[1+\cos(n\phi-\delta)] + \sum_{\text{non bonded pairs}}\left[\frac{A}{r^{12}} - \frac{B}{r^6} + \frac{q_1 q_2}{Dr}\right]. \quad (1)$$

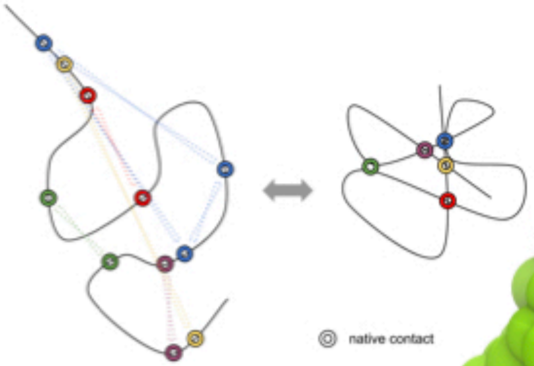
$$E(\mathbf{r}_a, \mathbf{r}_b) = \frac{C}{2}\left(\frac{\mathbf{r}^0_{a,b}\cdot\Delta\mathbf{r}_{a,b}}{|\mathbf{r}^0_{a,b}|}\right)^2$$



elastic network models

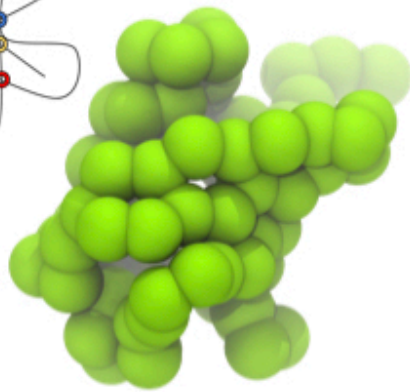
M. Tirion, PRL 1996

Figure by YH Sanejouand, from Monticelli and Saonen, Springer



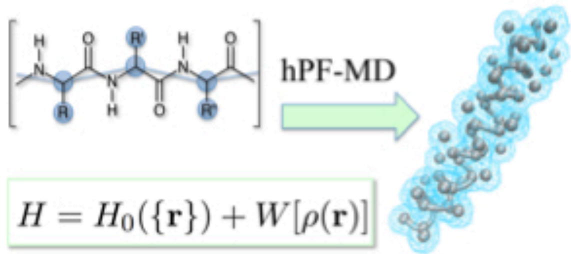
Go models

Taketomi *et al.*, Int J Pept Protein Res. 1975



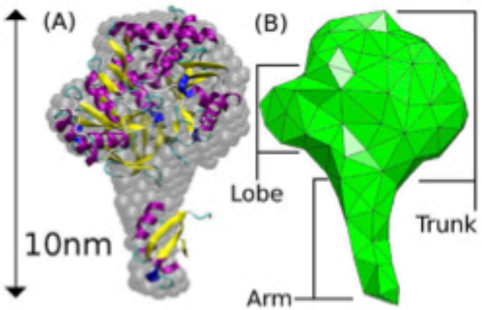
MARTINI

Monticelli *et al.*, J. Chem. Theory Comput. 2008



particles+continuum

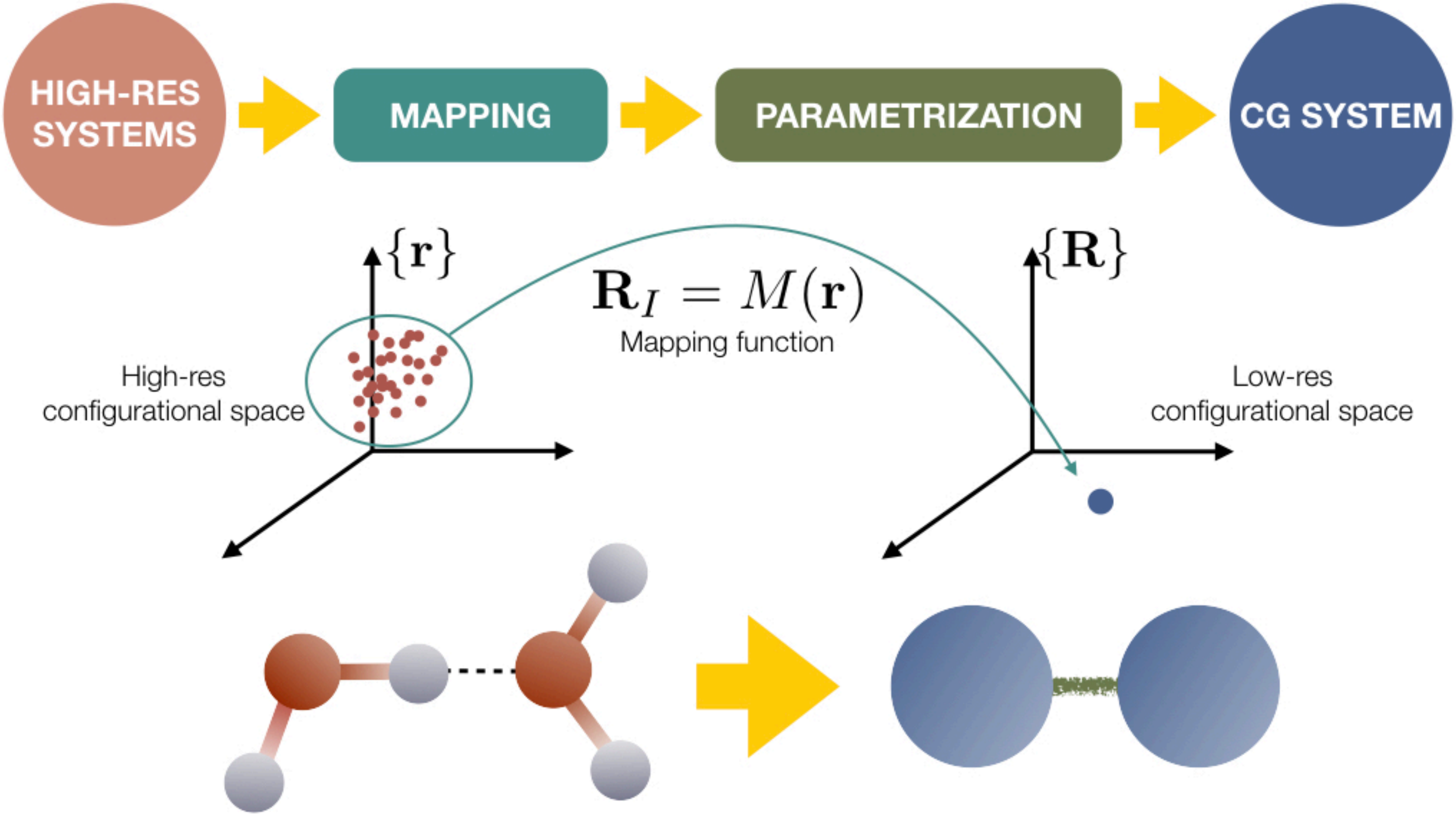
Bore *et al.*, J. Chem. Theory Comput. 2008



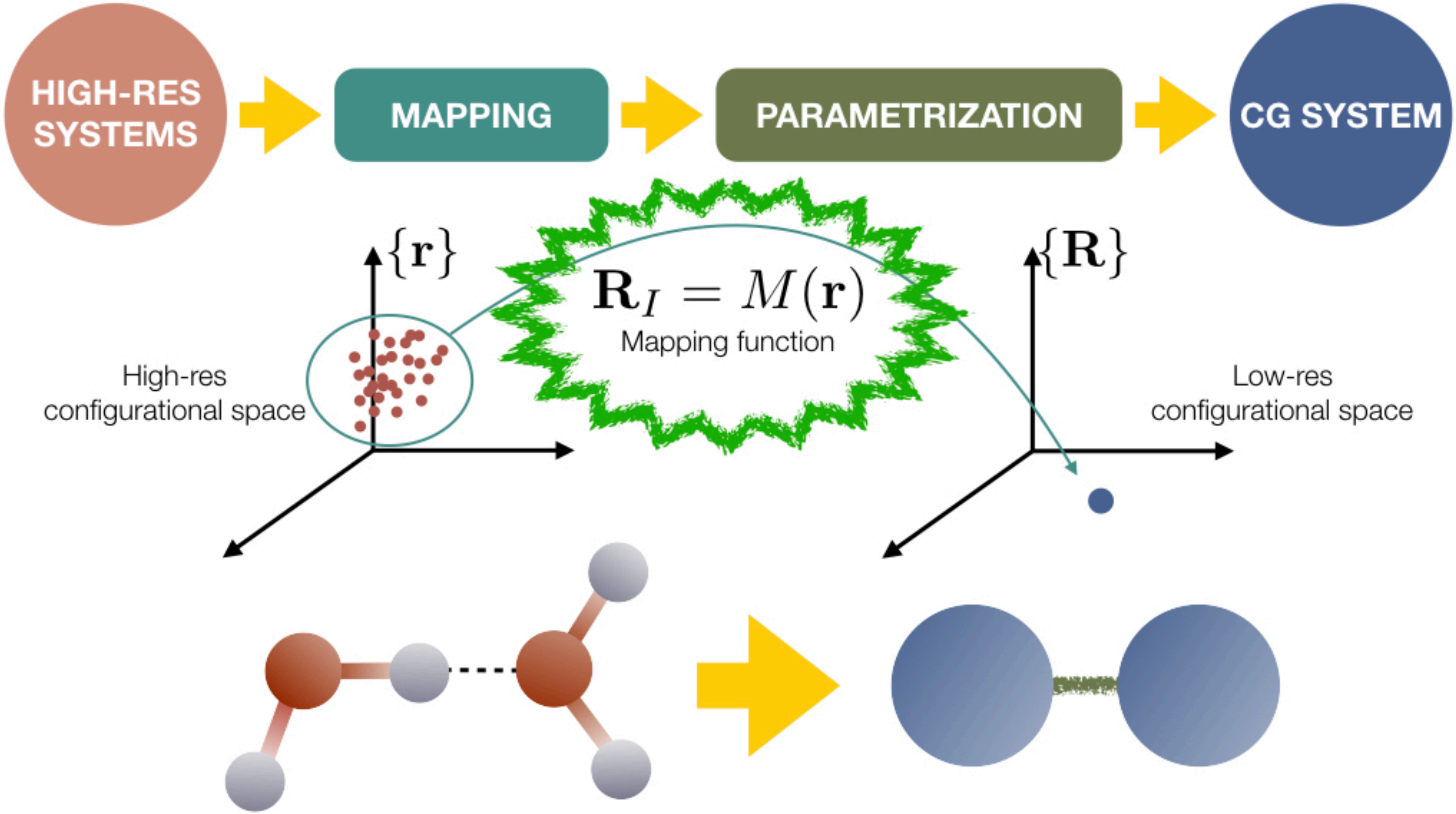
finite elements

Oliver *et al.*, J. Comp. Phys. 2013

Bottom-up (systematic) CG'ing



Bottom-up (systematic) CG'ing

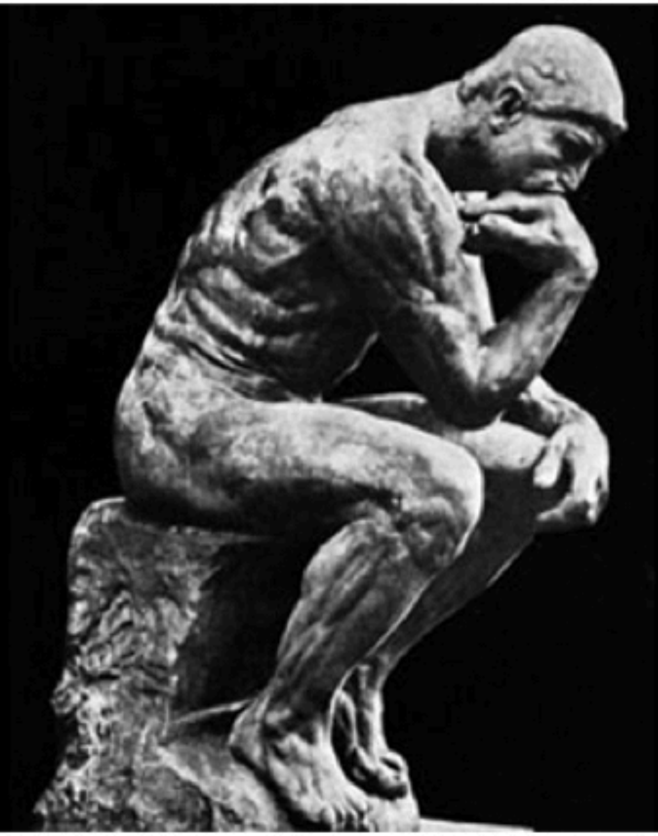


Philosophical (?) questions

Philosophical (?) questions

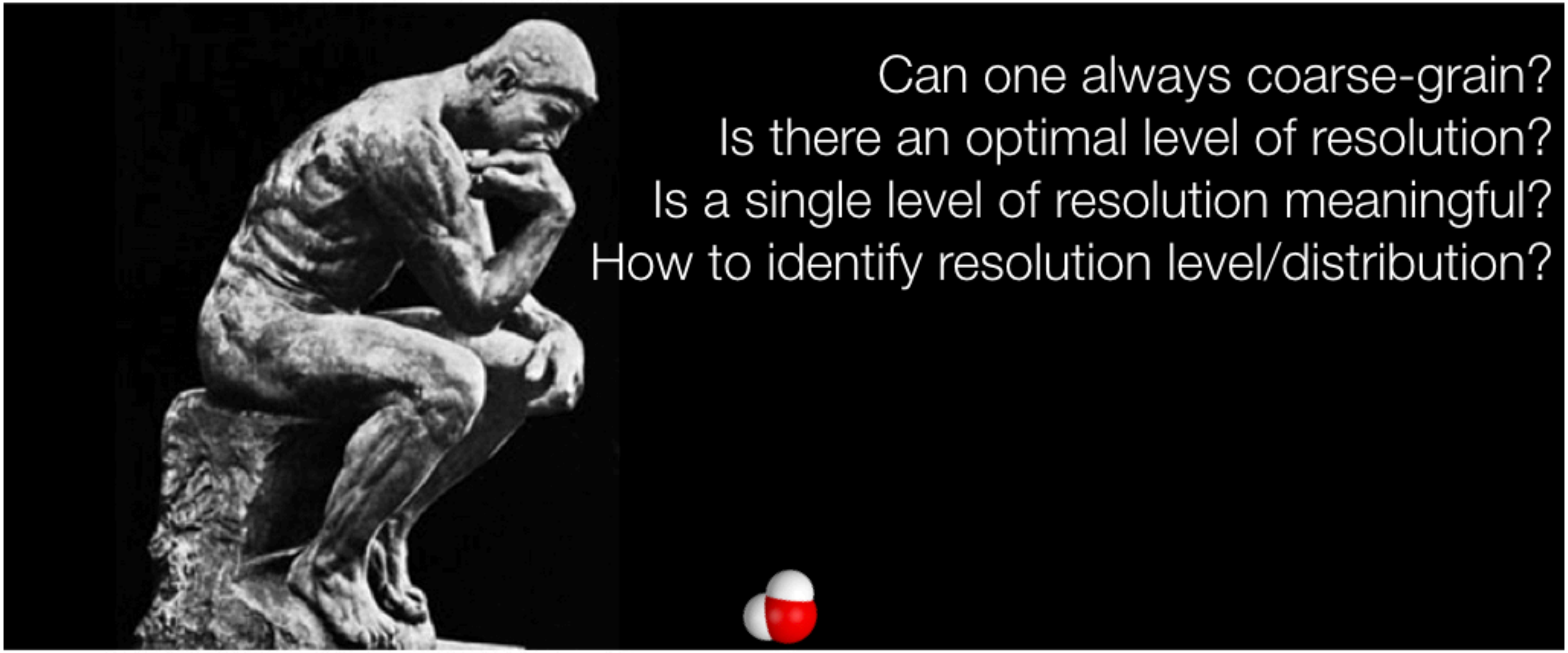
Philosophical (?) questions

Philosophical (?) questions

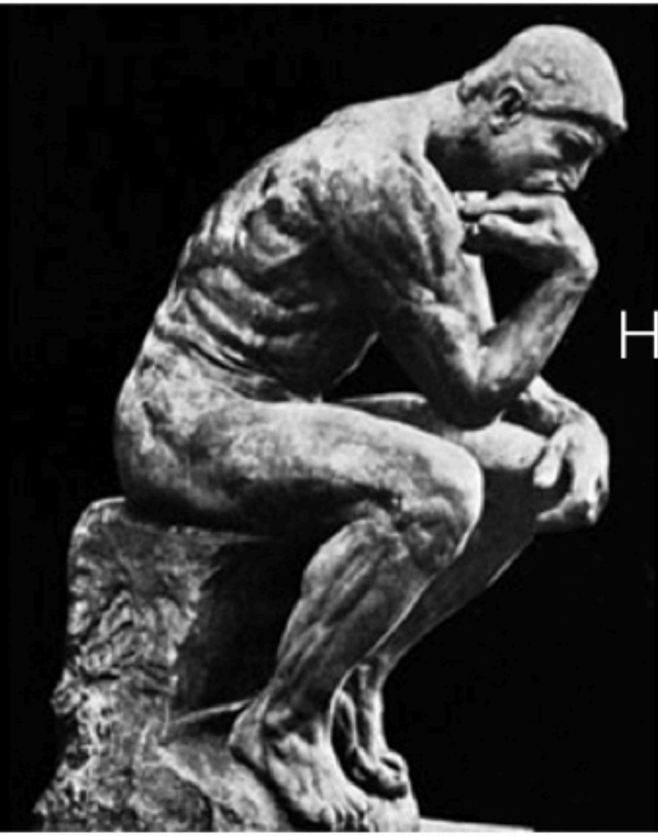


Can one always coarse-grain?
Is there an optimal level of resolution?
Is a single level of resolution meaningful?

Philosophical (?) questions



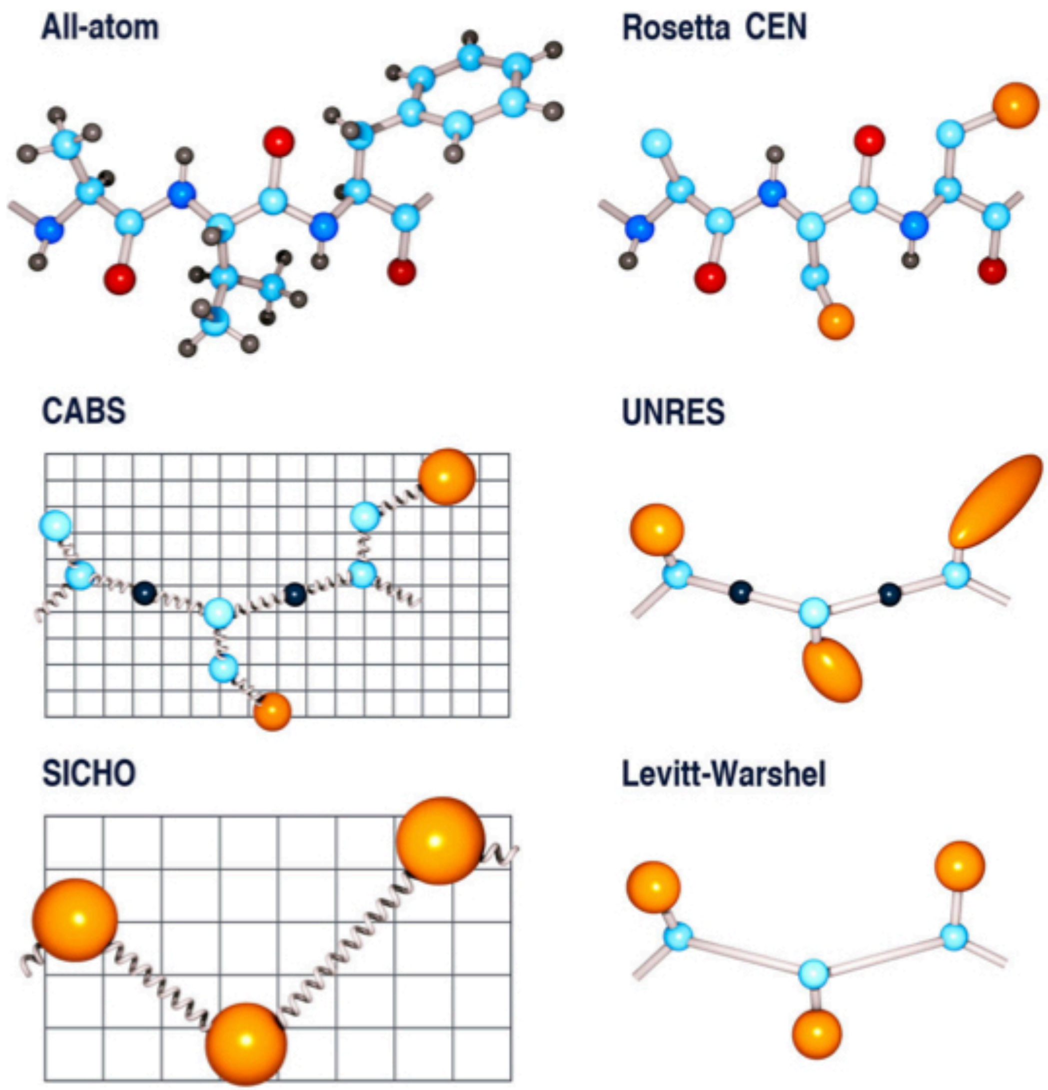
Philosophical (?) questions



Can one always coarse-grain?
Is there an optimal level of resolution?
Is a single level of resolution meaningful?
How to identify resolution level/distribution?
How to implement it in practice?

