Use of reactive force fields for the simulation of metallic nanoparticles

Adam Lahouari

▶ **To cite this version:**

Adam Lahouari. Use of reactive force fields for the simulation of metallic nanoparticles. Theoretical and/or physical chemistry. Sorbonne Université, 2024. English. NNT : 2024SORUS223 . tel-04773983

HAL Id: tel-04773983

https://theses.hal.science/tel-04773983v1

Submitted on 8 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Use of reactive force fields for the simulation of metallic nanoparticles

Utilisation de champs de forces réactifs pour la simulation de nanoparticules métalliques

Doctor of Philosophy of Chemistry at Sorbonne Université

Doctoral school 388: Physical-Chemistry and analytical chemistry of Paris Centre

Supervised by Johannes Richardi
Co-supervised by Jean-Philip Piquemal

Defended at Sorbonne University on the 27th September 2024 by

Adam LAHOUARI

Jury composition

Hannu Häkkinen Professor at University of Jyvaskyla, Finland	Reviewer
Christine Mottet Research Director at CNRS, Université Aix-Marseilles, France	Reviewer
Virginie Marry Professor at Sorbonne University, France	Examinator
Stephan Steinmann Research director at CNRS, ENS Lyon, France	Examinator
Thomas Plé Researcher at CNRS, Sorbonne University, France	Invited
Johannes Richardi Associated Professor at Sorbonne University, France	Thesis director
Jean-Philip Piquemal Professor at Sorbonne University, France	Thesis co-director

Acknowledgements

J’aimerais tout d’abord remercier Johannes d’avoir été là pour moi dès le début de ma thèse, de m’avoir encadré et permis de comprendre ce qu’est le métier d’enseignant-chercheur, et de m’avoir toujours montré la bonne route à suivre. Je remercie également Jean-Philippe de m’avoir offert l’opportunité de voyager et de comprendre l’importance de communiquer avec des chercheurs de tous horizons et de tous pays. C’est en partie pour cette raison que j’ai décidé de partir pour les États-Unis. Thomas, je te remercie de m’avoir épaulé dans ce domaine qu’est le machine learning, mais aussi pour les repas et les afterworks toujours agréables en ta compagnie.

Ensuite, j’aimerais remercier ma famille. Tout d’abord, ma mère, qui ne s’en souvient probablement pas, mais à l’époque, elle avait tiré les cartes pour lire mon avenir et m’avait dit que je deviendrais docteur. Personne n’aurait imaginé qu’elle parlait de chimie. Je tiens à la remercier de tout cœur ; elle a été là dans tous les moments difficiles durant ces trois années. Je sais que le fait que je parte à l’étranger lui pèse, mais ne t’inquiète pas, je serai toujours là, et ce voyage n’est pas un aller sans retour. Je t’aime de tout mon cœur. Je tiens aussi à remercier mon père. Il m’a permis de garder la tête froide dans de nombreux moments où je ne savais plus comment agir. Ses conseils m’ont permis d’atteindre tout cela. Sans lui, je ne serais pas là. Il m’a permis d’atteindre des sommets que je ne pensais pas pouvoir gravir. Je suis vraiment fier d’être son fils.

Anna, ma sœur, je te souhaite de trouver une direction qui te plaît, sans comparaison, avec pour seul objectif d’avoir ta propre vie. Tom, mon frère, je te souhaite de te retrouver pour redevenir le petit Tom mignon et rempli d’amour que tu étais.

Maintenant, parlons de mes supers amis. D’abord, Monsieur le Docteur Samuel Del Fré. On s’est rencontrés sur les bancs de la licence et depuis, on ne se quitte plus. Ça fait des années qu’on se suit, même jusqu’en Pologne. On a soutenu nos thèses à un jour d’écart et célébré ensemble. Pendant ces trois ans, nous avons vécu cette aventure ensemble, en se parlant tous les jours pour plaisanter, râler ou jouer. Sache que tu resteras dans mon cœur jusqu’à la fin de ma vie et que je serai toujours là pour toi, peu importe la distance. Mais comme je te l’ai dit, ce n’est pas parce que je pars qu’il faut m’oublier et ne plus répondre à mes messages.

Lise, j'aimerais aussi te remercier pour tout ce que tu as fait. Tu as été présente dans les moments compliqués, ce qui m'a permis de garder la tête froide et de me débloquer sur bien des choses.

Je pense aussi à mes autres amis qui m'ont accompagné durant toutes ces années : Clara, la future propriétaire de toutes les villes de France ; Paul, mon meilleur compagnon de soirées et de festivals ; Alexandre, le coureur fou ; Brian, le professionnel d'Apex ; Annaëlle, la reine de la raclette ; Emma, la tranquillité et la gentillesse incarnées ; Maurine, ma coach ; Phil, la rockstar du sud-ouest ; Charles, mon ami de longue date qui a toujours été là ; Louise, ma sœur, qui me manque depuis qu'elle est partie ; Bart, le Breton, futur coureur professionnel ; Alice, la Niçoise que je vois trop peu ; Laurine, la rugbywoman ; et pour finir Nathan, mon frère.

Ensuite, mes amis du laboratoire. Je tiens tout d'abord à remercier ma marraine Johanna, qui m'a aidé durant ma première année et m'a permis de m'intégrer facilement à la vie du laboratoire. Timothée, mon ami photographe, futur garde forestier ; j'imagine venir te voir dans 20 ans à la montagne pour que tu me fasses découvrir les animaux qui vivent autour de toi. César, mon ami parisien, qui m'a montré que tous les stéréotypes sur les Parisiens ne sont pas vrais. J'avais une image un peu négative de cette ville, mais il a su me prouver le contraire. Nous serons presque voisins l'année prochaine si tu décides aussi de partir. Hâte de découvrir les États-Unis avec toi. Marina, avec qui j'ai passé toute ma deuxième année de thèse, qui m'a montré qu'il existe des personnes capables de jouer la nuit et de travailler le jour, et qui essaie tant bien que mal de répandre de la joie autour d'elle par de petites attentions.

Je pense aussi à d'autres belles rencontres faites au cours de ces années : Jessica, qui boit son thé avec du lait ; Trinidad, the office girl ; Andrea, la danseuse ; Mary, la joueuse d'échecs; et Mika, mon frère d'arme, qui a commencé sa thèse en même temps que moi et qui, j'en suis sûr, deviendra le futur doyen de la faculté.

Je tiens donc à vous remercier tous une dernière fois pour tout ce que vous m'avez apporté durant ces trois belles années passées en votre compagnie. Je ne vous oublierai jamais

A) Introduction

A.1) Metal nanoparticles and their self-assembled-monolayers

Nanoparticles (NPs), have emerged as a basis in the field of nanotechnology due to their unique physical, chemical, and biological properties.[1–3] They are composed of transition metals and exist as clusters from two to thousands of atoms. In this thesis, we focus on coinage metal that refers to metallic elements in the 11th group of the periodic table, namely gold, silver, and copper. These metals have been traditionally utilized in coin minting due to their exceptional malleability and ductility.[4] Their high ionization potential and positive standard electrode potential make them valuable for currency, owing to their resistance to corrosion and oxidation.[5] Over time, further properties have been uncovered, leading to diverse applications.

From a biological perspective, silver has demonstrated remarkable efficacy as an antibacterial agent as presented in Figure A-1. Its inhibitory effect on bacterial growth is attributed to the disruption of the bacterial electron transport chain system. Initially, silver NPs (Ag-NPs) form bonds with ions on the bacterial membrane, rich in sulfur components. Subsequently, Ag^+ ions are released within the bacterial cell membrane, traversing it effectively.[6]

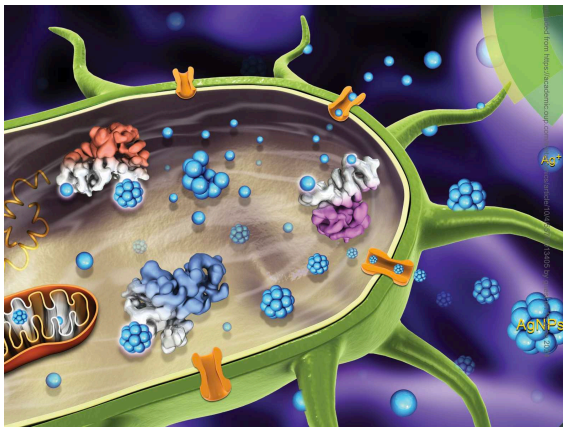


Figure A-1: Schematic of the antibacterial effect of silver nanoparticles adapted from Xueting Yan et al. (2018) [7]

Therefore, two main mechanisms happen–the disruption of the cell membrane functions–and–the generation of intracellular reactive oxygen species. Furthermore, the sustained influx of silver ions induces oxidative stress within the bacterial cell, thereby increasing the antibacterial effect and ultimately leading to bacterial death. A recent study showed the bacterial viability and

silver adsorption in Pseudomonas aeruginosa when treated with various concentrations of silver nanoparticles and silver ions.[7] It demonstrates a concentration-dependent inhibition of bacterial growth by both forms of silver as represented on Figure A-2.

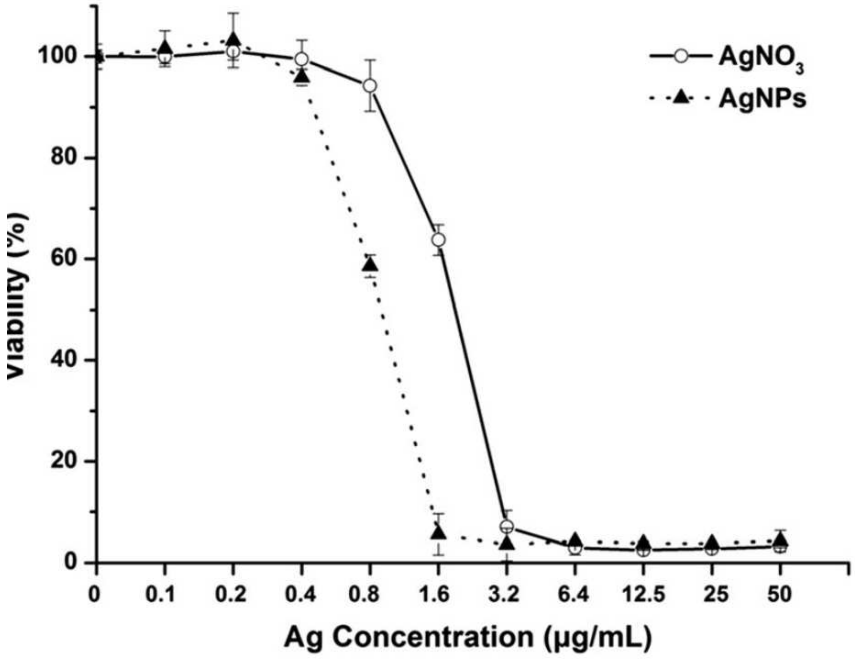


Figure A-2: Viability of a bacteria depending on the concentration of silver nanoparticles and silver ions adapted from Xueting Yan et al.[7]

The modulation of antibacterial effectiveness can be achieved through alterations in the shape, size, and concentration of silver nanoparticles. For instance, increasing the surface area of nanoparticles can enhance the release of Ag^+ ions, intensifying the antibacterial impact. Additionally, smaller-sized nanoparticles with a higher surface area exhibit faster Ag^+ ion release, further enhancing their antibacterial efficacy.[8]

Beyond antibacterial properties, Ag-NPs exhibit diverse biological functionalities, including antifungal[9], antiviral[10], antiparasitic[11], and anti-inflammatory[12] effects as represented in Figure A-3. Gold nanoparticles can be used for photothermal therapy[13], biocatalysis[14], X-ray imaging[15] , and drug delivery[16].

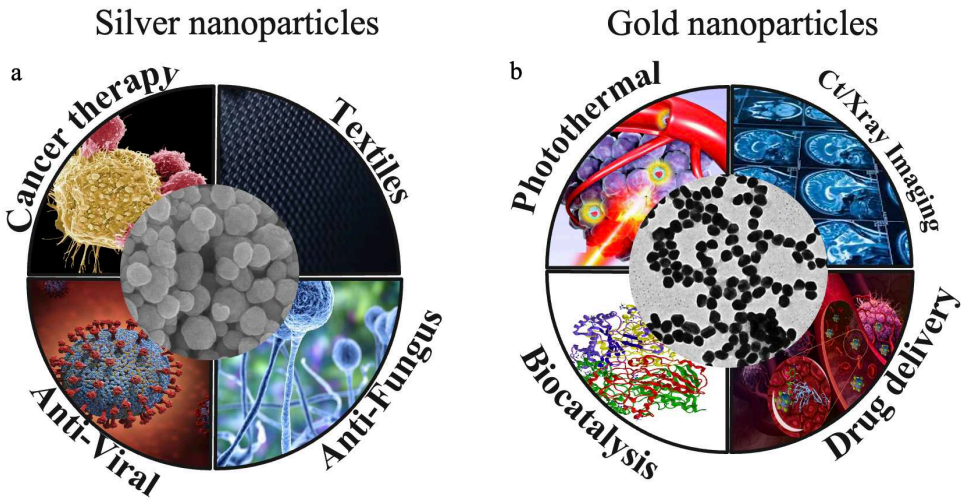


Figure A-3: Applications of (a) silver and (b) gold nanoparticles.

Experimentally, there exist various ways to prepare those NPs. The chemical, physical, and biological approach as presented on Figure A-4. Chemically, three components are needed, a metal precursor ($AgNO_3$), a reducing agent ($NaBH_4$), and then depending on the wanted size a stabilizing agent (Polyvinyl pyrrolidone).[17] A chemical reduction then occurs and depending on the agents used one can produce rather small or large nanoparticles. For example, sodium borohydride and *surfactin* can be used respectively to produce rather "small" 3-28 nm. For large ones, Ascorbic acid produces nanoparticles up to 600 nm.[18] The chemical reduction can be used also for the gold with Turkevich's method. Allowing to get spherical gold nanoparticles from 10 to 20 nm. Based on the reduction of HAuCl4 as the metal precursor by the use of trisodium citrate which plays both agent role.

