

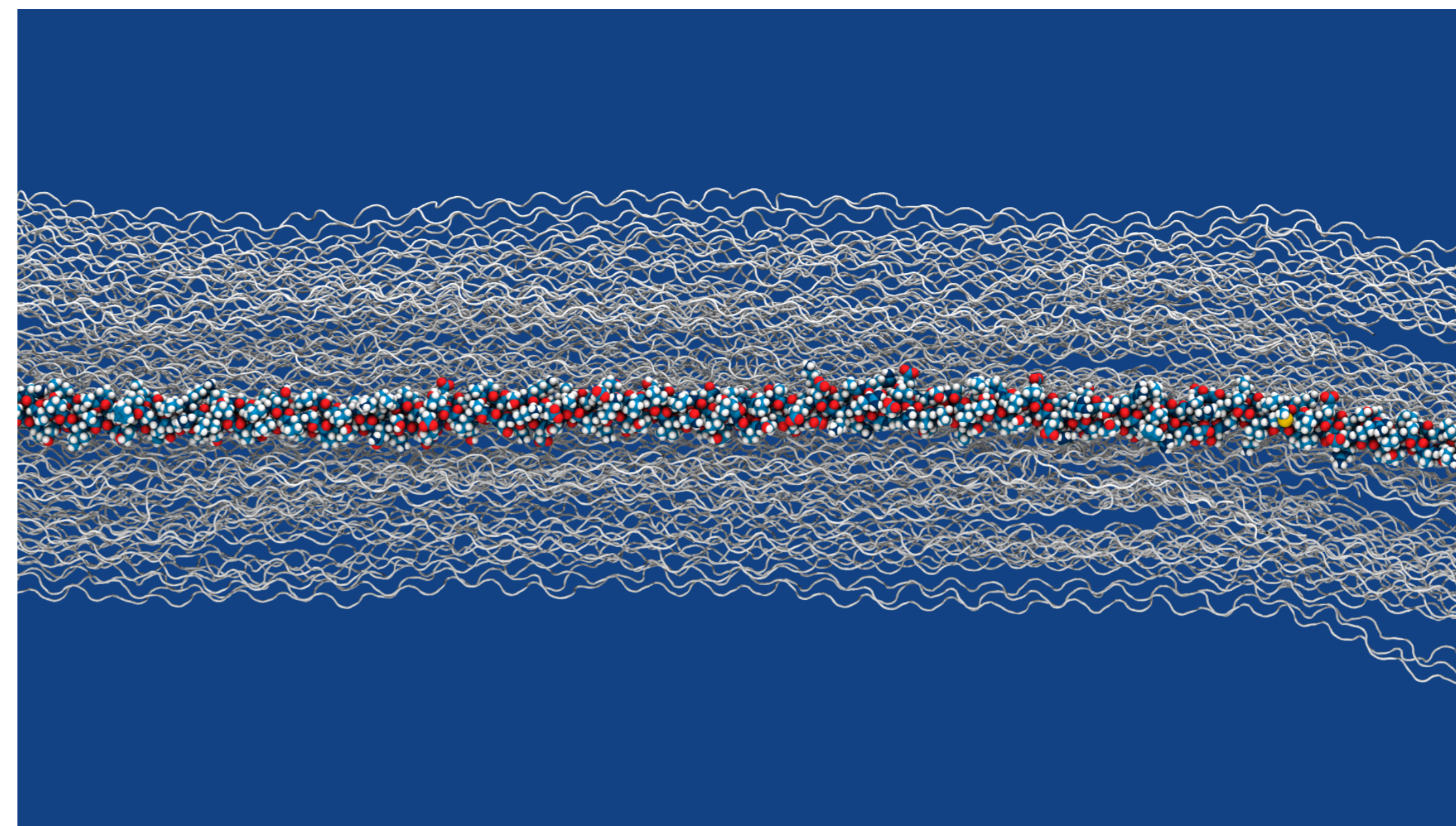
# KIMMDY 2.0

## Kinetic Monte Carlo Reactive Molecular Dynamics Framework

Jannik Buhr<sup>1,2</sup>, Eric Hartmann<sup>1,2</sup>, Benedikt Rennekamp<sup>1,2</sup>, Kai Riedmiller<sup>1,2</sup>, Frauke Gräter<sup>1,2</sup>

