

Scientific report for the project `Computational design of self-assembling supramolecular motors and quasicrystals` (*MOTOQUASI*)

Summary of project activities and deliverables

Activity	Deliverable
Exploitation/dissemination	2 publications in international peer reviewed journals
1.1. Adapting the global optimization algorithm for quasi-bidimensional systems	Rigid body model adapted for GMIN and OPTIM
1.2. Adapting the model rigid body potentials for GPUs in OPTIM	
1.3. Implementing the model potentials into a molecular dynamics package	
2.1. Selecting cooperative single transition state rearrangements	Model of self-assembled colloidal motors implemented in computational software
2.2. Finding the most feasible colloidal motor design	
2.3. Adapting the model for explicit solvent	
2.4. Exploring different energy transfer methods to drive the colloidal motors	
3.1. Optimization and parameter refinement of the minimal quasicrystal model	Model of self-assembled quasicrystals implemented in computational software
3.2. Dynamics of interconversion between crystalline and quasicrystalline phases	
3.3. Exploring the growth of quasicrystals with MD or MC	

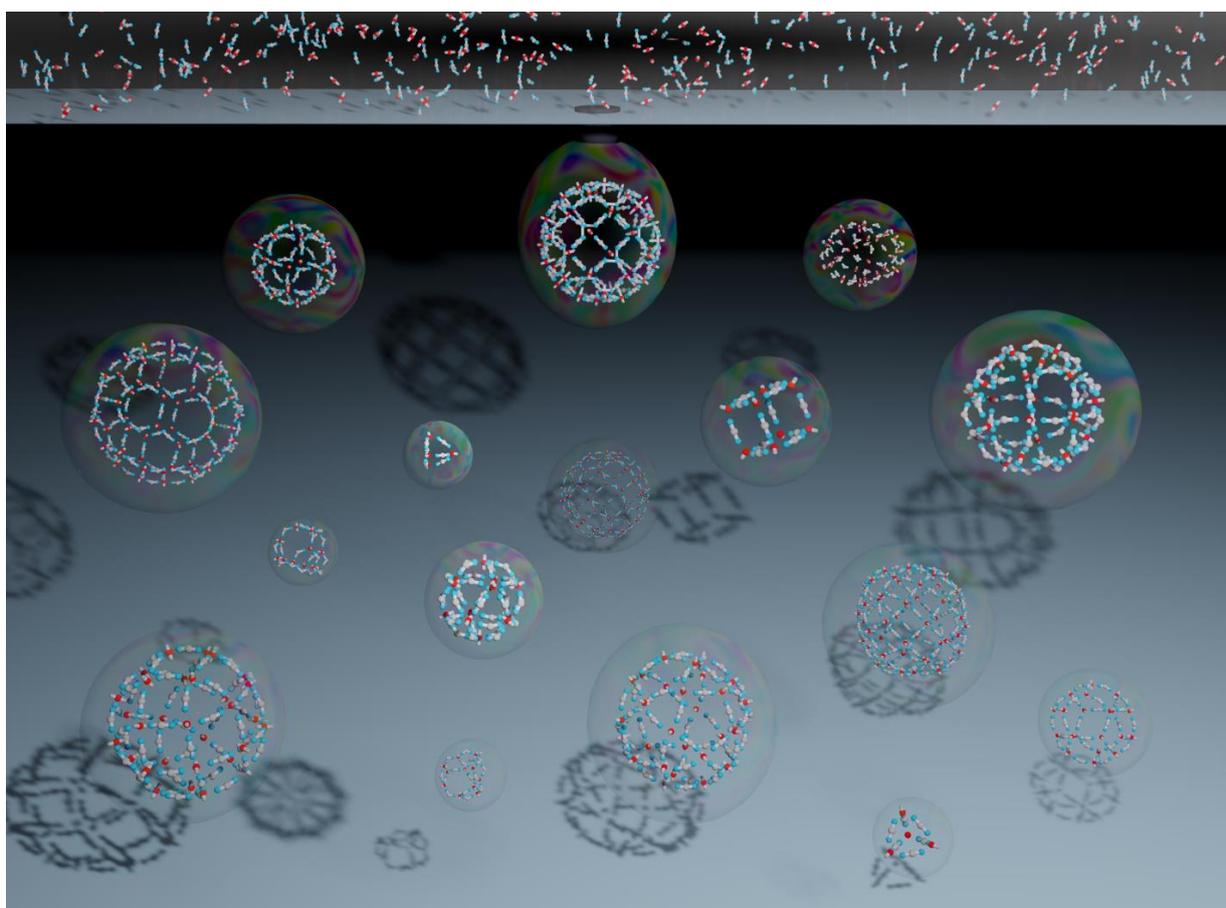
All project goals have been successfully met and the deliverables have been completed.

Project website: <https://szilard.ro/motoquasi>

Exploitation/Dissemination results:

Journal cover – Nanoscale Advances

Our artwork below has been selected as the inside front cover of Nanoscale Advances in their upcoming (September or October 2022) issue:



Publications of the group in international Q1 or Q2 journals during the project implementation period (publications with acknowledgement of the current project are highlighted):

1. M Baibarac, M Daescu, SN Fejer, *Coatings* 11 (2), 258 (2021). Optical evidence for the assembly of sensors based on reduced graphene oxide and polydiphenylamine for the detection of epidermal growth factor receptor.
2. M Oprica, M Iota, M Daescu, SN Fejer, C Negrila, M Baibarac, *Scientific reports* 11 (1), 1-12 (2021). Spectroscopic studies on photodegradation of atorvastatin calcium

3. G Hudák, G Farkas, B Vajik, B Sinka, K Rákosi, O Csákány, L Terza, Z Jenei, SN Fejer, *Clinica Chimica Acta* 523, 169-171 (2021). Extremely rare “daisy-like” crystals in urinary sediment can be due to a sampling artifact
4. J Szoverfi, SN Fejer, *Scientific reports* 12 (1), 1-11 (2022). Dynamic stability of salt stable cowpea chlorotic mottle virus capsid protein dimers and pentamers of dimers
5. M Korodi, I Horváth, K Rákosi, Z Jenei, G Hudák, M Kákes, K Dallos-Fejér, E Simai, O Páll, N Staver, V Briciu, M Lupșe, M Flonta, A Almaș, V Birlutiu, CD Lupu, AM Ghibu, D Pianoschi, L Terza, SN Fejer, *Vaccine* 40 (37), 5445-5451 (2022). Longitudinal determination of BNT162b2 vaccine induced strongly binding SARS-CoV-2 IgG antibodies in a cohort of Romanian healthcare workers
6. I Horvath, DJ Wales, SN Fejer, *Nanoscale Advances* (2022, Advance article). Design of self-assembling mesoscopic Goldberg polyhedra

Conference poster – Barcelona MMSML Workshop – Methods in Molecular Simulations and Machine Learning (14-16 July 2022)

I Horvath, SN Fejer. Using neural networks for geometry optimization of simple molecular structures

Phase 1 activities

Expansion of computing capabilities of the group

We originally planned to acquire graphics cards to expand the GPU computing capabilities of our compute nodes. However, the appropriate graphics cards are in short supply and we could not buy standalone GPUs that would fit our need. We therefore decided to get two high-end desktops which contain nVidia RTX 2060 GPUs, and add them to our computing nodes in a mesh configuration, with computing tasks distributed by a central node. We also acquired DDR4 memory modules to increase the available memory of the computing nodes so that larger systems can be simulated.

Setting up the software framework for efficient exploration of potential energy surfaces in model systems

The scripts needed throughout the course of this research project will mainly be written in the Python programming language. Python is a high level programming language, for which multiple different packages are available. The main packages which we will use include:

- **HOOMD-Blue**, a general-purpose particle simulation toolkit optimized for execution on both GPUs and CPUs. Website: <http://glotzerlab.engin.umich.edu/hoomd-blue/>

- **PELE** (Python Energy Landscape Explorer), a package that contains tools for global optimization and energy landscape exploration. Github: <https://github.com/pele-python/pele>
- **TensorFlow**, a platform for developing and testing machine learning models. Website: <https://www.tensorflow.org/>

The environment was created and deployed with the Docker platform (<https://www.docker.com/>). Docker provides the ability to package and run an application in a loosely isolated environment called a container and thus eases the process of deploying and maintaining computational environments. Once the images of these containers are created or downloaded, they can be moved to any other machine, which has docker installed, without worrying about the applications dependencies.