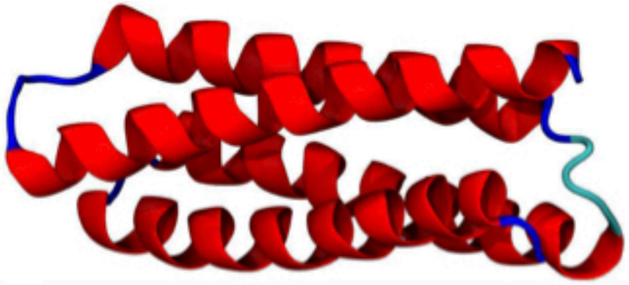
The problem of resolution **level**

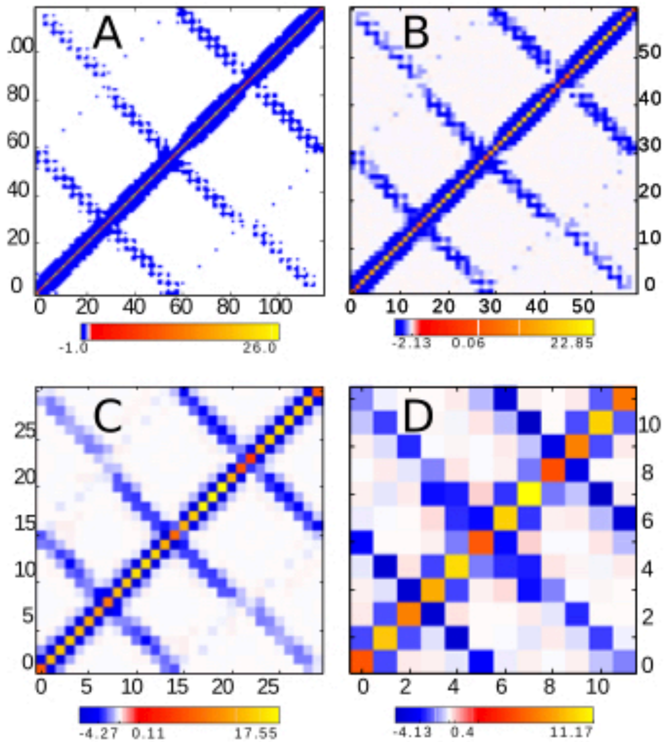
- Atomistic: a bead per atom
- CEN: all-atom backbone, centroid site for side chain
- CABS: Calpha, Cbeta, side chain
- UNRES: united atom residues, ellissoid side chains
- SICHO (side chain only): pseudoatoms placed near the centres of the retained side chains
- Levitt-Warshel: Calpha, Cbeta

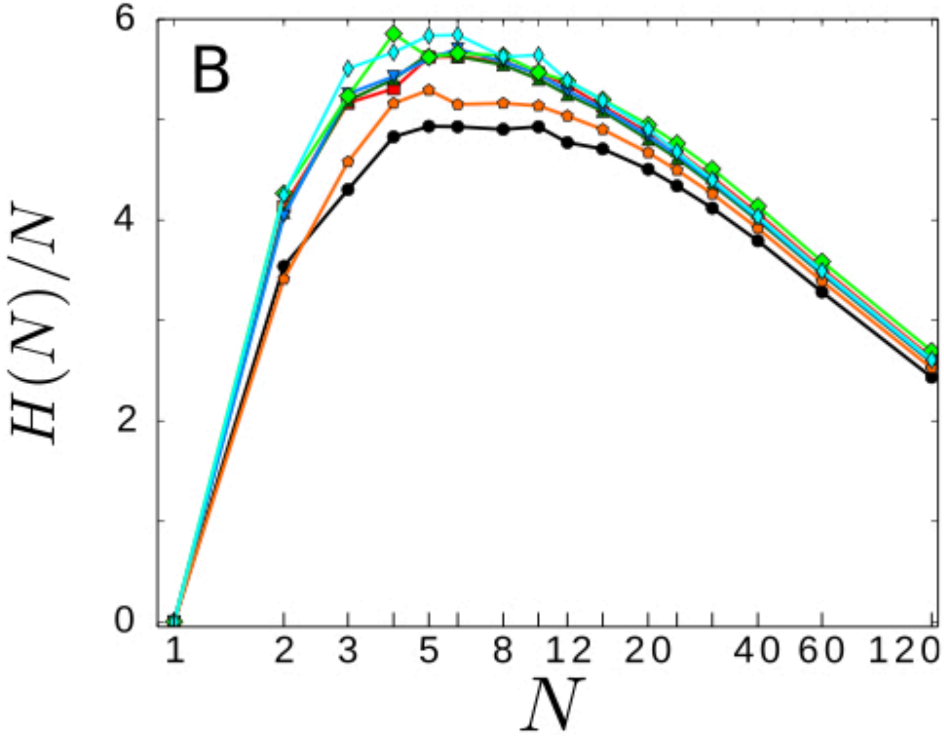


Kmiecik et al., Chem. Rev. 2016, 116, 7898−7936

The problem of resolution **level**



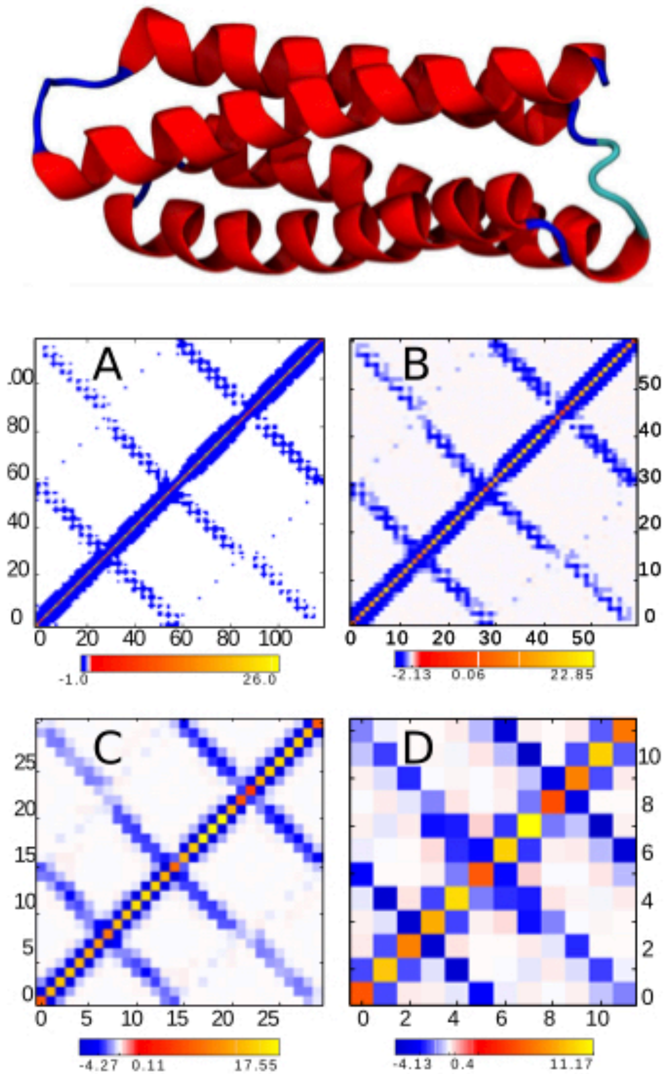


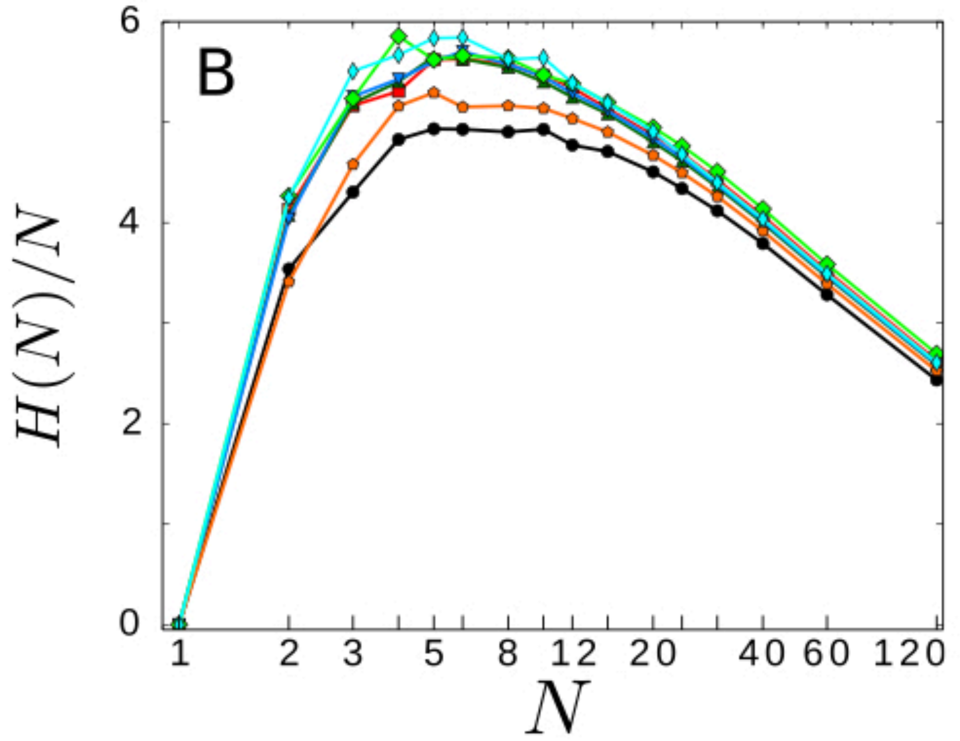


$$H(N) = -(s_R(N) - S_{R;c}(N))/k_B$$

Foley, Shell, Noid, JCP 2015

The problem of resolution **level**





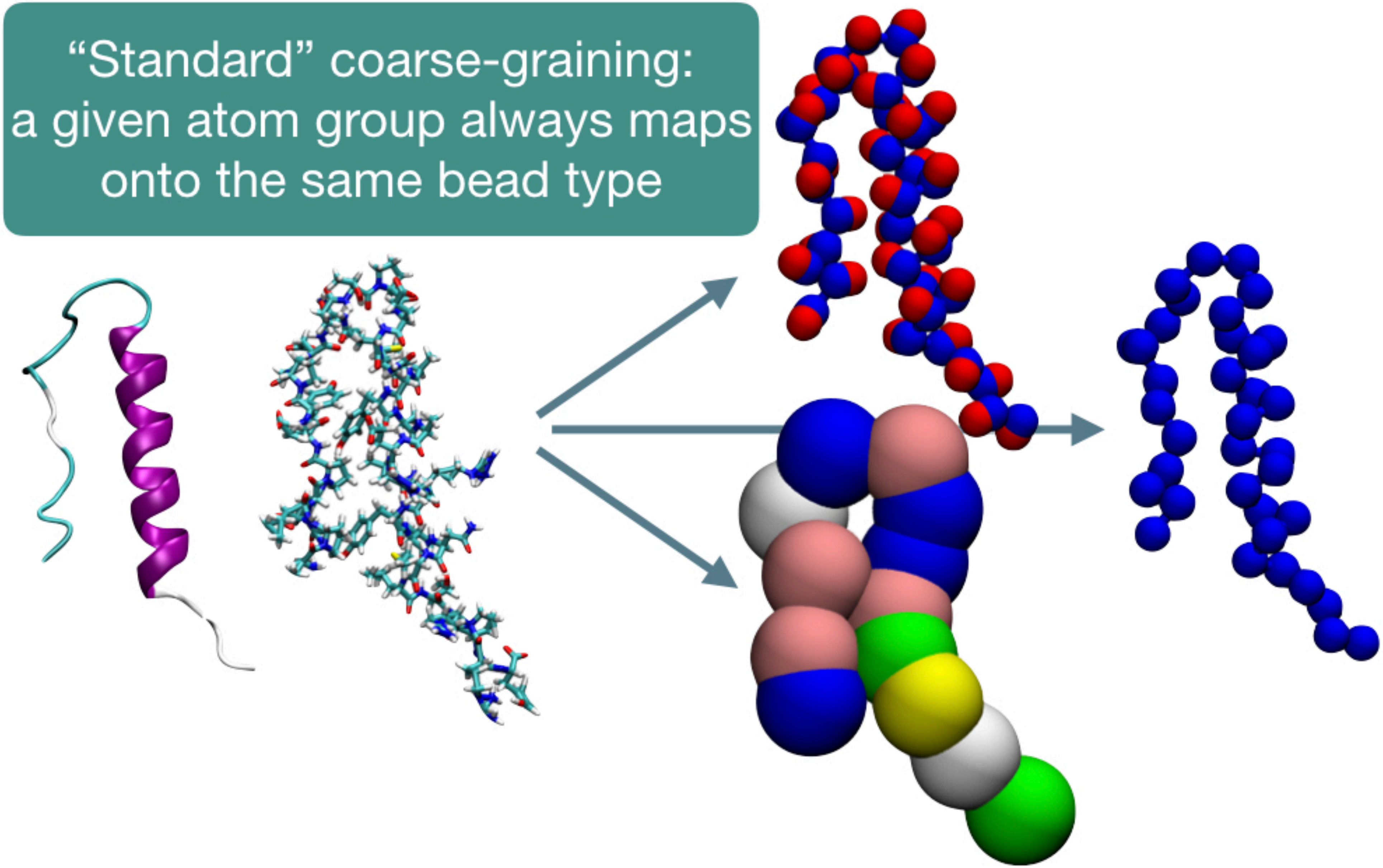
$$H(N) = -(s_R(N) - S_{R;c}(N))/k_B$$

An optimal resolution exists which balances coarsening and preservation of information

Foley, Shell, Noid, JCP 2015

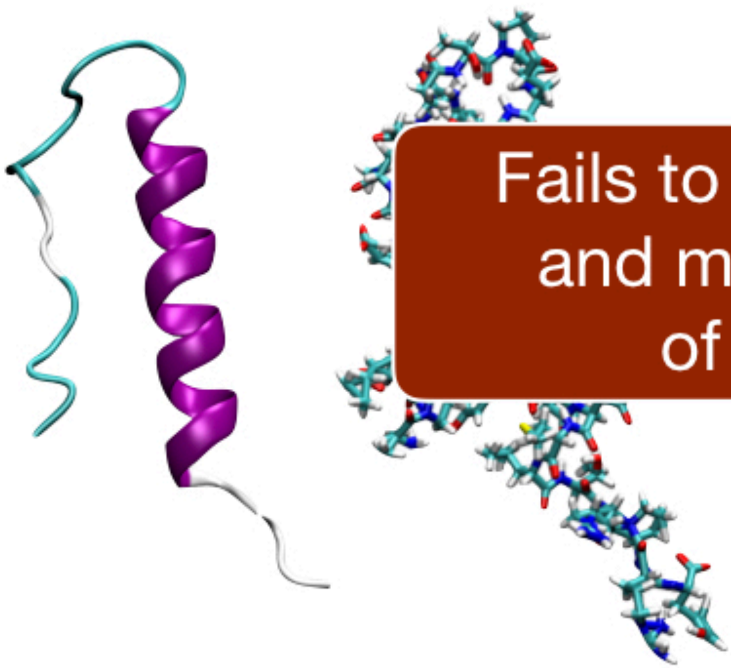
The problem of resolution **distribution**

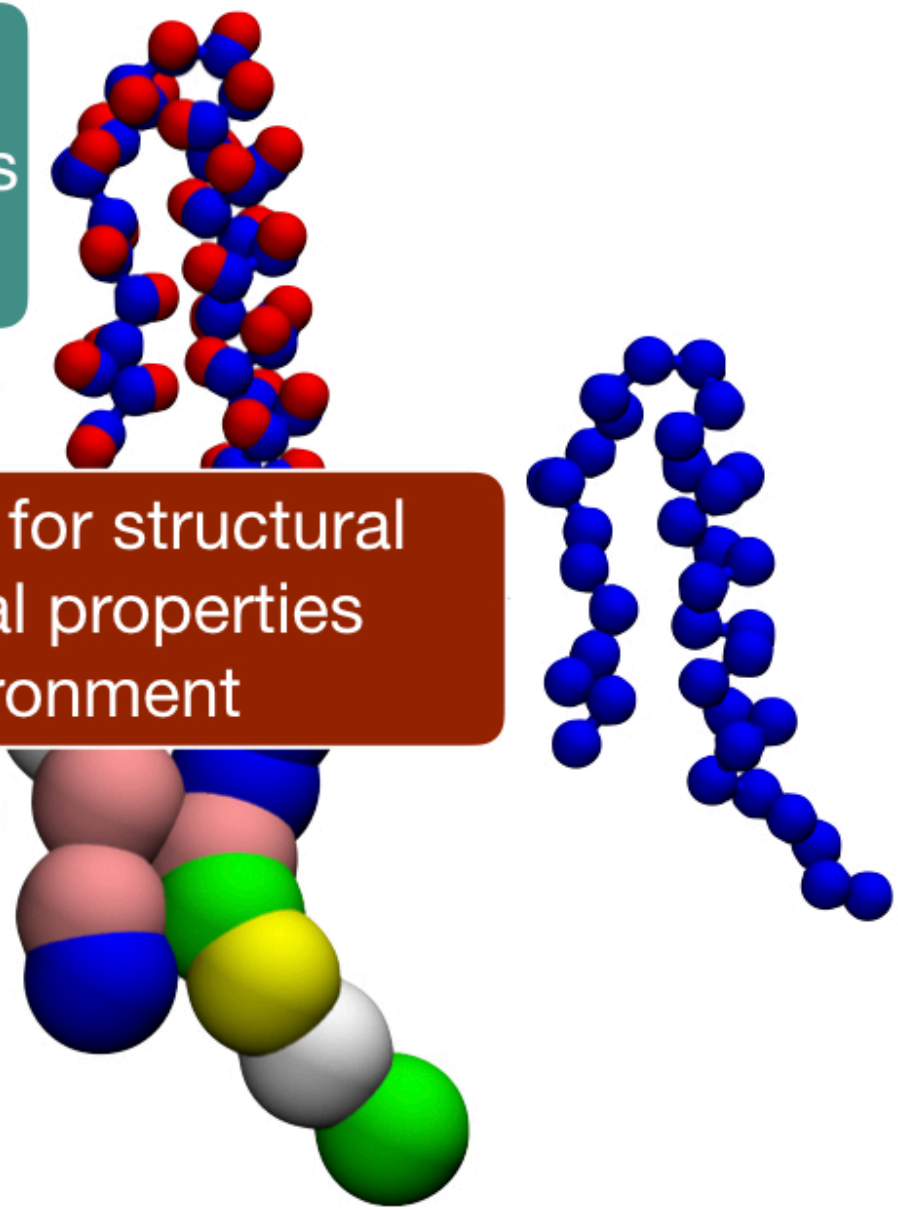
"Standard" coarse-graining:
a given atom group always maps onto the same bead type