This method gives high yields. Physically, the Ag-NP can be made by evaporation-condensation and laser ablation. However, it requires too much energy and is time-consuming.[18] Biologically, Ag-NP can be produced with an eco-friendly pathway without any toxic by-products and without consuming too much energies. They can be formed using algae, yeast, or plants by bio-reduction of silver ions into SNP.[19]

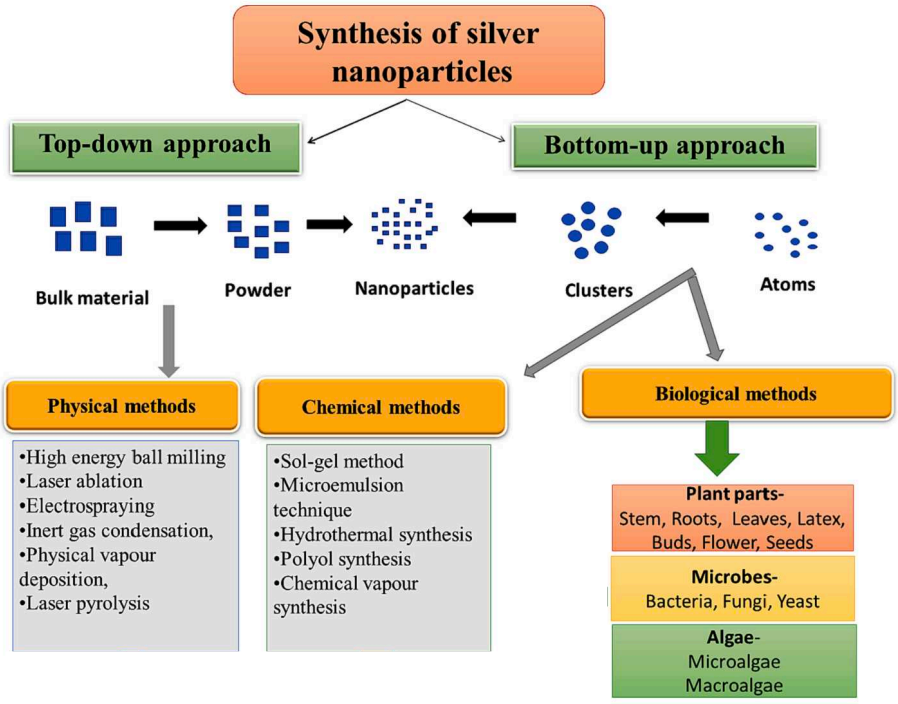


Figure A-4: Process of the synthesis of silver nanoparticles adapted from Dhaka et al. (2023)[17]

During the process, one can also easily choose the desired shape of the NP. Like spheres, pyramids, cubic, bars or prims as represented in Figure A-5.

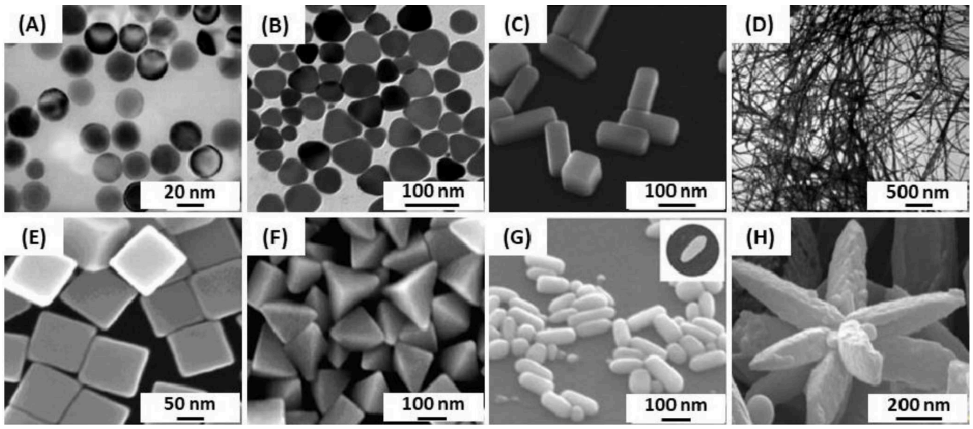


Figure A-5: Images of silver nanoparticles with different shapes and sizes made by TEM: (A) nanospheres, (B) nanoprisms, (C) nanobars and (D) nanowires. By SEM: (E) nanocubes, (F) pyramids, (G) nanorice, (H) nanoflowers Adapated from Loiseau et. Al (2019) [20]

The shapes have an impact on the reactivity of the NPs. As the reactivity depends mainly on a large proportion of facets. An important feature is that the NPs can be functionalized which involves the modification of a nanoparticle's surface with specific molecules or ligands, gives

access to desirable properties, and enables their controlled interaction with the external environment. The functionalized NP can for example be used for the immobilization of histidine-tagged proteins.[21] Silver can also be mixed with different metals in order to create alloys, that have different applications like catalysis for example.[22]

Chemical Functionalization of Nanoparticle Functionalization serves several key purposes; it can enhance the solubility of nanoparticles in various solvents and can be amphiphilic, hydrophilic, and/or hydrophobic depending on the tail of the ligands that it's attached to it.[23,24], prevent aggregation[25], create specific interaction sites for targeted binding[26], and modify the particle's optical and electronic properties.[27] In essence, the functionalization of NPs can be used to tailor specific properties as discussed before. NP can be bound with amino acids to get a biomedicine application for example.[28] Some recent studies show also an interest in the functionalization of aromatic compounds.[29] Here we focus on the study of sulfur compounds in NPs functionalization due to their strong covalent bonds with gold and silver.[30,31]

Sulfur in Nanoparticle Functionalization Sulfur atoms have six electrons in their valence shell and are highly potent to form stable compounds, particularly with transition metals, where the formation of metal-sulfur bonds is both energetically favorable and kinetically stable. One sulfur group that is used is the thiol (-SH) which are characterized by the presence of a sulfhydryl group, a functional group consisting of a sulfur atom bonded to a hydrogen atom. When thiols are exposed to the surface of a metal there is a formation of thiolates passing from R-SH to R-S-Metal, notably gold and silver, they form self-assembled monolayers (SAMs), which are highly ordered, densely packed layers of molecules attached to the metal surface. Thiolate-metal bonds are particularly strong due to the formation of a stable covalent bond between the sulfur atom and the metal atom on the nanoparticle surface. This robust interaction enables the creation of stable and functionalized nanoparticle systems.

Sulfur's Role Across Various Fields The relevance of sulfur-based functionalization extends across a wide range of fields, each showing the unique interactions between sulfur and nanomaterials for distinct purposes.

In biomedicine, for example, the functionalization of nanoparticles with sulfur-containing compounds has been important in the development of targeted drug delivery systems, bioimaging agents, and biosensors. Functionalized nanoparticles can navigate biological barriers, recognize

specific biological targets, and respond to biological stimuli, thereby enhancing the efficacy of therapeutic agents while minimizing systemic side effects.[32]

For catalysis, sulfur-functionalized nanoparticles are employed as catalysts due to their high surface area-to-volume ratio and the capacity for surface chemical modification. The presence of sulfur-based ligands can influence the electronic properties of the metal core, thereby enhancing its catalytic activity. The silver nanoparticles embedded in the sulfur-functionalized sponge act as a catalyst. For example, in the reduction of 4-nitrophenol to 4-aminophenol, the presence of silver nanoparticles speeds up the reaction.[33]

Sulfur-functionalized nanoparticles have also found significant applications in the development of chemical and biological sensors. The thiolates can be used to detect specific analytes, which are measurable changes in the optical, electrical, or magnetic properties of the nanoparticles. These changes allow for sensitive and selective detection.[34]

In summary, the functionalization of nanoparticles using sulfur-based chemistry is a versatile and powerful tool in nanotechnology. The ability of sulfur to form strong and specific bonds with metals, coupled with tunability, provides a platform for the design of advanced nanomaterials. The applications of these materials are diverse and impactful, with their use in biomedicine, catalysis, and sensing. As research progresses, the scope of sulfur in nanoparticle functionalization continues to expand, paving the way for new innovations in materials science and technology.

Introduction to Self-Assembled Monolayers (SAMs) Self-assembled monolayers (SAMs) represent a class of ordered molecular assemblies that form spontaneously on surfaces by the adsorption of an active surfactant or molecule.[35] This self-assembly process is driven by a balance of intermolecular forces and interactions with the substrate. SAMs are of great scientific and technological interest due as said before to their ability to form well-defined, functional surfaces with nanometer (nm) precision.[36] Understanding the principles of SAM formation is critical for designing surfaces with specific properties for applications in sensor technology, catalysis, and biocompatible coatings, among others.[37]

Spontaneous Organization The spontaneous organization of molecules into SAMs is essentially an energetically favorable process that occurs when molecules with a specific headgroup that has

an affinity for a particular substrate are introduced to that substrate under appropriate conditions. Molecules capable of forming SAMs have a tail group, which usually consists of an alkyl chain, and a headgroup, which interacts with the substrate. One of the main specificities of SAMs is their self-limiting growth; the formation of the monolayer stops once a single layer has formed. This characteristic is due to the saturation of available bonding sites on the substrate's surface and the equilibrium established between the adsorbed molecules and those in the surrounding phase (liquid or gas). This results in a monolayer with a uniform thickness that is determined by the length of the molecule's tail group.[38]

The definition of the assembly is based on Wood's notation.[39] Where the adsorbate unit cell depends on the substrate unit cell. They are defined by two vectors b_1 and b_2 that are based on the vectors of the unit cell a_1 and a_2 where in the fcc (111) we have $|a_1| = |a_2|$ and the assembly is written as $\left(\frac{|b_1|}{|a_1|} \times \frac{|b_2|}{|a_2|}\right) R_{angle}$ where R_{angle} is the one from a_1 and a_2. For the sake of comprehension let's give an example. For gold surfaces capped with alkythiolates, it is known that one of the most stable assembly is the $(\sqrt{3} \times \sqrt{3})R30°$, which we will call the 33 until the end of the study.[40] Where $|b_1| = \sqrt{3}|a_1|$ and $|b_2| = \sqrt{3}|a_2|$ with and R_{angle} of 30° as represented on Figure A-6. This is how we define what we call the superlattice or the adsorbate unit cell.

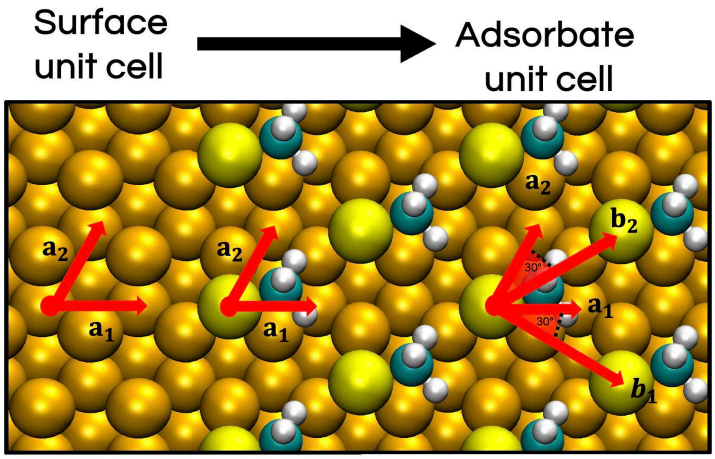


Figure A-6: Schematic of the difference between the surface and the adsorbate unit cell for gold capped with thiolates-SAMs. a1 and a2 corresponds to the unit cell vectors and b1 and b2 to the adsorbate vectors.

Moreover, for gold the 33 is not the only one that has been found. A competition between the 33 structure and the $c(4 \times 2)$ which is accompanied by an important reconstruction of the

surface.[40] Gold atoms are going out of the surface which means that there is reactivity occurring where there is the breaking of their bond followed by the formation of new ones with sulfurs as depicted in Figure A-7. It forms what we call staples.

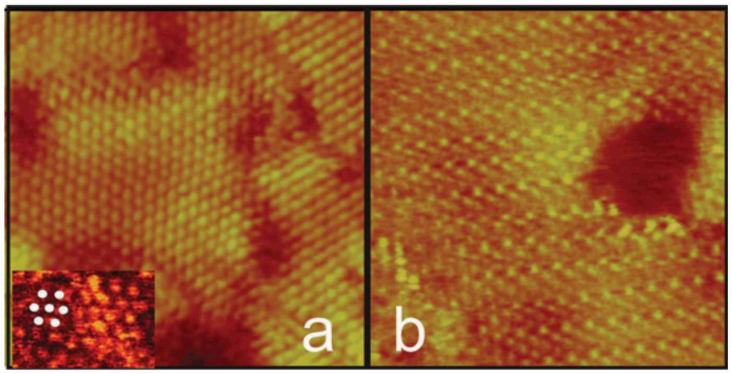


Figure A-7: STM images of (a) 33 decanethiol-SAM lattice and (b) c(4x2) hexanethiol lattice adapted from Vericat et al. (2010)[40]

Nevertheless, one should be careful about the length of the alkane-chain. Lin tang et al. found that for methane and butane thiolates no trace of 33 is found but at low temperatures, it is the (3 x 4) assembly that is covering the surface.[41] By annealing the temperature until 330K this assembly is totally replaced by a stripe phase where Au-adatoms are formed ("staples"), a binding between CH_3-S-Au-S-CH_3 but there is no trace of 33 superlattices. In any cases, the total coverage θ_S is the same for the three assemblies with a value of 0.33.

For the silver with a not-excessive concentration of alkyl-thiolates the $(\sqrt{7} \times \sqrt{7})R19.1°$ assembly is found. Nevertheless, when one increases the concentration of alkyl-thiolates we have a reactivity happening on the surface called sulfidation which creates an Ag_2S core-shell on the NP as presented in Figure A-8.[42]

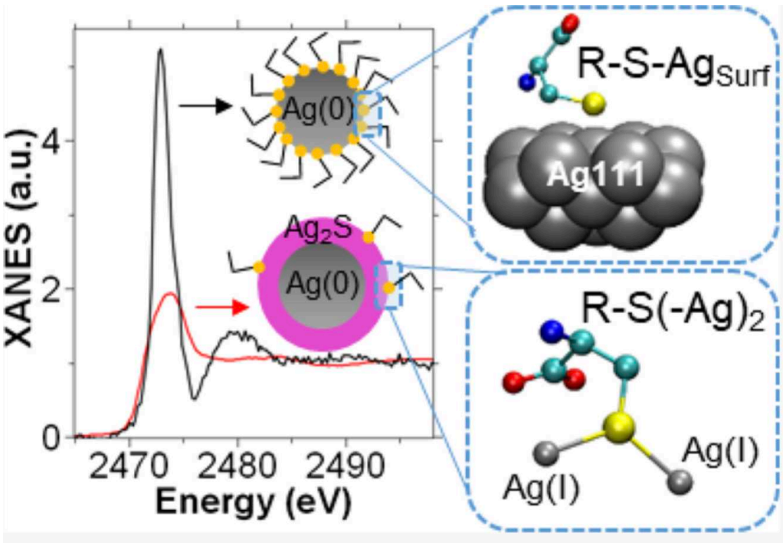


Figure A-8: XANES spectrum of silver nanoparticles capped with R-S. Showing both assembly that are formed during the capping. Adapted from Marchioni et al. (2020)[42]

Driving Forces Behind SAM Formation Several intermolecular and surface forces drive the formation of SAMs. These forces govern the ordering, orientation, density, and overall stability of the monolayer. One of the strongest interactions involved in the formation of SAMs is covalent bonding between the headgroup of the assembling molecule and the substrate. These metal-sulfur bonds are robust and result in highly stable monolayers. [43] The tail groups of the SAM-forming molecules, being long hydrocarbon chains, typically engage in van der Waals interactions, which include London dispersion forces and dipole-dipole interactions. Though individually weak, these forces collectively play a significant role in the lateral organization and packing of the molecules within the monolayer.[43]

Hydrophobic Interactions: One can use hydrophobic alkyl chains to avoid contact with polar solvents contributes to the self-assembly process. When SAMs form on substrates immersed in a solvent, the hydrophobic tails aggregate to minimize their exposure to the solvent, while the polar or reactive headgroups interact with the substrate. These hydrophobic interactions drive the vertical orientation of the molecules and are key to forming a dense, ordered monolayer.[44]

Electrostatic Interactions: For SAMs forming on charged or polar substrates, electrostatic interactions can also be influential. Molecules with charged or polar headgroups can be attracted

to a substrate with opposite charge or polarity, leading to an ordered arrangement of the molecules on the surface.[45]

Hydrogen Bonding: In some cases, hydrogen bonding can contribute to the organization and stability of SAMs. This is particularly relevant when functional groups capable of forming hydrogen bonds are present in the headgroups or at the terminal end of the tail groups.[46]

Thermodynamics and Kinetics of SAM Formation The thermodynamics of SAM formation are governed by a decrease in free energy associated with the process, making it spontaneous. This decrease in free energy comes from the various interactions described above and can include enthalpic and entropic contributions. The kinetics, on the other hand, describe the rate at which SAM formation occurs. Factors influencing the kinetics include the concentration of the assembling molecules, the temperature of the system, and the nature of the substrate. Kinetic barriers may need to be overcome for the formation of well-ordered SAMs. For example, the initial adsorption of molecules onto the substrate can occur in a disordered way. With time, thermal motion allows the molecules to reorganize into a more thermodynamically stable, ordered state. This process is often referred to as "annealing" in the context of SAMs and can result in increased crystallinity and better-defined structural properties.[47]

Structural Considerations The final structure of a SAM is influenced by the intrinsic properties of the assembling molecules and the substrate, as well as the conditions under which self-assembly takes place. The length and flexibility of the tail groups, the size and reactivity of the headgroups, and the cleanliness and crystallography of the substrate surface all factor into the resultant SAM structure. Furthermore, environmental conditions, such as solvent, temperature, and atmosphere, play roles in the quality of the SAMs formed. Through manipulations of these variables, researchers can exert control over the SAM properties, tailoring the monolayers for various specific applications. The precise control over the interfacial properties offered by SAMs is one of the key reasons why they have become a fundamental tool in nanotechnology and surface science. The rich interplay between molecular interactions, surface affinities, and assembly conditions has made the study of SAMs a dynamic field of research, revealing fundamental insights into molecular self-assembly processes. As we move forward to examine the diversity among SAMs and their functional capabilities, it becomes clear that the capacity to tune these interfaces has

profound implications for the advancement of nanotechnology, particularly in the area of nanoparticle functionalization.