Docker was installed on a local machine to develop and test any images used and on the server to carry out the calculations. Alongside docker the Nvidia run time was installed, so that we can utilize the GPUs of each system. The three main Docker containers, which we will use:

1. The molecular dynamics environment:

This mainly consists of the HOOMD Blue library mentioned earlier, and some custom scripts which will be used to automatize the workflow of guessing and perfecting the simulation parameters, also a basic database which will record all the data, written in Pandas. This container can be accessed through the built in Jupyter notebook, which is an interactive Python scripting environment.

2. The structure optimization environment:

This Docker image is built on the Python 2.7 image, with PELE and some other Python packages installed, like NumPy and SciPy for performing the required calculations and Pandas to manage data. The main role of this container is to optimize simple molecular structures, to produce input for the machine learning environment and to study the energy landscapes of molecules.

3. The machine learning environment:

This container is based on the official Tensorflow-Jupyter image, with some extra packages to support the creation of 3D images about the molecular structures. The purpose of this container is to create a neural network, which takes as input a randomly generated molecular structure and predicts a structure, which on the energy landscape is located closer to the global minima of the respective structure.

Testing of the software framework

MD simulations with different LJ potentials

A system of 125 particles at equal distances in a simple cubic lattice has been set up for molecular dynamics (MD) simulations (Figure 1). Three MD simulations were run using different LJ potentials: the value of the potential well depth (epsilon) was increased from 0.1 through 1.0 to 10.0 while keeping the other components of the potential constant.

The differences in the LJ potentials clearly affect the way how the particles interact with each other (Figure 2). The physical parameters of the systems are also clearly different (Figure 3).

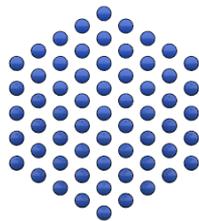


Figure 1. Starting position of the MD simulations.

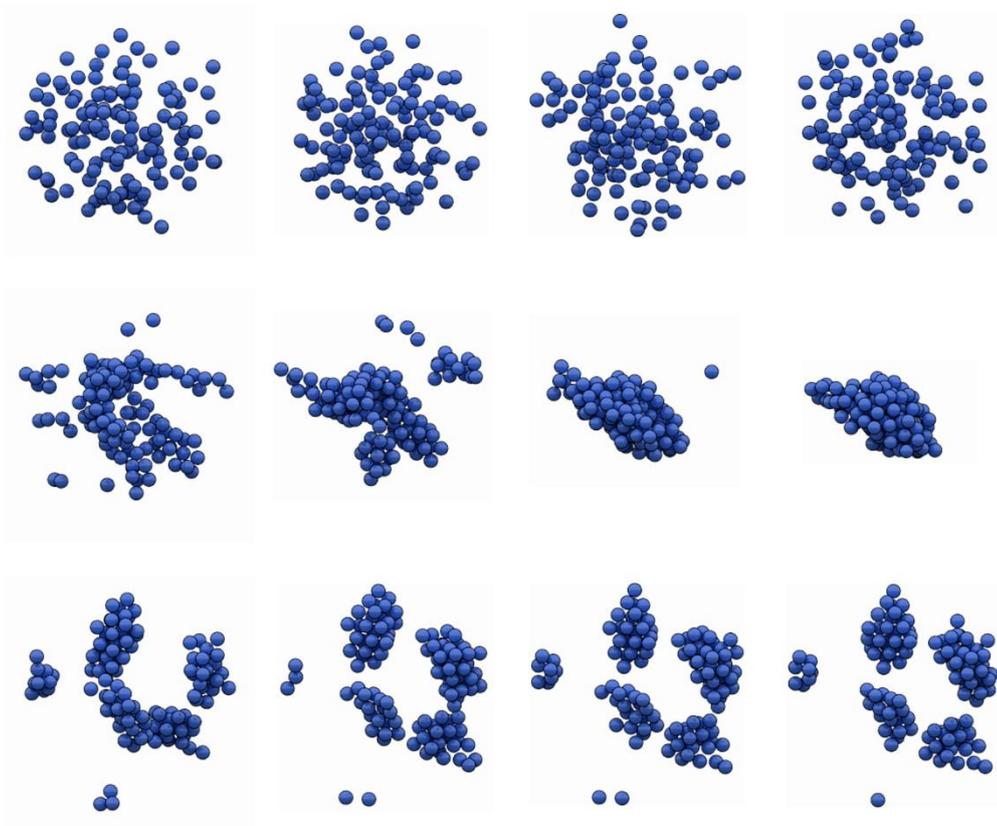


Figure 2. MD snapshots of the three different simulations taken at equal time intervals.

Top: epsilon = 0.1. *Center:* epsilon = 1.0. *Bottom:* epsilon = 10.0.

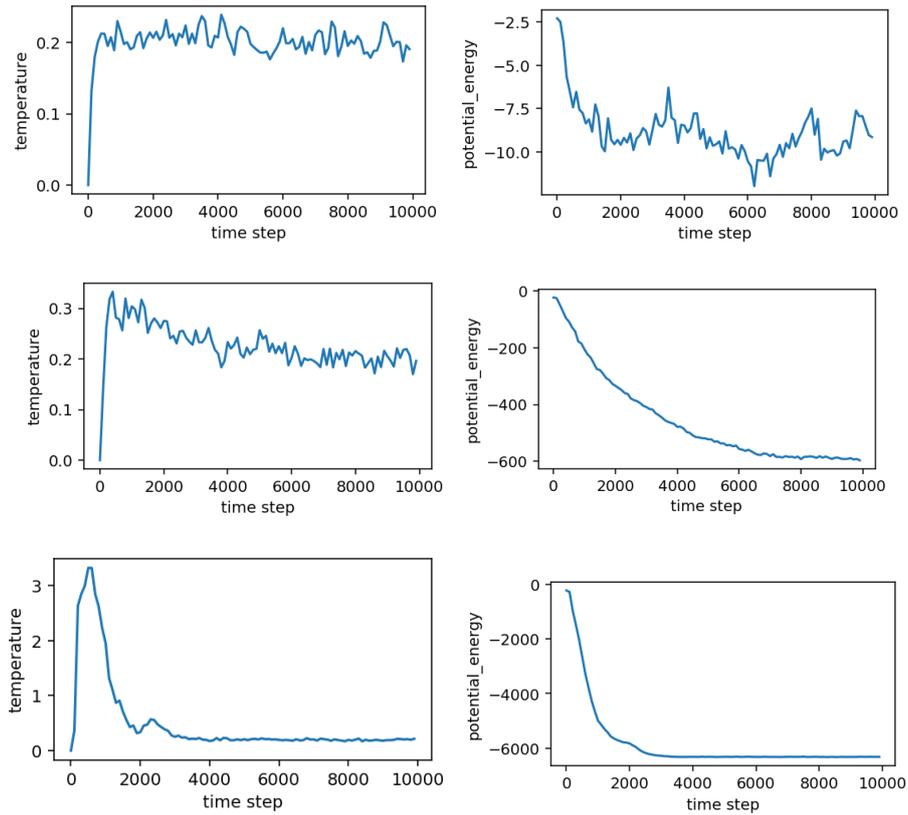


Figure 3. Evolution of the temperature and potential energy during the three MD simulations.

Top: epsilon = 0.1. *Center:* epsilon = 1.0. *Bottom:* epsilon = 10.0.

MD simulations with two types of particles

64 particles of two different type have been set up in a simple cubic lattice for MD simulations. The interaction strength between the different particle types can be controlled by changing the pair potential well depth (epsilon) for each particle type pair. In the first simulation, the interaction between different types of particles was forced (Figure 4, left). In the second simulation, the interaction between similar types of particles was preferred (Figure 4, right).