The problem of resolution **distribution**

“Standard” coarse-graining:
a given atom group always maps onto the same bead type

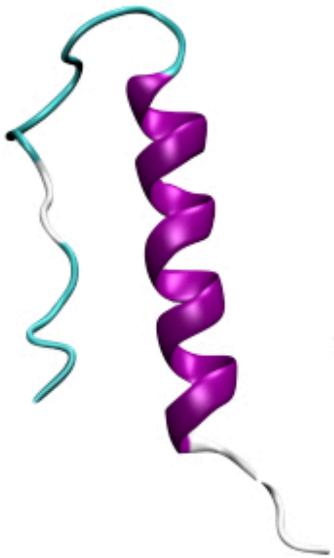


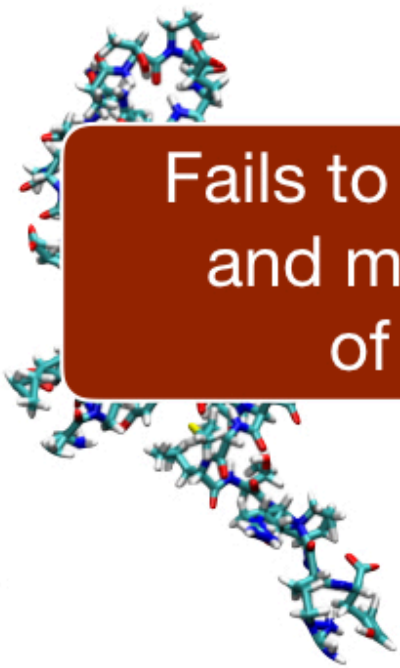


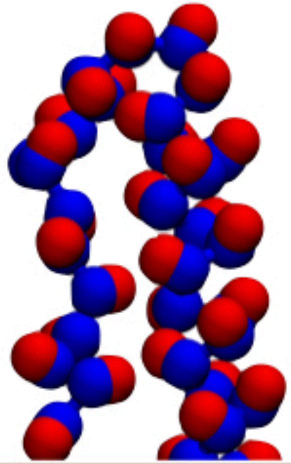
Fails to account for structural and mechanical properties of the environment

The problem of resolution **distribution**

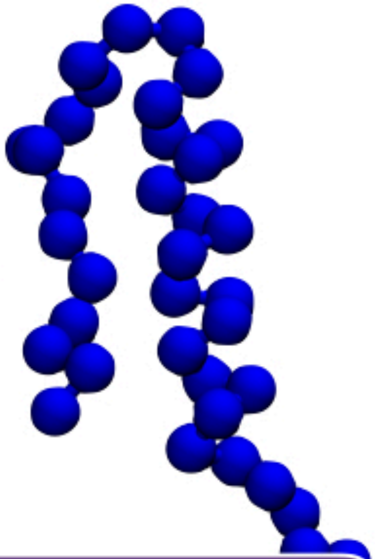
“Standard” coarse-graining:
a given atom group always maps
onto the same bead type





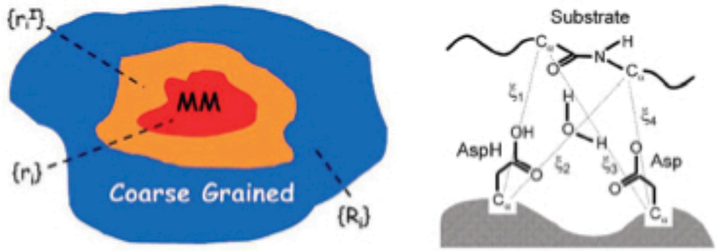


Fails to account for structural
and mechanical properties
of the environment

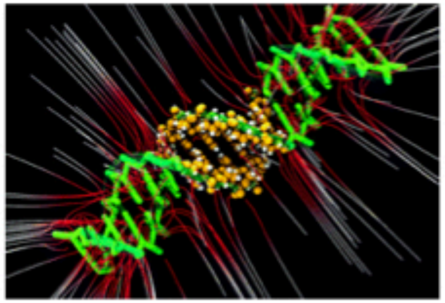


The appropriate level of resolution
depends on local features

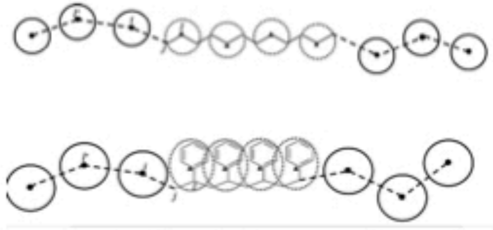
Cut-and-paste macromolecular models



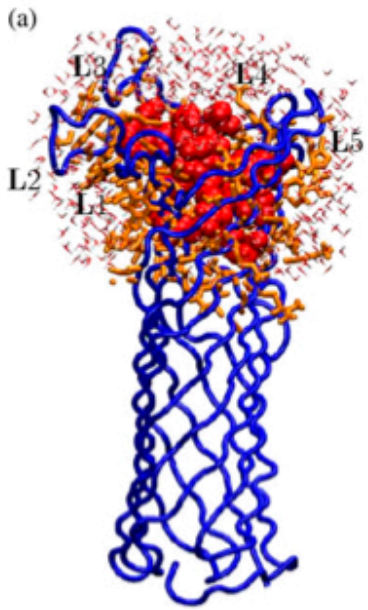
Neri, Anselmi, Cascella, Maritan, Carloni, PRL 2005



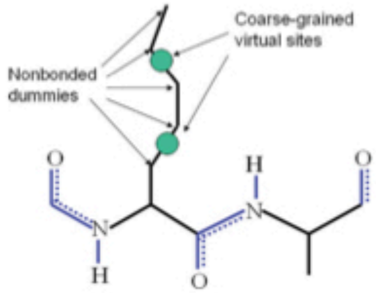
Machado, Dans, Pantano, PCCP 2011



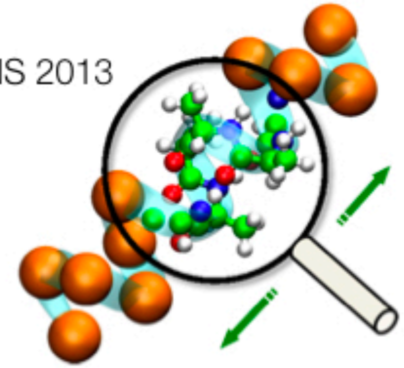
di Pasquale, Marchisio, Carbone, JCP 2012



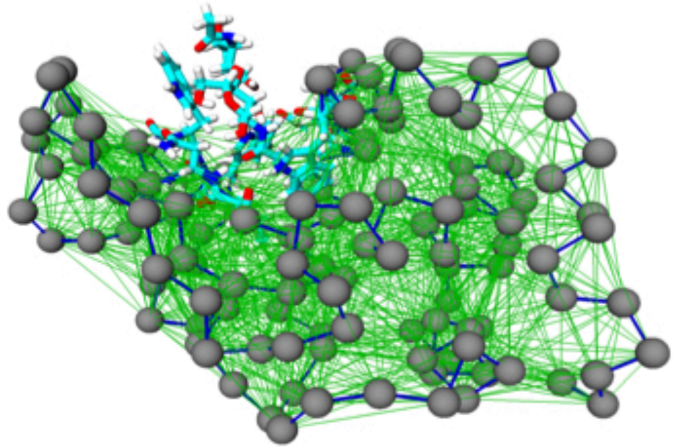
Neri, Baaden, Carnevale, Anselmi, Maritan, Carloni, BJ 2008



Zacharias, PROTEINS 2013

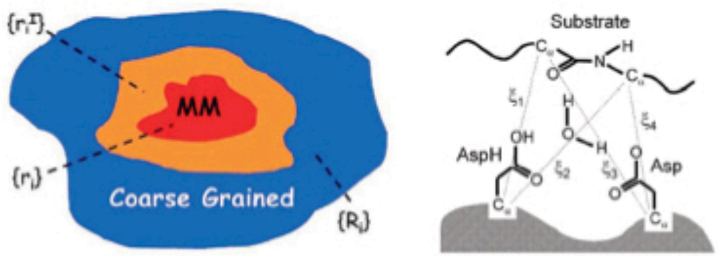


Shen and Hu, JCTC 2014

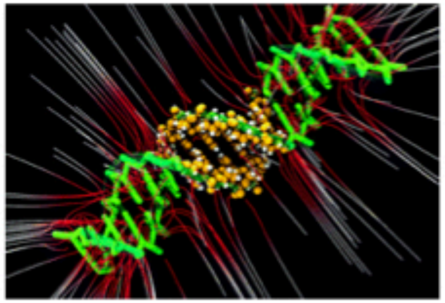


A.C. Fogarty, RP, K. Kremer, Proteins 2016
R. Fiorentini, K. Kremer, RP, Proteins 2020

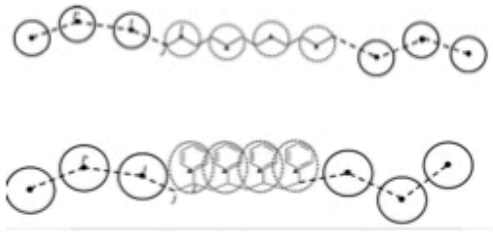
Cut-and-paste macromolecular models



Neri, Anselmi, Cascella, Maritan, Carloni, PRL 2005

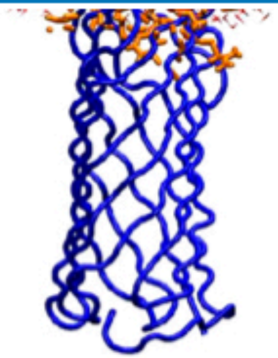


Machado, Dans, Pantano, PCCP 2011



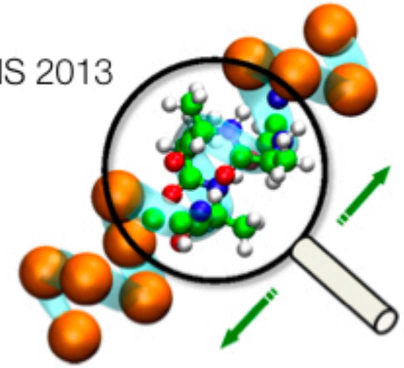
di Pasquale, Marchisio, Carbone, JCP 2012

What do we do if we do not know which parts are important?

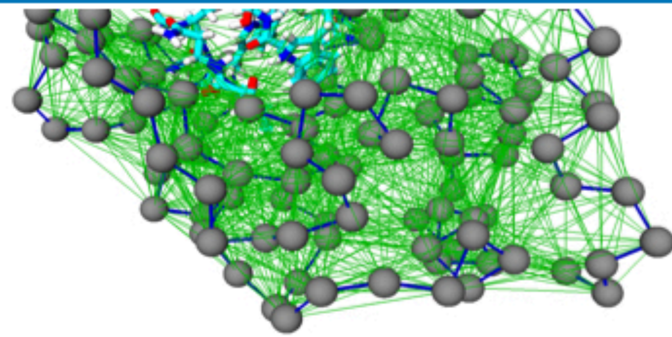


Neri, Baaden, Carnevale, Anselmi, Maritan, Carloni, BJ 2008

Zacharias, PROTEINS 2013



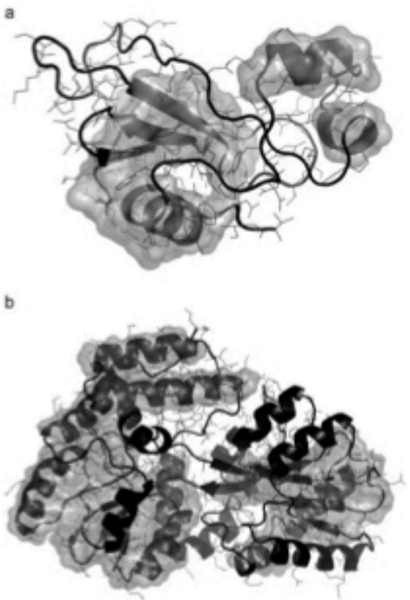
Shen and Hu, JCTC 2014



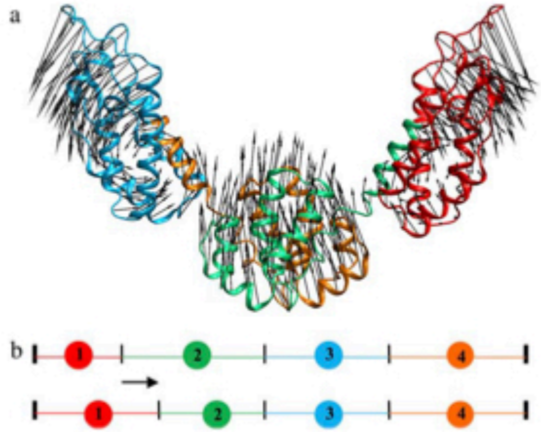
A.C. Fogarty, RP, K. Kremer, Proteins 2016
R. Fiorentini, K. Kremer, RP, Proteins 2020

Emergent coarse-grained representations

- Low res models as probes for mechanical features
- Identify groups of AA's based on collective dynamics
- No obvious definition of a CG "blob"

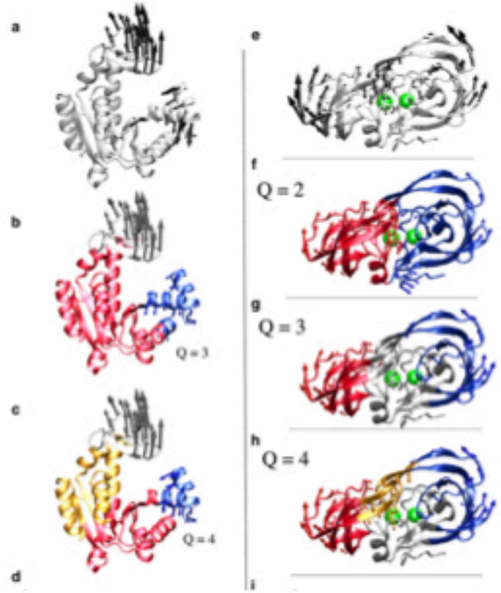


Gohlke and Thorpe
BJ 2006



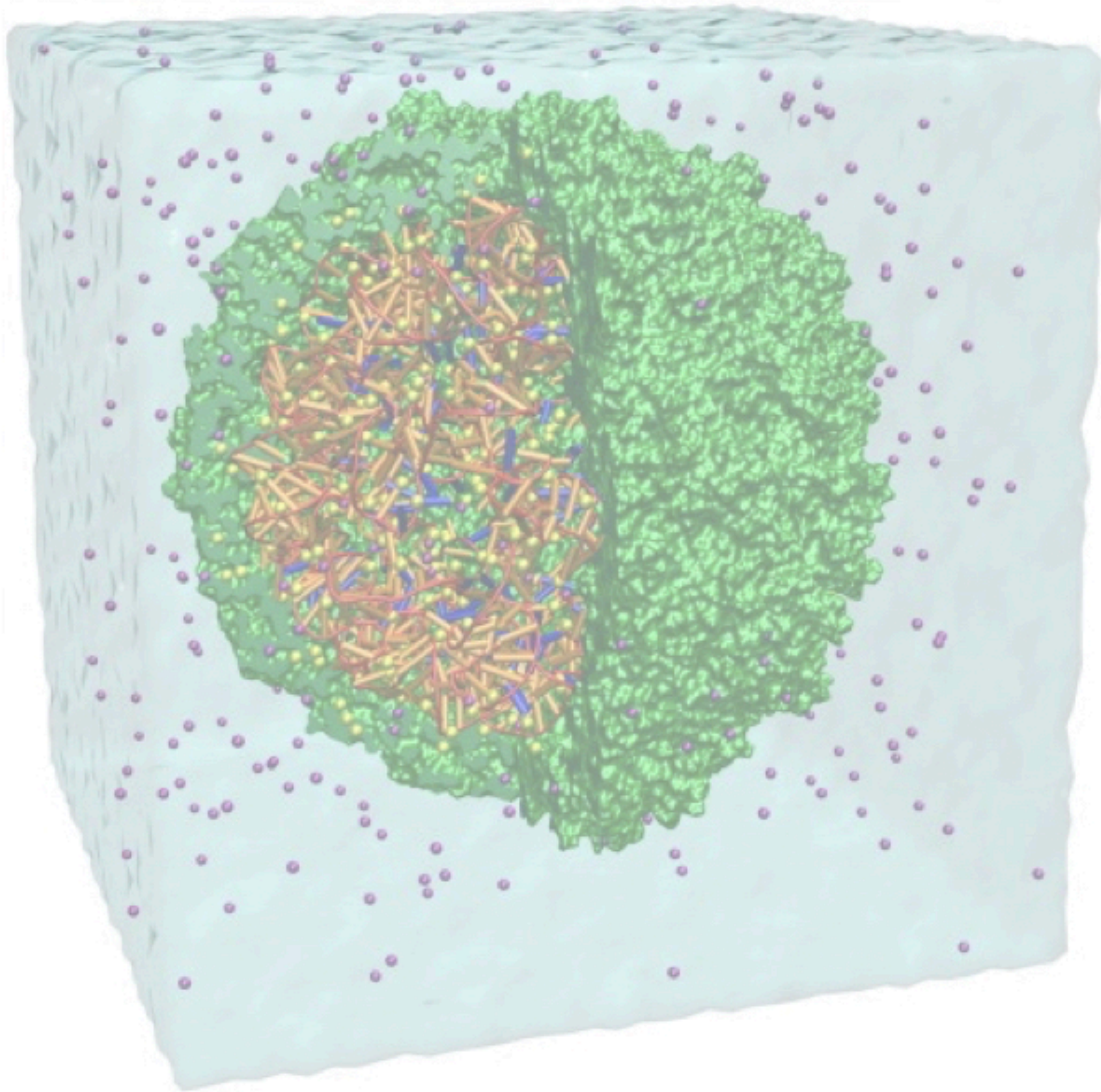
$$\chi^2 = \frac{1}{3N}\sum_{I=1}^{N}\frac{1}{n_t}\sum_{t=1}^{n_t}\left(\sum_{i\in I}\sum_{j\geq i\in I}|\Delta\mathbf{r}_i^{ED}(t) - \Delta\mathbf{r}_j^{ED}(t)|^2\right)$$