1. Heidelberg Institute for Theoretical Studies (HITS)
2. Interdisciplinary Center for Scientific Computing, Heidelberg University

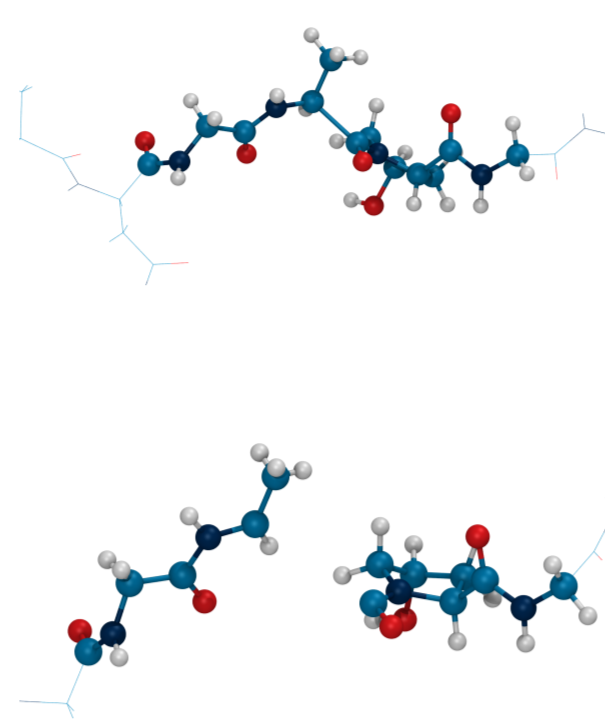
Underlined authors are at this conference and would love to talk to you in front of the poster.



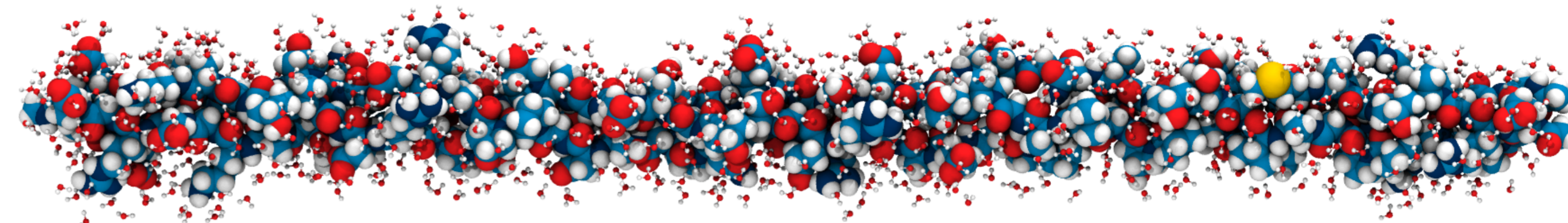
## Motivation and Biological Context

Chemical **reactions** occur on **many timescales**. And while those happening within femtoseconds can feasibly be simulated by quantum mechanical calculations, those that reach up to micro-, milli- or even full seconds can not. Forcefield-based molecular mechanics (MM) gets us into the microsecond timescale, but can not execute reactions. Even if it could, some reactions would still be inaccessible to our sampling due to their low rate.

KIMMDY **skips ahead** in time by querying possible reactions about their rates and choosing a reaction based on **kinetic Monte Carlo**. It then adjusts the topology and coordinates of the system according to the reaction results and continues the MD simulation. This allows us to **bridge timescales** and simulate trajectories that would otherwise be inaccessible.



Our specific use case is on **bond rupture** reactions in **collagen**, the most abundant mammalian protein. As the protein that provides structure and stability to our connective tissue, collagen is exposed to **mechanical stress**. The **mechano radicals** resulting from homolytic bond cleavage can propagate through the fibrillar structure through **Hydrogen Atom Transfer** (HAT).



