Supplementary Information: Alchemical Molecular Dynamics for Inverse Design

Pengji Zhou[a,*], James C. Proctor[b,*], Greg van Anders[a,c,d], and Sharon C. Glotzer[a,b,c,e]

[a]Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan, USA; [b]Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan, USA; [c]Department of Physics, University of Michigan, Ann Arbor, Michigan, USA; [d]Department of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, Ontario, Canada; [e]Biointerfaces Institute, University of Michigan, Ann Arbor, Michigan, USA; * P. Z. and J. C. P. contributed equally to this work.

1. Intermediate Steps in Alchemostat Derivation

Following Eq. (10) and Eq. (11) in the main text and incorporating the virtual variables for both the canonical and alchemical degrees of freedom with the traditional Nosé-Hoover conventions (i.e. $t' \equiv \frac{t}{s}$, $p' \equiv \frac{p}{s}$, $q' \equiv q$, and $s' \equiv s$) we arrive at:

$$\begin{aligned}
\frac{dq'}{dt'} &= \frac{p'}{m} \\
\frac{dp'}{dt'} &= -\frac{\partial U(q',\alpha')}{\partial q'} - p'\left(\frac{s'p'_s}{Q}\right) \\
\frac{d\ln s'}{dt'} &= \frac{s'p'_s}{Q} \\
\frac{d}{dt'}\left(\frac{s'p'_s}{Q}\right) &= \frac{1}{Q}\left(\frac{p'^2}{m} - \frac{L}{\beta}\right) \\
\frac{d\alpha'}{dt'} &= \frac{p'_\alpha}{M} \\
\frac{dp'_\alpha}{dt'} &= -\mu_\alpha N - \frac{\partial U(q',\alpha')}{\partial \alpha'} - p'_\alpha\left(\frac{{s_\alpha}'p'_{s_\alpha}}{Q_\alpha}\right) \\
\frac{d\ln {s_\alpha}'}{dt'} &= \frac{{s_\alpha}'p'_{s_\alpha}}{Q_\alpha} \\
\frac{d}{dt'}\left(\frac{s'_\alpha p'_{s_\alpha}}{Q_\alpha}\right) &= \frac{1}{Q_\alpha}\left(\frac{p'^2_\alpha}{M} - \frac{L_\alpha}{\beta_\alpha}\right)
\end{aligned} \tag{1}$$

Following Nosé-Hoover scheme, we define the relation $\xi \equiv \frac{s'p'_s}{Q}$ for both thermostat and 'alchemostat' variables. We also introduce X such that $\beta_\alpha \equiv \frac{\beta}{X}$. This introduces a scaling of the energy in the alchemical space relative to the canonical space. This

CONTACT Sharon C. Glotzer. Email: sglotzer@umich.edu

leads to the following equations:

$$\begin{aligned}
\frac{dq'}{dt'} &= \frac{p'}{m} \\
\frac{dp'}{dt'} &= -\frac{\partial U(q',\alpha')}{\partial q'} - p'\xi \\
\frac{d\ln s'}{dt'} &= \xi \\
\frac{d\xi}{dt'} &= \frac{1}{Q}\left(\frac{p'^2}{m} - \frac{L}{\beta}\right) \\
\frac{d\alpha'}{dt'} &= \frac{p'_\alpha}{M} \\
\frac{dp'_\alpha}{dt'} &= -\mu_\alpha N - \frac{\partial U(q',\alpha')}{\partial \alpha'} - p'_\alpha \xi_\alpha \\
\frac{d\ln s'_\alpha}{dt'} &= \xi_\alpha \\
\frac{d\xi_\alpha}{dt'} &= \frac{1}{Q}\left(\frac{p'^2_\alpha}{M} - \frac{L_\alpha X}{\beta}\right)
\end{aligned} \tag{2}$$

