

# Coarse-Graining Theory

THE VOTH GROUP

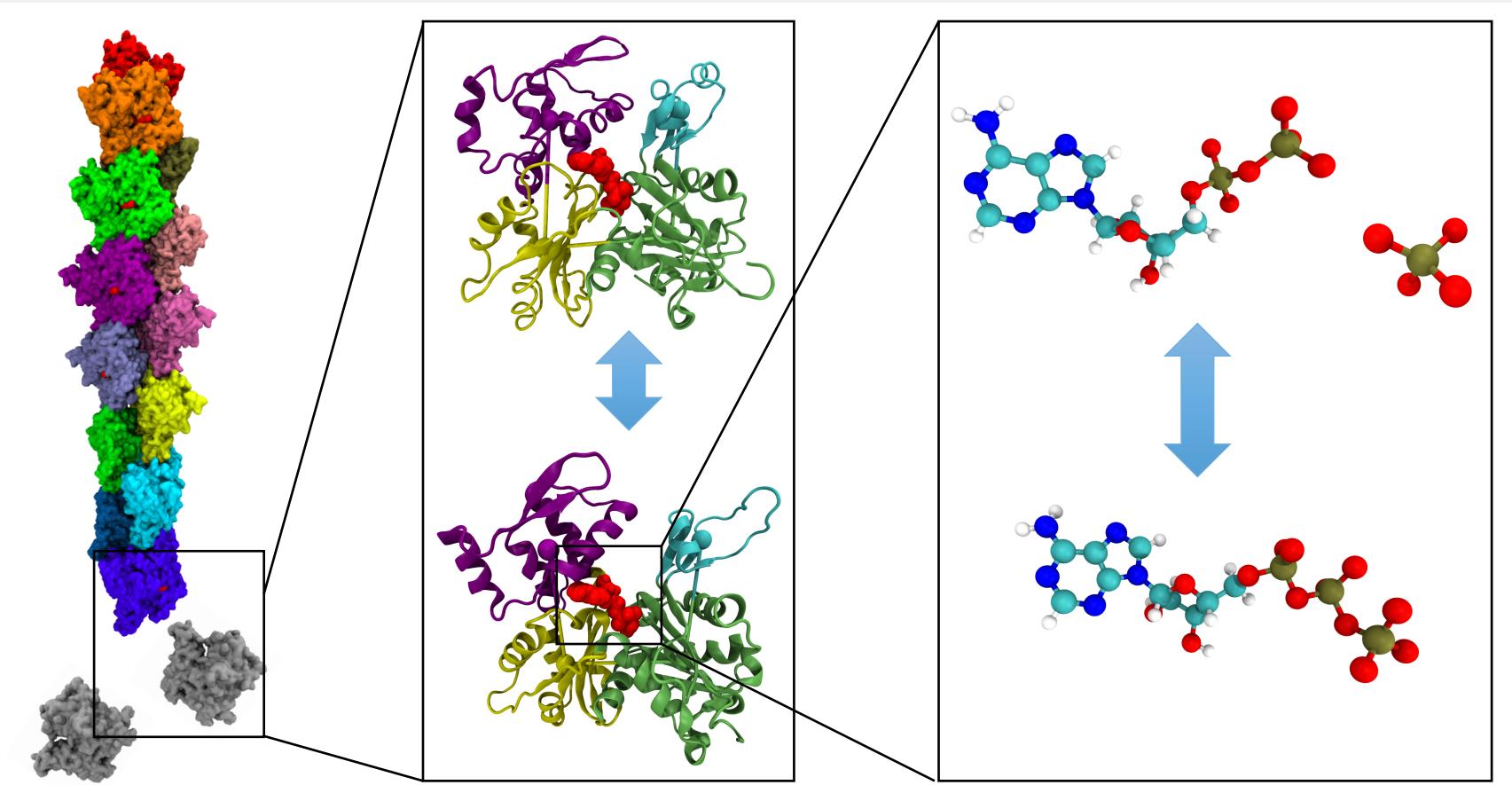
THE UNIVERSITY OF CHICAGO | DEPARTMENT OF CHEMISTRY

## Coarse-Graining

Molecular simulation is a tool used by chemists to explore chemical systems at atomic resolutions. However, current computer hardware limits system sizes to about  $10^6$  interaction sites. This makes simulating large systems, such as actin filaments, unfeasible. Coarse-graining provides a way to investigate systems at longer length and timescales.

