

POLITECNICO DI TORINO

Master of Science in Nanotechnologies for ICTs

Master Thesis

**Assessment of Molecular Dynamics
simulations for self-assembly based
electronics**



Supervisors

prof. Mariagrazia Graziano
prof. Gianluca Piccinini, M.Eng. Yuri Ardesi

Candidate

Mattia Siviero

2021/12/17

Contents

Acknowledgements	5
Introduction	6
molecular QCA	7
Molecular dynamics simulations[6]	11
The core loop	11
Integration schemes for controlling temperature and pressure	17
Important concepts and criticalities	21
Thermodynamic quantities and postprocessing	22
Force Fields	22
Software	25
LAMMPS, an open-source molecular dynamics package	25
Visualizing LAMMPS trajectories: the VMD software package	25
ORCA, a quantum chemistry software package	25
Part 1: hexanethiol deposition on gold on Lammmps	26
Simulation with reaxFF force field	26
Simulation with a united atom force field	30
Conclusions of part 1	35
Part 2 : deposition of 6-(Ferrocenyl)hexanethiol on a surface	36
Modeling the bonds	37
computing a dihedral angle with Orca	43
Deposition of a 6-(Ferrocenyl)hexanethiol droplet on a gold surface	45
Deposition atom by atom	49
Molecular dynamics study of 6(Ferrocenyl)hexanethiols atoms anchored on a gold substrate	55
MD results	59
MD Postprocessing	65
Conclusions of part 2	77
Part 3: deposition of 6(ferrocenyl)hexanethiol on trenches and wires	80
6(ferrocenyl)hexanethiol on a wire: oxidation and electric field application	88
Obtaining a wire by etching a 2D SAM	107
Conclusions of part 3	111
Deposition of bisferrocene atoms and electrical analysis on Scerpa	112
Bisferrocene molecule	112
Scerpa	114
Experiments	115

Matlab postprocessing	123
Information transmission on Scerpa	130
Scerpa results and conclusions	133
Conclusion and future perspectives	133
Appendix A: Lammmps implementation	133
Part 1: FcC6S code	133
Part 2: anchoring molecules to a substrate	142
Part 3: FcC6S wire	143
wire	143
Oxidation and electric field application	143
conclusion	144
Bibliography	145

Acknowledgements

Over the course of the writing of this dissertation I received a great deal of support and I wish to thank those who allowed me to reach the final goal of my graduation. First I wish to thank my relator prof. Mariagrazia Graziano and corelator Prof. Gianluca Piccinini for directing me on my master thesis work. A special thanks goes to corelator M.Eng. Yuri Ardesi whose constant guidance and advice helped me progress through my work when I encountered difficulties.

I would like to give my sincerest thank you to my parents who always unconditionally supported me and were always available to listen and advise whenever I encountered material or emotional difficulties not only during my dissertation, but throughout my studies. I would not have reached this point without them.

Finally I would like to thank my friend, both from before my university and encountered along the way, with whom I shared the beautiful moments and the worries along the way and created many precious memories, and whose company and support helped me overcome some of the most stressful moments. Even when a global pandemic imposed physical isolation on everyone, we found ways to spend time together with the means available. That was crucial to overcoming those trying times.

Computational resources were provided by HPC@POLITO, a project of Academic Computing within the Department of Control and Computer Engineering at the Politecnico di Torino (<http://hpc.polito.it>), and VLSI lab, Research group of the Department of Electronics and Telecommunications at Politecnico di Torino (<https://www.vlsilab.polito.it/>)

Introduction

In the last decades the interest, both at the industry and research level, for processes, structures and systems at the nanoscale has steadily grown in importance. Micro and nanotechnologies (specifically "micro/nano-electronics and photonics") are considered by the European commission among the key enabling technologies for industry[2] and the nanotechnology market has experienced explosive growth, with a predicted compound annual growth rate of more than 10% and is expected to reach a total \$70 billion value by 2026[1].

However, some criticalities may hinder progress, especially toward commercialization. Namely the high cost of state-of-the-art equipment (including cleanrooms, reactors, high resolution characterization technology and machinery for lithography), the time required to iterate between successive experimental trials, and the difficulty in the characterization at the nanoscale and especially of molecular systems.

Even more critical is the presence of barriers toward nanotech commercialization among which there are time and cost issues:

- the time between research and commercialization can easily span between 3 and 10 years or even longer and the long time before any return of investment can thus discourage investment;
- the costs to create competitive products from basic research results is two orders of magnitude greater than the costs of research;
- the expensive infrastructures and the fast advancement of the state of the art hinder the ability of small businesses and research groups to conduct the needed technological steps for research.[3]

Software simulations may help bridge the time and resource gap. With the growth of computational power and the increased availability of computer clusters to industry and research groups software simulation of systems and processes has become a viable avenue for research and design in multiple fields. Cad software for example allows iterating on the design of structures and devices of all scales, from bridges to appliances to the most recent generations of transistors, with multiphysics simulations integrating several domains in a single software suite.

For molecular systems especially, three main classes of techniques are of great importance: ab initio method, molecular dynamic techniques and Monte Carlo techniques.

Ab initio methods attempt to find an approximate representation of the electronic configuration and the electrostatic potential of a system through methods like Hartree-Fock equations and density functional theory.

Kinetic Monte Carlo techniques study the time evolution of processes where the transition rate between states, or the probability of the occurrence of a process over time, is known. They determine at random, according to the transition probabilities, the kind of transition occurring over a certain time interval and then apply the relevant physical laws to determine the state of the system after that interval has elapsed.

Molecular dynamics is a simulation technique that applies the numerical solution of the

Newton's equations of motion for a system of atoms to study the evolution of that system from a specified initial configuration and under a set of predefined constraints.[4] The latter will be the main focus of this dissertation.

One difficulty in studying molecular electronics systems is that the small scale limits the ability to investigate the system geometry because of the resolution limit. Thus simulation approaches like molecular dynamics might represent a way to obtain an insight into how molecules arrange themselves and also allow to better interpret experimental results. This work features a section describing molecular dynamics as a method followed by four experiments using molecular dynamics investigating systems of increasing complexity.

molecular QCA

One aspect of nanoscale device that can be exploited for beyond-CMOS technology is the charge confinement that occurs in structures that are small enough to show potential-well effects in the three dimensions. These structures are said to be zero-dimensional (0D) structures or quantum dots. If a pair of quantum dots is close enough that tunneling can occur, charge can be transferred between quantum dots. This has important technological consequences: if parallel pairs of quantum dots are taken so that the dots in each pair are close enough to allow for charge transfer, and nearby pairs are too far for charge transfer but close enough that the dipole field of one pair can induce charge transfer in the nearby pair[5], then, in order to minimize energy, nearby dot pairs will have the mobile charges localized in diagonally opposite positions. (see figure 1). This means that if a pair is polarized in one direction, the successive pair will have opposite polarization. Thus the basic unit is composed by two coupled pairs of quantum dots. A sequence of these basic units constitutes a molecular wire. Assuming "1" is the state where the second molecule of the pair has the positive charge pointing upwards, figure 2 shows the following: an input forces the first cell in an "1" state. To minimize energy the following cells are all forced into a "1" state.

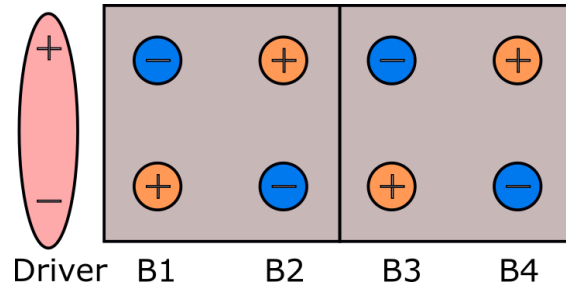


Figure 1: QCA cells with an external driver

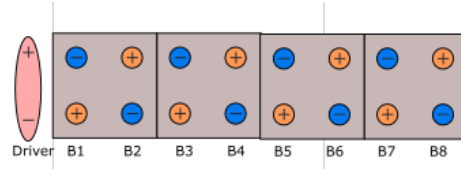


Figure 2: QCA wire

Molecular QCA logic

Other than information transmission, also logic operations can be performed through molecular QCA technology.

inverter

If a cell is substituted with an half cell, a "1" state in the previous cell will return a "0" state in the output of the half cell. This means a half cell behaves as an inverter. However using half cells means a wire with an inverter is never the same length as a wire without an inverter, which creates problems in complex circuits.

A different strategy is to have an input split into two lines and then put together in a line placed at 45° angle with the two ones, generating an inversion in the signal at the end (see figure 4)

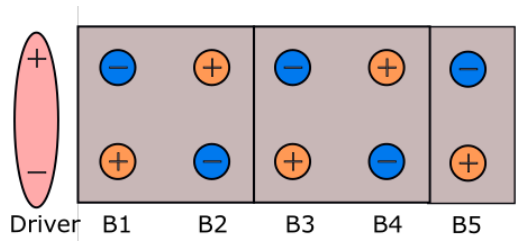


Figure 3: QCA inverter with half-cell

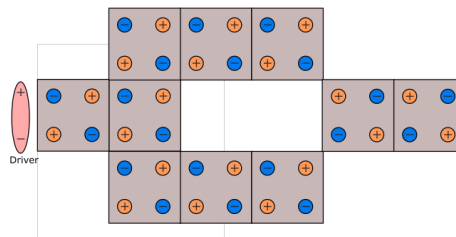


Figure 4: QCA inverter with full cells

Majority-voter, AND and OR

The majority voter is cell placed among three inputs on three sides and with an output

on the remaining side. The cell at the center will align with the majority of the inputs in order to minimize energy and thus the output will be equal to the majority. If one input is fixed to a logical "1" then the output will be "1" if at least one of the other inputs is a "1". Which corresponds to the function of an OR gate. If one input is fixed to a logical "0" then the output will be "0" if at least one of the other inputs is a "0". Which corresponds to the function of an AND gate.

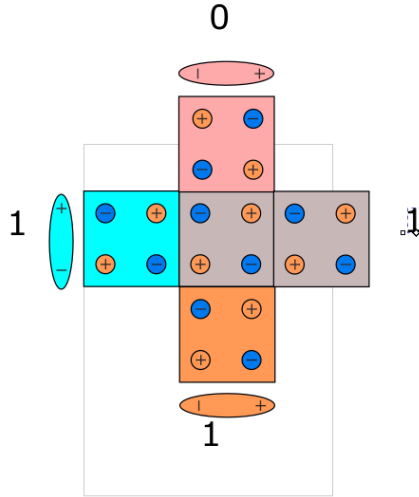


Figure 5: QCA majority voter

Implementation of molecular QCA

Any implementation of molecular QCA technology must have a basic unit with two regions where the charge can be confined, so that the charge can be transferred between these two regions under the action of an electric field. Equally important however is the presence of a reset state: if a field is applied orthogonally to the cell, the molecule must be forced in a state where the charge is drained from both the active regions and the molecule is in a null state. This is fundamental in digital applications to operate synchronously with a clock and to prevent information backpropagation. A reset phase where the clock prevents information transmission is alternated with a set phase, where the charge is transferred back to the two dots and information can be transmitted.

However a single clock phase doesn't work for molecular QCA circuits: if a wire is particularly long, after the clock is released, since the reset state is not stable, one molecule toward the end might collapse in one of the two polarized states and this information backpropagates through the wire, clashing with the information forward propagating. Thus multiple clock phases are actually needed: the clock is released first in a small section of the wire and then, after the information has reached the end, it is released in the following portion and then in the next one. After the information has reached the third clock phase, the first clock phase can be reset again and a new input can be accepted.

One way to achieve this type of system is to have molecules with two functional groups acting as quantum dots and an anchoring group binding the molecules to the substrate and hosting the charge in the reset state.

Molecular dynamics simulations[6]

The purpose of molecular dynamic simulations is to study a system evolving according to Newton's equations of motion, from an initial configuration and under a set of predefined conditions. This allows to analyze the time evolution and study the properties of the equilibrium configuration of said system.

The simulation consists of the following steps: the creation of a model system with N particles, its time evolution through a discrete-time solution of the equation of motion and the analysis of the equilibrium configuration. Such analysis can include the study of the final geometry and the computation of the values of physical quantities from time averages over the final configuration (constituting a good approximation of ensemble averages under the ergodic principle). The theoretical background on molecular dynamics has been mostly derived from "Understanding Molecular Simulation 2nd Edition From Algorithms to Applications" by D. Frenkel and B. Smit[6]

The core loop

The basic steps for running an MD simulation are the following:

1. The system is initialized: the particle initial positions, velocities, masses, charges are defined together with all other properties that need to be known a priori.
2. Forces on all particles are calculated. This can be done by running an ab-initio calculation at each time step to compute the gradient of the potential acting on each particle (accurate but extremely computationally expensive) or specifying a set of empirical equations defining two-body and multibody potential terms.
3. The equations of motion are time integrated with a discrete-time approximation to compute the velocities and the positions after a determined length of time (timestep). Force computation and time integration are reiterated until a maximum specified number of iterations is reached.
4. Measurables can be computed for the new timestep either on the fly or by postprocessing the output trajectory file.

The output can include a file with the measurement results and a trajectory file detailing positions and velocities every n timestep. Here n is chosen so that the time for I/O operations and the size of the output file are manageable. Other measurables can be computed from positions and velocities.

Initialization

Positions and velocities

Proper initialization is the key to ensuring that the time evolution of the system occurs as desired. For that reason initial positions should be set so that force terms, and thus displacements between timesteps, are not too high, driving the system into instability. This can occur for example if the initial positions of different particles are too close or if

two atoms connected by a spring force are too far away from the equilibrium distance. For the basic Verlet integrator, velocities are not used directly. However, the integration requires knowledge of both the previous ($x_{t-1}(i)$) and current $x_t(i)$ position of every particle. Thus velocities are used to compute an estimate of the positions before the initial condition through

$$x_{-1}(i) = x_0(i) - v(i) * dt \quad (1)$$

This is a very rough estimate that only accounts for conservation of linear momentum ignoring forces, but is good enough for bootstrapping the simulation.

Simulation box and spatial boundary conditions

When preparing a molecular dynamics simulation, a simulation box must be defined. It is a parallelepiped identified by three vectors corresponding to the length and orientation of the three pairs of parallel edges. It represents the simulation domain.

The size of the box can be fixed (constant volume) or changing to enforce constant pressure.

Once the box is set, boundary conditions can be defined. The main types are fixed and periodic. With fixed boundary conditions a particle leaving the simulation box is no longer considered part of the system and neglected for force computations. Having particles leave the box may be considered a fatal error and the simulation stopped if mass conservation is enforced or if the missing atom is part of a bond. Otherwise, the lost particles are just ignored. Periodic boundary conditions are enforced with the following two operations:

- If a particle leaves the simulation box it is wrapped back into the box by subtracting one period from the coordinate it exceeded. As a result, the particle is seen as reentering the box from the other side.
- To compute the forces on any atom, the algorithms account both for neighbors inside the box and their periodic replicas in the periodic directions. This can lead to an atom interacting with both a second atom and some of its replicas. For this reason a cutoff radius for each pairwise interaction has to be specified and that interaction is neglected if the interacting pair is at a greater distance than the cutoff radius. If the box is larger than twice the greatest cutoff, it becomes impossible for an atom and its replica to both be near enough to interact with any other atom.

A particular type of fixed boundary conditions are the reflecting ones: a plane can be defined with the property that particles on one specific side of the plane are moved at the following timestep to the position of their mirror image, with opposite momentum. An alternative to reflection is defining a repulsive force acting on atoms crossing a certain coordinate. In both cases, one must take care that the particle doesn't travel a long enough distance in a timestep to escape the simulation domain before it's reflected.

Force computation

Once the potential profile for the interaction between particle i and j is known, the force relative to that interaction can be easily obtained by computing the gradient. For the

simplest case of two-body interactions

$$F_{ij} = -\nabla U_{ij} \quad (2)$$

$$f_{x,ij} = -\left(\frac{\partial U_{ij}}{\partial x_{ij}}\right)_{r_{ij}} = -\left(\frac{\partial U_{ij}}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial x_{ij}}\right)_{r_{ij}} \quad (3)$$

For example, if the interaction under analysis is a Lennard Jones potential,

$$U_{ij}^{LJ} = 4\epsilon\left(\frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6}\right) \quad (4)$$

then the corresponding fx directed force is

$$f_{x,ij}^{LJ} = -\frac{\partial U_{ij}^{LJ}}{\partial x_{ij}} = -\frac{x_{ij}}{r_{ij}} \frac{\partial U_{ij}^{LJ}}{\partial r_{ij}} = \frac{24\epsilon}{r_{ij}} \left(2\frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6}\right) * \frac{x_{ij}}{r_{ij}} \quad (5)$$

with $F_{ij} = -F_{ji}$

Some interactions may involve more than two atoms. What is important to know though is that the force computation step is very computationally intensive if all pairs are considered, since the number of pairwise interactions scale as $N * (N - 1)/2$ (the number of possible pairs). It can be sped up with some techniques but it is still the limiting step for the simulation speed.

Integrating the equations of motions

Once the forces have been determined, a discrete-time integration scheme is applied. Time is discretized with a fixed step Δt and the information from the configuration of the system at previous moments in time is used to approximate the configuration after a time step has elapsed. A common integration scheme is the Verlet algorithm.

The Verlet algorithm

Taking a small time variation dt , the displacement of the particle can be expanded in Taylor series around the time t

$$r(t + dt) = r(t) + r'(t) * dt + \frac{r''(t)}{2} dt^2 + \frac{r'''(t)}{6} dt^3 + \mathcal{O}(dt^4) \quad (6)$$

Let $f(t)$ being the force acting on the particle at time t , m the mass and $v(t)$ the velocity at time t , then

$$\begin{aligned} r(t + dt) &= r(t) + v(t) * dt + \frac{f(t)}{2m} dt^2 + \frac{r'''(t)}{6} dt^3 + \mathcal{O}(dt^4) \\ r(t - dt) &= r(t) - v(t) * dt + \frac{f(t)}{2m} dt^2 - \frac{r'''(t)}{6} dt^3 + \mathcal{O}(dt^4) \end{aligned} \quad (7)$$

By just taking the first equation and truncating it at the second order, one obtains the Euler algorithm.

$$r(t + dt) = r(t) + v(t) * dt + \frac{f(t)}{2m} dt^2 + \mathcal{O}(dt^3) \quad (8)$$

However the standard algorithm for molecular dynamics, the Verlet algorithm, can be obtained instead by summing the two equations instead

$$\begin{aligned} r(t + dt) + r(t - dt) &= 2r(t) + \frac{f(t)}{m}dt^2 + \mathcal{O}(dt^4) \\ r(t + dt) &= 2r(t) - r(t - dt) + \frac{f(t)}{m}dt^2 + \mathcal{O}(dt^4) \end{aligned} \quad (9)$$

Thus if the position at two successive timesteps is known together with the forces acting at the second timestep, the successive position is computed up to an error term that scales with the fourth power of the timestep.

It's worth noting that the algorithm doesn't explicitly use the velocity and the initial velocity is only used to approximate the $r(-dt)$ configuration in order to bootstrap the algorithm. It can however be derived by subtracting the equations at (7)

$$\begin{aligned} r(t + dt) - r(t - dt) &= 2v(t) * dt + \mathcal{O}(dt^3) \\ v(t) &= \frac{r(t + dt) - r(t - dt)}{2dt} + \mathcal{O}(dt^2) \end{aligned} \quad (10)$$

The leapfrog ad velocity verlet schemes

Modified schemes that can produce trajectories equivalent to the Verlet scheme are the leapfrog algorithm and the velocity Verlet algorithm. The leapfrog algorithm is obtained from the Verlet scheme: first, the velocities are evaluated at half-integer timesteps, then the positions at integer timesteps are computed. The velocities at half-integer timesteps are defined as the average velocities between two timesteps assuming constant forces.

$$\begin{aligned} v(t - dt/2) &= \frac{r(t) - r(t - dt)}{dt} \\ v(t + dt/2) &= \frac{r(t + dt) - r(t)}{dt} \end{aligned} \quad (11)$$

inverting the latter

$$r(t + dt) = v(t + dt/2) * dt + r(t) \quad (12)$$

with the velocity updates derived by substituting the velocity Verlet expression at (9) after subtracting the half step velocities:

$$\begin{aligned} v(t + dt/2) - v(t - dt/2) &= \frac{r(t + dt) + r(t - dt) - 2r(t)}{dt} \\ v(t + dt/2) - v(t - dt/2) &= \frac{f(t)}{m}dt \\ v(t + dt/2) &= v(t - dt/2) + \frac{f(t)}{m}dt \end{aligned} \quad (13)$$

This algorithm produces identical trajectories but velocities are not synchronous with the position. Thus kinetic and potential energies are computed at half a timestep difference and the total energy cannot be computed.

The velocity Verlet algorithm starts from the Taylor expansion truncated at the second order

$$r(t + dt) = r(t) + v(t) * dt + \frac{f(t)}{2m} dt^2 + \mathcal{O}(dt^3) \quad (14)$$

but uses a velocity update scheme that's similar to leapfrog Verlet. However velocities are synchronous with the positions and forces. The velocity update term is taken as the average force between the current and new timestep.

$$v(t + dt) = v(t) + \frac{f(t + dt) + f(t)}{2m} dt \quad (15)$$

Thus the loop is as follows:

1. Compute the forces at time t
2. Compute the velocities at time t using (15) and the forces at t-dt
3. Compute the positions at t+dt and reiterate

To prove the equivalence to the Verlet algorithm one can start by writing down the truncated Taylor expansions around $r(t)$ and $r(t+dt)$ r is the approximate position derived by the algorithm:

$$\begin{aligned} r(t + dt) &= r(t) + v(t) * dt + \frac{f(t)}{2m} dt^2 \\ r(t) &= r(t + dt) - v(t) * dt - \frac{f(t)}{2m} dt^2 \\ r(t + 2dt) &= r(t) + v(t + dt) * dt + \frac{f(t + dt)}{2m} dt^2 \\ r(t + 2dt) &= r(t + dt) - v(t) * dt - \frac{f(t)}{2m} dt^2 + v(t + dt) * dt + \frac{f(t + dt)}{2m} dt^2 \\ r(t + 2dt) &= r(t + dt) + [v(t + dt) - v(t)] * dt + \frac{f(t + dt) - f(t)}{2m} dt^2 \\ r(t + 2dt) &= r(t + dt) + \frac{f(t + dt) + f(t)}{2m} dt^2 + \frac{f(t + dt) - f(t)}{2m} dt^2 \\ r(t + 2dt) &= r(t + dt) + \frac{f(t + dt)}{m} dt^2 \\ r(t + dt) &= r(t) + \frac{f(t)}{m} dt^2 \end{aligned} \quad (16)$$

Where the first truncated Taylor expansion has been substituted in the second and the velocity update term has been substituted to $[v(t+dt)-v(t)]$, showing that this update scheme yields an equivalent position update to the original Verlet scheme

Other algorithms

Other Verlet-like algorithms allow to obtain better estimates of the velocity. They are the Beeman algorithm and the velocity corrected Verlet algorithm.

However, if higher precision is needed to use a longer timestep without losing short-term accuracy, higher order integration schemes like the predictor-corrector algorithm can be used.

Choosing the algorithm

There are a few elements to account for when choosing an integration scheme. The time evolution of a real system obeys some properties. First of all energy must be conserved. Given a system with energy E , energy conservation means that, taking a hypervolume in phase space containing all and only the phase space points corresponding to states with a certain energy E , the trajectory in phase space corresponding to the time evolution of the system is entirely contained in that volume.

Secondly, taking a volume in the phase space having as coordinates the position and velocities of the classical point-like atoms in the system (representing a set of states the system is in) and applying Hamilton's equations of motion to that volume, the volume it is mapped to remains constant over time (Liouville's theorem). The phase space volume containing all points at a certain energy E is thus mapped to itself by Hamilton's equations of motion.

Lastly, the equations of motion are time-reversible, meaning that when a system is evolved in time through newtonian or equivalently hamiltonian dynamics, if a system follows a trajectory in phase space from point A to point B over a certain interval of time, by taking the system in point B, reversing all momenta and determining its time evolution according to Hamilton's equation, the same phase space trajectory is followed from B to A.

For an algorithm we care about whether it is area-preserving and time-reversible: if those two properties are absent the algorithm is not compatible with energy conservation and won't conserve the phase space volume over time.

The velocity Verlet algorithm is fast and has minimal memory requirements. Most importantly however it has low long-term energy drift because it's time-reversible [7] and as a consequence symplectic (phase space volume conserving), [7]. Thus there exist a Hamiltonian called "shadow hamiltonian" \tilde{H} : such that the phase space points describing the system at each timestep of velocity-Verlet integration lie on the trajectory described by the time evolution of the system under \tilde{H} . [8] This Hamiltonian converges to the true hamiltonian of the system at the limit of infinitely small timesteps. This means that in infinite arithmetic no energy drift occurs even with moderate timesteps.

In finite arithmetics drift might occur due to the accumulation of numerical roundoff errors and potential functions defined in such a way that they are not analytic at the cutoff. However if the potential is corrected to be analytic, the experienced drift of the Verlet algorithm is shown to be small enough to be neglected in most typical simulation lengths

Higher order algorithms like the predictor-corrector algorithm are generally neither reversible nor area-preserving so they should experience stronger drift. However, they are more accurate in the short term.

Lyapunov instabilities

A system has a Lyapunov instability when the upper bound on the error in the phase space trajectory due to an error in the initial conditions of the system increases exponentially with time. The equations of motion have this kind of instability so even the smallest error in the initial conditions will lead to diverging phase space trajectories. The advantage

of the velocity Verlet algorithm is that it approximates a true dynamical trajectory of the studied system thus it produces steady-state conditions and time averages that are representative of the real system.

Integration schemes for controlling temperature and pressure

Integrating Newton's equation of motion leads to sampling a microcanonical (constant energy) ensemble in the limit of no energy drift. However, the assumption of constant energy only holds if the system being studied is closed. The system should also be able to exchange energy with the environment. This is done to maintain boundary conditions in line with the experiment. We want in particular, in MD simulation, to be able to work at constant temperature (thus conserving the kinetic energy) or constant pressure.

In the case of Monte Carlo simulations particle insertion or deletion can allow simulating constant chemical potential ensembles, but that's outside the scope of this work.

To do this, boundary conditions called "thermodynamic boundary conditions" must be added to the system. If the system has either periodic or vacuum boundary conditions, the Hamiltonian is time-independent and invariant under translation or rotation of the system. In some cases however the boundary conditions add stochastic forces that limit the validity of some equations.

Ensemble

In the standard molecular dynamics simulations, the conserved quantities are the extensive ones: number of particles, volume and energy. This leads to a trajectory in a microcanonical ensemble.

The canonical ensemble is generated by coupling the system with a thermal reservoir: a subsystem such that the heat capacity of the original system is negligible with respect to that of the reservoir and both individual subsystems conserve volume and number of particles while the system including both also conserves energy (while the subsystems can exchange it). The two subsystems are at thermal equilibrium and thus together they form a microcanonical ensemble (since volume, energy and number of particles are conserved). By computing the evolution of such an ensemble and projecting the resulting trajectory over the $6N$ coordinates of the original system, one obtains the state of the original system and can compute averages and fluctuations. In the canonical (NVT) ensemble the temperature has a specified average value and the energy observable can fluctuate. There are fluctuations of the Hamiltonian (the total energy of the system) and the rms fluctuations depend on the heat capacity.

$$\sigma_E^2 = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = k_b T^2 c_v \quad (17)$$

The instantaneous temperature fluctuates too.

$$\sigma_T^2 = \langle \mathcal{T}^2 \rangle - \langle \mathcal{T} \rangle^2 = N_{df}^{-1} T^2 \quad (18)$$

N_{df} is the number of internal degrees of freedom. While fluctuations can be neglected for computations on large systems, MD simulations are rarely large enough. So any

thermostatted MD simulation at equilibrium will deal with instantaneous fluctuations of both T and E. What matters is the conservation of the time average of those quantities. If the system and reservoir are coupled in the isothermal-isobaric ensemble, the same discussion as above applies to the volume-pressure conjugated variables: the reservoir can exchange volume with the system (as if they were connected with a piston). The total volume is however conserved so the system + the reservoir still form a canonical ensemble at equilibrium. The volume fluctuations are:

$$\sigma_V^2 = \langle \mathcal{V}^2 \rangle - \langle \mathcal{V} \rangle^2 = V k_b T^2 \beta_T \quad (19)$$

The thermodynamic potential is the enthalpy, the Legendre transform of the Hamiltonian with respect to the pressure-volume conjugate pair:

$$H = \mathcal{H} + P\mathcal{V} \quad (20)$$

with the following fluctuations on the volume proportional to the constant pressure specific heat:

$$\sigma_V^2 = \langle (\mathcal{H} + P\mathcal{V})^2 \rangle_{NPT} - \langle \mathcal{H} + P\mathcal{V} \rangle_{NPT}^2 = V k_b T^2 \beta_T = k_B T^2 c_p \quad (21)$$

Thermostats

A thermostat algorithm is a modification of the MD integration scheme aiming at generating a canonical ensemble. Other than allowing the study of temperature-dependent process and allowing to mimic experimental conditions where the temperature is kept constant, it aims to avoid energy drifts. The practical way to implement it is by simulating the coupling of the system to a heat bath.

The average kinetic energy, according to the equipartition principle is

$$K = \langle \mathcal{K} \rangle = \frac{k_B}{2} N_{df} T \quad (22)$$

thus for this to hold we define the instantaneous temperature as:

$$\mathcal{T} = \frac{2}{k_B N_{DF}} \mathcal{K} = \frac{k_B}{2} N_{df} T \quad (23)$$

to have the average temperature be equal to the thermodynamic one. [9]

The Langevin equation

Most thermostats are based on the numerical solution of the Langevin equation:

$$\frac{\partial^2 r_i}{\partial t^2} = m_i^{-1} F_i(t) + \gamma_i(t) r_i(t) + m_i^{-1}(t) R_i(t) \quad (24)$$

which adds one atomic friction term with γ_i positive coefficient and a stochastic force. When a thermostat neglects the stochastic term and assumes a single friction coefficient for all atoms, then the equation becomes:

$$\frac{\partial^2 r_i}{\partial t^2} = m_i^{-1} F_i(t) + \gamma(t) r_i(t) \quad (25)$$

Here the friction term indicates the heat loss or gain due to the interaction with the thermal bath. This generates a smooth (continuous velocity) and deterministic trajectory. If $\gamma(-t) = -\gamma(t)$, also time reversibility is verified. The advantage of deterministic algorithms is the existence of conserved quantities and the reproducibility. The above equation can be integrated with a modification of Newton's equations integrators. In the leapfrog algorithms this is achieved by rescaling the velocity at every iteration:

$$\left(t + \frac{\Delta t}{2}\right) = \lambda(t; \Delta t) \left[\frac{\partial r_i}{\partial t} \left(t + \frac{\Delta t}{2}\right) + m_i^{-1} F_i(t) \Delta t \right] \quad (26)$$

which recovers the previous equation with an infinite timestep and $\lambda(t; 0) = 1$ with

$$\gamma(t) = \left(-\frac{\partial \lambda(t; \Delta t)}{\partial (\Delta t)} \right)_{\Delta t=0} \quad (27)$$

Such a form doesn't specify a unique form of the scaling factor since given one solution, any solution differing by a constant plus a term $o(\Delta t)$ It's possible to find a form of the type:

$$\lambda_i(t; \Delta t) = 1 - \gamma(t) \Delta t + \left[\frac{\gamma^2(t)}{2} + \gamma(t) F_i(t) + \frac{\gamma(t) F_i(t)}{2 m_i r'_i(t)} \right] (\Delta t)^2 \quad (28)$$

for which the algorithm preserves the accuracy of the NVE leapfrog algorithm. Integrating the Langevin equation of motion produces, at constant volume a trajectory in phase space constrained to a canonical distribution of microstates if the frictional coupling term is nonvanishing. The amplitude of temperature oscillation is independent of the friction but higher friction dictates a shorter time scale of the oscillation and a faster convergence to the equilibrium distribution if the system is outside the canonical equilibrium distribution. It can be proven that on an intermediate timescale:

$$\gamma = (2\zeta_T)^{-1}$$

where ζ_T is the temperature relaxation time, the time constant of temperature equilibration. [9]

The Berendsen thermostat

The simplest form of velocity scaling that imposes energy conservation is to rescale the velocity computed after each timestep by a factor

$$\lambda = \sqrt{T/T_0} \quad (29)$$

This doesn't allow the temperature to fluctuate and the system doesn't sample the canonical ensemble.

The Berendsen thermostat is a weak formulation of the approach: velocities are rescaled by a factor

$$dT(t)/dt = \frac{1}{\tau} (T_0 - T(t)) \quad (30)$$

where τ represents the strength of the coupling.

This produces over a timestep a change in temperature of :

$$\frac{\Delta t}{\tau}(T_0 - T(t)) \quad (31)$$

The scaling factor is:

$$\lambda^2 + \frac{\Delta t}{\tau} \left\{ \frac{T_0}{T\left(t - \frac{\Delta t}{2}\right)} - 1 \right\} \quad (32)$$

In the limit of $\tau = \Delta t$ the algorithm is simple velocity scaling. If $\tau \rightarrow \infty$ the thermostat is inactive and the temperature fluctuations reach the microcanonical ensemble limit. The thermostat however doesn't sample a canonical ensemble. [10]

The flying ice cube problem

One limit of rescaling thermostats is a simulation artifact where, for reasons not fully clear, the equipartition theorem is violated and higher frequency modes drain their energy into low-frequency mode, resulting in a system with a high center of mass linear momentum and frozen internal motion (a flying ice cube). [11]

The Nosé-Hoover thermostat

While the Berendsen thermostat relaxes well the system to the target temperature, for equilibrium simulation where ensemble averages are obtained, the Nosé-Hoover thermostat allows for correct sampling of the canonical ensemble. It is based on a representation of the heat bath by adding an artificial variable to the system, \tilde{s} associated with a mass Q and a velocity \tilde{s}' . It represents a time scaling parameter such that $d\tilde{t}' = \tilde{s}dt$. The resulting system has therefore a Lagrangian

$$\mathcal{L} = \sum_i \frac{m_i}{2} \tilde{s}^2 \dot{\tilde{r}}_i^2 - U(\tilde{r}) + \frac{1}{2} Q \tilde{s}'^2 - g k_b T_0 \ln(\tilde{s}) \quad (33)$$

With the first two terms representing the kinetic and potential energy of the real system, and the other two the additional kinetic and potential energy needed to guarantee that a canonical ensemble is produced. The obtained virtual phase space is sampled by the equations of motion from a microcanonical ensemble and the projection onto the real 6N dimensional phase space sample from a canonical ensemble since energy is exchanged with the additional variables. The smaller the Q the stronger the coupling. [10]

Thermo-barostats

Barostatting can be conceptually similar to thermostatting.

Scaling barostats bring the pressure to a predefined equilibrium by scaling all physical dimensions and thus the volume. The Berendsen barostat uses a scaling factor

$$\mu = \left[1 - \beta \frac{\Delta t}{\tau} \right]^{1/3} \quad (34)$$

where β is the compressibility, which doesn't need to be known since the ratio with the relaxation constant τ is what is defined a priori.

The Parrinello-Raman barostat operates with the same principle as the Nosé-Hoover thermostat, by allowing the exchange of volume with an extra dimension.[12]

Important concepts and criticalities

Choosing the time scale

One must pay caution to choose a proper time scale for molecular dynamics. For example if one wants vibrational stretching frequencies of water to be sampled adequately one should have a timescale of approximately 1 fs at most since the oscillation period is approximately 10 fs[13].

Moreover, a too large timestep can lead to numerical instability with some potentials. For example, if a spring force has been used as a potential to model a bond, and the timestep is greater than the oscillation period, overshoot will occur with the atoms involved crossing the energy minimum and ending up at a greater distance the following timestep. That will lead to an even greater force and thus an even greater overshoot at the following step. The consequence is diverging positions, forces and velocities.

On the other hand, a longer timestep will allow for longer simulation times for the same computation time. And some phenomena require long computation times, so a tradeoff between the fidelity of the model and the length of the simulation is needed.

It's however impossible for example to simulate phenomenons that occur on a timescale of milliseconds or greater in a reasonable computation time. Sometimes tricks may be used to reach the desired equilibrium in a shorter time scale.

Ergodicity

One aspect one must pay attention to in MD simulations is ergodicity, the ability of the system to sample over time all the parts of the configuration space in a uniform and random way. This is fundamental for ergodic sampling: if an ergodic simulation runs for long enough, the uniform exploration of the phase space means the time average of thermodynamic quantities approximate the ensemble averages.

In deterministic methods based on Hamilton's equation, below a certain energy threshold, a Hamiltonian system will display regular trajectories which are not ergodic and should not be used for ergodic sampling.

In general, deterministic MD integrators at constant energy can't move between regions of the phase space which are separate but not connected. Plus they can take a long time converging to equilibrium, for example when the system crosses a saddle point near a phase transition temperature. So particular care must be taken in choosing initial conditions to sample the required portion of the phase space instead of getting stuck near a local minimum.

Adding stochastic elements, like the stochastic force in the Langevin equation, removes the validity of short-term trajectory but may be useful to guarantee ergodic sampling at constant energy and thus better averages.[14]

Thermodynamic quantities and postprocessing

When one uses molecular dynamics to compute equilibrium properties, assuming ergodic sampling, time averages can allow determining ensemble averages.

Temperature

The temperature is allowed to fluctuate in all ensembles, with fluctuations depending on the specific of the thermostat. The instantaneous temperature is computed as

$$\sum_i \sum_{\alpha} \frac{1}{2} m_i v_{i,\alpha}^2 = \frac{3}{2} k_B T \quad (35)$$

Thus the arithmetic average of the instantaneous temperatures computed this way can be used to obtain an average temperature that's consistent with the equipartition principle.

Pressure

The pressure can be computed from the virial equation:

$$p = \frac{2}{3V} \langle E_{kin} \rangle + \frac{1}{3V} \sum_{n=1}^N \sum_{m=n+1}^N (r_n - r_m) * F_{nm} \quad (36)$$

[15]

Force Fields

A force field is the set of equations that determine the potential acting in the system and thus the force on the individual particles. Generally speaking force field terms are of two types: bonded interactions and nonbonded interactions.

While the distinction is dependent on how the software package handles bonding internally, nonbonded interactions occur between atoms belonging to the same or different molecules (e.g. the coulombic and Lennard-Jones potential (see eq. 4)) and can be fully attractive, repulsive or both. They become negligible (or are neglected by setting a cutoff) at long range.

Bonded interactions are terms that are meant to constrain specific pairs of atoms around equilibrium positions and often bonded atoms getting too far away or leaving the simulation domain is considered an error.

Two body bond potentials

They are meant to give simple approximations of the bond near the energy minimum, so the most practical choice is generally the harmonic bond:

$$U_{har}^{ij} = \frac{K}{2} (|r_j - r_i| - r_0)^2 \quad (37)$$

Here the fitting parameters are a spring constant K and the equilibrium distance r_0 . For biomolecules, K is generally on the order of a few hundred kJ/mol to allow a timestep

of the order of the femtosecond while preserving a reasonable stiffness. Accurate values are needed whenever you need to reasonably approximate the vibrational energy of the system, otherwise a rough approximation can be enough.

Three-body angle potential

Generally, whenever three molecules are linked by two consecutive bonds, the angle between the two bond axes also has an equilibrium value, and constraints are set so that the energy increases when deviating from that equilibrium. Typical examples are harmonic and cosine/harmonic potentials. The former has the same form of the harmonic bond but with angles instead of distances, while the latter has the form:

$$U_{cos}^{ijk} = \frac{K}{2}(\cos(\theta_{ijk} - \theta_0))^2 \quad (38)$$

and has the advantage of guaranteeing continuity over a 2π rotation and having energy minima centered on both complementary angles.

Four body dihedral angles

Given three consecutive bonds it's also possible to set the angle between the planes on which the two angles defined by these bonds lie. Such an angle, called dihedral angle, tends to be more weakly constrained. The chosen potential must be periodic over 2π and transitions smoothly when the angle rotates by more than one period. Thus the general form of an empirical dihedral potential is a truncated Fourier series:

$$U(\phi) = a_0 + \sum_{i=1}^n [a_i * \cos(n\phi) + b_i * \sin(n\phi)] \quad (39)$$

Another form for that is:

$$U(\phi) = K_0 + \sum_{i=1}^n K_i * [1 + \cos(n\phi)] \quad (40)$$

Many other dihedral potentials can be cast into a Fourier series by applying trigonometric formulas. One of those is the cosine power series:

$$U(\phi) = \sum_{i=0}^n k_i * \cos(\phi)^i \quad (41)$$

.

Bond order and reactive force fields

A force field can have a more complex formulation including bond order: the strength of a bond between two atoms depends in this case not only on the relative position of atoms, but also on their coordination, and update rules can be set to modify the strength and number of bonds on an atom when its coordination or that of its neighbor changes. This allows simulating reactions: given for example two separated ethene molecules with

a bond order based force fields for hydrocarbons, initially, carbons have three neighbors (two hydrogens and another carbon) so the C-C bond in each molecule is set to bond order 2 and the model for the carbon double bond is used to compute the forces. If a pair of carbons from two different molecules gets within the cutoff length for the bond order computations, then the coordination of each of these carbons changes to four and the update rules take care of establishing the bonds for the new carbon chain.

An example of bond-order based force field is the reaxFF force field, based on a theory of bond order that uses distance to compute the coordination of an atom, creates bonds between neighbors based on the estimated bond order for each pair and applies energy corrections to penalize under and overcoordination.

Current reaxFF models can describe systems containing few atom types. For example, a reaxFF force field exists for systems containing only Au, S, C and H.[\[16\]](#)

Software

LAMMPS, an open-source molecular dynamics package

The main software used in this thesis work is Lammmps (available at <http://lammmps.sandia.gov>) [17]. Lammmps is a classical MD open-source software written in C++. It parses an input file line by line and executes sequentially the instructions provided by calling the relevant subroutines.

This software package has several strong points. One is the flexibility in the definition of the system (from atomic to solid-state, polymeric and coarse-grained system and from a few atoms to larger systems with millions of atoms) and the force field (even supporting user-defined potentials through interpolation of the data from a lookup table). Moreover, it is optimized for running on parallel machines. It runs spatial decomposition by partitioning the simulation box and assigning each subdomain to a processor then running parallel computations using neighbour lists to determine in each subdomain the atoms outside that subdomain that need to be accounted for to compute forces. Dynamic load balancing can be enabled to ensure that the system is partitioned so that all processors run on a similar number of atoms, in order to reduce idle time.

One core element of the Lammmps code is the use of the above-mentioned neighbor lists to determine the pairs of atoms for force computation. The software marks as neighbors pairs of atoms at a distance within the cutoff of their pairwise interactions plus an extra "skin distance". This neighbor list is updated at a user-defined frequency. If the update is done less frequently it's good practice to specify a larger cutoff distance for the neighbor list construction by increasing the "skin distance". This is done in order to avoid losing force field terms in the steps where the list is not updated.

The downside of Lammmps is the reliance on external tools for building complex initial configurations and visualizing simulation outcomes. However, the code is available as a python library. Thus one can include Lammmps in a python interface.

Visualizing LAMMPS trajectories: the VMD software package

VMD (available at <https://www.ks.uiuc.edu/Research/vmd/>)[18] is used in this dissertation to produce the visualization of the trajectory files. VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting. In this work, it has been mainly used to load the trajectory files obtained by Lammmps and create a 3D visual representation. It also allows for scripting to modify the visualization and add or remove atoms from the simulation.

ORCA, a quantum chemistry software package

ORCA[19][20] is a fully parallelized quantum-chemical software package featuring a range of methods from semiempirical ones to density functional theory-based ones. With these methods, the ground state energy of a molecule or system can be computed from a list of atom types and coordinates of said molecule or system.