We define the Liouvillian, iL, and decompose it into a set of operators $iL \equiv iL_q + iL_p + iL'_p + iL_{\xi_\alpha} + i\tilde{L}_q + i\tilde{L}_p + i\tilde{L}'_p + iL_\xi$ using the conventions of $\sim$ for real space operators and $'$ for those which rescale time, we define the following:

$$\begin{aligned}
i\tilde{L}_q \equiv& \frac{p'}{m}\partial_q \\
i\tilde{L}_p \equiv& -\left(\frac{\partial U(q',\alpha')}{\partial q'}\right)\partial_{p'} \\
iL_\xi \equiv& \frac{1}{Q}\left(\frac{p'^2}{m} - \frac{L}{\beta}\right)\partial_\xi \\
i\tilde{L}'_p \equiv& -p'\xi\partial_{p'} \\
iL_q \equiv& \frac{p'_\alpha}{M}\partial_\alpha \\
iL_p \equiv& -\left(\mu_\alpha N + \frac{\partial U(q',\alpha')}{\partial \alpha'}\right)\partial_{p'_\alpha} \\
iL'_p \equiv& -p'_\alpha \xi_\alpha \partial_{p'_\alpha} \\
iL_{\xi_\alpha} \equiv& \frac{1}{Q}\left(\frac{p'^2_\alpha}{M} - \frac{L_\alpha X}{\beta}\right)\partial_{\xi_\alpha}
\end{aligned} \tag{3}$$

Now, to determine the time evolution of the system we use a Trotter factorization [**?**] since we desire the alchemical degrees of freedom to evolve significantly slower

than the canonical degrees of freedom:

$$\begin{aligned} e^{iL\Delta t} = e^{i\Delta t L_q/2} e^{i\Delta t L_p/2} e^{i\Delta t L'_p/2} \\ \Big(e^{i\Delta t \tilde{L}_p/2n} e^{i\Delta t \tilde{L}'_p/2n} \\ e^{i\Delta t \tilde{L}_q/n} e^{i\Delta t L_\xi/n} \\ e^{i\Delta t \tilde{L}'_p/2n} e^{i\Delta t \tilde{L}_p/2n} \Big)^n \\ e^{i\Delta t L_{\xi_\alpha}} e^{i\Delta t L'_p/2} e^{i\Delta t L_p/2} e^{i\Delta t L_q/2} \end{aligned} \tag{4}$$

Using the above Trotter factorization, we can derive the finite difference equations that implement the alchemical ensemble as described in main text Eq. (12) to Eq. (17).

2. Supplementary Figures

SI Fig. 1 shows the impact of nanocrystal size on per particle potential energy, optimized potential parameter k and ϕ in the 3D oscillating pair potential (OPP). The energy is compared against the calculated minimum for a bulk ideal BCC crystal which forms the minimum on the plot. The k and ϕ values are presented in the context of the range observed when also targeting bulk modulus. It can be seen that the alchemical values do not change significantly in this context, and fall along an asymptotic curve. We can say that our chosen system size, 4096, denoted by the vertical black line, is sufficiently on the asymptotic curve. Also, because the difference in value of optimal k and ϕ at droplet size around 4000 to more than 25000 for unbiased Alch-MD is relatively small compare to the change caused by Alch-MD simulation with bulk modulus design bias, the nanocrystal size beyond 4000 is unlikely to have significant effects on the optimized value in this work.