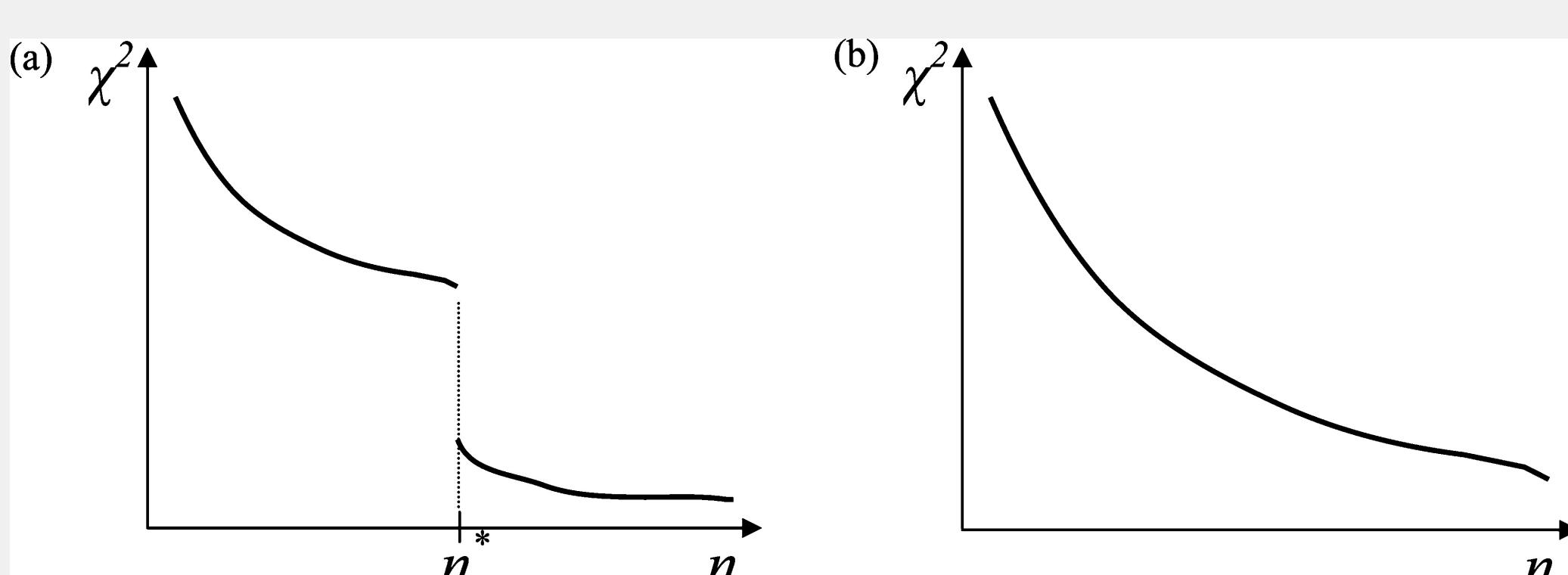
### Increasing system size and timescale

Coarse-graining is a technique that reduces the representation of a molecular system by integrating out “unnecessary” degrees of freedom from a more detailed representation. However, by removing these degrees of freedom, information about the system is lost. Thus, the challenge of coarse-grained modeling is to determine what information is removed and how the remaining information is represented and interpreted.



### Mapping

The mapping, or which atoms are represented by which coarse-grained site, provides one of the first fundamental challenges to coarse-graining. Ideally, one would wish to find a mapping where the information per site is optimized. However, using renormalization group theory, the information content per site appears to increase monotonically with the number of sites.