Zhang, Lu, Noid, Krishna, Pfaendtner, Voth
BJ 2008



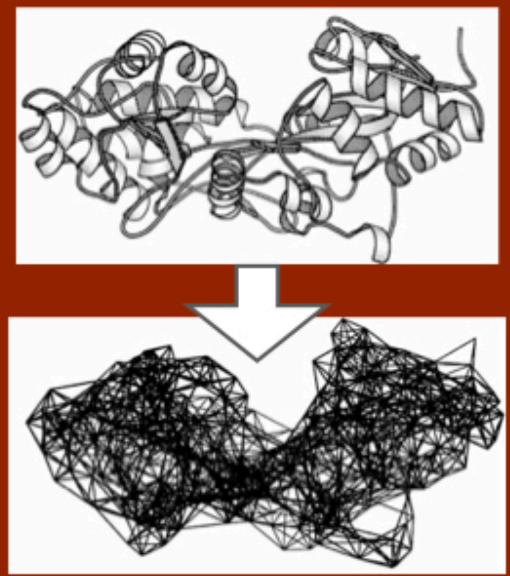
RP, Pontiggia, Micheletti
BJ 2009

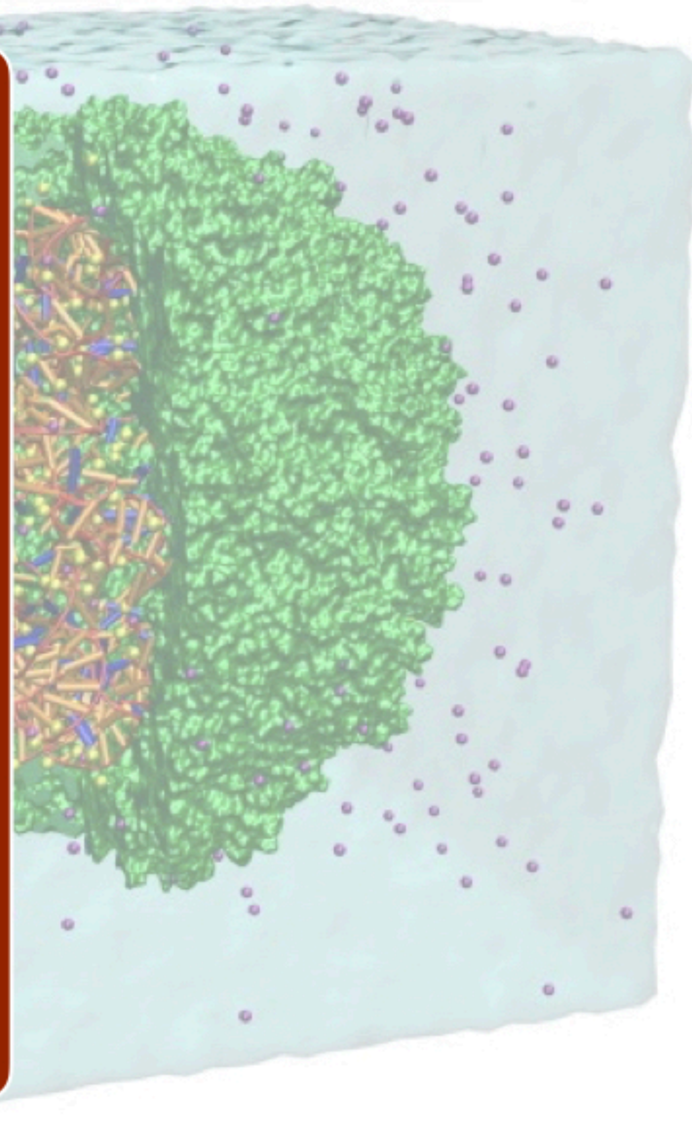
The two faces of the coarse-grained medal



The two faces of the coarse-grained medal

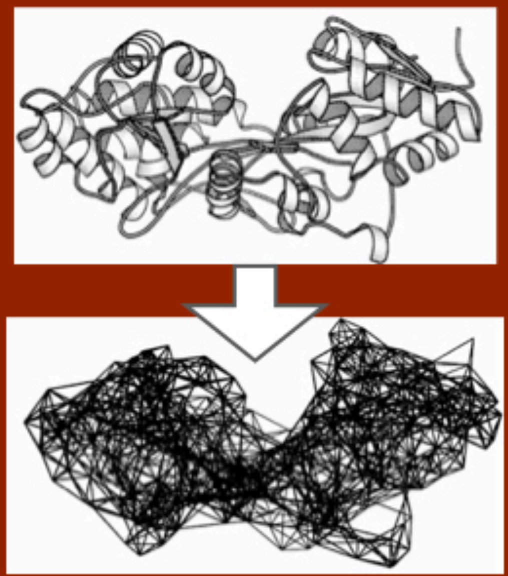
Coarse-grained model
to
reproduce
a property or behaviour



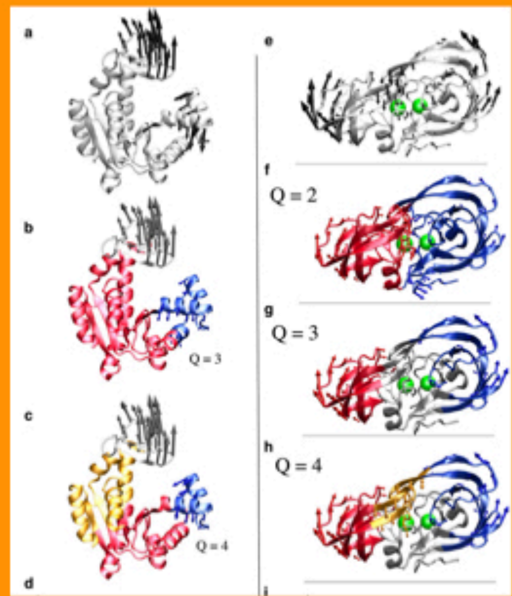


The two faces of the coarse-grained medal

Coarse-grained model
to
reproduce
a property or behaviour



Coarse-grained model
to
understand
a property or behaviour



Coarse-graining as an analysis tool

Coarse-graining as an analysis tool

We can simulate larger and larger systems for longer and longer times

Coarse-graining as an analysis tool

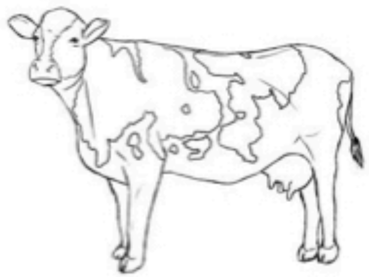
We can simulate larger and larger systems for longer and longer times

The amount of data we gain is redundant with respect to the amount of information we need

Coarse-graining as an analysis tool

We can simulate larger and larger systems for longer and longer times

The amount of data we gain is redundant with respect to the amount of information we need



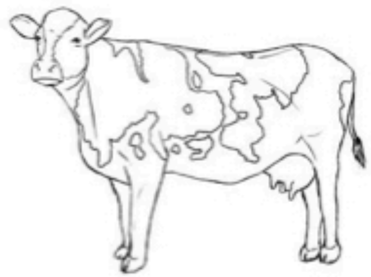
A coarse-grained picture of the system implies a resolution loss

Coarse-graining as an analysis tool

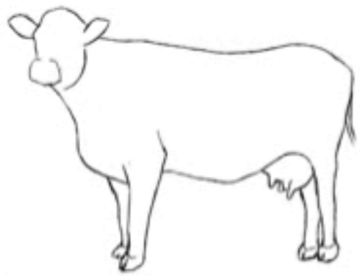
We can simulate larger and larger systems for longer and longer times

The amount of data we gain is redundant with respect to the amount of information we need

A coarse-grained picture of the system implies a resolution loss



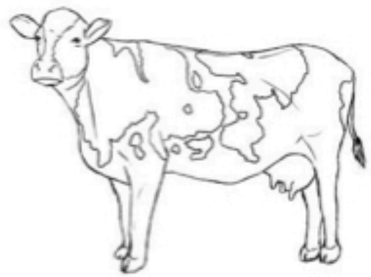
We can look for the most informative simplified description that provides the larges amount of understanding with the smaller amount of data



Coarse-graining as an analysis tool

We can simulate larger and larger systems for longer and longer times

The amount of data we gain is redundant with respect to the amount of information we need

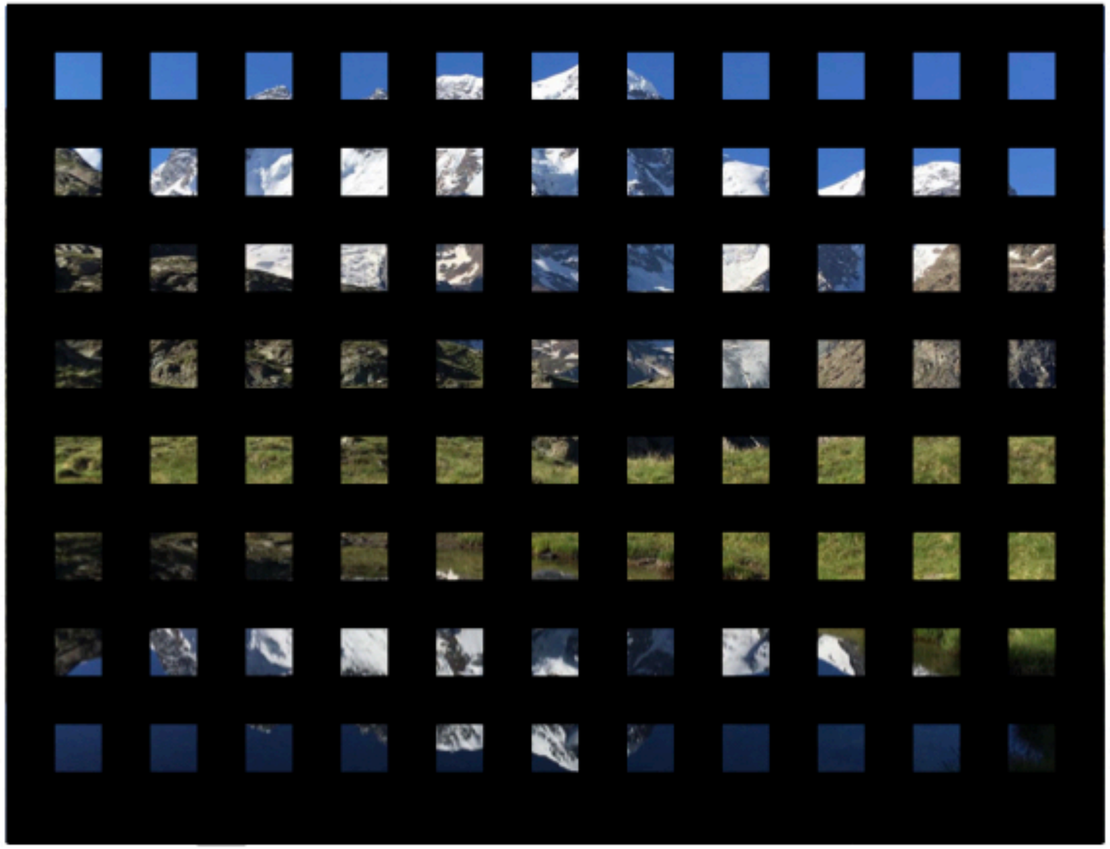


A coarse-grained picture of the system implies a resolution loss

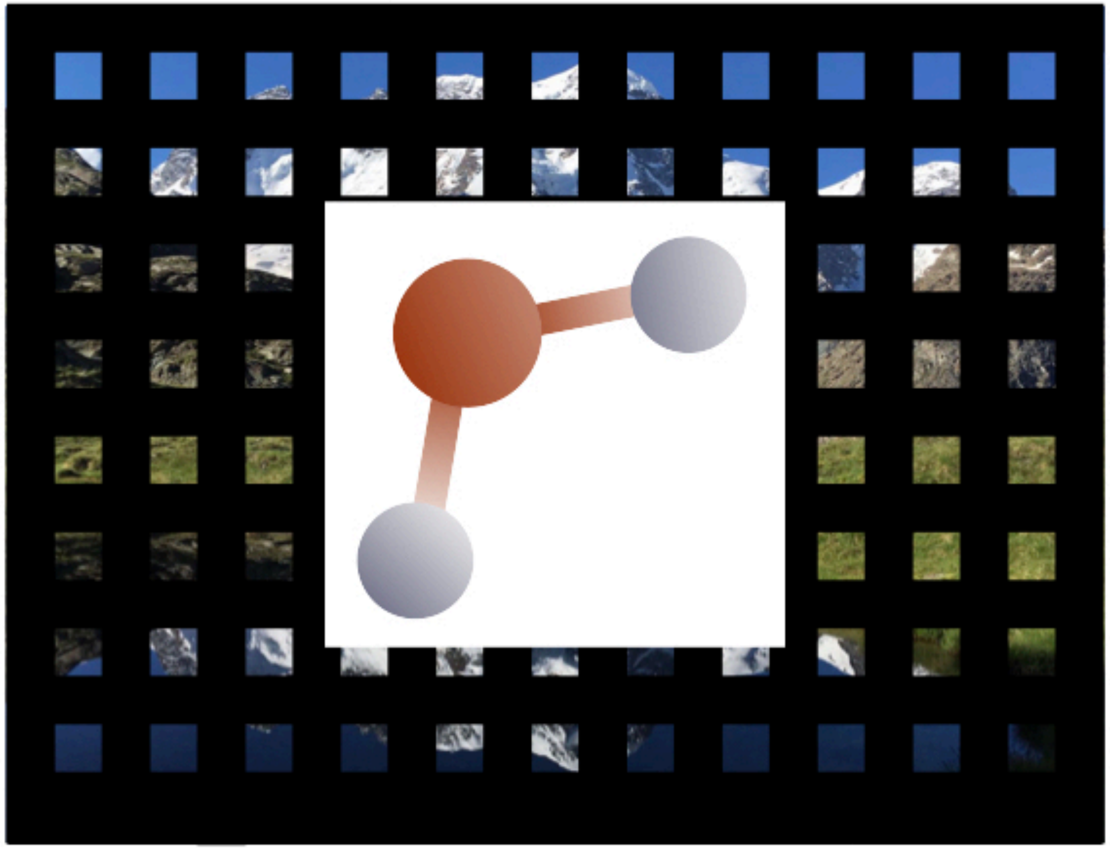
We can look for the most informative simplified description that provides the larges amount of understanding with the smaller amount of data