Part 1: hexanethiol deposition on gold on Lammmps

This experiment is an attempt to simulate the formation of a self-assembled monolayer of 1-hexanethiol molecules on gold. Self-assembly is a process where molecules arrange themselves in a highly ordered structure thanks to their chemical properties. If that ordered structure is composed of chemisorbed molecules on a substrate, a self-assembled monolayer (SAM) is formed.

the molecules used in a SAM process generally are composed of three different parts:

- a headgroup chemically interacting with the substrate and binding the molecule to it. It must be strong enough to prevent molecules from desorbing but at the same time allow for translation parallel to the substrate to allow rearrangement
- a backbone dictating how the molecules are oriented, their interaction and properties such as the monolayer thickness and conductivity
- a tailgroup representing the outer surface of the monolayer and providing the necessary functionalities (electrical properties, catalysis, binding sites for other molecules among the others)

The specific class of molecules treated in this work are based on a 1-hexanethiol molecule as the binding group and a backbone with a functional group substituting one of the terminal hydrogens. The 1-hexanethiol has the following formula: $\text{CH}_3(\text{CH}_2)_4\text{CH}_2\text{SH}$ and the substitution occurs on a hydrogen in the CH_3 end unit.

A typical technique for SAM consists in immersing the substrate in a solution containing the molecules to be deposited. The process occurs in several steps: initially the molecules diffuse to the substrate and the headgroups start reacting to form a bond with it. In this case, the molecules might be lying horizontally on the substrate if the backbone nonbonding interaction is strong enough. After the surface concentration reaches a critical value, islands start to form where the contribution of adatom-adatom nonbonded interactions can prevail on the substrate-adatom ones, allowing the molecules to stand upright.

Since the process described above happens in a time window of hours, in this work a simplified model system is studied, depositing the SAM molecules from vacuum with the hydrogen in the thiol group already removed. That enables the formation of a SAM in the nanosecond time scale typical of molecular dynamics simulations. It also contributes to keeping the complexity of the simulation manageable.

The first molecule of which deposition has been attempted is the simple 1-hexanethiol molecule.

Simulation with reaxFF force field

A first attempt to simulate a deposition has been using the reaxFF force field[16] with parameters taken from Bae and Aikens[22].

Larger substrate

First, a gold (111) substrate consisting of three atomic layers with dimension (5.99 Å, 5.19 Å) is created. The lattice is FCC with a constant of 4.08 Å.

Then a layer of 1-hexanethiols with the SH group hydrogens removed and with size, (3.33 Å, 2.89 Å) is positioned so as to have 0.5 nm of distance between the substrate and the sulfur. A total of 37 molecules are generated with twice the lattice spacing compared to the gold.

After energy minimization, 20 ps of dynamics are run with a timestep of 0.1 ps and using the velocity-Verlet integrator with the Nosé-hoover thermostat. The temperature is ramped up from 0 K to 300 K over the first 10 ps.

As seen in figure 32, over the course of the simulation, the molecules spread over the substrate, lying horizontally without an evident chemical interaction between sulfur and gold.

The simulation runs slowly due to the complexity of the force field. Thus it is unsuited to simulate larger systems and time scales of the order of nanoseconds in a reasonable simulation time.

Smaller substrate

In this second experiment, the total size of the substrate is (2.66 Å, 5.77 Å). 45 thiol molecules are generated in 3 layers above the substrate and dynamics is run with the same parameters as in the previous experiment. After 20 ps part of the molecules are attached to the substrate and part have flown off. As seen in figure 63, some of the sulfurs manage to bind the gold and the molecules lie horizontally, however sulfur-sulfur bonds between different thiols occur because removing the hydrogen leaves the sulfur end of the thiol able to react with other sulfur to satisfy its coordination number. While the timescale is too short to show if a SAM forms if the system is given time to equilibrate, the impossibility to make the unsaturated sulfur to be selectively reactive only toward gold is another limit to using the reaxFF force field for thiol SAM formation.

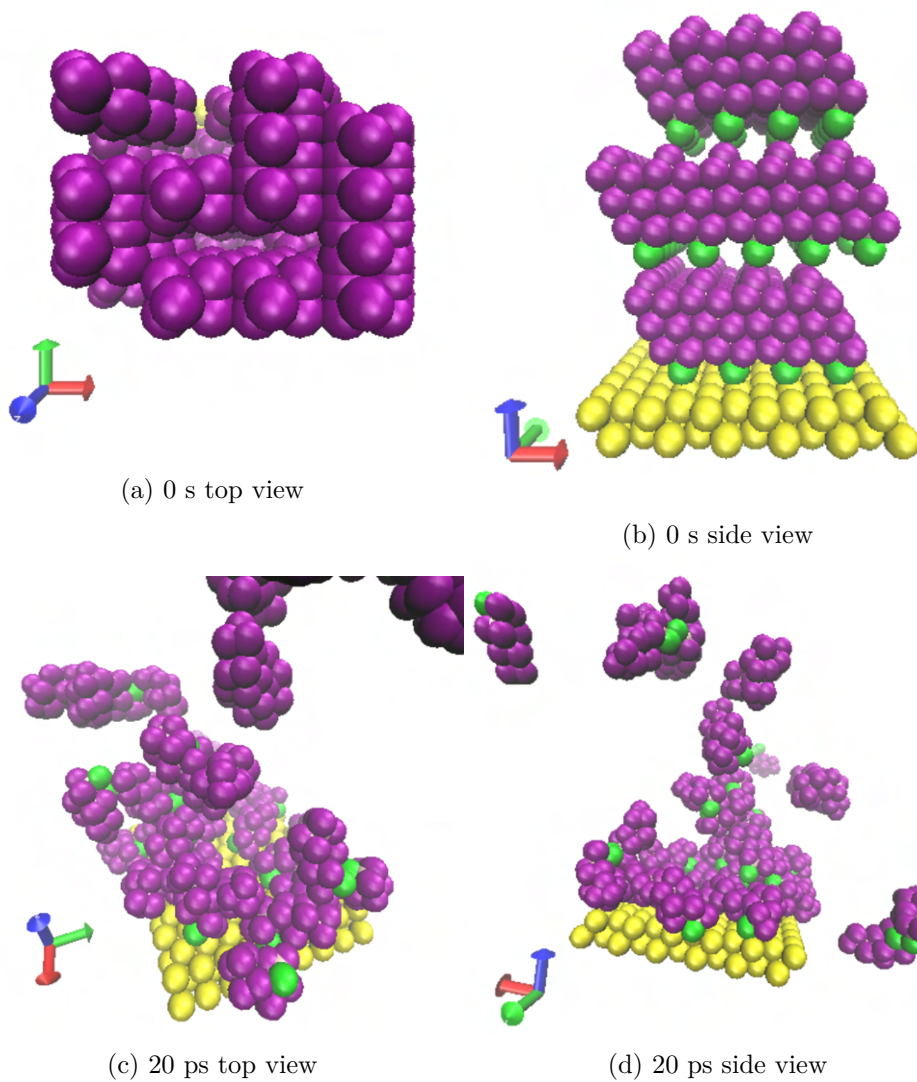
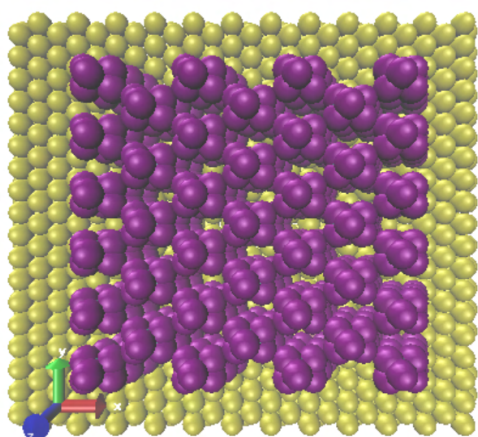
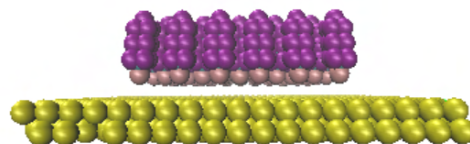


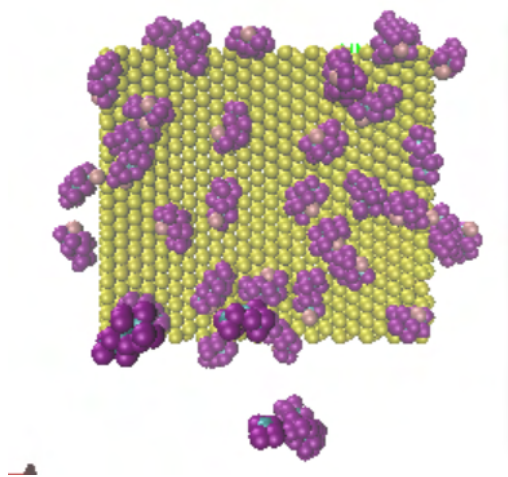
Figure 6: Deposition of 37 thiols on a large gold surface



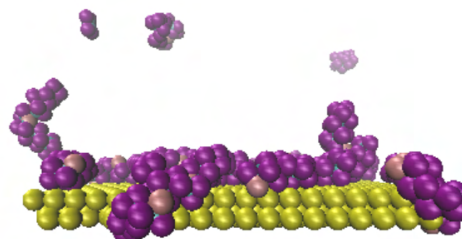
(a) 0 ps top view



(b) 0 ps side view



(c) 20 ps top view



(d) 20 ps side view

Figure 7: Deposition of 45 thiols on a small gold substrate

Simulation with a united atom force field

Force field parameters

Due to the unfeasibility of modeling the thiols with a reaxFF forcefield, a simpler united atom force field is adopted, where the hydrogens are implicit and lumped together with the carbon they are attached to. Thus the CH₃, CH₂ and SH groups were represented each by a single lumped element. For the force field all bonded and nonbonded interactions are listed individually.

The procedure is derived from H Kim et al. [23] A Lennard Jones potential with parameters taken from Saha et. al.[26] and listed in table 1 is adopted for nonbonded interactions. Lorentz-Berthelot[24][25] combination rules are used to compute the nonbonded interaction between different groups The S-Au bond is represented by a nonbonded Morse potential.

$$D_0 \left[e^{-2\alpha(r-r_0)} - 2e^{\alpha(r-r_0)} \right] \quad (42)$$

with $D_0 = 8.763$ kcal/mol, $\alpha = 1.47 \text{ \AA}^{-1}$ and $r_0 = 2.65 \text{ \AA}$ [27]. Both potentials are truncated with a 1.2 nm cutoff. For the bonded potential, adjacent atoms are connected by harmonic bond potentials, with parameters listed in table 2. All the angles are represented by the harmonic cosine potential in eq.(38) with $K = 60$ kcal/mol and $\theta_0 = 109.5^\circ$ and all the dihedral by the cosine power series in eq.(41) with $n = 4$ and coefficients 1.57, -4.05, 0.86 and 6.48 kcal/mol. Data for the bonded potential is taken from Hagita et al.[28]

groups	$\epsilon(kcal/mol)$	$\sigma(\text{\AA})$
CH ₃	0.175	3.905
CH ₂	0.118	3.905
S	0.395	4.250
Au	0.0390	2.935

Table 1: Wan der Waals potential parameters

groups	$K(kcal/mol)$	$r_0(\text{\AA})$
CH ₃ - CH ₂	350	1.54
CH ₂ - CH ₂	350	1.54
CH ₂ - S	350	1.82

Table 2: Bond harmonic potential parameters

Simulation

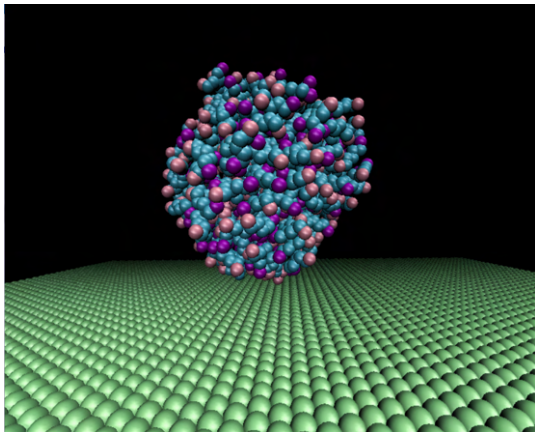
First, a thiol droplet is generated by packing 300 thiols on a cubic box and running energy minimization first, then velocity-Verlet dynamics with a Nosé-Hoover thermostat. The temperature is ramped up from 0 K to 300 K over 2 ns, with a timestep of 1 fs, then constant temperature dynamics is run at 300k for 3 ns with a timestep of 1 fs.

The equilibrated droplet is then placed at less than 1 nm from a [1,1,1] gold surface,

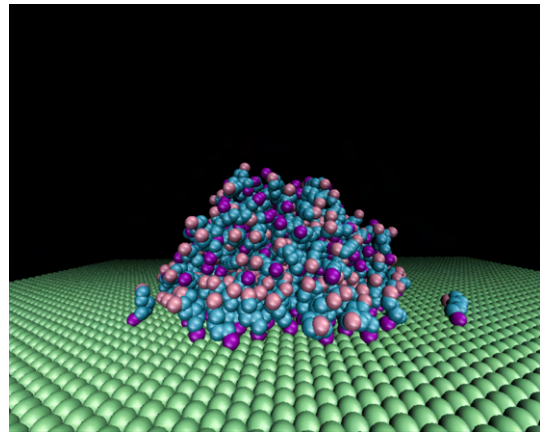
composed by a double layer of gold of size 12 nm*10.4 nm. The initial distance is chosen to be smaller than the force cutoff for the droplet to feel the interaction with the substrate. The substrate is frozen in position. After energy minimization, Velocity Verlet dynamics is run with a Nosé-Hoover thermostat for 80 ns.

- at 250 ps the droplet is already spreading, forming a half dome on the gold substrate.
- at 1 ns the thiols at the bottom of the droplet start standing in an upright position, with the sulfur binding the substrate. A thiol island starts to form.
- at 2 ns part of the thiols on the top slide down, reaching the surface. This, combined with the displacement of the thiols already attached, spreads the thiol island.
- at 5 ns few thiols remain on the top of the thiol island, which forms a single cluster.
- at 12.5 ns all the thiols are deposited on the substrate.
- over the remaining time the only change is that the island becomes rounder, minimizing energy. One single thiol remains upside-down, with the CH₃ end facing the substrate.

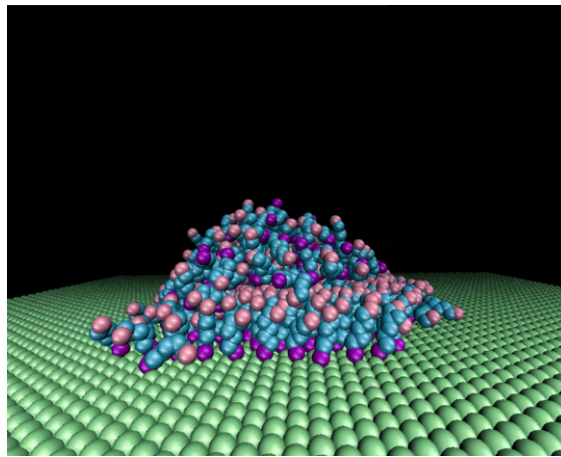
The approach described succeeds in producing a molecular self-assembly of thiols on a gold substrate.



(a) 0 s

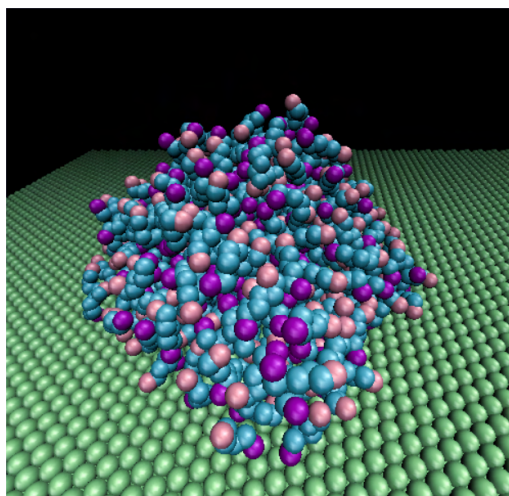


(b) 250 ps

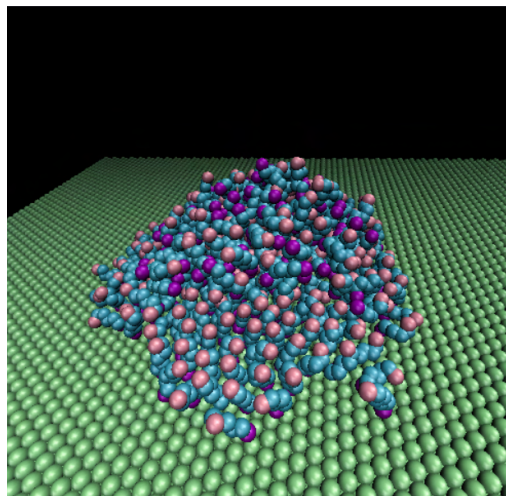


(c) 1 ns

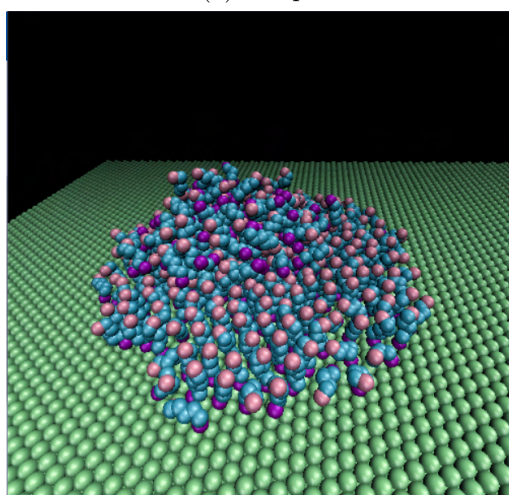
Figure 8: Side view of the droplet reaching the substrate



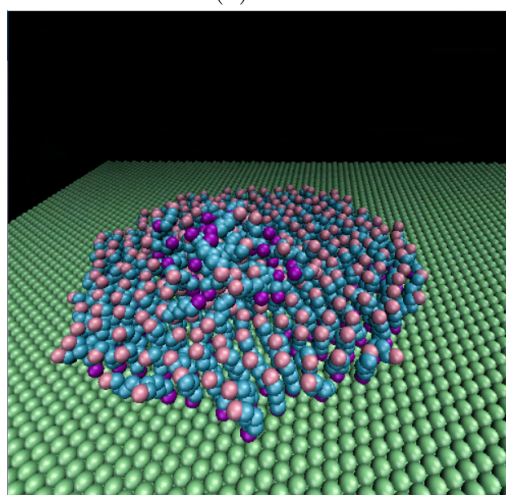
(a) 250 ps



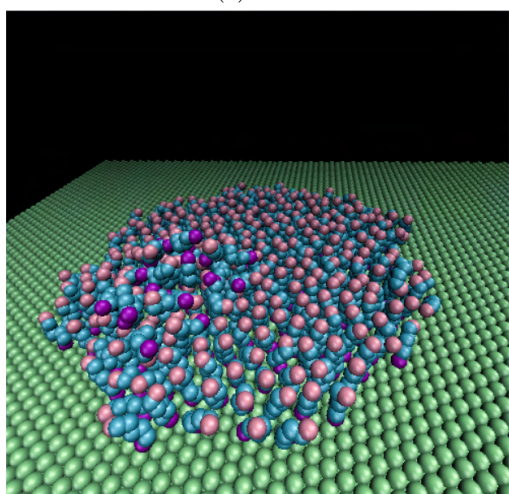
(b) 1 ns



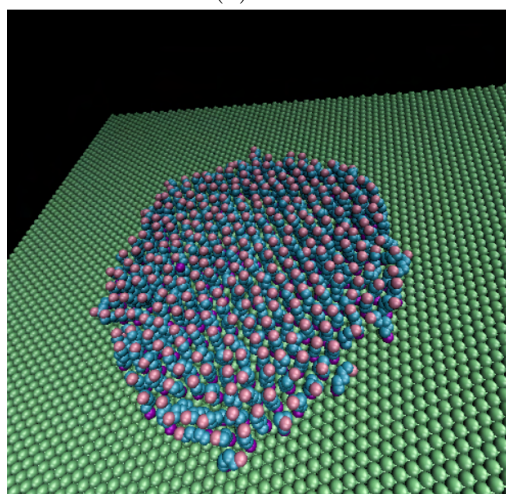
(c) 2 ns



(d) 5 ns



(e) 10 ns



(f) 12.5 ns

Figure 9: Close-ups of the SAM formation

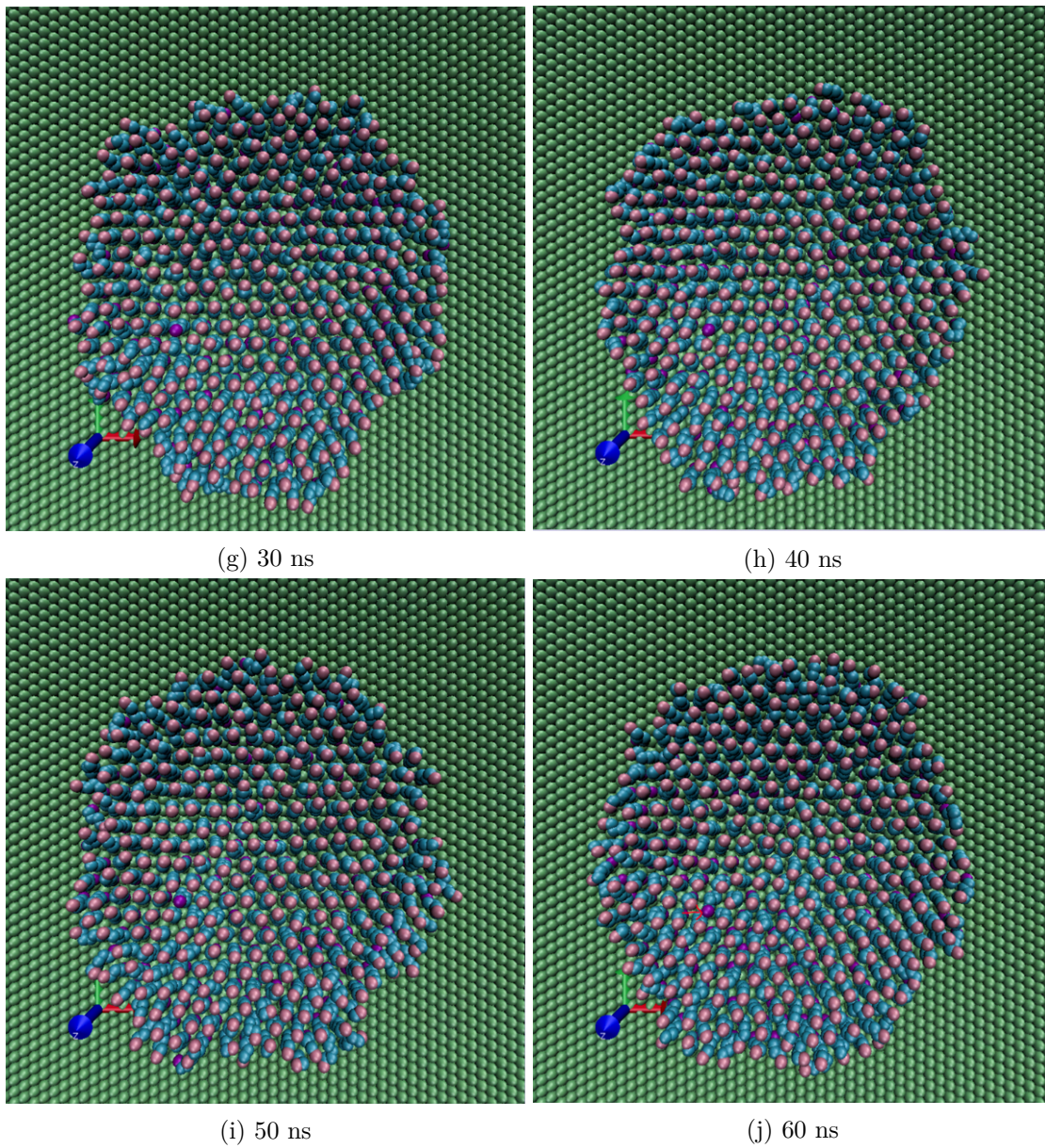


Figure 9: Top view of the SAM formation

Conclusions of part 1

The outcome of the experiments shows that, while a reactive force field like reaxff is unsuccessful in properly representing the formation of thiol SAMs, a united atom model, using a more simplified representation with implicit hydrogens, can instead generate a molecular SAM of thiols. Figure 10 show quantitative data from the cumulated analysis of the thiols in 100 timesteps at 50 ps intervals. Observing the first three nearest neighbor distances, they are compatible with a distance of $2.885 \text{ \AA} * \text{sqrt}3 = 4.996 \text{ \AA}$ ($\text{sqrt}3$ multiplied by the surface lattice vector), with slight differences that can be justified by fluctuations. This is compatible with a surface lattice measuring $\sqrt{3} \times \sqrt{3} R30^\circ$. The chain tilt angles show values almost totally lower than 30° , with a prevalence of values close to 0° . Small variations in the tilt angles also mean that the molecular height values with respect to the substrate are narrowly distributed around an average of approximately 9 \AA . The frames displayed in figure 9 show how the border is where the outliers are generally located, with the notable exception of an upside-down molecule.

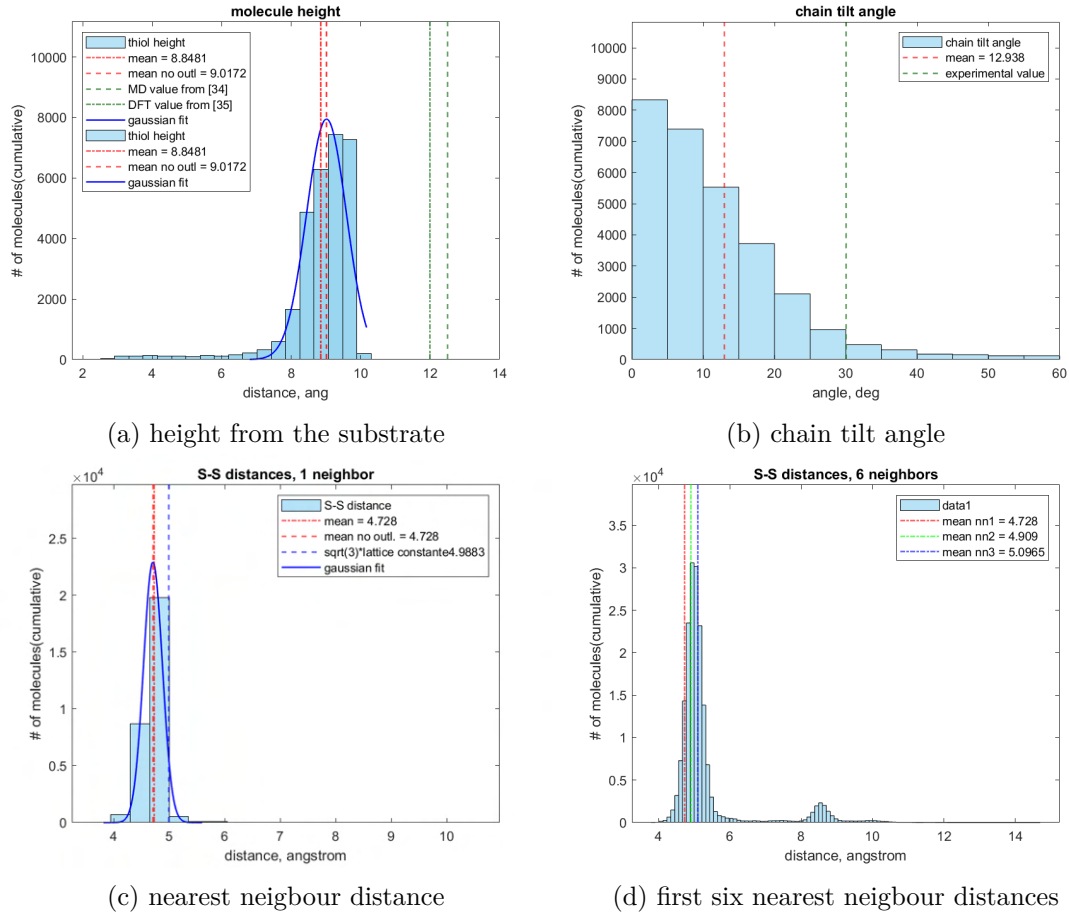


Figure 10: Thiol sam: quantitative analyses, cumulative data from 100 frames at equilibrium at 50 ps intervals

Part 2 : deposition of 6-(Ferrocenyl)hexanethiol on a surface

The 6-(Ferrocenyl)hexanethiol (FcC6S) molecule is composed of a 1-hexanethiol unit where a hydrogen in the tail carbon has been substituted by a ferrocenyl group, a radical obtained removing a hydrogen from a ferrocene molecule.

The ferrocene molecule is composed of two cyclopentadiene (Cp) rings bound to a central carbon. The bond is a form of fivefold coordination, typical of organometallic compounds, named η^5 coordination. A localized representation of the bond attributes it to a sigma bond, replacing one of the hydrogens from the carbon without double bonds, and the pi electrons involved in the formation of the Cp rings. However this results in five configurations that are equivalent by rotational symmetry and, as quantum chemistry studies confirm, the effective bond is instead their resonance hybrid, symmetric by rotation and delocalized over the rings.[30] In the current experiments, the ferrocene is functionalized with a hexanethiol anchoring group in order to form a SAM.

There are two goals:

- firstly, by comparing with theoretical values the geometrical data from different experiments, with both anchored and mobile thiols, the reliability of MD as a tool for investigating the properties of a FcC6S SAM is studied.
- secondly, the FcC6S molecules are deposited on wires of different thickness and their behavior after oxidation and application of an electric field is studied with a first-order simulation with fixed charge values not accounting for polarizability.

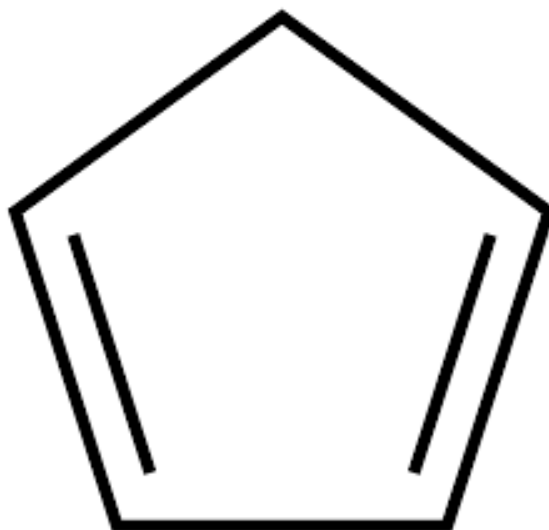


Figure 11: a cyclopentadiene ring[30]

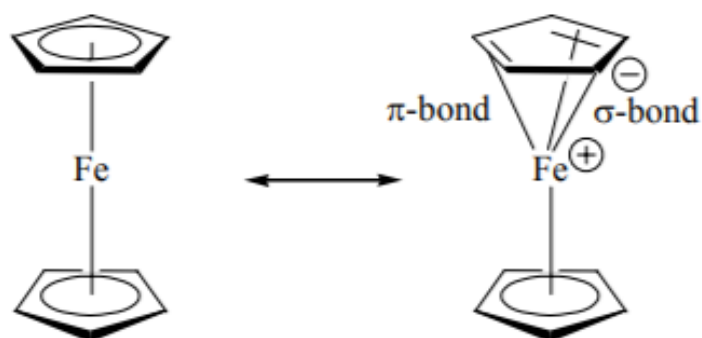


Figure 12: Structure of the ferrocene and resonance forms of the bond[29]

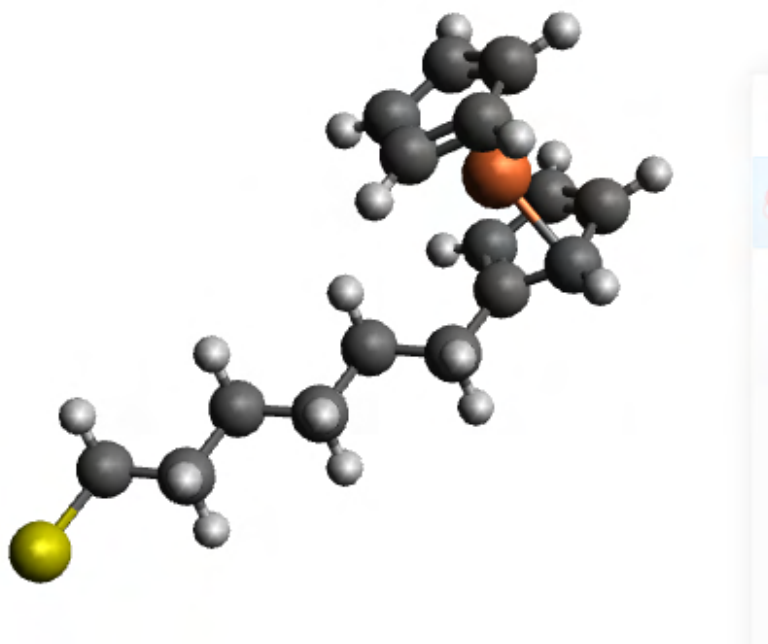


Figure 13: A 3D model of the 6-Ferrocenyl-1-Hexanethiol. The colors are yellow for sulfur, black for carbon, white for hydrogen and orange for iron. The single bond shown between iron and a Cp ring is an artifact from the visualization software Avogadro. (structure courtesy of VLSI lab in Politecnico di Torino)

Modeling the bonds

Two models have been studied for the force field parameters of the free thiols. The second one has been adopted for the anchored thiols too, varying the Au-S interaction in order to examine multiple anchoring configurations.

Model 1

In the first model the following has been done: For the thiol chains the force field terms are the same as in part 1: the hydrogens of the thiol chain have been lumped with the carbon atom and those united atom elements are referred to as CH_2 . The head is treated as a rigid body when dynamic is run, since its internal vibrational modes are assumed to have little relevance to the SAM formation. However, force field terms are included for energy minimization (since the algorithm doesn't account for rigid bodies) and the head geometry is constructed by placing a dummy atom in the position of the center of the Cp ring and linking it to both the iron and the carbons in the ring.

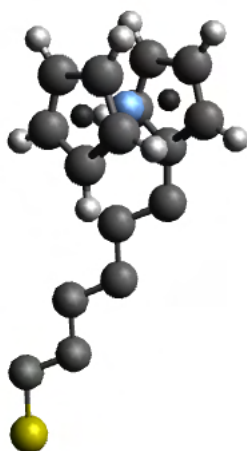


Figure 14: 3D model of the ferrocene molecules adapted from the precedent one with the centroids and implicit hydrogens in the chain

In the following, Cp will refer to the ring centroid, C to the ring carbons and CH_2 to the lumped elements including carbons and hydrogens in the thiol tail.

- The nonbonded interactions of choice are the Lennard-Jones potential with Lorentz Berthelot combination rules plus the Coulomb interaction using charge values on the atoms computed via Orca.
- The interactions between atoms of the same elements are reported in table 3.
- The values for C and H are taken from the AutoDock user guide on the Richards Center at Yale University website.[\[31\]](#)
- The additional bonds with respect to the simple thiol are harmonic bonds, with the values in table 4.
- The angle is again the squared cosine angle, with values in table 5. The value of 126° for the $\text{CH}_2\text{-C-C}$ angle is taken from Lopes et al.[\[32\]](#).

- The dihedrals in the Cp head are Fourier dihedrals truncated at the first order, with the values in table 6.
- The CH₂-CH₂-CH₂-CH₂ and S-CH₂-CH₂-CH₂ dihedrals are modeled with cosine power series in eq.(41) with n = 4 and coefficients 1.57, -4.05, 0.86 and 6.48 (kcal/mol).
- For the CH₂-CH₂-C-C dihedral representing the rotation angle of the head compared to the thiol chain, no data is available to the author's knowledge, so the potential is computed with ORCA, loaded in a lookup table and interpolated with the built-in spline interpolation tool in LAMMPS.
- The CH₂-C-C-C is modeled with an OPLS dihedral [37], as indicated by Lopes et al.[32]

$$U(\theta) = \frac{1}{2} [K_1(1 + \cos(\theta)) + K_2(1 - \cos(2\theta)) + K_3(1 + \cos(3\theta)) + K_4(1 + \cos(4\theta))] \quad (43)$$

groups	$\epsilon(kcal/mol)$	$\sigma(\text{\AA})$
CH ₂	0.118	3.905
S	0.395	4.25
H	0.02	2
C	0.07	3.55
Fe	0.4816	3.11
Cp	0	0
Au	0.0390	2.935

Table 3: Lennard_Jones force field parameters for the ferrocene head

groups	$K(kcal/mol)$	$r_0(\text{\AA})$
CH ₂ - CH ₂	350	1.54
CH ₂ - S	350	1.82
C - C	400	1.424
C - H	375	1.08
C - Cp	350	1.211
Cp - Fe	205	1.649

Table 4: Bond harmonic potential parameters

groups	$K(kcal/mol)$	$\phi_0(deg)$
CH ₂ - CH ₂ - CH ₂	60	109,5
S - CH ₂ - CH ₂	60	109.5
CH ₂ - CH ₂ - C	60	109.5
C - C - C	100	108
C - C - H	30	126
Cp - Fe - Cp	80	180
C - Cp - Fe	100	90
CH ₂ - C - C	140	126

Table 5: angle cosine/harmonic parameters

groups	n	$K_0(kcal/mol)$	$K_1(kcal/mol)$	$\phi_0(deg)$
C-C-C-C	1	6	2	180
H-C-C-H	1	1.5	2	180
H-C-C-C S	1	4	2	180
C-Cp-Cp-C	1	0.36	5	180

Table 6: dihedral Fourier parameters

Model 2

A second model for the molecule combines the parameters for the thiol part of the FcC6 molecule from Rai et al.[36] with the bonds and angles for the ferrocene head from model 1 and the dihedrals from Lopes et al.[32] with the addition of the CH₂ - CH₂ - C - C dihedral computed with ORCA. While bonds are always harmonic and angles are cosine/harmonic, the dihedral used in this context both for the chain and for the head is the OPLS dihedral in eq. 43. The parameters are in the tables from 7 to 10

groups	$\epsilon(kcal/mol)$	$\sigma(\text{\AA})$
CH ₂	0.0935	4.411
S	0.2506	3.987
H	0.0302	2.42
C	0.105	3.57
Fe	0.484	3.11
Cp	0	0
Au	0.0390	3.293

Table 7: Lennard_Jones force field parameters for the ferrocene head

groups	$K(kcal/mol)$	$r_0(\text{\AA})$
CH ₂ - CH ₂	900	1.54
CH ₂ - S	350	1.82
C - C	1000	1.440
C - H	1000	1.08
C - Cp	350	1.211
Cp - Fe	205	1.649

Table 8: harmonic bond potential parameters

groups	$K(kcal/mol)$	$\phi_0(deg)$
CH ₂ - CH ₂ - CH ₂	124	114
S - CH ₂ - CH ₂	60	109.5
C - Cp - C	100	72
CH ₂ - CH ₂ - C	124	114
C - C - C	100	108
C - C - H	30	126
Cp - Fe - Cp	80	180
C - Cp - Fe	100	90
CH ₂ - C - C	140	126

Table 9: cosine/harmonic angle potential parameters

groups	n	$K_0(kcal/mol)$	$K_1(kcal/mol)$	$\phi_0(deg)$
S-CH ₂ -CH ₂ -CH ₂ 1.412	-0.2712	3.147	0	
C-C-C-C	0	10.79	0	0
H-C-C-H	0	10.79	0	0
H-C-C-C S	0	10.79	0	0
C-Cp-Cp-C	0	10.79	0	0
CH ₂ -C-C-C 0	10.79	0	0	

Table 10: OPLS dihedral potential parameters

computing a dihedral angle with Orca

Using the Orca software package, one can compute the ground state energy of a system over a range different values of a generalized coordinate. Said coordinate can be a bond length, an angle or a dihedral angle. This allows to obtain a coordinate vs energy curve that can be interpolated to find the stable equilibria of the system with respect to that parameter and implement the bond, angle or dihedral model in the force field of a molecular dynamics simulation.

In this work, the computed dihedral angle is the $\text{CH}_2\text{-CH}_2\text{-C-C}$ dihedral between the thiol chain and the ferrocene head, not available in the literature to the best of the author’s knowledge.

To compute it a DFT calculation is run with the B3LYP hybrid functional[38] for the restricted Kohn-Sham equations[39] and 58 equispaced points are acquired over a 360° rotation of the angle. Finally those points are interpolated.

A key requirement to obtaining a physical force field is that the resulting function has to be periodic with period 2π and with continuous first derivative at 0 and 2π . This ensures consistent left and right limits over a 2π rotation of the bonds. This can be achieved with the FFT interpolation method.

This method allows to interpolate periodic functions and it consists in transforming the original data to the Fourier domain through the FFT algorithm and then reconstructing the function having as Fourier coefficient the coefficients computed through the FFT algorithm. That function interpolates the original data points and is periodic with continuous first derivative, being a sum of Fourier components that are themselves periodic and continuous with 2π periodicity.

The interpolation is exact at infinite numerical precision if the original function is sampled above the Nyquist rate (the reciprocal of the spacing between samples has to be twice the frequency of the highest frequency Fourier component of the original function). In practice, if the original function is not bandlimited, it will have frequency components that are outside the spectrum of the interpolant. These frequency components are the source of the interpolation error. However, due to the decay of Fourier coefficients, in practice, these higher-order components are negligible if enough samples are taken. After the interpolation of the original 58 data points, the interpolant function is evaluated at 1083 different points spaced between 0 and 2π . The resulting data are loaded in a lookup table and a continuous potential function is reconstructed by Lammmps through spline interpolation. [34][35]

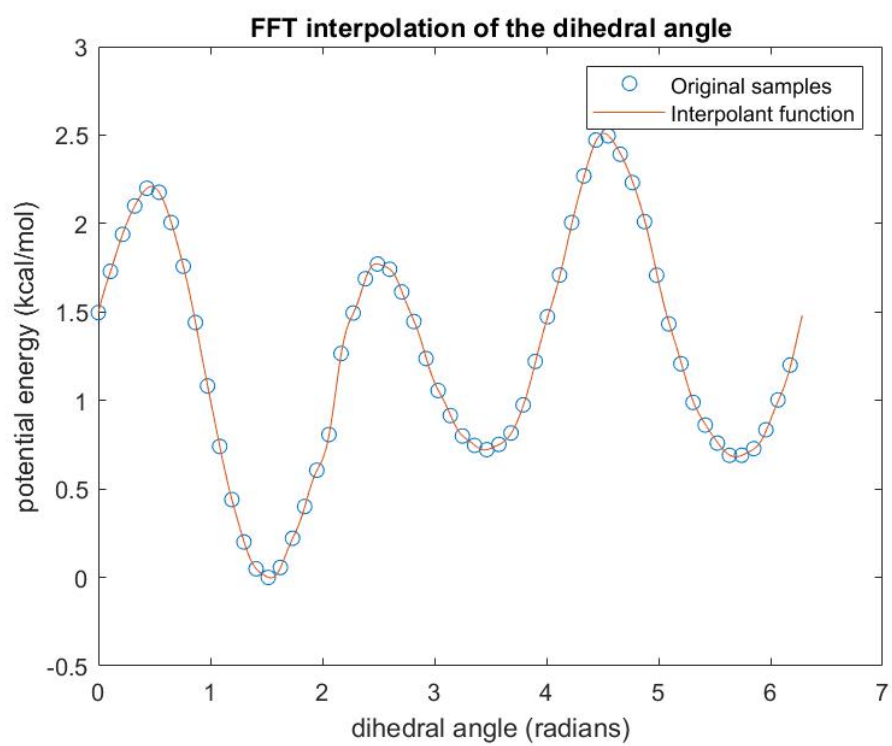


Figure 15: FFT interpolation

Deposition of a 6-(Ferrocenyl)hexanethiol droplet on a gold surface

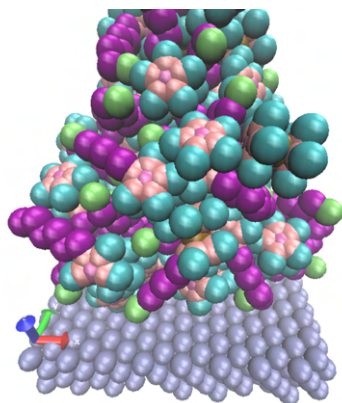
Similarly to the 1-hexanethiol case, a droplet of 50 FcC6S molecules is created. First, the molecules are generated on a cubic grid. Then, after running energy minimization, the thiol head is set as a rigid body and the droplet is equilibrated at 300K for 1ns after ramping up the temperature for 20 ps from 0K to 300K, using a velocity-Verlet integrator with Nosé-Hoover thermostating.

