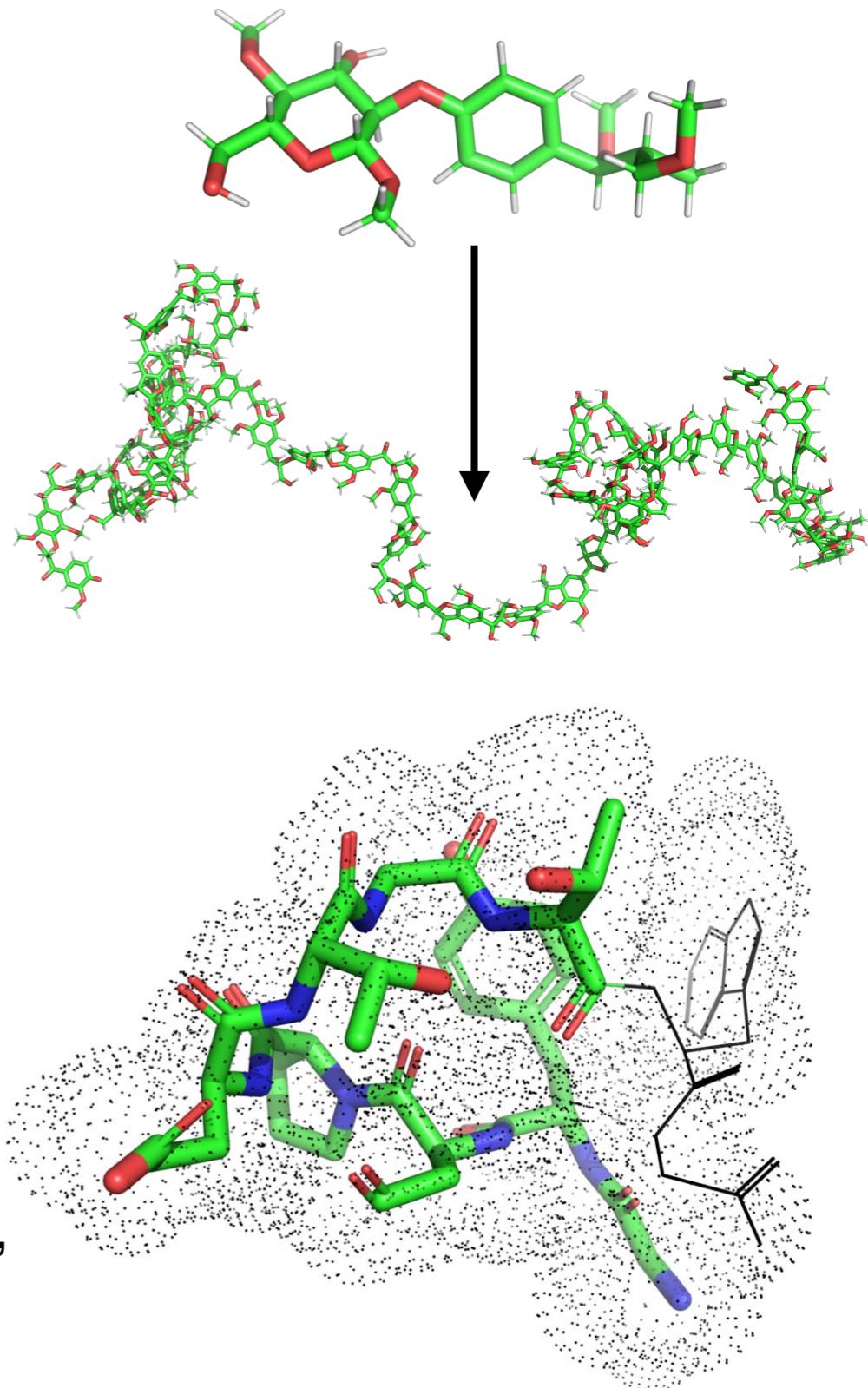


Complexity Reduction Overview

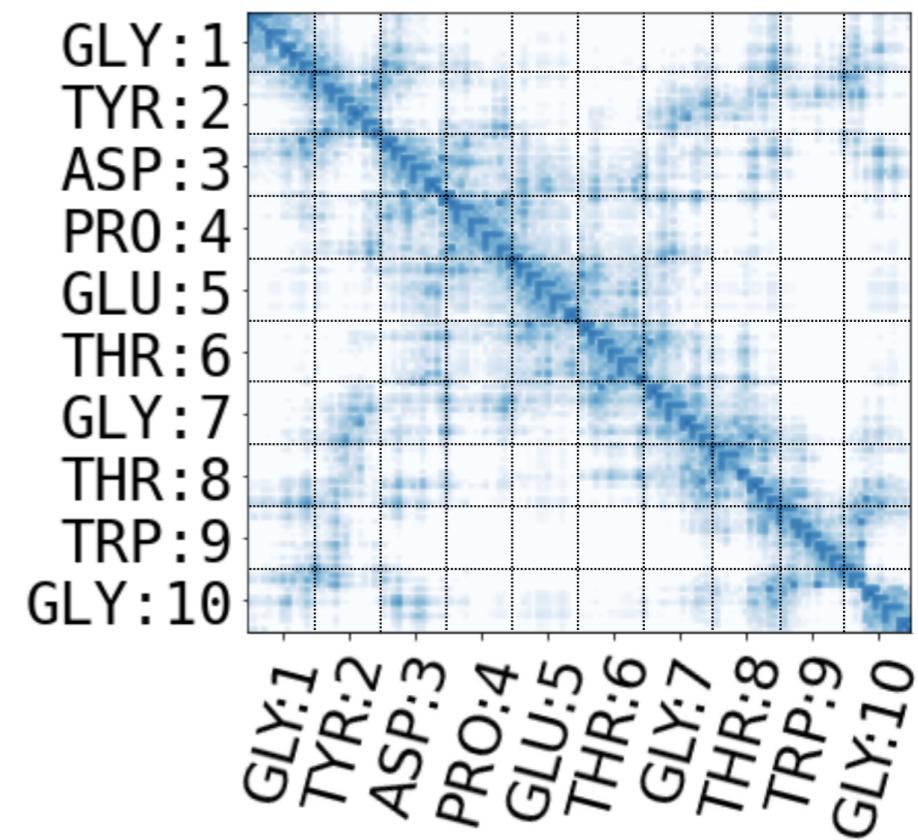
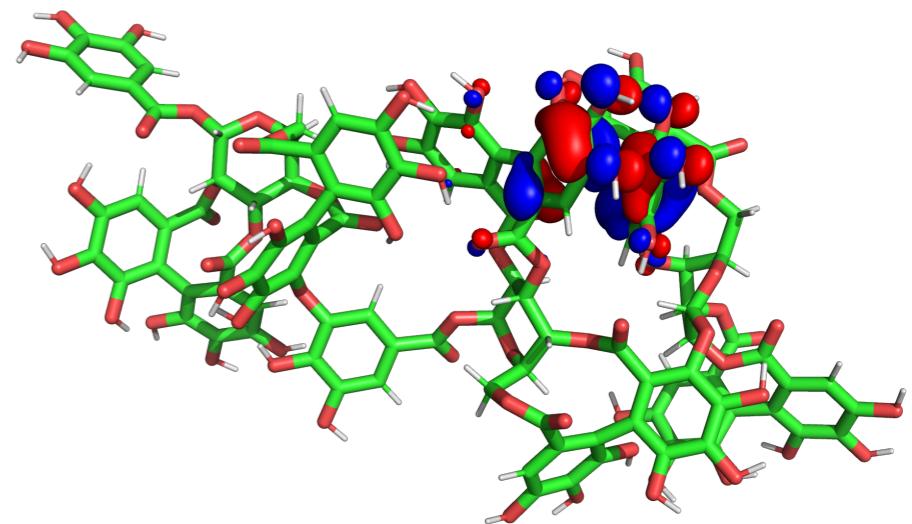
Motivation - Barriers to Calculation

- When DFT could only be applied at the small scale, you knew that at best you were treating a model system, and planned your study accordingly.
- But with a code like BigDFT, you might have the ambition to perform a simulation at the actual length scale of experiments.
- Setting up a complex system