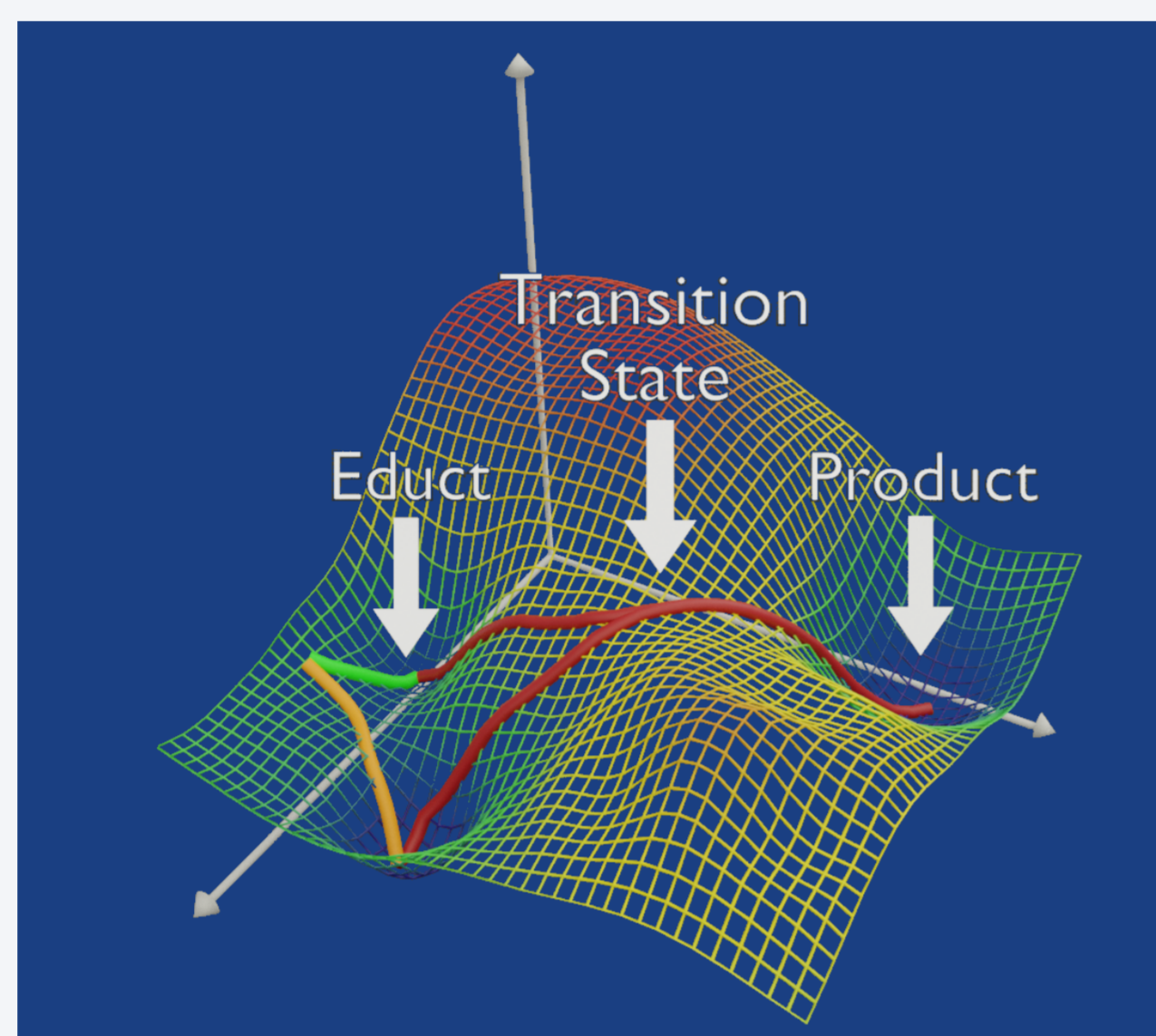
## Physics

### Initialize System

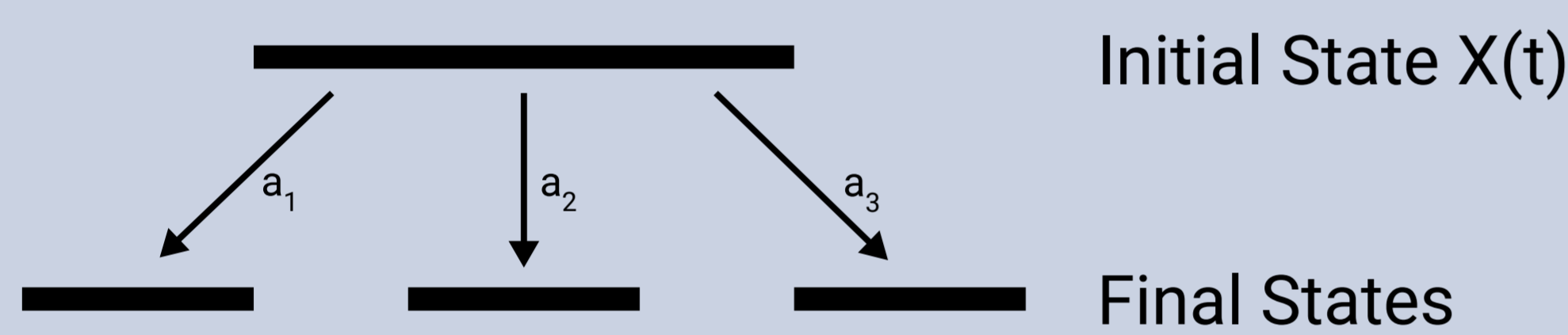
The **initial state** of the system is given by its **topology** and **coordinates** that are typically generated by equilibration with MM Molecular Dynamics (MD).