Similar to the 3D system, SI Fig. 2 shows the impact of nanocrystal size on the 2D Lennard-Jones-Gauss (LJG) square lattice system. For this system, in the main text, we use a particle number of 1024, denoted by the black line in SI Fig. 2. By thermalizing idealized crystals at specific sizes, other variations of the surface, such as grain boundaries, can be removed. Such ideal crystals were formed in square lattices at ideal numbers for symmetry from 4 to 12,853. The analogous bulk system was previously studied alchemically by van Anders et al.[**?**] and found to have an alchemical minimum around $r_0 = 1.383$. The presented energy minimum, converted using our normalization, can be seen to be offset as well as the largest system's average alchemical r_0. This shows that while the surface does have an effect, it is non-vanishing at large crystal sizes. Therefore approximating a bulk system would require alternative methods, such as an $NPT\mu_\alpha$ implementation. It is, however, unclear if the bulk or nanocrystal system is best for inverse design of systems which will undergo self-assembly. Above the threshold where the isoperimetric quotient (IQ) jumps to reflect rounder droplets, we observe consistent behavior approximating an asymptotic curve. This suggests that above a sufficient threshold the surface effects becomes roughly constant. The convergence of this curve is faster than the 3D OPP system, which is expected given the percentage of particles near the perimeter decreases faster in 2D. We therefore need to be on this curve, but need not compute using excessively large systems.

[Figure 1 about here.]

[Figure 2 about here.]

SI Fig. 3 shows the λ parameter dependence for Alch-MD with bulk modulus bias. λ has flexible working range with a negligible variation for values ranging from 0.1 to 0.5, for larger λ values, we see significant instability of the method leading to diverge results. We also observe significantly different behavior of the optimization process when our target bulk modulus value is significantly different from the self-assembly achievable range, suggesting that such values of bulk modulus could be physically impossible within this design space, thus leading to failure in our optimization scheme.

[Figure 3 about here.]

References

[1] G.J. Martyna, M.E. Tuckerman, D.J. Tobias and M.L. Klein, Molecular Physics **87** (5), 1117–1157 (1996).

[2] G. van Anders, D. Klotsa, A.S. Karas, P.M. Dodd and S.C. Glotzer, ACS Nano **9**, 9542–9553 (2015).