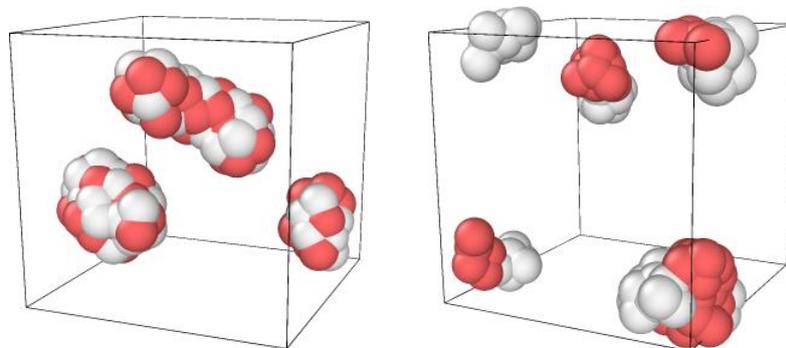


Figure 4. Final snapshots of the MD simulations using two types of particles (shown in different colours). *Left:* interaction between different particles was forced. *Right:* interaction between similar particles was favoured.

Figure 5 shows example results of a simulation on rigid body particles that has been set up with the developed Docker images. The script takes as input only a few parameters (definition of rigid body types and numbers, cell size) and then the system setup, interaction matrix and starting configuration are determined automatically. After that step, input files are generated to be run on GPUs on our computing cluster. Trajectory files can be collected from the compute nodes by a script and visualized by the user.

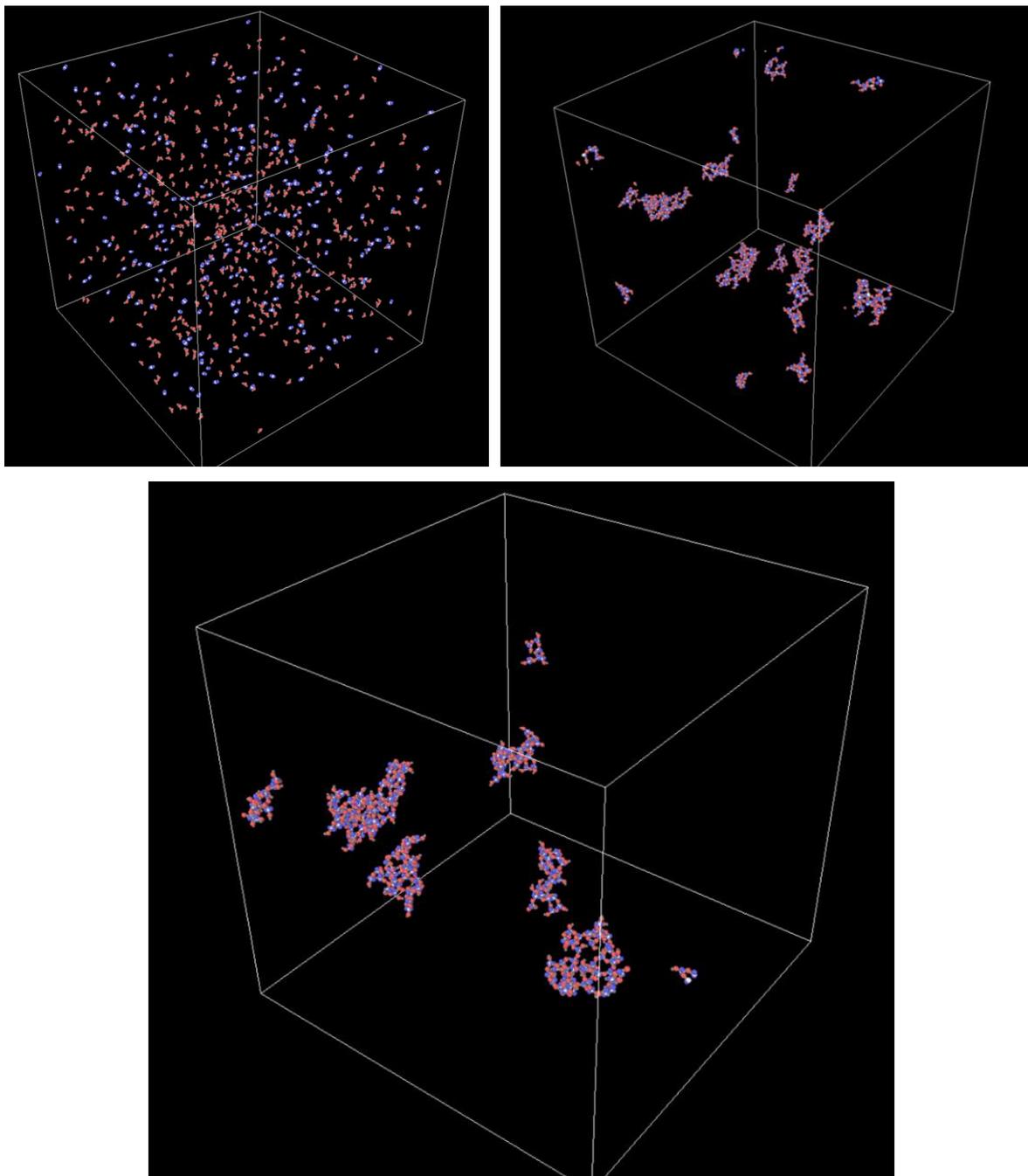


Figure 5. Snapshots along the trajectory of a two-component rigid body system using the automated script generation platform.

Phase 2 activities

Selecting single-transition state rearrangements

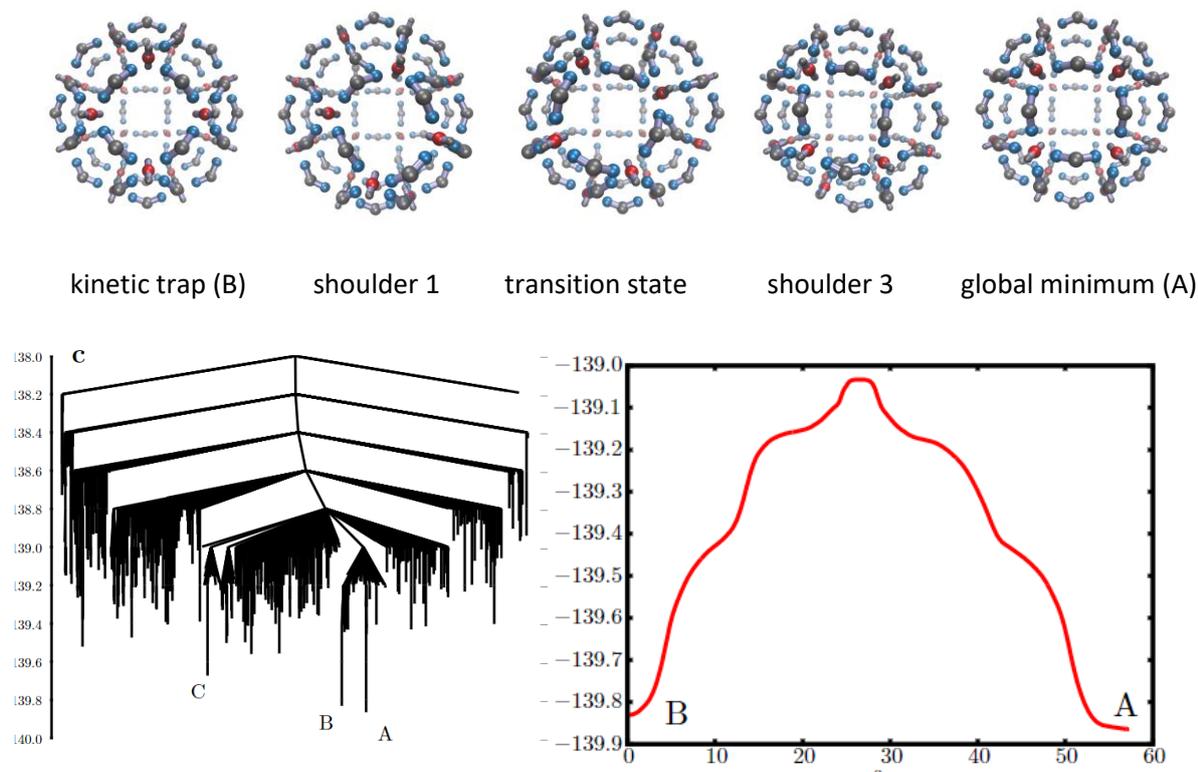


Figure 6. Top panel: snapshots of the single transition state rotatory motion involving two shoulders on each side for the fastest transition pathway between two lowest-energy structures on the landscape. Bottom left panel: extensively sampled low-energy region of the landscape for a ligand bend angle of about 136° . Bottom right panel: the energy profile of the transition.