Different Types of SAMs To fully understand why thiolates hold in the self-assembled monolayers (SAMs) formation an important role, it is crucial to first explore the broader landscape of SAMs, their diverse chemical structures, and the resulting properties which influence their selection for specific applications. Different types of SAMs differ in their chemical nature, stability, ease of formation, and their potential to modify nanoparticles.

Amine-Based SAMs SAMs containing amine groups at their terminal end can be adsorbed onto various substrates, forming either electrostatic or covalent bonds depending on the substrate material. Amine-based SAMs are often used on gold surfaces, where they can form relatively stable layers. However, they do not provide the same level of stability as thiolate-gold bonds. On the other hand, amine-based SAMs have found significant use compared to thiolates-SAMs as they have a greater resistance to oxidation.[48]

Thiolates-SAMs Thiolate-based SAMs on gold are not only stable but also exhibit ease of formation. They can assemble from either solution or gas phase, and the process can be effectively controlled to yield well-defined monolayers with desirable properties such as controlled thickness, terminal functionality, and uniformity. Moreover, as said before the densely packed property of the SAM is led by the Van-Der-Waals forces that are on the alkyl-chain as presented in Figure A-9.

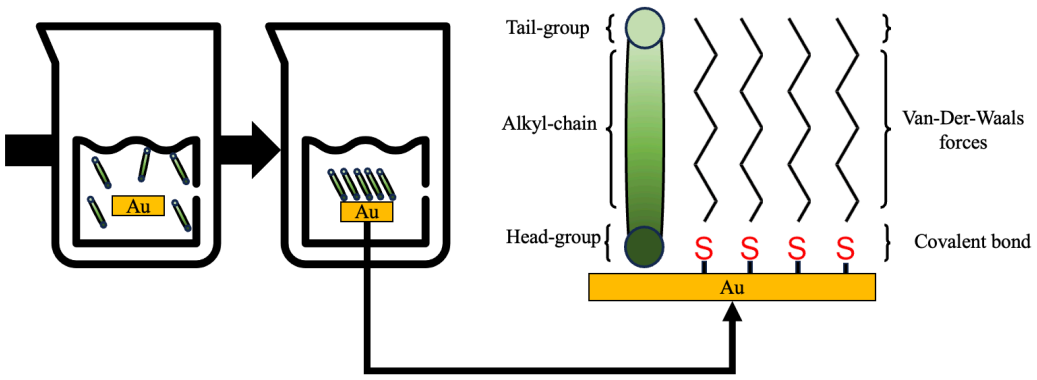


Figure A-9: schematic of the reaction between alkyl-thiolates and gold surface.

The tail group is a site that can be reactive – depending on its composition – and that can either bind itself to other surfaces (or nanoparticles). A crucial part as it gives the specific desired solvability of the nanoparticles.

Synthesis and Characterization of Thiol-SAMs Building upon their unique properties, the synthesis of thiol-based SAMs is relatively straightforward, involving the immersion of the metal substrate into a solution containing thiol molecules. The solvent not only serves as a medium for the thiols to reach the surface but also acts to solvate the alkanethiol chains and promote an ordered assembly as the SAM forms. The process can take anywhere from a few minutes to 24 hours, depending on the desired quality and thickness of the SAM, and usually at room temperature, although temperature can be used to control the kinetics of SAM formation.

However, the conditions such as solvent type, concentration, temperature, and time, need to be optimized for the formation of a well-ordered SAM. Upon adsorption, the S-H bond is cleaved, and the sulfur atom binds covalently to the gold surface, resulting in the liberation of hydrogen. This interaction not only anchors the thiol to the gold substrate but also leads to the self-assembly of an ordered monolayer due to the strong affinity between gold and sulfur, which is manifest in a high binding energy. It has been found that the binding energy for sulfur in the form of bound thiolate was around 42 kcal/mol with no restructuration and 50 kcal/mol with the formation of staples.[49] The strong binding energy shows the stability of the thiol-gold bond, and it is this robustness that gives a monolayer with remarkable thermal and chemical stability. The resulting monolayer, characterized by its homogeneity and orderliness, presents a lot of possibilities for nanotechnology applications. A well-defined monolayer is crucial for applications where a predictable and uniform surface chemistry is required.

Following synthesis, characterization techniques such as X-ray photoelectron spectroscopy (XPS), S K-edge X-ray absorption near edge structure (XANES), and atomic force microscopy (AFM) are employed to verify the presence, quality, and order of the SAMs.[50]

XPS is a surface-sensitive quantitative spectroscopic technique that measures the elemental composition, chemical state, and electronic state of the elements that exist within a material. In the context of thiol-SAMs, XPS can be utilized to verify the presence and quantify the coverage of sulfur on the metal surface, offering insight into the binding efficiency and integrity of the monolayer. The binding energies associated with the sulfur 2p and gold 4f orbitals are of particular

interest.[51] Shifts in these binding energies can indicate the formation of the gold-thiolate bond, which is a definitive sign of successful SAM formation. Furthermore, XPS allows for the detection of contaminants or oxidation products that may be present on the SAM as presented in Figure A-10.[30]

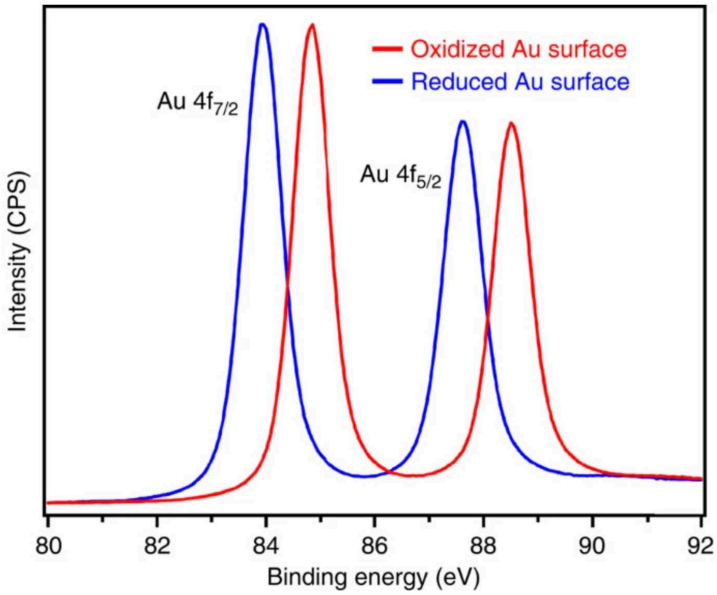


Figure A-10: Oxidization and reduction of gold surfaces by the use of XPS. Adapted from Xue et al. (2014)[30]

By analyzing the chemical environment of the sulfur atoms, researchers can distinguish between different oxidation states of sulfur, enabling the identification of specific reactivity with for example the disulfide binding interactions.[52]

XANES allows the study of the stability of thiolates-SAMs on gold or silver. It allows us to analyze the coordination environment of sulfur atoms within the thiolate groups, thus understanding specific interactions like the formation of Ag@Ag2S core@shell or staple formation for gold. It examines absorption near the core-level binding energies. The investigation of the adsorption site can also be made as presented in Figure A-11.

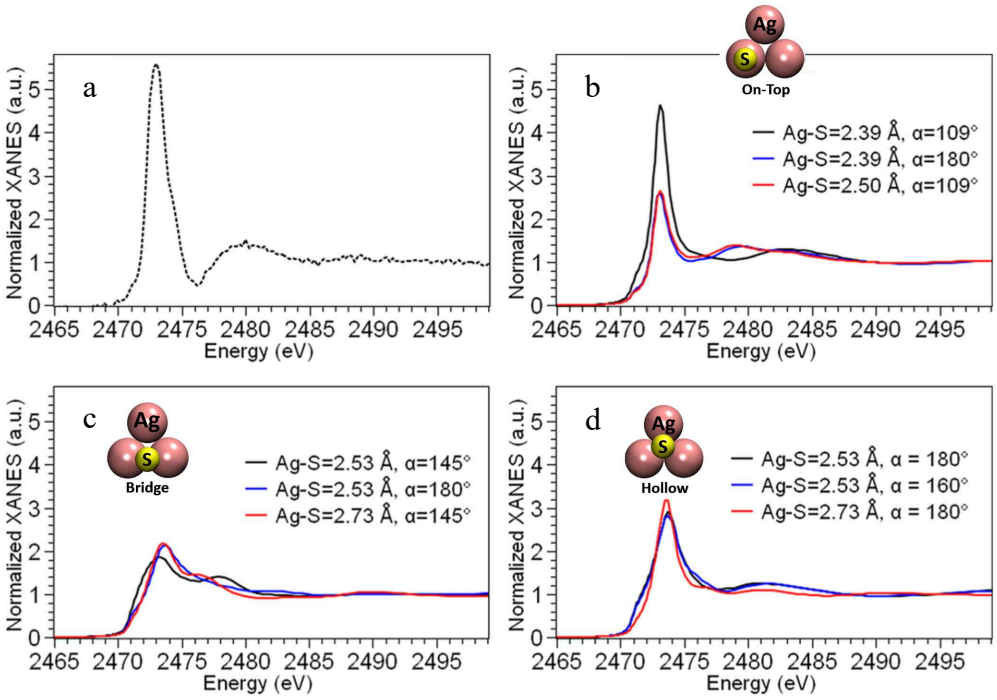


Figure A-11: XANES spectra of the thiolates-SAMs on Ag (111) surface. (a) corresponds to the experimental result compared to (b), (c), and (d) which are the DFT-based spectra for the on-top, bride, and hollow sites with variable Ag-S distance and bending angle. Adapted from Marchioni et al. (2020) [42]

Shows that the adsorption sites for the thiolate-SAMs on the Ag(111) are the on-top, bridge, and hollow with a preference for the former.[42]

AFM, on the other hand, provides a topographical map of the surface, offering nanoscale-resolution images that can reveal the physical structure of the SAM. Through AFM, one can determine the uniformity and homogeneity of the monolayer, detect the presence of domain boundaries, and measure the thickness and roughness of the SAM.[53] A well-ordered thiol-SAM typically presents a smooth and uniform surface, while disruptions in the homogeneity may suggest incomplete surface coverage or the presence of multilayers. Gao et al. did a measurement using AFM of the interaction between ultra-thin thiol SAM and a gold surface. Also, forces can be determined using AFM allowing to see the influence of the pH on the strength of S-Au contacts as presented in a study by Xue et al. in Figure A-12.[30]

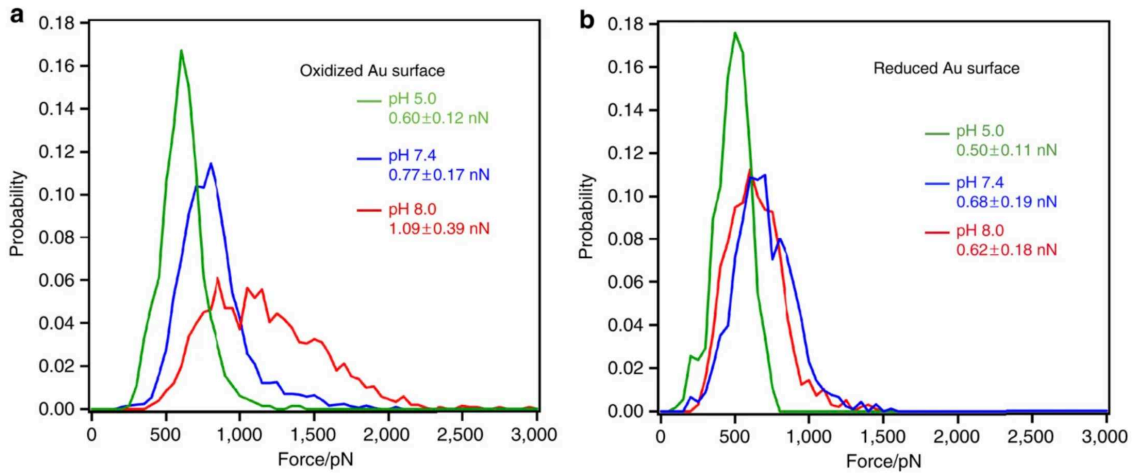


Figure A-12: Analyze of the forces/pN for the S-Au bonds for oxidized and reduce gold surfaces depending on the pH. Adapted from Xue et al. (2014)[30]

Other techniques are also employed in conjunction with XPS and AFM to gain a more comprehensive understanding of thiol-SAMs. For instance, ellipsometry can be used to monitor adsorption/desorption processes and thickness. Moreover, it can be combined with AFM data presented by Solano et al. it can help in calculating the SAM's packing density. Contact angle goniometry provides information on the hydrophilic or hydrophobic nature of the SAM, which is critical for applications involving interfaces with biological systems or environmental factors.[54]

One should also note the importance of infrared spectroscopy (IR) and surface plasmon resonance (SPR) in the characterization of SAMs. IR spectroscopy can identify the functional groups present on the SAM and confirm their orientation and binding state with the detection in the vibrational modes of thiols. SPR, meanwhile, is a sensitive method for studying the optical properties of SAMs and can be utilized to monitor the process of monolayer formation in real time.[55]

The synthesis and characterization of thiolate-SAMs thus represent a synergy between simple chemisorption processes and advanced analytical techniques. Through this, researchers are able to fabricate monolayers with tailored functionalities and verify their quality and structure, ensuring their suitability for the intended nanotechnology applications. As the field continues to mature, these synthesis and characterization techniques will undoubtedly evolve, potentially enabling even greater control over the properties and performance of thiolate-SAMs.

