LAMMPS Driven Quantum Simulations of Prebiotic Chemistry in Impacting Icy Materials

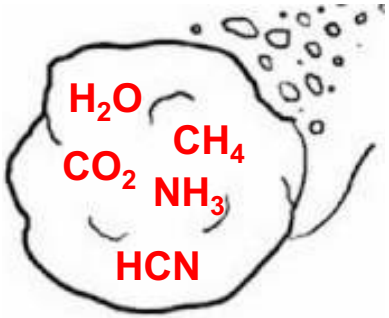
Matt Kroonblawd, Rebecca Lindsey, Nir Goldman
Lawrence Livermore National Laboratory

Lawrence Livermore National Laboratory

*This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Could impacts of interstellar ices on early Earth have created simple prebiotic compounds (e.g., amino acids, proteins, etc.)?

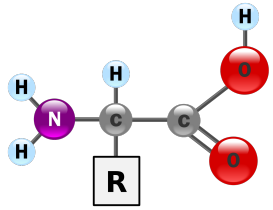
Key issue: Impacting icy materials could be a relevant source of life building compounds.



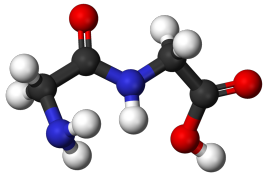
T < 5000 K

P < 100 GPa

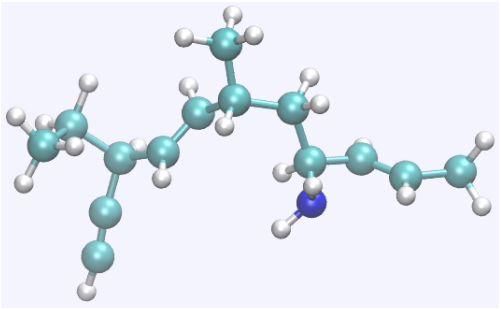
(1 GPa=10 kbar)



Amino acids



Peptides



Amines: prebiotic backbones?

Oblique impacts deliver energy for prebiotic chemistry.

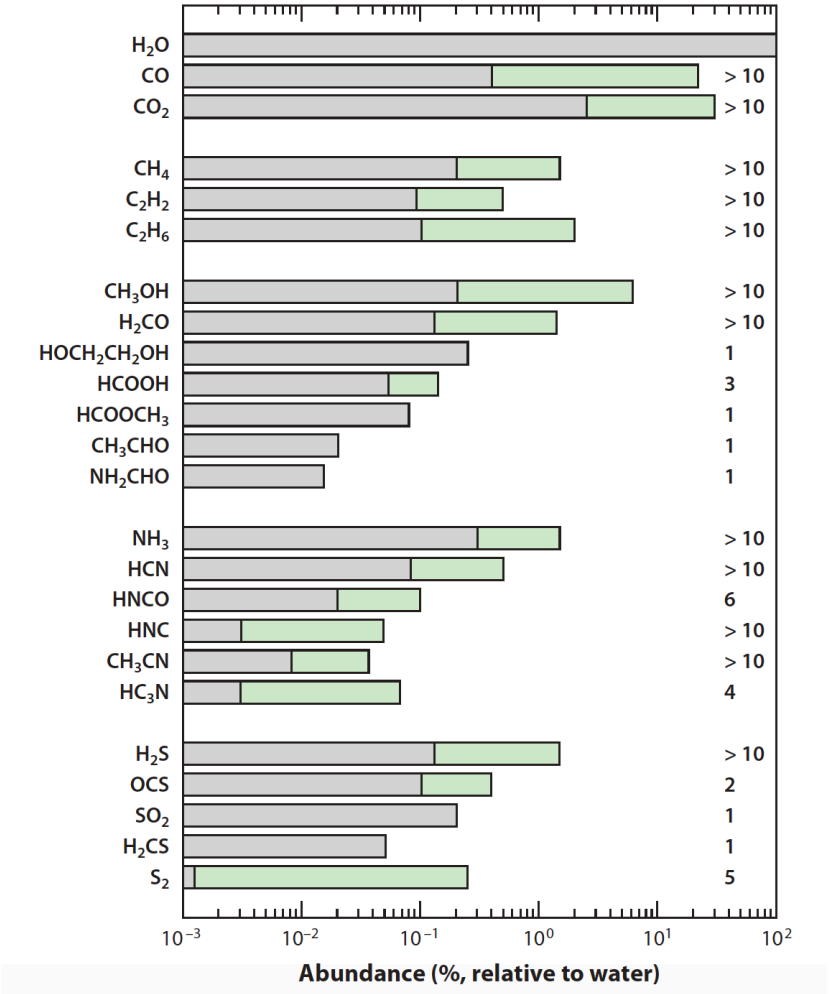
- **Prebiotic synthesis from terrestrial materials:**
  - Organic synthesis require reducing environment (H_2 and CH_4).

    Miller, *Science*, 1953, 1959.
  - Early atmospheric conditions were likely more oxidizing (CO_2 rich)

    Brack, *Chem. Biodiversity*, 2007.

- **Extra-terrestrial sources:**
  - Quantum simulations show amino acids and aromatics could be shock-synthesized from CO_2-rich mixtures: **Goldman et al., Nature Chem., 2010; Goldman et al., JPC A, 2013.**
  - Shock synthesis of amino acids observed in experiments: **Martins, et al., Nat. Geoscience, 2013**.
  - HCN polymerization can form simple and complex prebiotics: **Koziol et al., ApJ, 2015.**

Carbon containing compounds can comprise as much as 22% by concentration of a comet.



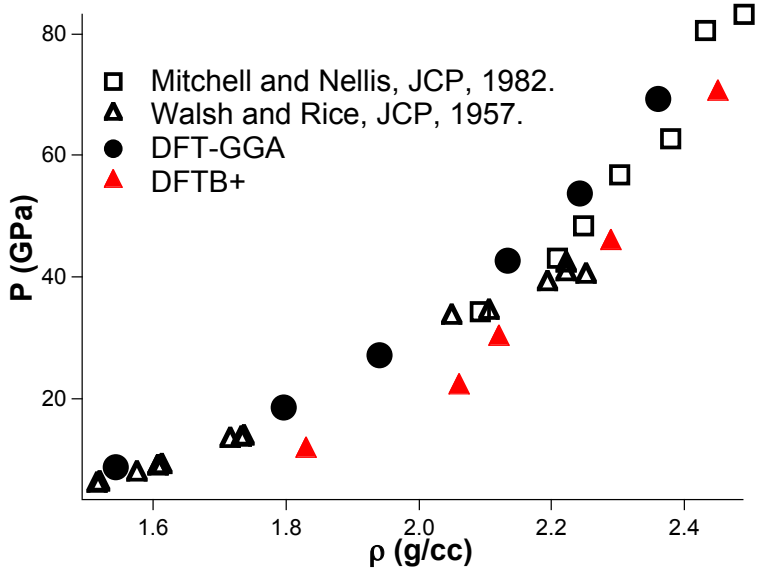
Mumma and Charnley, Annu. Rev. Astron. Astrophys. 2011. 49:471–524

Impacts result in cycling of materials through distinct thermodynamic regions:

(1) Impact (high-P, high T)