The droplet is then suspended 0.5 nm over a 4.00 Å by 3.46 Å gold surface. The surface is frozen so that the surface atoms are fixed in position. While this prevents the occurrence of second-order effects like the reconstruction of the surface layer due to the strength of the S-Au bond, those effects would only be reproduced correctly through a full quantum chemical treatment and the presence of the bulk of gold. The droplet is then drawn to the surface due to intermolecular forces. Similarly to the simple thiol case, the sulfur heads bind the surface, the droplet starts spreading and, thanks to the smaller number of atoms, it distributes over the surface quickly, already covering it at 1nm. This is true for both the force field models. Figures 16 and 17 show frames from the trajectory of the deposition with the first model, but the second one is very similar.

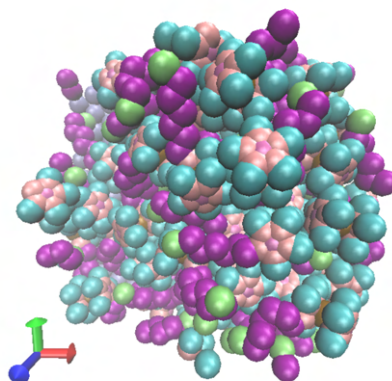
Since the theoretical maximum coverage for FcC6S on gold is $4.5\text{e-}10 \text{ mol/cm}^2$, not all molecules bind to the surface. Indeed after cleaning the unbound molecules and rerunning NVT dynamics at 300k for 5 ns to equilibrate the systems, 38 molecules bind the substrate in the first model and 30 in the second. The first leads to a higher density than the theoretical one, $4.68\text{e-}10 \text{ mol/cm}^2$, likely due to the presence of molecule heads stuck below the monolayer, in contact with the substrate, which can be observed both from the qualitative images and the quantitative data shown in section . With the second the density is $3.84\text{e-}10 \text{ mol/cm}^2$, which corresponds to a looser packing.

What is clear is that the packing is limited by the hindrance of the head. While thiol chains are free to arrange in a $\sqrt{3}x\sqrt{3}R30^\circ$ superlattice, with significant chain-chain interactions, which correspond to a distance of 4.99 Å between anchoring points. FcC6S molecules show an average first neighbor distance of 4.75Å. However, while the heads are packed tightly, the chains are free to move and so more closely packed molecules and hollows can be both seen.

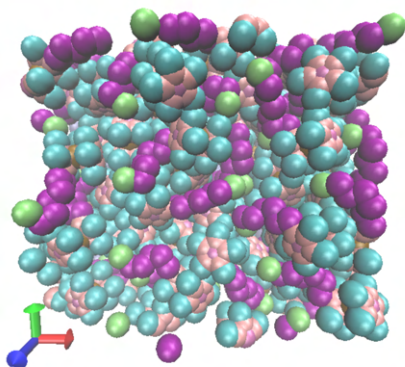
The hindrance of the thiol head might limit the interactions of the chains, preventing them from attaining a regular arrangement side by side. This might explain the strong variability in the geometrical positions of the bound thiols as shown in section . In fact both trans and gauche isomers can be observed and the arrangement of the thiol chains is highly irregular. Constraining the anchoring angle may allow the formation of a greater number of trans isomers and, as a consequence, molecules lying at higher tilt angles.



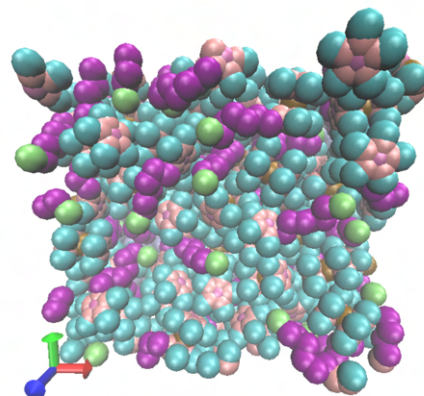
(a) 0 s top view



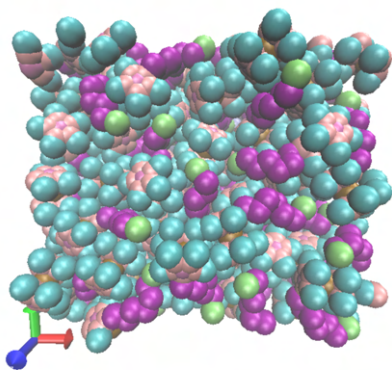
(b) 250 ps top view



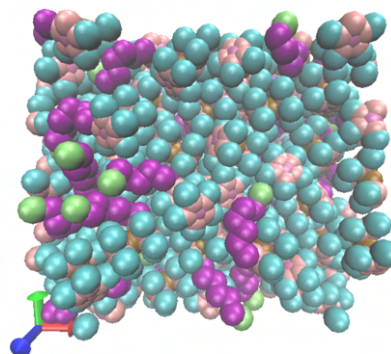
(c) 1 ns top view



(d) 2.5 ns top view

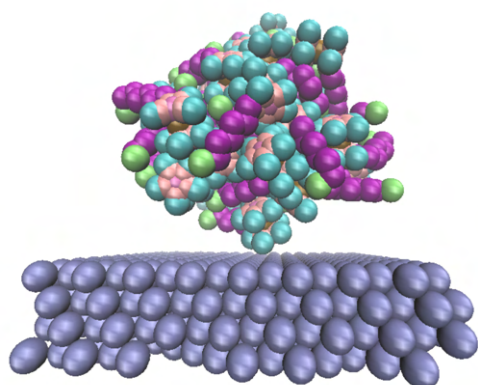


(e) 5 ns top view

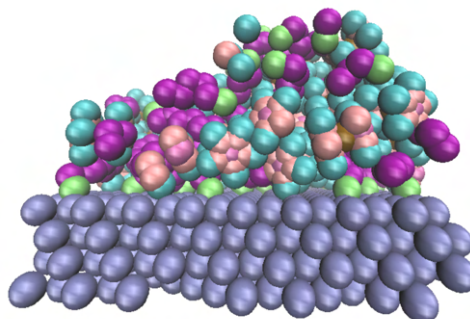


(f) 20 ns top view

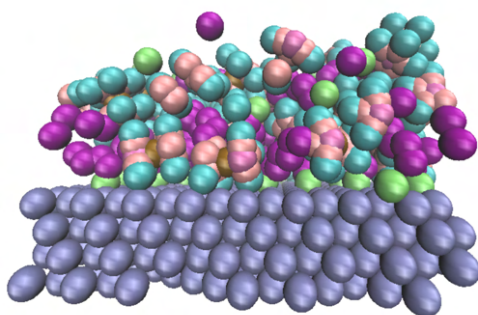
Figure 16: Deposition of 50 6(ferrocenyl)hexanethiol on gold, top view



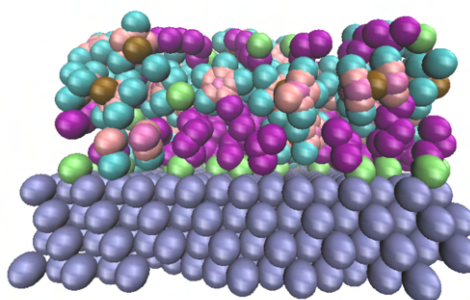
(a) 0 s side view



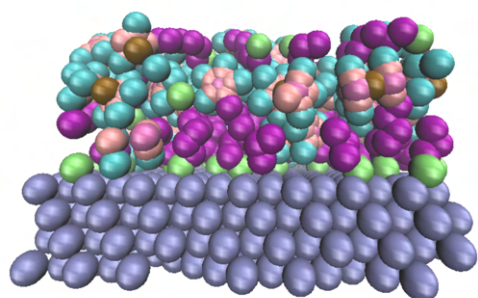
(b) 250 ps side view



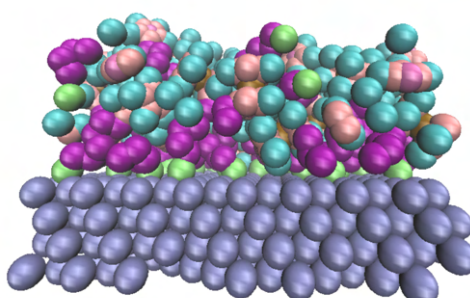
(c) 1 ns side view



(d) 2.5 ns side view



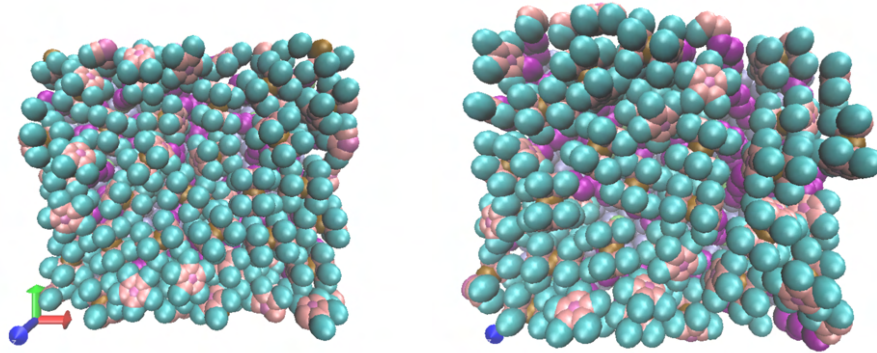
(e) 5 ns side view



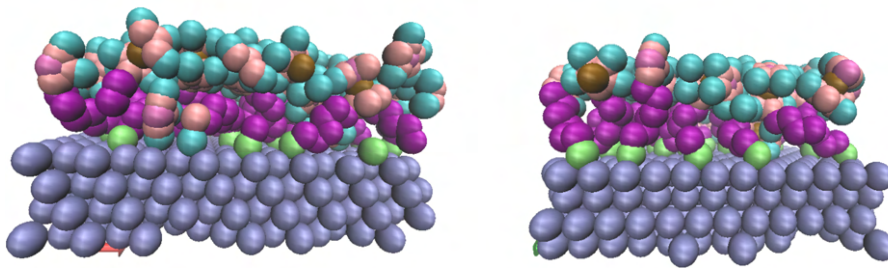
(f) 20 ns side view

Figure 17: Deposition of 50 6(ferrocenyl)hexanethiol on gold, side view

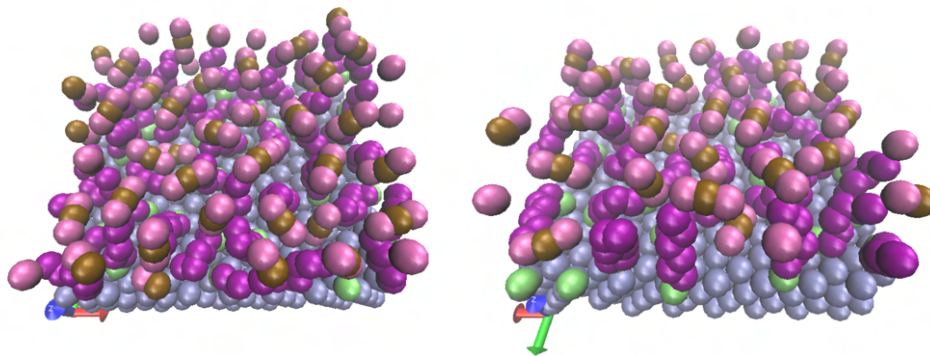
The final configurations reached after cleaning the substrate from unbound molecules can be compared.



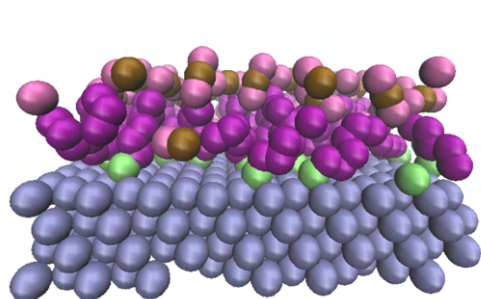
(a) Model 1: final configuration, top view (b) Model 2: final configuration, top view



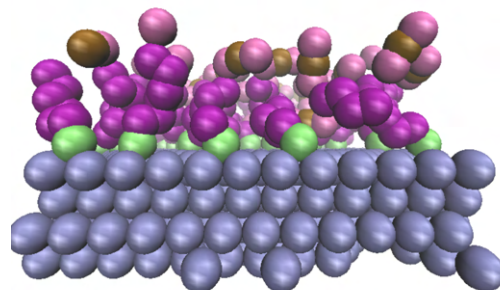
(c) Model 1: final configuration, side view (d) Model 2: final configuration, side view



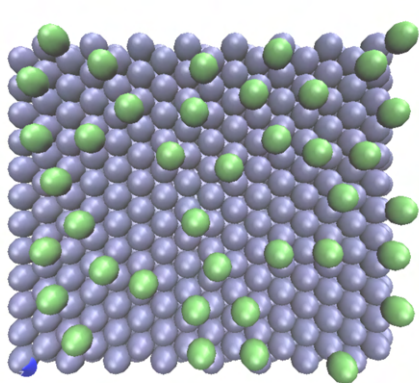
(e) Model 1: final configuration, simplified top view (f) Model 2: final configuration, simplified top view



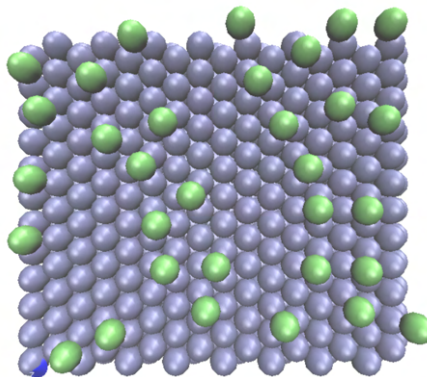
(g) Model 1: final configuration, simplified side view



(h) Model 2: final configuration, simplified side view



(i) Model 1: final configuration, sulfur ligands distribution



(j) Model 2: final configuration, sulfur ligands distribution

Figure 18: Comparison between the two free depositions of thiols on gold

Deposition atom by atom

The deposition above is attempted again, with the force field from model 1, but this time the molecules are added to the system atom by atom.

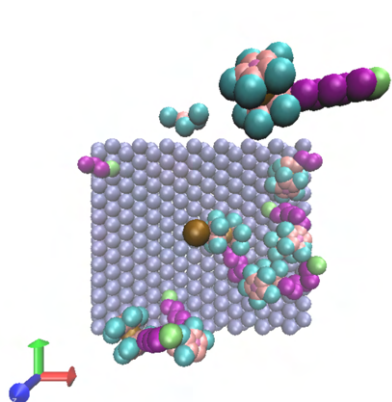
For the deposition step, a timestep of 0.1 fs is chosen and velocity Verlet dynamics is run with a Nosé Hoover thermostat at 300K for 0.2 ns (2'000'000 timesteps) with a molecule being inserted every 3 ps until a total of 50 molecules are inserted.

The molecules are inserted in a region 4.5 nm above the substrate with an initial center of mass velocity selected randomly between (-20,-20,-80) and (20, 20, -70) m/s. This range, lower than the thermal velocity, is selected to reduce the probability for the molecule to bounce off the substrate after impacting on it. The head of the molecules are not rigid

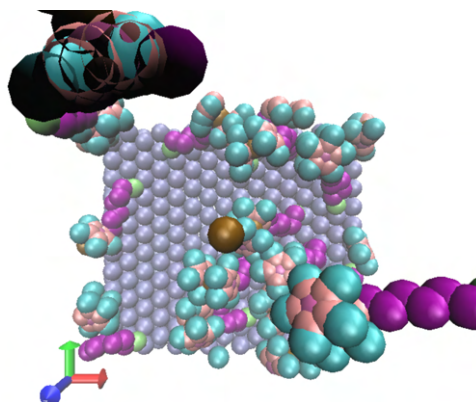
due to the inability to set newly created groups of atoms dynamically as a rigid body while the simulation is running.

With the ferrocene geometry, a challenge in molecular dynamics is the difficulty of constraining the oscillation of hydrogen atoms. For that reason whenever possible the head is set as a rigid body after energy minimization, since the internal oscillations in the head are not expected to significantly affect the dynamics of the deposition. In this case, since this is not possible, a lower timestep had to be used to prevent the hydrogen oscillations from becoming unstable.

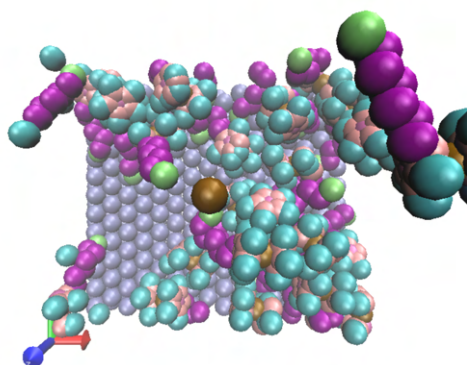
After the deposition, a minimization step is executed and the ferrocene head of the molecules are set as rigid bodies, then 10 ns of velocity Verlet dynamics are run with a Nosé Hoover thermostat at 300K, with a timestep of 1 fs. A full can of the molecules over the substrate could already be observed at the end of 200 ps from the start of the deposition. Subsequently the molecules in excess are removed leaving 36 molecules anchored to the substrate, with a resulting density of 4.32×10^{-10} mol/cm² after a further 5 ns of dynamics under the same conditions as before.



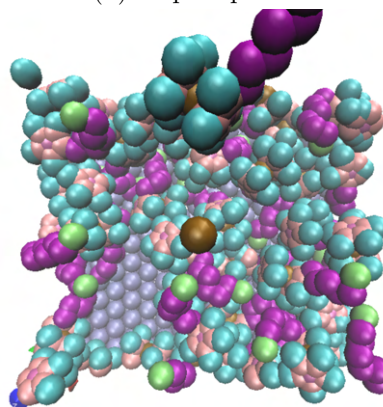
(a) 25 ps top view



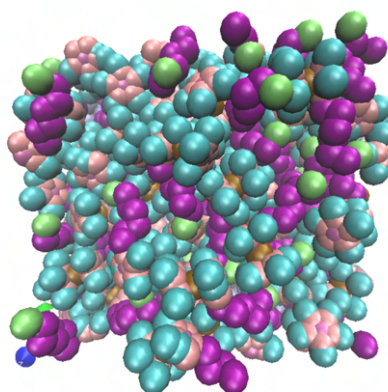
(b) 50 ps top view



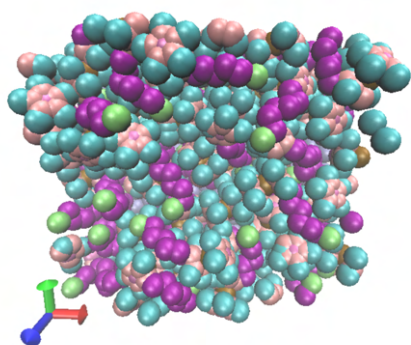
(c) 75 ps top view



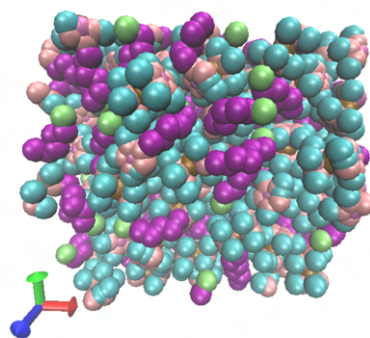
(d) 100 ps top view



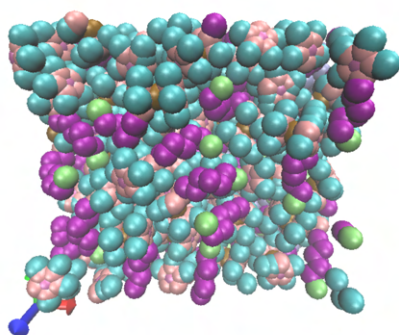
(e) 200 ps top view



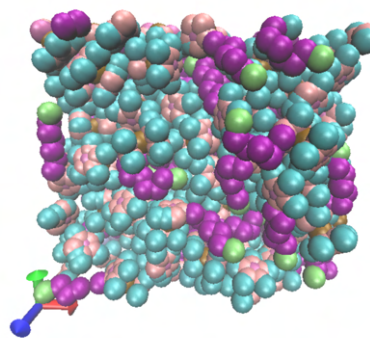
(f) 1ns top view



(g) 2.5ns top view

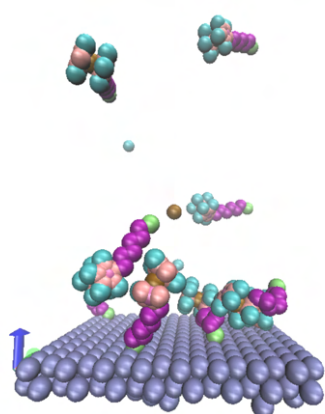


(h) 5ns top view

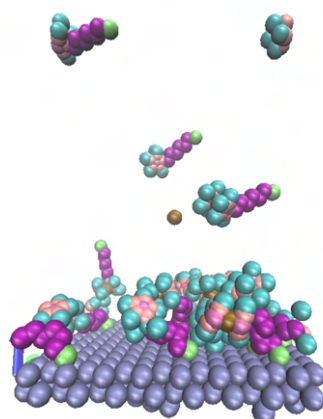


(i) 10ns top view

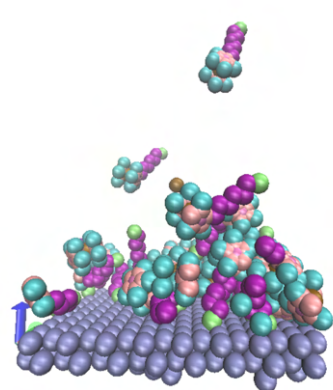
Figure 19: Highlights of the atom by atom deposition of 50 6(Ferrocenyl)hexanethiols on gold, top view



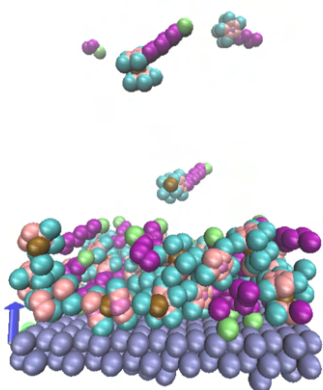
(a) 25 ps side view



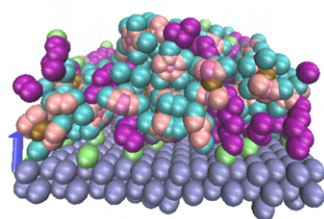
(b) 50 ps side view



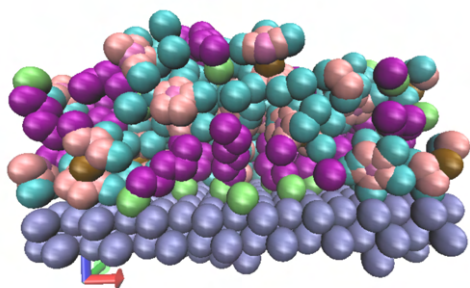
(c) 75 ps side view



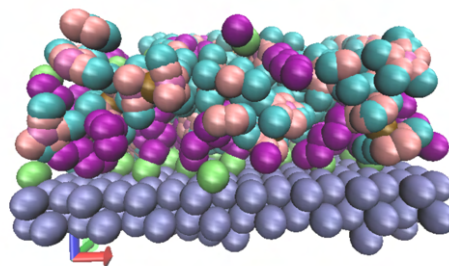
(d) 100 ps side view



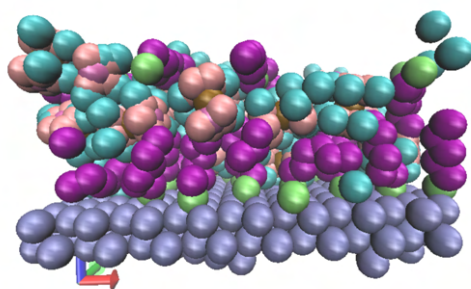
(e) 200 ps side view



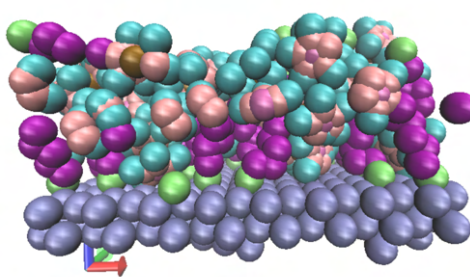
(f) 1ns side view



(g) 2.5ns side view



(h) 5ns side view



(i) 10ns side view

Figure 20: Highlights of the atom by atom deposition of 50 6(Ferrocenyl)hexanethiols on gold, side view

Molecular dynamics study of 6(Ferrocenyl)hexanethiols atoms anchored on a gold substrate

There are several limits with the MD analysis of a deposition from a droplet: with such a small sample size of atoms, limited to prevent a too high computation time, there may be a strong variability of the final distributions depending on the initial conditions. Moreover, it's not clear that the Morse potential is effective in modeling the Au-S bond, especially in the case where the thiol chains are too loosely packed to guarantee that the anchoring point is maintained stable by the presence of the nearest neighbors.

Anchoring the sulfur atoms in a known pattern allows to remove the variability of the binding site, guaranteeing a regular distribution of the molecules, as it should happen in experimental conditions where the monolayer forms in normal deposition times.

The force field parameters are the same as in model 1 for the free thiols in section ?? except for the sulfur to gold interaction. Two different geometrical patterns are studied. Figure 21 highlights both of them. Choosing two vectors following the hexagonal lattice direction at an angle of 120° , for instance respectively $(-0.5, \sqrt{3}, 0)$ and $(1, 0, 0)$, the two superlattices have as vectors $[1\ 3\ 0]$ and $[-2\ 1\ 0]$ for the first pattern and $[2\ 0\ 0]$ and $[3\ 1\ 0]$ for the second one, as seen in figure 21

For both patterns, sulfur atoms are placed at the top of gold atoms as done by Rai and colleagues[36] or bridge position, as suggested by the experiments by Yourdshahyan and Rappe[41]. The Au-S bond length is 2.89 \AA for the bridge position (sulfur bonded to the two bridge atoms) and 2.40 \AA for the top position. In the top position three cases are studied: two where respectively a 108° and 100° Au-S-C angle is fixed, as suggested by the literature[36][41], and one with no constrained bond angle. In the bridge position an angle of 50° for the S-C bond with respect to the surface normal, and a configuration with only the bonds have been investigated instead. In figure 22 a model of the FcC6S molecule on gold is shown both in the top (a) and the bridge (b) configurations. The three critical angles are highlighted. θ is the fixed angle for the S-C bond, β the angle of the fc head with the surface and ϕ the tilt angle of the head with respect to the surface normal.

As for the surface size, in order to ensure a correct periodic wrapping of the superlattice, 35.0×40.4 was chosen for the first arrangement and 60.0×34.6 for the second one. This means that the molecules were deposited at a density of $3.29e - 10\text{mol}/\text{cm}^2$ in the first case and $3.84e - 10\text{mol}/\text{cm}^2$ in the second case.

The molecules are initially positioned at an approximate angle of 30° with the surface normal and energy minimization is run. Then a timestep of 1 fs is selected, the temperature is ramped up over 1 ns of velocity-Verlet dynamics with a Nosè-Hoover thermostat and then kept at 300k for 4 more ns of dynamics.

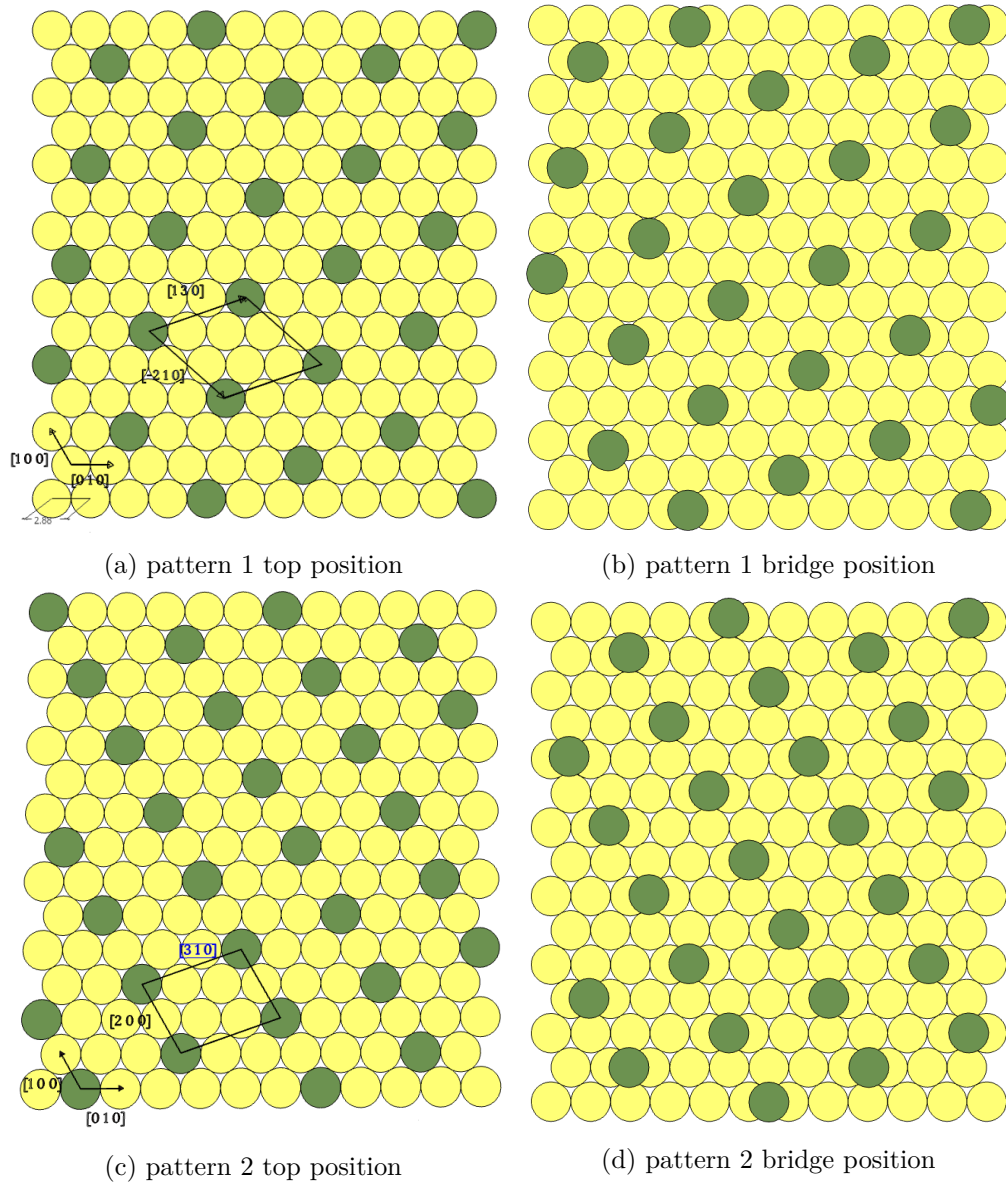
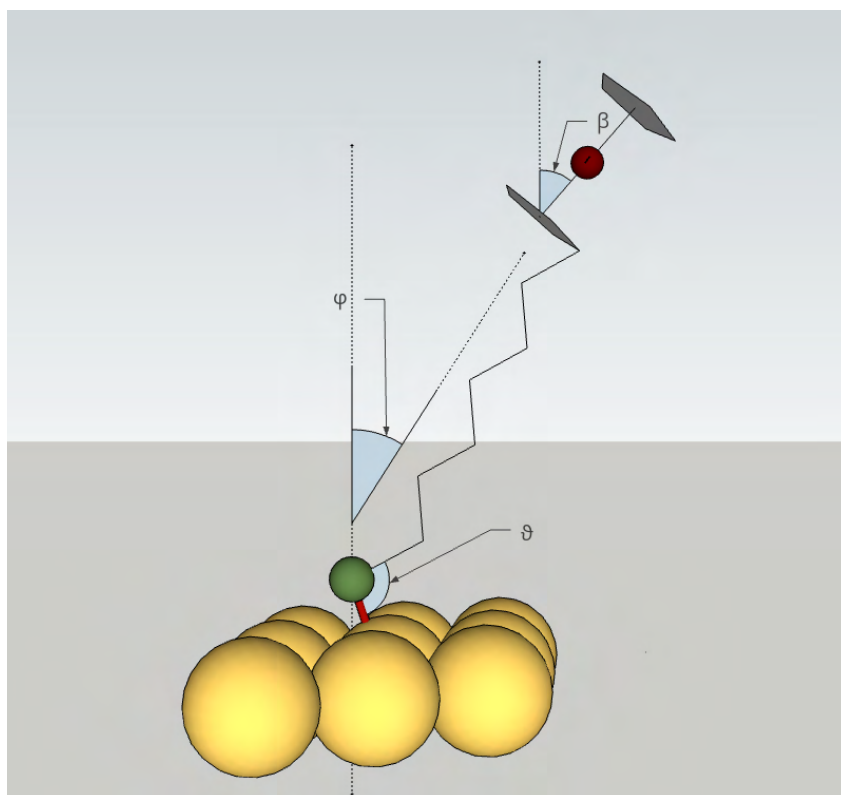
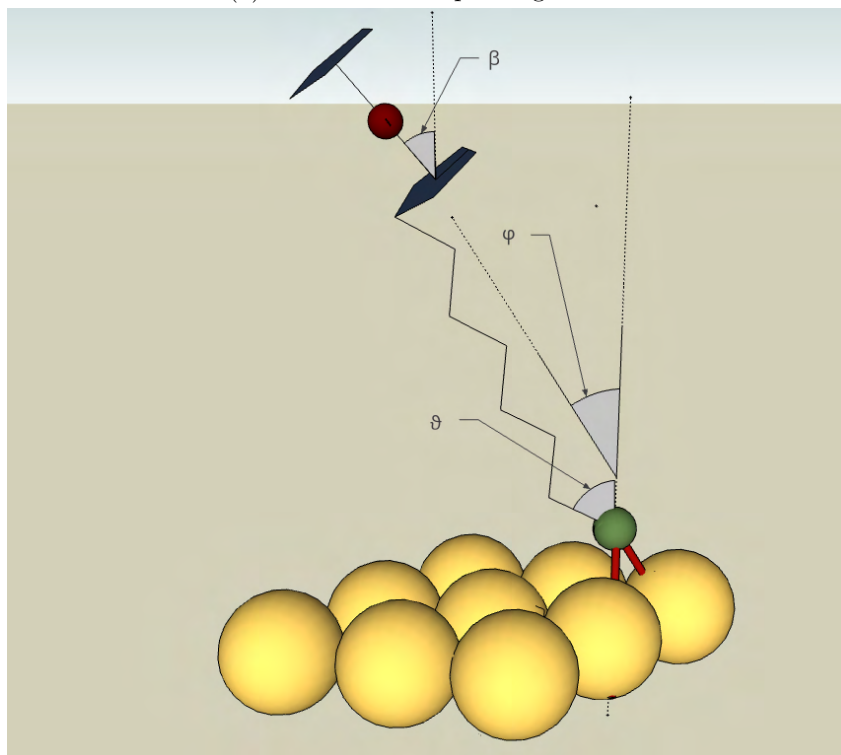


Figure 21: Arrangement of the anchoring points of the FeC6S superlattice



(a) FcC6S in the top configuration



(b) FcC6S in the bridge configuration

Figure 22: Model of the FcC6S molecule on gold. ⁵⁷ ϕ is the molecule tilt angle, β is the head tilt angle and θ is the Au-S-C bond for the top configuration and the S-C bond with the surface normal in the bridge configuration

Theory

Obtaining theoretical parameters for the study of the FcC6S molecule is difficult as the number of papers focusing on the geometry of such a system is scarce. In many cases the focus is instead the electrochemical properties. A research through the scarce literature available however allowed to find some geometrical quantities to be used in order to validate the molecular dynamics experiments.

Molecule tilt angle From the literature the estimated tilt angle for the alkyl chain is in the neighborhood of 30° [44]. [46] estimates $34^\circ \pm 8^\circ$ for SC11Fc

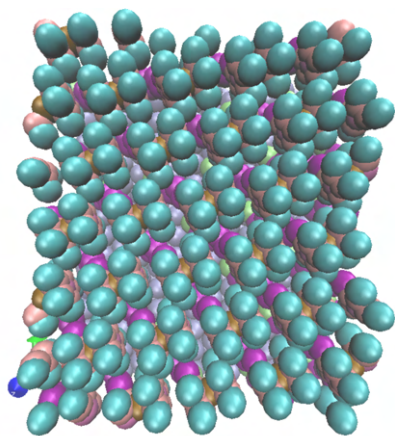
Height of the monolayer For the height of the monolayer, the best guesses come from other simulations and geometrical calculations, giving values between 12 and 12.5 \AA [43][44]. Experiments tend to strongly overestimate the values, returning heights above 20 \AA [45], which is not geometrically feasible.

Head tilt angle NEXASf experiments allow to guess what the tilt angle of the head is with respect to the horizontal. However unfortunately no data is available directly on FcC6S. Nerngchamnong et al. [43] found values from 54° to 64° for alkyl chains with a length of 8 to 13 carbons, highlighting an odd-even effect: chains with an odd number of carbons have higher values, between approximately 61° and 63° and chains with an even number of carbons occupy the lower end of the spectrum, with values between 54° and 57° . The results of the MD experiment performed by the same authors show 56° as the value for FcC6S. All values are assumed to have a relative error of approximately 5° . Nerngchamnong et al. [45] report α values of 60.2° for the SC3Fc SAMs and 45.4° for the SC4Fc SAMs. 61.55° and 44.06° are the data derived by the same authors from XRD experiment. Again an odd-even effect is apparent. Watcharinyanon et al. [46] report $46^\circ \pm 5^\circ$ for the SC11Fc molecule. Overall, while it's not possible to find theoretical values with confidence, a range of 45 to 60° is acceptable for the values of the molecule tilt angle.