Beyond the stability of the thiolate-gold bond, the chemical versatility provided by the thiol group plays an equally significant role for the applications of SAMs. The terminal end of the thiol, can be modified with various functional groups, leading to different surface properties to the SAM. This enables the engineering of surfaces that are hydrophilic or hydrophobic, positively or negatively charged, or reactive towards specific biomolecules. Such versatility is imperative when designing nanoparticles for different environments, such as in aqueous biological systems or non-polar organic solvents.

For instance, the presence of a carboxylic acid group (-COOH) at the end of a thiol molecule provides reactive sites for the conjugation of peptides, proteins, or DNA, turning the nanoparticle into a platform for biosensing or gene delivery.[56] Hydroxyl (-OH) terminated thiolate can be used to anchor polymers or to create surfaces that resist protein adsorption, a feature desirable in medical implants.[57] Additionally, the thiolate group itself can serve as a synthetic handle for the attachment of more complex molecules, such as PEG (polyethylene glycol), increasing their circulation time in the bloodstream.[45]

In summary, the chemistry of thiols and their interaction with gold surfaces result in SAMs that are not only stable but also versatile in their functionality. The strength and specificity of the thiol-gold interaction led to well-defined monolayers that are essential for the reliable performance of nanotechnology applications. The functional diversity provided by the thiols enables the creation of surface chemistries tailored to a wide array of scientific and industrial needs, consolidating the position of thiol-based SAMs as a cornerstone in nanoparticle functionalization.

Advantages of Thiolates-SAMs over Other SAMs Advantages of thiol-SAMs are abundant, as expressed before – the strong covalent bond – the van der Waals forces and its head-group allowed thiol-SAMs to have specific properties. Nevertheless, it is not only a matter of bond strength. The nature of the bond imparts stability to the monolayer, allowing it to withstand a range of environmental conditions, including variations in pH, temperature, and the presence of solvents.[58,59, 60] In contrast for other functional groups used in SAMs, such as silanes, even if they are more stable at high temperatures the strength of the bond is lower than thiolate-SAMs with weaker attachment and reduced monolayer stability.[61]

By contrast, other types of SAMs may require more complex syntheses involving multiple steps, stricter environmental controls, or more specialized equipment. For example, silane-based

SAMs often demand anhydrous and oxygen-free environments since the presence of water or oxygen can prematurely hydrolyze or oxidize the silane, complicating the preparation process.[62] In addition, silane SAMs may require curing at elevated temperatures to achieve optimal adhesion, further complicating their use.[63,64] Moreover, thiol-SAMs can serve as templates for further modification. For instance, the thiol group is an ideal precursor for reactions with maleimides[65] or disulfides[66], allowing for the attachment of more complex molecules or the formation of multilayer structures. Consequently, thiol-SAMs can be engineered to present active biomolecules, such as enzymes or antibodies, in an oriented manner, vastly increasing the efficiency of biosensors or biochips.

The ability to precisely control the density and orientation of the functional groups on a thiol-SAM is another important aspect. By adjusting synthesis parameters or employing mixed SAMs (combinations of different thiols), researchers can dictate the spatial arrangement of functional groups at the nanoscale, influencing the properties of the surface in a controlled manner.

In comparison, other SAM-forming groups may not offer the same level of functional diversity or may not be as compliant with post-assembly modifications. For instance, silanes have a limited scope of functionalizable end groups due to steric hindrance and susceptibility to hydrolysis[67], while phosphonates, despite having a robust attachment, lack the ease of formation of the thiols.[68]

The ability to modify the surface of a thiolate-SAM post-assembly extends its advantages well beyond the initial surface treatment. This post-functionalization flexibility is important for applications that require dynamic or multi-step surface modifications. For example, in drug delivery systems, thiol-SAM-coated nanoparticles can be conjugated with targeting ligands and therapeutic agents.[69] The biocompatibility and ease of introducing bioactive molecules on the thiol-functionalized surface also make thiol-SAMs ideal for medical device coatings and tissue engineering scaffolds.[70]

Additionally, thiol-SAMs can be utilized in fundamental research to study intermolecular interactions at interfaces. Due to the control over surface chemistry that thiol-SAMs afford, they can act as model systems for exploring the principles governing molecular recognition, adhesion, and electron transfer even in recent studies.[71,72,73]

It's worth mentioning that while other SAM-forming compounds can also be used for subsequent modifications, the straightforward reactivity of thiols combined with their strong adherence to metal surfaces typically makes them the preferred choice for most applications that rely on post-assembly modifications. Overall, the inherent chemistry of thiols with metals, their simplicity of synthesis, and the multifunctional nature of thiol-SAMs define them as a superior choice for the functionalization of nanoparticles and other nanotechnological applications. These advantages make thiol-based SAMs indispensable tools in the expanding field of nanotechnology and are a testament to their unparalleled utility in both industry and academia.

Challenges and Limitations of Thiol-SAMs Despite the numerous advantages that thiol-SAMs offer in nanoparticle functionalization, they are not without their limitations and challenges. As with any technology, understanding these restrictions is essential for developing strategies to overcome them and improving the overall functionality of these systems. They are then discussed in the following.

1. **Desorption of Thiol-SAMs** Desorption, the process by which molecules detach from the surface they are bound to, represents another significant challenge. Thiol-SAMs can desorb from surfaces under certain conditions, particularly when subjected to mechanical stress[74], exposure to solvents[75], or elevated temperatures.[76] Desorption compromises the stability of the functionalized surface and is of particular concern in applications involving fluidics or where the nanoparticles are used in biological environments. The development of more complex SAM architectures is one approach to enhance the resilience against desorption.[49] The use of bidentate or multidentate ligands, which bind to the surface at multiple points, can reduce the likelihood of desorption. For example, dithiols (with two thiol groups) can form more stable SAMs than their monothiol counterparts.[77] Crosslinking on the surface of the nanoparticles[78] or between nanoparticles can also enhance stability.[79]

2. **Assembly Defects and Quality Control** Uniformity and defect control in SAM assembly are additional challenges. SAMs are typically formed under the assumption that each thiol molecule will find and bind to an available gold site in an orderly way. However, in practice, the assembly is often imperfect, leading to defects such as pinholes, overpacking,

or incomplete surface coverage. These imperfections can adversely affect the SAM's physical and chemical properties and can be particularly problematic when monolayer perfection is required. For silver, it can lead to the sulfidation phenomena which is accompanied by a high concentration of thiols that are creating a core-shell structure of silver sulfide on the silver nanoparticles.[80] For gold nanoparticles, different assemblies are found with the formation of staples that gives enhancement of stability.[81] Those phenomena are explained in more detail in section B-1. Advances in synthetic chemistry have allowed for the creation of thiol compounds with greater purity and uniformity, which has helped in reducing defects. Moreover, sophisticated characterization techniques such as AFM, X-ray photoelectron spectroscopy (XPS)[82], and scanning tunneling microscopy (STM)[83] enable researchers to detect and quantify defects in SAMs. By identifying the types of defects present and understanding their formation mechanisms, one can fine-tune the assembly process to minimize their occurrence.

3. **Environmental Sensitivity and Chemical Selectivity** Thiolate-SAMs are also sensitive to the chemical environment. The presence of specific ions or molecules can interfere with the assembly process or affect the monolayer's stability post-formation. For example, certain metal ions can displace the thiol from the surface, or catalyze unwanted side reactions, leading to degradation of the SAM.[84] To overcome this sensitivity, the use of protective coatings[85] or encapsulation strategies has been proposed. Encapsulation of the functionalized nanoparticles within a biocompatible matrix can shield them from harsh environmental conditions while still allowing for the intended interactions with the target analytes.[86] Additionally, the design of SAMs with specific chemical selectivity can prevent unwanted interactions.

4. **Challenges in Multilayer Formation** For certain applications, the construction of multilayer SAMs is desirable. However, building subsequent layers on top of an existing thiol-SAM can be challenging due to the already-passivated surface. Additional layers may not adhere as strongly or may form with less control over the molecular orientation and density, impairing the multilayer's performance.[87] Approaches to facilitate multilayer formation include the use of 'linker' molecules that can bind to the initial SAM and provide

functional groups for subsequent layer attachment.[79] Layer-by-layer techniques, where alternating layers of positively and negatively charged molecules are assembled, offer a method for building multilayer structures with more control over the interlayer interfaces than for example co-adsorption method.[88]

5. **Integration with Other Materials** Lastly, integrating thiol-SAMs with other materials can present compatibility issues. For example, thiol-SAMs formed on gold surfaces may not interact well with silicon-based electronics or certain polymer matrices.[89] Researchers are working on developing intermediate layers and surface treatments to enable better integration of thiol-SAMs with a wider range of materials. In spite of these challenges, thiol-SAMs continue to be a focus of intense research due to their remarkable potential in nanoparticle functionalization. By understanding and addressing their limitations, researchers can expand the utility of these powerful systems in various technological and scientific applications.

Current Research and Future Directions Recent years have witnessed a significant surge in research focused on enhancing the robustness and functionality of thiol-SAMs, which have been foundational in the continued advancement of nanoparticle functionalization technologies.[90,8] A cross-disciplinary approach for chemistry, materials science, and nanotechnology has led to innovative strategies to improve the stability, selectivity, and overall efficacy of these SAMs. This continued research is critical for the development of next-generation devices across a range of fields, including biosensing, catalysis, and molecular electronics.

Surface Coverage and Binding Efficiency One of the fundamental areas of research pertains to the optimization of surface coverage and binding efficiency of thiolate-SAMs. Efficient surface coverage is pivotal for the homogeneity and integrity of SAMs. Scientists have been developing new quantification methods for surface coverage of self-assembled thiol ligands bound onto GNPs.[91,92] Such accurate quantification is essential for consistency in applications such as drug delivery and biosensing. These quantification techniques, such as those using nuclear magnetic resonance (NMR) spectroscopy, are revealing insights into the precise control of thiol-gold binding stoichiometry, providing a clearer understanding of the molecular arrangement and allowing for more accurate predictions of SAM behavior in various environments.[93]

Thiol-Gold Bond Strength A particularly significant area of research delves into the thiol-gold bond itself both experimentally and theoretically.[94] Thiol-Au interactions are the cornerstone of thiol-SAM stability, yet the exact nature of this bond has been a topic of ongoing debate. Recent studies employing advanced spectroscopy and computational modeling have attempted to quantify the strength of thiol-gold interactions, aiming to fabricate more robust SAMs.[30] Investigations into the electronic structure of the thiol-gold interface, for example, may contribute to the development of SAMs that maintain their structural integrity even in the presence of competing molecules or under extreme conditions.[95]

Novel Thiol Derivatives and Multifunctional Ligands A substantial portion of current research is aimed at designing novel thiol derivatives that can overcome the intrinsic limitations of traditional thiols. Such research includes the synthesis of multifunctional thiol ligands that offer multiple points of contact with the metal surface, thus enhancing stability. There is also considerable interest in creating responsive thiol-SAMs that can change their properties in reaction to external stimuli such as light[96], pH[97], or redox conditions[98], thus opening the door to new nanoparticle systems with switchable functionalities.

Biocompatibility and Biomedical Applications In the realm of biomedicine, the biocompatibility of thiol-SAMs is a pressing concern. With the goal of minimizing immune reactions and improving the integration of nanoparticles within biological systems, researchers are investigating the biofunctionalization of thiolate-SAMs with peptides, proteins, and other biomolecules. This research is moving toward creating SAMs that mimic natural biological processes, facilitating interactions within the human body.[79,99,100] For example, the functionalization with dithiols is being studied for the construction of more stable linkages that could withstand the dynamic environments encountered in vivo.[101]

Environmental Stability Environmental stability continues to be a major area of interest. New methodologies are being explored to shield thiol-SAMs from oxidative and thermal degradation.[102] Innovations such as the development of antioxidant-rich SAM formulations that protect against oxidative species are being assessed for their efficacy in enhancing SAM longevity.[79]

Synthesis Techniques and Scalability On the production side, the scalability of thiol-SAM synthesis is a central issue being addressed. Many thiol-SAMs are currently synthesized through batch processes that may not be suitable for large-scale production. Researchers are investigating

continuous flow synthesis and other scalable techniques that could meet the demands of industrial applications.[103] Improvements in synthesis methods not only aim to enhance scalability but also seek to reduce defects in SAMs, which could greatly improve their reliability and performance in real-world applications.[104]

Predictive Modeling and Simulation Advanced predictive modeling and computational simulations are becoming increasingly integral to thiol-SAM research. These tools are being used to predict the behavior of SAMs under various conditions, to design new SAM architectures, and to simulate the assembly process, potentially reducing the need for extensive empirical testing. Molecular dynamics (MD) using reactive force fields appeared to be the best computational tool to simulate such systems and get information on various properties.[105,106] ReaxFF is one of the methods that is used in order to see the reactivity of gold and silver nanoparticles capped with alkyl-thiolates.[107] Nevertheless, Machine learning algorithms are also beginning to make their way into the field, offering the possibility to create interatomic potential that would be used to do simulations on those systems with quantum accuracy at the cost of MD.[108,109,110] Those specific tools are described in more detail in sections A-2.

Integration with Emerging Technologies Finally, the integration of thiol-SAMs with emerging technologies is an avenue for future research. Developments in two-dimensional materials, such as graphene, present new opportunities for SAM applications.[111] Additionally, the interfacing of thiol-SAMs with silicon-based electronics and other non-metal substrates is an area of ongoing innovation, which involves the development of new strategies to create compatible interfaces and harness the unique properties of SAMs on these novel platforms.[112]

From an overall point of view, thiols-SAM capped on silver and gold are of strong use and researchers have still some information to get from them either experimentally or theoretically. In this thesis, we investigate the results and limits of the current existing methods and we are starting to create a new one that should be more precise than others that already exist using as said previously machine learning methods.

A.2) Theoretical approaches for self-assembled monolayers on nanoparticles

Different methods exist in order to simulate metallic structures for the prediction of the reactivity of SAMs on a metallic surface. For high accuracy, the use of first-principles methods like Density Functional Theory (DFT) with periodic boundary conditions is possible using

different codes like VASP or GAUSSIAN.[113] Such an accurate strategy intends to compute the electronic density via an approximation of the Schrodinger equation for many electron systems. The main drawback of this method is its cost, as DFT scales poorly with the number of atoms considered thus it makes it impractical considering the simulation of nanoscale systems.[114] Nevertheless, DFT- DFT-tight-binding (DFTB) has been introduced to simplify the problem. Simplification is introduced in the Hamiltonian as it is used restricting it to a minimal basis set and via the inclusion of empirical parameters. From that, DFTB scales more favorably with the number of atoms passing from a N^3 complexity to N. Thus, it is possible to study larger systems while keeping some essential physics of DFT. DFTB still provides information about electronic structure properties, a feature not shared by classical force fields. It can be used to study solid-state systems and complex nanostructures where a full DFT calculation would require too many resources.[115] Nevertheless, for larger-scale simulation, DFTB has its limit and the use of molecular dynamics (MD) becomes necessary.[116] MD models interaction between atoms at longer nanoscale capturing the dynamics behavior of atoms. It is based on force fields parametrized to reproduce essential physical properties without requiring too many resources.[117] Moreover, in order to strengthen the accuracy of the simulation MD reactive force fields are used.[118] Contrary to classical force fields which assume fixed bonding configurations and predefined interaction parameters, they allow the bond formation and breaking of bonds. It dynamically adjusts the bonding configurations and interaction parameters as atoms approach or move apart.