(2) Adiabatic expansion (low-P, high T)

(3) Cooling & equilibration (low-P, low-T)



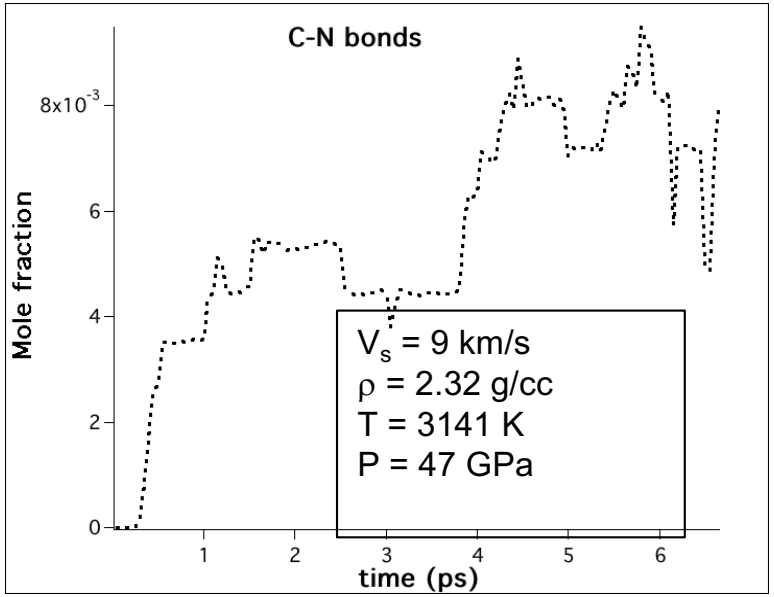
Water P vs. ρ Hugoniot

Goldman et al., JCP, 2009.

Quantum (DFT) approaches tend be very accurate under these conditions: little dependence of chemistry on specific functionals.

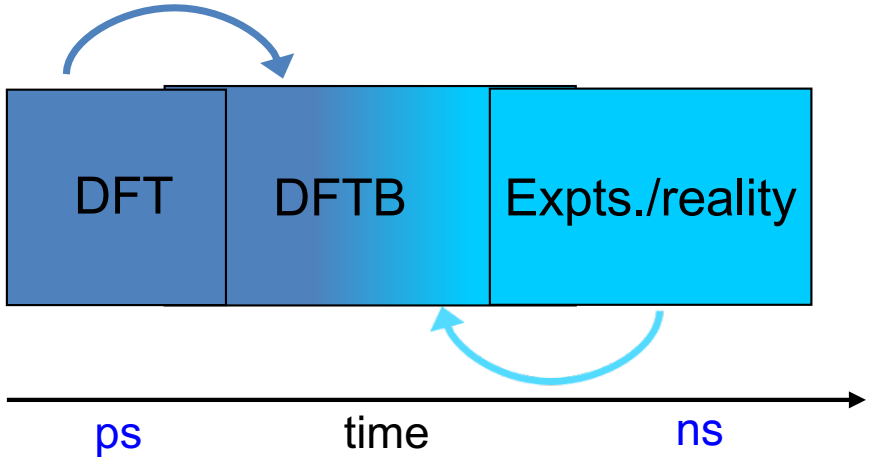
Materials under dynamic conditions can take over 1 ns to equilibrate, even at 10's of GPa and 1000's of K

- **Problem:** standard quantum simulations (Kohn-Sham DFT) cannot capture experimental time and length scale for dynamic processes.



Experiments probe nanosecond time scales and beyond, whereas DFT is limited to 10's of picoseconds.

- **Solution: Density functional tight binding (DFTB)** is a quantum mechanical semi-empirical method that can bridge these gaps with experiments.



- *DFTB*: underpinnings from Kohn-Sham Density Functional Theory (DFT):
  - *Approximate quantum mechanics combined with empirical functions.*
  - Yields 2-3 orders of magnitude increase in computational efficiency.

DFTB holds promise as a quantum method that can approach experimental timescales while retaining the accuracy of higher order methods.

The DFTB Hamiltonian assumes a neutral charge density $n_0(r)$ and expands the Density Functional Theory Hamiltonian to second order in charge fluctuation $\delta n(r)$.

DFTB total energy:

$$E_{DFTB} = \sum\sum \left\langle \phi_\mu^\alpha \middle| \hat{H}_{\mu\nu}^0 \middle| \phi_\nu^\beta \right\rangle + E_{Coul} + E_{rep}$$

Band structure energy from the Hamiltonian

E_{coul} is computed self-consistently

E_{rep} is computed via fitting to an empirical function; *usually pair-wise and short range*.

$$H_{\mu\nu}^0 = \begin{cases} \varepsilon_\mu^{\text{neutral free atom}} & \text{if } \mu = \nu \\ \langle \varphi_\mu^\alpha | \hat{T} + V_0^\alpha + V_0^\beta | \varphi_\nu^\beta \rangle & \text{if } \alpha \neq \beta \\ 0 & \text{otherwise.} \end{cases}$$

Quantum mechanical calculations exclude costly many-body interactions.

H and S matrix elements are pretabulated!

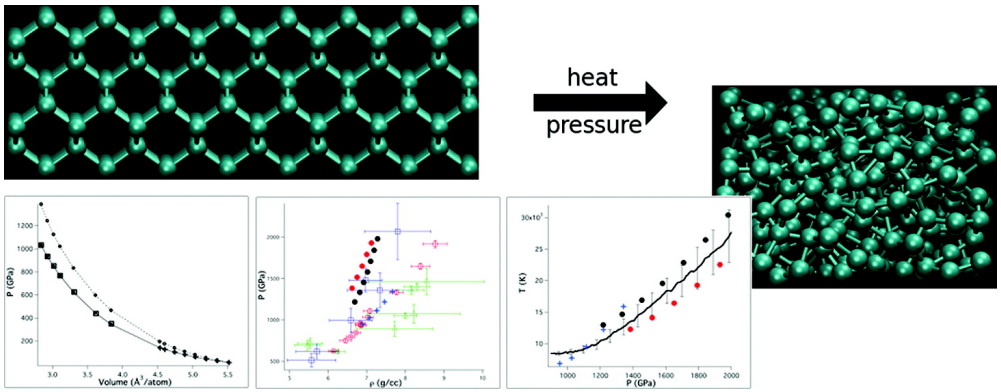
E_{rep}: We take this to be a linear combination of polynomials of very short range (≥ 2 Å).

$$E_{\text{rep}} = -\frac{1}{2}\int V_H[n_0](\boldsymbol{r})n_0(\boldsymbol{r}) + E_{xc}[n_0] + E_{II} - \int V_{xc}[n_0](\boldsymbol{r})n_0(\boldsymbol{r})$$

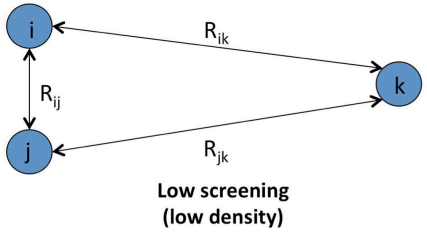
E_{rep} augments "bonded" interactions only, and all longer range interactions are taken care of in $H^0_{\mu\nu}$

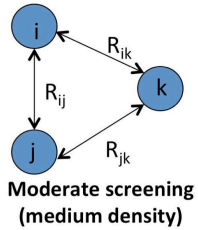
We have created a number of simulation approaches for materials under reactive conditions

- DFTB for materials under extreme pressures and temperatures (P = 10-1000 GPa):
  - Goldman, et al., JPC C, 2012.
  - Goldman, et al., JCTC, 2015.