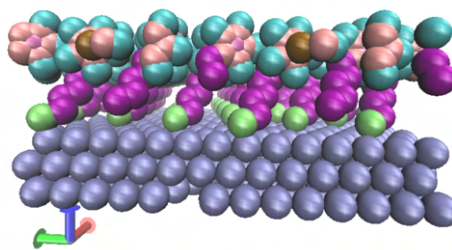
Density of the monolayer and structure The literature [46] suggests a maximum monolayer density of $4.5 \times 10^{-10} \text{ mol/cm}^2$. Several experiments show a range of values: 3.98×10^{-10} [43], $3.32 \times 10^{-10} \text{ mol/cm}^2$ [44], $2.98 \times 10^{-10} \text{ mol/cm}^2$ [45], $3.32 \times 10^{-10} \text{ mol/cm}^2$ [47] and $3.0 \times 10^{-10} \text{ mol/cm}^2$ [48]. For what concerns the packing, different packings have been proposed by Nerngchamnong et al. [45] and Müller-Meskamp et al. [42], as shown in section

MD results

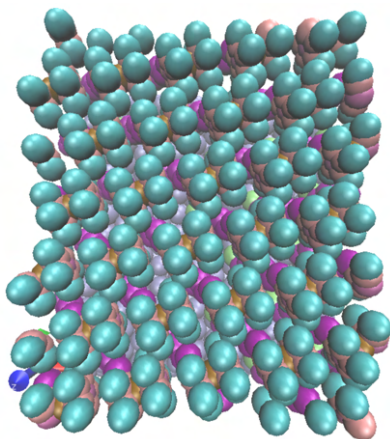
figure 23 shows the initial position of the FcC6S molecules, while figure 24 shows the end result of the simulation. In all cases, despite the regular anchoring, a significant amount of disorder can be observed. The first geometry, with the lower density, shows some hollows caused by the molecule heads packing together, while the second geometry, with the higher density, shows a more uniform filling. A good amount of the disorder can be explained by the significant presence of trans isomers in the thiol chain, affecting both the chain tilt and the head tilt. In general, the chains are too far apart to interact significantly to the point of aligning. This is a significant distinction compared to the anchoring of simple 1-hexanethiols, with the spacing forced by the head hindrance, reducing the regularity of the outcome



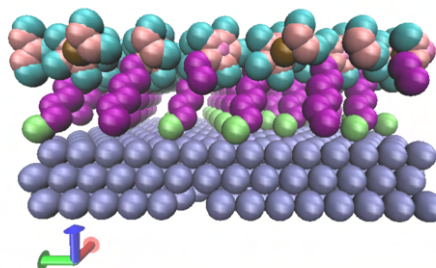
(a) Initial distribution, geometry 1, top anchoring



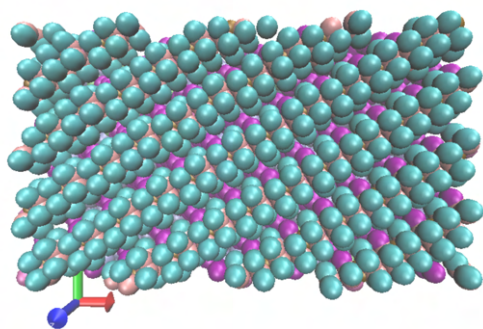
(b) side view



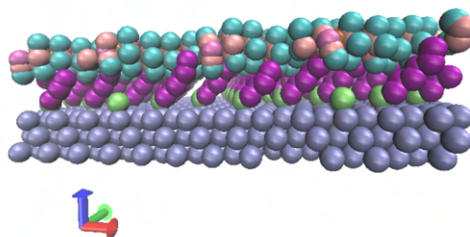
(c) Initial distribution, geometry 1, bridge anchoring



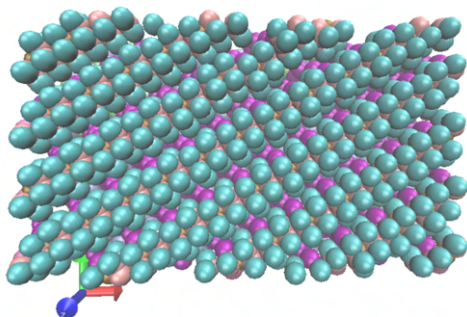
(d) side view



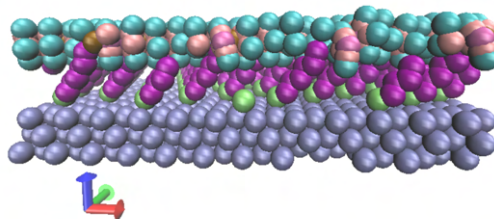
(e) Initial distribution, geometry 2, top anchoring



(f) side view

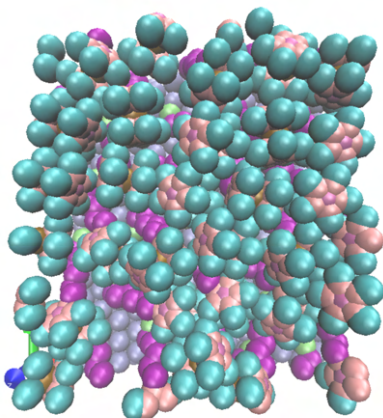


(g) Initial distribution, geometry 2, bridge anchoring

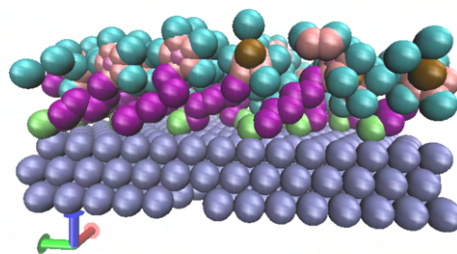


(h) side view

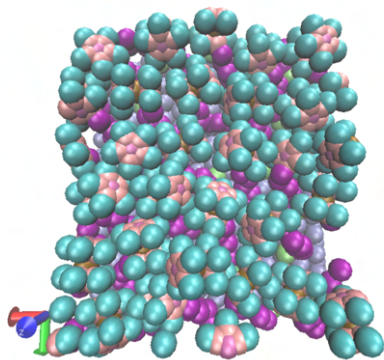
Figure 23: Initial distribution of the thiols on the top and bridge configurations



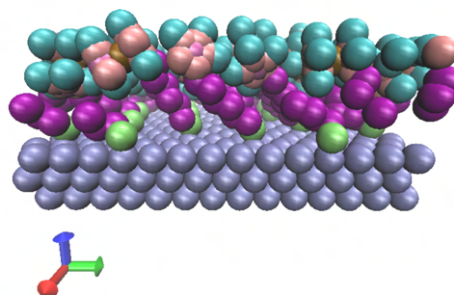
(a) Final distribution, geometry 1, top anchoring, 100°



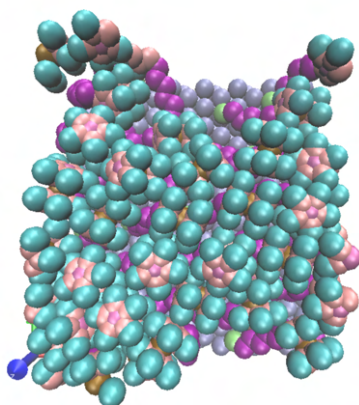
(b) side view



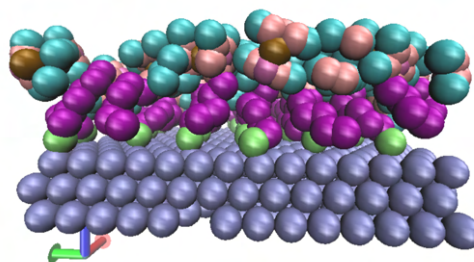
(c) Final distribution, geometry 1, top anchoring, 108°



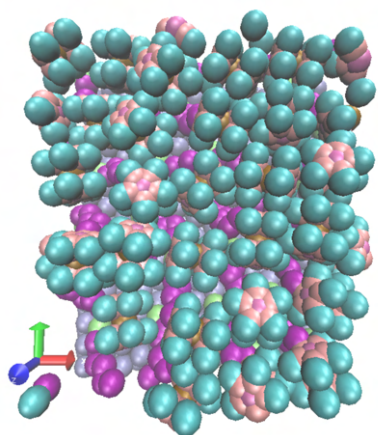
(d) side view



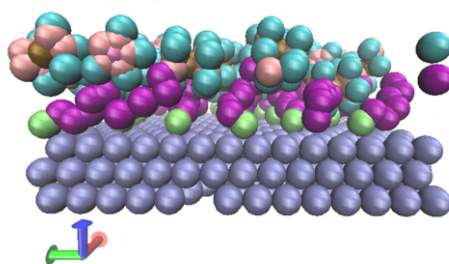
(e) Final distribution, geometry 1, top anchoring, no angle



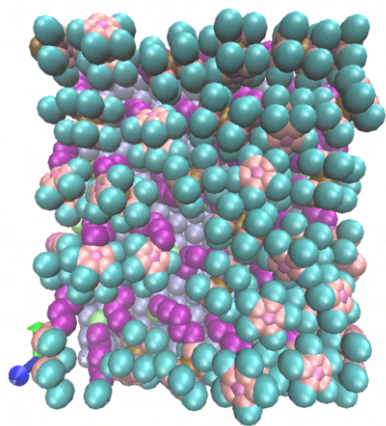
(f) side view



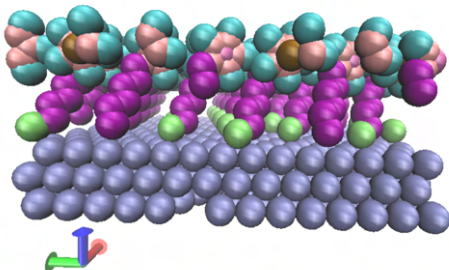
(g) Final distribution, geometry 1, bridge anchoring, fixed angle



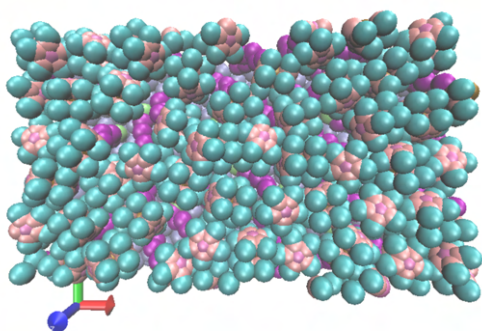
(h) side view



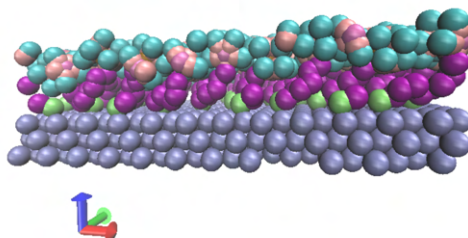
(i) Final distribution, geometry 1, bridge anchoring, no angle



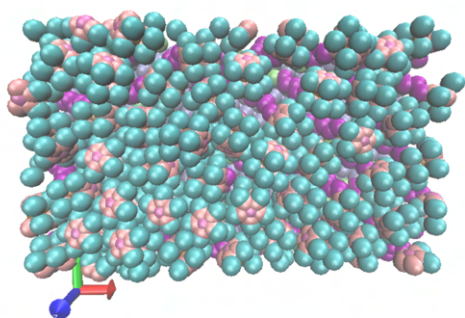
(j) side view



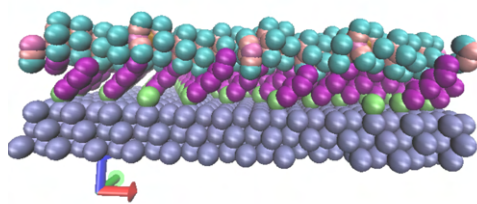
(k) Final distribution, geometry 2, top anchoring, 100°



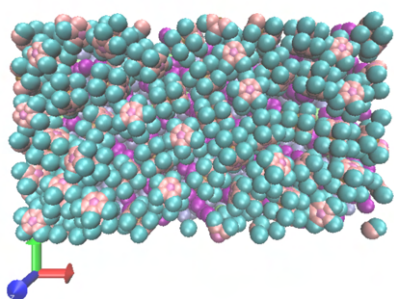
(l) side view



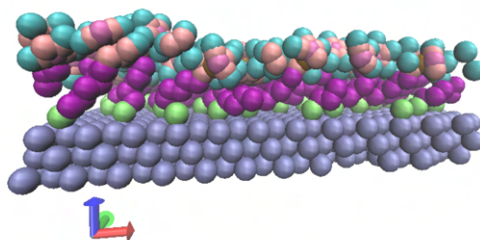
(m) final distribution, geometry 2, top anchoring, 108°



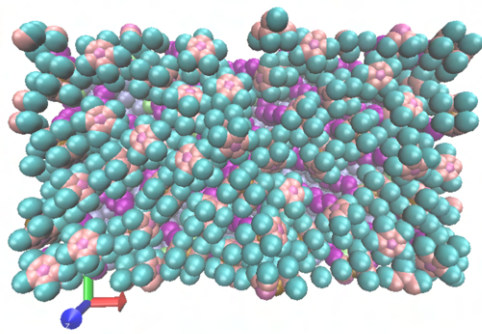
(n) side view



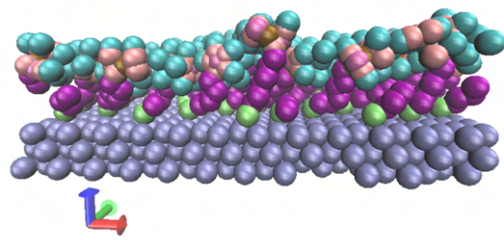
(o) Final distribution, geometry 2, top anchoring, no angle



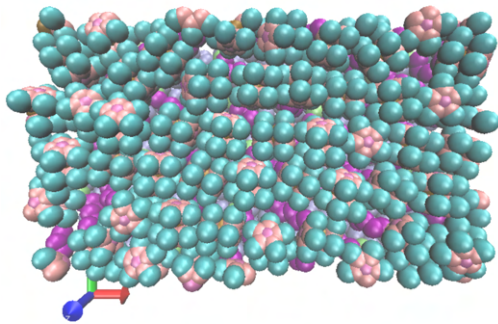
(p) side view



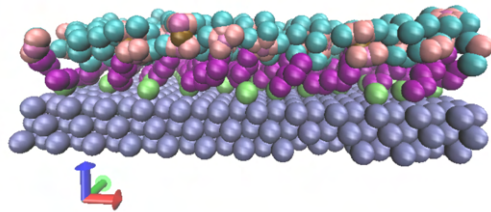
(q) Final distribution, geometry 2, bridge anchoring, fixed angle



(r) side view



(s) final distribution, geometry 2, bridge anchoring, no angle



(t) side view

Figure 24: Final distribution of the thiols on the top and bridge configurations

MD Postprocessing

Free thiols The raw data from the trajectory files at equilibrium are postprocessed taking 100 frames at 50 ps intervals for the deposited FeC6S and 100 frames at 20 ps intervals for the anchored molecules. The following plots show the results. The table in figure [25](#) shows the averages for the computed metrics, obtained by averaging over all the per molecule values in the 100 frames:

- "head avg tilt angle" is the average tilt angle β between the line connecting the Cp centroids and the surface plane;
- "mol avg tilt angle" is the average tilt angle of the molecule measured as the angle with the surface normal of the line connecting the sulfur atom with the chain carbon atom connected to the cp ring;
- "sam density" is the surface density of the monolayer;
- "sam height" is the average z distance between the surface gold atoms and the top of the molecule;
- "average S nn distance" is the average distance between each molecule and its first neighbour;
- "sam height no extremes" is the average z distance between the surface gold atoms and the top of the molecule, with values at more than two standard deviations from the average removed;
- "average S nn dist no extremes" is the average distance between each molecule and its first neighbour, with values at more than two standard deviations from the average removed;
- "second neighbor" and "third neighbor" are respectively the distance of each molecule from its second and third neighbor.

For the SAM height and the first neighbour distance, standard deviation values are also reported, since the distribution is central with values tapering off toward the extremes.

	Free thiols force field 1	Free thiols force field 2	Free thiols deposition
method	MD	MD	MD
head avg tilt angle	56.7°	58.1°	54.7°
mol avg tilt angle	33.3°	40.5°	35.6°
sam density (mol/cm ²)	4.68E-10	3.84E-10	4.32E-10
sam height (Å)	11.78±2.02	11.9±1.75	12.48±1.97
average S nn distance (Å)	4.75±0.41	4.96±0.63	4.82±0.52
sam height no extremes (Å)	12.22±1.44	12.3±1.16	12.98±1.22
avg s nn dist no extremes (Å)	4.68±0.32	4.90±0.53	4.68±0.32
second neighbor (Å)	5.28	5.79	5.28
third neighbor (Å)	5.92	6.65	5.92

(a) Free thiols

	geom. 1	geom. 1, top, 108° angle	geom. 1, top, no angle	geom 1, bridge, angle	geom 1, bridge no angle
method	MD	MD	MD	MD	MD
head avg tilt angle	54.0°	54.78°	57.6°	55.61°	55.74°
mol avg tilt angle	46.3°	48.13°	28.2°	48.24°	54.94°
sam density (mol/cm ²)	3.29E-10	3.29E-10	3.29E-10	3.29E-10	3.29E-10
sam height (Å)	11.24±0.99	11.30±1.00	11.75±1.11	11.46±0.92	11.41±0.97
average S nn distance (Å)	7.03±0.28	7.08±0.24	7.03±0.27	7.25±0.22	7.25±0.23
sam height no extremes (Å)	11.27±0.91	11.34±0.91	11.87±0.85	11.51±0.79	11.47±0.86

(b) geometry 1

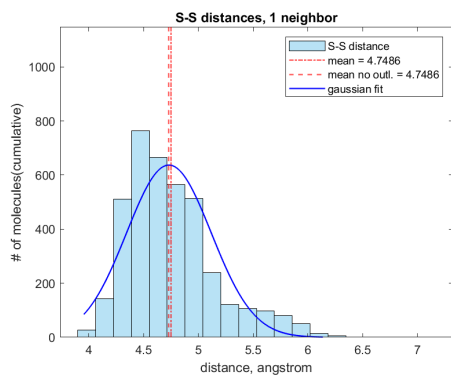
	geom. 2	geom. 2, top, 108° angle	geom. 2, top, no angle	geom 2, bridge, angle	geom 2, bridge no angle
method	MD	MD	MD	MD	MD
head avg tilt angle	57.11°	56.41°	57.94°	56.71°	63.4°
mol avg tilt angle	44.97°	40.1°	39.90°	40.1°	45.1°
sam density (mol/cm ²)	3.84E-10	3.84E-10	3.84E-10	3.84E-10	3.84E-10
sam height (Å)	11.43±0.94	11.55±1.10	11.62±1.08	11.68±1.04	11.59±0.73
average S nn distance (Å)	5.40±0.39	5.45±0.36	5.48±0.33	5.49±0.35	5.47±0.33
sam height no extremes (Å)	11.33±0.80	11.53±0.89	11.65±0.82	11.56±0.82	11.51±0.62

(c) geometry 2

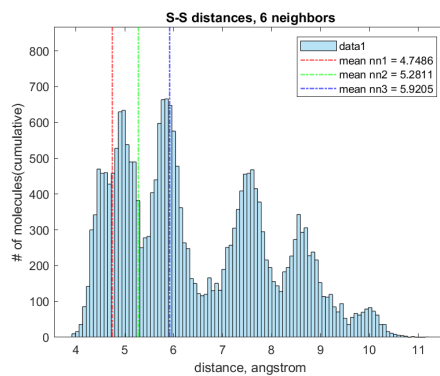
Figure 25: Tables reporting the experimental results

The following graphs report the histograms, showing the distribution of the variables under analysis aggregating the values from all 100 frames the graphs are reported first for the systems with the mobile S-C bond and then for the ones with the anchored sulfurs. The reported plots are first neighbour distances, distances of the first six neighbours, head tilt angle, chain tilt angle and height of the monolayer. Comparing the free distributions the following apply:

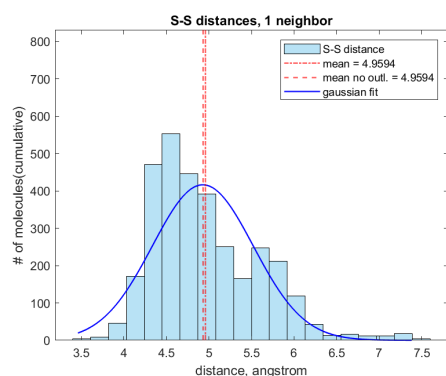
1. the nearest neighbour distances are between 4.75 and 5, the latter of which correspond to the minimum distance of $\sqrt{3} \cdot 2.88$ achievable by placing the molecules in non-contiguous three-fold hollows. The distribution is left-skewed. This bias might be the consequence of the fact that if more neighbours are at a similar distance from one atom, the closest one is always picked.
2. The first force field, both in the deposition and in the droplet case, show the denser packing, even greater than the theoretical maximum of $4.5 \cdot 10^{-10} \text{ mol/cm}^2$.[\[47\]](#)
3. The head tilt angle distributions are similar, increasing in frequency the greater the angle, but with a significant tail at low tilt angles.
4. As for the chain tilt angles, the second model for the force field is more skewed toward higher values, likely due to the presence of a greater number of trans isomers.
5. the molecule heights are more or less normally distributed, with a tail due to some molecules lying horizontally, stuck below the other molecule in the monolayer.
6. in general the molecules stuck in a horizontal position will generate spurious data with respect to a defect-free monolayer.



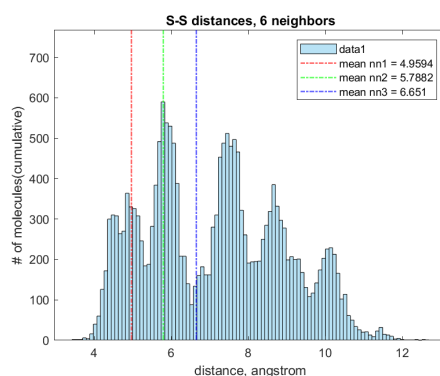
(a) model 1



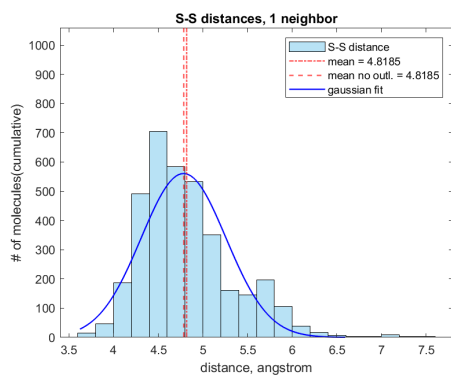
(b) model 1



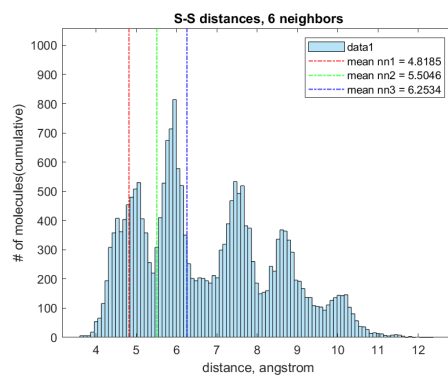
(c) model 2



(d) model 2

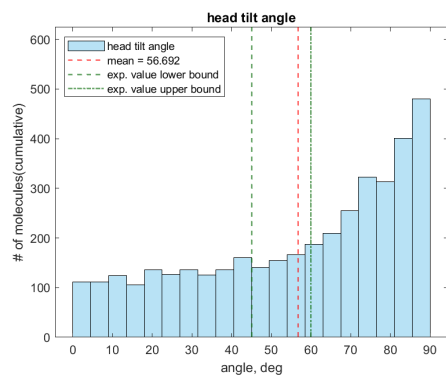


(e) model 1, deposition

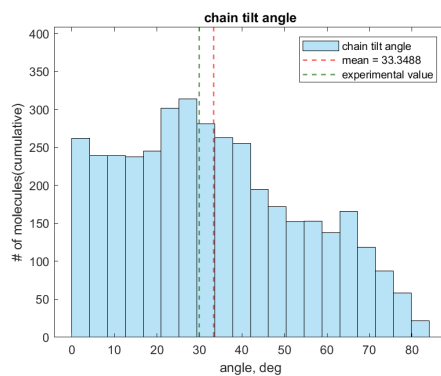


(f) model 2, deposition

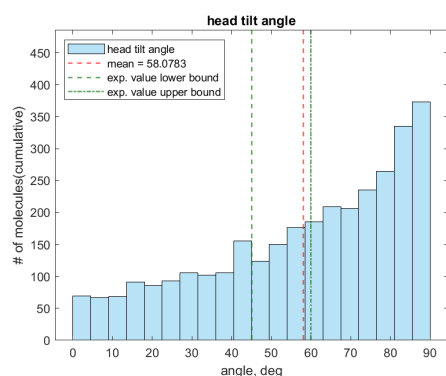
Figure 26: Free thiols, sulfur nearest neighbour distances on the left, the first neighbour distances first six neighbour distances on the right



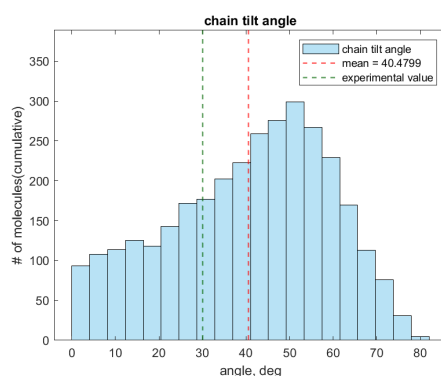
(a) model 1



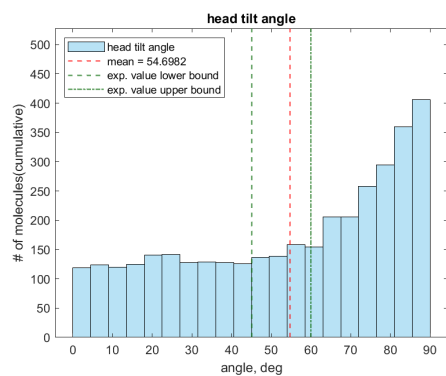
(b) model 1



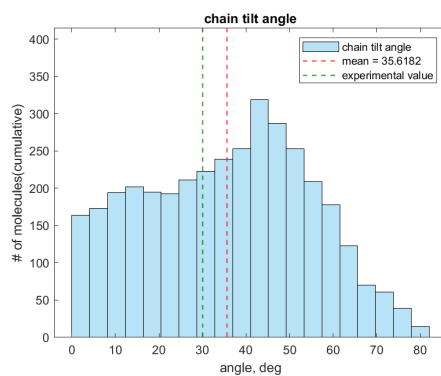
(c) model 2



(d) model 2

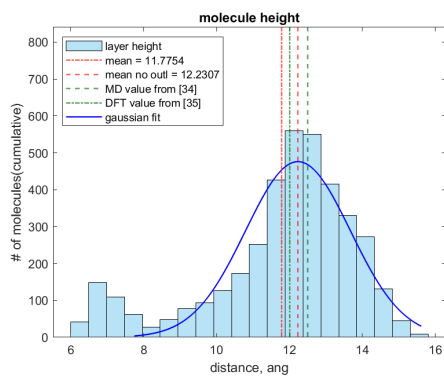


(e) model 1, deposition

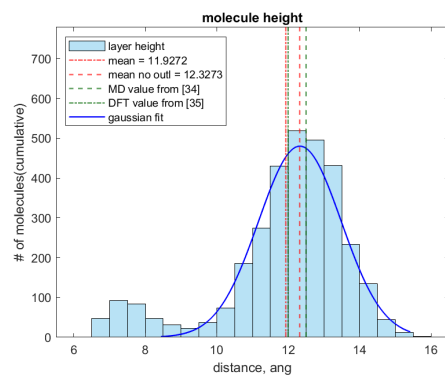


(f) model 1, deposition

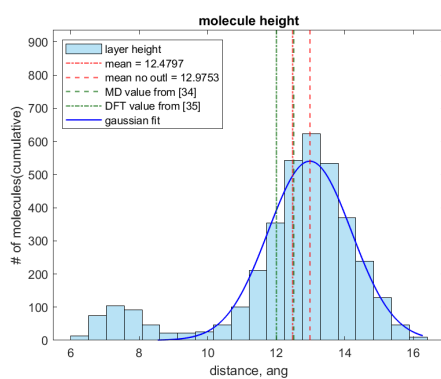
Figure 27: Free thiols, head tilt angles on the left, chain tilt angle on the right



(a) model 1



(b) model 2

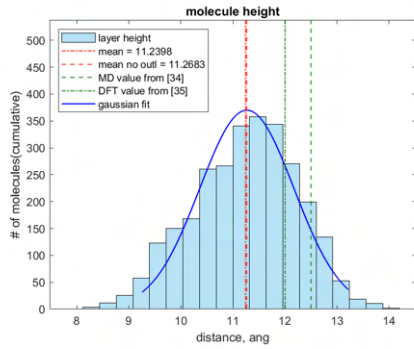


(c) model 1, deposition

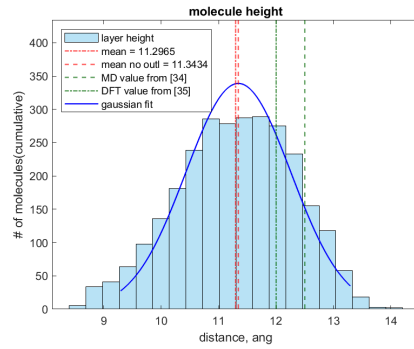
Figure 28: Free thiols, layer height

Anchored molecules

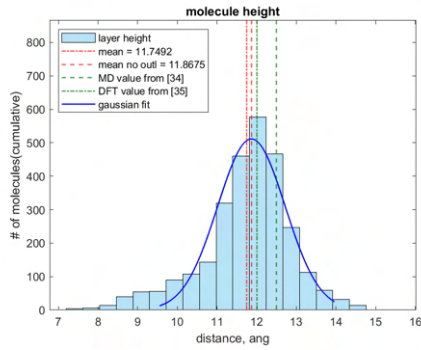
Layer height The molecule height data is approximately normally distributed in the case of the anchored molecule too. Surprisingly, the molecules in the top configuration without the fixed angles exhibit the more sharply peaked distribution. However, not fixing the angle means that there is a tail of strongly bent molecules that have lower values of height compared to the other cases. However in all cases the spread is significant. The average height is between 11 and 12 Å, lower than the theoretical one but without high variability across all the different models.



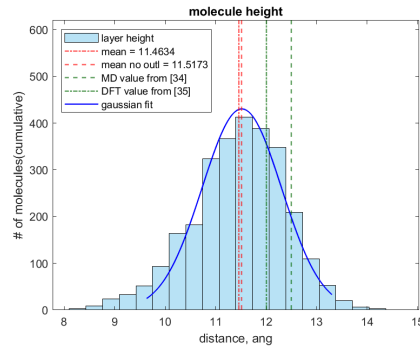
(a) model 1 top, 100° angle



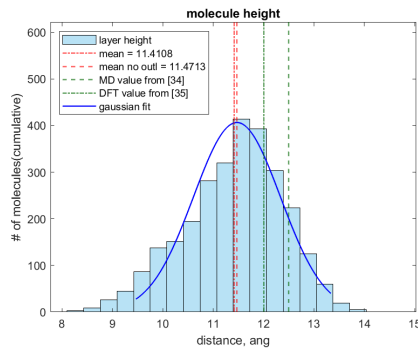
(b) model 1 top, 108° angle



(c) model 1 top, no angle

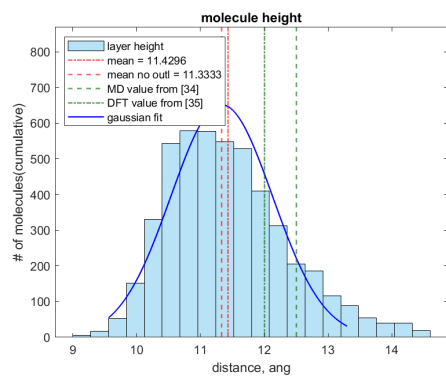


(d) model 1 bridge, angle

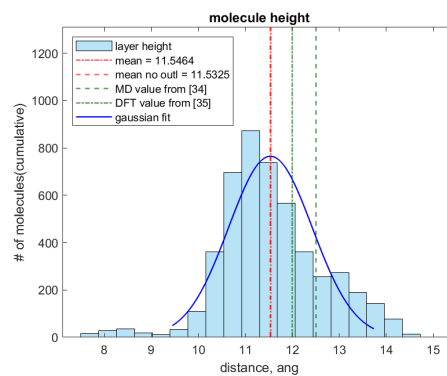


(e) model 1 bridge, no angle

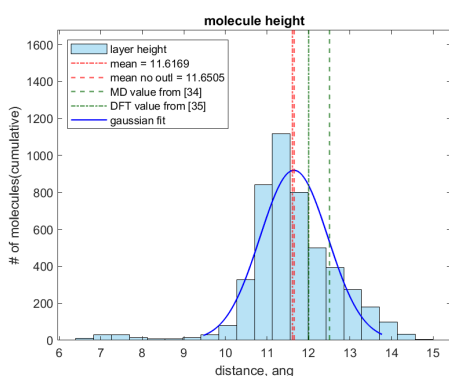
Figure 29: Anchored thiols, layer heights, model 1



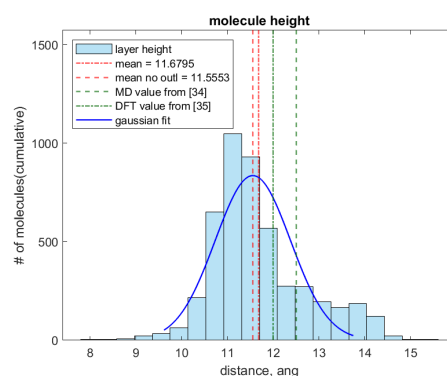
(f) model 2 top, 100° angle



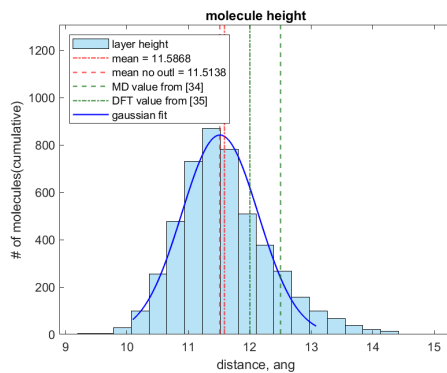
(g) model 2 top, 108° angle



(h) model 2 top, no angle



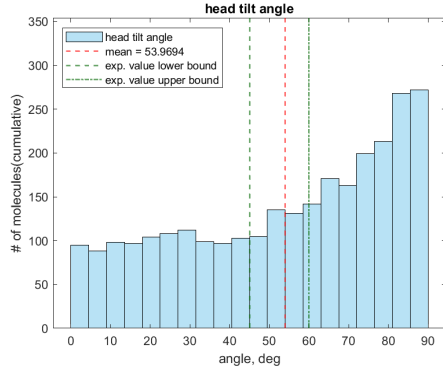
(i) model 2 bridge, angle



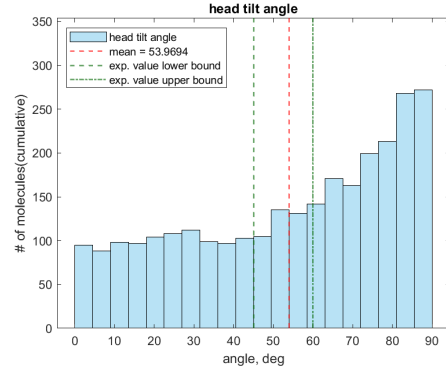
(j) model 2 bridge, no angle

Figure 29: Anchored thiols, layer heights, model 2

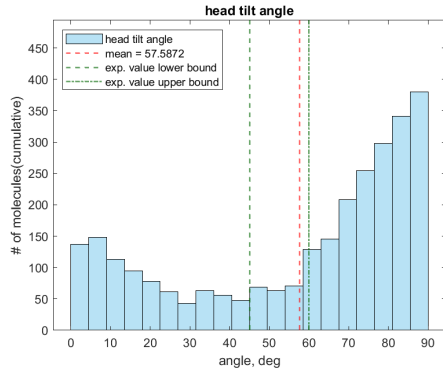
Head tilt angle The head tilt angle shows average values between 54° and 57° , in line with the predictions, especially from Nerngchamnong et al. [43]. However the distribution is again spread among all different angles, growing toward 90° and peaking at 90° . Moreover, in some cases, there is a minimum close to the average. Thus one would need to know the distribution underlying the theoretical data to make a more accurate comparison.



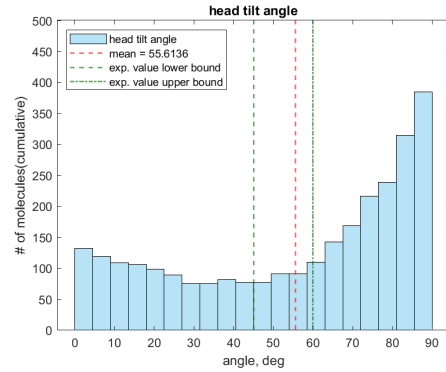
(a) model 1 top, 100° angle



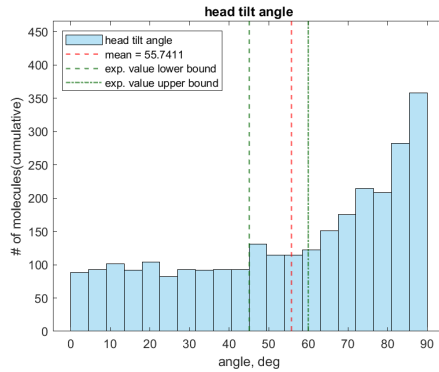
(b) model 1 top, 108° angle



(c) model 1 top, no angle

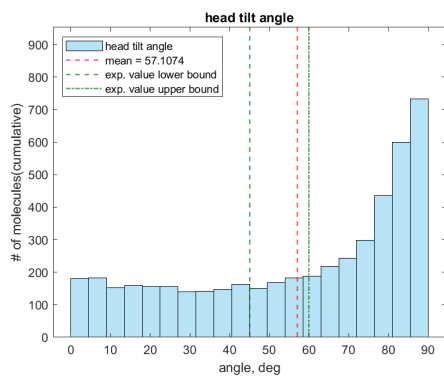


(d) model 1 bridge, angle

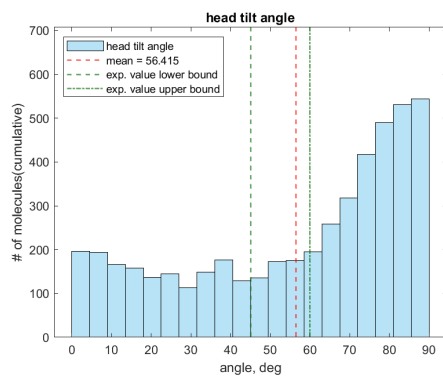


(e) model 1 bridge, no angle

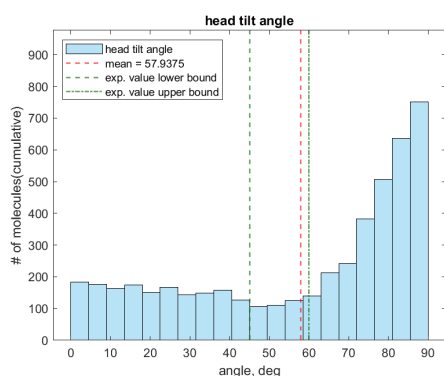
Figure 30: Anchored thiols, head tilt angle, model 1



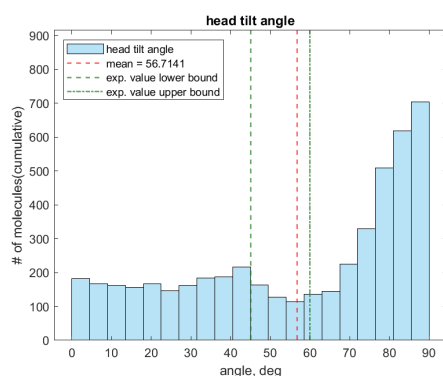
(f) model 2 top, 100° angle



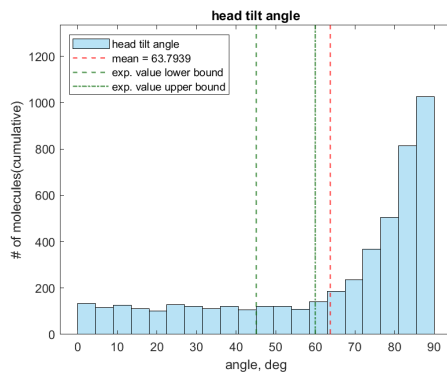
(g) model 2 top, 108° angle



(h) model 2 top, no angle



(i) model 2 bridge, angle



(j) model 2 bridge, no angle

Figure 30: Anchored thiols, head tilt angle, model 2

Chain tilt angle The chain tilt angles tend to have distribution biased toward higher values, well above the 30° tilt from the literature. The exception is the case of the first geometry, with no angle fixed, that exhibits a spread biased toward the low values and thus an average close to the 30° from the literature.

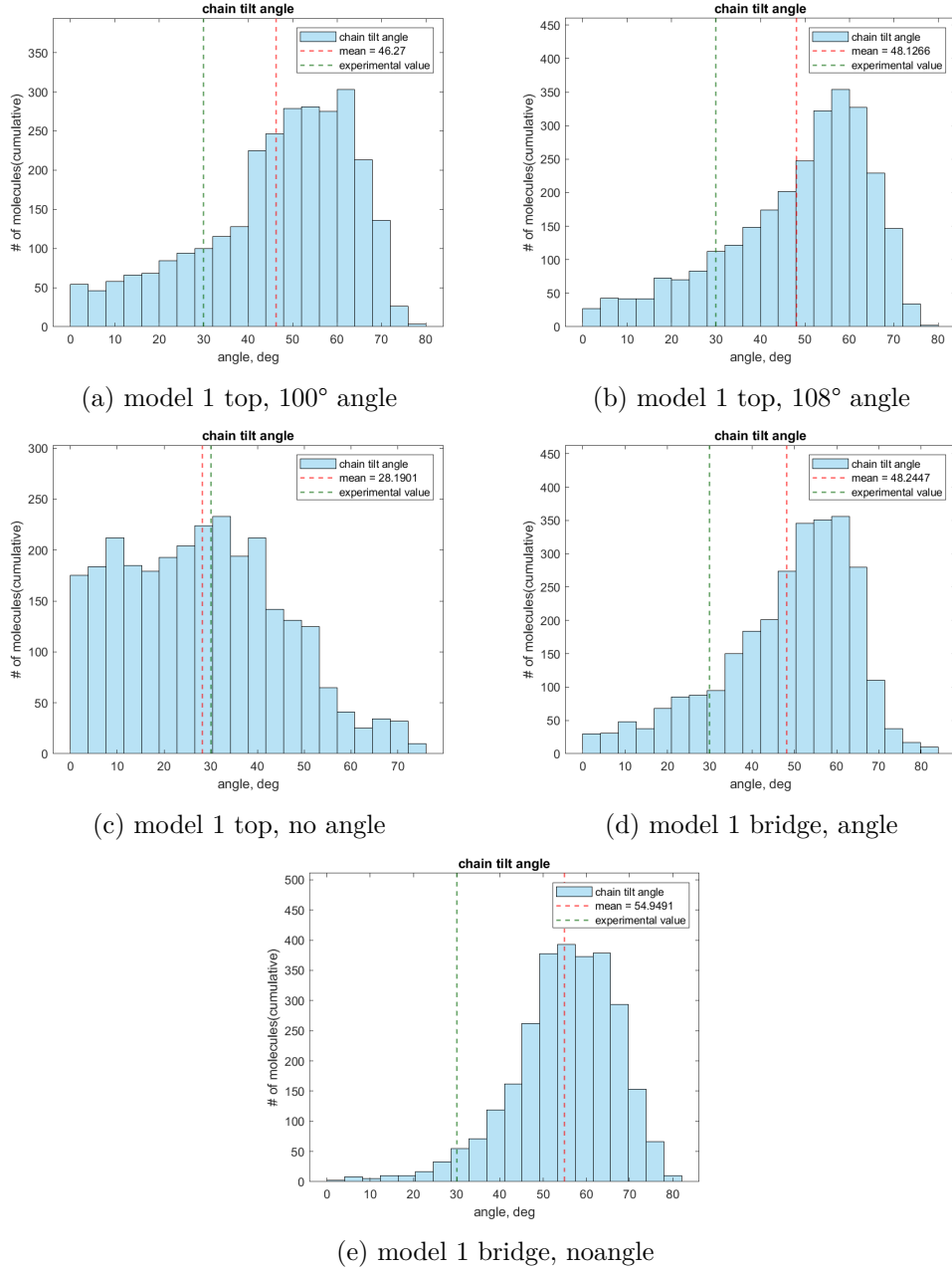
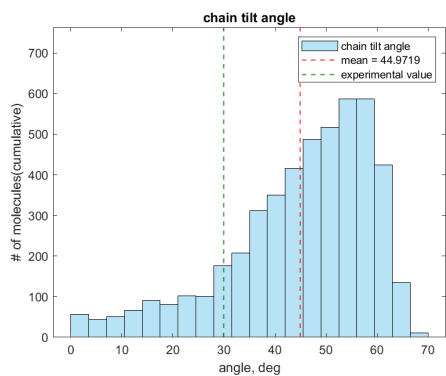
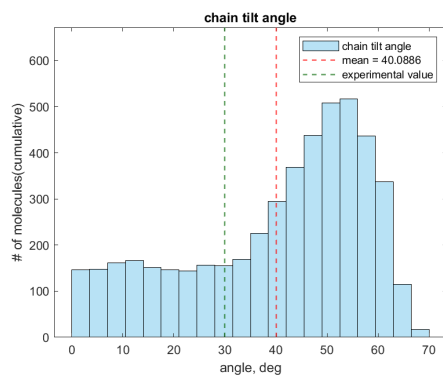


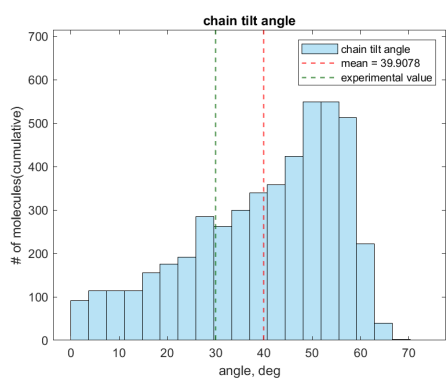
Figure 31: Anchored thiols, chain tilt angle, model 1