 - Preparation of complex input geometry.
 - DFT requires all atoms to be there, knowledge of the charge of the system, reasonable geometries to avoid convergence issues.
 - It is not feasible to run ab-initio molecular dynamics of such systems, so multi-scale modeling is essential.
- Don't imagine DFT as a “super forcefield”. Instead, use it to complement your favorite techniques.



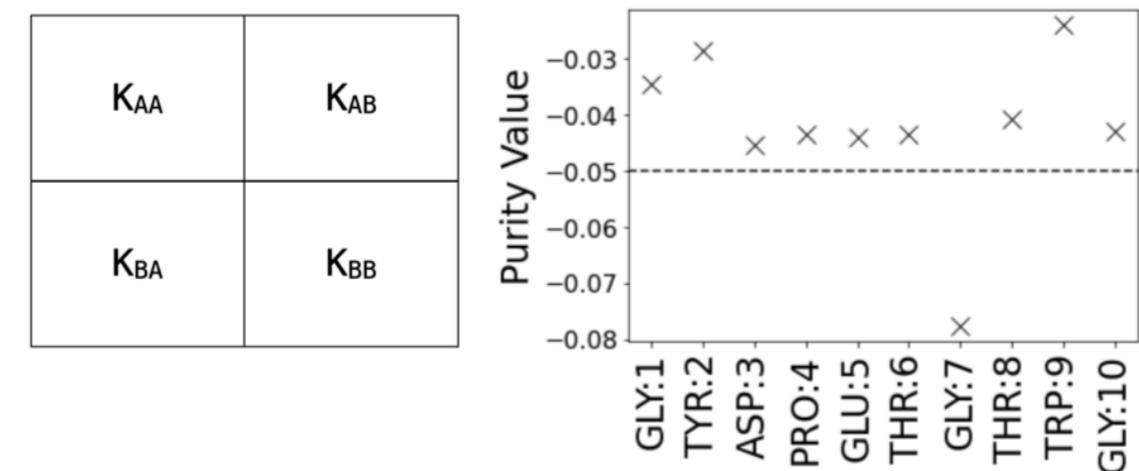
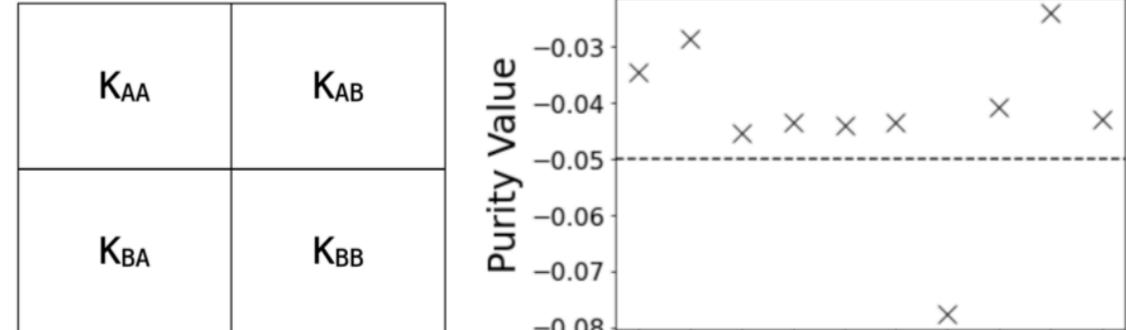
Analysis Techniques At Scale