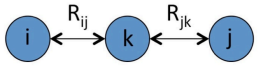


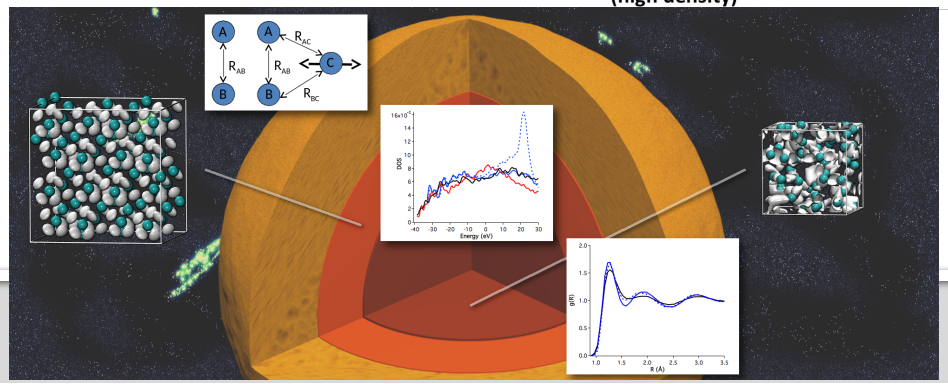
- DFTB with expanded basis set with three-body interactions for extremely high pressure e.g., P > 1000 GPa, (DFTB-p3b):
  - Goldman, et al., JPC C, 2013.
  - Srinivasan, et al., JPC A, 2014.
  - Goldman, CPL Frontiers article, 2015.



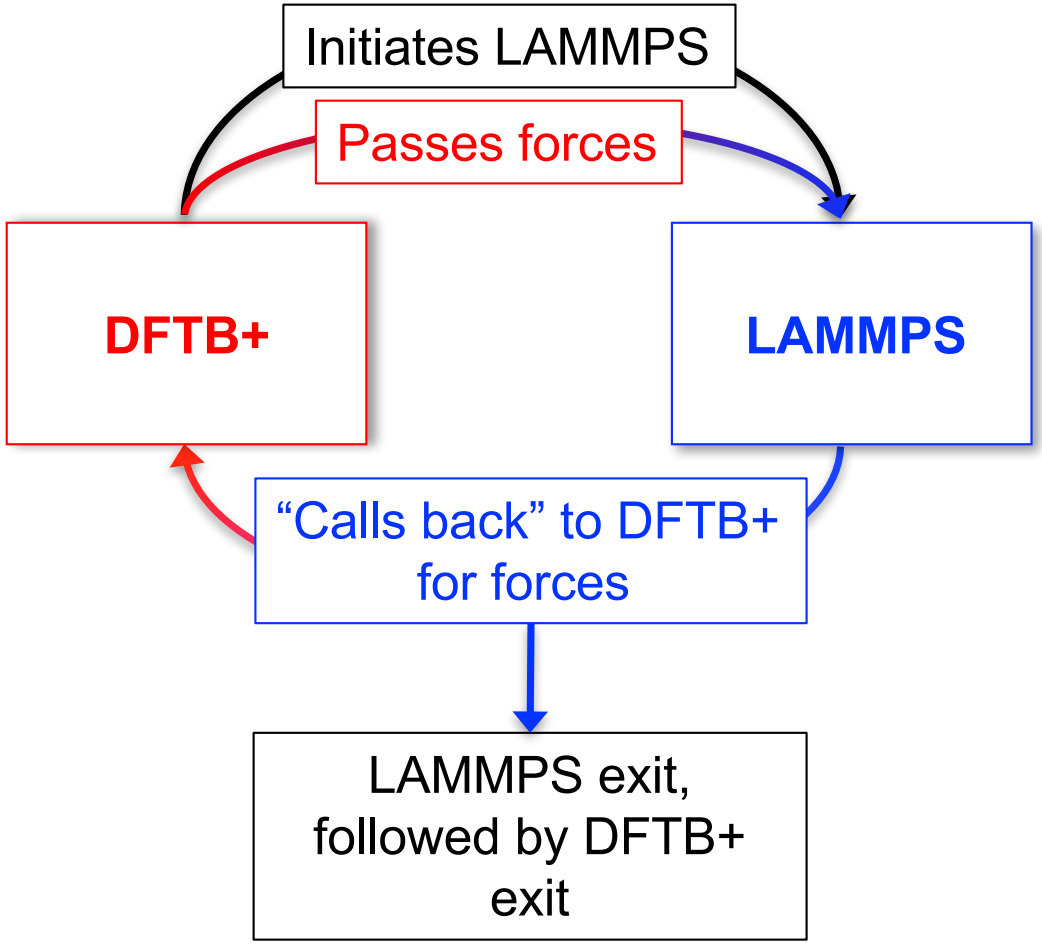


- Reactive many-body force-fields based on linear combinations of Chebyshev polynomials:
  - Koziol, et al., JCTC, 2017.
  - Lindsey, Fried, Goldman, JCTC, in prep.





We run our calculations through a hybrid code using LAMMPS and DFTB+, linked via *callback function.*



- **DFTB+:** "Master" code
  - Sets up LAMMPS input file and then instantiates LAMMPS
- **LAMMPS:** Dynamics/optimiz. driver
  - Linked in to DFTB+ as library (static or shared object).
  - Communicates with DFTB+ through a *callback* function: *fix external pf/callback*
  - Sends over positions, cell lattice vectors; DFTB+ saves atomic forces into a LAMMPS array *via pointer*.

This approach is one of the most efficient approaches to leverage all of the features of LAMMPS without having to write them directly into DFTB+ code.

Fortran code can be linked to LAMMPS through wrapper functions

- Wrapper functions (linked in via static library):
  - Fortran 2003 module which wraps all functions in src/library.h so they can be used directly from Fortran-encoded programs.
- User is required to write callback function in code to be linked.
  - Facilitated by use of ISO_C_BINDINGS
- Two example directories in examples/COUPLE: fortran2 and fortran3.
  - Our efforts (fortan3) leveraged previous work by Karl Hammond (Mizzou).
  - simple.f90: fortran code that calls DFTB+ executable through callback.
- External callback function is invokes through a fix, e.g., in the LAMMPS input deck, e.g.:
  - fix 2 all external pf/callback 1 1

Use of fix external callback is easier with a C++ code (our own in house MD code for force field development).