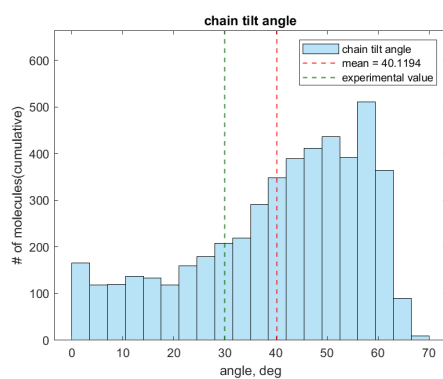
(f) model 2 top, 100° angle



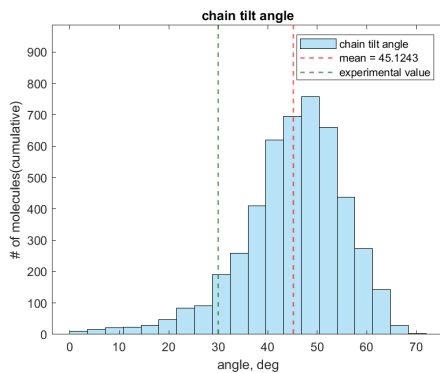
(g) model 2 top, 108° angle



(h) model 2 top, no angle



(i) model 2 bridge, angle

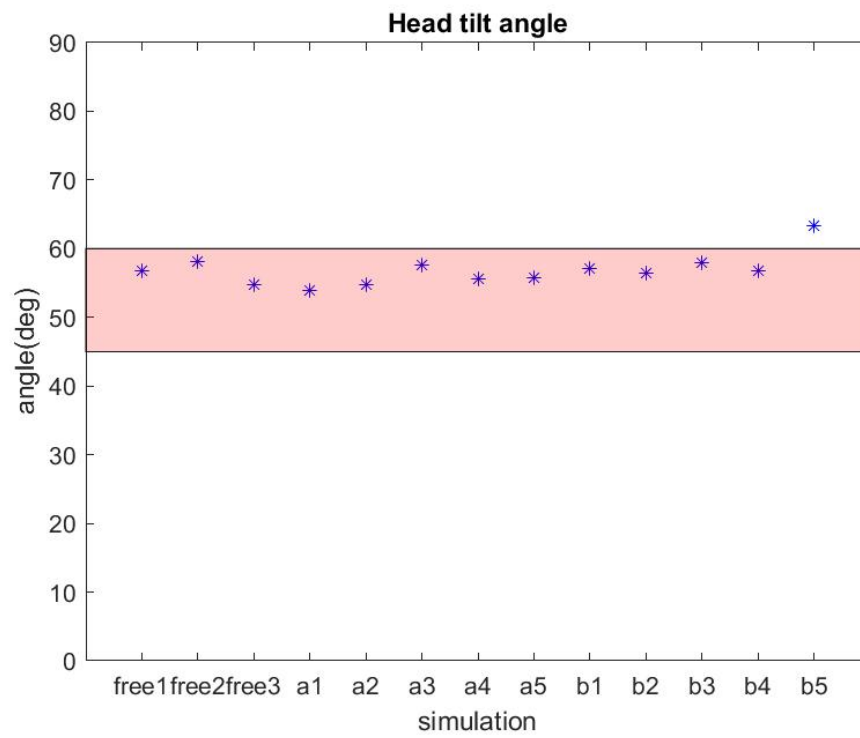


(j) model 2 bridge, no angle

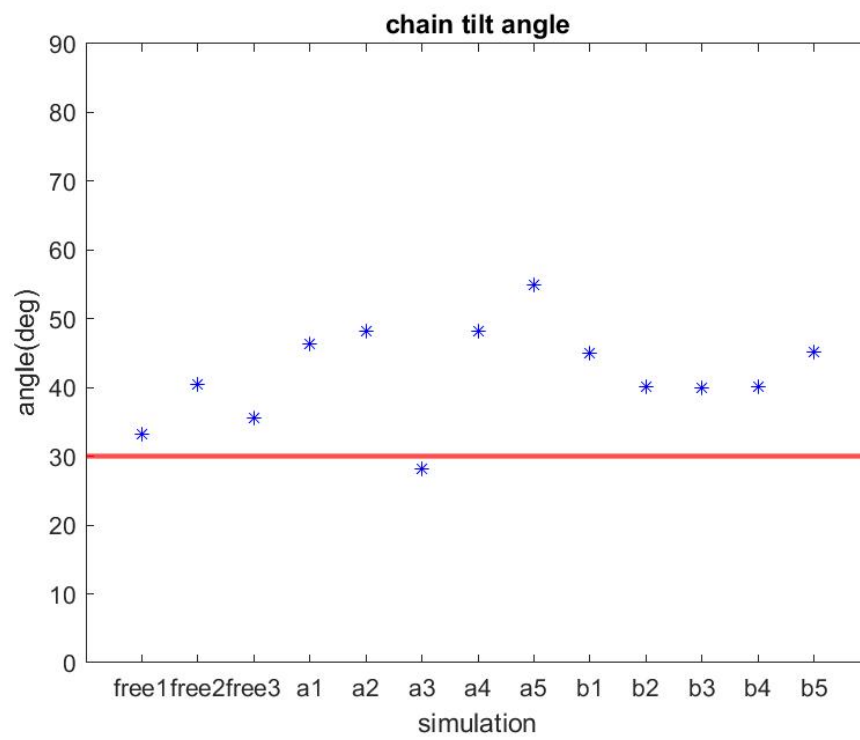
Figure 31: Anchored thiols, chain tilt angle, model 2

Conclusions of part 2

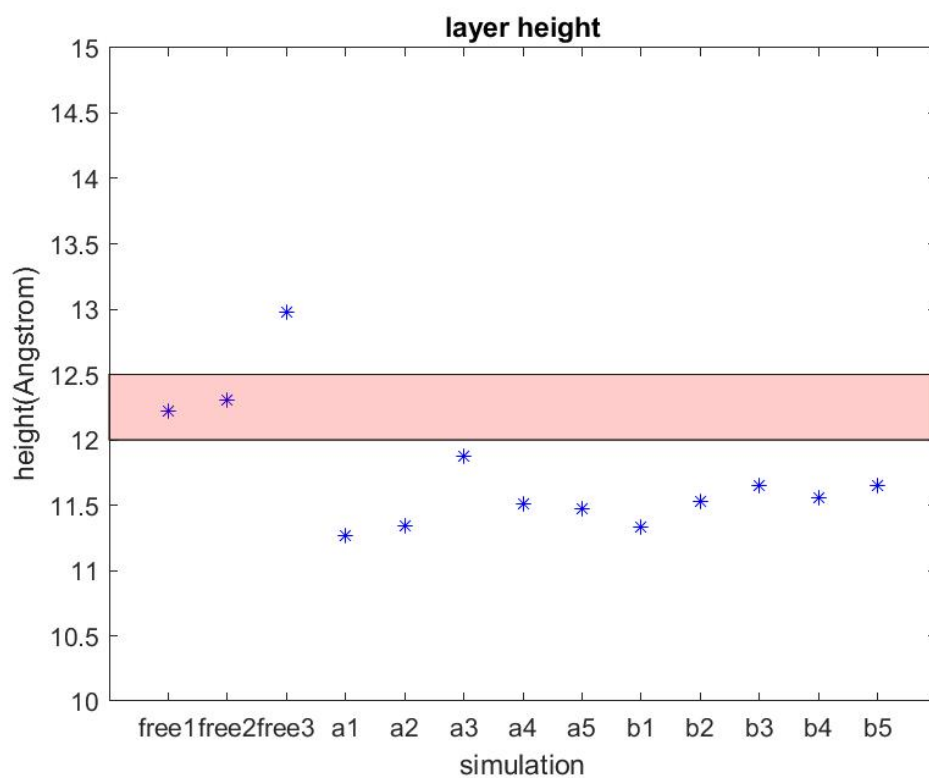
Two types of simulation are performed in part 2. On the one hand, FcC6S molecules are deposited on the gold surface either one by one or as a droplet. For the droplet two different force field models are adopted and the results compared. In this case, the models feature bonds allowing the movement of thiols on the gold surface, based on the nondirectional Morse potential (eq. 42). However, since evidence appears to suggest that a strong directional bond is involved in the thiol bond[40][41], for a second round of experiments the molecules are anchored on the surface in predetermined positions, according to two different periodic patterns, previously referred to as "pattern 1" and "pattern 2" in figure 21 of section . For each of those patterns both the top and the bridge configuration and with or without fixed bond angles are simulated. All the average results regarding head tilt angle, chain tilt angle and monolayer height have been plotted to compare them to the experiments. While the average head tilt angle produces reasonable values and the monolayer height tends to show a slight underestimation, the tilt angle of the chains is not effectively reproduced, with results about 10 or more degrees greater than the 30° expectation result for thiol chains bound to a gold substrate. Thus, while the results obtained provide an encouraging starting point that partly matches the experimental results, there are still several challenges to overcome. More experimental measures are needed to improve the confidence in the validation of the results. Moreover, the force field parameters need further refinement to better represent real systems, especially the Au-S bond model, for which force fields generally use very simplified representations. The bonding nature combined with the presence of differently coordinated sites (top, bridge, three-fold hollow) means that an accurate model for the anchoring should account for bond order, resulting in a non-trivial programming challenge to integrate the source code for the force field within the LAMMPS platform.



(a) head tilt angles



(b) chain tilt angles



(c) layer height

Figure 32: Plot reporting the average values for the head and chain tilt angles and the layer heights compared to the theoretical values. The models "free1" and "free2" are model 1 and 2 for the force field of the molecules. "free3" is the model where the molecules are deposited one by one. "a" refers to the first and "b" to the second geometry. 1 to 3 refer to the top configuration with 100°, 108° and free Au-S-C angle. 4 and 5 to the bridge configuration with fixed and free angle.

Part 3: deposition of 6(ferrocenyl)hexanethiol on trenches and wires

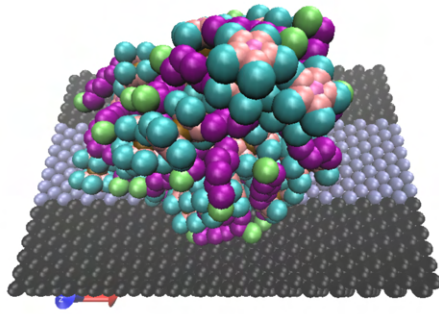
In all the cases considered in this section, a second substrate material is added and the gold is shaped as a wire embedded within. For the new material the same lattice constant as gold is used as a simplification to make the geometry more manageable. The experiments use for the mass and Lennard-Jones parameters of this material those of silicon with $\epsilon = 0.5997$ KJ/mol and $\sigma = 3.92\text{\AA}$ and Lorentz-Berthelot combination rules for the interaction with other atoms. All the other force field parameters are those from the first model for the FcC6S in section

In all the simulations NVT dynamics is run with a velocity-Verlet integrator and a Nosé-Hoover thermostat, with a timestep of 1 fs unless otherwise specified. The substrate atoms are frozen in their lattice positions.

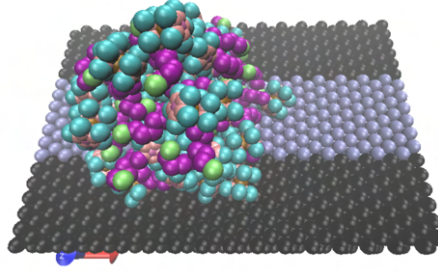
Deposition of 6(ferrocenyl)hexanethiol on a large wire

Using the same deposition process as the previous droplet experiment, a droplet of 40 FcC6S molecules is suspended on a substrate, constituted by three atomic layers. Those represent a surface formed by silicon surrounding on the side a gold wire, 1.73 nm in width and 7.99 nm in length. Two different situations are studied by running NVT dynamics for 20 ns. In the first case, the simulation includes the attractive component of the Lennard-Jones interaction between the molecules and the supporting material. In the second case, such component is removed to simulate a non-interacting material.

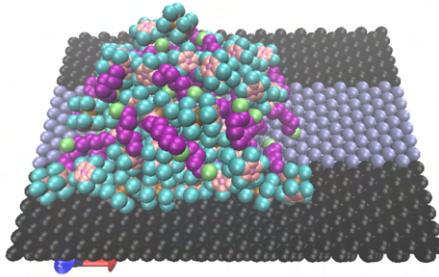
In the interactive case, the ferrocene groups stick to the silicon and the sulfur binds the wire. As a result, the molecules lie horizontally as shown in fig. 33. In the second case, the molecules are tightly packed on the wire in an upright position. Some of them at the side only manage to bind with the substrate by bending diagonally because of the steric hindrance of the ferrocene group, as shown in fig. 35.



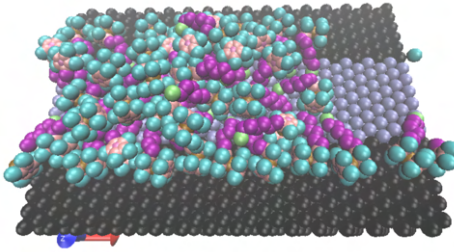
(a) 0 s



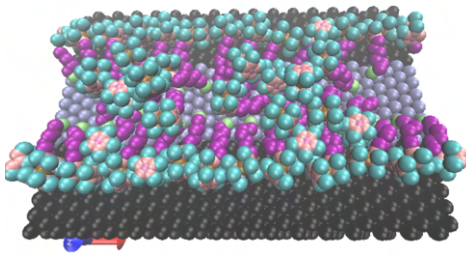
(b) 250 ps



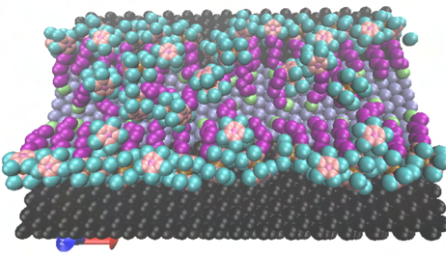
(c) 1 ns



(d) 2.5 ns

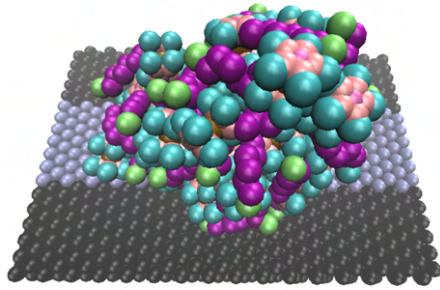


(e) 5 ns

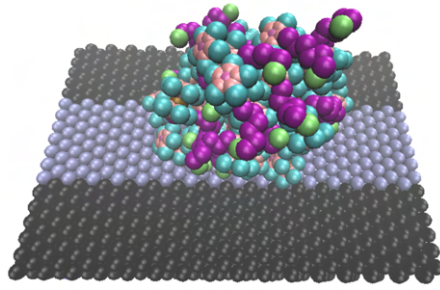


(f) 20 ns

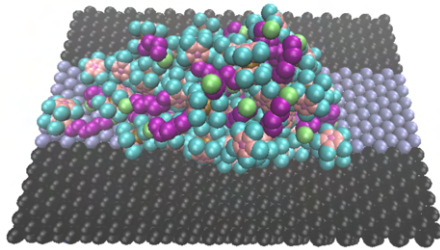
Figure 33: Deposition of 40 FcC6S on a 1.7 nm gold wire. Attractive interactions are included between the substrate and the headgroups



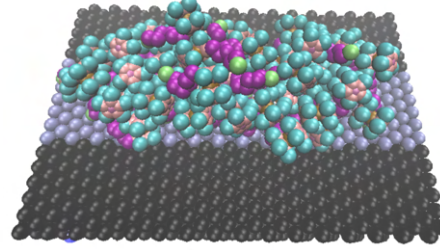
(a) 0 s top view



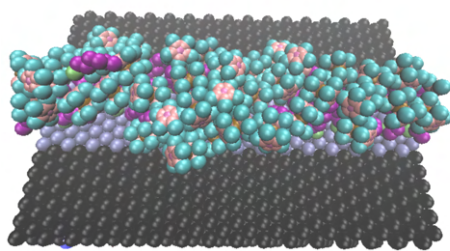
(b) 250 ps top view



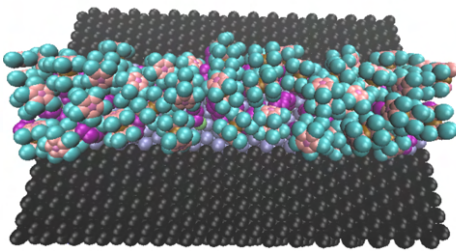
(c) 1 ns top view



(d) 2.5 ns top view

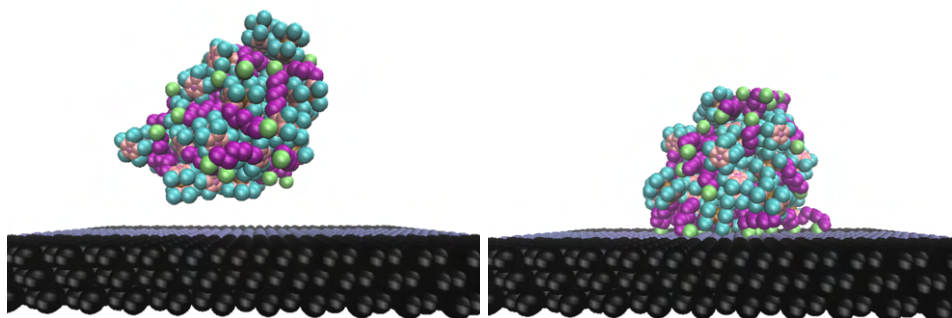


(e) 5 ns top view



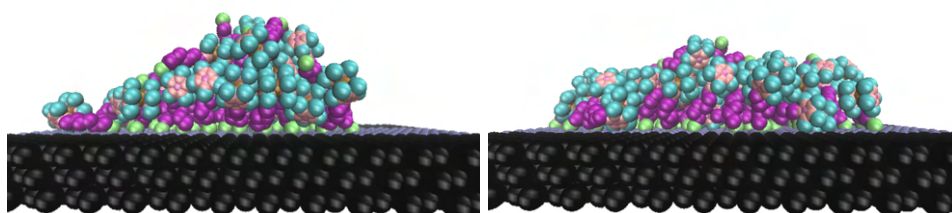
(f) 10 ns top view

Figure 34: Deposition of 40 FcC6S on a 1.7nm wide gold wire, top view. Attractive interactions are neglected between the silicon-like substrate and the headgroups



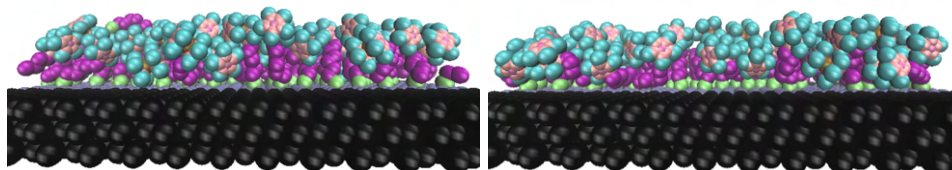
(a) 0 s side view

(b) 250 ps side view



(c) 1 ns side view

(d) 2.5 ns side view



(e) 5 ns side view

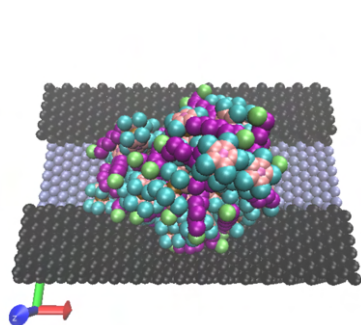
(f) 20 ns side view

Figure 35: Deposition of 40 FcC6S on a gold wire, side view. Attractive interactions are neglected between the silicon-like substrate and the headgroups

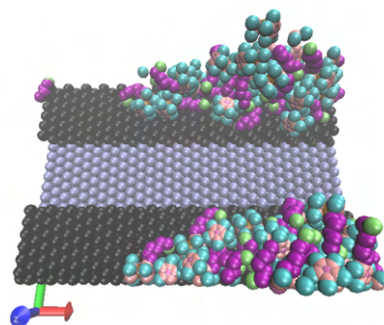
Deposition of 6(ferrocenyl)hexanethiol in a shallow trench

40 FcC6S molecules are deposited in droplet form onto a gold wire of 1.7 nm width and 7.99 nm length, confined in a shallow trench 0.5 nm deep made with the silicon-like material. The sidewall atoms have the same lattice constant as gold for simplicity and dynamics is run once for 20 ns neglecting the interactions of the FcC6S with the sidewalls and once again for 40 ns including them. In the case of the interaction, most of the molecules tend to stick outside the trench and can't easily diffuse inside. the end result is that the trench is not fully occupied by 40 ns and any additional molecules need more time to enter the more the sidewall coverage increases.

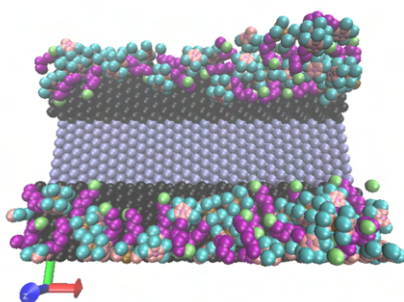
If the attractive interaction with the sidewalls is neglected instead, the molecules enter the trench and begin forming an ordered self-assembly.



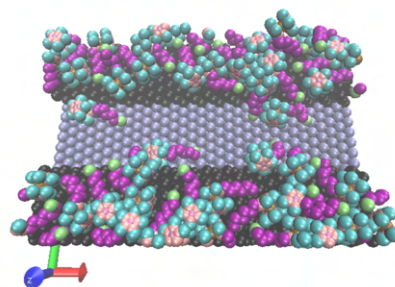
(a) 0 s



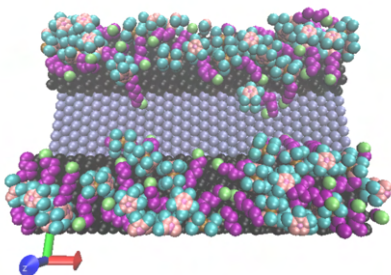
(b) 250 ps



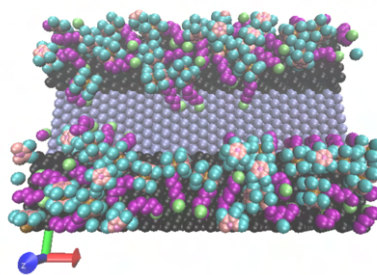
(c) 1 ns



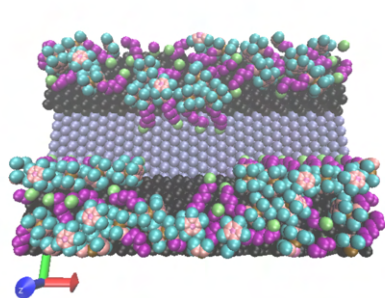
(d) 2.5 ps



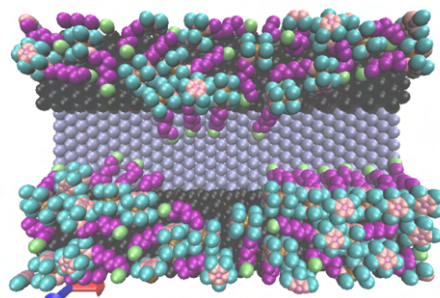
(e) 5 ns



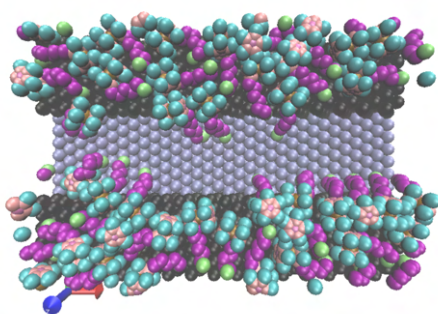
(f) 10 ns



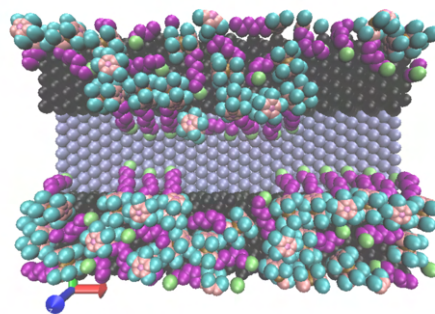
(g) 20 ns



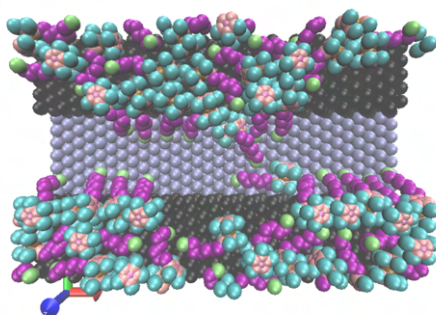
(h) 25 ns



(i) 30 ns

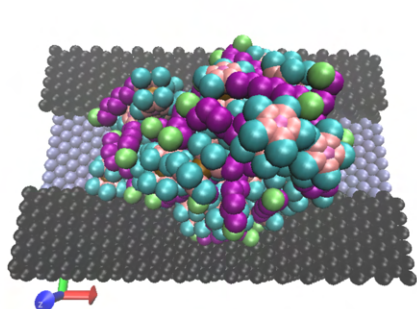


(j) 35 ns

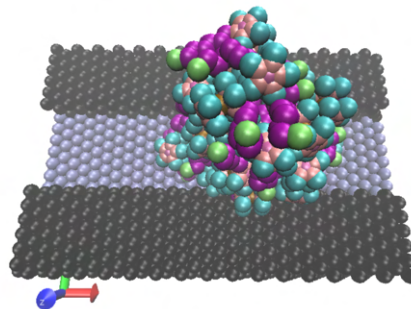


(k) 40 ns

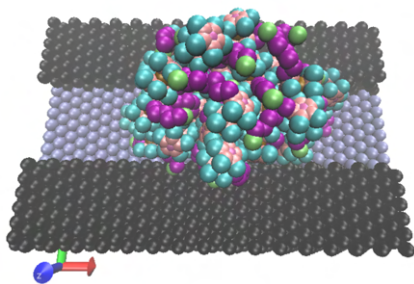
Figure 36: Deposition of 40 FcC6S on a gold wire in a trench. Attractive forces are included between the molecules and the sidewalls



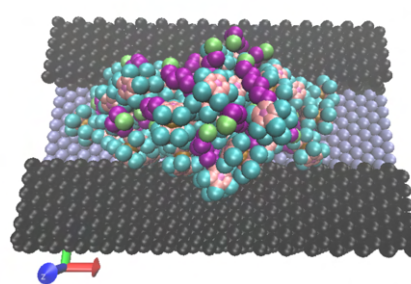
(a) 0 s



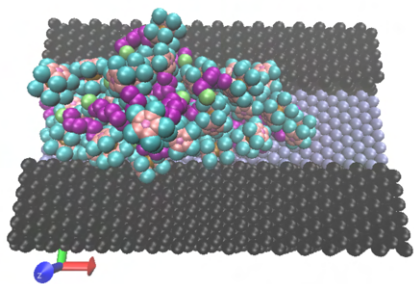
(b) 250 ps



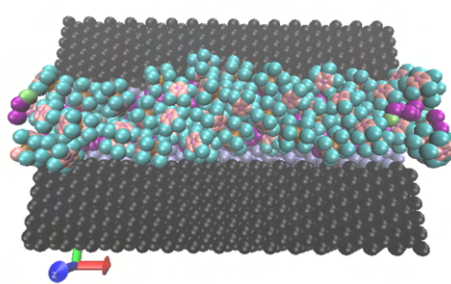
(c) 1 ns



(d) 2.5 ns



(e) 5 ns



(f) 20 ns

Figure 37: Deposition of 40 FeC6S on a 1.7 nm gold wire in a trench. Attractive forces are neglected between the molecules and the sidewalls

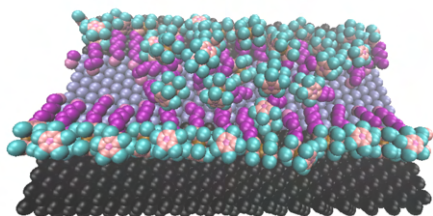
6(ferrocenyl)hexanethiol on a wire: oxidation and electric field application

1.7 nm wire

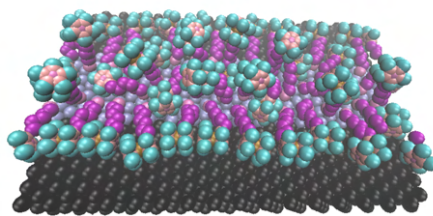
Starting from the simulation at the previous point, where the molecules were deposited on an interacting wire, the following steps are executed after removing the molecules that are not bound to the substrate:

- the per-atom charge of the deposited molecules is replaced with the charge computed by ab-initio DFT simulation on the software Orca assuming a +1 net charge on the whole 6(ferrocenyl)hexanethiol molecule. Due to the electropositivity of iron, the positive charge is almost totally located on the head and thus repulsion pushes the heads away from each other. NVT dynamics is run for 5 ns and, as shown in figure 39, a few molecules sticking to the border are pushed into an upright position and moved to the middle of the wire. Thus the configuration changes from only having a few molecules occupying the middle of the wire while the bulk of them are stuck to the border, to having thiolated ferrocene molecules populating the middle of the wire.
- A 2 V/nm electric field is applied pointing toward the positive z direction and dynamics is run for 6 ns, as shown in figure 41. The field causes most of the molecules to lift in an upright position, distributing quite uniformly over the surface of the wire. However not all the molecules lift because of hindrance and so even after 6 ns of dynamics, when the system stops evolving, several molecules remain in a horizontal position, spread regularly across the wire.
- The electric field is removed and the system is relaxed by running NVT dynamics for 6 ns. Most of the molecules relax by occupying again a horizontal position with the heads sticking to the substrate. However a few of them remain upright or lying down in the center of the wire, similarly to the system before the application of the electric field.

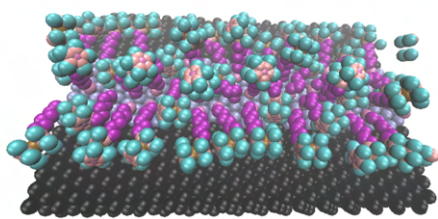
Oxidation



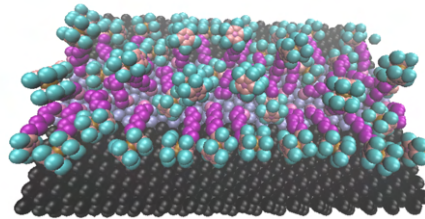
(a) 0 s top view



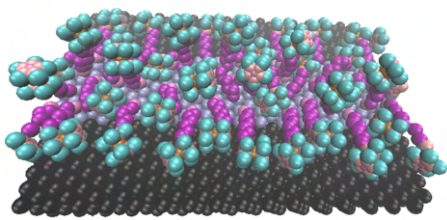
(b) 250 ps top view



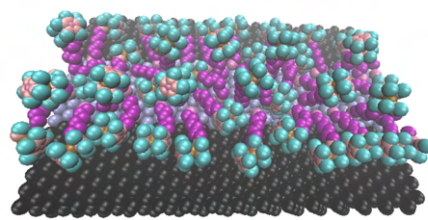
(c) 1 ns top view



(d) 2.5 ns top view

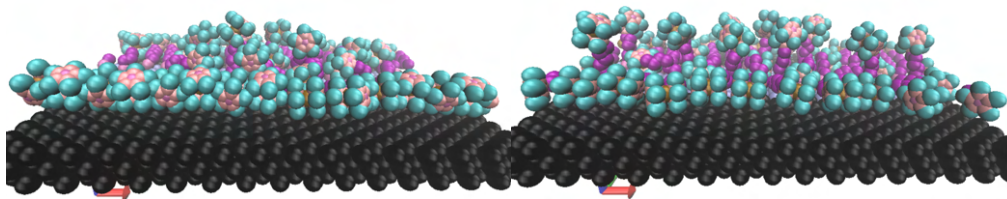


(e) 5 ns top view



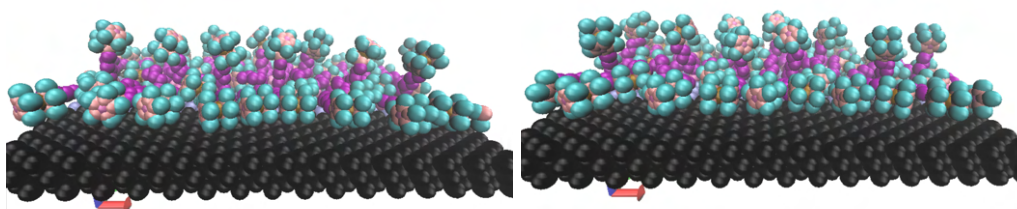
(f) 10 ns top view

Figure 38: FcC6S, 1.7nm wide, oxidized with a +1 charge, on a gold wire, top view.



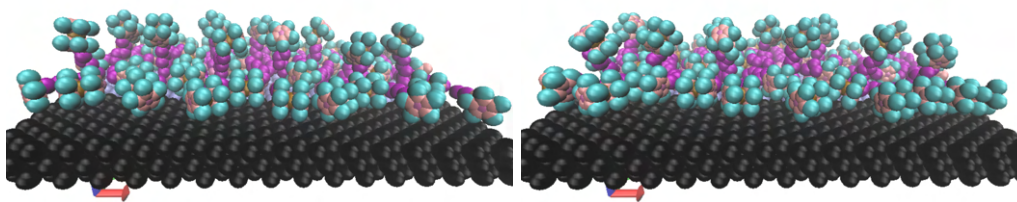
(a) 0 s side view

(b) 250 ps side view



(c) 1 ns side view

(d) 2.5 ns side view

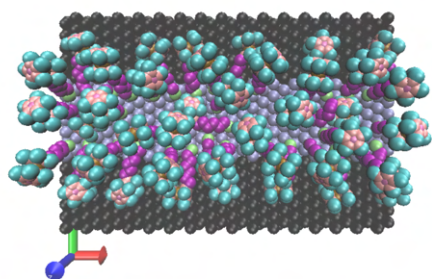


(e) 5 ns side view

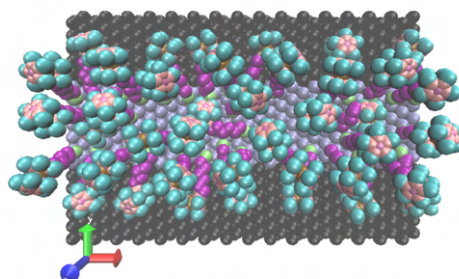
(f) 10 ns side view

Figure 39: FcC6S, oxidized with a +1 charge, on a 1.7 nm wide gold wire, side view.

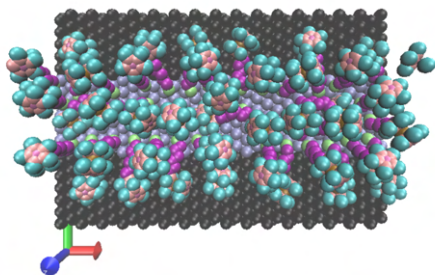
Applying an electric field



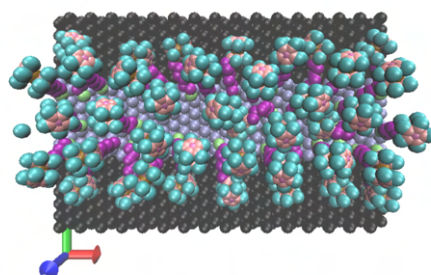
(a) 0s top view



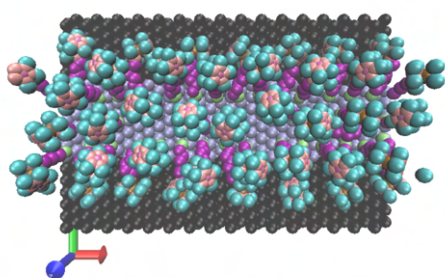
(b) 250 ps top view



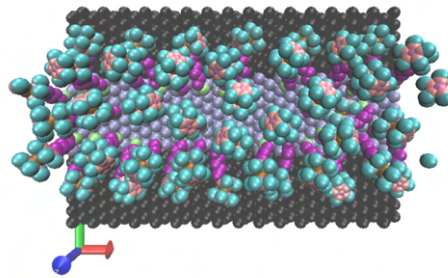
(c) 500 ps top view



(d) 1 ns top view

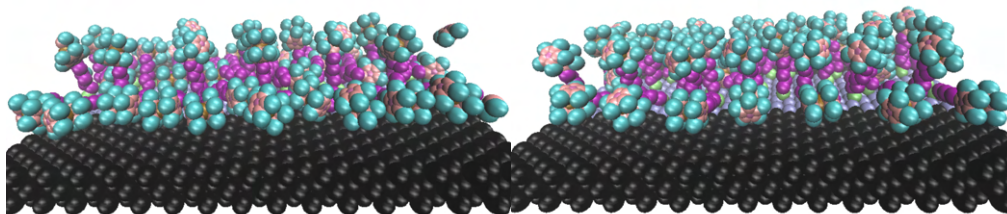


(e) 3 ns top view



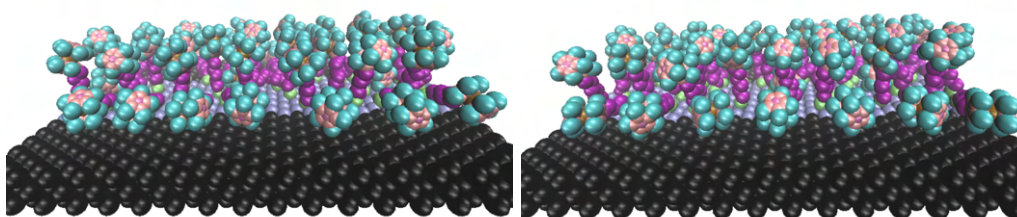
(f) 6ns top view

Figure 40: FcC6S, oxidized with a +1 charge, on a 1.7 nm gold wire, top view. The molecules are subjected to a 2 V/nm electric field.



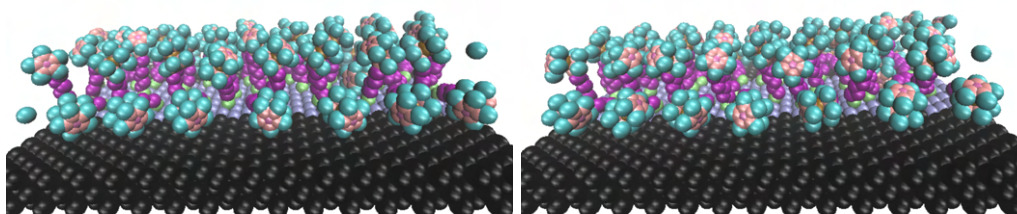
(a) 0 s side view

(b) 250 ps side view



(c) 500 ps side view

(d) 1 ns side view

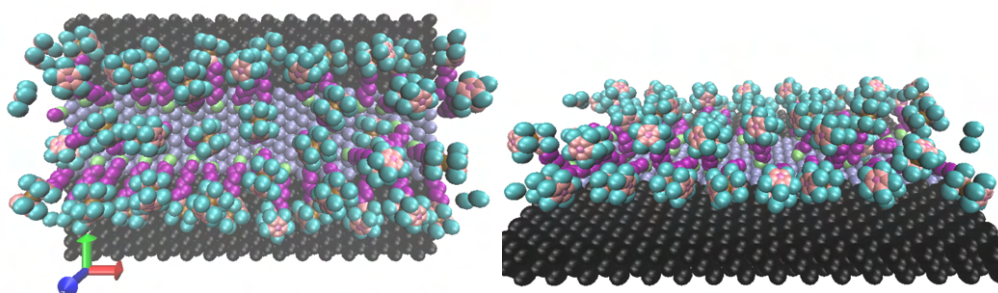


(e) 3 ns side view

(f) 6 ns side view

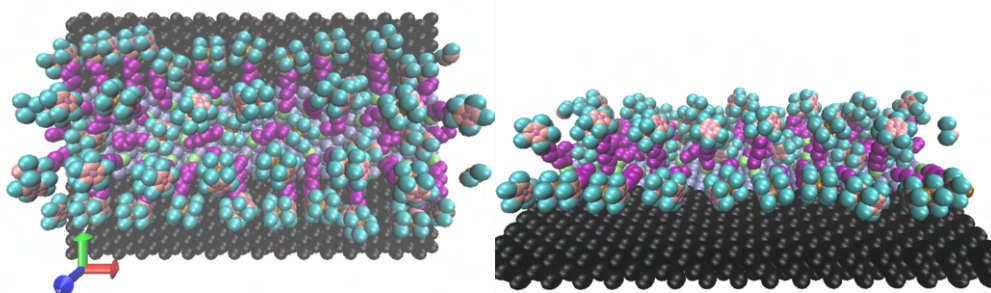
Figure 41: FcC6S, oxidized with a +1 charge, on a 1.7 nm gold wire, top view. The molecules are subjected to a 2V/nm electric field.

Removing the electric field



(a) 0 s top view

(b) 0 s side view



(c) 6 ns top view

(d) 6 ns side view

Figure 42: FeC6S, oxidized with a +1 charge, on a 1.7 nm gold wire. Relaxation after the removal of the 2 V/nm electric field.

1.15 nm wire

The previous steps have been repeated. However, the wire is reduced in width by one surface lattice cell and thus the total width becomes 1.15 nm. This means that the wire can't host more than two rows of molecules due to hindrance.

- The deposition stage is the same as the 1.7 nm wire, with a thiol droplet suspended above the substrate. Already, at 250 ps, one can see some molecules arranging themselves horizontally in two rows. After a total time of 15 ns the molecules arrange themselves in an orderly way, with the molecule lying down with the head bound to the support material outside the wire.
- The molecules failing to bind to the substrate are removed and 25 molecules remain. When the charges are replaced with the oxidized charge, some of the molecules are lifted due to the repulsion acting among the heads and by 10 ns they arrange themselves in the center of the wire.
- With the application of the electric field part of the molecules lift and part stay lying horizontally due to hindrance, approximately in an alternating pattern.
- After the removal of the electric field the system relaxes back to the configuration with most atoms lying down horizontally and a few, four of them in this case, standing upright.

Deposition

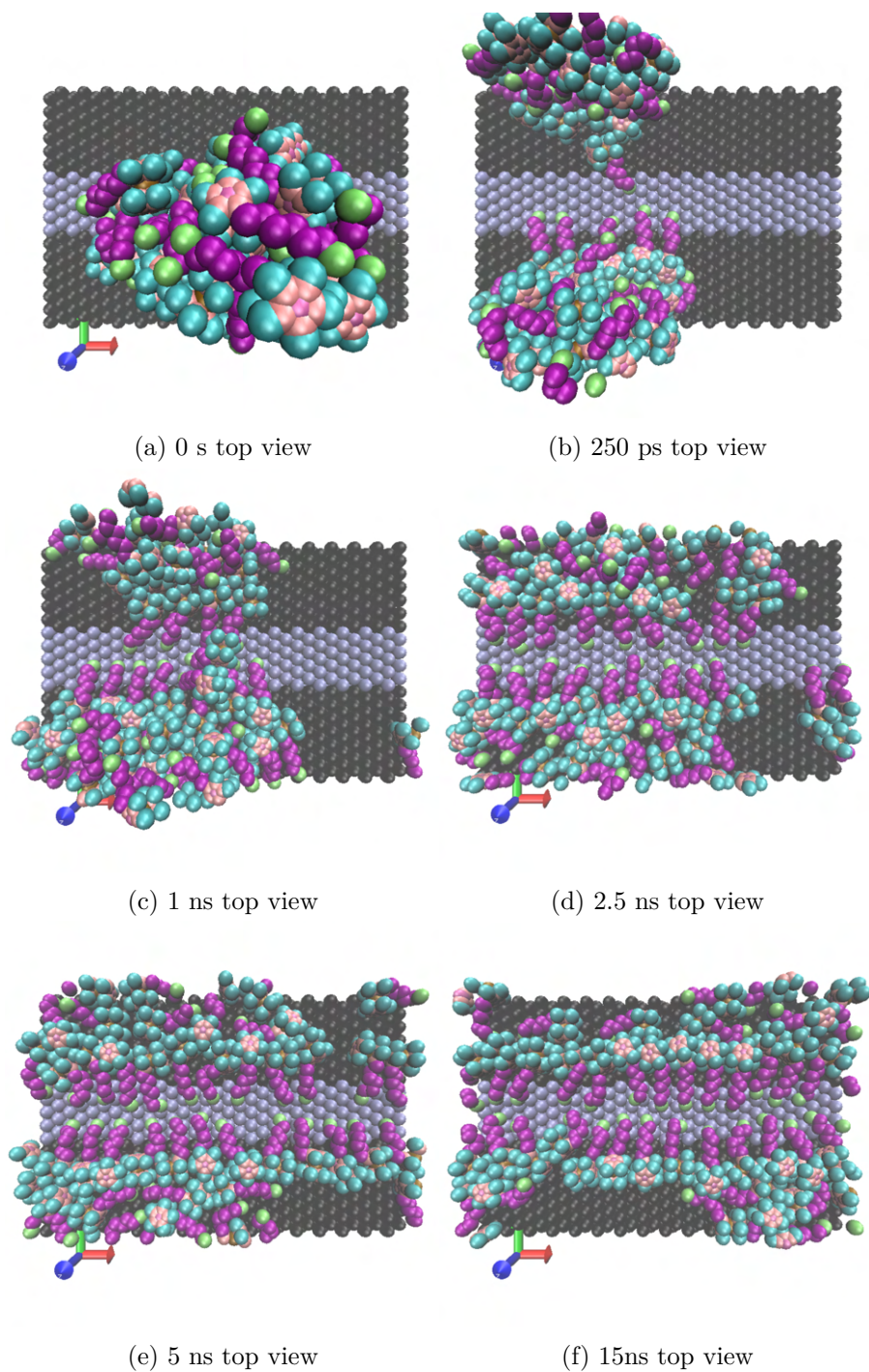


Figure 43: FcC6S deposited on a 1.15nm wide gold wire, top view.