### Calculate Rates

Molecular dynamics simulations are used to **sample configurations** of our system. From these configurations, different events (reactions) are possible. According to transition state theory, the minimum energy path between two minima can be used to obtain the **energy barrier of an event  $\Delta E$**  (Andersen, Panosetti, and Reuter (2019)). Usage of the Arrhenius equation leads to the reaction rate constant  $k = Ae^{-\frac{\Delta E}{k_B T}}$  and thus the **reaction rate  $a$** . The fundamental premise of chemical kinetics (Gillespie (1977)) states that the probability that reaction  $k$  takes place in the time interval  $[t, t + \Delta t]$  is  $a_k(X(t))\Delta t + o(\Delta t)$ . Given **time-varying reaction rates**, an  $a_k$  can be found such that  $a_k \Delta t = \int_t^{t+\Delta t} a_k(X(t), s) ds$ .

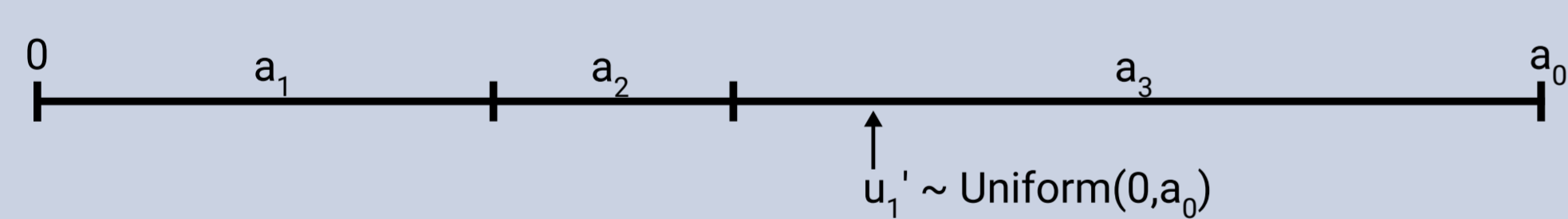


### Select Event



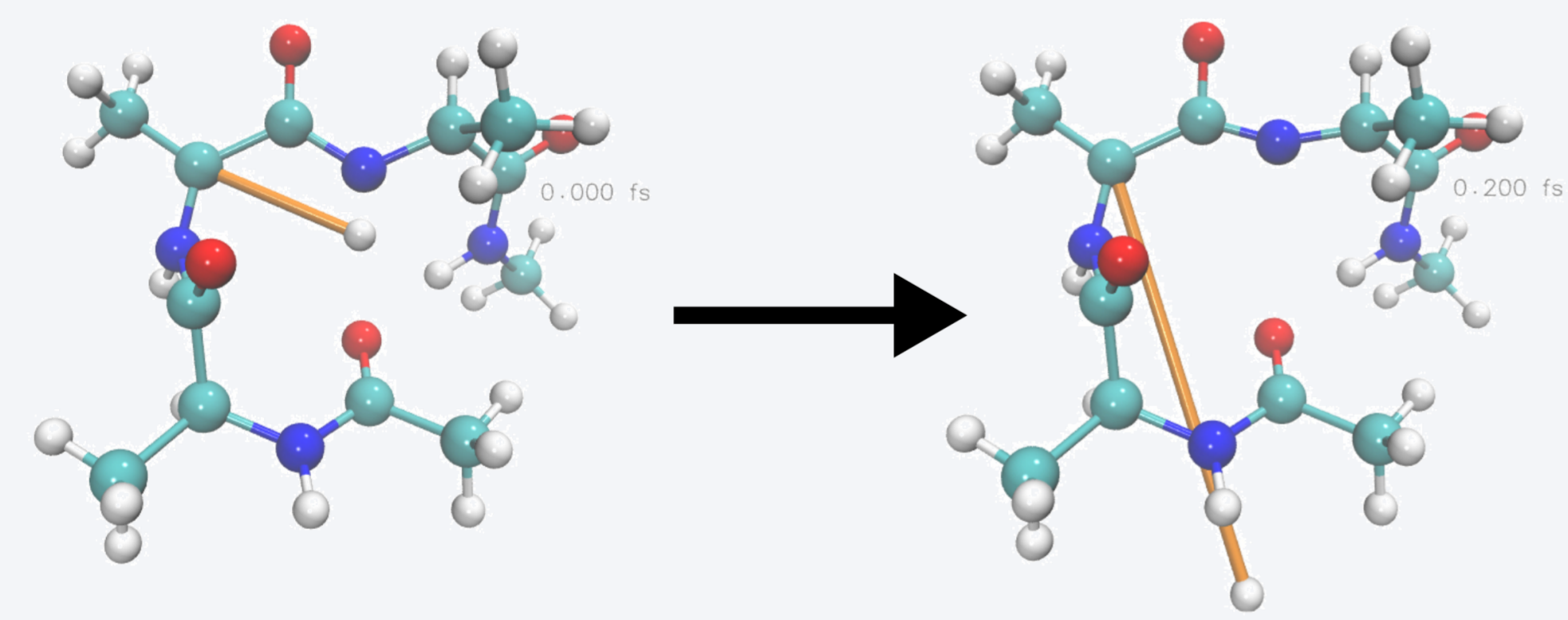
To describe the time evolution of a chemical system in the stochastic simulation formalism, it is crucial to know a) **which reaction** takes place next and b) **when** that is.