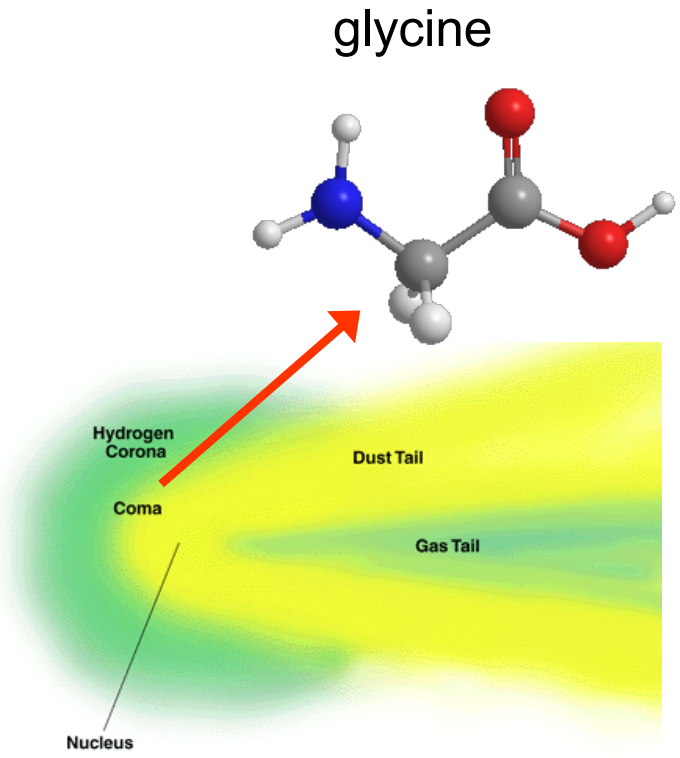
Callback function/hybrid code features

Advantages	Disadvantages
~40% more efficient (for small systems) than our previous implementation (communication through disk i/o).	Writing a mixed language code can be tricky: *function names, arrays, argument calls and types*. • Can leverage the Fortran 2003 ISO_C_BINDINGS module
Allows DFTB+ users to run LAMMPS driven calculations (mostly) through the regular DFTB+ input environment.	Usability is somewhat limited to DFTB+ users.
Arguably easier to implement that creating a DFTB+ library and linking in to LAMMPS.	

Glycine chemistry under extreme conditions could have played a significant role in the synthesis of prebiotic materials on early Earth.

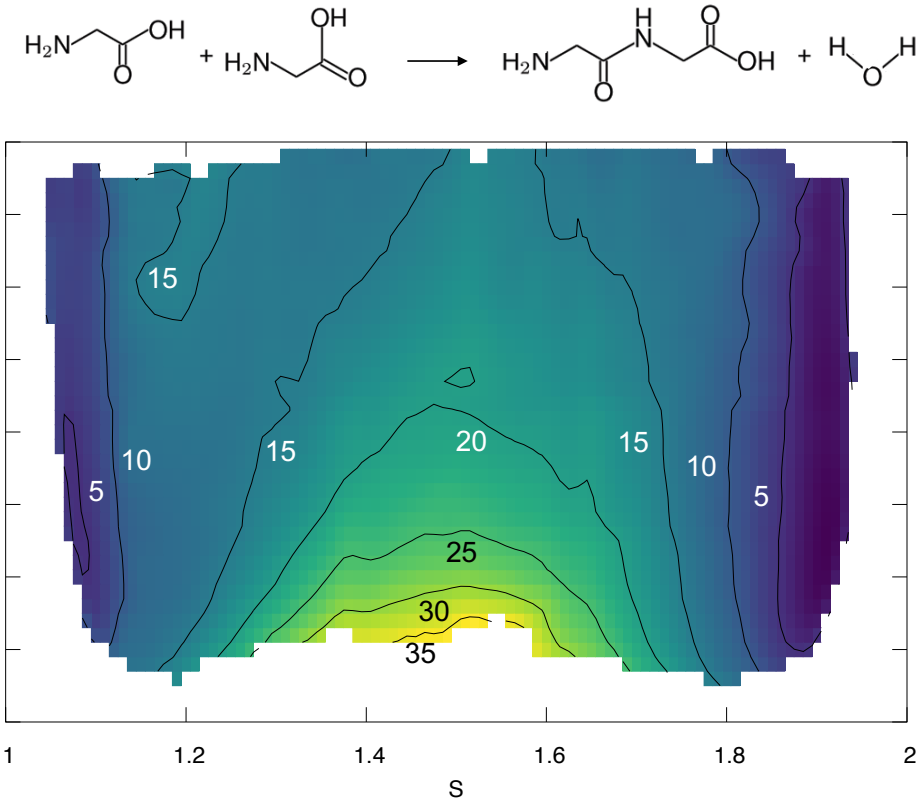
- Amino acids such as glycine could have been abundant in the Archaean ocean, possibly introduced by extraterrestrial means.
- Stepwise synthesis in submarine hydrothermal vents could have readily facilitated the multiplicative oligomerization of glycine (Matsuno, 1999).
- Glycine has been found in cometary material (Elsila, 2009), and could have been formed in the process of comet or meteoritic impact (Furukawa, 2009; Goldman, 2010; Martins, 2013).
- However, the survivability and reactivity of amino acids and under material when subjected to extreme conditions is still an open question.

We have studied the autocatalytic nature of *aqueous glycine* subjected to rapid cycling from high thermodynamic conditions.



Our DFTB models indicate that glycine can oligomerize easily in aqueous environments under ambient conditions (1 g/mL, 298 K).

Free energy surface for two solvated glycine molecules to form diglycine
(DFTB+/LAMMPS hybrid, using PLUMED plugin)

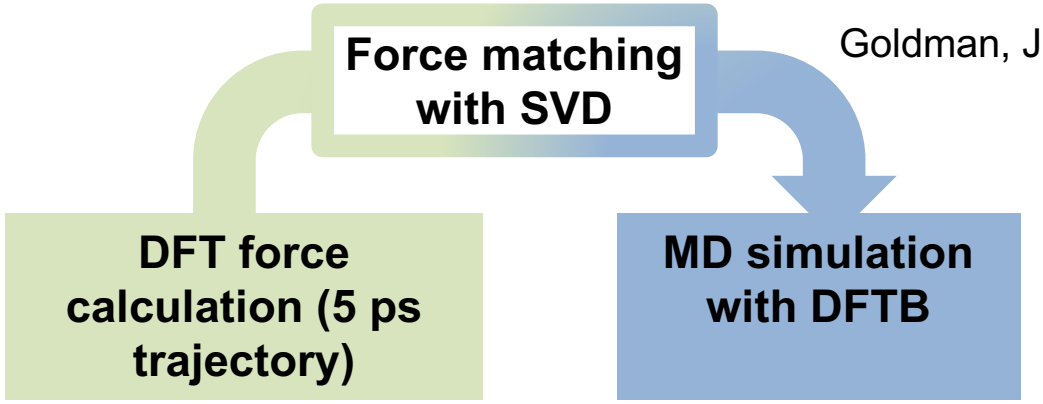


Free energy barrier = ~10 kcal/mol

- Computed at 300 K and 1 atm using our DFTB models
- Constructed from 259 independent single-core simulations, 5.2 ns combined (1 week of wall clock time)
- Cost for a similar DFT calculation would be prohibitive
- **Can feasibly compute free energy surfaces for many state points simultaneously**

Calculations are underway to determine free energy surfaces for oligomerization under a wide range of P-T conditions.