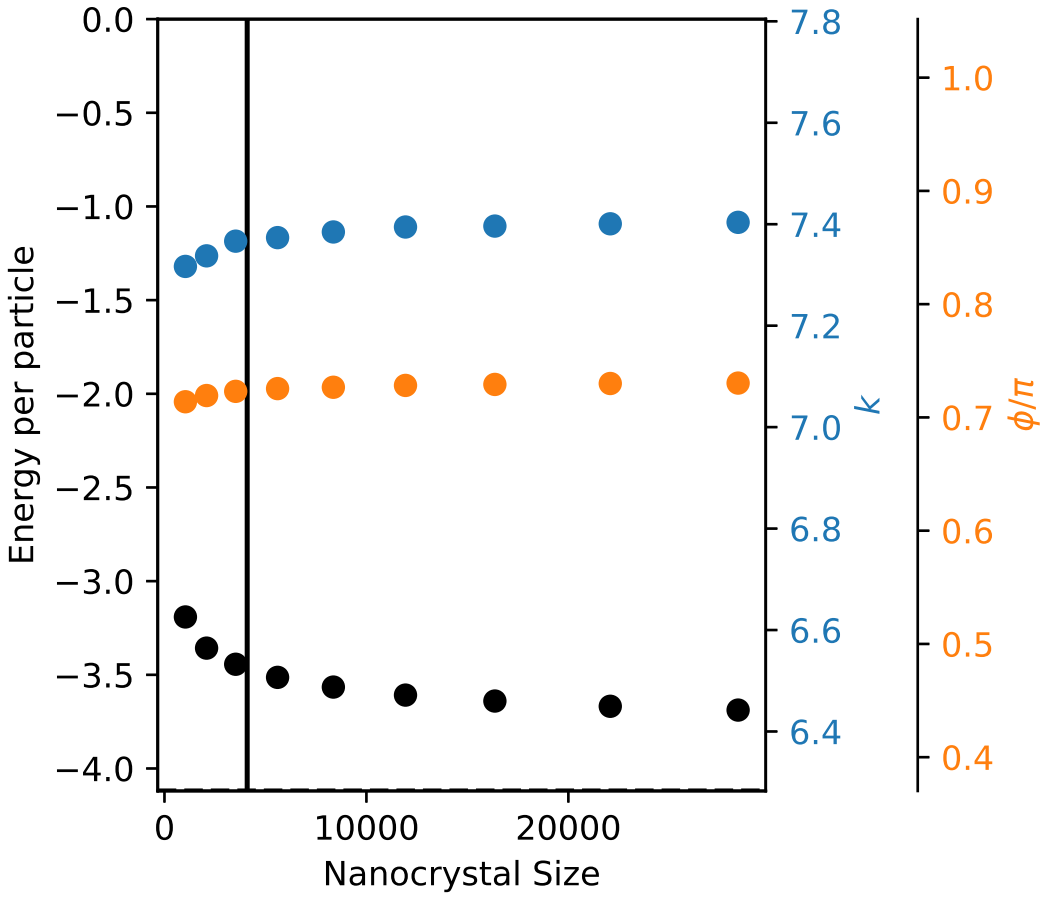

Figure 1. The effect of changing nanocrystal size in the 3D OPP system. The horizontal dashed line represents the per particle potential energy from calculations of the ideal bulk BCC crystal. The vertical line represents the nanocrystal size used in the main text.