Oxidation

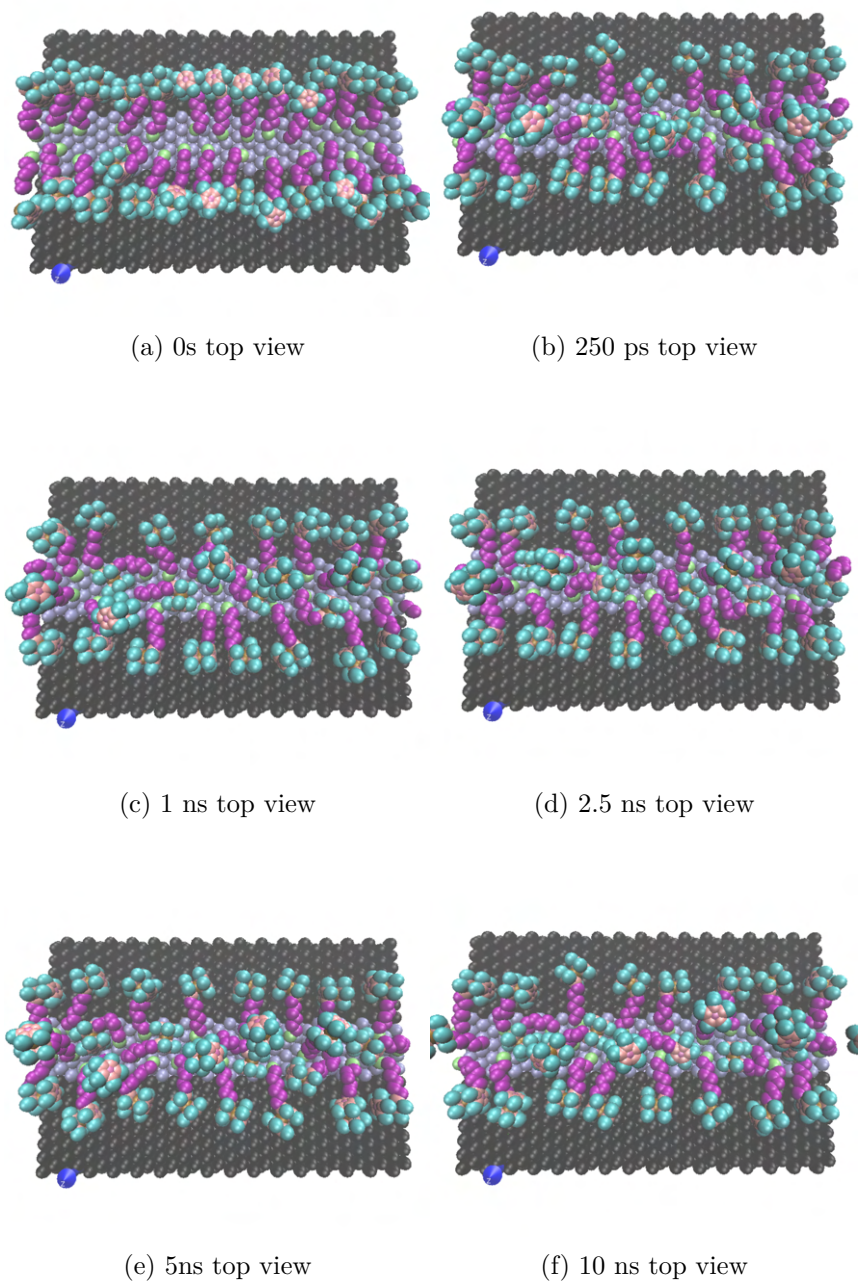
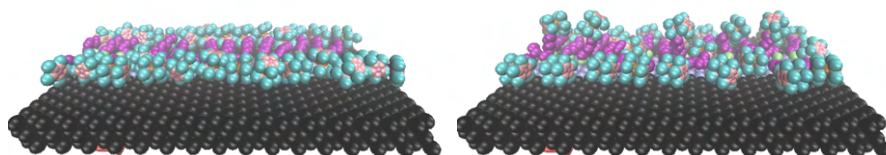
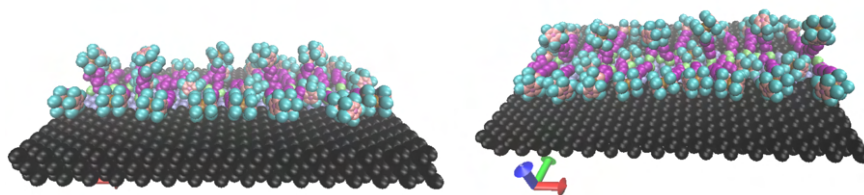


Figure 44: FcC6S, deposited on a 1.15 nm wide gold wire, and oxidized (+1 charge), side view.



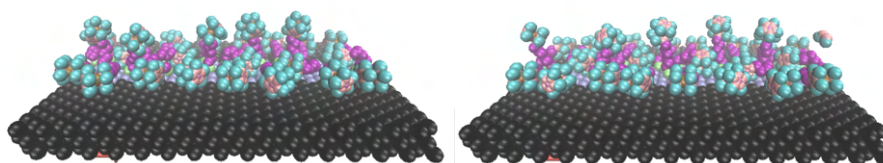
(a) 0 s side view

(b) 250 ps side view



(c) 1 ns side view

(d) 2.5 ns side view



(e) 5 ns side view

(f) 10ns side view

Figure 45: FcC6S, deposited on a 1.15 nm wide gold wire, and oxidized (+1 charge), side view.

Applying an electric field

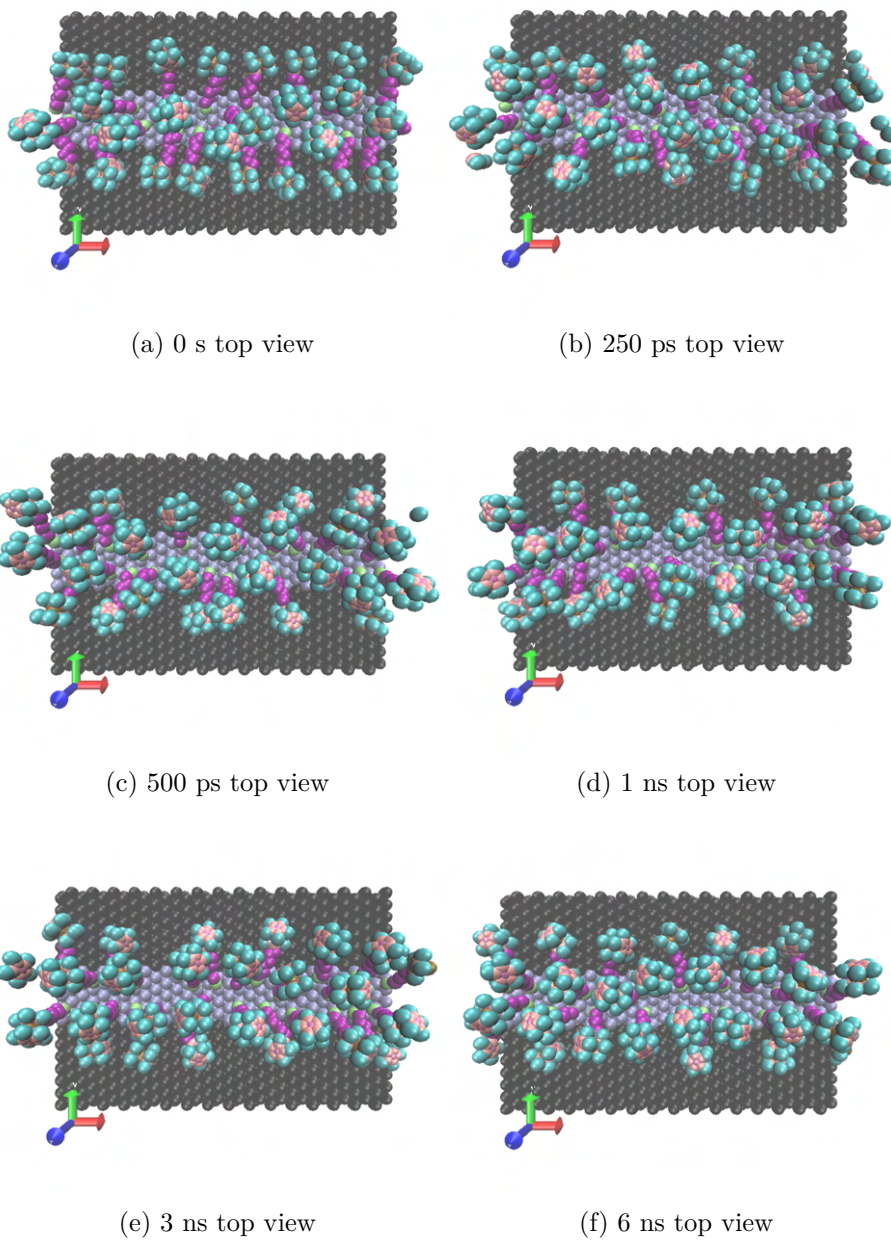


Figure 46: FcC6S, deposited on a 1.15 nm wide gold wire, oxidized (+1 charge) and subject to a 2V/nm electric field pointing upwards. top view.

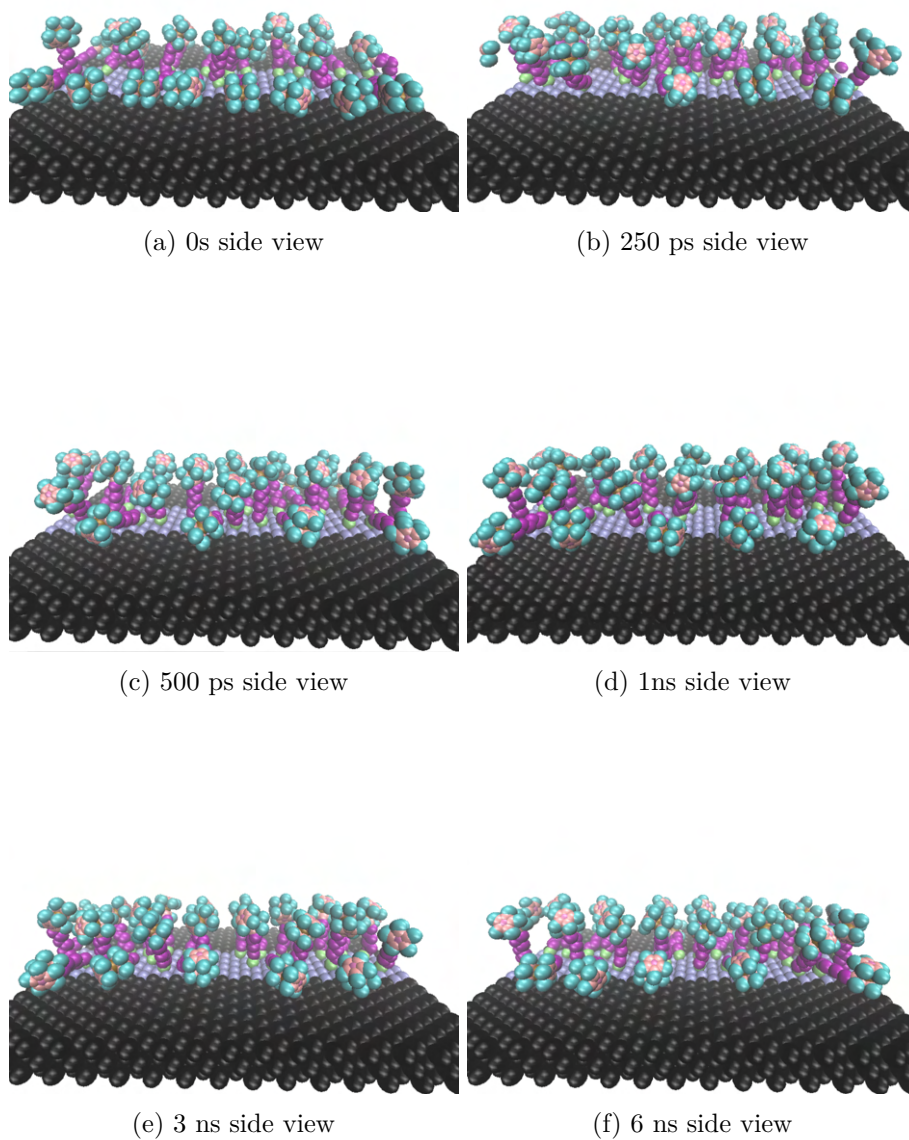


Figure 47: FcC6S, deposited on a 1.15 nm wide gold wire, oxidized (+1 charge) and subject to a 2 V/nm electric field pointing upwards. side view.

Removing the electric field

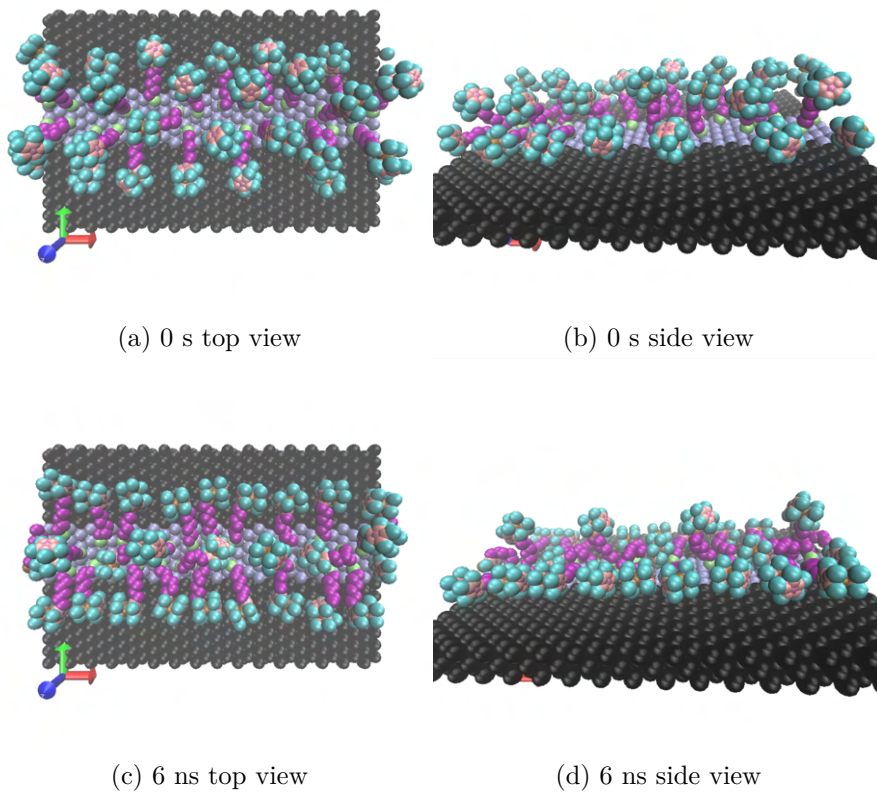


Figure 48: FcC6S, oxidized with a +1 charge, on a 1.1 nm gold wire. Relaxation after the removal of the 2 V/nm electric field, side view.

0.58 nm wire

The steps above have been again reproduced with a wire of gold of 0.58 nm width in the x direction, corresponding to a single cell. Regardless of the achievability of such a resolution in real systems, the following can be observed:

- Similarly to the previous case the molecules manage to deposit horizontally, alternating between the two sides, in this case, because of the narrower space available for anchoring. After just 5 ns the whole length of the wire is occupied by the deposited molecules.
- After removing the molecules not bound to the wire, 30 molecules are left. Then substituting the original charge with the values after oxidation, a few molecules lift occupying the middle of the wire due to the coulombian repulsion.
- The application of a 2 V/nm electric field forces the molecule to lift. However approximately half tends to remain horizontal in an alternating pattern.
- The removal of the electric field restores the configuration existing prior to its application.

Deposition

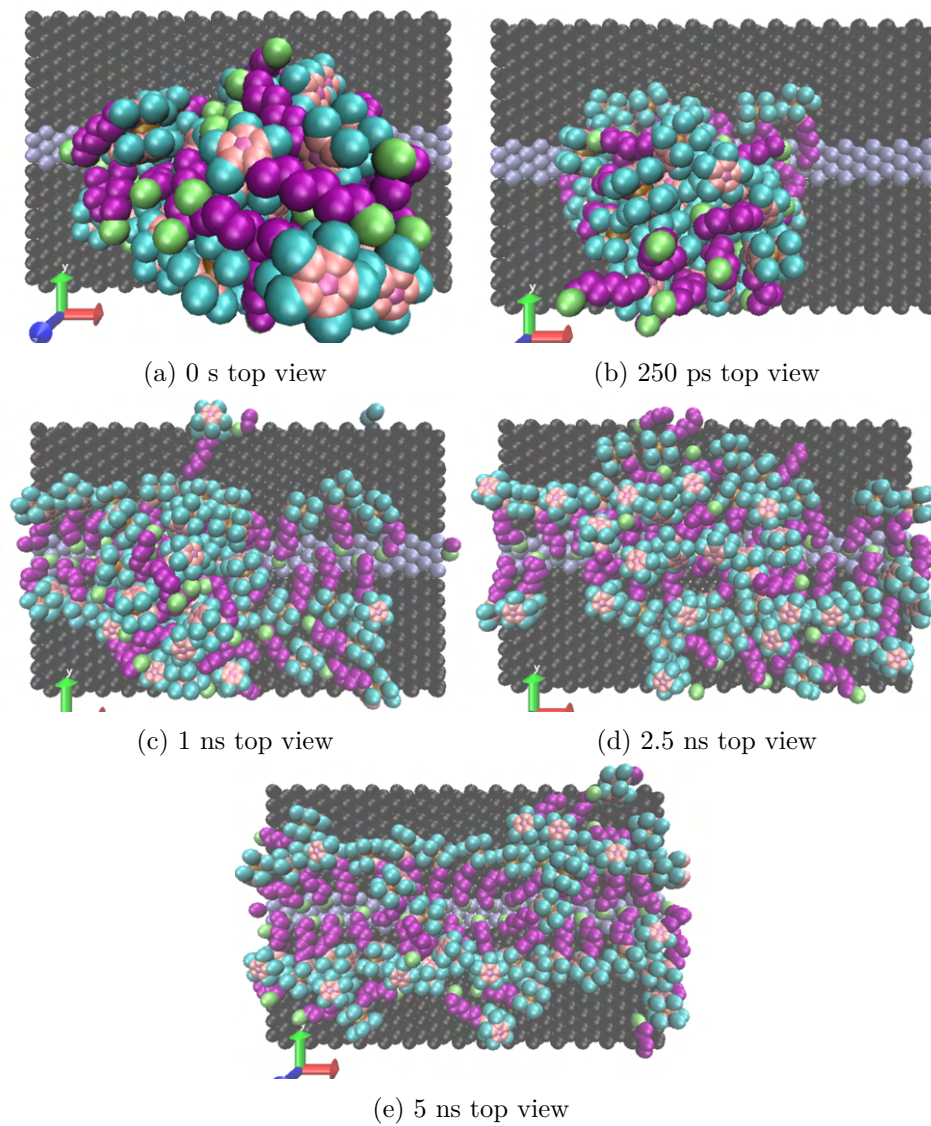


Figure 49: FcC6S, deposited on a 0.57 nm wide gold wire, top view.

Oxidation

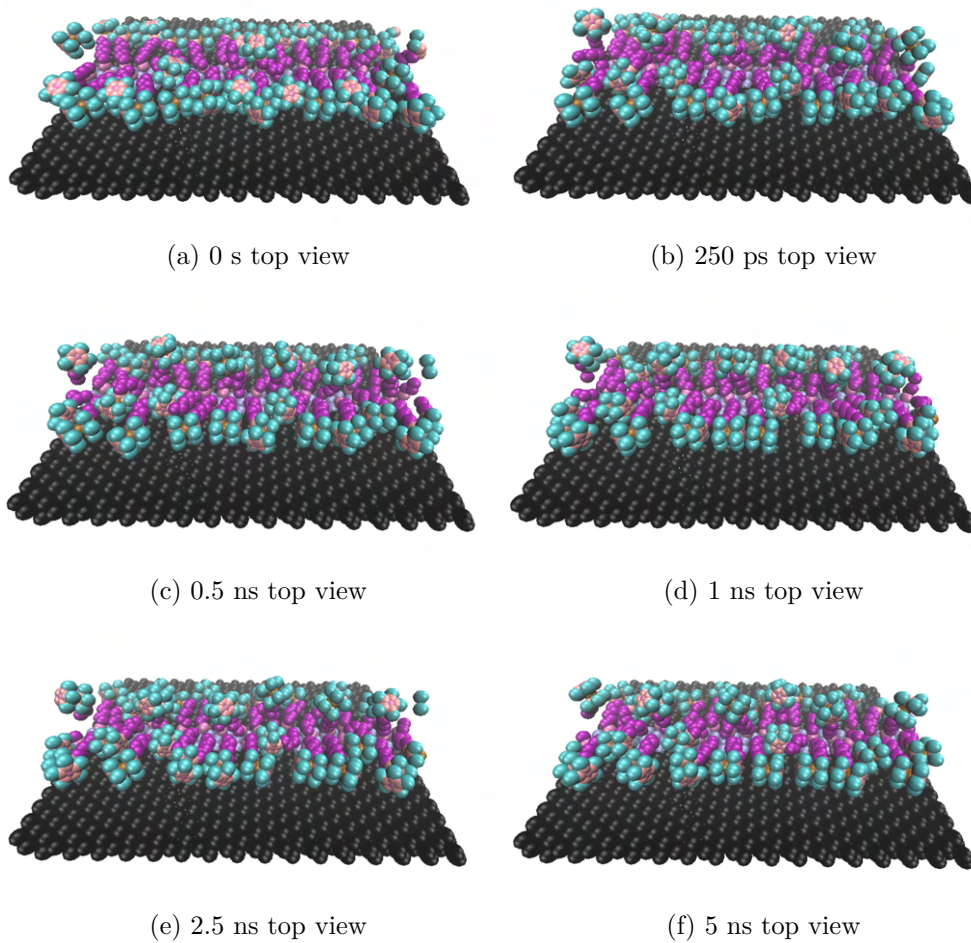


Figure 50: FcC6S, deposited on a 0.57 nm wide gold wire and oxidized (+1 charge). top view.

Applying an electric field

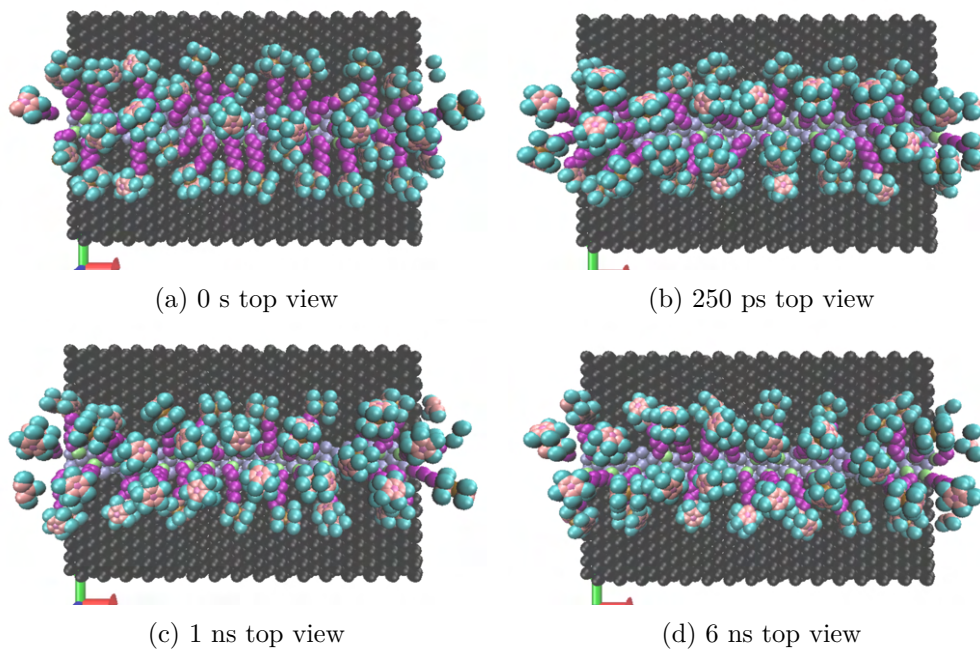


Figure 51: FcC6S, with the head oxidized (+1 charge) on a 0.57 nm wide gold wire, with a 2 V/nm electric field, top view.

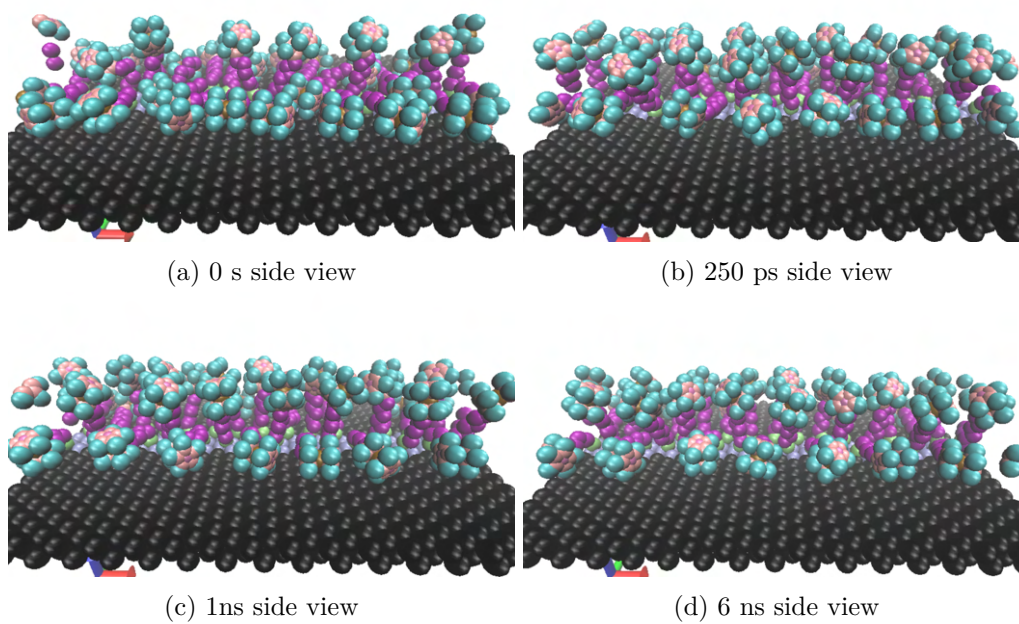


Figure 52: FcC6S, with the head oxidized (+1 charge) on a 0.57 nm wide gold wire, with a 2 V/nm electric field, side view.

Removing the electric field

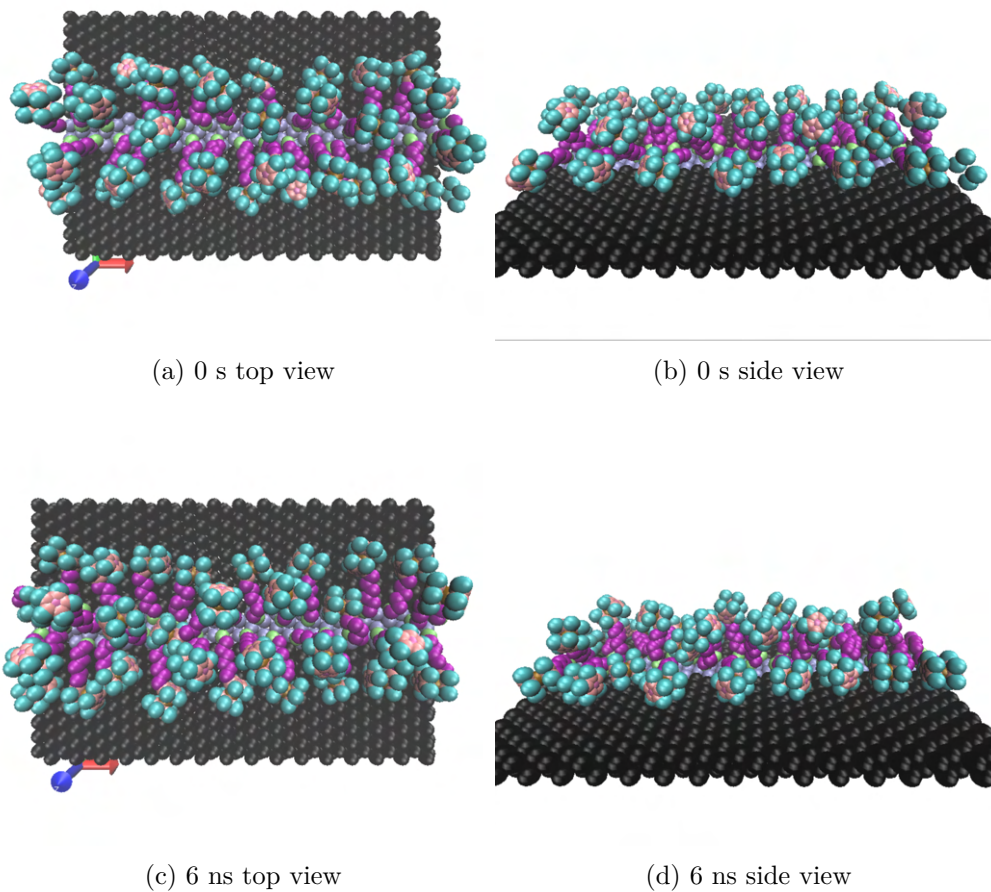


Figure 53: FeC6S, oxidized with a +1 charge, on a 0.58 nm gold wire. Relaxation after the removal of the 2 V/nm electric field.

Obtaining a wire by etching a 2D SAM

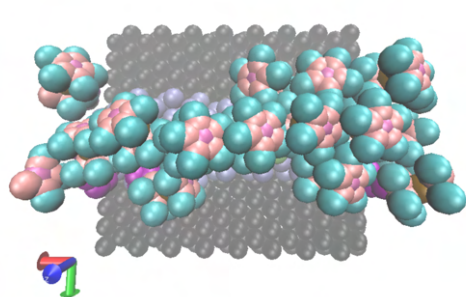
Having the molecules drop coat directly a surface with a patterned wire is impractical from one point of view: since the heads are responsible for most of the steric hindrance and the potential between the sulfur and the gold has been set to a strong bonded potential as if the chemisorption has already occurred, it's easy for the sulfur headgroups to be captured by the gold and the molecules lying horizontally occupy the gold wire with a higher density than they could maintain by remaining upright.

A different attempt to create a gold wire can be a deposition of a 2D SAM on a gold surface followed then by an etching. In this case the etching is simulated by adding a silicon support for the gold wire then removing all gold and molecules outside of a strip 1.15 nm wide in the x direction and repeating the above steps after running NVT dynamics to equilibrate the new wire. A difference from before is that this 2 layers thick strip lies above the support instead of being coplanar to it.

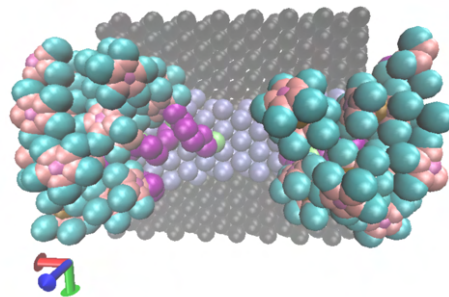
- In order to equilibrate the new wire, nvt dynamics is been run after the etching for 20 ns. The resulting wire, however, does not seem to be stable and at different time periods one can observe the opening and closing of gaps. For example at 1 ns a gap is open on the wire and at 3 ns the wire is continuous. However, this continuity is broken and reformed multiple times: for example the wire is gapped at 13.5 ns, continuous at 14 ns, gapped at 15ns and 18ns and continuous at 16 and 20 ns.
- After that the charge is replaced with the oxidized values for the 6(ferrocenyl)hexanethiol. With a suspended strip, however, the repulsion is enough to push the thiol chain inside the gold strip, embedding it. This behavior may show the limit of a united atom potential for the thiol chain. While the Lennard Jones parameters for the CH₂ group are chosen to simulate correct force values at a long enough range, at short range this model doesn't account for the greater hindrance caused by the hydrogen and carbon model acting as separate repulsion centers. Thus it may happen that the forces generated by the interatomic repulsion are enough to push this atom wide chain through the minima in the surface potential of the gold, given also that a lifted strip has edges where there are fewer repulsive contributions. Since this phenomenon persists at very small timesteps of the order of 0.05 fs, one can exclude a too large timestep as the cause of this phenomenon. This becomes more evident if one relaxes the frozen substrate constraint: in that case a movement of the gold atoms can be observed, with the sulfur heads sinking below the first gold layer. Thus it appears that this structure is not stable under the effect of strong local forces.

This suggests that the issue might be that the gold potential used in these experiments is inadequate to represent correctly lowly coordinated structures such as small, limited size gold wires[49]. Moreover freezing the gold structure is not realistic because the forces acting at the thiol-gold interface have a strong perturbative effect on gold nanostructures. A correct representation of that bond would require a model depending parametrically both on the coordination of the sulfur and that of the gold atoms. A model with these characteristics is the one proposed by Olmos-Asar et al.[50]. However such a potential does not have a Lammmps implementation

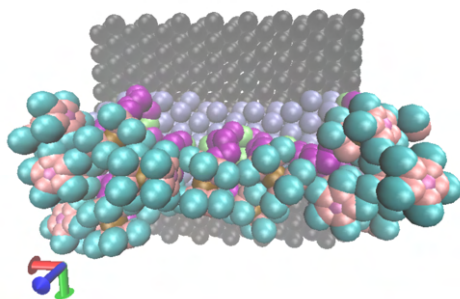
and any implementation endeavour is not trivial, requiring more than basic C++ programming expertise.



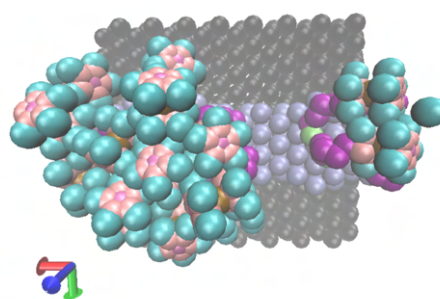
(a) 0s top view



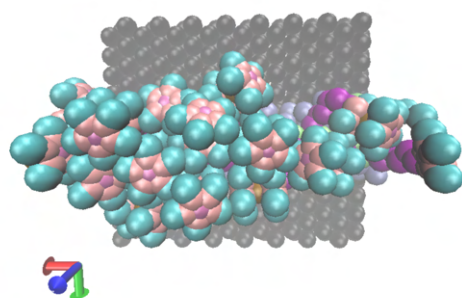
(b) 1ns top view



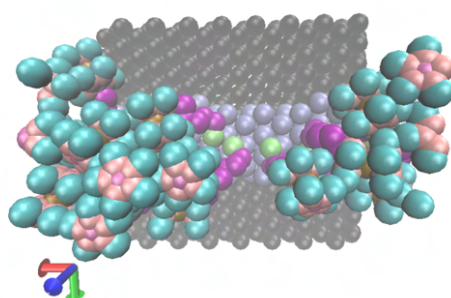
(c) 3ns top view



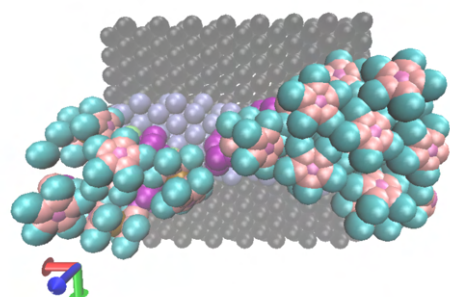
(d) 13.5ns top view



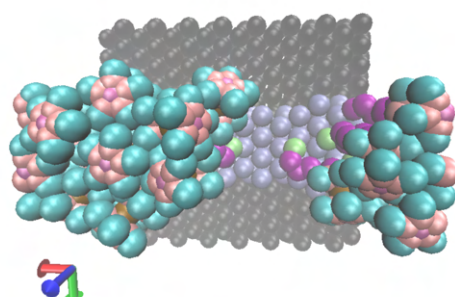
(e) 14ns top view



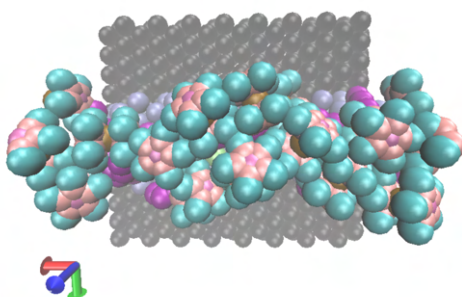
(f) 15ns top view



(g) 16ns top view



(h) 18ns top view



(i) 20ns top view

Figure 54: FeC6S, on a lifted gold wire obtained from cutting a 1.1 nm wide strip off the thiol SAM on a gold surface

Oxidation

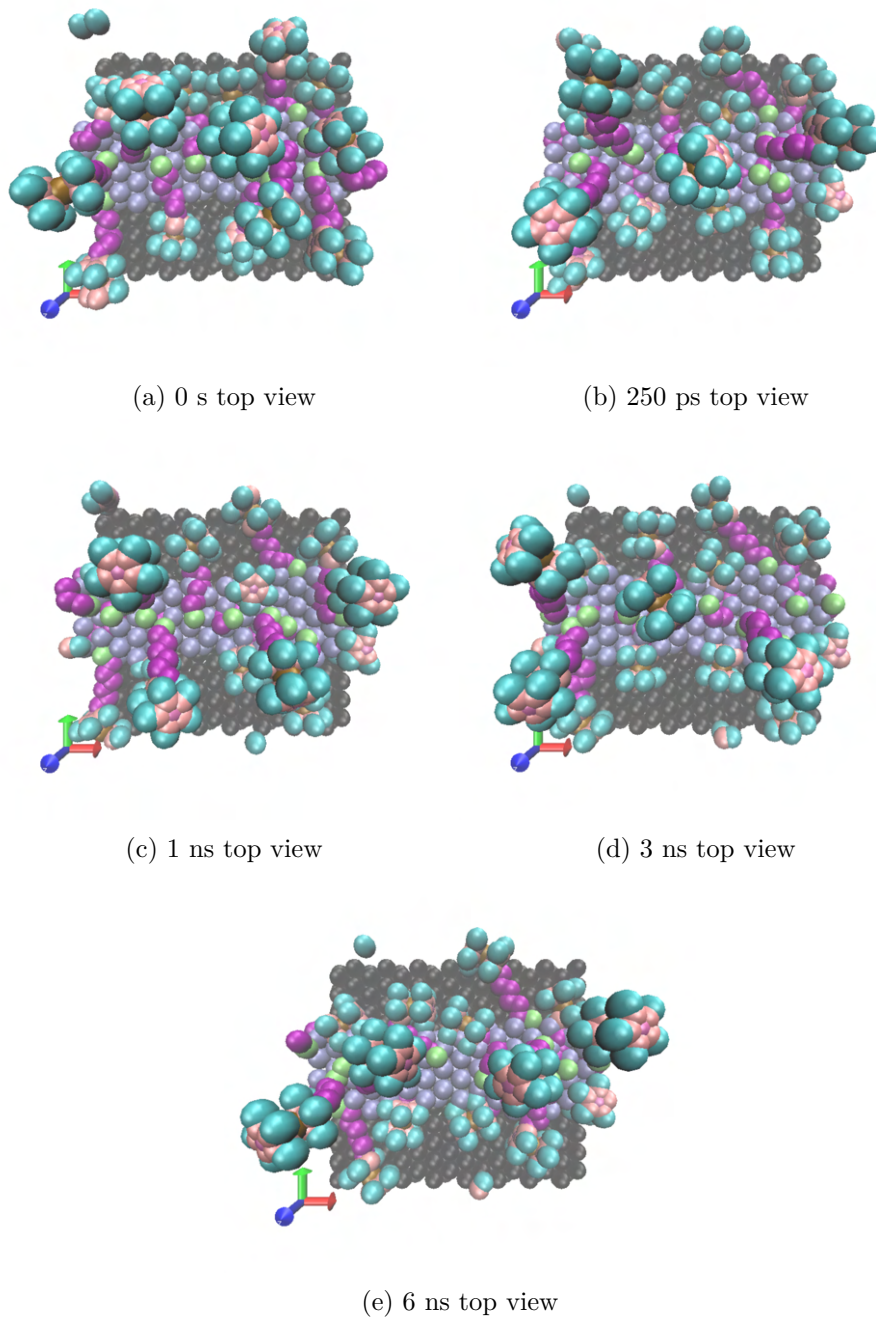


Figure 55: FcC6S, deposited on a 1.1 nm suspended gold wire and oxidized (+1 charge). top view.

Conclusions of part 3

results

The case of the lifted wire, where the lack of a reasonable force field model for the system produces strongly unrealistic results, will not be discussed.

For the remaining wire models what can be observed is the following: if a material with a strong enough interaction with the molecule heads is chosen (silicon might be a valid candidate), the obtained system presents in all cases two states once oxidized and then subjected or not to a 2 V/nm: one with almost all the molecules lying horizontally, except the few that are kept upright by the coulombian repulsion, and one with molecules on the border lifting more or less in an alternating pattern to balance the external field with the coulombian repulsion. The two states are more well defined in the two cases of the narrower wire, where's there's no room in the middle of the wire for hosting part of the molecules so the anchoring point doesn't shift much from the original position when the molecule lifts.

limits

The current results are based on a naive model of the molecule and applied field, with static charges on the molecule atoms and no polarization effects in the molecule the substrate. Thus it can provide a first preliminary idea on the molecules' behavior, but to obtain a more accurate computation charges should be recomputed with computation models that include at least charge transfer and polarization effects if not full quantum chemical approaches. However, those methods are outside the domain of standard molecular dynamics and require at the very least a hybrid approach with more sophisticated techniques.

Another element as in the previous cases is the simplifications in modeling the S-Au bond with the Morse potential, to which one has to add the fact that in this case the gold substrate is finite in width which is outside of the scope of the model of the S-Au interface. Also, the gold potential is the one calculated for the bulk material. Finally, the substrates are frozen and matched and this prevents rearrangement effects due to mismatch and rearrangement due to the forces acting on the wire atoms. Thus further improvement of the results within the domain of molecular dynamics should go in the direction of improving the force fields, especially for what concerns the gold substrate and its interaction with gold. Moreover experimental results on systems such as this one are crucial to enable a comparison that permits a validation of the simulation and guides the iterative process of improvement of the model and verification of the results that is crucial for the field of molecular dynamics to provide an effective guide for technology.

Deposition of bisferrocene atoms and electrical analysis on Scerpa

Bisferrocene molecule

A bisferrocene molecule[51] is composed of two ferrocene units, acting as redox centers, linked by a carbazole group. The latter is a tricyclic aromatic carbon composed of two benzene rings (C_6H_6) linked together by a pyrrole ring (C_4H_5N) as shown in figure 56. The two ferrocene bond the carbazole on carbon 3 and 6 through one of the two carbons of an ethane molecule (see figure 56). The nitrogen atom is linked to the terminal carbon of a 1-hexanethiol. Figure 57 shows the whole bisferrocene molecule.