We have created semi-empirical quantum DFTB models for water/glycine mixtures using a force matching approach

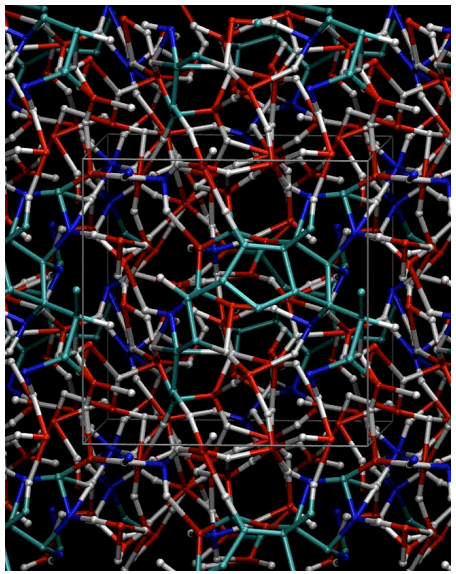


Goldman, JCTC, 2015.

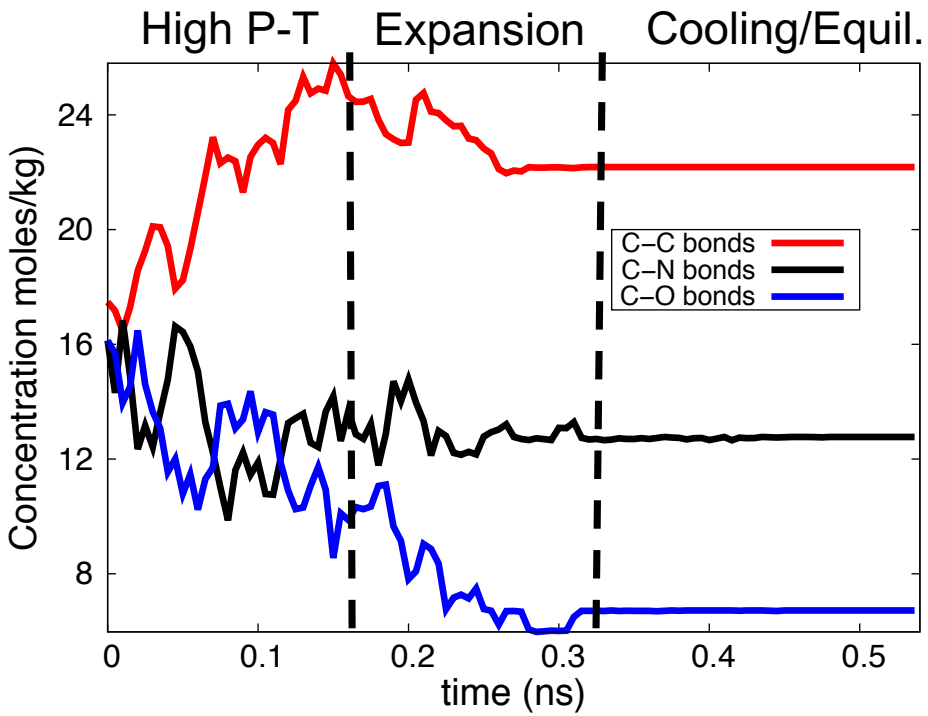
- 1:1 H_2O/glycine mixtures: 16 water molecules mixed with 16 glycine molecules, 208 atoms total (biased towards reactivity and repeated impacts). *We include dispersion interactions.*
- 10 statistically independent impact simulations: allows for some degree of ensemble averaging of chemistry.
- One DFTB simulation yields ~15 ps/day on a single CPU.
  - Standard QMD trajectories yielded ~1 ps/day using 128 processors total.
- Extreme conditions: 2.5 g/mL and 3000 K (10 systems total), followed by expansion and cooling to 1.0 g/mL and 298 K.

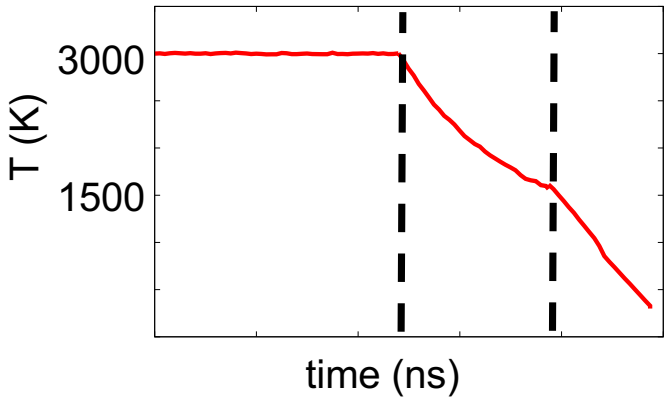
Extreme conditions yield the condensation of C-N extended structures