Single transition state rearrangements that comprise of highly cooperative moves can in theory be exploited as colloidal motors. We therefore explored our previous simulations to see if such rearrangements are a common occurrence or not when the rigid building blocks feature only Coulomb and excluded volume interactions. Using our model of self-assembling Goldberg polyhedra, we found that these systems are capable of cooperative moves of long path lengths, involving the motion of a large proportion of the cluster particles themselves, but the low-energy single transition state rearrangements are mostly deformations of bistable hollow shells. However, we found an ideal range of bend angles in M_nL_{2n} model systems where the single transition state rearrangement having a rotatory motion described in page 7 of the project proposal can occur and is the lowest energy pathway: approx. $120\text{-}140^\circ$ for the pseudorhombicuboctahedral-rhombicuboctahedral transition.

The rearrangement shown in Figure 6 can be seen at <https://youtu.be/redvzHn34IQ>. We attempted to create conditions which are appropriate for spontaneous observation of the same single transition

state rearrangement with molecular dynamics using the HOOMD Blue MD environment set up in phase 1. However, for this particular example, only motions leading up to the first shoulder were observed, and on higher temperature the vibrations of the system made observing such rotatory motions extremely difficult. In fact, once the system reached the kinetic trap (which seems to be the entropically most favoured state to reach starting from random building block configurations), we did not observe transition towards the global minimum over the selected range of temperatures. On this landscape, the event is too rare to be observed with the simulation lengths accessible to us. We therefore modified the parameters to lower the barrier for this rearrangement, but that destabilized the structure. Further simulations are ongoing.

Finding the basic design of a self-assembling colloidal motor

As the parameter space of the building blocks to be explored is huge, we had to design a somewhat automated framework that can autonomously navigate this parameter space by selecting the fittest set of parameters for a particular task. In this way, a basic rigid body model designed with common sense can be further enhanced to further stabilize the desired structure or enhance its rigidity or self-assembling properties. In this framework molecular dynamics simulations are created automatically by a program from a starting input parameter set and run on the HOOMD Blue MD environment of the group.

Basic notions that will be used in the next sections:

- particle = a basic spherical element with: charge (q), mass (m) and radius (r)
- rigid body = a structure formed from multiple particles, which have fixed position relative to each other
- simulation = the observation of the interaction between multiple rigid bodies confined to a simulation box
- mutation = a random change in any of the characteristics of the studied element
- mutation distance = a measure to somehow quantify the difference between two elements of the same type
- fitness = a measure of how well the simulation result adheres to the desired configuration

All structure data is stored in a database, the simulations are saved in gsd format.

Mutations

The program is designed to create new simulations from existing ones by mutating them. There are multiple mutations which can occur for a simulation, these are:

- increasing the number of a specific rigid body in the simulation
- decreasing the number of a specific rigid body in the simulation
- adding a new type of rigid body to the simulation
- removing a rigid body from the simulation
- mutation of a rigid body type

The possible mutations of a rigid body are:

- adding a new particle
- removing a new particle
- swapping a particle for another one registered in the database
- changing the position of one of the particles

Each of the listed mutations can occur with a given probability, which is defined by the user beforehand. The user can also define a maximum mutation distance, which limits the change of the simulations between generations.

The process is started from an initial simulation (S_0), which will have a fitness, F_0 . A new simulation is created by mutating S_0 with a probability of P_M , which can be defined as:

$$P_M(i) \propto \frac{1}{F_i}$$

or

$$P_M(i) = 1 - F_i$$

depending on the range of the fitness function values. If there is no mutation then S_0 will be continued from where it was left off. In further iterations a simulation S_i is started, with probability:

$$P_C(i) \propto F_i$$

Once again the selected simulation is either mutated into a new one or continued based on its fitness value. The process is stopped when the number of desired simulations is reached.

Basic density fitness function

To test the procedure, we chose to implement a simple fitness function based on the density of a subsection of the simulation box. In this approach a fractional box size (α) is chosen (we use $\alpha=0.3$) and densities are calculated for all the possible fits:

$$\rho_{ijk}(\alpha) = \frac{n_{ijk}}{V(\alpha)}$$

where $V(\alpha)$ is the volume of the cell and n_{ijk} is the number of particles in the cell. The fitness value is given by comparing the maximal $\rho_{ijk}(\alpha)$ value to the theoretical maximum density

$$D(\alpha) = \frac{N}{V(\alpha)}$$

Where N is the total number of particles in the simulation. The fitness function will be computed as

$$F = \frac{\max(\rho_{ijk}(\alpha))}{D(\alpha)} = \frac{n_{ijk}}{N}$$

Technical details

The program is implemented with the use of Docker containers. There are 3 types of containers:

- *glotzerlab*, these containers run the molecular dynamical simulations
- *head*, this container is responsible for creating new simulations, handing out simulations to the *glotzerlab* containers and writing to the database
- *database*, this container holds data about the structures

The containers communicate between each other on the default docker networks and Python sockets.

Exploring the parameter space of rigid building blocks

To see if we can get further from our original concept of binary charged systems, we considered several rigid body configurations and ran MD simulations to explore their behaviour. Some examples are illustrated below, with their respective MD trajectories:

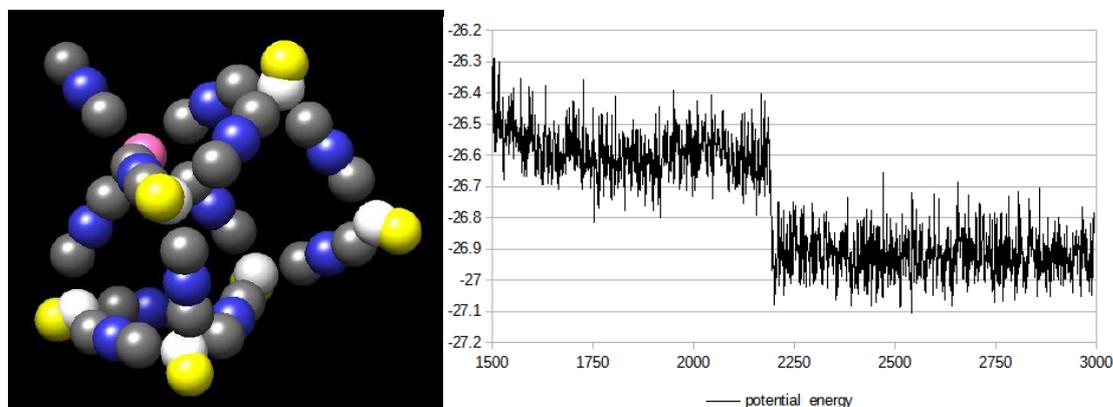


Figure 7. Assembly intermediate of a system preferring face-sharing polytetrahedral arrangement