Decimation mapping

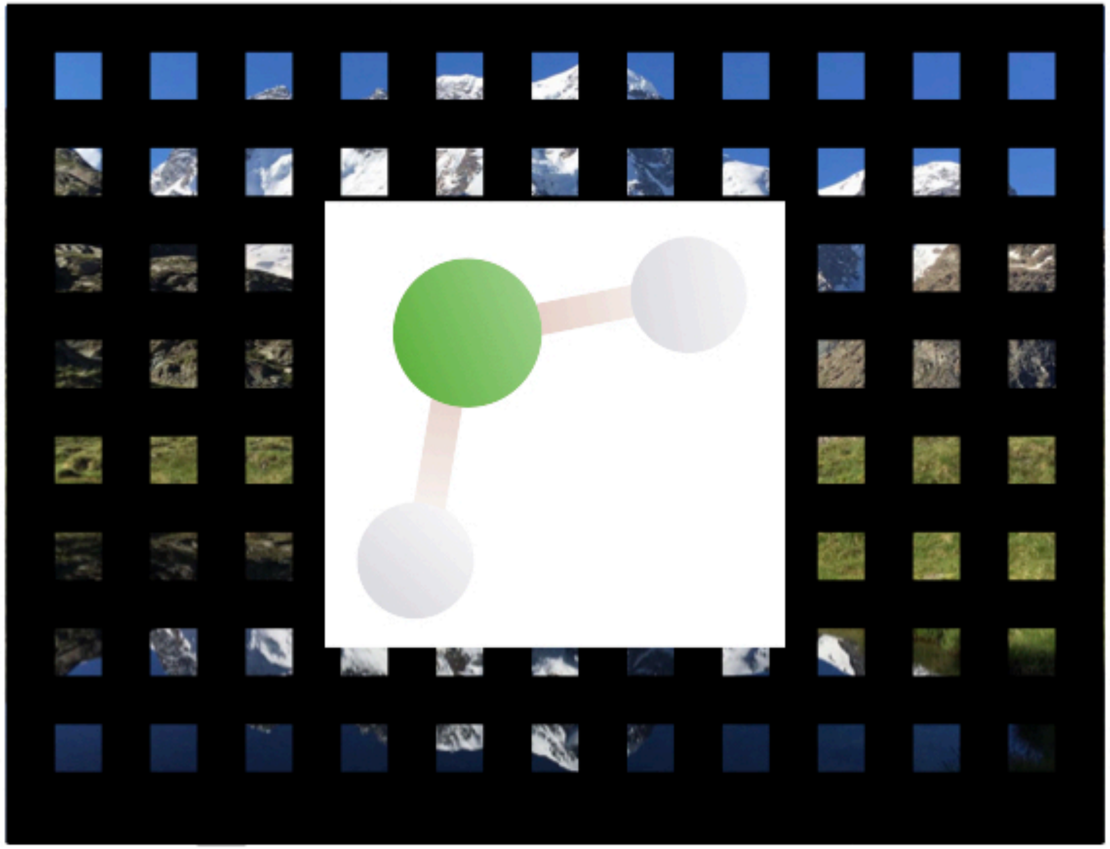
Decimation mapping



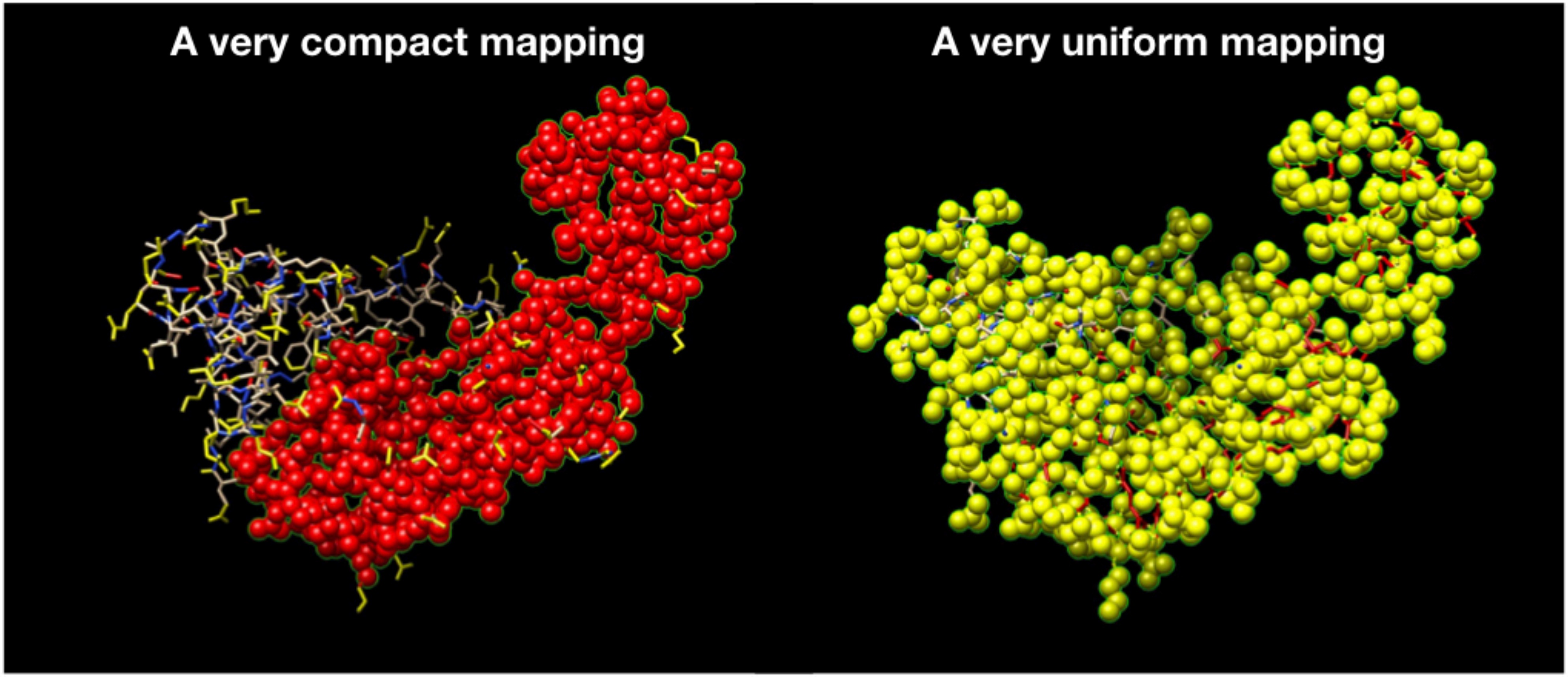
Decimation mapping



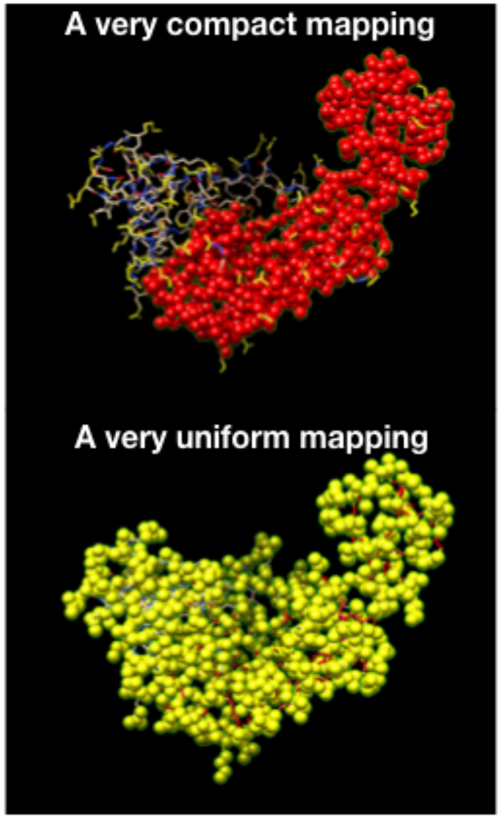
Decimation mapping



Decimation mapping



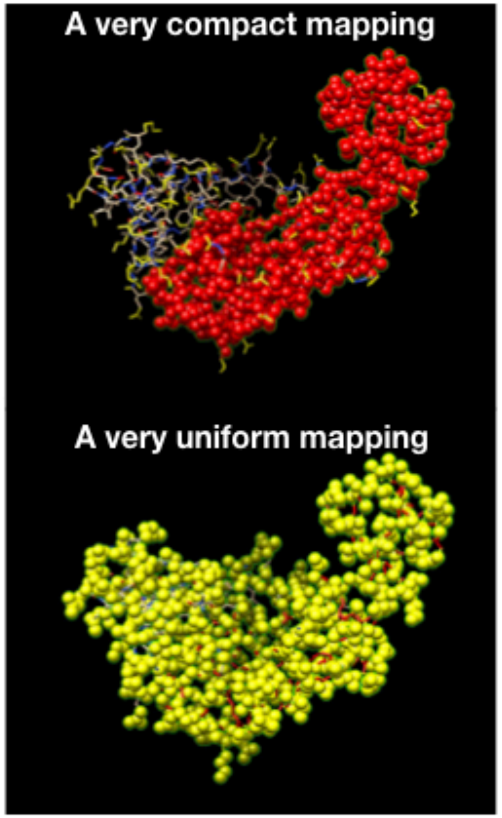
Exploring the mapping space



n atoms, N retained sites

$$\sum_{N=0}^{n} g(n,N) = \sum_{N=0}^{n} \frac{n!}{N!\,(n-N)!} = 2^n$$

Exploring the mapping space

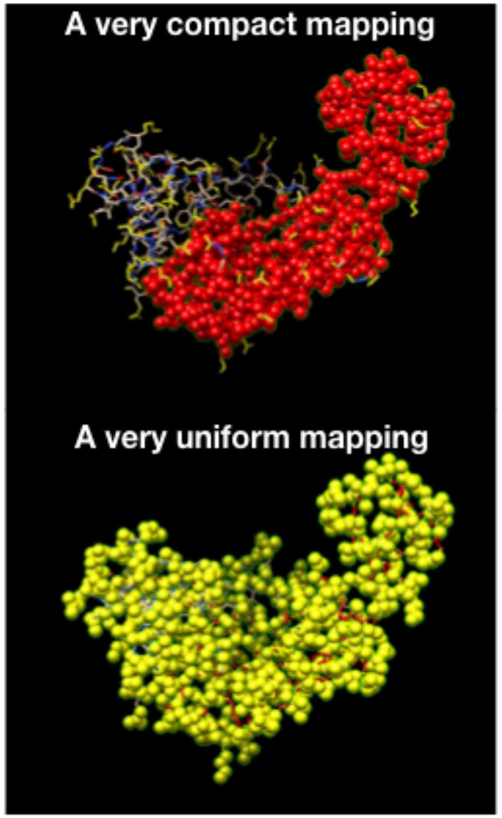


n atoms, N retained sites

$$\sum_{N=0}^{n} g(n,N) = \sum_{N=0}^{n} \frac{n!}{N!\,(n-N)!} = 2^n$$

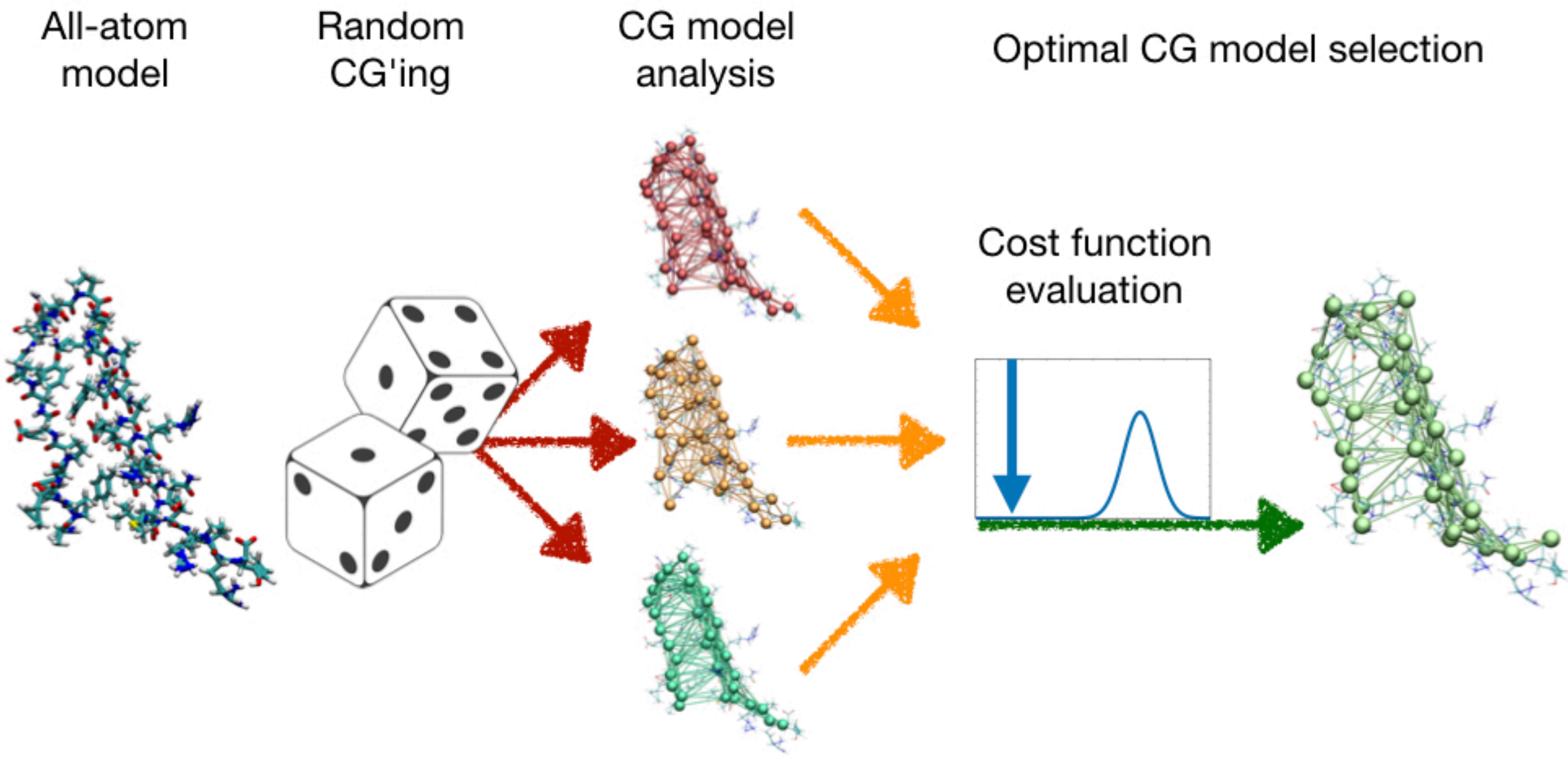
50 residues / 950 atoms $\Rightarrow \sim 10^{286}$ possible mappings
(number of atoms in the Universe: $\sim 10^{80}$)

Exploring the mapping space



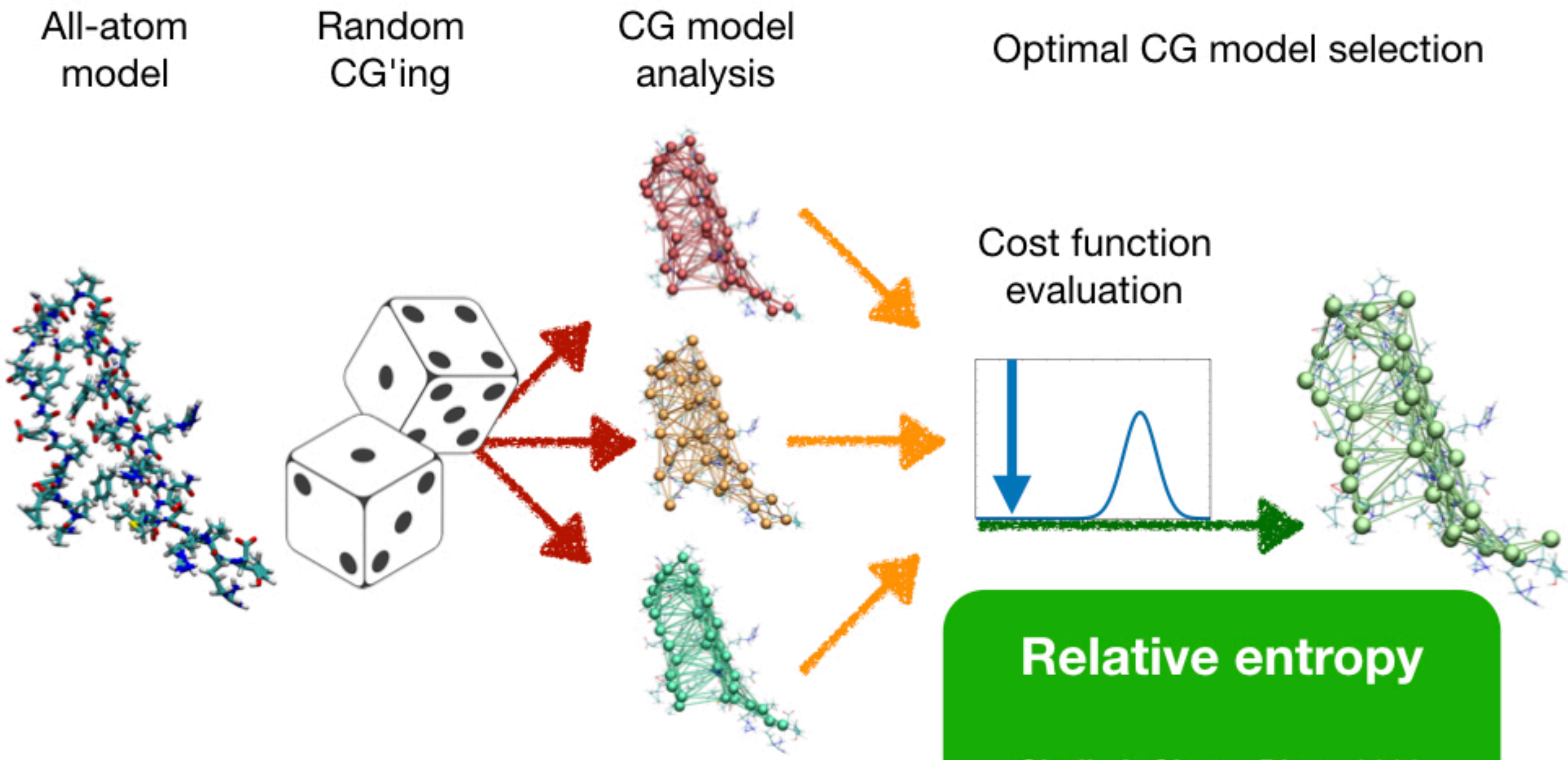
A stochastic search scheme is needed

Optimisation of the mapping: general strategy



Diggins, Liu, Deserno, RP, JCTC 2018

Optimisation of the mapping: general strategy



Relative entropy

Shell, J. Chem. Phys. 2008

Diggins, Liu, Deserno, RP, JCTC 2018

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln \left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)} \right]$$

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

Probability to sample from the all-atom model an all-atom configuration

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r}) || P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

Probability to sample from the all-atom model an all-atom configuration

Probability to sample from the coarse-grained model an all-atom configuration

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

$$= \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{V^n}{V^N}\frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r})|U)}\right] + S_{map}$$

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

$$= \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{V^n}{V^N}\frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r})|U)}\right] + S_{map}$$

Probability to sample from the coarse-grained model a coarse-grained configuration that maps onto a given all-atom configuration

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

$$= \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{V^n}{V^N}\frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r})|U)}\right] + S_{map}$$

$$S_{map} = \left\langle \ln\left[\frac{p_r(\mathbf{r})}{\bar{p}_r(\mathbf{M}(\mathbf{r}))}\right]\right\rangle = D_{KL}(p_r(\mathbf{r})||\bar{p}_r(\mathbf{M}(\mathbf{r})))$$

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

$$= \int d\mathbf{r}\ p_r(\mathbf{r}) \ln\left[\frac{V^n}{V^N}\frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r})|U)}\right] + S_{map}$$

$$S_{map} = \left\langle \ln\left[\frac{p_r(\mathbf{r})}{\bar{p}_r(\mathbf{M}(\mathbf{r}))}\right]\right\rangle = D_{KL}(p_r(\mathbf{r})||\bar{p}_r(\mathbf{M}(\mathbf{r})))$$

Average probability to sample
from the all-atom model
a coarse-grained configuration
that maps onto a given all-atom configuration

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$S_{rel} = D_{KL}(p_r(\mathbf{r})||P_r(\mathbf{r}|U)) = \int d\mathbf{r}\; p_r(\mathbf{r}) \ln\left[\frac{p_r(\mathbf{r})}{P_r(\mathbf{r}|U)}\right]$$

$$= \int d\mathbf{r}\; p_r(\mathbf{r}) \ln\left[\frac{V^n}{V^N}\frac{p_r(\mathbf{r})}{P_R(\mathbf{M}(\mathbf{r})|U)}\right] + S_{map}$$

$$S_{map} = \left\langle \ln\left[\frac{p_r(\mathbf{r})}{\bar{p}_r(\mathbf{M}(\mathbf{r}))}\right]\right\rangle = D_{KL}(p_r(\mathbf{r})||\bar{p}_r(\mathbf{M}(\mathbf{r})))$$

$$\Omega_1(\mathbf{R}) = \int d\mathbf{r}\; \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R})$$

$$\bar{p}_r(\mathbf{R}) = \frac{1}{\Omega_1(\mathbf{R})}\int d\mathbf{r}\; p_r(\mathbf{r})\delta(\mathbf{M}(\mathbf{r}) - \mathbf{R})$$

Shell, J. Chem Phys. 2008

Relative entropy / mapping entropy

$$p(\mathbf{r}) = \bar{p}_r(\mathbf{M}(\mathbf{r})) \Leftrightarrow S_{map} = 0$$

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Relative entropy / mapping entropy

$$p(\mathbf{r}) = \bar{p}_r(\mathbf{M}(\mathbf{r})) \Leftrightarrow S_{map} = 0$$

$$p(\mathbf{r}) = \frac{\exp(-\beta E(\mathbf{r}))}{Z}$$

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Relative entropy / mapping entropy

$$p(\mathbf{r}) = \bar{p}_r(\mathbf{M}(\mathbf{r})) \Leftrightarrow S_{map} = 0$$

$$p(\mathbf{r}) = \frac{\exp(-\beta E(\mathbf{r}))}{Z}$$

$$E(\mathbf{r}) = E(\mathbf{r}') \; \forall \; \mathbf{r}, \mathbf{r}' : \mathbf{M}(\mathbf{r}) = \mathbf{M}(\mathbf{r}') = \mathbf{R}$$

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Relative entropy / mapping entropy

$$p(\mathbf{r}) = \bar{p}_r(\mathbf{M}(\mathbf{r})) \Leftrightarrow S_{map} = 0$$

$$p(\mathbf{r}) = \frac{\exp(-\beta E(\mathbf{r}))}{Z}$$

$$E(\mathbf{r}) = E(\mathbf{r}') \; \forall \; \mathbf{r}, \mathbf{r}' : \mathbf{M}(\mathbf{r}) = \mathbf{M}(\mathbf{r}') = \mathbf{R}$$

The mapping entropy vanishes if:

- The mapping acts through decimation
- The macrostates ***R*** are composed by microstates ***r*** with the same energy

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Relative entropy / mapping entropy

$$p(\mathbf{r}) = \bar{p}_r(\mathbf{M}(\mathbf{r})) \Leftrightarrow S_{map} = 0$$

$$p(\mathbf{r}) = \frac{\exp(-\beta E(\mathbf{r}))}{Z}$$

$$E(\mathbf{r}) = E(\mathbf{r}') \ \forall\ \mathbf{r}, \mathbf{r}' : \mathbf{M}(\mathbf{r}) = \mathbf{M}(\mathbf{r}') = \mathbf{R}$$

The mapping entropy vanishes if:

- The mapping acts through decimation
- The macrostates ***R*** are composed by microstates ***r*** with the same energy

So what?