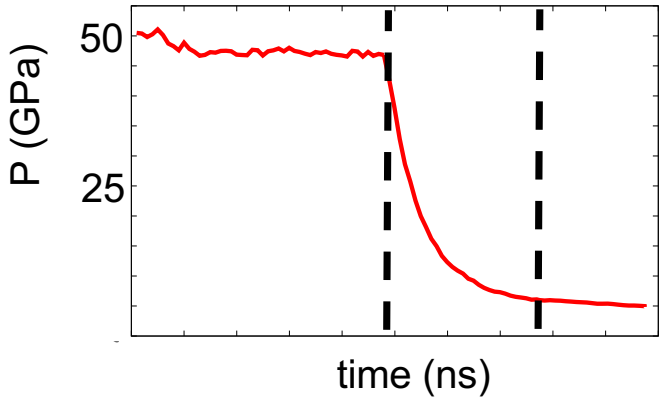
1:1 H_2O/glycine



High P-T conditions

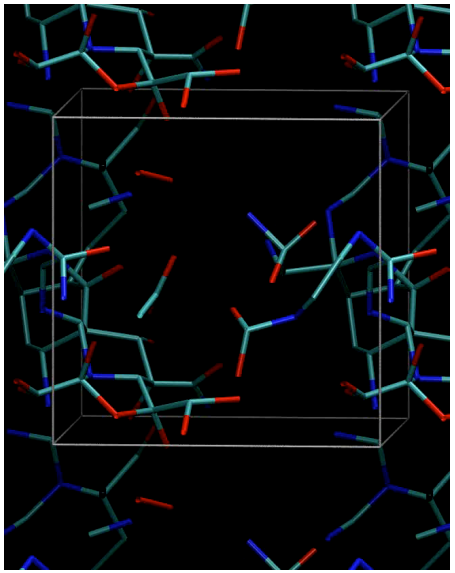




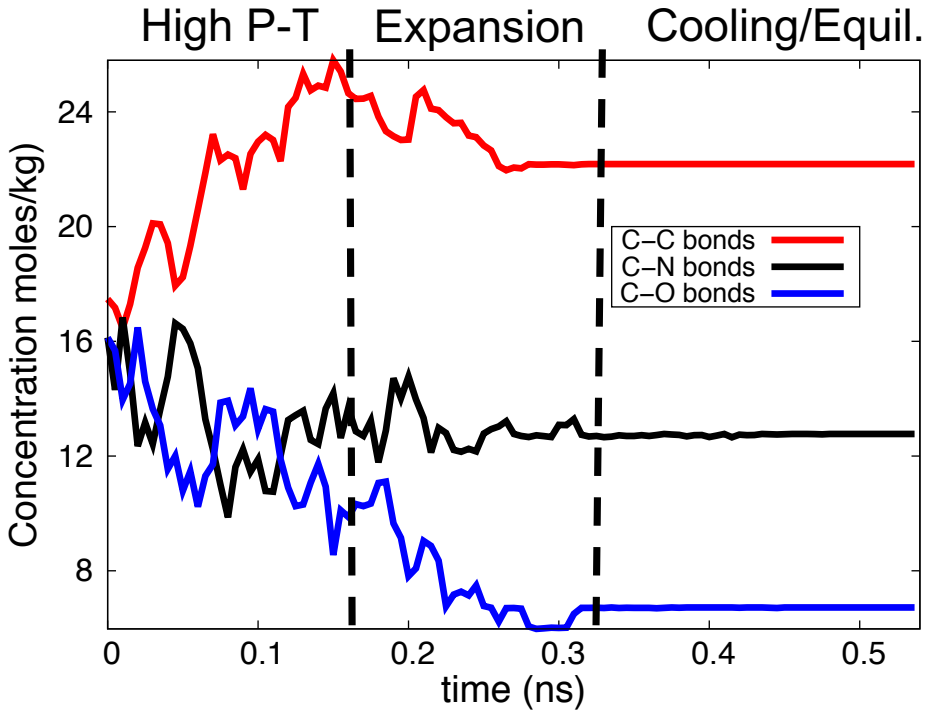


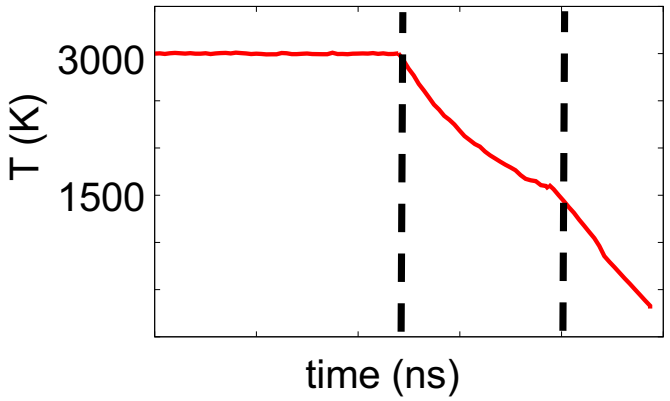
Three dimensional C-N-O structures form within the hot compressed system.

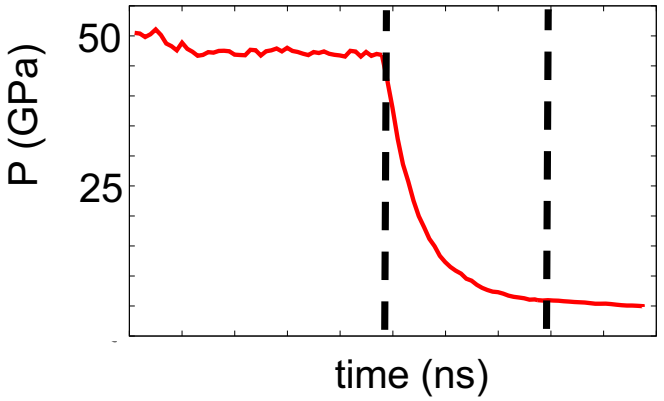
1:1 H_2O/glycine



C-N-O bonds, only (high P-T)

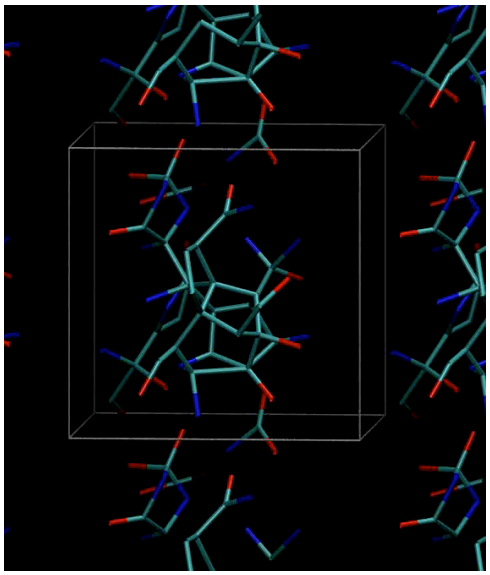




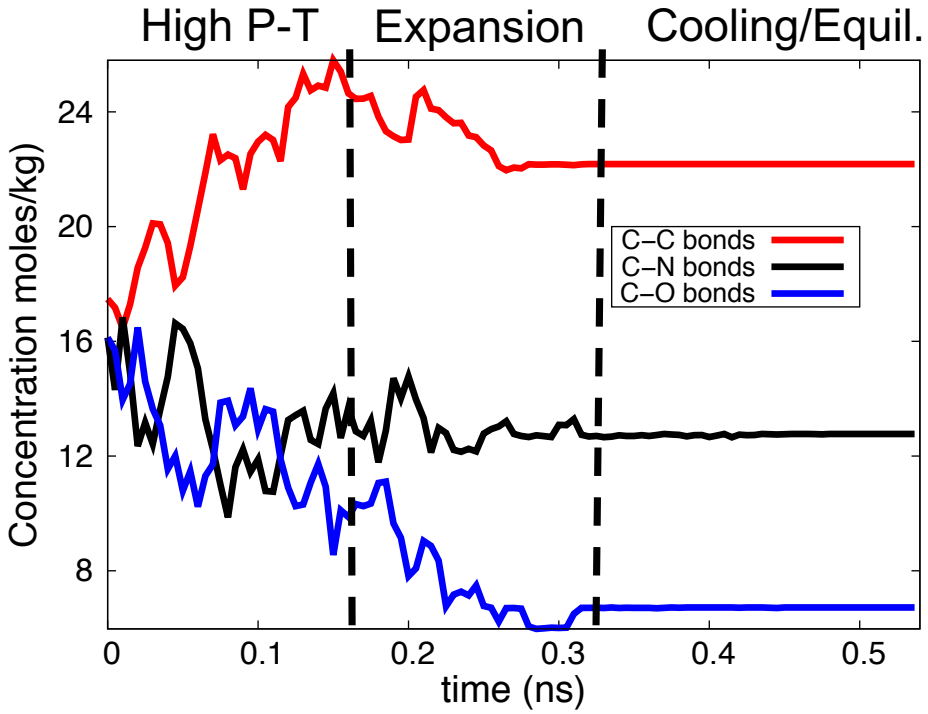


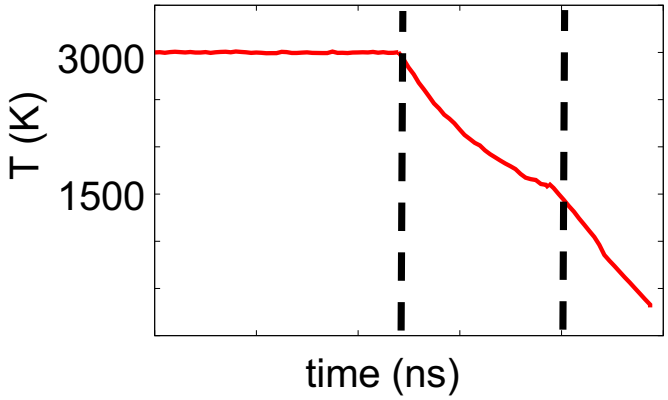
Expansion (low P, high T) yields an unfolding of 3D structures into planar sheets.