The classical rejection-free KMC algorithm uses the sum of rates  $a_0 = \sum_{k=1}^M a_k$  and **uniform random numbers**  $u_1, u_2$  for both tasks. Using  $u_1$ , a reaction  $\mu \in [1, \dots, M]$  can be found such that  $\sum_{k=1}^{\mu-1} a_k < u_1 a_0 \leq \sum_{k=1}^{\mu} a_k$ , which is equivalent to choosing from reactions  $[1, \dots, M]$  with the  **$k$ th reaction having probability  $a_k/a_0$** . The time step  $\Delta t = 1/a_0 \ln(1/u_2)$  is equivalent to drawing an exponential random variable with parameter  $a_0$ .



### Execute Event

Once an event has been selected, the **topology** and **coordinates** of the system need to be updated. This includes placing atoms at their final coordinates and adjusting the velocities of a hydrogen that has just undergone a HAT to match its new heavy atom binding partner. Before advancing the production simulation, a **relaxation** is performed. This can include localized changes in the case of HAT or global movements for more disruptive events such as bond rupture.



## Implementation

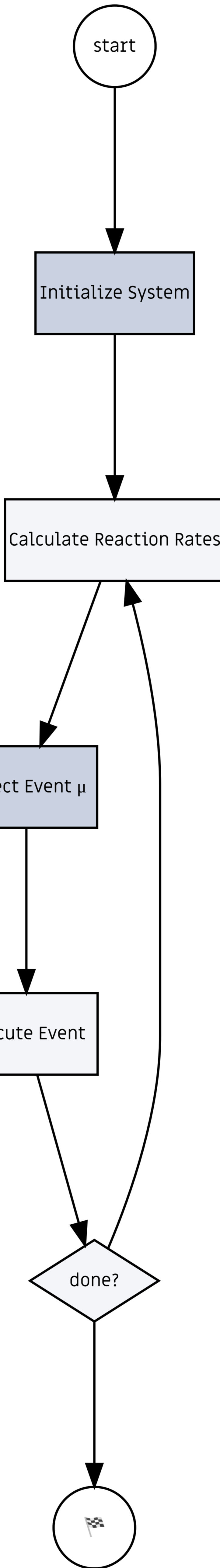
### Initialize System

KIMMDY is configured via a **yaml file** whose options are merged with the defaults and started via the `kimmdy` command. In here, a user may specify the **MD step** to be performed and the **reactions** to be evaluated. The defined MD simulations and reactions are used to create a sequence of tasks that fill the **Task queue**.