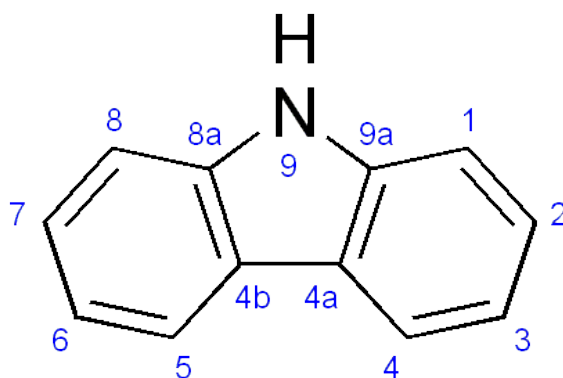


Figure 56: a carbazole group[52]

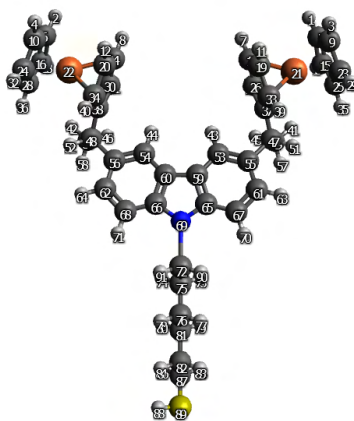


Figure 57: a bisferrocene molecule. Carbons are in grey, hydrogens in white, sulfur in yellow, nitrogen in blue and iron in orange

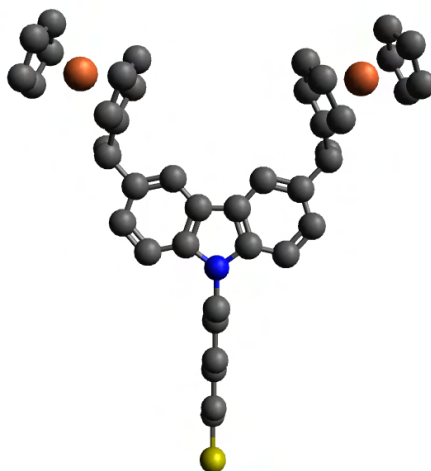


Figure 58: a bisferrocene molecule. Carbons are in grey, hydrogens in white, sulfur in yellow, nitrogen in blue and iron in orange

The geometry of the bisferrocene is subordinate to its functionality: the easy oxidation of the bisferrocene group favors a charge transfer between the bisferrocene units. This allows the molecule to be used in molecular QCA technology: whenever a strong enough field is applied between the bisferrocene units, they form a dipole with a partial negative charge on a ferrocene unit and a partial positive charge on the other. The polarization is opposite to the external field. This means that if the external field is strong enough to saturate the polarization, a '0' or '1' state for binary logic is generated.

If a polarized bisferrocene is put parallel to another bisferrocene molecule, the second dipole arranges itself antiparallely to minimize repulsion. A pair of parallel bisferrocene form a basic molecular QCA cell generating an output field that's parallel to the input. The input information can propagate between successive molecular QCA cells assuming that one can deposit bisferrocene molecules in an ordered way.^[53]

Figure 59 shows a scheme of a molecular QCA: every vertical couple of dots represent the two ferrocenes. When an external dipole driver is applied, the molecules are polarized and the polarization propagates between molecules, inverted on odd positions and with the same direction as the driver on even positions.

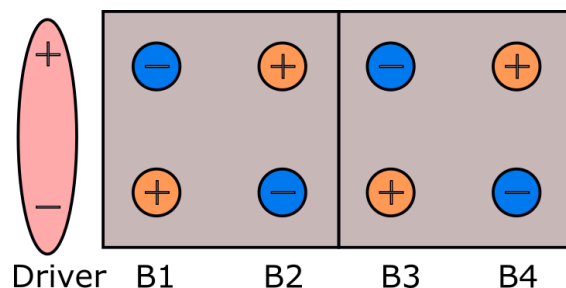


Figure 59: bisferrocene QCA cells with an external driver

Scerpa

Scerpa[54], developed by the VLSI laboratory in Politecnico di Torino, is a Matlab software to compute signal propagation in molecular electronics through a self-consistent algorithm. It consists of three blocks: the layout generator, the self-consistent loop and the viewer.

- The layout generator block allows to specify the geometry. To do so one can insert drivers, a block applying static dipole electric field, and molecules, represented as lumped field-dependent point-like charges, with their relative position and response to an external field specified in an input file.
- The algorithm follows the following steps: it initializes the system with the charges specified in the builder and then at each step it computes the field $E(x_i, y_i, z_i)$ at the coordinates of every lumped charge due to all the other charges and updates the charges according to their field response, until the system converges self-consistently to a certain tolerance (no charge update is greater than a certain tolerance between successive iterations).
- The viewer displays plots of the charge values for each molecule at each iteration, both as 2D molecule vs charge plots and as a 3D spatial representation, allowing to conveniently visualize the simulation results.

Experiments

The force field parameters for the bisferrocene have been taken from the thesis work of Daniele Volpintesta [55]. A united atom model with implicit hydrogens has been used. This means that the hydrogen is lumped with the carbon in point particles representing CH₃, CH₂ or CH₁ blocks. The simplified model is shown in figure 60. Moreover, the ferrocene and carbazole groups have very high elastic constant to restrain the movement near to the equilibrium position. For the gold to surface anchoring, this model uses the same Morse potential model as the one adopted previously in this dissertation.

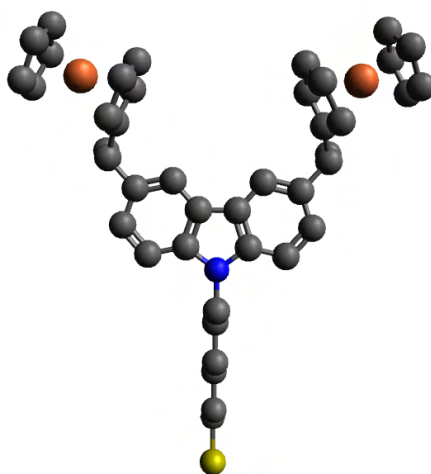


Figure 60: a bisferrocene molecule with implicit hydrogens. Carbons are in grey, sulfur in yellow, nitrogen in blue and iron in orange

For the simulation parameters, in every case, a velocity-Verlet algorithm with Nosé-Hoover thermostat at 300K has been used except in the initial temperature ramp-up phase of droplet equilibration.

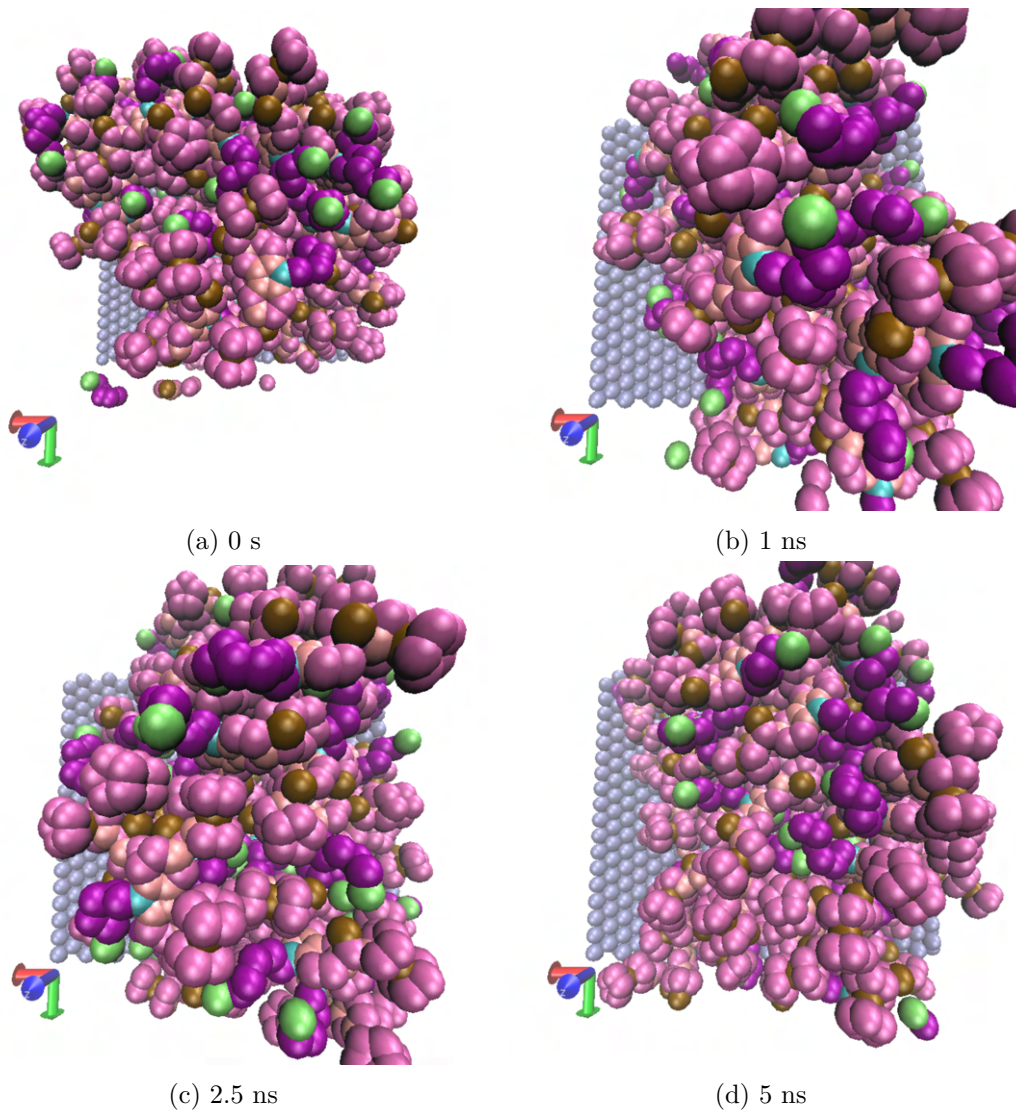
Deposition o bisferrocene molecules on a surface

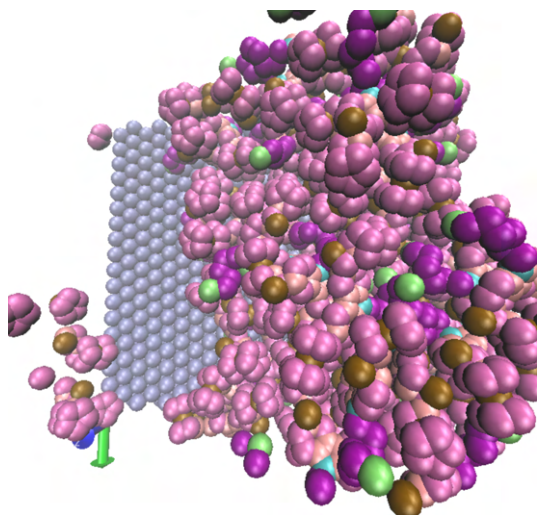
A droplet of 50 bisferrocene molecules has been heated up to 300K for 0.2ns before equilibrating a 300K at 0.8ns. This droplet has then been positioned at around 0.5 ns above 3 layers of an FCC gold lattice with constant 4.08 Å and the [1 1 1] surface pointing upwards. This has been followed by 20 ns of dynamics with the gold surface kept frozen and the same potential model as the previous depositions for the gold-gold van der Waals parameters and gold-sulfur interaction.

As the bisferrocene-bisferrocene forces are high, the molecules take a long time to spread over the substrate. By 10ns there still is a strip of the substrate that hasn't been covered and by the end, at 20 ns, an uncovered spot still remains.

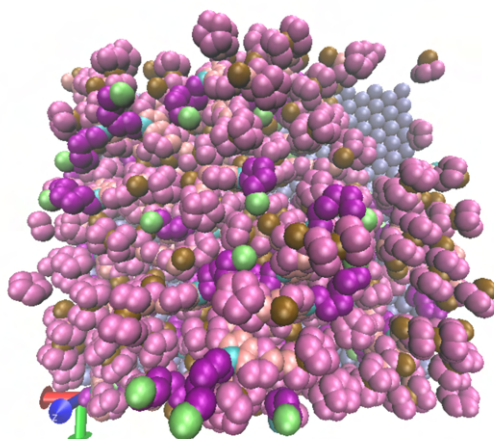
The following step is removing the unattached molecules and reequilibrating. During the

equilibration the remaining gap partially closes but there remains still a small uncovered area. Moreover, the deposition shows a strong degree of disorder, with molecules having all degrees of orientations and tilt angles and failing to achieve an ordered assembly.

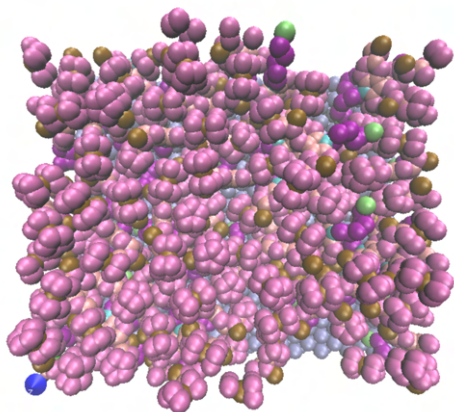




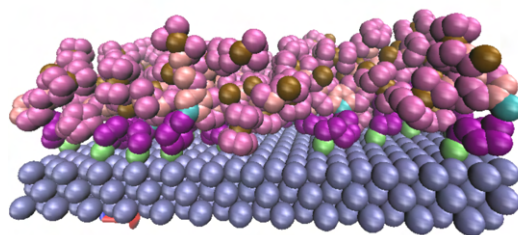
(e) 10 ns



(f) 20 ns



(g) Final cleaned top



(h) final cleaned side

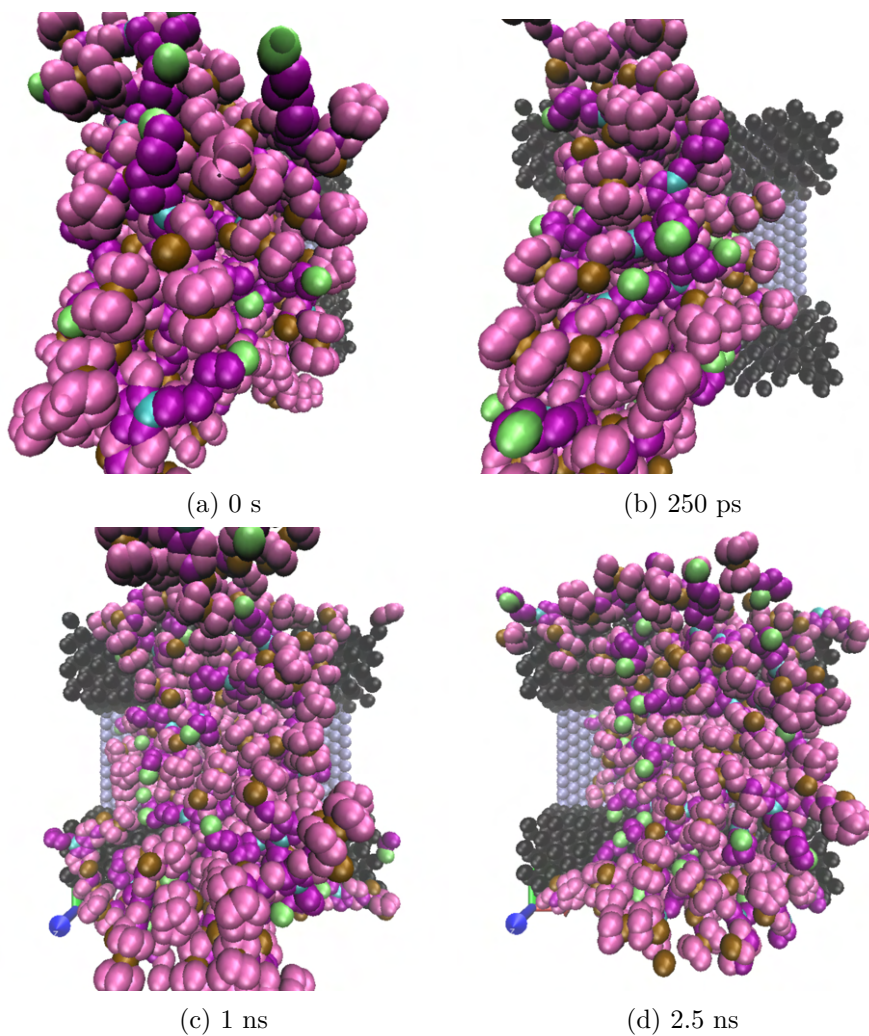
Figure 61: Deposition of 50 bisferrocene molecules on a gold substrate

Deposition of bisferrocene molecules in a trench

The same process as before is repeated but this time two 1.4 nm tall shoulders are placed 3.5 nm apart, with the same lattice constant and as and a purely repulsive Van der Waals term to simulate the lack of attractive interactions with the molecules.

Due to the strong bisfe-bisfe interaction, by 10 ns the trench is not completely filled and it takes the full 20 ns of simulation time for a sufficient number of thiols to fill it.

After removing the unattached molecules and reequilibrating, 14 bisferrocene molecules remain inside. Their orientation shows a high degree of disorder, highlighting the inability to properly orient the bisferrocenes through confinement.



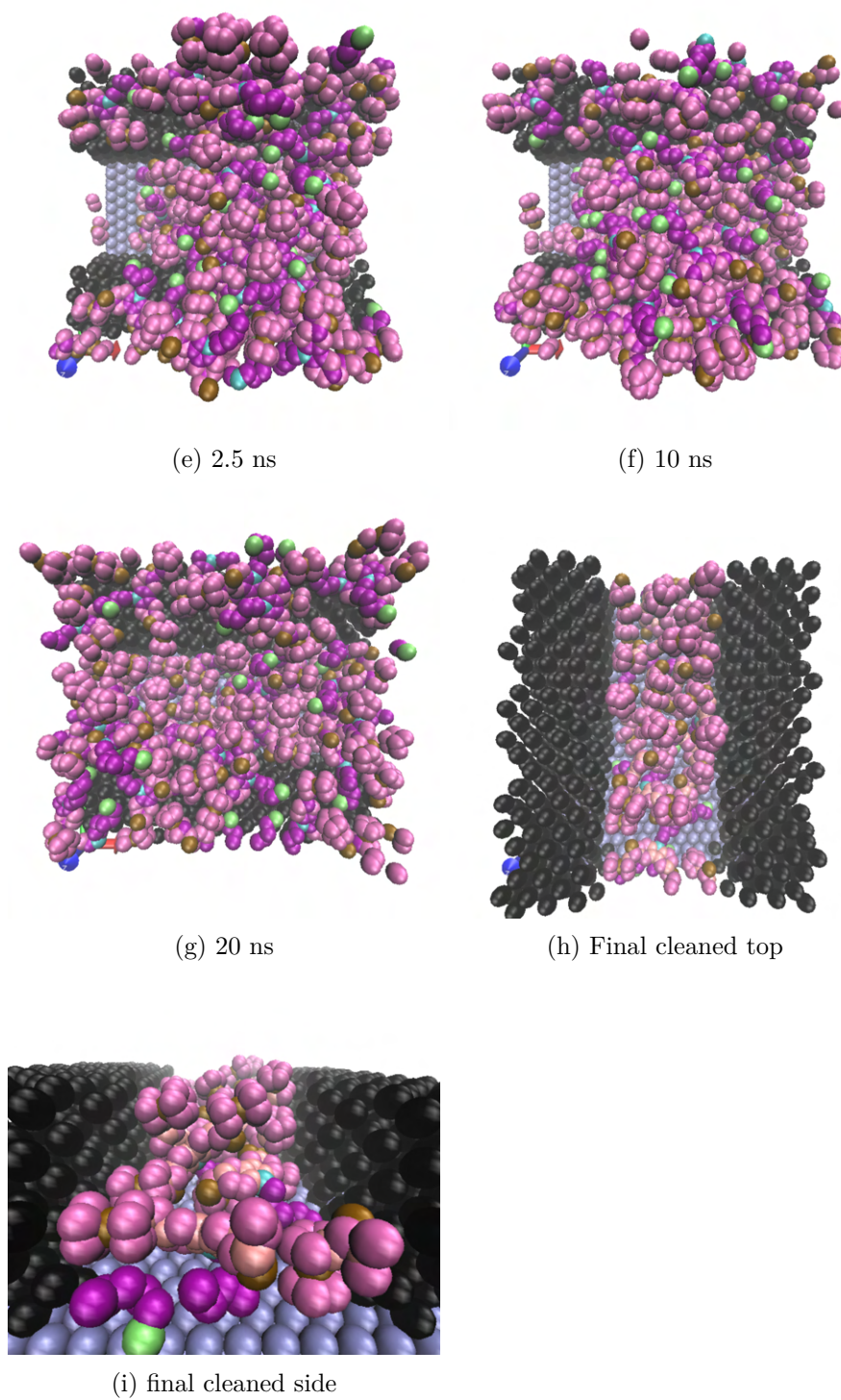


Figure 62: Deposition of 40 bisferrocene molecules on a gold substrate and confined in a 3.5 nm wide trench

Etching of a line of bisferrocene molecules

One last attempt to achieve an ordered distribution of bisferrocene molecules to analyze in SCERPA is cutting out a wire from the previous gold deposition. A wire obtained this way will have variability in the spacing of positioning of the anchoring points, helping to simulate process variability.

Initially, after cutting out the wire, eight molecules remain and they are equilibrated by running dynamics for 1ns. One can however observe how the molecules tend to aggregate leaving a gap in the middle that hinders signal transmission. An attempted solution is to oxidize the ferrocene molecules and add a Coulomb interaction term to the force field. To do that, however, one more misaligned atom is removed leaving seven atoms in the system.

After calculating the oxidized charges with ORCA, using the same methods as for the FcC6S molecules, these charges are applied to the atoms of the system before running 1 ns of dynamics. At this point the Coulomb repulsion is not sufficient to obtain an ordering of the molecules, thus a 2 V/nm electric field is applied. This pulls the molecules upright and some degree of ordering is obtained.

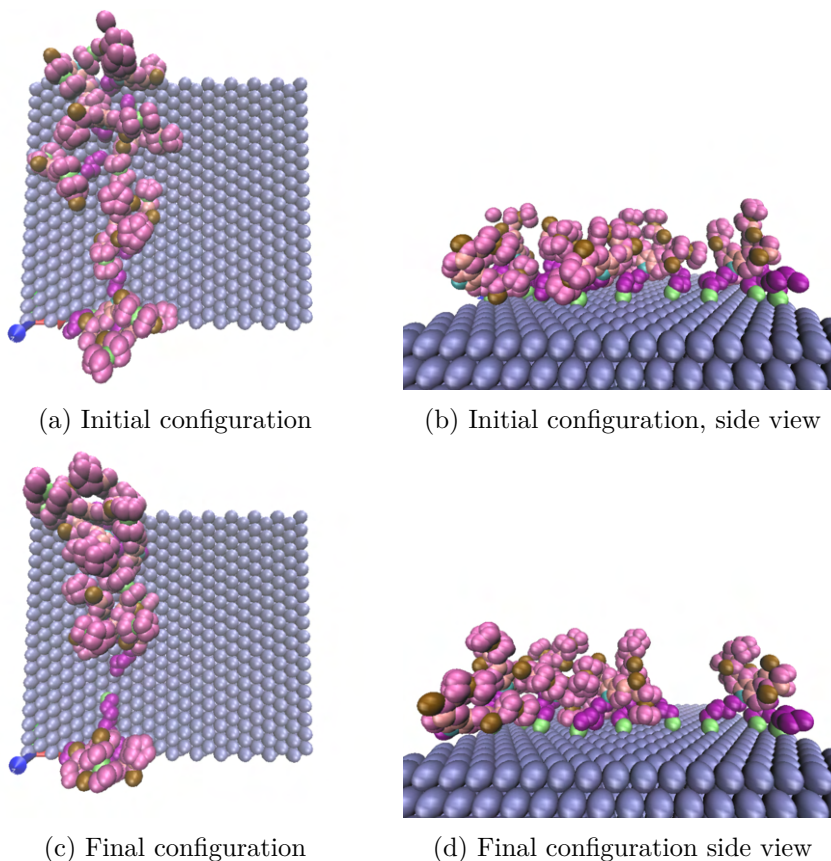
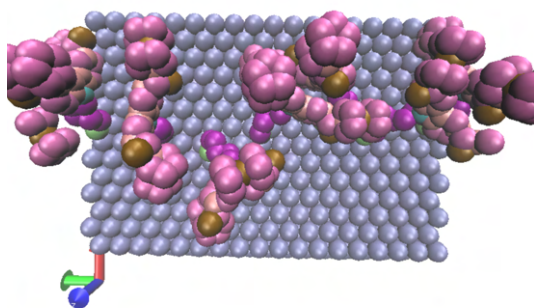
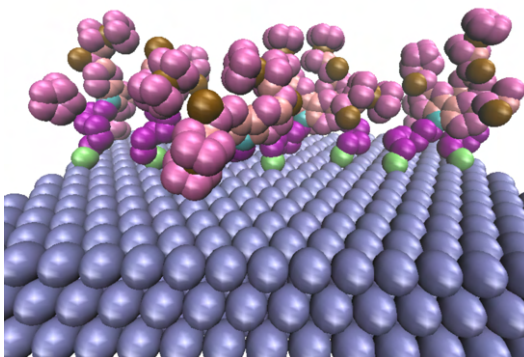


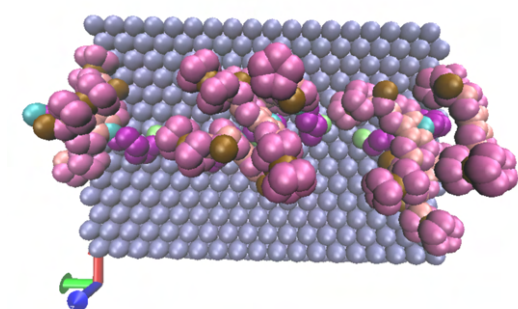
Figure 63: Equilibration of a bisferrocene wire



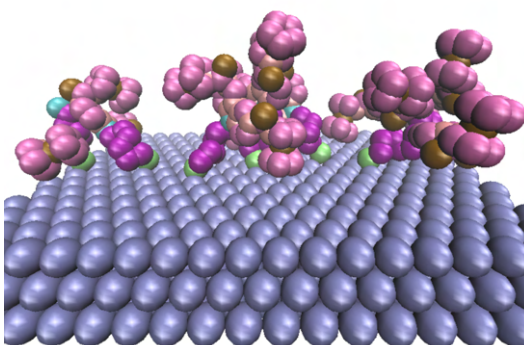
(a) Initial configuration



(b) Initial configuration, side view

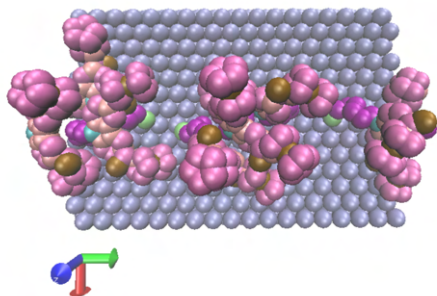


(c) Final configuration

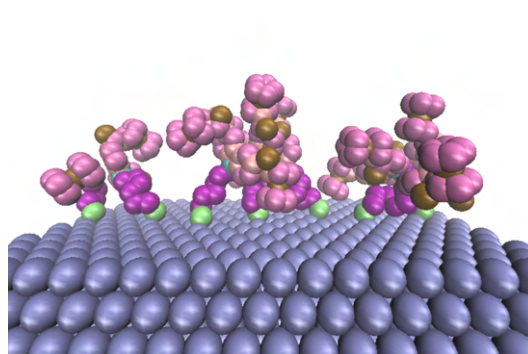


(d) Final configuration side view

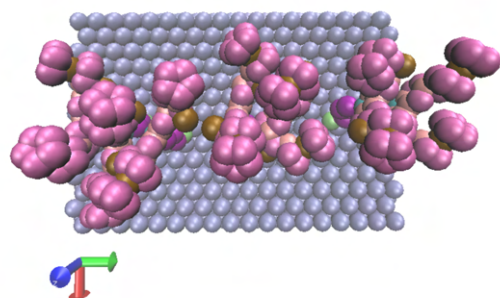
Figure 64: Equilibration of a bisferrocene wire after oxidation



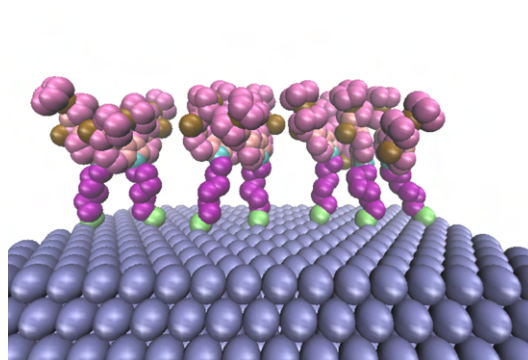
(a) Initial configuration



(b) Initial configuration, side view



(c) Final configuration



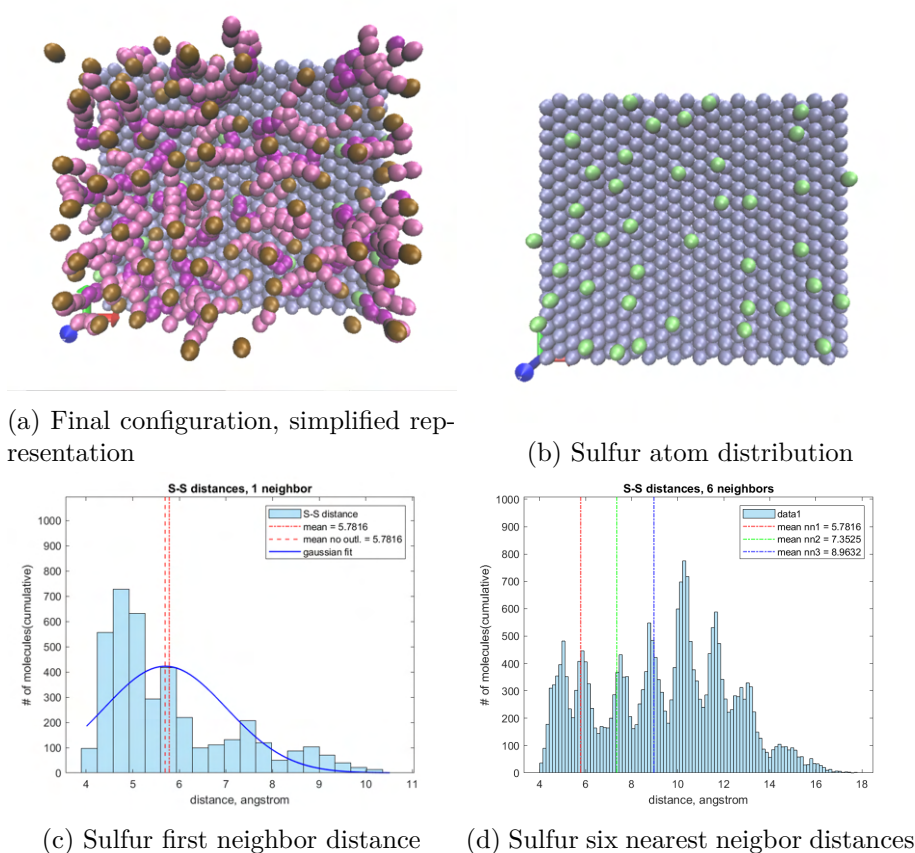
(d) Final configuration side view

Figure 65: Equilibration of a bisferrocene wire after oxidation and application of a 2 V/nm electric field

Matlab postprocessing

Postprocessing on Matlab is performed to show some information about the geometry of the system. In particular the tilt of both the whole molecule and just of the thiol chain, The tilt of the axis linking the two iron atoms with respect to the horizontal and the rotation of the molecule with respect to the x axis. Finally the sulfur to sulfur distances are also shown. A schematic representation showing just the thiol chains and the atoms linking them to the iron allows to observe the positioning of the molecules in the system. Another image shows the distribution of the sulfur atoms.

Bisferrocene on a gold surface



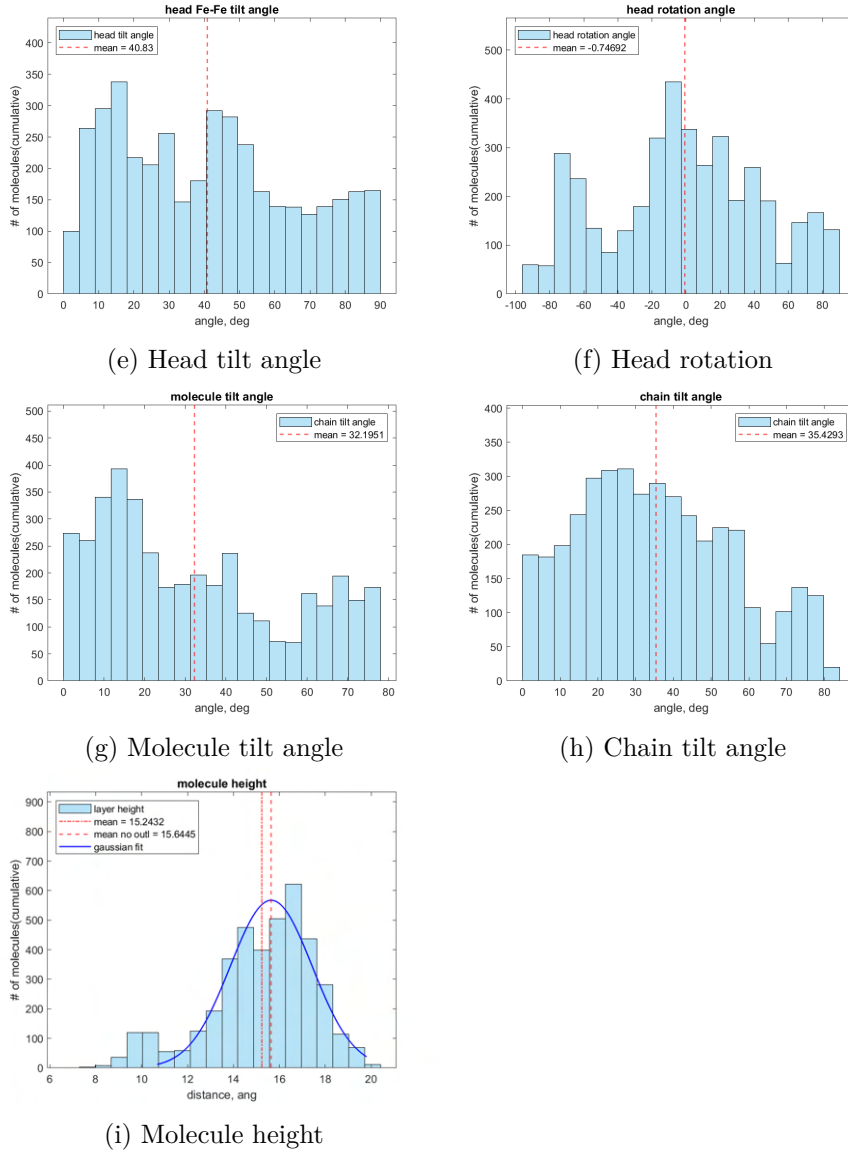
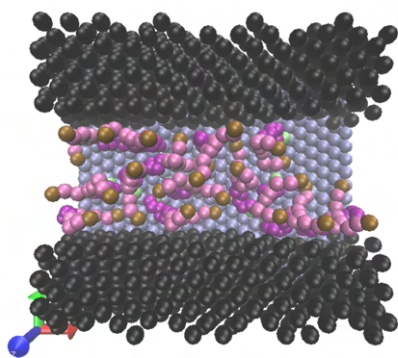


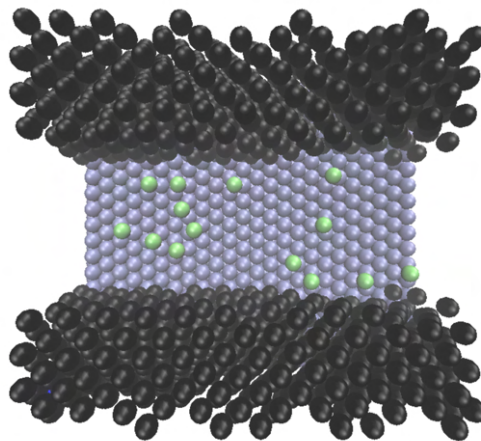
Figure 66: Bisferrocene on gold, final configuration

The plots highlight a strong degree of disorder in the molecule distribution. First, the spread in the nearest neighbour distance shows that, while there is a preferred distance around 6Å, there are areas of lower density. The hindrance of the head means that every sulfur atom has multiple positions it can bind and is not constrained to a regular pattern. The tilt angles occupy the whole range from a fully horizontal molecule to a fully vertical one. The former are responsible for the tail in the molecule height distribution, while the latter represent the majority of the atoms.

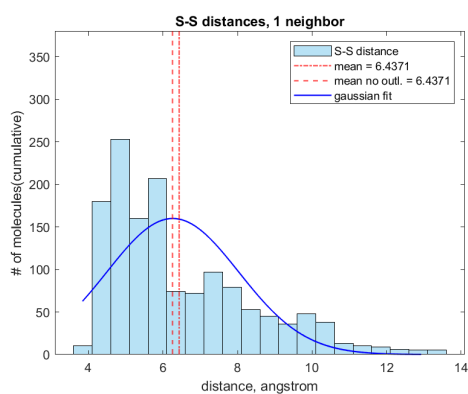
Bisferrocene in a trench



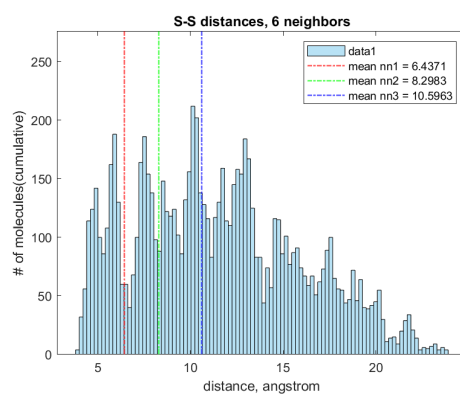
(a) Final configuration, simplified representation



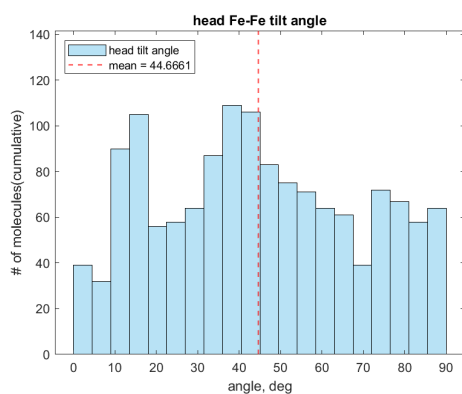
(b) Sulfur atom distribution



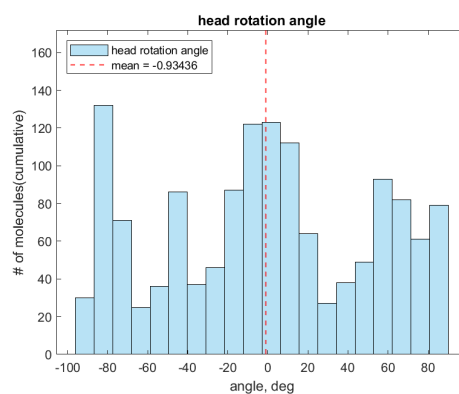
(c) Sulfur first neighbour distance



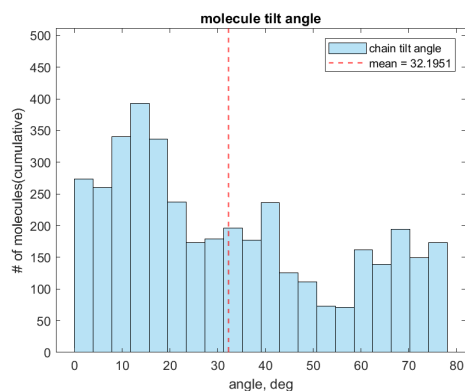
(d) Sulfur six nearest neighbour distances



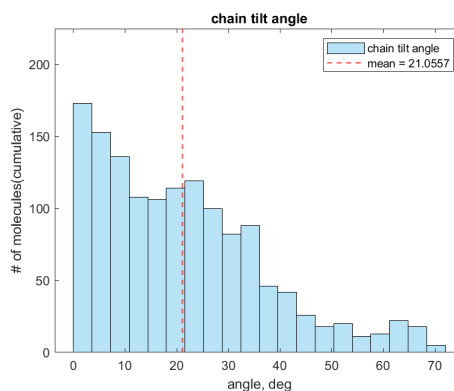
(e) Head tilt angle



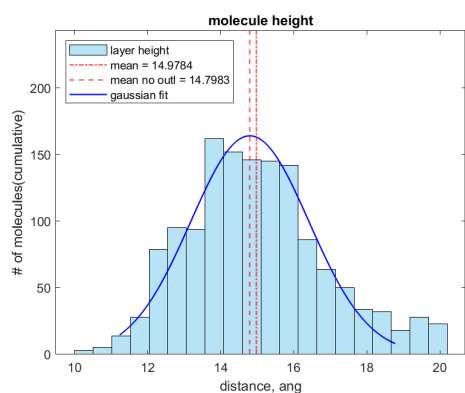
(f) Head rotation



(g) Molecule tilt angle



(h) Thiol chain tilt angle



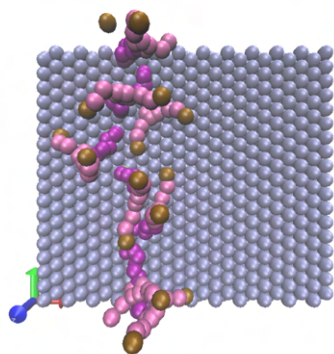
(i) Molecule height

Figure 67: Bisferrocene in a trench, final configuration

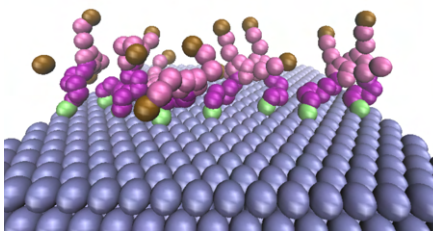
The plots show how using a trench to guide the deposition doesn't facilitate obtaining an ordered distribution of the ferrocene molecules: molecules have very different orientations and some are lying horizontally on the substrate.