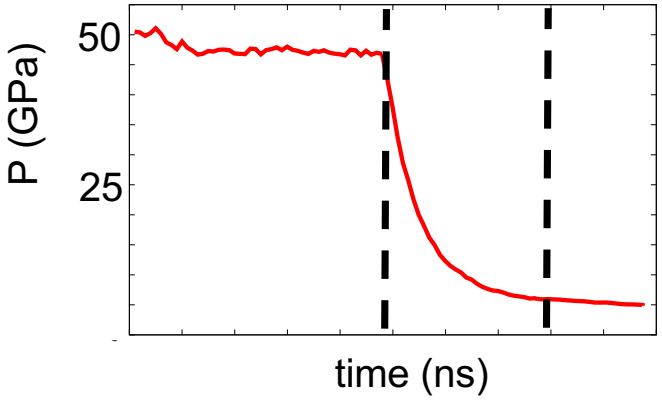
1:1 H_2O/glycine



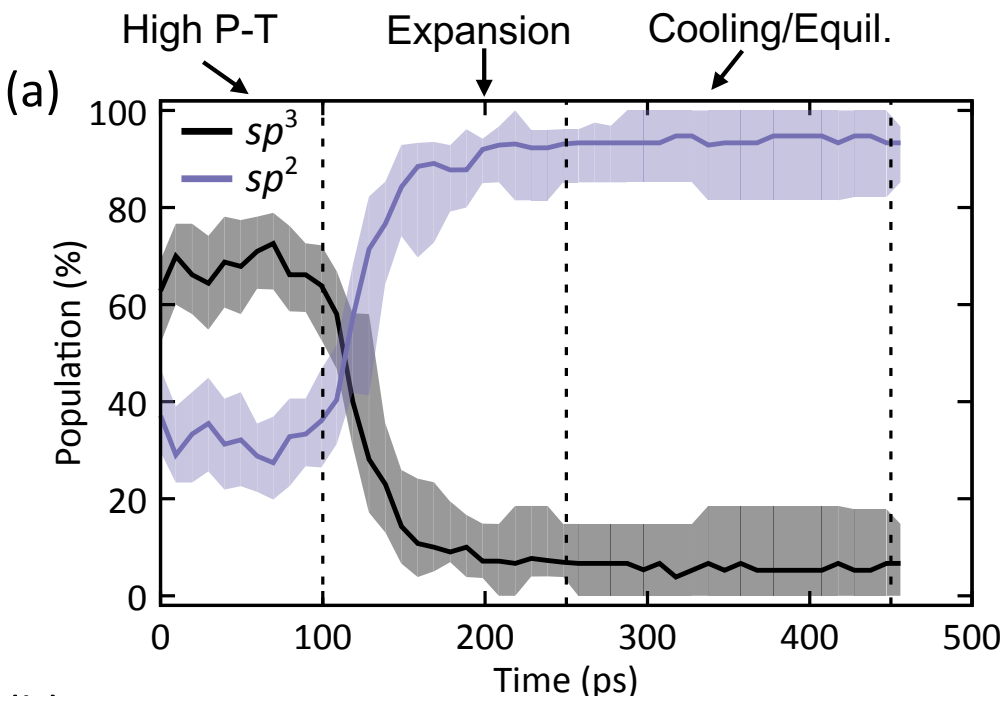
C-N-O bonds, only (expansion)