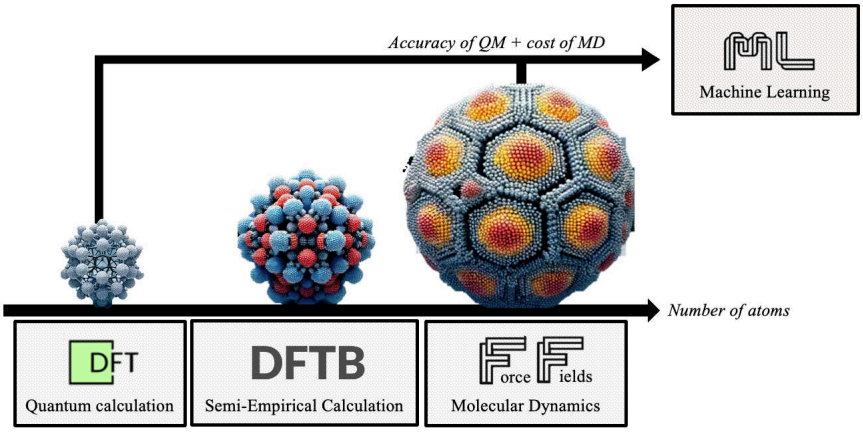


Figure A-13: Schematic showing the limit of each method for the simulation of large systems.

An important feature when one wants to study the reactivity of thiol-SAMs on metallic surfaces as a restructuration of the surface can appear in some cases. ReaxFF is a reactive force

field described in more detail in section B) that provides insights into complex mechanisms that can occur.[119] This thesis focuses on the use of this force field and studies its efficiency to accurately describe large nanoscale metallic systems capped with SAM-thiols.

Furthermore, it is noteworthy to point out that with the advent of AI and machine learning, it is now possible to create interatomic potentials from first-principles datasets. This approach facilitates the development of accurate models that combine the precision of QM with the computational efficiency of molecular dynamics (MD), as illustrated in Figure A-13.[120] A key advantage of employing machine learning potentials is their flexibility and ability to predict specific reactivity—such as staple formation—that may occur during the interaction of thiolate-SAMs with gold, as described in the previous section.

A.3) Objectives

The first goal of the thesis is to use the ReaxFF method in order to see if the method is well suited for the study of silver and gold nanoparticles (NPs) capped with alkylthiolates. For silver, from experimental studies, only one stable superstructure has been found called the $\sqrt{7x7}R19.1°$ (that we will call 77).[121] On the other hand, for gold, different superstructures have been found the $\sqrt{3x3}\ R30°$ (called 33), and the c(4x3) for alkyl-chain longer than butane. The last conformation is accompanied by a specific phenomenon favored by gold surface defects - the restructuration of the metallic surface gives the formation of staples as represented in Figure A-14.[122]

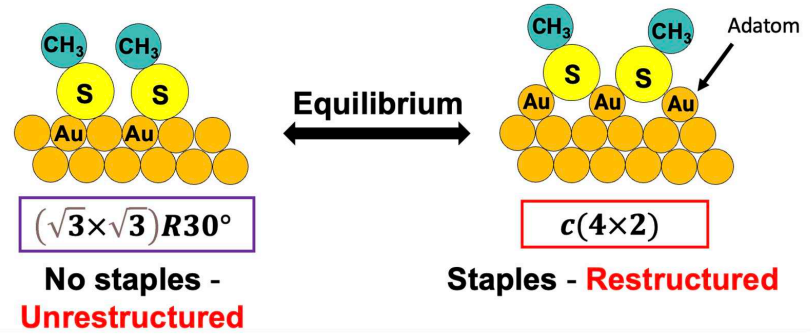


Figure A-14: Assemblies found for thiolates-SAMs on gold surfaces.

Staple formation involves the disruption of gold atom bonds, prompting these atoms to relocate between two sulfur atoms. To accurately predict this reactivity, careful selection of the computational method is essential. As previously mentioned, quantum mechanics (QM) or hybrid

QM methods are unsuitable for such large systems, which may include surfaces and nanoparticles comprising thousands of atoms. Consequently, molecular dynamics simulations using force fields become necessary. However, classical force fields inadequately describe the Ag-S bond, as they typically favor hollow adsorption sites, contradicting XANES measurements that show a preference for on-top sites.[123] Therefore, we have turned to reactive force fields, specifically employing ReaxFF to simulate bond breaking and formation during the simulation. This method also facilitates surface restructuring, striking a favorable balance between time consumption and accuracy.

Moreover, due to the limit of the current parametrization of ReaxFF (explained in section B), a new approach was needed, merging quantum accuracy and MD costs for gold systems. ML was the best option as it proved itself to be more consistent than ReaxFF in terms of complexity for the reactivity of a small system.

References

(1) De, M.; Ghosh, P. S.; Rotello, V. M. Applications of Nanoparticles in Biology. *Advanced Materials* **2008**, *20* (22), 4225–4241. https://doi.org/10.1002/adma.200703183.

(2) Salata, O. Applications of Nanoparticles in Biology and Medicine. *J Nanobiotechnol* **2004**, *2* (1), 3. https://doi.org/10.1186/1477-3155-2-3.

(3) Stark, W. J.; Stoessel, P. R.; Wohlleben, W.; Hafner, A. Industrial Applications of Nanoparticles. *Chem. Soc. Rev.* **2015**, *44* (16), 5793–5805. https://doi.org/10.1039/C4CS00362D.

(4) Lee, H. M.; Ge, M.; Sahu, B. R.; Tarakeshwar, P.; Kim, K. S. Geometrical and Electronic Structures of Gold, Silver, and Gold−Silver Binary Clusters: Origins of Ductility of Gold and Gold−Silver Alloy Formation. *J. Phys. Chem. B* **2003**, *107* (37), 9994–10005. https://doi.org/10.1021/jp034826+.

(5) Talarska, P.; Boruczkowski, M.; Żurawski, J. Current Knowledge of Silver and Gold Nanoparticles in Laboratory Research—Application, Toxicity, Cellular Uptake. *Nanomaterials* **2021**, *11* (9), 2454. https://doi.org/10.3390/nano11092454.

(6) Yin, I. X.; Zhang, J.; Zhao, I. S.; Mei, M. L.; Li, Q.; Chu, C. H. The Antibacterial Mechanism of Silver Nanoparticles and Its Application in Dentistry. *IJN* **2020**, *Volume 15*, 2555–2562. https://doi.org/10.2147/IJN.S246764.

(7) Yan, X.; He, B.; Liu, L.; Qu, G.; Shi, J.; Hu, L.; Jiang, G. Antibacterial Mechanism of Silver Nanoparticles in Pseudomonas Aeruginosa: Proteomics Approach†. *Metallomics* **2018**, *10* (4), 557–564. https://doi.org/10.1039/c7mt00328e.

(8) Li, J.; Liu, F.; Ye, J. Boosting the Brightness of Thiolated Surface-Enhanced Raman Scattering Nanoprobes by Maximal Utilization of the Three-Dimensional Volume of Electromagnetic Fields. *J. Phys. Chem. Lett.* **2022**, *13* (28), 6496–6502. https://doi.org/10.1021/acs.jpclett.2c01741.

(9) Radhakrishnan, V. S.; Reddy Mudiam, M. K.; Kumar, M.; Dwivedi, S. P.; Singh, S. P.; Prasad, T. Silver Nanoparticles Induced Alterations in Multiple Cellular Targets, Which Are Critical for Drug Susceptibilities and Pathogenicity in Fungal Pathogen (Candida Albicans). *IJN* **2018**, *Volume 13*, 2647–2663. https://doi.org/10.2147/IJN.S150648.

(10) Yang, X. X.; Li, C. M.; Huang, C. Z. Curcumin Modified Silver Nanoparticles for Highly Efficient Inhibition of Respiratory Syncytial Virus Infection. *Nanoscale* **2016**, *8* (5), 3040–3048. https://doi.org/10.1039/C5NR07918G.

(11) Vijayakumar, S.; Malaikozhundan, B.; Saravanakumar, K.; Durán-Lara, E. F.; Wang, M.-H.; Vaseeharan, B. Garlic Clove Extract Assisted Silver Nanoparticle – Antibacterial, Antibiofilm, Antihelminthic, Anti-Inflammatory, Anticancer and Ecotoxicity Assessment. *Journal of Photochemistry and Photobiology B: Biology* **2019**, *198*, 111558. https://doi.org/10.1016/j.jphotobiol.2019.111558.

(12) David, L.; Moldovan, B.; Vulcu, A.; Olenic, L.; Perde-Schrepler, M.; Fischer-Fodor, E.; Florea, A.; Crisan, M.; Chiorean, I.; Clichici, S.; Filip, G. A. Green Synthesis, Characterization and Anti-Inflammatory Activity of Silver Nanoparticles Using European Black Elderberry Fruits Extract. *Colloids and Surfaces B: Biointerfaces* **2014**, *122*, 767–777. https://doi.org/10.1016/j.colsurfb.2014.08.018.

(13) Parida, S.; Maiti, C.; Rajesh, Y.; Dey, K. K.; Pal, I.; Parekh, A.; Patra, R.; Dhara, D.; Dutta, P. K.; Mandal, M. Gold Nanorod Embedded Reduction Responsive Block Copolymer Micelle-Triggered Drug Delivery Combined with Photothermal Ablation for Targeted Cancer Therapy. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861* (1), 3039–3052. https://doi.org/10.1016/j.bbagen.2016.10.004.

(14) Plascencia-Villa, G.; Mendoza-Cruz, R. Gold Nanoparticles for Biocatalysis. In *Nanomaterials for Biocatalysis*; Elsevier, 2022; pp 377–434. https://doi.org/10.1016/B978-0-12-824436-4.00013-7.

(15) Rengan, A. K.; Jagtap, M.; De, A.; Banerjee, R.; Srivastava, R. Multifunctional Gold Coated Thermo-Sensitive Liposomes for Multimodal Imaging and Photo-Thermal Therapy of Breast Cancer Cells. *Nanoscale* **2014**, *6* (2), 916–923. https://doi.org/10.1039/C3NR04448C.

(16) Li, G.; Chen, Y.; Zhang, L.; Zhang, M.; Li, S.; Li, L.; Wang, T.; Wang, C. Facile Approach to Synthesize Gold Nanorod@Polyacrylic Acid/Calcium Phosphate Yolk–Shell Nanoparticles for Dual-Mode Imaging and pH/NIR-Responsive Drug Delivery. *Nano-Micro Lett.* **2018**, *10* (1), 7. https://doi.org/10.1007/s40820-017-0155-3.

(17) Dhaka, A.; Chand Mali, S.; Sharma, S.; Trivedi, R. A Review on Biological Synthesis of Silver Nanoparticles and Their Potential Applications. *Results in Chemistry* **2023**, *6*, 101108. https://doi.org/10.1016/j.rechem.2023.101108.

(18) Iravani, S.; Korbekandi, H.; Mirmohammadi, S. V.; Zolfaghari, B. Synthesis of Silver Nanoparticles: Chemical, Physical and Biological Methods. *Res Pharm Sci* **2014**, *9* (6), 385–406.

(19) Ameen, F.; Al-Homaidan, A. A.; Al-Sabri, A.; Almansob, A.; AlNAdhari, S. Anti-Oxidant, Anti-Fungal and Cytotoxic Effects of Silver Nanoparticles Synthesized Using Marine Fungus Cladosporium Halotolerans. *Appl Nanosci* **2023**, *13* (1), 623–631. https://doi.org/10.1007/s13204-021-01874-9.

(20) Loiseau, A.; Asila, V.; Boitel-Aullen, G.; Lam, M.; Salmain, M.; Boujday, S. Silver-Based Plasmonic Nanoparticles for and Their Use in Biosensing. *Biosensors* **2019**, *9* (2), 78. https://doi.org/10.3390/bios9020078.

(21) Abad, J. M.; Mertens, S. F. L.; Pita, M.; Fernández, V. M.; Schiffrin, D. J. Functionalization of Thioctic Acid-Capped Gold Nanoparticles for Specific Immobilization of Histidine-Tagged Proteins. *J. Am. Chem. Soc.* **2005**, *127* (15), 5689–5694. https://doi.org/10.1021/ja042717i.

(22) Front, A.; Mottet, C. Stress Effect on Segregation and Ordering in Pt–Ag Nanoalloys. *J. Phys.: Condens. Matter* **2021**, *33* (15), 154006. https://doi.org/10.1088/1361-648X/abe07a.

(23) Lim, Y.-S.; Lee, Y.-M.; Kim, J. H.; Jin, K. S.; Lee, H.-Y. Effect of Ratio of Hydrophilic and Hydrophobic Ligand Length on the Adsorption Behaviors of Amphiphilic Gold Nanoparticles at the Liquid-Liquid Interface. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **2023**, *666*, 131265. https://doi.org/10.1016/j.colsurfa.2023.131265.

(24) Guven, Z. P.; Silva, P. H. J.; Luo, Z.; Cendrowska, U. B.; Gasbarri, M.; Jones, S. T.; Stellacci, F. Synthesis and Characterization of Amphiphilic Gold Nanoparticles. *J Vis Exp* **2019**, No. 149. https://doi.org/10.3791/58872.

(25) Zhang, S.; Leem, G.; Srisombat, L.; Lee, T. R. Rationally Designed Ligands That Inhibit the Aggregation of Large Gold Nanoparticles in Solution. *J. Am. Chem. Soc.* **2008**, *130* (1), 113–120. https://doi.org/10.1021/ja0724588.

(26) Hock, N.; Racaniello, G. F.; Aspinall, S.; Denora, N.; Khutoryanskiy, V. V.; Bernkop-Schnürch, A. Thiolated Nanoparticles for Biomedical Applications: Mimicking the Workhorses of Our Body. *Advanced Science* **2022**, *9* (1), 2102451. https://doi.org/10.1002/advs.202102451.

(27) Wessels, J. M.; Nothofer, H.-G.; Ford, W. E.; von Wrochem, F.; Scholz, F.; Vossmeyer, T.; Schroedter, A.; Weller, H.; Yasuda, A. Optical and Electrical Properties of Three-Dimensional Interlinked Gold Nanoparticle Assemblies. *J. Am. Chem. Soc.* **2004**, *126* (10), 3349–3356. https://doi.org/10.1021/ja0377605.

(28) Tiwari, P. M.; Vig, K.; Dennis, V. A.; Singh, S. R. Functionalized Gold Nanoparticles and Their Biomedical Applications. *Nanomaterials (Basel)* **2011**, *1* (1), 31–63. https://doi.org/10.3390/nano1010031.

(29) Tandiana, R.; Sicard-Roselli, C.; Van-Oanh, N.-T.; Steinmann, S.; Clavaguéra, C. In-Depth Theoretical Understanding of the Chemical Interaction of Aromatic Compounds with a Gold Nanoparticle. *Phys. Chem. Chem. Phys.* **2022**, *24* (41), 25327–25336. https://doi.org/10.1039/D2CP02654F.