Bisferrocene wire

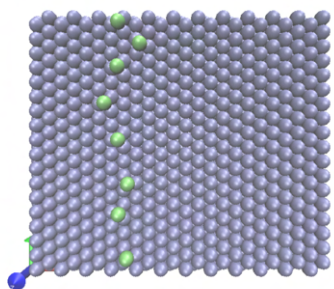
Equilibration



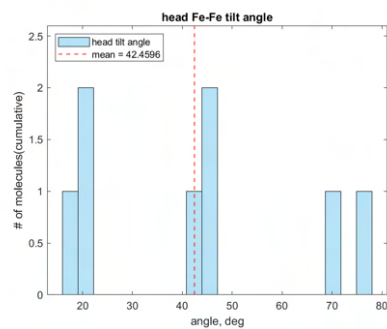
(a) Final configuration, simplified representation



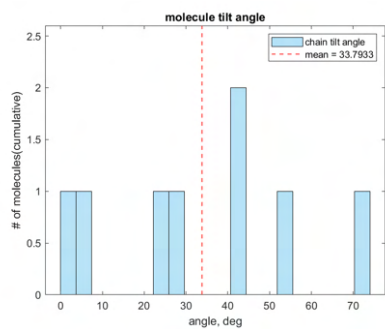
(b) Final configuration, simplified representation, side view



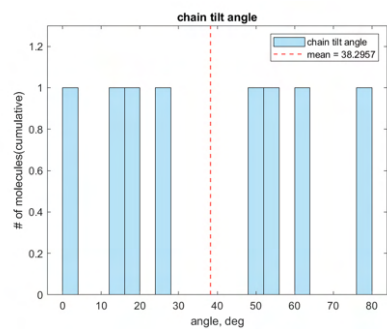
(c) Sulfur atom distributionI



(d) Head tilt angle



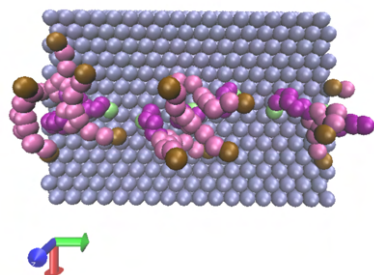
(e) Molecule tilt angle



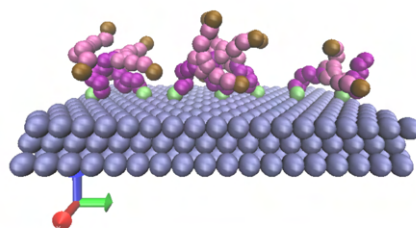
(f) thiol chain tilt angle

Figure 68: Wire equilibration, final configuration

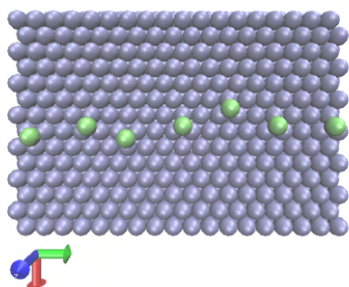
Oxidation



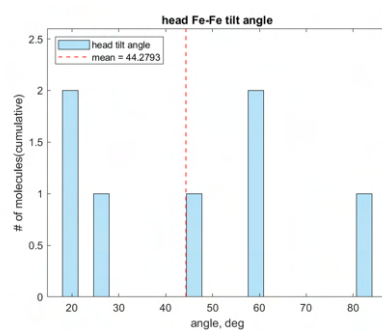
(a) Final configuration, simplified representation



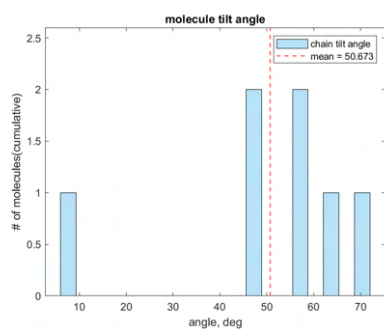
(b) Final configuration, simplified representation, side view



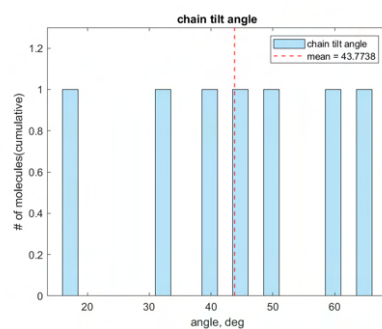
(c) Sulfur atom distribution



(d) Head tilt angle



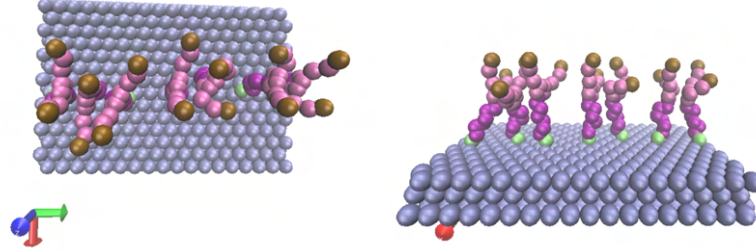
(e) Molecule tilt angle



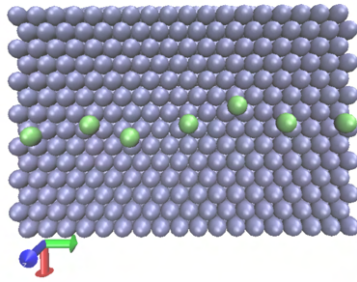
(f) Thiol chain tilt angle

Figure 69: Oxidized wire, final configuration

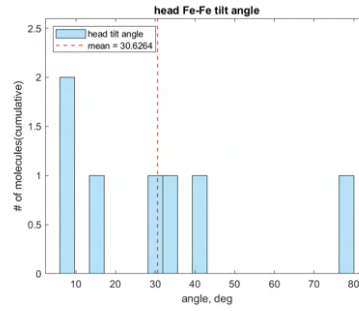
Electric field application



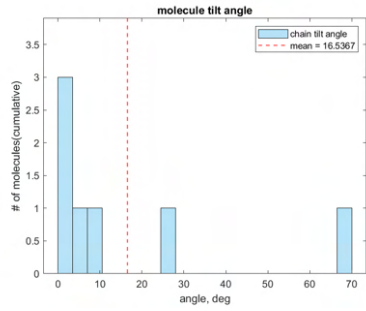
(a) Final configuration, simplified representation
(b) Final configuration, simplified representation, side view



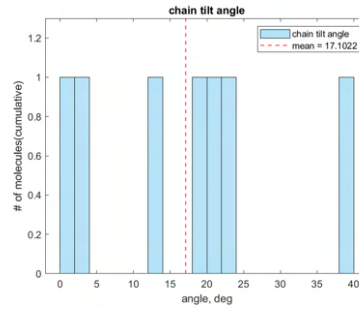
(c) Sulfur atom distribution



(d) Head tilt angle



(e) Molecule tilt angle



(f) Thiol chain tilt angle

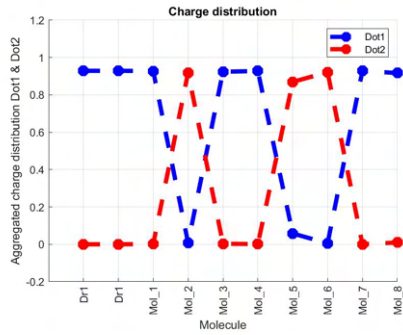
Figure 70: Final configuration of an oxidized bisferrocene wire after the application of a 2 V/nm electric field

In the wire the molecules are still highly disordered after substituting the oxidized charge. Applying an electric field restores a more ordered state. In particular, the molecule tilt angles are very low, with a single outlier in the second leftmost molecule.

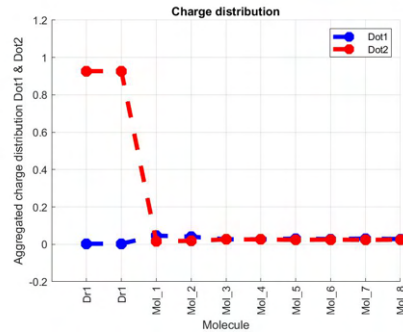
Information transmission on Scerpa

The three stages of the wire configuration are studied on Scerpa after extrapolating a reasonable approximation of the molecule orientations in space. The sequence of input is two timesteps with a positive driver and two with a negative driver, each with a negative clock (reset) in the first half and a positive clock in the second.

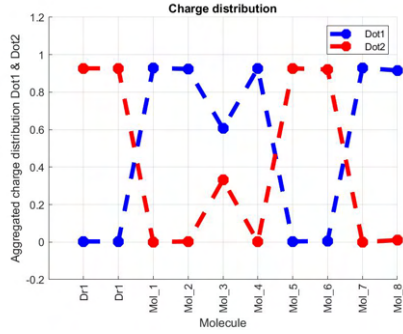
Simple wire



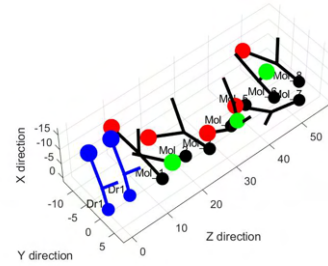
(a) Charge graph, positive driver



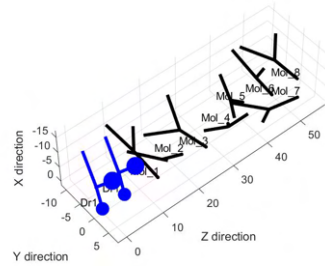
(b) Charge graph, reset



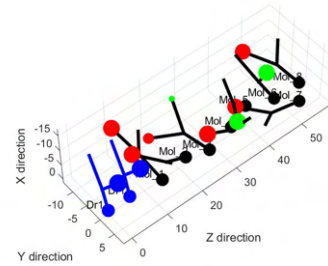
(c) Charge graph, negative driver



(d) 3D graph, positive driver



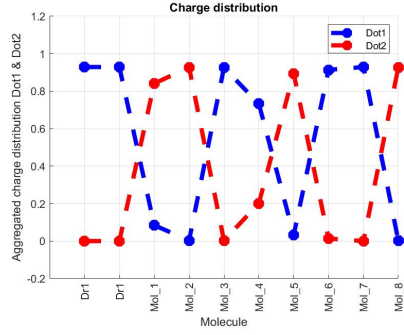
(e) 3D graph, reset



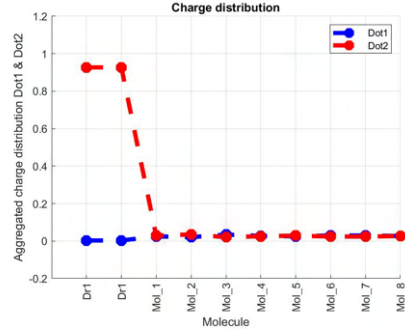
(f) 3D graph, negative driver

Figure 71: Simulation on Scerpa of the charge distribution on a bisferrocene wire under different driver conditions

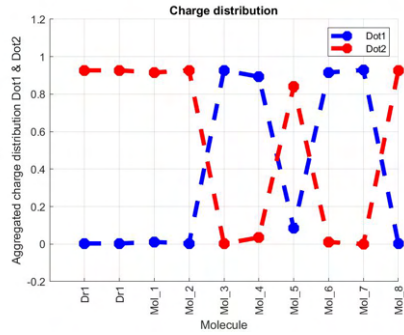
Oxidized wire



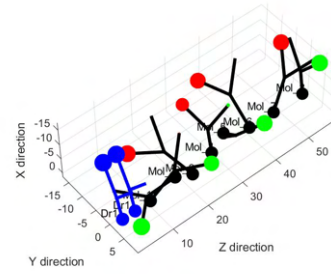
(a) Charge graph, positive driver



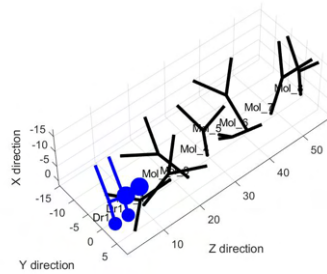
(b) Charge graph, reset



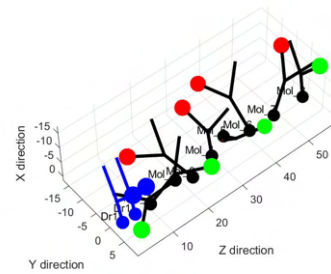
(c) Charge graph, negative driver



(d) 3D graph, positive driver



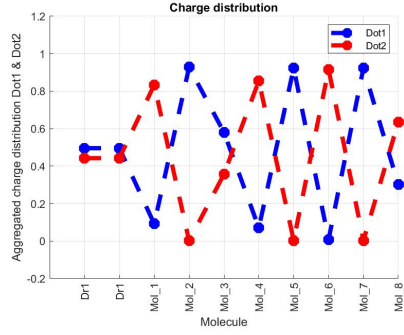
(e) 3D graph, reset



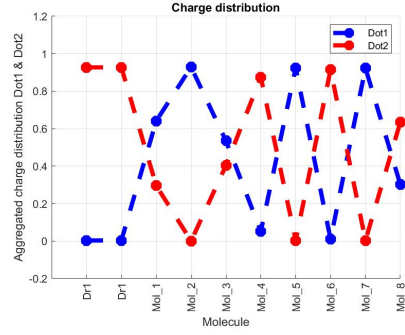
(f) 3D graph, negative driver

Figure 72: Simulation on Scerpa of the charge distribution on an oxidized bisferrocene wire under different driver conditions

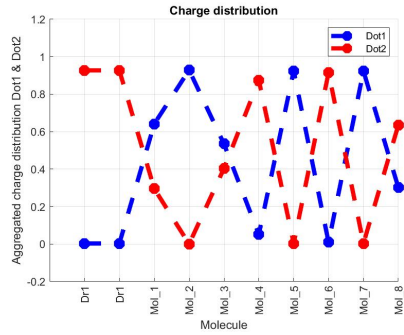
Oxidized wire with field



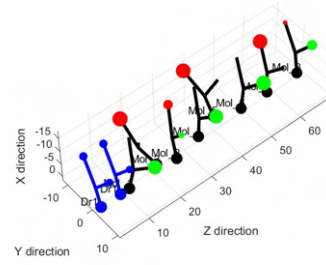
(a) Charge graph, positive driver



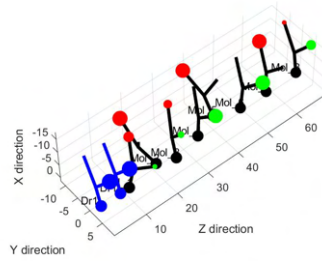
(b) Charge graph, reset



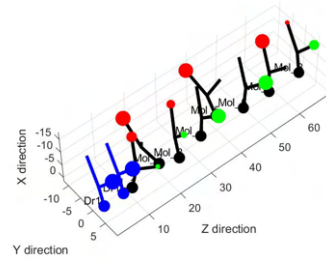
(c) Charge graph, negative driver



(d) 3D graph, positive driver



(e) 3D graph, reset



(f) 3D graph, negative driver

Figure 73: Simulation on Scerpa of the charge distribution on a bisferrocene wire, oxidized and subjected to a vertical electric field, under different driver conditions

Scerpa results and conclusions

The Scerpa simulations highlight the difficulty in obtaining a molecular QCA wire without controlling accurately the position of each molecule. Even in the most ordered case, the one where a field is applied to orient the molecules, one misalignment is enough to stop the information transmission.

Thus molecular QCA applications might have tolerances that are too small compared to the ability to control systems in MD simulations. Further work is necessary to ensure a functional workflow between MD and Scerpa simulations.

Conclusion and future perspectives

Molecular dynamics is a viable approach to provide insight into systems that are difficult to study with traditional experimental techniques. However, it cannot substitute the latter due to the inherent limits of simulations, where a tradeoff between accuracy and simulation time occurs and obtaining a treatable model while still capturing the main characteristics of phenomena is a complex task. This is particularly true when organic molecules are involved, which is the case in this work. In this case, multiple working hypotheses need to be made in order to have a treatable model for molecular dynamics and experimental validation is needed to test those hypotheses. In the cases where experimental data was available, albeit scarce, a degree of correspondence was obtained. However, there's still room for refinement of the model. In other cases, like the FcC6S wire and the bisferrocene molecules, the lack of experimental data on the geometry means a validation of the results is not yet possible.

In order to improve the simulations, experimental research must be paired with computational chemistry techniques so that the simulation models can be improved and in exchange provide more clarity in interpreting the experimental data.

What warrants more exploration and is particularly important in molecular electronics is the interaction between charge and geometry. In particular, the MD approach allows to compute the probable equilibrium geometry knowing the charge and the electric field. Tools like SCERPA allow to compute the charges and electric field knowing the geometry. Future research might investigate how these two approaches can be combined in order to reach a better understanding of electrically active molecular systems.

Appendix A: Lammmps implementation

Part 1: FcC6S code

Droplet equilibration

This code is a sample of the template used for droplet equilibration. Full documentation for the commands can be found on the LAMMPS website at <https://www.lammps.org/>. <...> indicates that the instruction has been split into multiple lines for space constraints but should be treated as a single line.

```
1)    units real
```

```

2)    dimension 3
3)    processors * * *
4)    boundary p p p
5)    atom_style full

```

This code sets some parameters for the simulation. The first line sets the physical units to "real". It's a convention where in particular the following units are set:

- [mass] = g/mol
- [distance] = Å
- [temperature] = K
- [time] = fs
- [charge] = multiple of electron charge (1.0 is a proton charge)
- [electric field] = V/Å

Line 2 specifies the number of dimensions for the simulation domain

Line 3 specifies the number of grid points for the partition of the simulation domain. Asterisk in any dimension means LAMMPS chooses the number of partitions per direction to minimize the surface to volume ratio of the subdomain.

The "boundary" command at line 4 specifies the boundary conditions for the three spatial dimensions. "p" stands for periodic boundary, "f" for fixed, aperiodic boundary. The atom_style command determines the quantities stored for each atom and is needed so that LAMMPS can preallocate memory space. LAMMPS is optimized for speed and thus doesn't allocate space on the fly, thus an assignment for a field with no allocated memory will return an error. The keyword "full" allocates space to store all the information needed to build molecules plus charge information.

```

6)    lattice fcc 4.08 orient x 2 -1 -1 orient y 0 1 -1 orient z 1 1 1
7)    region box block -15 15 -15 15 -15 20
8)    create_box 8 box bond/types 6 extra/bond/per/atom 4 angle/types 7 <...>
      extra/angle/per/atom 5 dihedral/types 7 extra/dihedral/per/atom 10 <...>
      extra/special/per/atom 26
9)    lattice custom 11 a1 0 1 0 a2 1 0 0 a3 0 0 1.2 basis 0 0 0
10)   region deprec block -1.5 1 -2 2 -1 2

```

At line 6 the "lattice" command creates a grid of lattice points used by other commands and sets the default distances for certain commands to the lattice periods in the cardinal directions. It specifies an FCC lattice with lattice constant 4.08 Å (a gold lattice) and lattice vectors oriented in the [2 -1 -1] [0 1 -1] and [1 1 1] directions

The "region" command identifies a portion of space. the first keyword, "box" is the region ID. "block" indicates that the region is a parallelepiped, followed by the minimum and maximum coordinates in the three cardinal directions in lattice units.

The "create_box" command, followed by an integer and the region ID creates a simulation

box spanning the specified region and with as many atom types as that integer value. Its following keyword-argument pairs specify the total different bond angles and dihedrals in the system and the maximum number of each force field element preallocated for each atom. It's necessary that the number of different bonds be exact and that the per-atom values be smaller or equal than the maximum or an error will be thrown.

The lattice command at line 9 changes the lattice into a custom one with lattice constant 11 and lattice vector components specified after a1, a2 and a3. "basis" followed by a triplet of numbers between 0 and 1 adds a basis site at the fractional lattice coordinates identified by the triplet.

The region command at line 10 creates a region where the molecules to deposit will be generated.

```
10)    mass 1 14.027 # CH2
11)    mass 2 1.08  #H
12)    mass 3 32.06 # S
13)    mass 4 12.0107 #C
14)    mass 5 0.1
15)    mass 6 55.845 # FE
16)    mass 7 196.967 # Au
17)    mass 8 28.0855 # Si
```

Lines 10 to 17 assign a mass value (in g/mol) to each atom type (1 to 8). # starts a comment

```
18)    pair_style hybrid lj/cut/coul/cut 10 morse 10
19)    pair_coeff 1 1 lj/cut/coul/cut 0.0935 4.411
...
44)    pair_coeff 3 7 morse 8.763 1.47 2.65
...
64)    pair_coeff 7 8 lj/cut/coul/cut 0.152932338 3.6065
```

Lines 18 to 64 specify the type and strength of the nonbonded interactions among the atom types. "pair_style hybrid" specifies that there is more than one type of nonbonded interaction and the following keywords declare that the nonbonded interaction types are Lennard-Jones plus Coulomb and Morse. The coulomb part of the potential uses the charges specified in the molecule file to compute coulomb interactions, treating the atoms as point like charges, and the strength of the Lennard-Jones and Morse interactions is set in the following lines. The "pair_coeff" command is followed by a pair of atom types, the type of interactions and the specific parameters for that interaction. An interaction must be specified for each pair of atom types.

```
65)    bond_style harmonic
66)    bond_coeff 1 1000 1.440 #C_C ring from paper
...
71)    bond_coeff 6 205 1.649
72)    angle_style cosine/squared
```

```

73)    angle_coeff 1 124 114
      ...
79)    angle_coeff 7 140 126
80)    dihedral_style hybrid opls table spline 1082
81)    dihedral_style hybrid nharmonic opls quadratic fourier table spline 1082
82)    dihedral_coeff 1 nharmonic 4 1.57 -4.05 0.86 6.48 #paul yoon
83)    dihedral_coeff 2 table look.txt DIH_TABLE
      ...
87)    dihedral_coeff 6 fourier 1 0.36 5 180 #1-5 dihedral
88)    dihedral_coeff 7 opls 0 10.79 0 0 #from linear metallocene paper

```

Here all bonding force field terms are assigned a value. The number of entries for bonds, angles and dihedrals must be equal to the preallocated amount in the "create_box" command. Lines 65, 72 and 80 specify the type of models used respectively for bonds angles and dihedral. It can be a single type of model or multiple. In the latter case the required keyword is "hybrid" followed by a list of the models and the necessary associated parameters.

Once the model is specified, all the coefficients are set with the commands "bond_coeff", "angle_coeff" and "dihedral_coeff", followed by the number associated with that set of coefficients, the model in case the hybrid keyword is used, and all the parameters required by that specific model. Further instructions need to be provided separately to list which pair, triplet or quadruplet of atoms are subject to these interactions.

```

89)    molecule thiol rigid_fc
90)    create_atoms 0 region depreg mol thiol 1234 subset 50 123

```

The instruction "molecule" loads the molecule file "rigid_fc" and assigns it to the variable "thiol". The "create_atoms" instruction creates atoms in the lattice sites contained in the above-mentioned region "depreg". The "mol" argument specifies that at each site the molecule "thiol" is created, with the random seed 1234 to determine the orientation. The "subset" keyword specifies that only a subset of 50 sites is occupied, with random seed 123 to choose them.

```

91)    group Thiol type 3
92)    group Thiol type 1
93)    group Fc type 4
94)    group Fc type 5
95)    group Fc type 2
96)    group nr type 1
97)    group nr type 3
98)    group nr type 6

```

This set of instructions assign atoms to different groups by type, to be referenced by following instructions. Group "Thiols" refers to the tail atoms. Group "Fc" to the ferrocene unit, which will be set as rigid later, and group "nr" to parts of the molecule that are left non-rigid.

```

99)    comm_modify cutoff 12
100)   timestep 1
101)   thermo_style custom step temp pe etotal press vol
102)   thermo 10
103)   fix rec all recenter INIT INIT INIT
104)   fix rig Fc rigid molecule

```

The "comm_modify" command at line 99 sets the global cutoff for force field terms to 12 Å. The "timestep" command sets the timestep to 1 fs, the "thermo_style" command with the "custom" keyword prints to the log a custom set of output value every x timestep where x is the argument of the "thermo" command at line 102. The printed values are the timestep number, the temperature, the potential energy, the total energy, the pressure and the volume.

The following part contains "fix" instructions, instructions that specify a constraint on a group of atoms or on group "all", containing all atoms on the system. The formatting is "fix" followed by the fix ID, the group ID, the type of fix and its argument.

Fix "recenter" resets the coordinates of the center of mass to the xyz values mentioned. "INIT" refers to the initial values of the center of mass coordinate. Fix "rigid" treats all atoms in the specified groups as a rigid body. The "molecule" keyword indicates that atoms belonging to different molecules form separate rigid bodies.

```

105)   dump 1 all custom 1000 equil_droplet.lammpstrj <...>
        id type x y z vx vy vz
106)   dump_modify      1 sort id

```

The "dump" keyword specifies the format of an output file containing per-atom values. The formatting can be predetermined or customized with the "custom" keyword. The latter is followed by the interval, in terms of number of timesteps, at which data is acquired, the output file name "equil_droplet.lammpstrj" and the values to dump. In this case, the format is that of a "trajectory file", containing the information to visualize the trajectory on third-part software like VMD. It includes the atom ID number followed by atom type, x y z coordinates and velocity components. The "dump_modify 1 sort id" instruction modifies "dump 1" to guarantee that the involved atoms are sorted in numerical order based on their ID.

```

107) minimize 1e-4 1e-6 1000 10000
108) fix 4 nr nvt temp 1 300 10
109) run 200000
110) unfix 4
111) fix 4 nr nvt temp 300 300 10
112) run 1000000
113) write_restart restart.equil_droplet

```

This section contain the commands that set up the MD loop. The instruction "minimize" line 107 launches energy minimization with the default lammps algorithm, gradient descent. The following values are the energy and force variations between iteration used as stopping criterion, the maximum number of iterations and the maximum number of

force/energy evaluations.

Line 108 adds a Nosé-Hoover thermostat to constrain the temperature of the nonrigid part of the system, assigned to group "nr". "fix rigid" handles the integration of the rigid portion of the system. The "temp" argument specifies the temperature and is followed by the initial and final temperature and the period of temperature fluctuations in number of timesteps.

The "run" command at line 109 executes the MD loop for 200.000 timesteps under the specified constraints. The "unfix" command removes the fix with ID 4 at line 108, which is reassigned at line 111 before running time integration for 1.000.000 more steps. The last command, "write_restart", creates a file with all the necessary information to continue the simulation in another LAMMPS instance.

The molecule file

A molecule template is included as an external file to be used with the "create_atoms" command. This template is described extensively at <https://docs.lammps.org/molecule.html>.

The file has a header and a body. The first line of the header is skipped and reserved for comments. The header contains a list of settings where values like the number of bonds, angles, dihedrals and atoms are specified. The following lines are formatted this way: a header describing the type of information provided, followed by an empty line which is ignored and a series of lines listing the atom numbers and the corresponding values:

```
1)    ###this line is skipped###
2)    29 atoms
3)    39 bonds
4)    34 angles
5)    37 dihedrals
6)
7)    Coords
8)
9)    1 2.219810339 -1.429069019 -3.487094387
...
37)   29 1.52367516825960 -0.95656490944412 -1.60298401094047
38)
39)   Types
40)
41)   1 4
...
69)   29 6
70)
71)   Bonds
72)
73)   1 1 1 2
...
```



```

111) 39 6 27 29
112)
113) Angles
114)
115) 1 1 26 25 24
...
147) 34 7 20 4 3
148)
149) Dihedrals
150)
151) 1 1 26 25 24 23
...
187) 37 7 5 4 1 20
188)
189) Charges
190)
191) -0.253819
...
219) 29 0.043574

```

The atom IDs go from 1 to the maximum number of atoms. The "Coords" section lists the atom coordinates associated with each ID; the "Types" section the atom types. The "Bonds" section the types of bonds associated with each bond ID followed by the pairs of atom IDs involved. The "Angles" section the type of angles associated with each angle ID, followed by the IDs of the triplet of atoms involved. Same for the dihedral sections, with a quadruplet of atoms involved.

From all the information in the molecule file, together with the list of atom types and bonds, angles and dihedral types on the main files, Lammmps can fully reconstruct the model of the molecule.

Depositing the molecules

Another simulation is run to deposit the molecules on a gold substrate, with some similarities and some major differences with the previous input file.

```

1) units real
2) dimension 3
3) processors * * *
4) boundary p p p
5) atom_style full

```

This section is the same as the previous simulation.

```

6) read_restart restart.equil_droplet
7) fix rig Fc rigid molecule

```

The instruction at line 6 read the system information stored as a restart file. This includes the atoms coordinates and velocities together with any information about the

masses, bonding structure, group belonging. Not included is information about the force field models and all the "fix", "compute" and "thermo" instructions.

The "rigid" fix has been immediately reassigned since it needs to precede any box-changing instruction.

```
8)  lattice fcc 4.08 orient x 2 -1 -1 orient y 0 1 -1 orient z 1 1 1
9)  change_box all z final -9 15 x final -3 3 y final -3 3 boundary p p f
10) region    gold block INF INF INF INF -3 -2
11) create_atoms 7 region gold
```

The "lattice" command defines an FCC lattice with the spacing of gold and oriented so that the [1 1 1] surface is pointing upwards; the "change_box" command alters the size of the simulation box and the boundary conditions so that the latter are aperiodic in the z direction. The "region" command defines a region for the creation of the gold substrate. The region has been carefully positioned in order to have less than 1nm from the thiol droplet.

The "create_atoms" instruction at line 11 creates atoms with type 7 (previously associated with the mass of gold) and at the lattice sites defined at line 8 (fcc gold lattice)

```
12) comm_style tiled
13) comm_modify cutoff 12
14) fix bal all balance 1000 1.05 rcb
```

The "comm_style" command defines the space partition between the processors as "tiled": instead of a regular grid of blocks the domain decomposition can be performed with generic tiles, parallelepipeds with varying sizes and shapes. The "balance" fix performs reassignment of atoms and changes in domain boundaries until the atoms are partitioned equally with a tolerance of 1.05 (a domain can have up to 5% more atoms than another domain).

```
15) pair_style hybrid lj/cut/coul/cut 10 morse 10
...
74) dihedral_coeff 7 oplis 0 10.79 0 0
```

The force field models associated with each bond, angle and dihedral type and the non-bonding interactions are defined as before.

```
75) group Gold type 7
76) velocity Gold zero linear
77) velocity Fc zero linear
78) velocity nr zero linear
79) fix frz Gold setforce 0 0 0
```

After assigning the gold atoms to a group ("Gold") the initial velocity of the atoms in the system is set to 0 and the gold atoms are frozen in position with the "setforce" fix, which overrides any previous instruction concerning the forces acting on gold atoms and sets them to the assigned value (0 0 0).

```

80)  timestep 1ns
81)  compute Tth nr temp
82)  fix 4 nr nvt temp 300 300 10
83)  dump 1 all custom 2000 thfc_gold1.lammpstrj id type x y z vx vy vz
84)  dump_modify 1 sort id
85)  thermo 100
86)  thermo_style custom step c_Tth pe etotal press vol
87)  thermo_modify lost warn
88)  minimize 1.0e-4 1.0e-6 100 1000
89)  run 5000000
90)  write_restart restart.thiolfc_drop_gold_1
...

```

This set of instructions is similar to the previous case and is used to run energy minimization followed by NVT dynamics. However, a fundamental difference is that the temperature computation is restricted to the FcC6S molecules. To that purpose, a "compute" instruction at line 80 is used to compute the temperature of the molecular subsystem. That's because the temperature is computed as the average kinetic energy and thus the frozen gold substrate must be excluded from the temperature computation. Otherwise, the molecular subsystem reaches a higher temperature than 300K to keep the average T constrained at 300K.

The "thermo_modify lost warn" instruction prevents the simulation from stopping if a particle leaves the simulation domain due to the aperiodicity in the z direction. Instead a warning is thrown.

Using the ReaxFF force field

In the case of a simple thiol, the deposition has been attempted using the reaxff force field. In this case a timestep of 0.1 fs or a little more can be chosen. No bonds, angles and dihedrals must be allocated with the "create_box" command and the force field must be set as:

```

pair_style reax/c control
pair_coeff * * ffield.reax.AuSCH C H S Au
fix 3 all qeq/reax 10 0.0 10.0 1.0e-6 reax/c

```

Where "reax/c" is a c++ implementation of the reaxFF force field, "control" is the control file described at https://www.smcm.iqfr.csic.es/docs/lammps/pair_reax_c.html. For simplicity "pair_style reax/c NULL" can be used in which case default values are adopted.

The qeq/reax fix performs a charge equilibration. The arguments are the number of timesteps between equilibrations, the minimum and maximum cutoff radius and the tolerance. The "reax/c" keyword specifies that the parameters used for charge equilibration are the ones associated with the reaxFF force field.

deposition molecule by molecule

for the deposition simulation, the "create_atoms" instruction must be replaced with a "deposit" fix.

```
fix dep new deposit 60 0 30000 321 region depreg mol thiol <...>
near 3 attempt 9999 vz -70e-5 -80e-5 vx -20e-5 20e-5 vy -20e-5 20e-5
```

The arguments are in order the number of deposited molecules, the offset to add to the atom types in the molecule file, the timestep interval between successive depositions, the random seed for the insertion position and orientation. The "region" keyword specifies the insertion region, the "mol" keyword the molecule to insert, the "near" keyword the minimum distance from existing atoms. If a molecule is inserted closer than that distance from existing atoms, the insertion fails. The "attempt" keyword specifies the number of times the system attempts to insert a molecule after a failure. If all attempts fail the molecule is no longer inserted. The last three keywords, "vx", "vy" and "vz" specify the minimum and maximum velocity in the x, y and z directions that are assigned to the newly created molecules.

Part 2: anchoring molecules to a substrate

Some crucial steps are required to anchor the molecules to a substrate in a regular pattern. First of all it's necessary to generate the molecules on superlattice sites.

```
region thiols block 0 5.2 0 7 -1.0 -0.85
lattice custom 2.885 a1 0.866205 2.5 0 a2 2.598076211 0.5 0 a3 0 0 0.25 basis 0.22 0.154
molecule thiol rigid_fc
create_atoms 0 region thiols mol thiol 1234 rotate 90 0 1 0
```

Some special care is needed, when creating the periodic box with the "create_box" command, to ensure that the periodicity of the box coincides with that of the superlattice. The superlattice is defined with the "lattice custom" command followed by the definition of the lattice vectors a1, a2 and a3 and the coordinates of the basis points. Then the molecule is imported and the "create_atoms" command generates the superlattice. Some fine-tuning is required to have the molecules oriented correctly and positioned correctly above the surface. a correction can be applied by measuring the distance between the actual position of the sulfur and its expected one and shifting the superlattice with the "displace_atoms" command after changing the lattice so that the lattice vectors are unitary and coinciding with the cartesian directions.

```
lattice custom 1 a1 1 0 0 a2 0 1 0 a3 0 0 1 basis 0 0 0
displace_atoms Fc move 8.540000 1.04999 -1.30000 units lattice
```

Finally, the most straightforward way to establish a fixed bond between the superlattice and the gold surface is generating a data file with the "write_data" command with the filename as first argument.

```
write_data data.config_2 nocoeff
```

The "nocoeff" keyword means that the force field coefficients are not written directly in the data file and need to be set before importing it in the following simulation with the "read_data" command. Before reading the data file, the necessary additional bonds and angles can be added manually to the bonds and angles section of the data file.

Part 3: FcC6S wire

wire

for the wire and trench part a triple region is used for the create-atoms command

```
change_box all z final -9 15 x final -6 6 y final -4.5 4.5 boundary p p f
region fill1 block INF INF -4.5 -1.7 -3.33 -2.33
region gold block INF INF -1.6 1.6 -3.33 -2.33
region fill2 block INF INF 1.7 4.5 -3.33 -2.33
create_atoms 7 region gold
create_atoms 8 region fill1
create_atoms 8 region fill2
```

, the regions are contiguous, for the trench simulation. The region "gold" is filled with gold atoms while the regions "fill1" and "fill2" with silicon atoms. Both the gold and the silicon must be added to a group and frozen before the simulation is run.

Oxidation and electric field application

After that the molecule is oxidized and the original charges are replaced with the charges for the oxidized thfc molecule. This is performed with the following set of instructions:

```
variable qat atomfile charge_atomfile_ox.txt
set atom * charge v_qat
```

The first creates an atomfile type variable. This variable reads a file formatted this way:

```
2888
1 -0.160556
2 -0.072563
3 -0.085419
4 -0.124474
...
2888 0
```

The first line contains the number of following lines that must be loaded into the atomfile variable. The following lines are pairs of atom ID and values. The set command with the "atom" keyword reads from the specified atomfile variable the values corresponding to the list of atom ID following the "atom" keyword (an asterisk means all the atoms referenced in the atomfile variable are considered). The "charge" keyword means those values are written into the atom charge.

conclusion

This is a set of methods used in this work for the simulation. To have a more complete idea of the functions of the different commands and for doubts about lammmps programming, free resources like the official documentation <https://docs.lammps.org/Manual.html> and the mailing list lammps-users@lists.sourceforge.net available through <https://lists.sourceforge.net> can help in setting up a simulation.

Bibliography

- [1] Bcc Research, accessed on 2021/07/22, <https://www.bccresearch.com/market-research/nanotechnology/nanobiotechnology-market.html>
- [2] European Commission website, accessed on 2021/07/22, https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/key-enabling-technologies_en
- [3] Sponsorship, acknowledgement of. "barriers to nanotechnology commercialization final report." (2007).
- [4] Synopsys, Inc., 2021, *quantumatk.com* accessed on 2021/07/22 https://docs.quantumatk.com/tutorials/md_basics/md_basics.html
- [5]
- [6] D. Frenkel and B. Smit. Understanding Molecular Simulation: From Algorithms to Applications, Computational Science Series Academic Press, San Diego, Second edition (2002)
- [7] Technische Universitat Munchen, accessed on 2021/07/ https://www5.in.tum.de/lehre/vorlesungen/sci_compII/ss13/uebungen/blatt9solution.pdf
- [8] Toxvaerd, Søren, Ole J. Heilmann, and Jeppe C. Dyre. "Energy conservation in molecular dynamics simulations of classical systems." The Journal of chemical physics 136.22 (2012): 224106.
- [9] Hünenberger, Philippe H. "Thermostat algorithms for molecular dynamics simulations." Advanced computer simulation (2005): 105-149.
- [10] Rühle, Victor. "Berendsen and nose-hoover thermostats." Am. J. Phys (2007).
- [11] Braun, Efrem, Seyed Mohamad Moosavi, and Berend Smit. "Anomalous effects of velocity rescaling algorithms: the flying ice cube effect revisited." Journal of chemical theory and computation 14.10 (2018): 5262-5272.
- [12] Finnerty, Justin "Molecular dynamics meets the physical world: Thermostats and barostats" German Research School for Simulation Sciences, accessed on 2021/07/24 http://www.grs-sim.de/cms/upload/Carloni/Tutorials/FMCP/Thermostats_and_Barostats.pdf
- [13] Purdue university, department of chemistry, accessed on 2021/07/24 <https://www.chem.purdue.edu/gchelp/vibs/h2o.html>
- [14] Calvo, F., J. Galindez, and F. X. Gadéa. "Sampling the configuration space of finite atomic systems: How ergodic is molecular dynamics?." The Journal of Physical Chemistry A 106.16 (2002): 4145-4152.
- [15] van den Akker, E. A. T., et al. "Thermodynamic analysis of molecular dynamics simulations of evaporation and condensation." molecules 1001 (2008): 1-8.

- [16] Van Duin, Adri CT, et al. "ReaxFF: a reactive force field for hydrocarbons." *The Journal of Physical Chemistry A* 105.41 (2001): 9396-9409.
- [17] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, *J Comp Phys*, 117, 1-19 (1995)
- [18] Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38
- [19] Neese, F. (2012) The ORCA program system, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2, 73-78.
- [20] Neese, F. (2017) Software update: the ORCA program system, version 4.0, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 8, e1327.
- [21] Watcharinyanon, Somsakul. Structure of Self-Assembled Monolayers on Gold Studied by NEXAFS and Photoelectron Spectroscopy. Diss. Karlstads universitet, 2008.
- [22] Bae, Gyun-Tack, and Christine M. Aikens. "Improved reaxff force field parameters for au-s-c-h systems." *The Journal of Physical Chemistry A* 117.40 (2013): 10438-10446.
- [23] Kim, Hyojeong, et al. "Molecular dynamics study on the self-assembled monolayer grown from a droplet of alkanethiol." *The Journal of Physical Chemistry C* 118.20 (2014): 11149-11157.
- [24] Lorentz, Hendrik Antoon. "Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase." *Annalen der physik* 248.1 (1881): 127-136.
- [25] Daniel Berthelot "Sur le mélange des gaz", *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 126 pp. 1703-1855 (1898)
- [26] Saha, Joyanta K., et al. "Small size limit to self-assembled monolayer formation on gold (111)." *The Journal of Physical Chemistry C* 115.27 (2011): 13193-13199.
- [27] Zhao, Xiongce, Yongsheng Leng, and Peter T. Cummings. "Self-assembly of 1, 4-benzenedithiolate/tetrahydrofuran on a gold surface: a Monte Carlo simulation study." *Langmuir* 22.9 (2006): 4116-4124.
- [28] Hagita, Katsumi, Susumu Fujiwara, and Nobuyuki Iwaoka. "Structure formation of a quenched single polyethylene chain with different force fields in united atom molecular dynamics simulations." *AIP Advances* 8.11 (2018): 115108.
- [29] it.wikipedia.org accessed on on 21/08/2021 <https://it.wikipedia.org/wiki/File:Ciclopentadiene.svg>
- [30] Frison, Gilles, and Alain Sevin. "eta (5)-and eta (6)-Coordinations Revisited: An ELF Study of Ferrocene and Dibenzenechromium." *Internet Electronic Journal of Molecular Design* 3.5 (2004): 222-232.
- [31] Richards Center at Yale University, accessed on 21/08/2021, http://www.csb.yale.edu/userguides/datamanip/autodock/html/Using_AutoDock_305.a.html
- [32] Lopes, Jose N. Canongia, P. Cabral do Couto, and Manuel E. Minas da Piedade. "An all-atom force field for metallocenes." *The Journal of Physical Chemistry A* 110.51 (2006): 13850-13856.
- [33] Faccts (Fast & Accurate Computational Chemistry Tools), accessed on 14/09/2021 <https://www.faccts.de/orca/>
- [34] Mathwork, accessed on 14/09/2021 <https://it.mathworks.com/help/matlab/ref/interpft.html#bvmwngw-5>

- [35] Fraser, Donald. "Interpolation by the FFT revisited-an experimental investigation." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.5 (1989): 665-675.
- [36] Rai, Beena, et al. "Molecular dynamic simulations of self-assembled alkylthiolate monolayers on an Au (111) surface." *Langmuir* 20.8 (2004): 3138-3144.
- [37] Jorgensen, William L., David S. Maxwell, and Julian Tirado-Rives. "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids." *Journal of the American Chemical Society* 118.45 (1996): 11225-11236.
- [38] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* 1994, 98, 11623-11627
- [39] Kohn, Walter, and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects." *Physical review* 140.4A (1965): A1133.
- [40] Majumder, Chiranjib, Hiroshi Mizuseki, and Yoshiyuki Kawazoe. "Thiophene thiol on the Au (111) surface: Size-dependent adsorption study." *The Journal of chemical physics* 118.21 (2003): 9809-9813.
- [41] Yourdshahyan, Yashar, and Andrew M. Rappe. "Structure and energetics of alkanethiol adsorption on the Au (111) surface." *The Journal of Chemical Physics* 117.2 (2002): 825-833.
- [42] Müller-Meskamp, Lars, et al. "Structural ordering of ω -ferrocenylalkanethiol monolayers on Au (1 1 1) studied by scanning tunneling microscopy." *Surface science* 603.4 (2009): 716-722.
- [43] Nerngchamnong, Nisachol, et al. "The role of van der Waals forces in the performance of molecular diodes." *Nature nanotechnology* 8.2 (2013): 113-118.
- [44] Méndez De Leo, Lucila P., et al. "Molecular and electronic structure of electroactive self-assembled monolayers." *The Journal of chemical physics* 138.11 (2013): 114707.
- [45] Nerngchamnong, Nisachol, et al. "Supramolecular structure of self-assembled monolayers of ferrocenyl terminated n-alkanethiolates on gold surfaces." *Langmuir* 30.44 (2014): 13447-13455.
- [46] Watcharinyanon, Somsakul, Ellen Moons, and Lars SO Johansson. "Mixed self-assembled monolayers of ferrocene-terminated and unsubstituted alkanethiols on gold: surface structure and work function." *The Journal of Physical Chemistry C* 113.5 (2009): 1972-1979.
- [47] Yao, Xin, et al. "Study of the ferrocenylalkanethiol self-assembled monolayers by electrochemical surface plasmon resonance." *Sensors and Actuators B: Chemical* 122.2 (2007): 351-356.
- [48] Feng, Yanqi. "Odd-Even Effects in Electroactive Self-Assembled Monolayers." (2018).
- [49] Pu, Qing, et al. "Molecular dynamics simulations of stretched gold nanowires: The relative utility of different semiempirical potentials." *The Journal of chemical physics* 126.14 (2007): 144707.
- [50] Olmos-Asar, Jimena A., Arnaldo Rapallo, and Marcelo M. Mariscal. "Development of a semiempirical potential for simulations of thiol-gold interfaces. Application to thiol-protected gold nanoparticles." *Physical Chemistry Chemical Physics* 13.14 (2011): 6500-6506.

- [51] Arima, Valentina, et al. "Toward quantum-dot cellular automata units: thiolated-carbazole linked bisferrocenes." *Nanoscale* 4.3 (2012): 813-823.
- [52] Wikipedia, <https://it.wikipedia.org/wiki/Carbazolo#/media/File:Carbazole.png>, visited on 1/12/2021, image courtesy of user Edgar181.
- [53] Pulimeno, Azzurra, et al. "Understanding a bisferrocene molecular QCA wire." *Field-coupled nanocomputing*. Springer, Berlin, Heidelberg, 2014. 307-338.
- [54] Ardesi, Yuri, et al. "SCERPA: A self-consistent algorithm for the evaluation of the information propagation in molecular field-coupled nanocomputing." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 39.10 (2019): 2749-2760.
- [55] Volpintesta, Daniele. *Molecular Dynamics Simulations of Nanofabrication Processes*. Diss. Politecnico di Torino, 2021.