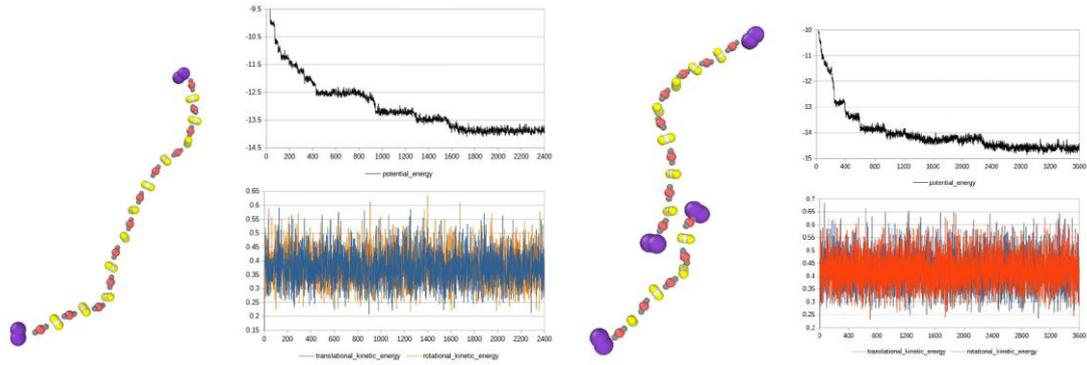


Figure 8. Capped strings which tend to form dimers if the number of caps is increased

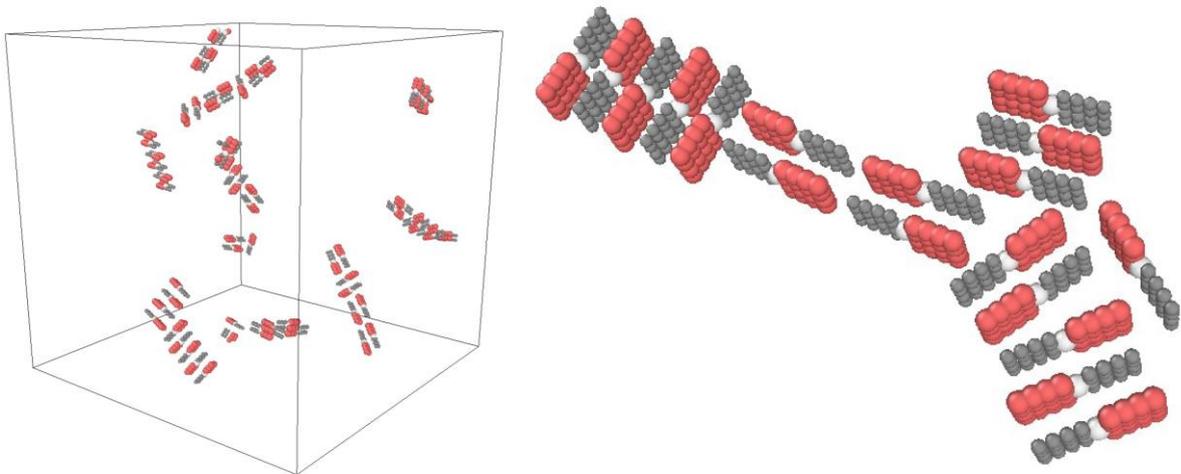


Figure 9. Lennard-Jones interactions with anisotropic plate-like excluded volume areas, having no charges in the system