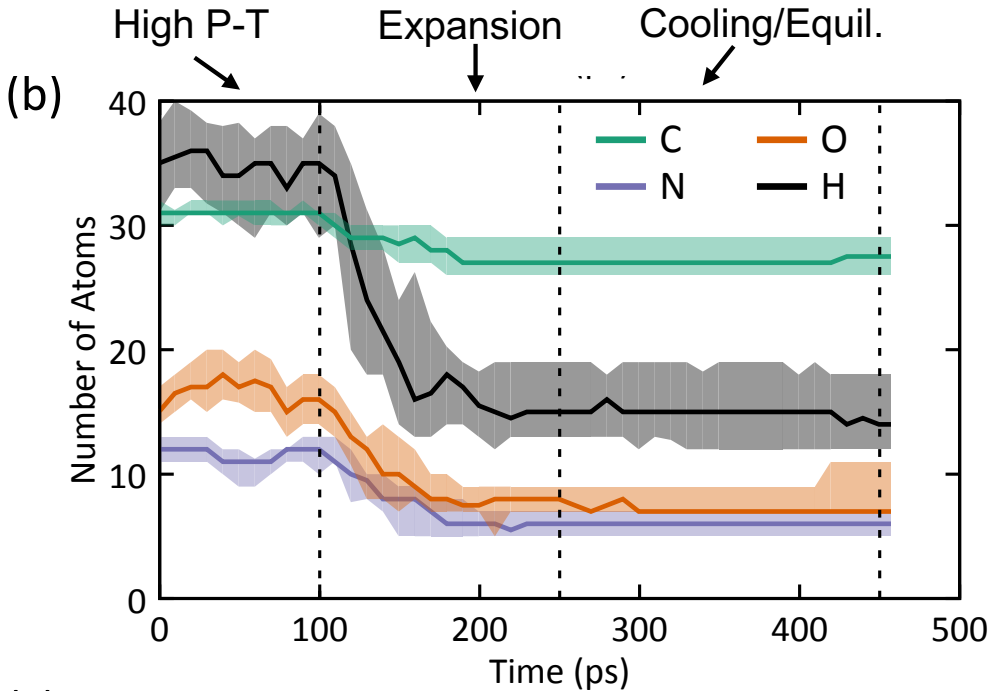






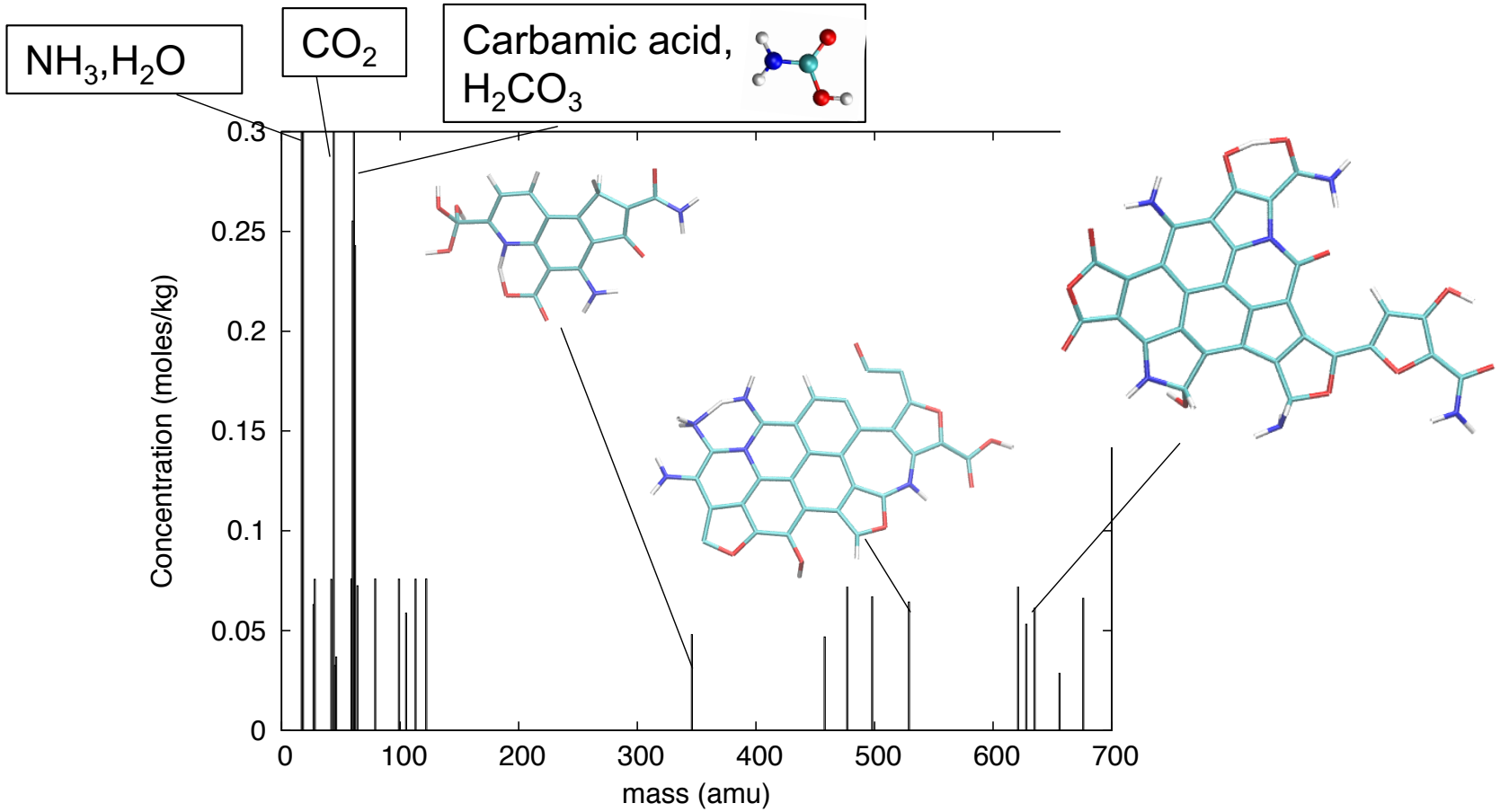
We have analyzed the changes in chemistry and structural morphology of the C-N bonded oligomers





Quenching of the chemistry appears to coincide with unfolding of the 3D C-N structures into planar sheets.

Cooling and equilibrating to ambient conditions (1 g/mL, 298 K) yields a wide variety of stable, highly functionalized graphite-like sheets

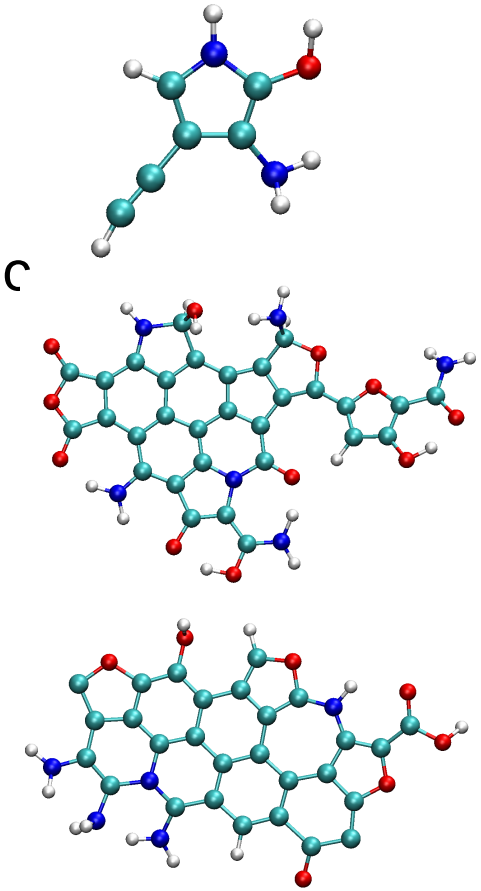


Aggregate results from 10 statistically independent simulations

Nitrogen substituted PAH's represent the missing link between nitrogen bearing acyclic molecules and prebiotic nucleobases along with vitamins found in meteorites

Aggregate Simulation Results

- All 10 simulations yield NPAHs, with one simulation also yielding a pyrrole cluster with a C-C triple bond.
- 8 simulations contain planar or nearly planar polycyclic sheets, whereas 2 sets had bridged-ring NPAHs that lead to complex 3D geometries.
  - NAPHs resemble both enzymes and vitamin related products.
  - 5 NPAHs contain an oxygen substituted five-membered ring (sugar chemistry).
- All simulations yield prebiotic small molecules, e.g., *carbamic acid* (NH_2COOH), *guanidine* (functional group in arginine and guanidine), *urea* (amino acid synthesis).



Conclusions and Future Work

- High P-T conditions likely played a role in the synthesis of chemically complex prebiotic compounds from simple precursors.
- We have developed a suite of atomistic simulation methods (quantum and classical) that can accurately model equilibrium chemical processes while retaining the accuracy of higher order methods.
- New research areas:
  - *Method Development:* Further DFTB method development, implementation of new Order-N eigensolvers, reactive classical force field development, *neural networks for reactivity.*
  - *Catalytic systems:* Carbohydrate chemistry in the presence of minerals.

Acknowledgments

- **LLNL:** Rebecca Lindsey, Matt Kroonblawd, Larry Fried
- **Funding:** NASA Exobiology program, LLNL's Lab Directed Research and Development program (LDRD)

Couple method comparisons

Method	Advantages	Disadvantages
Communicate through disk i/o: lammps initiates and executes DFTB+ every time step	Very easy to implement via lammps pair style.	Can be inefficient: wasted time with restarting code, i/o
Couple via LAMMPS library/callback function	Removes previous inefficiencies (~40% faster for small systems).	Writing a mixed language code can be tricky: *function names, arrays, argument calls and types*.
Couple via DFTB+ library	Likely removes previous inefficiencies (~40% faster for small systems)	Mixed code issues plus arguably more difficult than callback function approach

Example of Fortran wrapper functions

- Example callback function:
  - subroutine dftb_callback (lmp, timestep, nlocal, ids, c_pos, c_fext)
  - c_pos and c_fext are pointers to arrays in LAMMPS
- Pertinent wrapper functions/subroutines:
  - In main program/subroutine:
    - call lammps_open() -> instantiates LAMMPS
    - call lammps_file() -> read input file into LAMMPS
    - call lammps_set_callback() -> set pointer to callback function
    - call lammps_command() -> send specific commands to LAMMPS
  - In fortran callback function:
    - call lammps_extract_global() -> create pointers to relevant LAMMPS data, like box lattice vectors
    - call lammps_set_user_energy() -> set energy for fix external
    - call lammps_set_user_virial() -> set pressure virial for system