Thus, the question of mapping becomes one of determining what information to retain in a coarse-grained model, instead of merely maximizing the information. One such way to select the information that is retained is using different criteria for determining mapping weights. Different information, as detailed below, can be captured between center of mass (COM) and center of charge mappings (COC).

1 Site COM	2 Site COM	3 Site COM	4 Site COC
No Dipole	Dipole Captured	Dipole Captured	Best Dipole
Good RDF	Poor RDF	Fair RDF	Good RDF

## Interaction Parameters

Determining the interaction parameters between CG sites is another fundamental challenge of coarse-grained modeling. In order to determine a possible coarse-grained potential energy, one must use the fundamental consistency equation of coarse-graining. This equation states that probability of finding a coarse-grained model in a given configuration should be the same as a mapped atomistic configuration.

$$p_R(\mathbf{R}^N) = \int d\mathbf{r}^n \delta(M_{RI}(\mathbf{r}^n) - \mathbf{R}_I) p(\mathbf{r}^n)$$

### Multiscale Coarse-Graining

The above equation can be manipulated to get the equation below. The equation below says that the force on a coarse-grained particle should be equal to sum of the forces on its constituent atoms. By taking advantage of this relation, the potential energy can be determined variationally. By taking advantage of linear basis sets, one can determine the basis set coefficients by a linear least squares algorithm.

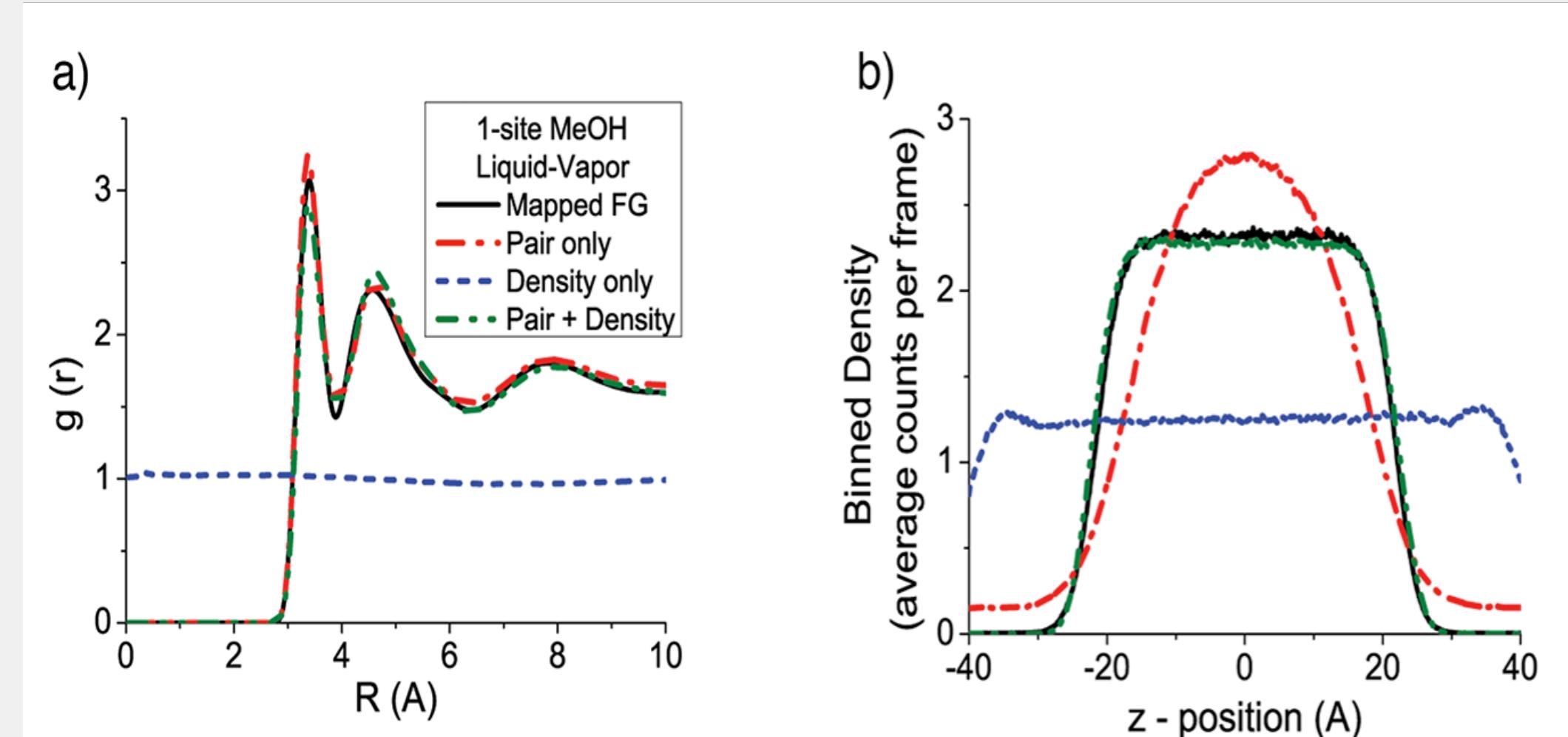
$$\mathbf{F}_I(\mathbf{R}^N) = \left\langle \sum_{i \in I} \mathbf{f}_i(\mathbf{r}^n) \right\rangle_{\mathbf{R}^N}$$
$$\chi^2 = \frac{1}{3N_t N} \sum_{t=1}^{N_t} \sum_{I=1}^N \left\| \mathbf{f}_I(\mathbf{r}_t^n) - \mathbf{F}_I(M(\mathbf{r}_t^n)) \phi \right\|^2$$