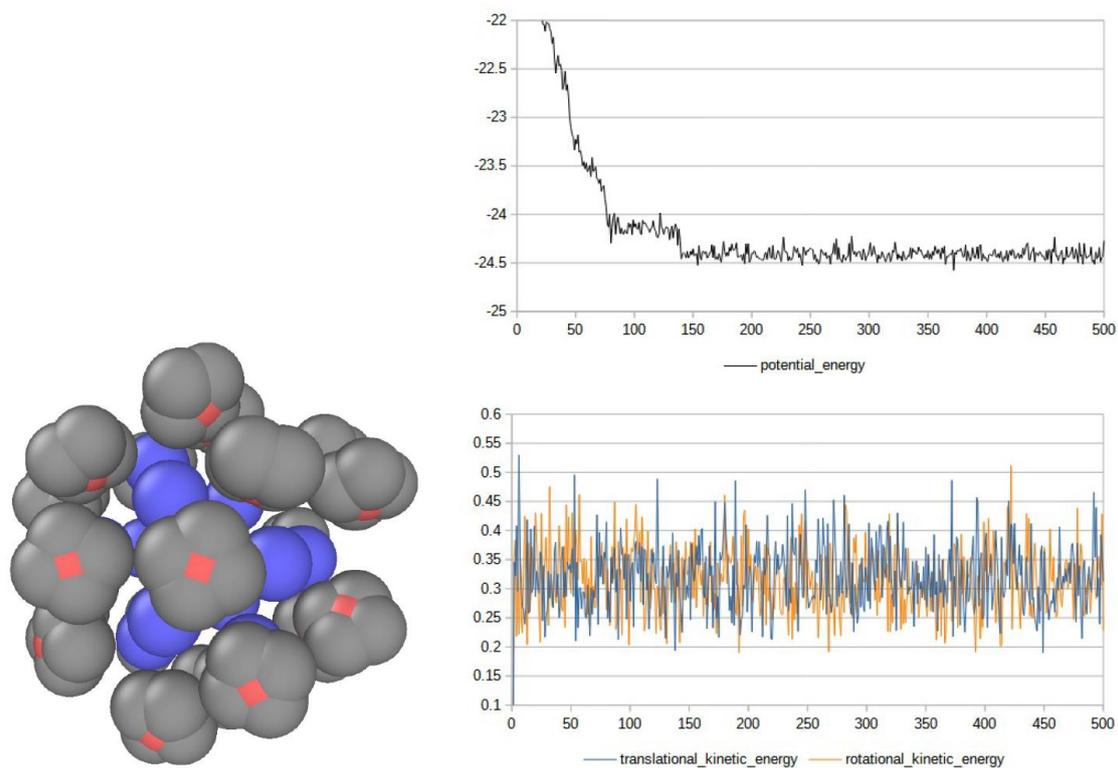


Figure 10. Icosahedral charged particle (+20) with 20 planar counterions

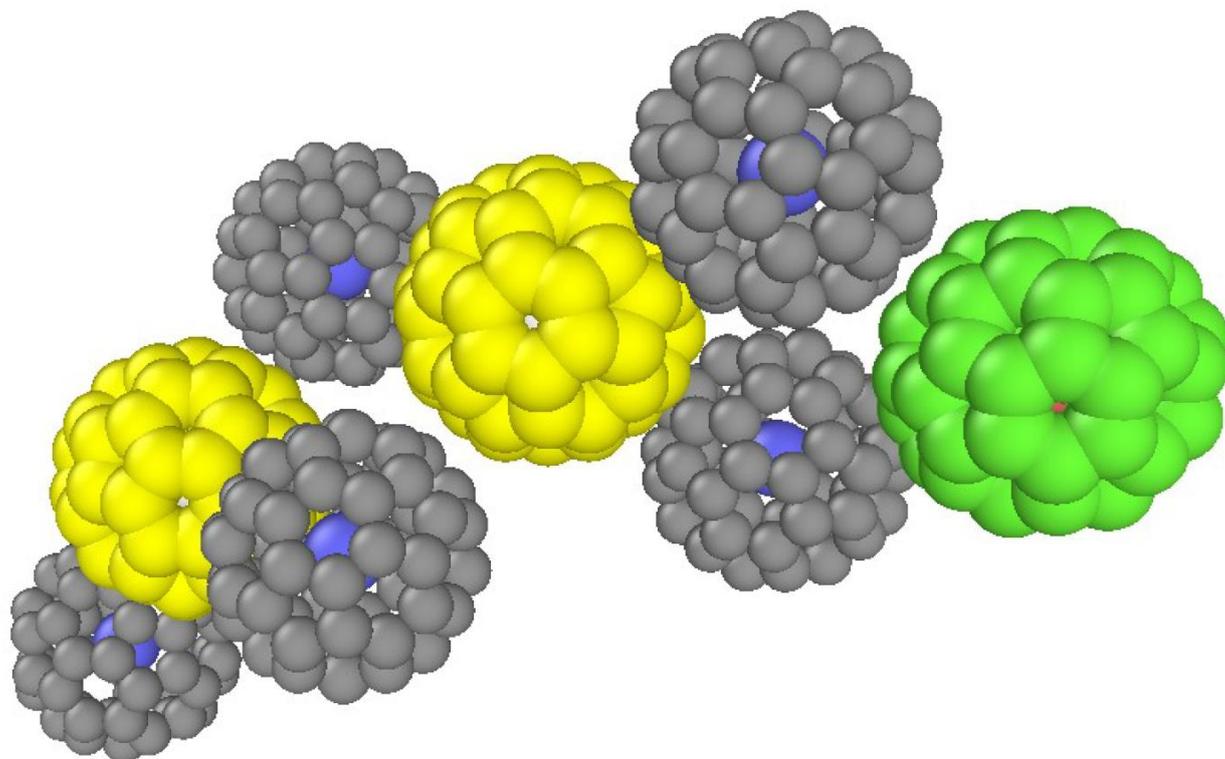


Figure 11. Icosahedral charged particles with various centre charges

The minimal model for a rotating colloidal motor

We found that in order to have self-assembling behaviour of a cluster which can then be capable of a rotational motion easily distinguishable during MD runs, we need to create artificial traps that act on at least two particles so that one axis gets fixed into a position in space, and this will become the axis of rotation. The system contains 3 types of particles: a central charged planar rigid body which will act as an axis (rotation around two sites, one of which is charged +6), two linear rigid bodies with a central charge of +4 and two apex sites with excluded volume interaction, and 7 'ligand' particles (linear rigid bodies with the two apex sites charged -1). If we put these building blocks in a box, they will readily self-assemble into an asymmetric triangle-shaped unit which attaches to the planar rigid body in a tilted manner, like the blade of a fan. The next obvious step was to see how the thermal energy can induce rotatory motion, therefore inert solvent molecules had to be added to the system. For better visualisation, the two axial particles get fixed to a region of the box by using three perpendicular planes that interact only with these particles (basic Lennard-Jones interaction).

Adapting the model for explicit solvent

Clearly, one cannot exploit directional rotatory motion using collisions of a blade with solvent molecules undergoing Brownian motion (see also Maxwell's demon). However, we found that starting the simulation far from equilibrium can induce directional rotatory motion, which will obviously turn into rotations in either directions once equilibrium has been reached. The plane perpendicular to the z axis, interacting with one site of the 'M' particle also interacts with the solvent particles (excluded volume, repulsive term of the LJ potential). This plane, together with another one at the bottom of the box is designed to keep the solvent particles on one side of the box so that eventually a pressure difference can be simulated.

Our minimal model building blocks readily self-assembled into the desired asymmetric fan blade structure even in the presence of solvent particles which interact with the rigid body sites using short-range excluded volume interactions. On low temperatures, the self-assembled blade began rotating.

Exploring different energy transfer methods to colloidal motors

We are aware of experiments in vacuum chambers where a levitating nonspherical particle was capable of directional rotation due to an anisotropic thermal gradient. In our case, we tried to replicate a similar behaviour by using two different thermostats, a low-temperature thermostat for the rotor, and a higher temperature thermostat for the solvent. However, as the thermostats work

by simply rescaling the velocities in simulations, with this approach it is not possible to create an anisotropic temperature gradient and therefore the simulation cannot produce directional rotation. However, we found out that by lowering the temperature of the rotor, the rotation is greatly enhanced, but eventually clockwise and counterclockwise rotations almost cancel out during long simulations.

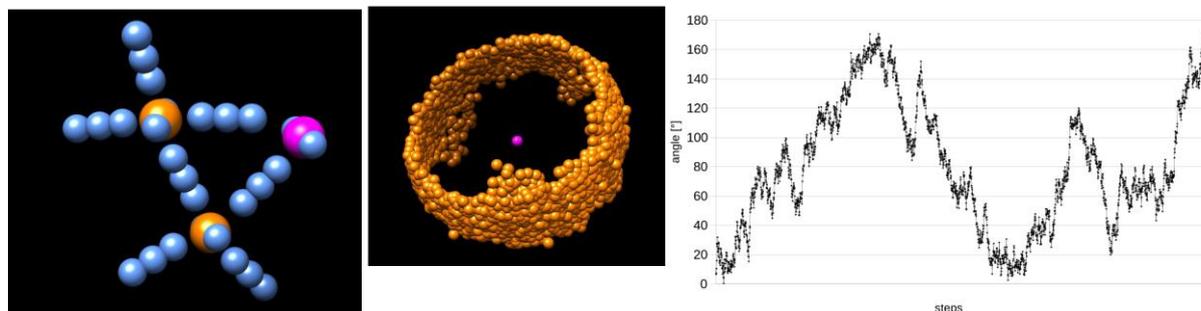


Figure 12. A basic rotor model and the points in space occupied by the rotor vertices during a simulation. The angle variation between a vertex and the rotation axis is also plotted.

A top view of a long simulation of the self-assembled rotor in explicit solvent can be accessed at https://youtu.be/t9GM_H9wXJo. That simulation contains two long, clockwise rotations (5 and 6 full rotations, respectively), while counterclockwise rotations take place on longer timescales but on shorter, more sequential distances. The system therefore has a likely asymmetric activation energy profile for the rotation, but that cannot be investigated with energy landscape methods due to the necessary presence of explicit solvent particles. The charges in that simulation were +12 for the two main 'M' particles, +18 for the 'M' particle within the axis, and -3 for the two charged sites on the 'L' particles.

We also explored different methods in order to produce an anisotropic thermal field by running NVE simulations on equilibrated systems with explicit solvent, attached to a low-energy 'heat sink' composed of a crystalline configuration of cold particles. An alternative method was explored as well, where there is a pressure difference between the two sides of the rotor, and therefore a net influx of solvent particles through a hole within the range of the rotor. However, due to the lack of proper initial equilibration and the very long computational time needed, we could not yet establish if these are viable methods to simulate unidirectional rotation.

Phase 3 activities

Optimization and parameter refinement of the minimal quasicrystal model

The minimal quasicrystal model contains four different particles in two rigid bodies (M-type and L-type, for a more in-depth description, see the publication in *Nanoscale Advances*). The main difference between the Goldberg model and the quasicrystal model is that in the latter, both the M and L particles are linear. Planar coordination around the center of the M particles is enforced by repulsive, uncharged sites located symmetrically along its main axis, with a multiply charged central particle. Coordination incompatible with crystalline order is driven by this central charge. The ends of the L (ligand) particles are oppositely charged, and overall neutrality of the system is ensured by the appropriate M:L particle ratio.

The charges set for the minimal quasicrystal model are: site A with +5 charge (this is the centre of the M rigid body), site C with -1 charge (the L particle contains two such sites). Lennard-Jones repulsion strength parameters were set in order to ensure a preference for planar coordination.

Dynamics of interconversion between crystalline and quasicrystalline phases

Starting from the optimised parameters for quasicrystal formation, we investigated how different M-L charge ratios influence the preferred assembly. We therefore set up systems with M-L ratios of 4:2, 5:2, 6:2 and 7:2, corresponding to preferred coordination numbers of 4, 5, 6 and 7. For each simulation we used 600 M particles, and the appropriate number of L particles for charge neutrality (1200 to 2100) for each charge ratio. Starting from randomly placed rigid bodies in a cubic box of 250 length units (300 length units for the biggest system with charge ratio 7:2), we then simulated these systems at low temperature ($kT=0.02$). Figure 13 shows the starting structure and first intermediates for the 7:2 charge ratio.