From the yaml file, a config class is initialized which can be requested by our modules. For example, information required by GROMACS for the MD simulation, like the location of the force field or mdp files to use, or files needed by a reaction are also specified here.

All **files are tracked** when they get created by a KIMMDY run. The most recent file of a given type can easily be requested to ensure the right files are chosen for consecutive steps.

```
kimmdy:
  daemon: false
  mpi: false
  ftl: none
  forcefield: gromacs-100k.ff
  top: topp.tpo
  gmx: gmx
  id: backbone.idx
  plumed:
    dist: plumed.dat
  distances: distances.dat
  equilibration:
    app: eq.mdp
  production:
    app: prod-100k.mdp
  reactions:
    none
  eddy: eddyoc.dat
  bonds: "bonded.txt"
  dummyreaction: default
  md:
    model: default
  sequence:
    equilibration
    md: 2
    relax
    production
    reactions
```



### Calculate Rates

Depending on your system of interest, different events (reactions) have to be considered in the KMC step. Thus, KIMMDY is designed to be **extensible** with new reactions via a simple python plugin architecture. All a new reaction plugin needs, is a way of returning rates for possible reactions and the corresponding changes that need to be performed on the topology on acceptance. The **interface** is formalized such that reaction plugins have to contain a subclass of our Reaction abstract base class. A class method returns rate calculations in the form of a ReactionResult object.

Current reaction plugin examples are the bond rupture reaction (**homolysis**), which estimates transition state energies with a simple bond stretching potential. Other reactions require more sophistication. Once a bond is ruptured, the remaining radicals can "jump" via **Hydrogen Atom Transfer** (HAT), for which you want to head over to the poster of Kai Riedmiller to see his **Machine Learning** solution for rate prediction.

### Select Event

To select a single reaction out of all possible ones, the ReactionResults are flattened and the cumulated rates calculated. A random number is used to determine the chosen rate based on the selection criterion. The recipe that belongs to the chosen rate is passed on to the next step.

### Execute Event

A reaction returns a **recipe** of changes that are executed by KIMMDY. KIMMDY comes with modules to perform arbitrary modifications of the topology. A recipe is any composition of the **two basic operations**: breaking a bond or forming a new bond.