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Relative entropy / mapping entropy

Second order cumulant expansion

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \;\; P_\beta(U|\mathbf{R}) \; U$$

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2} \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \; p_R(\mathbf{R}) \; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

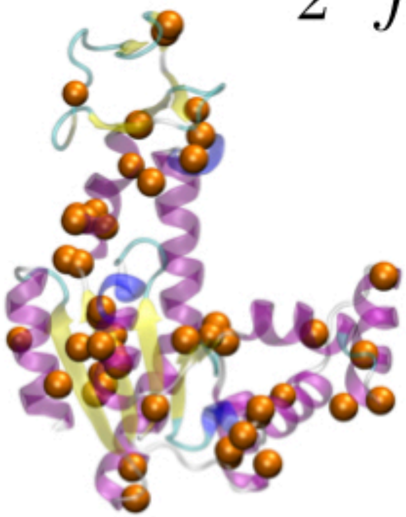
Relative entropy / mapping entropy

Second order cumulant expansion

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \ P_\beta(U|\mathbf{R}) \ U$$

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2} \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \ p_R(\mathbf{R}) \ \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

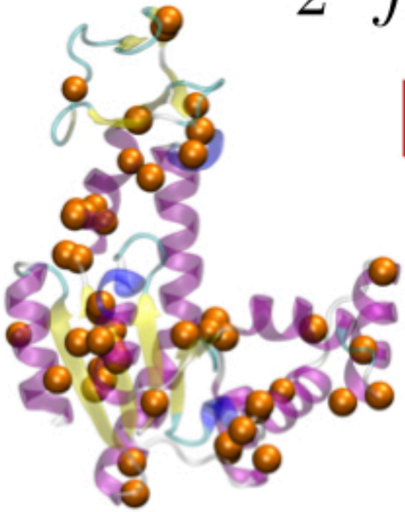
Relative entropy / mapping entropy

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \;\; P_\beta(U|\mathbf{R}) \; U$$

Second order cumulant expansion

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2}\langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2}\int d\mathbf{R} \; p_R(\mathbf{R}) \; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



simulation

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

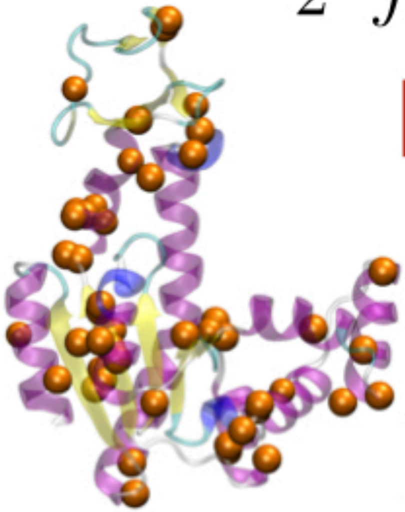
Relative entropy / mapping entropy

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \;\; P_\beta(U|\mathbf{R}) \; U$$

Second order cumulant expansion

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2}\langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \; p_R(\mathbf{R}) \; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



simulation

mapping

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

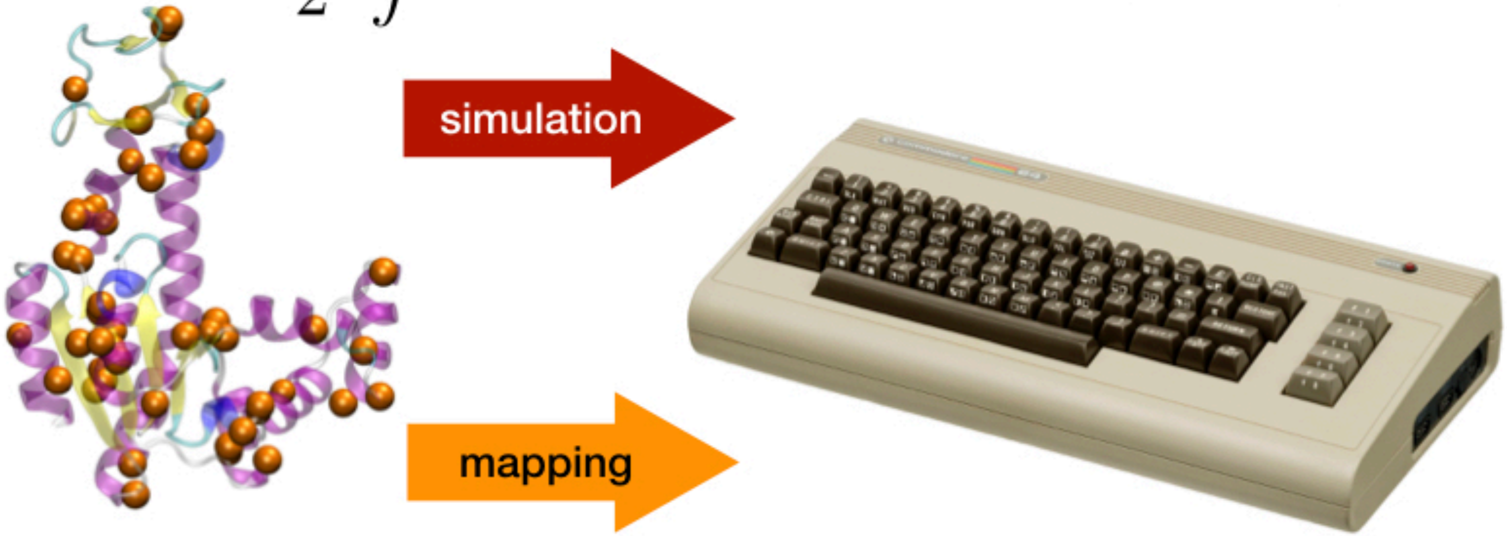
Relative entropy / mapping entropy

Second order cumulant expansion

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \;\, P_\beta(U|\mathbf{R}) \; U$$

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2}\langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \; p_R(\mathbf{R}) \; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

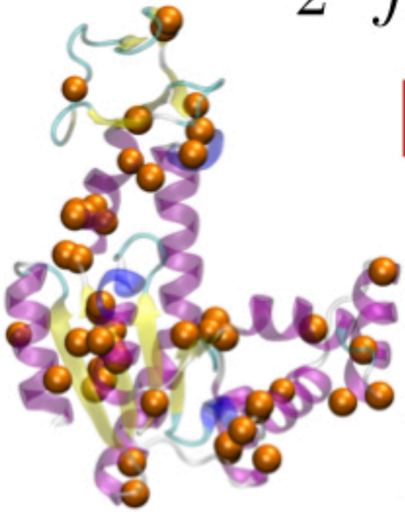
Relative entropy / mapping entropy

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \;\, P_\beta(U|\mathbf{R})\; U$$

Second order cumulant expansion

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2} \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R}\; p_R(\mathbf{R})\; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

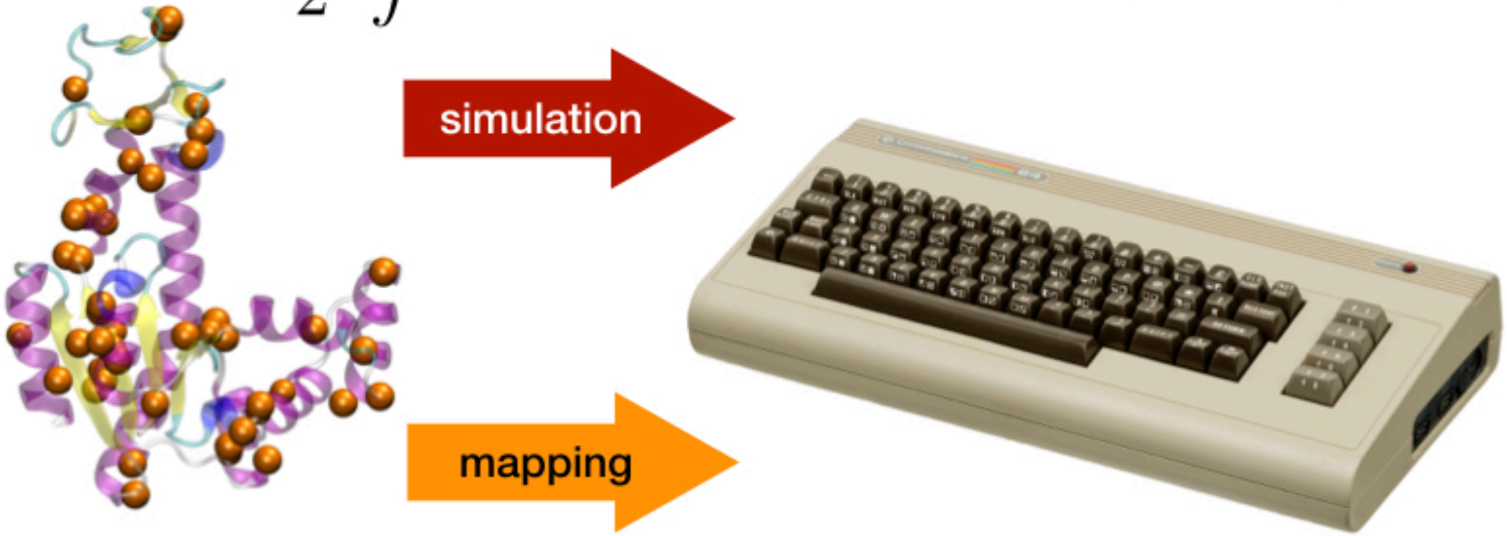
Relative entropy / mapping entropy

Second order cumulant expansion

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \ P_\beta(U|\mathbf{R}) \ U$$

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2} \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \ p_R(\mathbf{R}) \ \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$



Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

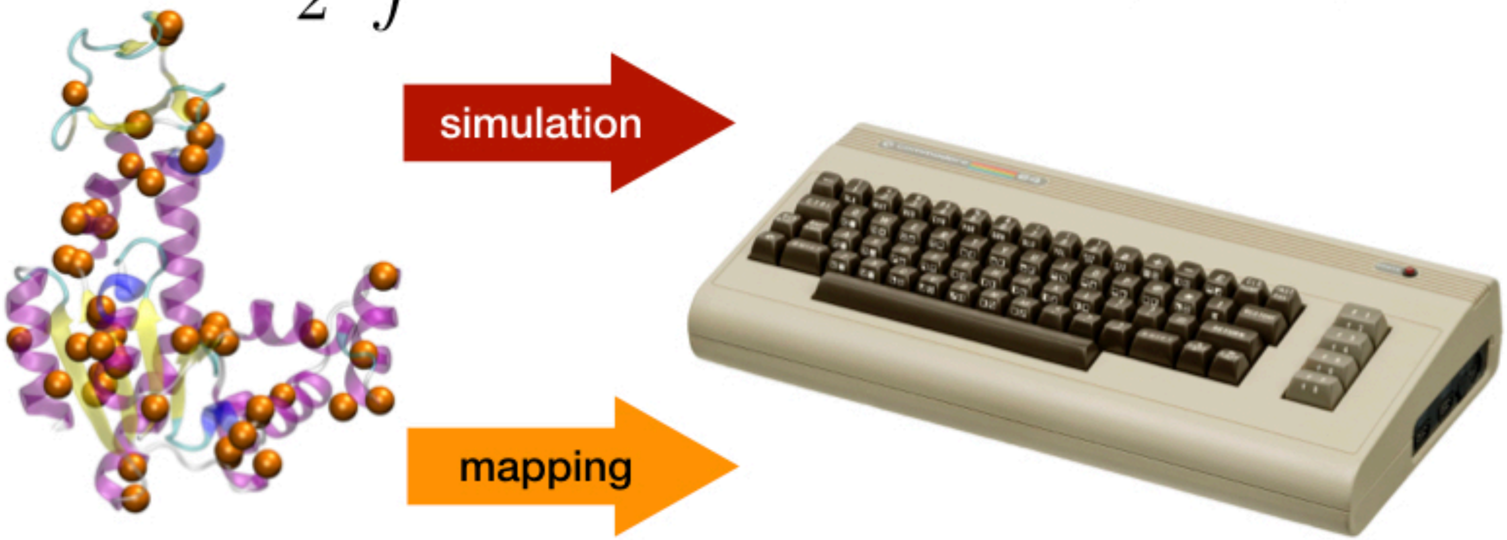
Relative entropy / mapping entropy

Second order cumulant expansion

$$\langle U \rangle_{\beta/\mathbf{R}} = \int dU \; P_\beta(U|\mathbf{R}) \; U$$

$$S_{map}(\mathbf{R}) \simeq -\frac{\beta^2}{2} \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

$$S_{map} \simeq \frac{\beta^2}{2} \int d\mathbf{R} \; p_R(\mathbf{R}) \; \langle (U - \langle U \rangle_{\beta/\mathbf{R}})^2 \rangle_{\beta/\mathbf{R}}$$

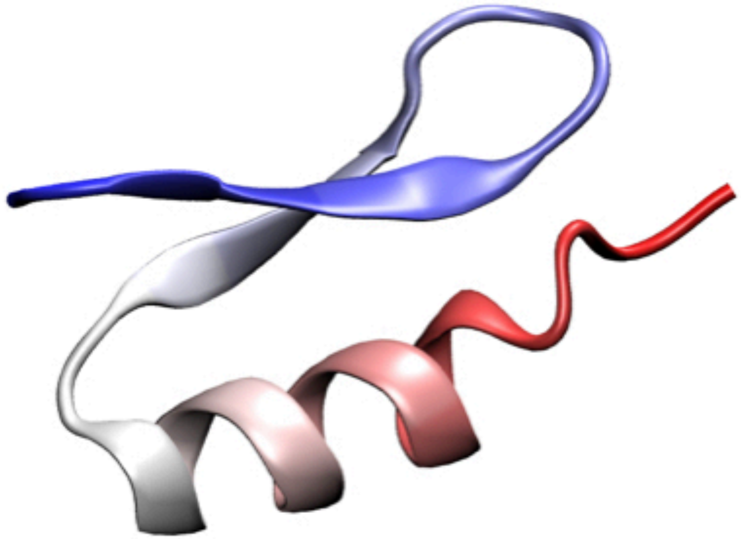


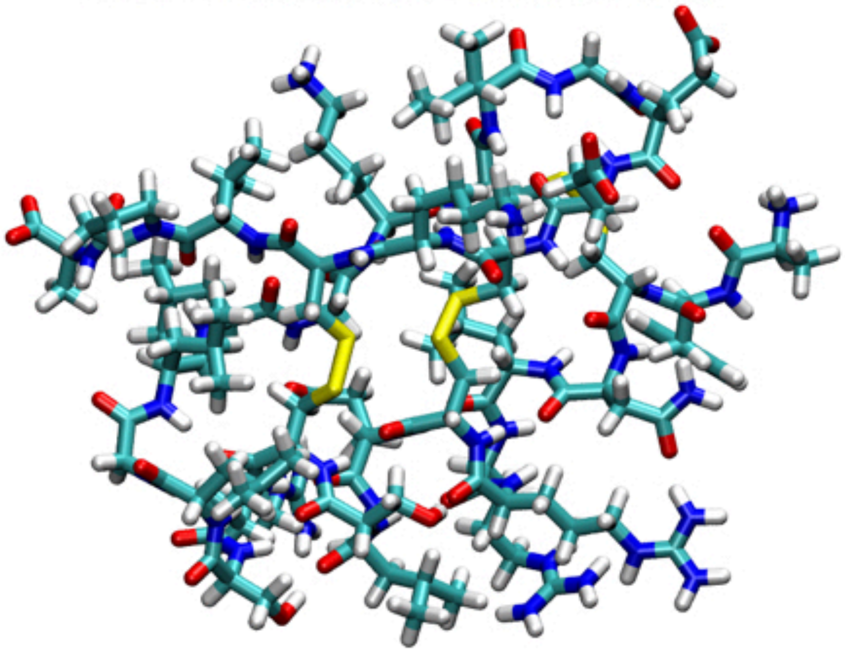
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation

Application to tamapin

Toxin from Indian red scorpion
Blocker of SK2 potassium channel
High pharmaceutical interest



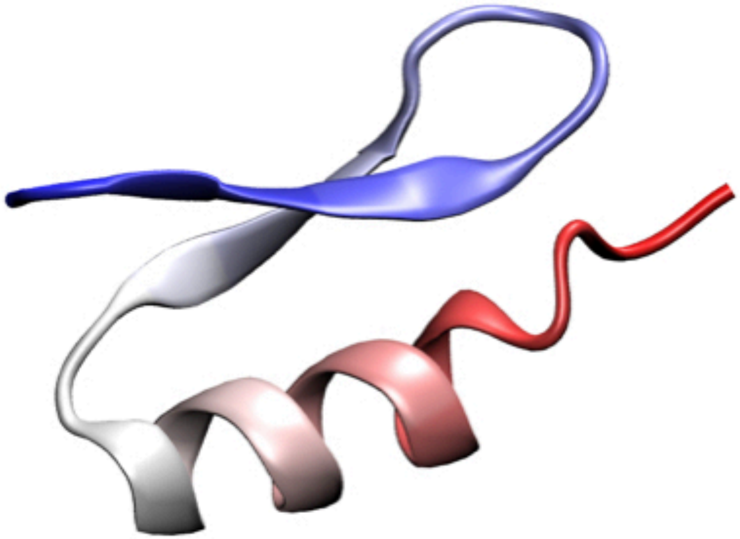


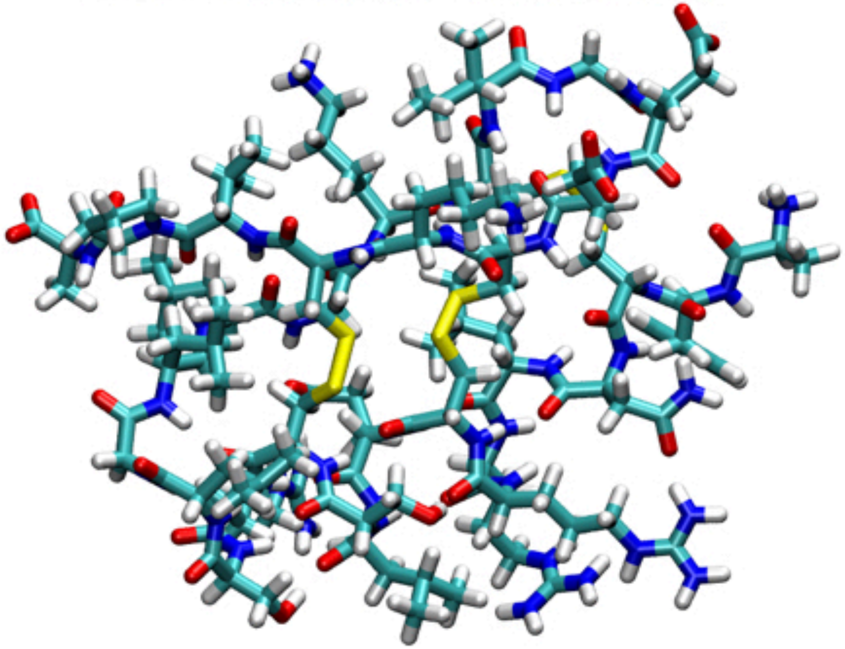
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation

Application to tamapin

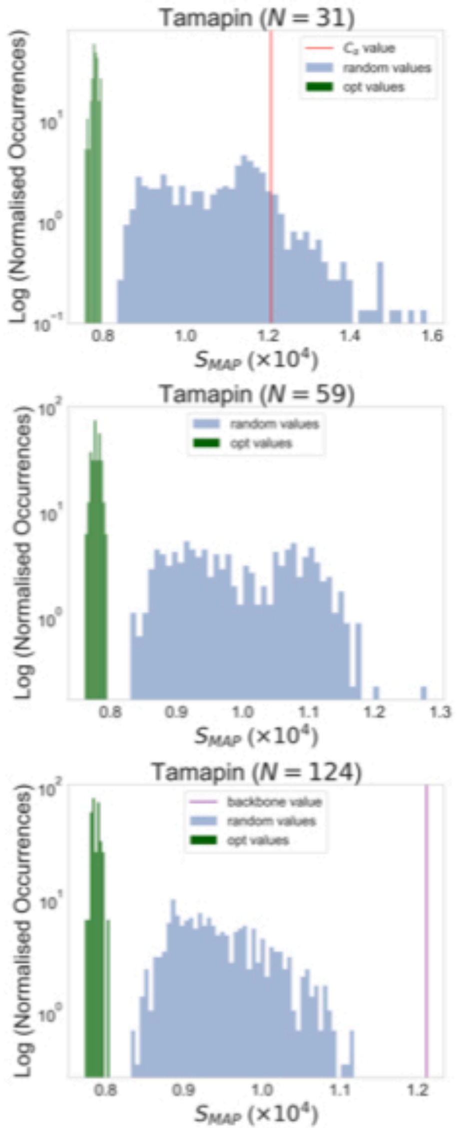
Toxin from Indian red scorpion
Blocker of SK2 potassium channel
High pharmaceutical interest





Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation



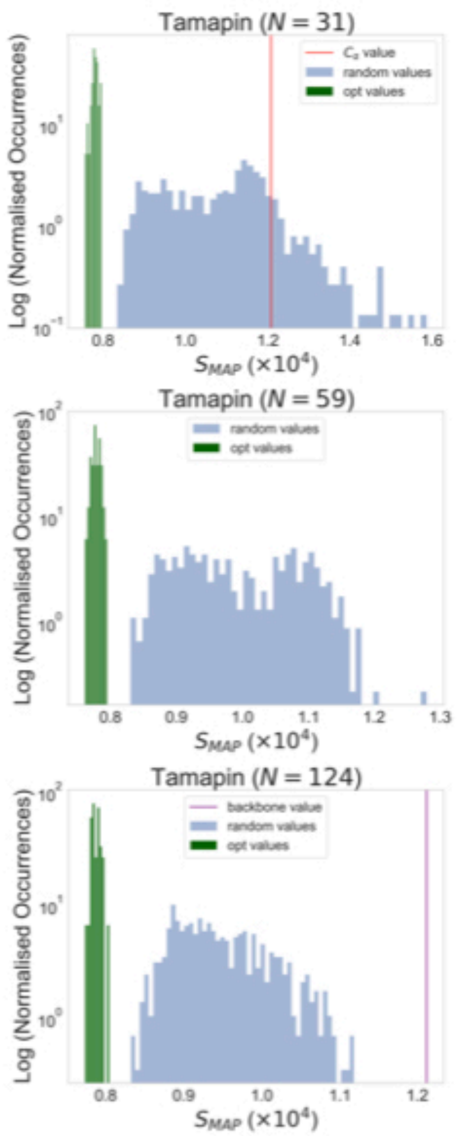
Mapping entropy is computed for:

- random mappings with fixed number N of CG sites
- Calpha atoms only
- backbone atoms only

48 optimisation runs are performed

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation



Mapping entropy is computed for:

- random mappings with fixed number N of CG sites
- Calpha atoms only
- backbone atoms only

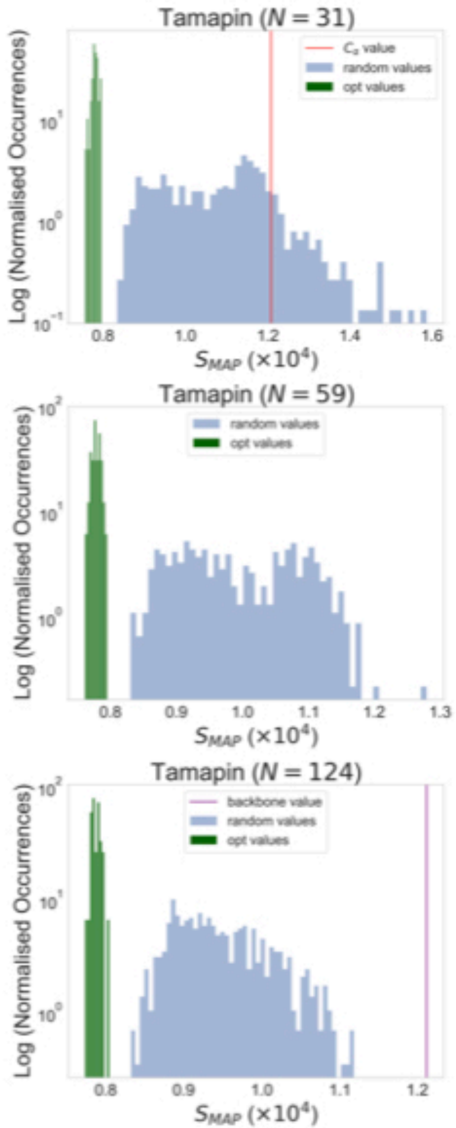
48 optimisation runs are performed

Notable observations

- Calpha atoms only mapping has S_{map} slightly above random average (z-score ~0.87)
- backbone atoms only mapping has S_{map} substantially above random average (z-score ~4.37)
- optimised mappings have similar and lower-than-average S_{map} (z-score ~ -2.38 for N = 59)

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation



Mapping entropy is computed for:

- random mappings with fixed number N of CG sites
- Calpha atoms only
- backbone atoms only

48 optimisation runs are performed

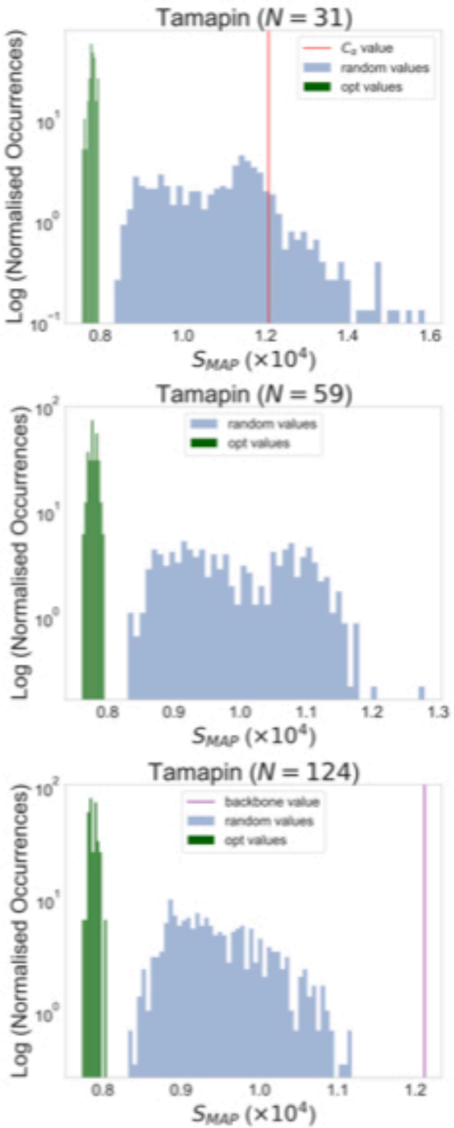
Notable observations

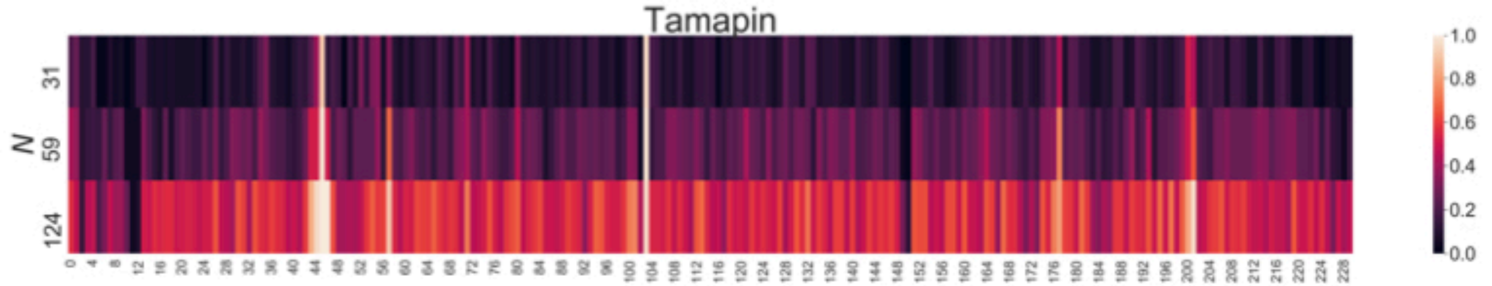
- Calpha atoms only mapping has S_{map} slightly above random average (z-score ~0.87)
- backbone atoms only mapping has S_{map} substantially above random average (z-score ~4.37)
- optimised mappings have similar and lower-than-average S_{map} (z-score ~ -2.38 for N = 59)

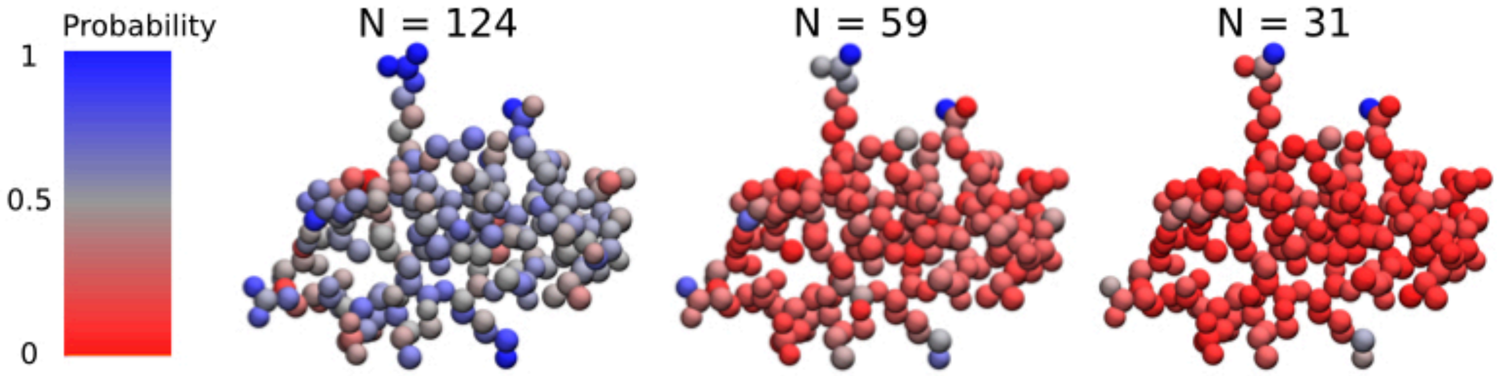
Conventional mappings (Calpha only, united atoms)
are very bad at providing a faithful representation of the system

Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation

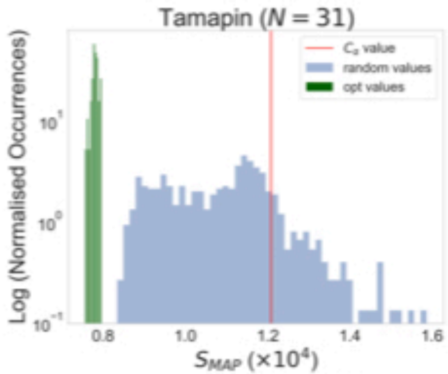


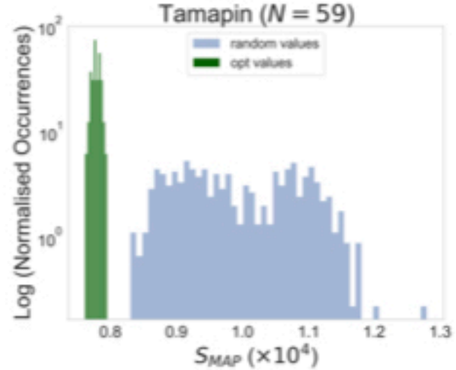


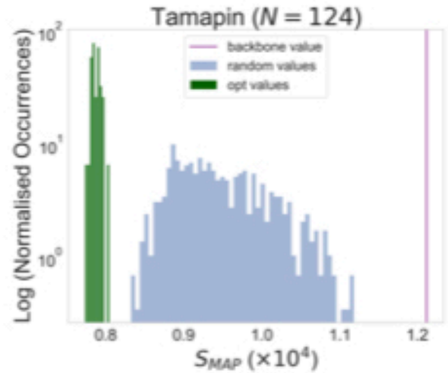


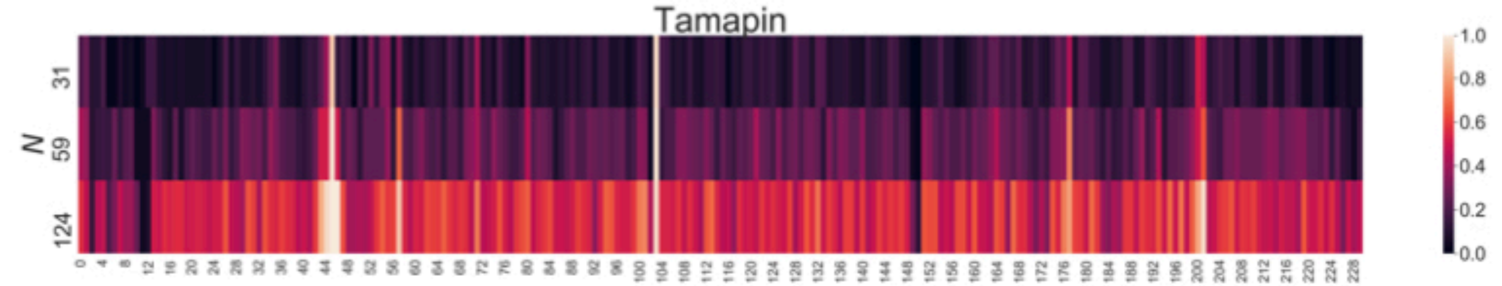
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

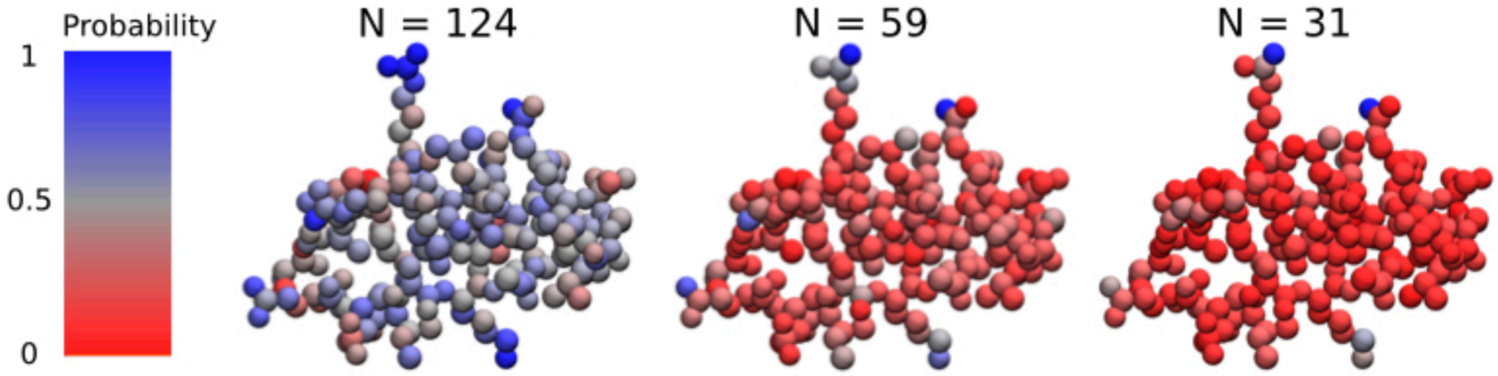
Mapping entropy minimisation











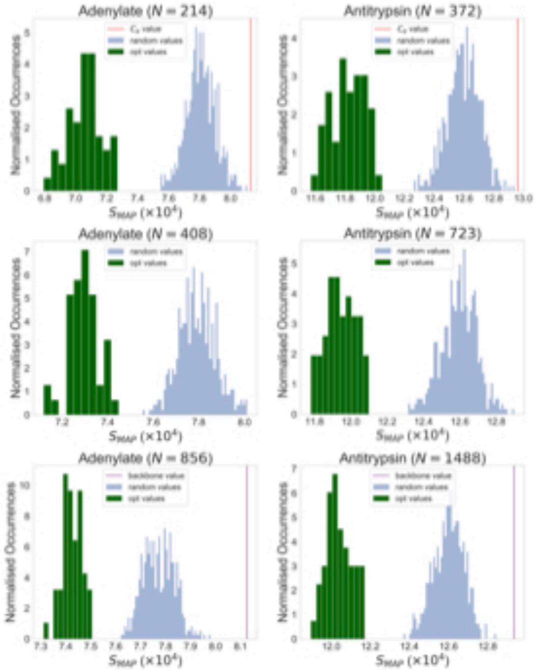
ARG6 and ARG13: main actors involved in the tamapin-SK2 channel interaction through electrostatics and hydrogen bonding

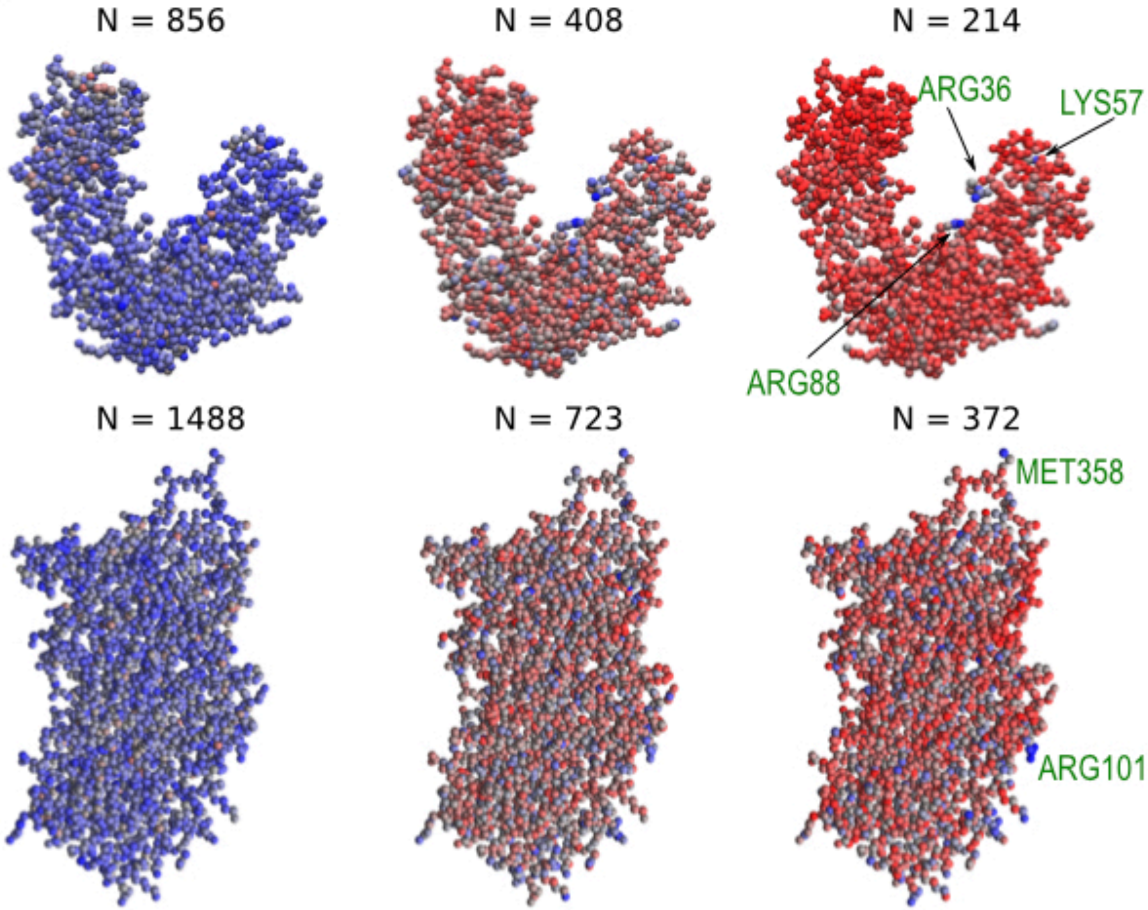
Mutation of an arginine dramatically decreases selectivity

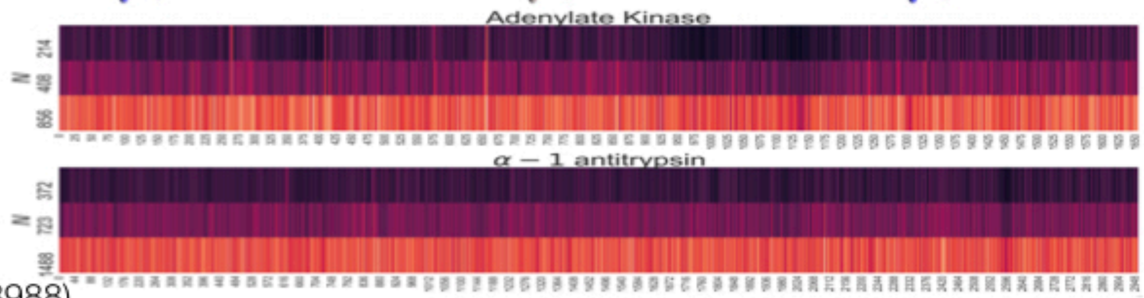
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Andreotti, di Luccio, Sampieri, Waard, Sabatier, Peptides 2005
Quintero-Hernndez, Jimnez-Vargas, Gurrola, Valdivia, Possani Toxicon 2013
Ramirez-Cordero, Toledano, Cano-Snchez, Herndez-Lopez, Flores-Solis, Saucedo-Yez, Chavez-Uribe, Brieba, and Rio-Portilla, Chemical Research in Toxicology 2014