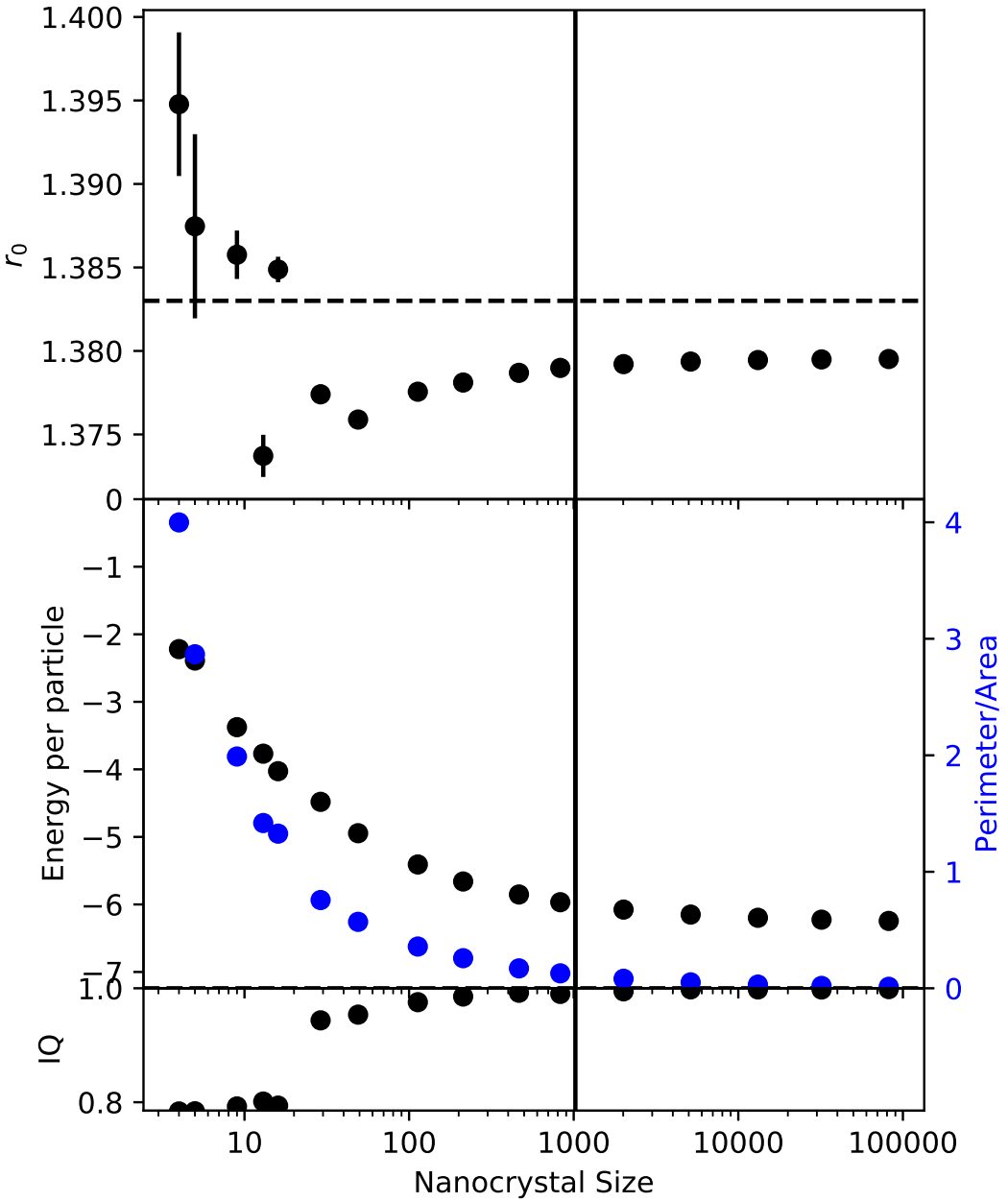

Figure 2. The effect of changing nanocrystal size is compared with previous bulk estimates by van Anders et al.[**?**], represented by dashed lines. This convex hull was used to calculate the perimeter to area ratio and the dimensionless isoperimetric quotient (IQ). The verticle line represents the system size used in the main text.

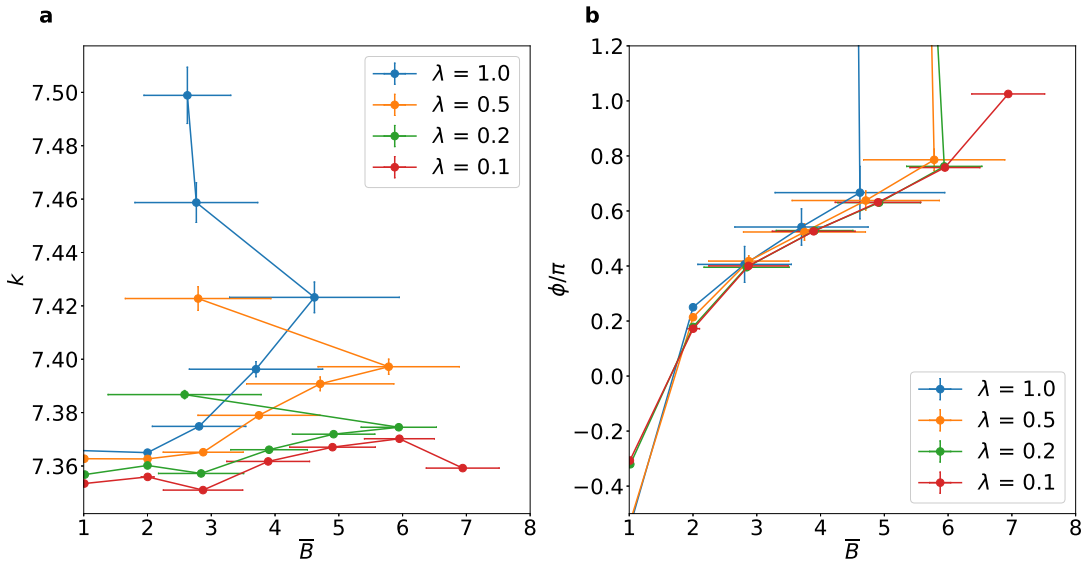

Figure 3. Design results from Alch-MD simulations with a bulk modulus bias with different Λ value.. **(a)** Correlation between targeted bulk modulus and design parameter k, with $\lambda = 0.1, 0.2, 0.5$ and 1.0 (targeting B_0 between 1.0 and 7.0) **(b)** Correlation between targeted bulk modulus and design parameter ϕ, with $\lambda = 0.1, 0.2, 0.5$ and 1.0 (targeting B_0 between 1.0 and 7.0)