### Order Parameters

While the equations above provide the framework for determining what the ideal forces should be, one must also determine the basis set to represent the forces. One common approach is to expand the basis set into interaction between pairs, triples, etc, until the structure is adequately reproduced. The Voth group has pioneered the use of order parameters as a basis set. This allows for the  $n$ -fold basis sets to be folded into a single set of basis functions and can be simulated at much less computational cost.

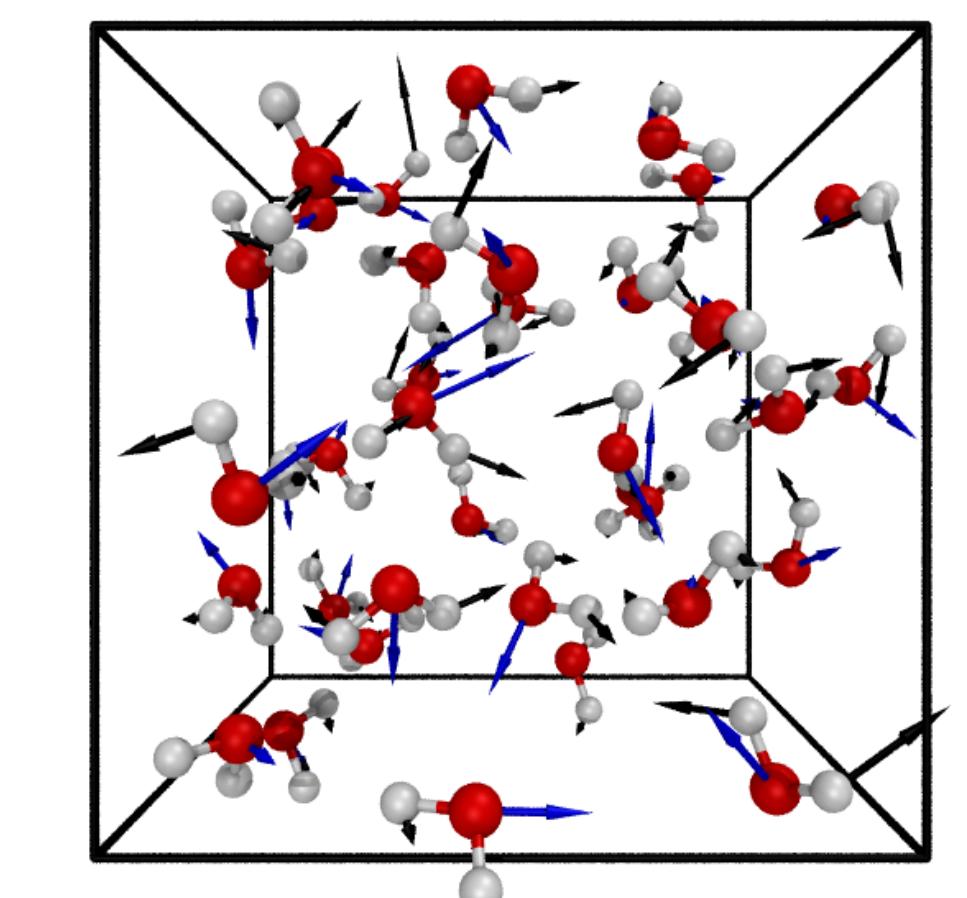
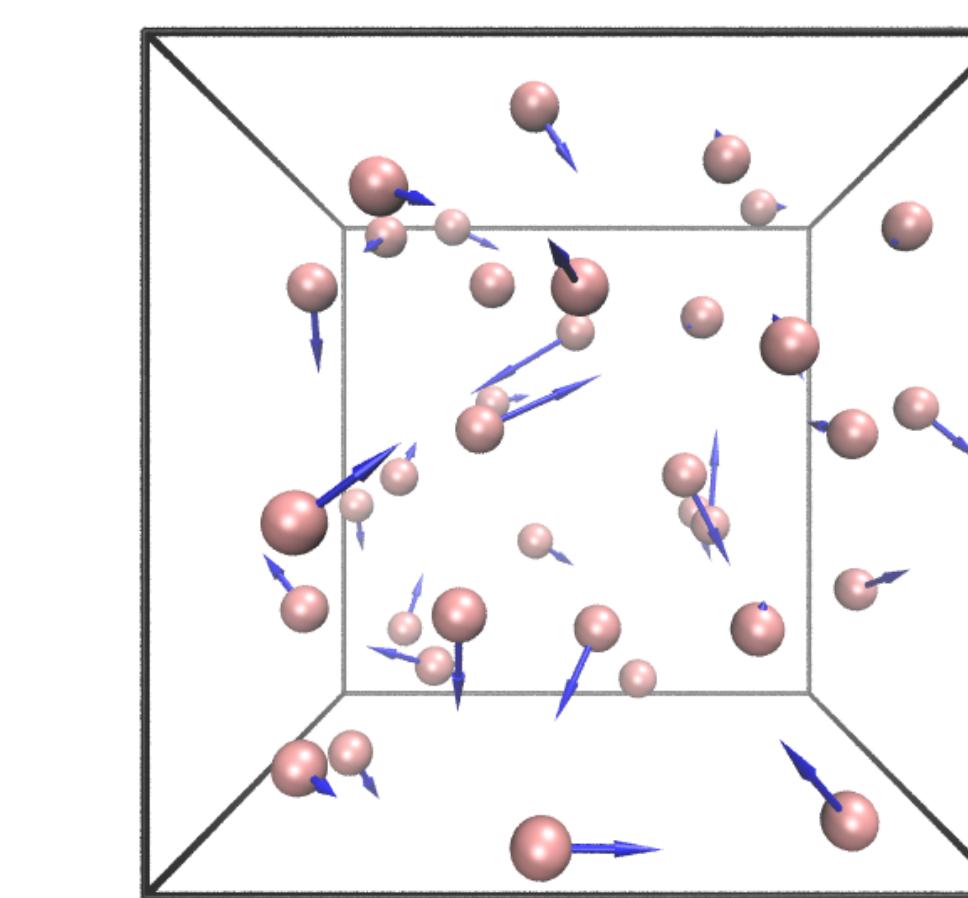
$$U_{CG} = U_1 + \sum_{i,j} u_2(r_i, r_j) + \sum_{i,j,k} u_3(r_i, r_j, r_k) + \dots$$

$$U_{CG} = U_1 + \sum_{i,j} u_2(r_i, r_j) + \sum_{\{i\}} u_{OP}(\mathbf{r}^n)$$



## Observables

Because of the reduction in representation, typical observable operators cannot be used to describe the expectation values of the coarse-grained system.