Mapping entropy minimisation

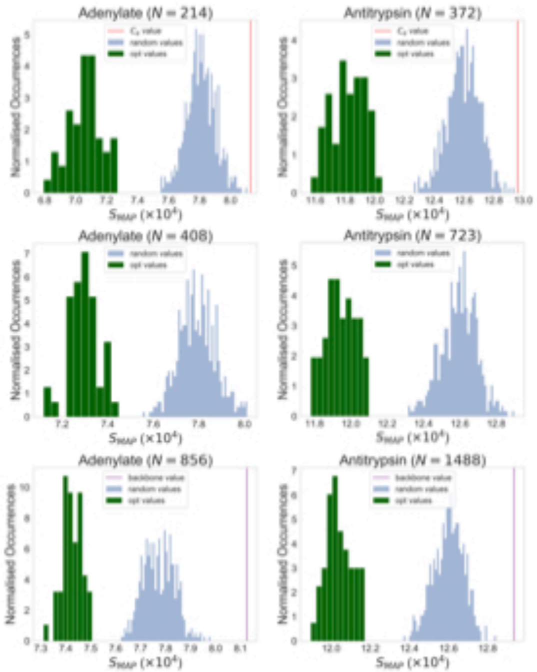




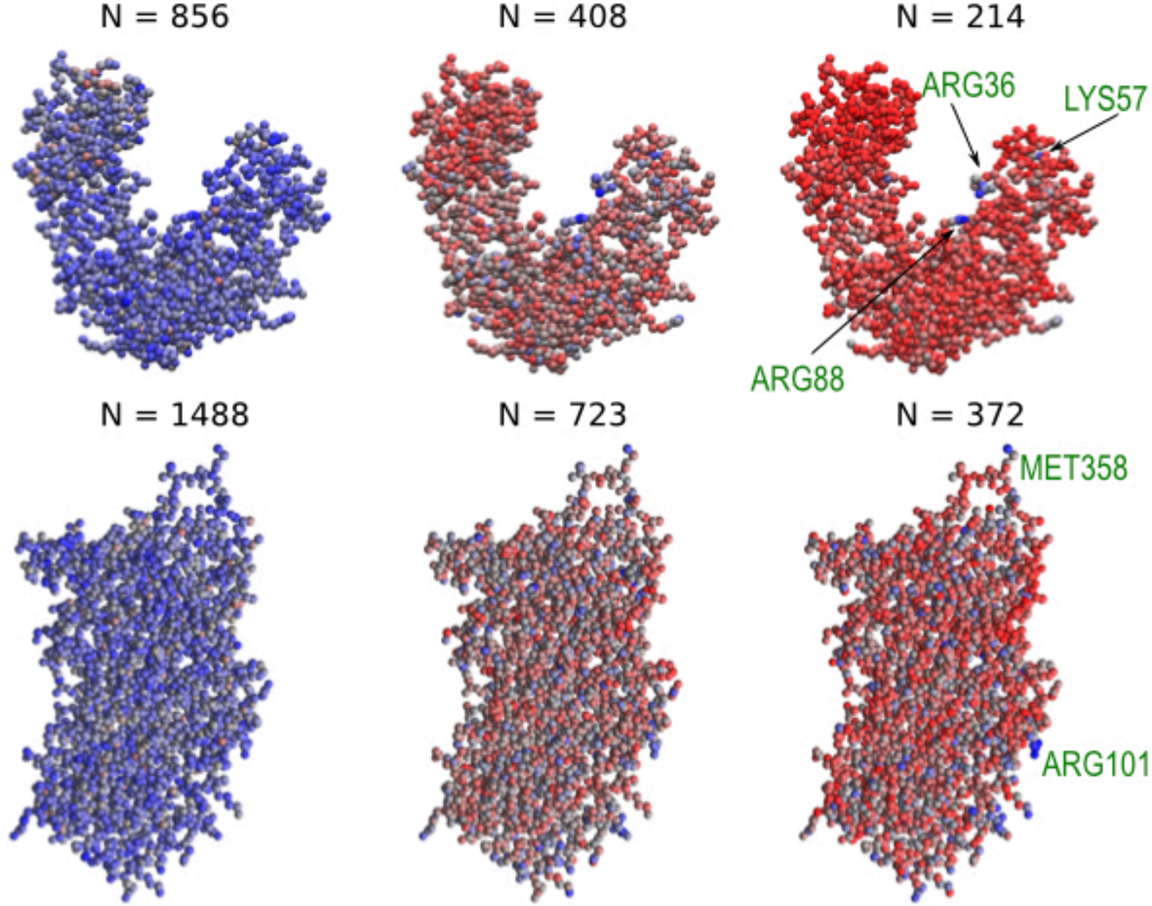


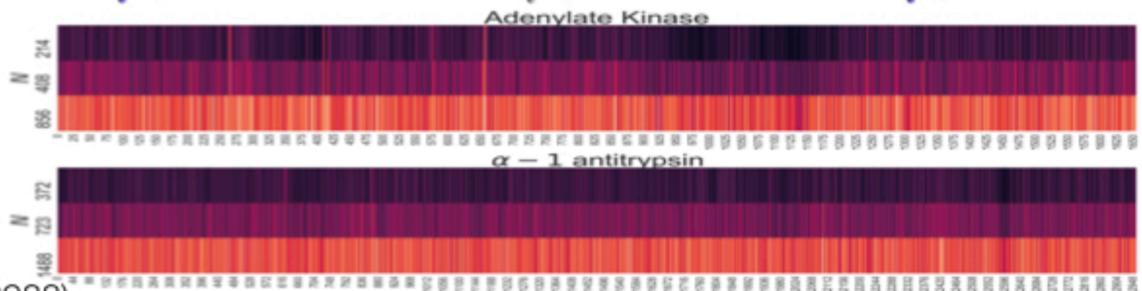
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation



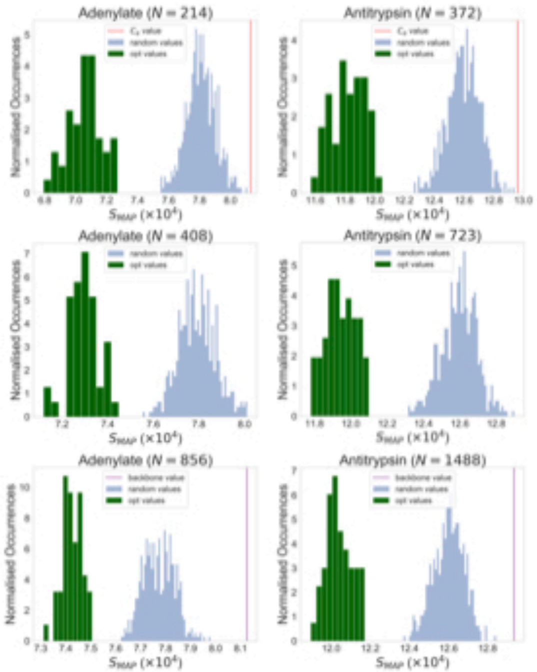
Calpha and backbone representation MEs systematically larger than random cases

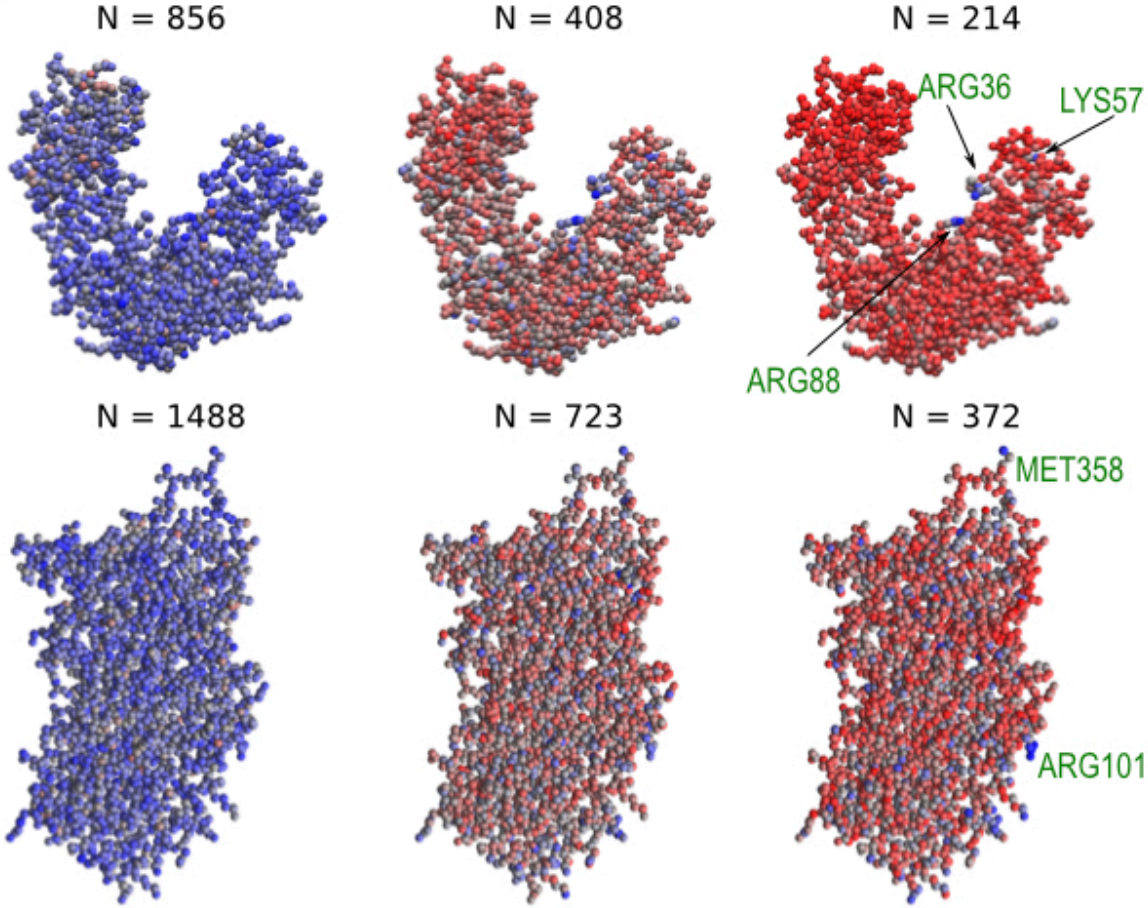


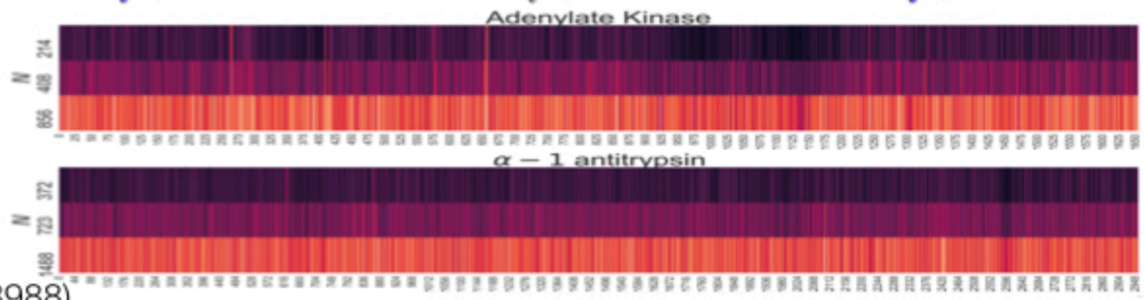


Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Mapping entropy minimisation





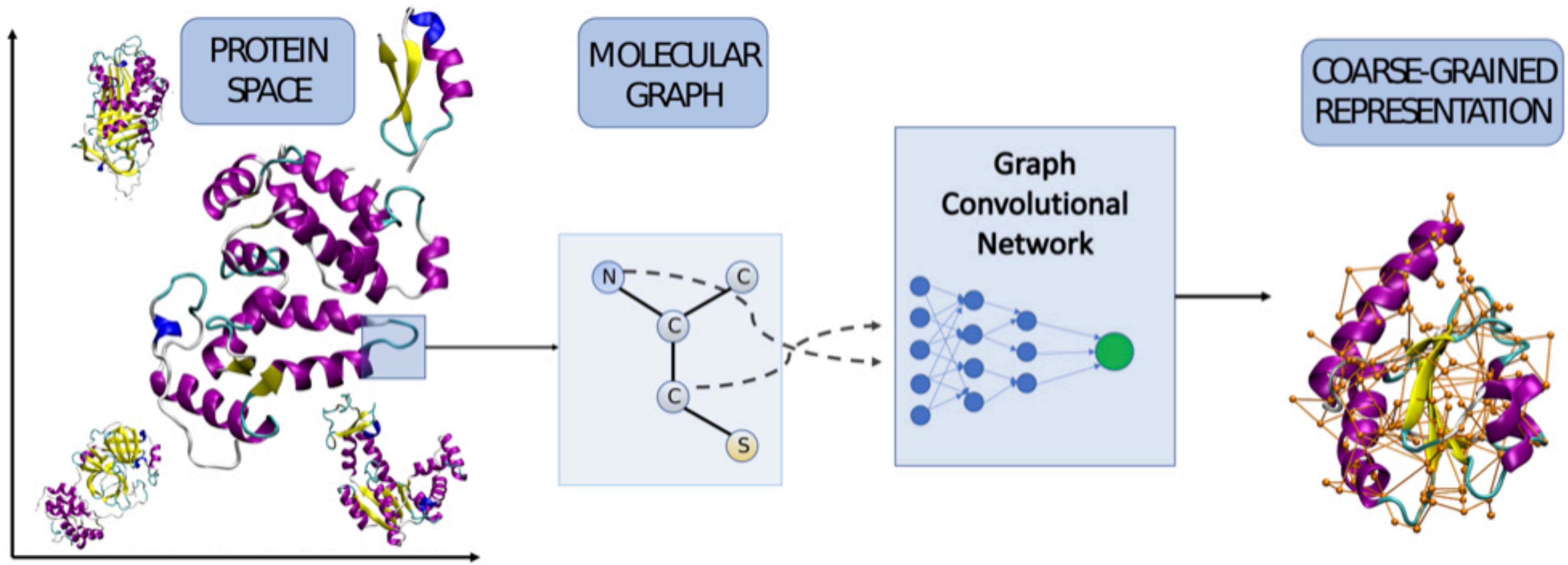


Calpha and backbone representation MEs systematically larger than random cases

Out of several highly-charged and exposed residues, only biologically relevant ones are singled out

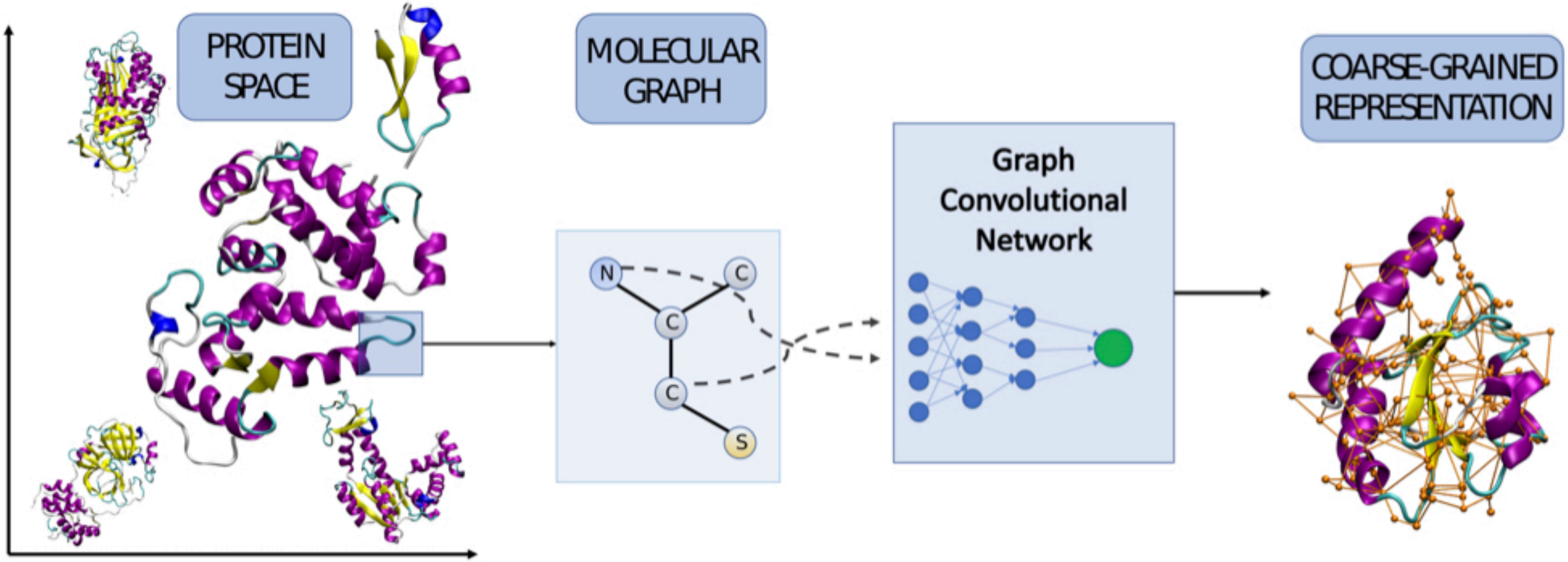
Giulini, Menichetti, Shell, RP, submitted (arXiv 2004.03988)

Perspective / ongoing work



Errica, Giulini, Bacciu, Menichetti, Rigoli, Tarenzi, Micheli, RP, in preparation

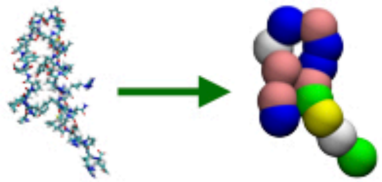
Perspective / ongoing work



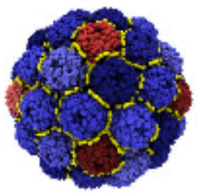
Speedup factor: ~10^5

Errica, Giulini, Bacciu, Menichetti, Rigoli, Tarenzi, Micheli, RP, in preparation

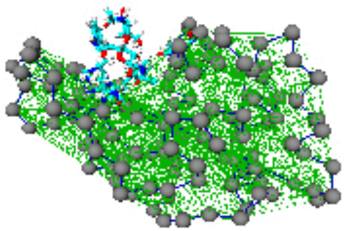
Food for thought



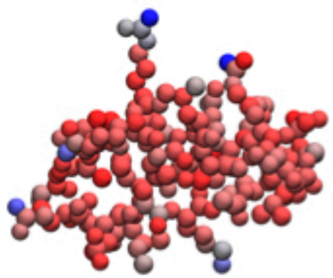
- Systematic coarse-graining methods provide speedup and valuable insight in the physics of biosystems



- Large (supra)molecules pose new challenges (large size, complex function)
- “One-size-blurs-all” CG models are too crude (different system parts require a tailored description)



- Different resolution distributions entail varying degrees of retained information



- Efficient method for mapping entropy computation
- Optimal mappings single out biologically relevant residues

Acknowledgments

Roberto Menichetti

Thomas Tarenzi

THE VARIAMOLS GROUP

Marta Rigoli

Marco Giulini

Statistical and Biological Physics group @UniTn

Pietro Faccioli

Gianluca Lattanzi

Luca Tubiana

External collaborators

Markus Deserno (Pittsburgh)

M. Scott Shell (Santa Barbara)

Alessio Micheli (Pisa)

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 758588)

Acknowledgments

Statistical and Biological Physics group @UniTn

Pietro Faccioli
Gianluca Lattanzi
Luca Tubiana

External collaborators

Markus Deserno (Pittsburgh)
M. Scott Shell (Santa Barbara)
Alessio Micheli (Pisa)

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 758588)