Nucleation of Protein Crystals and Strong Scaling on Summit

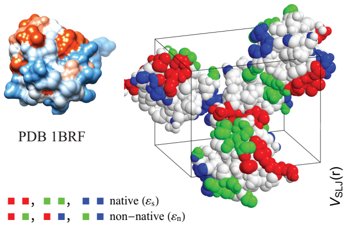
Jens Glaser, Peter Schwendeman, Joshua A Anderson, and Sharon C Glotzer
Department of Chemical Engineering, University of Michigan

ARO #W911NF-15-1-0185

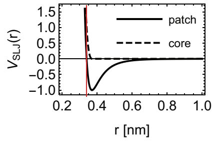
DOE OLCF INCITE (Titan)

Protein crystallization is a routinely practiced and important technique for structure elucidation of biological macromolecules, yet its mechanisms are poorly understood. Addressing the underlying fundamental question poses formidable theoretical challenges at the intersection of structural biology, chemical engineering and multiscale simulation. Here, we use molecular dynamics simulations and Metadynamics free energy methods to unravel pathways to the crystallization of a patchy shape model of rubredoxin. Our large scale simulations feature realistic molecular shape and reproduce the experimentally observed $P2_12_12_1$ structure. The model demonstrates the importance of non-specific shape and non-native contacts for crystallization, and includes a tunable pathway from direct to two-step nucleation.

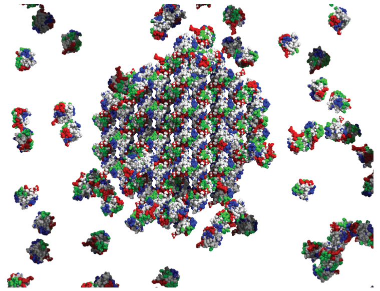
Patchy shape model for rubredoxin in $P2_12_12_1$ unit cell


PDB 1BRF
■ ■, ■ ■, ■ ■ native (ϵ_s)
■ ■, ■ ■, ■ ■ non−native (ϵ_n)

$$\begin{pmatrix} \epsilon_n \\ \epsilon_s & \epsilon_n \\ \epsilon_n & \epsilon_n & \epsilon_n \\ \epsilon_n & \epsilon_n & \epsilon_s & \epsilon_n \\ \epsilon_n & \epsilon_n & \epsilon_n & \epsilon_n & \epsilon_n \\ \epsilon_n & \epsilon_n & \epsilon_n & \epsilon_n & \epsilon_s & \epsilon_n \end{pmatrix}$$

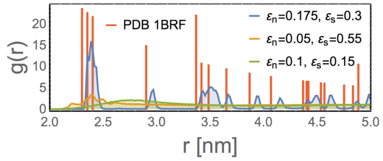


Self-assembly reproduces the experimental structure

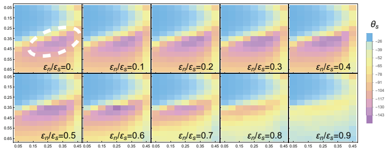


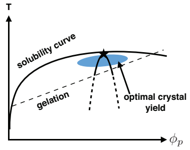
Rigid-body NVT in HOOMD-blue 2.2

Radial distribution function

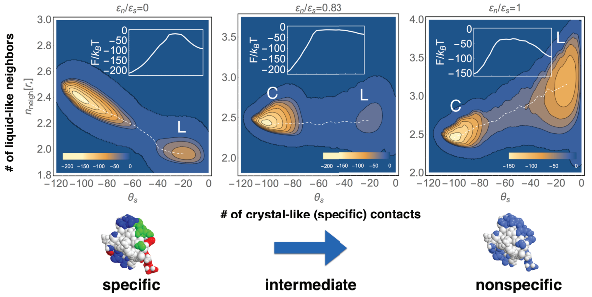


Phase diagrams exhibit narrow crystallization gap

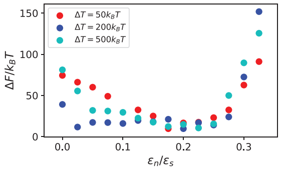




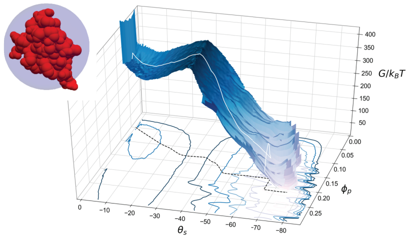
Free energy elucidates role of non-native contacts



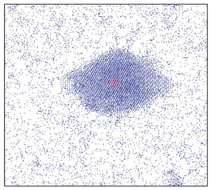
Non-native contacts facilitate spontaneous nucleation



3-phase coexistence is enhanced by isotropic attraction

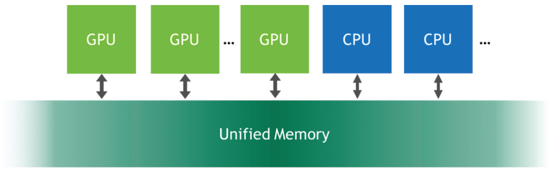


Outlook:
GFP assembly



Strong scaling of HOOMD-blue on Summit

CUDA Unified memory



Work distribution on multiple GPUs in a node

TwoStepNVTMTKGPU.cu

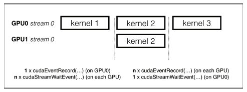
```
// iterate over active GPUs in reverse, to finish on GPU 0
for (int idev = gpu_partition.getNumActiveGPUs() - 1; idev >= 0; --idev)
    {
    auto range = gpu_partition.getRangeAndSetGPU(idev);

    unsigned int nwork = range.second - range.first;
    unsigned int offset = range.first;

    // setup the grid to run the kernel
    dim3 grid( (nwork/run_block_size) + 1, 1, 1);
    dim3 threads(run_block_size, 1, 1);

    // run the kernel, starting with offset
    gpu_nvt_mtk_step_one_kernel<<< grid, threads >>>(d_pos, .., nwork, offset);
    }
```

Explicit synchronization and memory usage hints



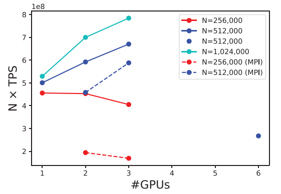
1 x cudaEventRecord(…) (on GPU0)
n x cudaStreamWaitEvent(…) (on each GPU)

n x cudaEventRecord(…) (on each GPU)
1 x cudaStreamWaitEvent(…) (on GPU0)

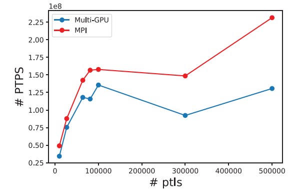
```
GlobalArray<unsigned int> member_idx(member_tags.size(), m_pdata->getExecConf());
m_member_idx.swap(member_idx);

#ifdef ENABLE_CUDA
if (m_pdata->getExecConf()->isCUDAEnabled())
    {
    cudaMemAdvise(m_member_idx.get(),
        m_member_idx.getNumElements()*sizeof(unsigned int),
        cudaMemAdviseSetReadMostly, 0);
    CHECK_CUDA_ERROR();
    }
#endif
```

Strong Scaling performance (LJ Liquid)



Summit/V100 (NVLINK 2)



2xP100 Workstation
(NVLINK 1)