For instance, if one were to use a 1-site CG representation of water, as illustrated above, the usual virial could not be used to describe the pressure, even if one could perfectly reproduce the dynamic evolution of the system. The Voth group has investigated the fundamental theory that explains what observables must satisfy for thermodynamic consistency to be achieved. Using the equation below, we can relate the CG and FG observables.

$$A_{CG}(\mathbf{R}^N) = \left\langle A_{FG}(\mathbf{r}^n) \right\rangle_{\mathbf{R}^N} = \frac{\int d\mathbf{r}^n A_{FG}(\mathbf{r}^n) \delta(M(\mathbf{r}^n) - \mathbf{R}^N) e^{-\beta U_{FG}(\mathbf{r}^n)}}{e^{-\beta U_{CG}(\mathbf{R}^N)}}$$

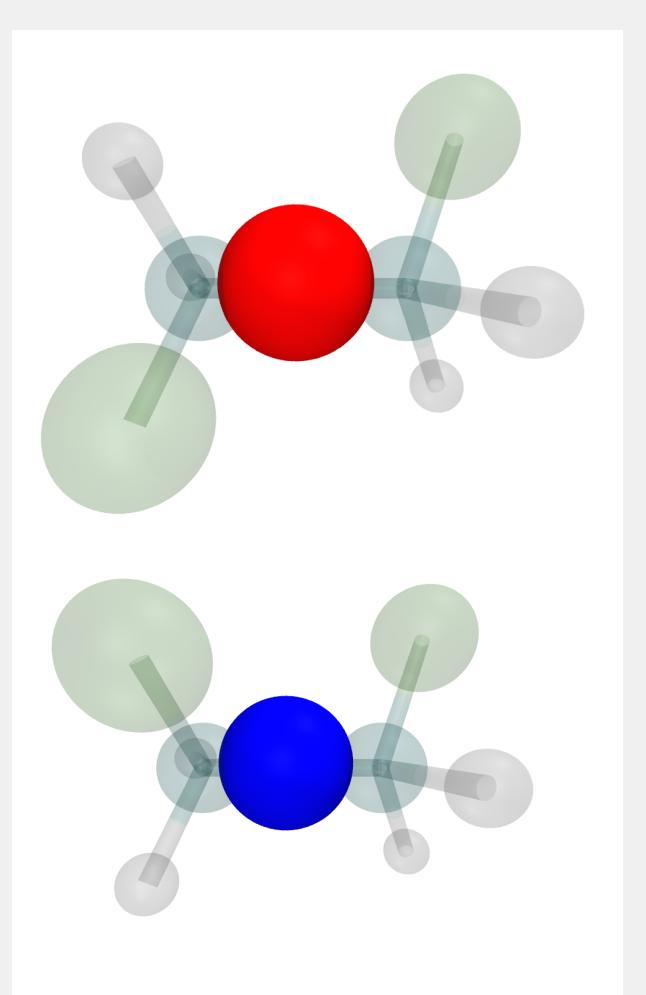
### State Dependence

Coarse-grained models usually do a good job at representing a single conformation of a protein or a single state of a molecule. But, multiple protein conformations or different bonding topologies cannot be captured by normal CG techniques. The Voth group has developed two frameworks to capture state dependence in CG models.

#### Ultra Coarse-Graining

Ultra-coarse-graining allows one to construct CG models at lower resolution while still capturing discrete conformational or chemical changes within CG sites. These discrete states are described using different interactions.

$$\frac{\partial U(v, \mathbf{R}^{Nv})}{\partial \mathbf{R}^N} = \left\langle M_{R,v}^{Nv} \left( \frac{\partial}{\partial \mathbf{r}^n} \left( u(\mathbf{r}^n) - \frac{1}{\beta} \ln p_\Sigma(v; \mathbf{r}^n) \right) \right) \right\rangle_{\mathbf{R}^{Nv}, v}$$



#### Reactive Coarse-Graining

Reactive coarse-graining allows one to capture chemical reactions with changing bonding topologies using principles from empirical valence bond theory.