(30) Xue, Y.; Li, X.; Li, H.; Zhang, W. Quantifying Thiol–Gold Interactions towards the Efficient Strength Control. *Nat Commun* **2014**, *5* (1), 4348. https://doi.org/10.1038/ncomms5348.

(31) Park, S. Y.; Chung, J. W.; Priestley, R. D.; Kwak, S.-Y. Covalent Assembly of Metal Nanoparticles on Cellulose Fabric and Its Antimicrobial Activity. *Cellulose* **2012**, *19* (6), 2141–2151. https://doi.org/10.1007/s10570-012-9773-6.

(32) Magee, E. A.; Curno, R.; Edmond, L. M.; Cummings, J. H. Contribution of Dietary Protein and Inorganic Sulfur to Urinary Sulfate: Toward a Biomarker of Inorganic Sulfur Intake123. *The American Journal of Clinical Nutrition* **2004**, *80* (1), 137–142. https://doi.org/10.1093/ajcn/80.1.137.

(33) Kim, J. G.; Cha, M. C.; Lee, J.; Choi, T.; Chang, J. Y. Preparation of a Sulfur-Functionalized Microporous Polymer Sponge and In Situ Growth of Silver Nanoparticles: A Compressible Monolithic Catalyst. *ACS Appl. Mater. Interfaces* **2017**, *9* (43), 38081–38088. https://doi.org/10.1021/acsami.7b14807.

(34) Toombs, C.; Insko, M.; Wintner, E.; Deckwerth, T.; Usansky, H.; Jamil, K.; Goldstein, B.; Cooreman, M.; Szabo, C. Detection of Exhaled Hydrogen Sulphide Gas in Healthy

Human Volunteers during Intravenous Administration of Sodium Sulphide. *British journal of clinical pharmacology* **2010**, *69 6*, 626–636. https://doi.org/10.1111/j.1365-2125.2010.03636.x.

(35) Woodruff, D. P. The Interface Structure of N-Alkylthiolate Self-Assembled Monolayers on Coinage Metal Surfaces. *Phys. Chem. Chem. Phys.* **2008**, *10*, 7211.

(36) Halik, M.; Hirsch, A. The Potential of Molecular Self-Assembled Monolayers in Organic Electronic Devices. *Advanced Materials* **2011**, *23* (22–23), 2689–2695. https://doi.org/10.1002/adma.201100337.

(37) Schwartz, D. K. MECHANISMS AND KINETICS OF SELF-ASSEMBLED MONOLAYER FORMATION. *Annual Review of Physical Chemistry* **2001**, *52* (Volume 52, 2001), 107–137. https://doi.org/10.1146/annurev.physchem.52.1.107.

(38) Casalini, S.; Augusto Bortolotti, C.; Leonardi, F.; Biscarini, F. Self-Assembled Monolayers in Organic Electronics. *Chemical Society Reviews* **2017**, *46* (1), 40–71. https://doi.org/10.1039/C6CS00509H.

(39) Wood, E. A. Vocabulary of Surface Crystallography. *Journal of Applied Physics* **1964**, *35* (4), 1306–1312. https://doi.org/10.1063/1.1713610.

(40) Vericat, C.; Vela, M. E.; Benitez, G.; Carro, P.; Salvarezza, R. C. Self-Assembled Monolayers of Thiols and Dithiols on Gold: New Challenges for a Well-Known System. *Chem. Soc. Rev.* **2010**, *39* (5), 1805–1834. https://doi.org/10.1039/B907301A.

(41) Tang, L.; Li, F.; Zhou, W.; Guo, Q. The Structure of Methylthiolate and Ethylthiolate Monolayers on Au(111): Absence of the (√3 × √3)R30° Phase. *Surface Science* **2012**, *606* (5–6), L31–L35. https://doi.org/10.1016/j.susc.2011.12.005.

(42) Marchioni, M.; Battocchio, C.; Joly, Y.; Gateau, C.; Nappini, S.; Pis, I.; Delangle, P.; Michaud-Soret, I.; Deniaud, A.; Veronesi, G. Thiolate-Capped Silver Nanoparticles: Discerning Direct Grafting from Sulfidation at the Metal-Ligand Interface by Interrogating the Sulfur Atom. *J. Phys. Chem. C* **2020**, *124* (24), 13467.

(43) Mandler, D.; Kraus-Ophir, S. Self-Assembled Monolayers (SAMs) for Electrochemical Sensing. *J Solid State Electrochem* **2011**, *15* (7), 1535–1558. https://doi.org/10.1007/s10008-011-1493-6.

(44) Szöri, M.; Tobias, D. J.; Roeselová, M. Microscopic Wetting of Mixed Self-Assembled Monolayers: A Molecular Dynamics Study. *J. Phys. Chem. B* **2009**, *113* (13), 4161–4169. https://doi.org/10.1021/jp8074139.

(45) Valtiner, M.; Donaldson, S. H. Jr.; Gebbie, M. A.; Israelachvili, J. N. Hydrophobic Forces, Electrostatic Steering, and Acid–Base Bridging between Atomically Smooth Self-Assembled Monolayers and End-Functionalized PEGolated Lipid Bilayers. *J. Am. Chem. Soc.* **2012**, *134* (3), 1746–1753. https://doi.org/10.1021/ja209653n.

(46) Ramin, M. A.; Le Bourdon, G.; Heuzé, K.; Degueil, M.; Belin, C.; Buffeteau, T.; Bennetau, B.; Vellutini, L. Functionalized Hydrogen-Bonding Self-Assembled Monolayers Grafted onto SiO2 Substrates. *Langmuir* **2012**, *28* (51), 17672–17680. https://doi.org/10.1021/la303805d.

(47) Azzam, W.; Al-Rashdi, A.; Subaihi, A.; Rohwerder, M.; Zharnikov, M.; Bashir, A. Annealing Effect for Self-Assembled Monolayers Formed from Terphenylethanethiol on Au(111). *Phys. Chem. Chem. Phys.* **2020**, *22* (24), 13580–13591. https://doi.org/10.1039/D0CP02374D.

(48) de la Llave, E.; Clarenc, R.; Schiffrin, D. J.; Williams, F. J. Organization of Alkane Amines on a Gold Surface: Structure, Surface Dipole, and Electron Transfer. *J. Phys. Chem. C* **2014**, *118* (1), 468–475. https://doi.org/10.1021/jp410086b.

(49) Cortés, E.; Rubert, A. A.; Benitez, G.; Carro, P.; Vela, M. E.; Salvarezza, R. C. Enhanced Stability of Thiolate Self-Assembled Monolayers (SAMs) on Nanostructured Gold Substrates. *Langmuir* **2009**, *25* (10), 5661–5666. https://doi.org/10.1021/la804251a.

(50) Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M. Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev.* **2005**, *105* (4), 1103–1170. https://doi.org/10.1021/cr0300789.

(51) Ishida, T.; Hara, M.; Kojima, I.; Tsuneda, S.; Nishida, N.; Sasabe, H.; Knoll, W. High Resolution X-Ray Photoelectron Spectroscopy Measurements of Octadecanethiol Self-Assembled Monolayers on Au(111). *Langmuir* **1998**, *14* (8), 2092–2096. https://doi.org/10.1021/la971104z.

(52) Castner, D. G.; Hinds, K.; Grainger, D. W. X-Ray Photoelectron Spectroscopy Sulfur 2p Study of Organic Thiol and Disulfide Binding Interactions with Gold Surfaces. *Langmuir* **1996**, *12* (21), 5083–5086. https://doi.org/10.1021/la960465w.

(53) Gao, X.; Zhang, H.; Li, S.; Wang, L.; Dai, X.; Hu, Y.; Xu, J.; Jing, G.; Fan, G. In Situ Differential Atomic Force Microscopy (AFM) Measurement for Ultra-Thin Thiol SAM Patterns by Area-Selective Deposition Technique. *Surfaces and Interfaces* **2024**, *46*, 104133. https://doi.org/10.1016/j.surfin.2024.104133.

(54) Solano, I.; Parisse, P.; Gramazio, F.; Cavalleri, O.; Bracco, G.; Castronovo, M.; Casalis, L.; Canepa, M. Spectroscopic Ellipsometry Meets AFM Nanolithography: About Hydration of Bio-Inert Oligo(Ethylene Glycol)-Terminated Self Assembled Monolayers on Gold. *Phys. Chem. Chem. Phys.* **2015**, *17* (43), 28774–28781. https://doi.org/10.1039/C5CP04028K.

(55) Salsabiila, N.; Morsin, M.; Nafisah, S.; Razali, N. L.; Mahmud, F.; Alip, A. A. Investigation of the Effect of Amine and Thiol as Functional Groups on Gold Nanobipyramids Properties. *Materialia* **2024**, *34*, 102072. https://doi.org/10.1016/j.mtla.2024.102072.

(56) Irfan, M. I.; Amjad, F.; Abbas, A.; Rehman, M. F. ur; Kanwal, F.; Saeed, M.; Ullah, S.; Lu, C. Novel Carboxylic Acid-Capped Silver Nanoparticles as Antimicrobial and Colorimetric Sensing Agents. *Molecules* **2022**, *27* (11), 3363. https://doi.org/10.3390/molecules27113363.

(57) Gonçalves, I. C.; Martins, M. C. L.; Barbosa, M. A.; Ratner, B. D. Protein Adsorption on 18-Alkyl Chains Immobilized on Hydroxyl-Terminated Self-Assembled Monolayers. *Biomaterials* **2005**, *26* (18), 3891–3899. https://doi.org/10.1016/j.biomaterials.2004.10.006.

(58) Kong, B.; Kim, Y.; Choi, I. S. pH-Dependent Stability of Self-Assembled Monolayers on G0|d.

(59) Garg, N.; Carrasquillo-Molina, E.; Lee, T. R. Self-Assembled Monolayers Composed of Aromatic Thiols on Gold: Structural Characterization and Thermal Stability in Solution. *Langmuir* **2002**, *18* (7), 2717–2726. https://doi.org/10.1021/la0115278.

(60) Käfer, D.; Witte, G.; Cyganik, P.; Terfort, A.; Wöll, C. A Comprehensive Study of Self-Assembled Monolayers of Anthracenethiol on Gold: Solvent Effects, Structure, and Stability. *J. Am. Chem. Soc.* **2006**, *128* (5), 1723–1732. https://doi.org/10.1021/ja0571592.

(61) Chandekar, A.; Sengupta, S. K.; Whitten, J. E. Thermal Stability of Thiol and Silane Monolayers: A Comparative Study. *Applied Surface Science* **2010**, *256* (9), 2742–2749. https://doi.org/10.1016/j.apsusc.2009.11.020.

(62) Herzer, N.; Hoeppener, S.; S. Schubert, U. Fabrication of Patterned Silane Based Self-Assembled Monolayers by Photolithography and Surface Reactions on Silicon - Oxide Substrates. *Chemical Communications* **2010**, *46* (31), 5634–5652. https://doi.org/10.1039/C0CC00674B.

(63) Dodiuk, H.; Keing, S.; Liran, I. Room Temperature Curing Epoxy Adhesives for Elevated Temperature Service. Part III. The Effect of Silane Coupling Agents. *The Journal of Adhesion* **1992**, *39* (2–3), 123–136. https://doi.org/10.1080/00218469208026544.

(64) Tchoquessi Doidjo, M. R.; Belec, L.; Aragon, E.; Joliff, Y.; Lanarde, L.; Meyer, M.; Bonnaudet, M.; Perrin, F. X. Influence of Silane-Based Treatment on Adherence and Wet Durability of Fusion Bonded Epoxy/Steel Joints. *Progress in Organic Coatings* **2013**, *76* (12), 1765–1772. https://doi.org/10.1016/j.porgcoat.2013.05.014.

(65) Park, C. S.; Lee, H. J.; Jamison, A. C.; Lee, T. R. Robust Maleimide-Functionalized Gold Surfaces and Nanoparticles Generated Using Custom-Designed Bidentate Adsorbates. *Langmuir* **2016**, *32* (29), 7306–7315. https://doi.org/10.1021/acs.langmuir.6b01299.

(66) Chechik, V.; Schönherr, H.; Vancso, G. J.; Stirling, C. J. M. Self-Assembled Monolayers of Branched Thiols and Disulfides on Gold: Surface Coverage, Order and Chain Orientation. *Langmuir* **1998**, *14* (11), 3003–3010. https://doi.org/10.1021/la971090x.

(67) Artusio, F.; Fumagalli, F.; Bañuls-Ciscar, J.; Ceccone, G.; Pisano, R. General and Adaptive Synthesis Protocol for High-Quality Organosilane Self-Assembled Monolayers as Tunable Surface Chemistry Platforms for Biochemical Applications. *Biointerphases* **2020**, *15* (4), 041005. https://doi.org/10.1116/6.0000250.

(68) Mani, G.; Johnson, D. M.; Marton, D.; Dougherty, V. L.; Feldman, M. D.; Patel, D.; Ayon, A. A.; Agrawal, C. M. Stability of Self-Assembled Monolayers on Titanium and Gold. *Langmuir* **2008**, *24* (13), 6774–6784. https://doi.org/10.1021/la8003646.

(69) Tang, J.; Howard, C. B.; Mahler, S. M.; Thurecht, K. J.; Huang, L.; Xu, Z. P. Enhanced Delivery of siRNA to Triple Negative Breast Cancer Cells in Vitro and in Vivo through Functionalizing Lipid-Coated Calcium Phosphate Nanoparticles with Dual Target Ligands. *Nanoscale* **2018**, *10* (9), 4258–4266. https://doi.org/10.1039/C7NR08644J.

(70) Kehr, N. S.; Atay, S.; Ergün, B. Self-Assembled Monolayers and Nanocomposite Hydrogels of Functional Nanomaterials for Tissue Engineering Applications. *Macromolecular Bioscience* **2015**, *15* (4), 445–463. https://doi.org/10.1002/mabi.201400363.

(71) White, K. E.; Avery, E. M.; Cummings, E.; Hong, Z.; Langecker, J.; Vetushka, A.; Dušek, M.; Macháček, J.; Višnák, J.; Endres, J.; Bastl, Z.; Mete, E.; Alexandrova, A. N.; Baše, T.; Weiss, P. S. Competing Intermolecular and Molecule–Surface Interactions: Dipole–Dipole-Driven Patterns in Mixed Carborane Self-Assembled Monolayers. *Chem. Mater.* **2024**, *36* (4), 2085–2095. https://doi.org/10.1021/acs.chemmater.3c03210.

(72) Azzam, W. Effect of Immersion Time on the Structural Order of the Self-Assembled Monolayers of Terphenylthiols on the Gold Surface: Coexistence of Different Molecular Adlayer Phases. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **2024**, *683*, 132956. https://doi.org/10.1016/j.colsurfa.2023.132956.

(73) Canepa, P.; Rotondi, S. M. C.; Cavalleri, O. Atomic Force Microscopy as a Nanolithography Tool to Investigate the DNA/Gold Interface. *Current Opinion in Electrochemistry* **2024**, *44*, 101444. https://doi.org/10.1016/j.coelec.2024.101444.

(74) Jain, P. K.; Qian, W.; El-Sayed, M. A. Ultrafast Cooling of Photoexcited Electrons in Gold Nanoparticle−Thiolated DNA Conjugates Involves the Dissociation of the Gold−Thiol Bond. *J. Am. Chem. Soc.* **2006**, *128* (7), 2426–2433. https://doi.org/10.1021/ja056769z.

