Avi Purkayastha
Computational Science Center

**Trichoderma reesei CBH I**

- Linker
- Catalytic Domain
- Carbohydrate-binding module
- Cellulose
REMD Method for Enhanced Sampling

- Useful for proteins to see where there stable states are.
- Multiples copies or *replicas* are run for the same systems but at different temperatures.
- Every few fixed timesteps, exchange between temperatures neighbors occur based on criterion for energy weighted by temperatures.
- Overall idea is to have the lowest temperature replica eventually be populated by the most stable state(s).
- Proteins whose stable states are known, are easier to identify.

Generalized Born using Multiple Vols

- The Generalized Born (GB) model is an approximation to the exact Poisson-Boltzmann equation.
- The GBMV method is faster, gives accurate estimation of the born radii, and more stable than the linearized PB equation.

\[
G_s = \frac{1}{8\pi} \left[ \frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right] \sum_{i,j} q_i q_j f_{GB}
\]

where
\[
f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}}
\]

and
\[
D = \left( \frac{r_{ij}}{2a_{ij}} \right)^2, \quad a_{ij} = \sqrt{a_i a_j}
\]
Computational Strategies

- The GBMV fortran90 module has about 20 functions/subroutines, and few 1000s lines of codes.
- Figure out the functions with the highest computational intensity.
- Design for the GPU implementations on those functions.
- One option is to use pgfortran for CUDA extensions.
- Another option is to have wrappers from the fortran functions to separately written CUDA/OpenCL codes.
- Iterate on the GPU design.