- Unlike classical forcefields, DFT explicitly simulates the electrons.
- The electron density is represented by the single particle density matrix, which an NxN matrix where N is the number of basis functions.
- For large systems, this matrix becomes sparse (i.e. matrix elements of basis functions separated by a large distance have a very small value).
- Since each basis function is centered on a given atom, we can assign blocks of the matrix to groups of atoms (fragments).
- How do we define these groups and what can we learn from them?



Complexity Reduction Framework

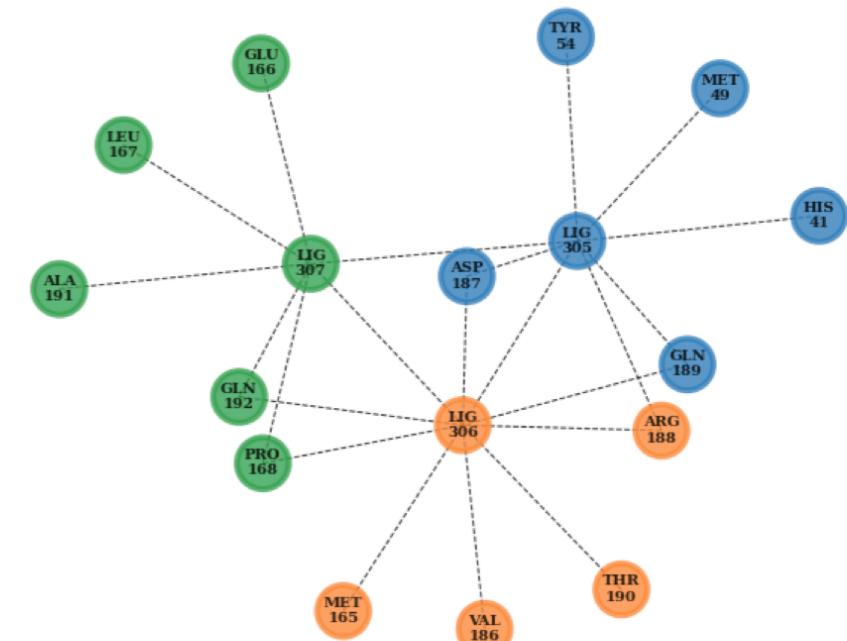
- One of the properties of the density matrix (ignoring the overlap matrix for simplicity) is that it is Idempotent, i.e. $K^*K = K$.
- We imagine splitting a system into two fragments: A and B.
- We define the fragment “Purity Value” π_{AA} as the normalized deviation from idempotency.
- Deviation from idempotency comes from neglecting the off diagonal blocks. Thus we define a second measure, the “Fragment Bond Order” B_{AB} , using the trace of off diagonal blocks.
- Prospect:
 - Automatically partition a system using the purity indicator as guide.
 - Generate a graph of interactions using the fragment bond order.
 - Project observables (interaction energies) on to fragment pairs.



$$q_{AA}\Pi_{AA} = \text{Tr}(\boxed{K_{AA}} \times \boxed{K_{AA}} - \boxed{K_{AA}})$$

$$B_{AB} = \text{Tr}(\boxed{K_{BA}} \times \boxed{K_{AB}})$$

$$q_K\Pi_K = q_{AA}\Pi_{AA} + B_{AB} + B_{BA} + q_{BB}\Pi_{BB} = 0$$



Case Study: Evolutionary Biology

- E484K is a characteristic mutation of the spike protein of the Beta and Gamma variant of SARS-CoV-2.
- By analyzing the electronic structure the spike bound to human ACE2, we can see the affect of a point mutation at 484.
- We identify E484 as a binding weak point with hACE2 due to electrostatic repulsion.
- Mutation E484K (negative to positive charge) is an evolutionary solution to the deleterious interaction.