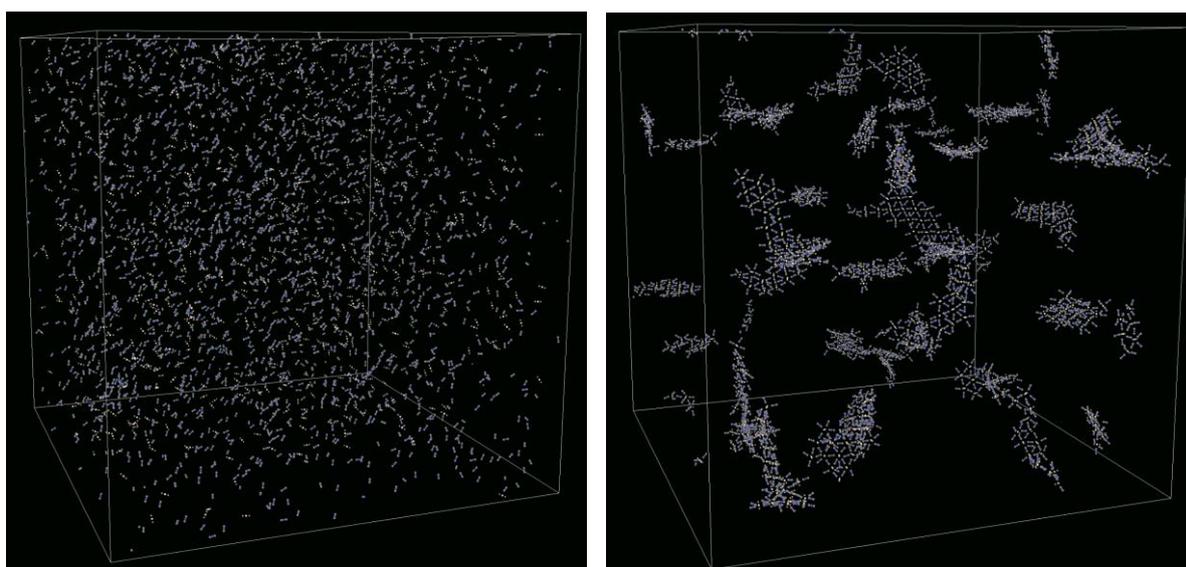


Figure 13. Starting structure and the formation of small assemblies for the charge ratio of 7:2.

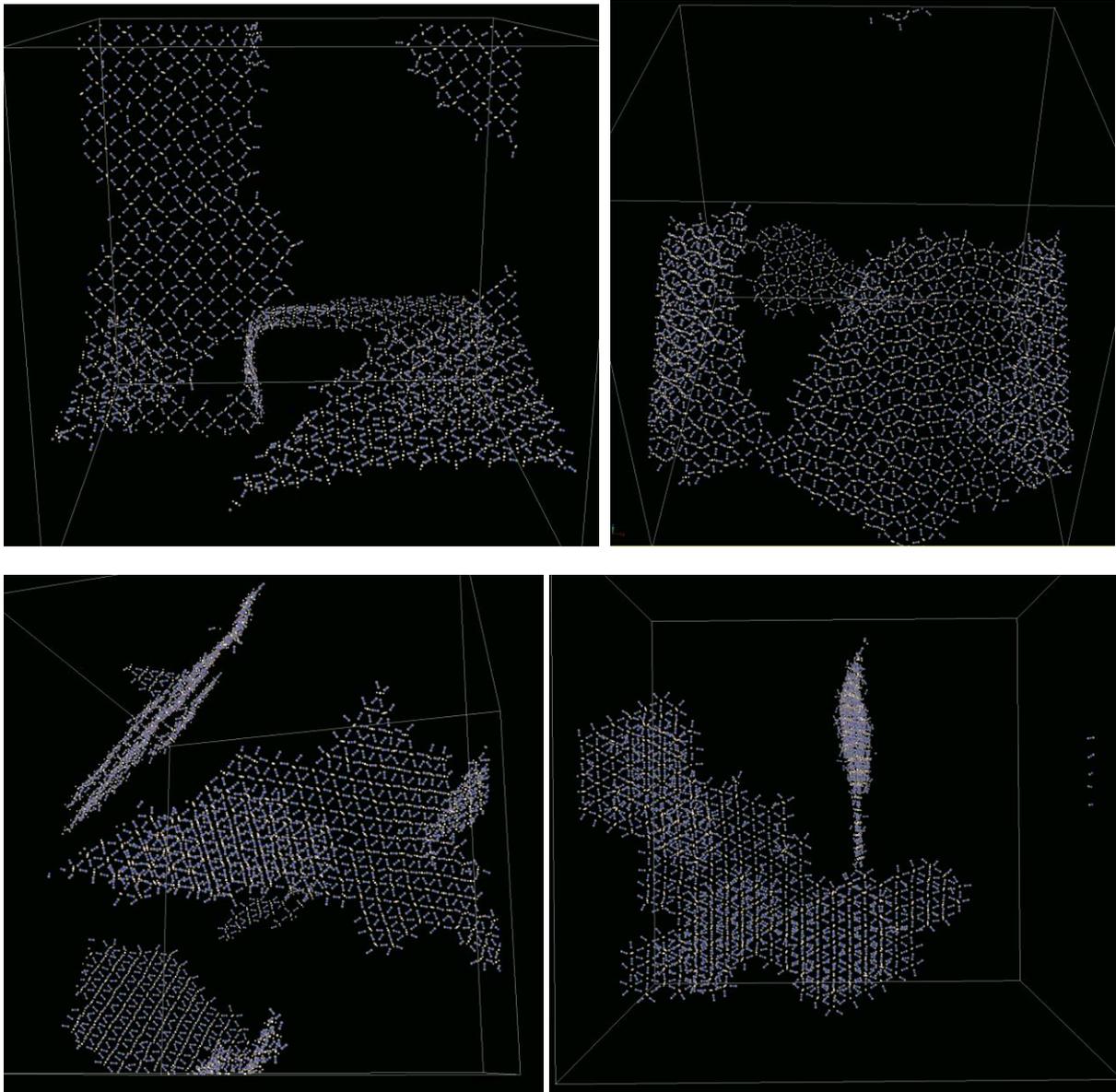


Figure 14. Final structures for simulations with charge ratios of 4:2 and 5:2 (top), 6:2 and 7:2 (bottom). The quasicrystalline structure is observed for charge ratio 5:2, the rest form sheets compatible with crystalline order.

Figure 14 shows the final assemblies for the four charge ratios considered, after extensive simulation. All systems prefer planar coordination, and the formed crystalline or quasicrystalline sheets can span the whole periodic box, forcing the sheets to bend.

Exploring the growth of quasicrystals with MD or MC

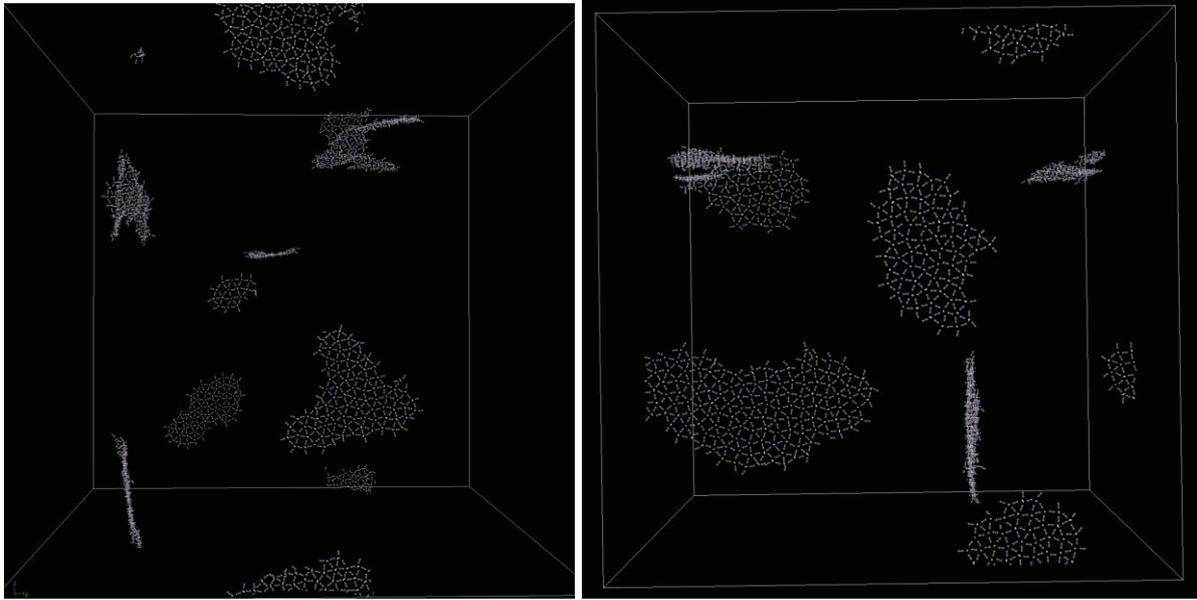


Figure 15. Assembly of single sheets with local quasicrystalline order at low densities (box lengths 500 and 400, respectively)