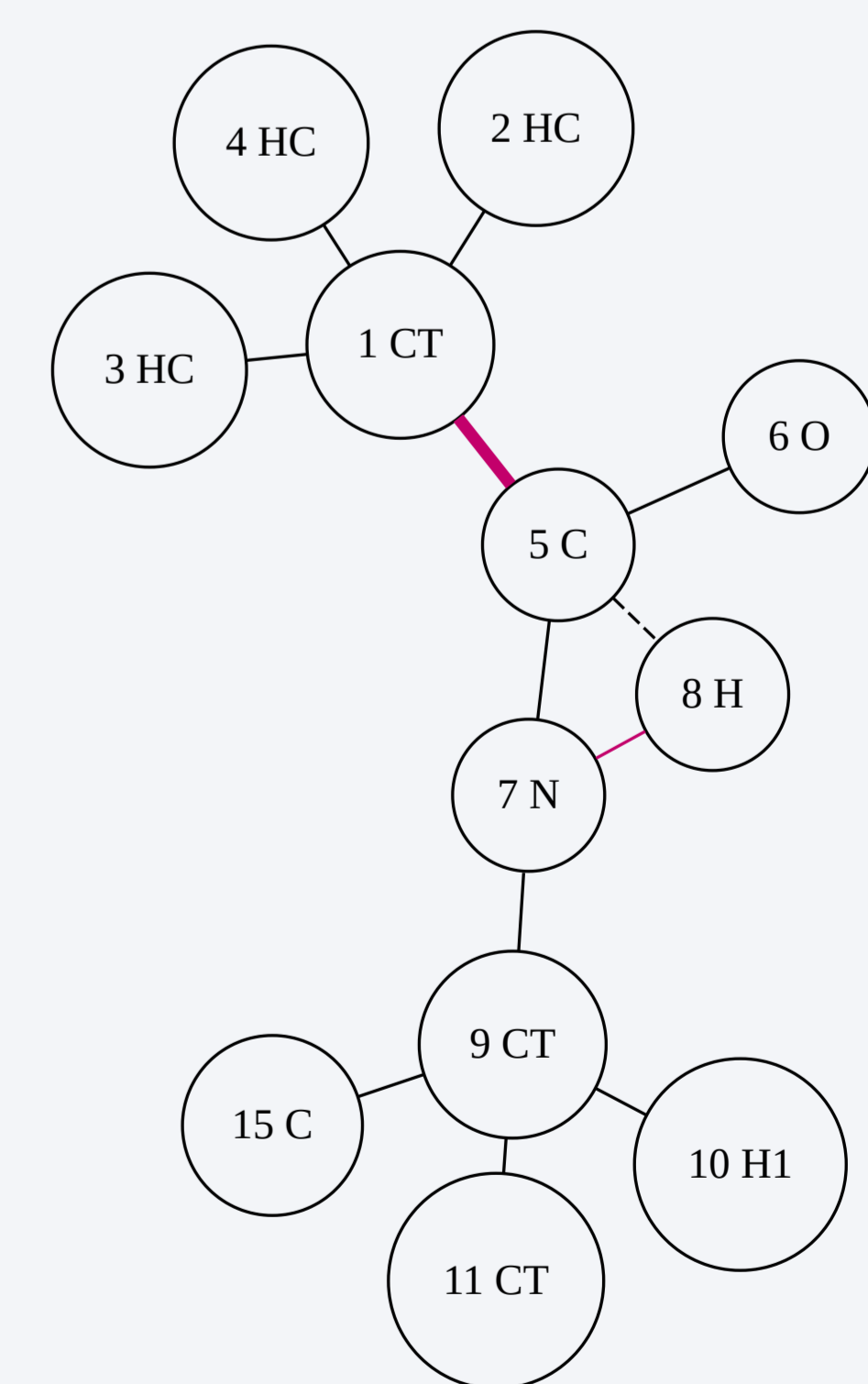
For example, to homolytically break the bond indicated in bold magenta, followed by a HAT of the nearby hydrogen 8 from 7 N to 5 C (dashed line), we perform the following operations:

```
(top.break_bond((1,5))
 .break_bond((7,8))
 .bind_bond((5,8))
)
```

KIMMDY performs the necessary changes to **all parts of the topology** such as bonds, angles and (im-)proper dihedrals.

Introducing radicals into the system leads to configurations for which no parameters exist in most force fields. We provide a way of **automatically adding or modifying parameters** in places where radicals are created. This is done by explicit patching rules or on the fly parametrisation through a specifically trained espaloma (Wang et al. 2022) model.

The ideal **coordinate relaxation scheme** depends on the reaction. Currently, slow growth simulations and an atom placement algorithm are implemented.



## References

- Abraham, Mark James, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. 2015. "GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers." *SoftwareX* 1–2 (September): 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
- Andersen, Mie, Chiara Panosetti, and Karsten Reuter. 2019. "A Practical Guide to Surface Kinetic Monte Carlo Simulations." *Frontiers in Chemistry* 7. <https://www.frontiersin.org/articles/10.3389/fchem.2019.00202>.
- Gillespie, Daniel T. 1977. "Exact Stochastic Simulation of Coupled Chemical Reactions." *J. Phys. Chem.* 81 (25): 2340–61. <https://doi.org/10.1021/j100540a008>.
- Rennekamp, Benedikt, Fabian Kutzki, Agnieszka Obarska-Kosinska, Christopher Zapp, and Frauke Gräter. 2020. "Hybrid Kinetic Monte Carlo/Molecular Dynamics Simulations of Bond Scissions in Proteins." *Journal of Chemical Theory and Computation* 16 (1): 553–63. <https://doi.org/10.1021/acs.jctc.9b00786>.
- Wang, Yuanqing, Josh Fass, Benjamin Kaminow, John E. Herr, Dominic Rufa, Ivy Zhang, Iván Pulido, Mike Henry, and John D. Chodera. 2022. "End-to-End Differentiable Molecular Mechanics Force Field Construction." *Chemical Science* 13 (41): 12016–33. <https://doi.org/10.1039/D2SC02739A>.



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386



This work was supported by  
the Klaus Tschira  
Foundation.

This project has received  
funding from the European  
Research Council (ERC)  
under the European Union's  
Horizon 2020 research and  
innovation program (grant  
agreement No. 101002812).