(75) Huang, Y.; Liu, W.; Cheng, H.; Yao, T.; Yang, L.; Bao, J.; Huang, T.; Sun, Z.; Jiang, Y.; Wei, S. Solvent-Induced Desorption of Alkanethiol Ligands from Au Nanoparticles. *Phys. Chem. Chem. Phys.* **2016**, *18* (23), 15927–15933. https://doi.org/10.1039/C6CP00480F.

(76) Borzenkov, M.; Chirico, G.; D'Alfonso, L.; Sironi, L.; Collini, M.; Cabrini, E.; Dacarro, G.; Milanese, C.; Pallavicini, P.; Taglietti, A.; Bernhard, C.; Denat, F. Thermal and Chemical Stability of Thiol Bonding on Gold Nanostars. *Langmuir* **2015**, *31* (29), 8081–8091. https://doi.org/10.1021/acs.langmuir.5b01473.

(77) Qu, D.; Kim, B.-C.; Lee, C.-W. J.; Uosaki, K. 1,n-Alkanedithiol (n = 2, 4, 6, 8, 10) Self-Assembled Monolayers on Au(111): Electrochemical and Theoretical Approach. *Bulletin of the Korean Chemical Society* **2009**, *30* (11), 2549–2554. https://doi.org/10.5012/bkcs.2009.30.11.2549.

(78) Yang, H.; Lazos, D.; Ulbricht, M. Thin, Highly Crosslinked Polymer Layer Synthesized via Photoinitiated Graft Copolymerization on a Self-Assembled-Monolayer-Coated Gold Surface. *Journal of Applied Polymer Science* **2005**, *97* (1), 158–164. https://doi.org/10.1002/app.21621.

(79) Yang, W.; Ella-Menye, J.-R.; Liu, S.; Bai, T.; Wang, D.; Yu, Q.; Li, Y.; Jiang, S. Cross-Linked Carboxybetaine SAMs Enable Nanoparticles with Remarkable Stability in Complex Media. *Langmuir* **2014**, *30* (9), 2522–2529. https://doi.org/10.1021/la404941m.

(80) Levard, C.; Hotze, E. M.; Colman, B. P.; Dale, A. L.; Truong, L.; Yang, X. Y.; Bone, A. J.; Brown, G. E.; Tanguay, R. L.; Di Giulio, R. T.; Bernhardt, E. S.; Meyer, J. N.; Wiesner, M. R.; Lowry, G. V. Sulfidation of Silver Nanoparticles: Natural Antidote to Their Toxicity. *Environ. Sci. Technol.* **2013**, *47* (23), 13440–13448. https://doi.org/10.1021/es403527n.

(81) Jiang, D.; Tiago, M. L.; Luo, W.; Dai, S. The "Staple" Motif: A Key to Stability of Thiolate-Protected Gold Nanoclusters. *J. Am. Chem. Soc.* **2008**, *130* (9), 2777–2779. https://doi.org/10.1021/ja710991n.

(82) Aagaard, N. D.; Zelaya, E.; Fonticelli, M. H.; Azcárate, J. C. Unraveling the Radiation Damage of N-Dodecanethiol Self-Assembled Monolayers on Au(100) and Au(111): A Prerequisite to Understand the Oriented Attachment of Au Nanoparticles. *J. Phys. Chem. C* **2021**, *125* (49), 27305–27313. https://doi.org/10.1021/acs.jpcc.1c08444.

(83) Pensa, E.; Azofra, L. M.; Salvarezza, R. C.; Carro, P. Effect of Ligands on the Stability of Gold Nanoclusters. *J. Phys. Chem. Lett.* **2022**, *13* (28), 6475–6480. https://doi.org/10.1021/acs.jpclett.2c01616.

(84) *Strong Resistance of Citrate Anions on Metal Nanoparticles to Desorption under Thiol Functionalization | ACS Nano*. https://pubs.acs.org/doi/10.1021/nn506379m (accessed 2024-05-23).

(85) Li, Z.; Narouz, M. R.; Munro, K.; Hao, B.; Crudden, C. M.; Horton, J. H.; Hao, H. Carboxymethylated Dextran-Modified N-Heterocyclic Carbene Self-Assembled Monolayers on Gold for Use in Surface Plasmon Resonance Biosensing. *ACS Appl. Mater. Interfaces* **2017**, *9* (45), 39223–39234. https://doi.org/10.1021/acsami.7b13114.

(86) Frasconi, M.; Tortolini, C.; Botrè, F.; Mazzei, F. Multifunctional Au Nanoparticle Dendrimer-Based Surface Plasmon Resonance Biosensor and Its Application for Improved

Insulin Detection. *Anal. Chem.* **2010**, *82* (17), 7335–7342. https://doi.org/10.1021/ac101319k.

(87) Tour, J. M.; Jones, L. I.; Pearson, D. L.; Lamba, J. J. S.; Burgin, T. P.; Whitesides, G. M.; Allara, D. L.; Parikh, A. N.; Atre, S. Self-Assembled Monolayers and Multilayers of Conjugated Thiols, .Alpha.,.Omega.-Dithiols, and Thioacetyl-Containing Adsorbates. Understanding Attachments between Potential Molecular Wires and Gold Surfaces. *J. Am. Chem. Soc.* **1995**, *117* (37), 9529–9534. https://doi.org/10.1021/ja00142a021.

(88) Yuge, R.; Miyazaki, A.; Enoki, T.; Tamada, K.; Nakamura, F.; Hara, M. Fabrication of TTF−TCNQ Charge-Transfer Complex Self-Assembled Monolayers: Comparison between the Coadsorption Method and the Layer-by-Layer Adsorption Method. *J. Phys. Chem. B* **2002**, *106* (27), 6894–6901. https://doi.org/10.1021/jp0135757.

(89) *Characterization of self-assembled monolayers (SAMs) on silicon substrate comparative with polymer substrate for Escherichia coli O157:H7 detection - ScienceDirect*. https://www.sciencedirect.com/science/article/pii/S0169433209009556?via%3Dihub (accessed 2024-05-23).

(90) Im, J.; Trindade, G. F.; Quach, T. T.; Sohaib, A.; Wang, F.; Austin, J.; Turyanska, L.; Roberts, C. J.; Wildman, R.; Hague, R.; Tuck, C. Functionalized Gold Nanoparticles with a Cohesion Enhancer for Robust Flexible Electrodes. *ACS Appl. Nano Mater.* **2022**, *5* (5), 6708–6716. https://doi.org/10.1021/acsanm.2c00742.

(91) Demissie, A. T.; Haugstad, G.; Frisbie, C. D. Quantitative Surface Coverage Measurements of Self-Assembled Monolayers by Nuclear Reaction Analysis of Carbon-12. *J. Phys. Chem. Lett.* **2016**, *7* (17), 3477–3481. https://doi.org/10.1021/acs.jpclett.6b01363.

(92) Hinterwirth, H.; Kappel, S.; Waitz, T.; Prohaska, T.; Lindner, W.; Lämmerhofer, M. Quantifying Thiol Ligand Density of Self-Assembled Monolayers on Gold Nanoparticles by Inductively Coupled Plasma–Mass Spectrometry. *ACS Nano* **2013**, *7* (2), 1129–1136. https://doi.org/10.1021/nn306024a.

(93) Smith, A. M.; Marbella, L. E.; Johnston, K. A.; Hartmann, M. J.; Crawford, S. E.; Kozycz, L. M.; Seferos, D. S.; Millstone, J. E. Quantitative Analysis of Thiolated Ligand Exchange on Gold Nanoparticles Monitored by 1H NMR Spectroscopy. *Anal. Chem.* **2015**, *87* (5), 2771–2778. https://doi.org/10.1021/ac504081k.

(94) Marchi, D.; Cara, E.; Lupi, F. F.; Hönicke, P.; Kayser, Y.; Beckhof, B.; Castellino, M.; Klapetek, P.; Zoccante, A.; Laus, M.; Cossi, M. Structure and Stability of 7-Mercapto-4-Methylcoumarin Self-Assembled Monolayers on Gold: An Experimental and Computational Analysis. *Phys. Chem. Chem. Phys.* **2022**, *24* (36), 22083–22090. https://doi.org/10.1039/D2CP03103E.

(95) Tero, T.-R.; Malola, S.; Koncz, B.; Pohjolainen, E.; Lautala, S.; Mustalahti, S.; Permi, P.; Groenhof, G.; Pettersson, M.; Häkkinen, H. Dynamic Stabilization of the Ligand–Metal Interface in Atomically Precise Gold Nanoclusters Au68 and Au144 Protected by Meta-Mercaptobenzoic Acid. *ACS Nano* **2017**, *11* (12), 11872–11879. https://doi.org/10.1021/acsnano.7b07787.

(96) Brewer, N. J.; Janusz, S.; Critchley, K.; Evans, S. D.; Leggett, G. J. Photooxidation of Self-Assembled Monolayers by Exposure to Light of Wavelength 254 Nm: A Static SIMS Study. *J. Phys. Chem. B* **2005**, *109* (22), 11247–11256. https://doi.org/10.1021/jp0443299.

(97) Breuer, R.; Schmittel, M. Redox-Stable SAMs in Water (pH 0–12) from 1,1′-Biferrocenylene-Terminated Thiols on Gold. *Organometallics* **2012**, *31*, 6642–6651. https://doi.org/10.1021/OM300718K.

(98) Shi, S.; Wang, Q.; Wang, T.; Ren, S.; Gao, Y.; Wang, N. Thermo-, pH-, and Light-Responsive Poly(N-Isopropylacrylamide-Co-Methacrylic Acid)--Au Hybrid Microgels Prepared by the in Situ Reduction Method Based on Au-Thiol Chemistry. *The journal of physical chemistry. B* **2014**, *118 25*, 7177–7186. https://doi.org/10.1021/jp5027477.

(99) Hu, Y.; Cheng, H.; Zhao, X.; Wu, J.; Muhammad, F.; Lin, S.; He, J.; Zhou, L.; Zhang, C.; Deng, Y.; Wang, P.; Zhou, Z.; Nie, S.; Wei, H. Surface-Enhanced Raman Scattering Active Gold Nanoparticles with Enzyme-Mimicking Activities for Measuring Glucose and Lactate in Living Tissues. *ACS Nano* **2017**, *11* (6), 5558–5566. https://doi.org/10.1021/acsnano.7b00905.

(100) Schollbach, M.; Zhang, F.; Roosen-Runge, F.; Skoda, M. W. A.; Jacobs, R. M. J.; Schreiber, F. Gold Nanoparticles Decorated with Oligo(Ethylene Glycol) Thiols: Surface Charges and Interactions with Proteins in Solution. *Journal of Colloid and Interface Science* **2014**, *426*, 31–38. https://doi.org/10.1016/j.jcis.2014.03.052.

(101) Neu, M.; Germershaus, O.; Mao, S.; Voigt, K.-H.; Behe, M.; Kissel, T. Crosslinked Nanocarriers Based upon Poly(Ethylene Imine) for Systemic Plasmid Delivery: In Vitro

Characterization and in Vivo Studies in Mice. *Journal of Controlled Release* **2007**, *118* (3), 370–380. https://doi.org/10.1016/j.jconrel.2007.01.007.

(102) Cometto, F. P.; Ruano, G.; Soria, F. A.; Calderón, C. A.; Paredes-Olivera, P. A.; Zampieri, G.; Patrito, E. M. Thermal and Chemical Stability of *n*-Hexadecanethiol Monolayers on Au(111) in O2 Environments. *Electrochimica Acta* **2016**, *215*, 313–325. https://doi.org/10.1016/j.electacta.2016.08.119.

(103) *Scalable and Precise Synthesis of Thiolated Au10–12, Au15, Au18, and Au25 Nanoclusters via pH Controlled CO Reduction | Chemistry of Materials*. https://pubs.acs.org/doi/10.1021/cm304098x (accessed 2024-05-23).

(104) Janssens, P.; Debrouwer, W.; Van Aken, K.; Huvaere, K. Thiol−Ene Coupling in a Continuous Photo-Flow Regime. *ChemPhotoChem* **2018**, *2* (10), 884–889. https://doi.org/10.1002/cptc.201800155.

(105) Dickers, M. D.; Verkhovtsev, A. V.; Mason, N. J.; Solov'yov, A. V. Atomistic Modelling and Structural Characterisation of Coated Gold Nanoparticles for Biomedical Applications. *Eur. Phys. J. D* **2023**, *77* (7), 141. https://doi.org/10.1140/epjd/s10053-023-00726-1.

(106) Shavalier, S. A.; Gezelter, J. D. Heat Transfer in Gold Interfaces Capped with Thiolated Polyethylene Glycol: A Molecular Dynamics Study. *J. Phys. Chem. B* **2023**, *127* (47), 10215–10225. https://doi.org/10.1021/acs.jpcb.3c05238.

(107) Schaefer, K.; Liu, C.-Y.; Meyer, A.; Schlicke, H.; Vossmeyer, T.; Herrmann, C. Cross-Linked Gold Nanoparticle Assemblies: What Can We Learn from Single Flat Interfaces? *J. Phys. Chem. C* **2024**, *128* (9), 3994–4008. https://doi.org/10.1021/acs.jpcc.3c07646.

(108) Ladygin, V. V.; Korotaev, P. Yu.; Yanilkin, A. V.; Shapeev, A. V. Lattice Dynamics Simulation Using Machine Learning Interatomic Potentials. *Computational Materials Science* **2020**, *172*, 109333. https://doi.org/10.1016/j.commatsci.2019.109333.

(109) Malola, S.; Nieminen, P.; Pihlajamäki, A.; Hämäläinen, J.; Kärkkäinen, T.; Häkkinen, H. A Method for Structure Prediction of Metal-Ligand Interfaces of Hybrid Nanoparticles. *Nat Commun* **2019**, *10* (1), 3973. https://doi.org/10.1038/s41467-019-12031-w.

(110) Fronzi, M.; Amos, R. D.; Kobayashi, R.; Matsumura, N.; Watanabe, K.; Morizawa, R. K. Evaluation of Machine Learning Interatomic Potentials for the Properties of Gold Nanoparticles. *Nanomaterials* **2022**, *12* (21), 3891. https://doi.org/10.3390/nano12213891.

(111) Wang, Y.; Wang, Y.; Wang, F.; Chi, H.; Zhao, G.; Zhang, Y.; Li, T.; Wei, Q. Electrochemical Aptasensor Based on Gold Modified Thiol Graphene as Sensing Platform and Gold-Palladium Modified Zirconium Metal-Organic Frameworks Nanozyme as Signal Enhancer for Ultrasensitive Detection of Mercury Ions. *Journal of Colloid and Interface Science* **2022**, *606*, 510–517. https://doi.org/10.1016/j.jcis.2021.08.055.

(112) *Electrically Conducting Nanopatterns Formed by Chemical e-Beam Lithography via Gold Nanoparticle Seeds | Langmuir*. https://pubs.acs.org/doi/10.1021/la204393h (accessed 2024-05-23).

(113) Newton, L.; Slater, T.; Clark, N.; Vijayaraghavan, A. Self Assembled Monolayers (SAMs) on Metallic Surfaces (Gold and Graphene) for Electronic Applications. *J. Mater. Chem. C* **2013**, *1* (3), 376–393. https://doi.org/10.1039/C2TC00146B.