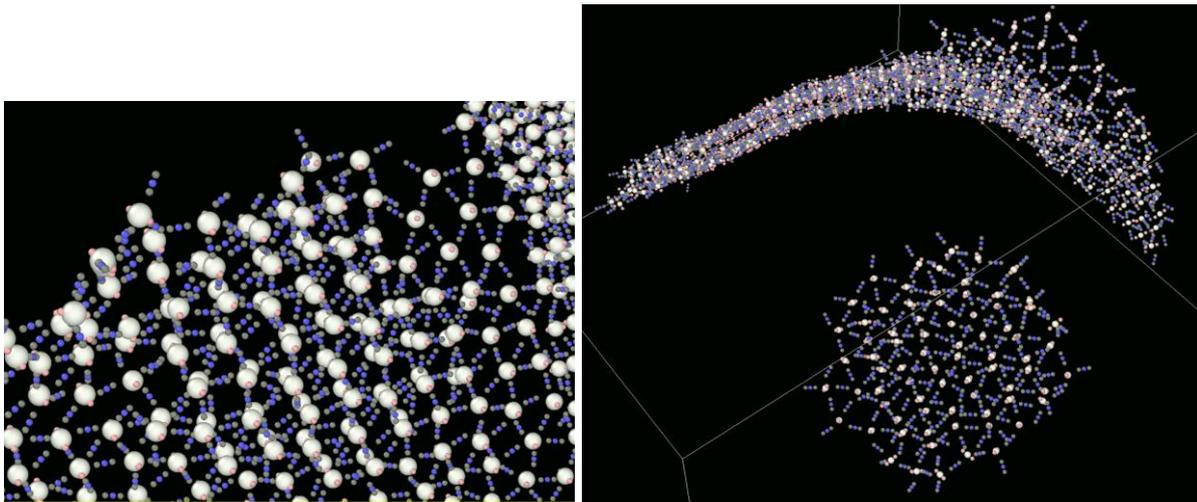


Figure 16. Example of parallel sheets in which the QC arrangement is copied over from one sheet to another (box length of 300).

We carried out molecular dynamics simulations in our framework set up in HOOMD-Blue, with fixed numbers of particles (600 M and 1500 L particles, charge ratio 5:2) and a variable density of particles (box lengths 250, 300, 400 and 500 distance units). All simulations were run at low temperature ($kT=0.02$). We found that the density greatly influences the formation of quasicrystalline sheets. At low densities, small clusters readily form which then slowly merge while keeping local quasicrystalline order. Only single sheets are formed if the density is low. Decreasing the box size decreases the number of sheets that form during the simulation, until we start observing multi-layered sheets forming. These sheets have quasicrystalline order along their main plane, and crystalline periodicity along the direction perpendicular to that plane. The multi-layered

configuration is similar to the structure of graphite, but with quasicrystalline order within the sheet. One sheet acts as a template for the other sheets that form around it, therefore the most part of the second sheet can be obtained by translation of the first sheet.

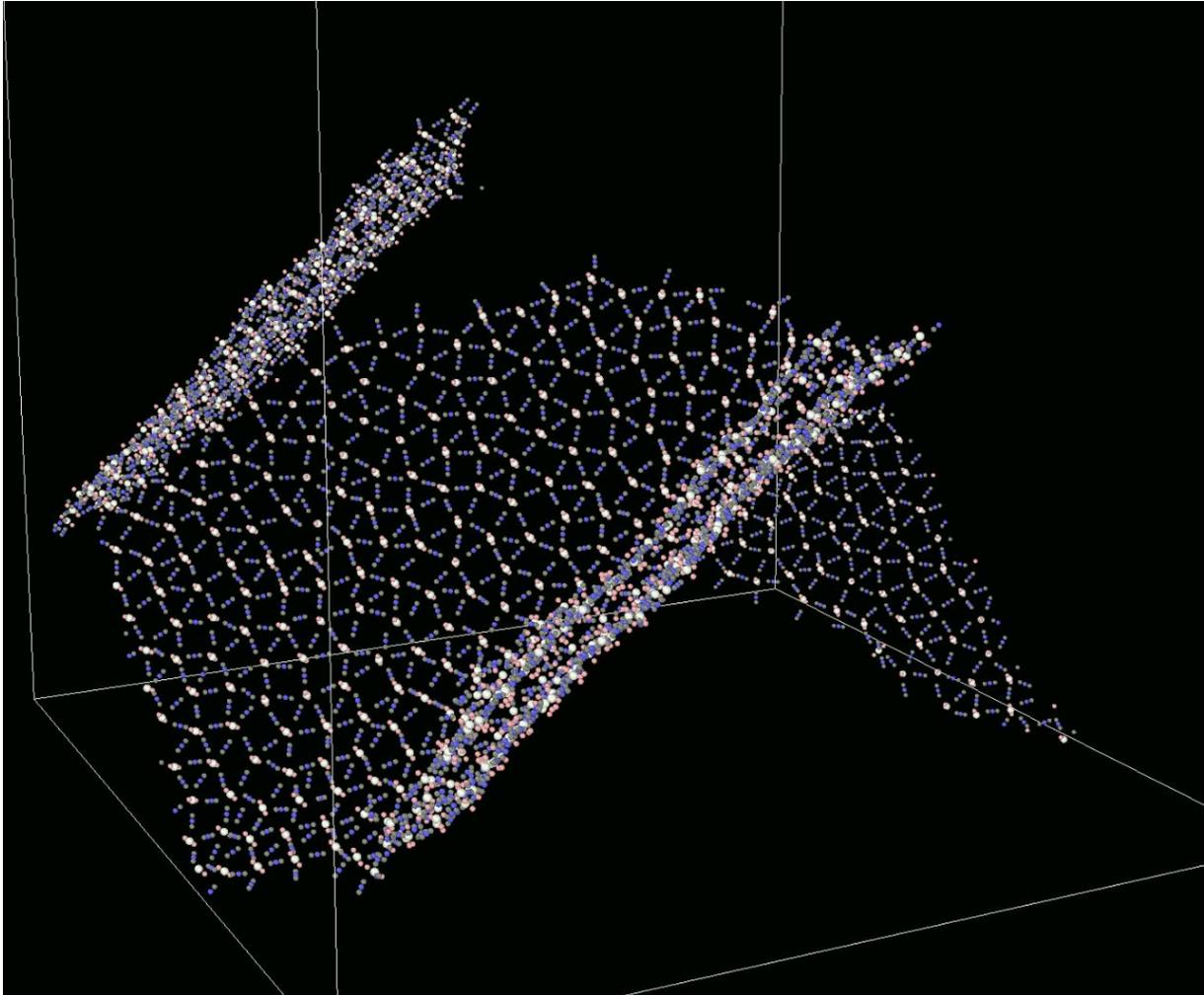


Figure 17. Undulating sheet with local quasicrystalline order and a smaller sheet stacked on it (box length 250).

Estimated impact of obtained results

Our results on mesoscale motors can aid rational design of experiments to replicate the observed behaviour and exploit the rearrangement mechanisms to transform energy gradients into rotatory motion. The minimal physics necessary to construct this system is very simple, therefore it can drive experiments in the field. We will also continue pursuing this direction of research and applying the models to mesoscopic systems in which the building blocks themselves are capable of non-random motion (bacterial motors).

Our most important result is the minimal model for quasicrystalline self-assembly. The formation of mixed crystalline-quasicrystalline order is an unexpected, but emergent behaviour of the model building blocks, and we expect very interesting physical properties from analogous, real-life quasicrystals on the mesoscale in the future.

Sfantu Gheorghe, on 14.09.2022

Dr Szilard Fejer