(114) Ghiorso, M. S.; Spera, F. J. Large Scale Simulations. *Reviews in Mineralogy and Geochemistry* **2010**, *71* (1), 437–463. https://doi.org/10.2138/rmg.2010.71.20.

(115) Fihey, A.; Hettich, C.; Touzeau, J.; Maurel, F.; Perrier, A.; Köhler, C.; Aradi, B.; Frauenheim, T. SCC-DFTB Parameters for Simulating Hybrid Gold-thiolates Compounds. *J Comput Chem* **2015**, *36* (27), 2075–2087. https://doi.org/10.1002/jcc.24046.

(116) O'Carroll, D.; English, N. J. Multi-Theory Comparisons of Molecular Simulation Approaches to TiO2/H2O Interfacial Systems. *Crystals* **2023**, *13* (7), 1122. https://doi.org/10.3390/cryst13071122.

(117) González, M. A. Force Fields and Molecular Dynamics Simulations. *JDN* **2011**, *12*, 169–200. https://doi.org/10.1051/sfn/201112009.

(118) Farah, K.; Müller-Plathe, F.; Böhm, M. C. Classical Reactive Molecular Dynamics Implementations: State of the Art. *ChemPhysChem* **2012**, *13* (5), 1127–1151. https://doi.org/10.1002/cphc.201100681.

(119) Chenoweth, K.; van Duin, A. C.; Goddard, W. A. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A* **2008**, *112*, 1040.

(120) Chan, H.; Narayanan, B.; Cherukara, M. J.; Sen, F. G.; Sasikumar, K.; Gray, S. K.; Chan, M. K. Y.; Sankaranarayanan, S. K. R. S. Machine Learning Classical Interatomic Potentials for Molecular Dynamics from First-Principles Training Data. *J. Phys. Chem. C* **2019**, *123* (12), 6941–6957. https://doi.org/10.1021/acs.jpcc.8b09917.

(121) Torres, D.; Carro, P.; Salvarezza, R. C.; Illas, F. Evidence for the Formation of Different Energetically Similar Atomic Structures in Ag (111) - (7 × 7)-R 19.1° - CH 3 S. *Phys. Rev. Lett.* **2006**, *97* (22), 226103.

(122) Guo, Q.; Li, F. Self-Assembled Alkanethiol Monolayers on Gold Surfaces: Resolving the Complex Structure at the Interface by STM. *Phys. Chem. Chem. Phys.* **2014**, *16* (36), 19074. https://doi.org/10.1039/C4CP00596A.

(123) Becerril, D.; Noguez, C. Adsorption of a Methylthio Radical on Silver Nanoparticles: Size Dependence. *J. Phys. Chem. C* **2015**, *119* (20), 10824.

B) Study of silver and gold systems using molecular dynamics

B.1) Computational Methodology

As previously discussed, studying theoretically large systems such as gold or silver nanoparticles using quantum mechanics (QM) is impractical due to the thousands of atoms involved in the simulations. Although more accurate, employing QM for such large-scale simulations could take decades to yield initial results. Molecular dynamics (MD), in contrast, offers a better approach not only by decreasing the computational cost but also by its ability to handle diverse conditions and materials. In this section, MD is first introduced with some fundamentals. Following this, we provide a detailed description of a reactive force field known as ReaxFF, which is here used in simulations of thiolate-SAMs on silver nanoparticles using the LAMMPS software. It allows us to study reactive processes in nanoparticles, like the breaking and the formation of bonds, which is challenging for non-reactive classical force fields.

B.1.a) Classical and reactive Molecular Dynamics

Basics of Molecular Dynamics Molecular Dynamics (MD) involve the computational modeling of atomic and molecular movements over time governed by Newton's law of motion. It allows one to observe the dynamical evolution of systems at atomic resolution, from femtoseconds to microseconds. The positions and velocities of atoms are updated iteratively using the force equation, $\vec{F} = \mathrm{m}\vec{a}$, where $\vec{F}$ is the force applied, m is the mass, and a the acceleration. Each timestep $\Delta\mathrm{t}$ in an MD simulation involves recalculating the forces acting on each atom based on their positions and using these forces to update velocities and positions. The choice of $\Delta\mathrm{t}$ is crucial, as too large timesteps can lead to numerical instabilities, while too small a time step can make the simulation unnecessarily computationally intensive. It should be adapted to the studied system depending on the elements that are considered in the system. The Nyquist theorem suggests that to accurately capture the dynamics of the fastest-moving atoms, the time step should be less than half the period of the fastest vibrations observed within the system. For more heavy nuclei such as gold or silver, 2 fs should be enough, but for lighter ones like hydrogens, carbons, or even sulfur, one should consider lowering the value to 0.25 fs to capture all the dynamics of the system. Then, for each timestep, the forces of each atom are calculated. From that, their positions are stored, allowing them to follow the system and analyze it by calculating different properties at the molecular level. For example, structural properties can be calculated from the radial distribution

function, giving how density varies as a function of distances or analyzing the different adsorption sites. We developed a program allowing us to calculate different properties at any point of the dynamics for gold or silver NP capped with thiolates-SAM that is presented in section 3b.

MD also includes temperature and pressure effects, depending on the ensemble that is set up beforehand. Thus, thermodynamics analysis can be performed relative to the stability along those intensive properties. From that, it is possible to simulate, for example, the thermalization of thousands of atoms for a certain simulation time and study the stability of the thiolates-SAM during the dynamics on a nanoparticle like in Figure B-1.

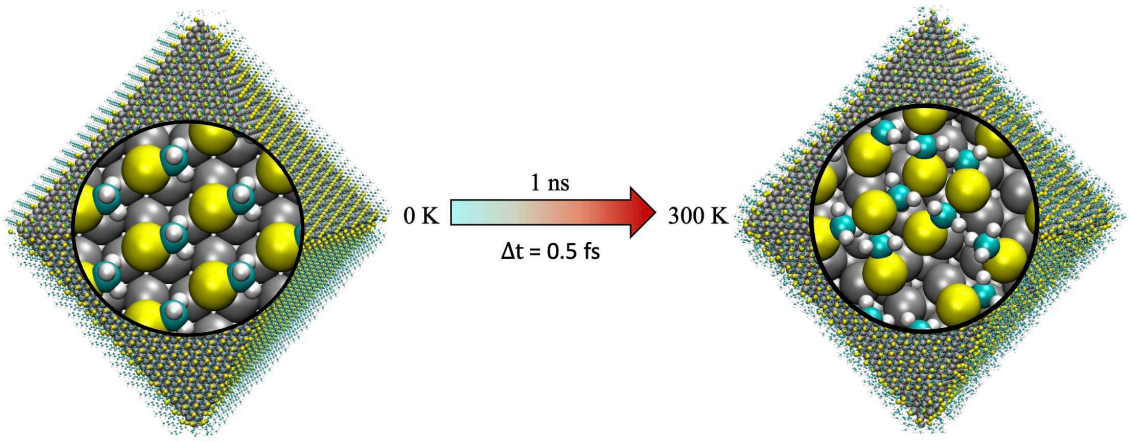


Figure B-1: Schematic of molecular dynamics of a silver nanoparticle capped with thiolates-SAMs for 1 ns by the increases of the temperature up to 300 K and a timestep of 0.5 fs

MD simulations are based on the Hamiltonian equation of motion, where the total energy is the sum of the kinetic and potential energy function of momenta (p) and the position (r) as defined in equation B-1.

$$H(r,p) = K(p) + U(r) \qquad \text{(B-1)}$$

First, the interaction model, also called force field, is chosen to compute the forces acting on each atom. Then, the atoms move during a specific time step, which updates the position and velocity using Newton's law. The velocity Verlet algorithm is used to integrate Newton's law.[1] It calculated velocities at mid-step and then calculated the positions at the next step. Followed by the calculation of the accelerations at the next step from the potential and update the velocities.

In MD simulations, different ensembles can be used to model systems under various thermodynamic constraints.

Ensembles and their applications Different ensembles are employed during simulations to replicate specific thermodynamic conditions accurately. The microcanonical ensemble (NVE), which stands for a constant number of particles (N), constant volume (V), and total energy (E), keeps the particles in an isolated system to stay in statistical equilibrium. It does not exchange energy or particles with its environment; it is the natural environment for MD. Then, the canonical ensemble, named NVT, kept a constant temperature (T), as presented in Figure B-2.

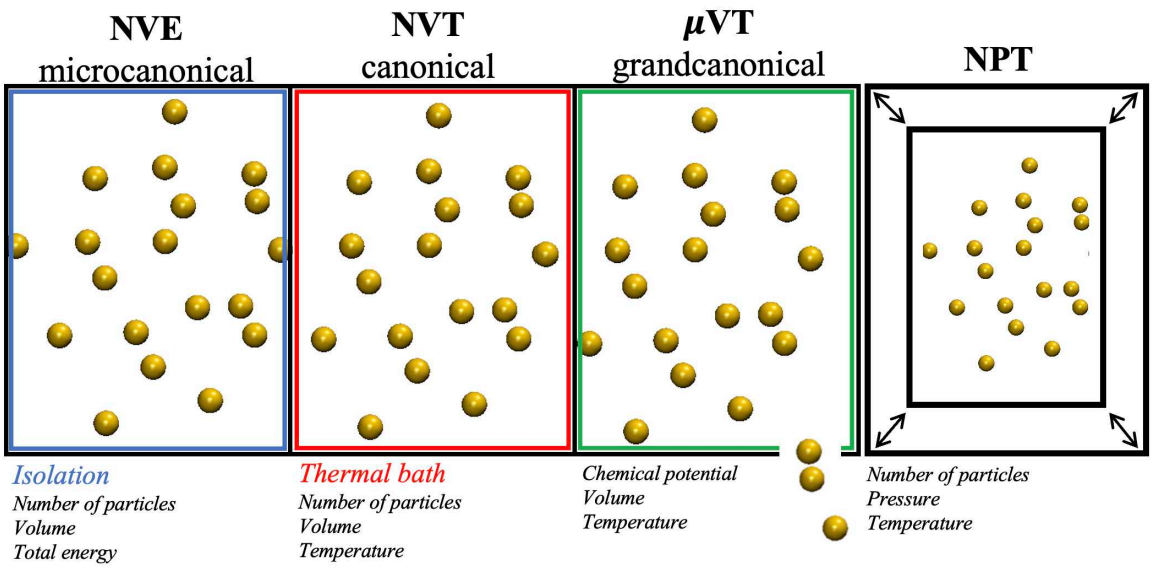


Figure B-2: Description of different ensembles that are used in MD simulations

This ensemble requires the use of a thermostat, which gives the desired temperature during the simulations. The kinetic energy for N atoms is defined by equation B-2 and depends on the Boltzmann constant and the temperature and can be rewritten in terms of velocities. From that, the thermostat makes sure that the velocities are giving a constant kinetic energy.

$$K = \frac{1}{2}\sum_{i} m_i v_i^2 = \frac{3}{2} N k_B T \qquad \text{(B-2)}$$

One possible way is to rescale the velocities at each timestep by a scaling factor called λ which reupdate its value at each timestep. It scales the velocities by multiplying the old velocities by a value superior to one when the velocities are too small while, for the opposite situation, by a

factor inferior to one. Such scaling factor has a dependence on the desired temperature (T_{bath}). Thus, at each timestep, when the temperature (T_{actual}) is larger or lower than the desired one, the scaling factor is adapting its value. To well reproduce the rescaling effect on velocities and thus to keep the right kinetic energy, different algorithms are used that are called thermostats. Choosing the appropriate thermostat is essential for ensuring that the simulation not only reaches but also maintains the target temperature while correctly sampling the desired statistical ensemble. The Berendsen thermostat uses an exponential decay of the temperature.[2]

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau}\left(\frac{T_{bath}}{T_{actual}} - 1\right)} \tag{B-3}$$

It is used during a simulation to initialize the equilibration of the system, allowing to reach the wanted temperature easily by relaxing the system by uniformly rescaling the velocities by the equation B-3. Where τ is the strength of the coupling between the bath and the system. Nevertheless, the Berendsen thermostat does not guarantee the correct sampling of the canonical ensemble.[3] A correction has been given by Bussi et al. where a stochastic rescaling is made, reproducing well the correct energy fluctuation for the kinetic energy matching Maxwell-Boltzmann distribution.[4]

The Nosé-Hoover (NH) thermostat is an extended system which is directly incorporating a fictious variable Q with a given mass. The magnitude of Q leads to a control of the temperature by an exchange between the heat bath and the system. The thermostat adjusts the velocity scaling by incorporating friction and random noise, generating a canonical ensemble. The main drawback appears when one is studying a small system using the NH thermostat since it does not reproduce an ergodic dynamic. Thus, the correct distributions are not generated.[5] The NH chains method has been proposed by Martyna et al. and corrects this problem by adding thermostat variables linked into a chain to the original system dynamics.[6]

The Langevin thermostat ensures that the energy distribution follows the canonical distribution.[7] It provides a more physically realistic model of the heat bath interaction, including random collisions assuring equipartition ergodicity and frictional forces. It is a robust tool that combines deterministic dynamics with stochastic elements to simulate the right temperature. It is based on a damping term that removes energy from the system and allows it to keep a realistic

state without gaining continuous kinetic energy. The random force is used to ensure the system does not cool down and freeze. Its stochastic nature allows the system to mimic the random thermal motions found in a true thermodynamic ensemble.[8] This method is particularly beneficial for simulations of biological systems where solvent interactions are significant.[9] Despite its ability to simulate the canonical ensemble, the Langevin thermostat can lose momentum transfer information due to random forces.[10]

Correctly setting the thermostat and the ensemble is crucial for simulating physical behaviors that closely mirror real-world phenomena, thus enhancing the reliability and applicability of simulation results.

Potential energy and the use of Force fields To keep the right pace of the temperature during the MD, we just saw how to handle the kinetic energy, but the total energy is also defined in terms of potential energy U(r). Understanding the nuances of potential energy within molecular dynamics simulations is necessary for accurately predicting the physical and chemical behaviors of molecular systems. It is handled through the use of force fields, which embody two main categories of energetic contributions: the intra and inter-molecular ones. Indeed, force fields serve as mathematical tools that describe the energy landscape of a molecular system, influencing everything from molecular stability to reaction dynamics. They are based either on empirical data or first-principle calculations (HF, DFT…) and allow to correctly describe the system. Empirical parametrizations rely on experimental data, while first-principle methods like Hartree-Fock (HF) and Density Functional Theory (DFT) are based on fundamental quantum calculations (QM) principles. The choice between using an empirical or a first-principle force field often depends on the specific requirements of the simulation, including the desired accuracy and computational resources available. Often, a hybrid approach containing both strategies is used. They enable accurate modeling of interatomic interactions, which is crucial for predicting physical and chemical properties at a lower cost than QM calculations. They allow for the description of the time evolution of bonded and non-bonded interaction between atoms. It is a collection of equations and constants that reproduce molecular geometries and energies. A typical potential energy function used for force fields is shown here:

$$U(\vec{r_1},\vec{r_2},\ldots,\vec{r_n}) = \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}}(\vec{r_i}-\vec{r_0})^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}}(\theta_i-\theta_0)^2}_{U_{\text{angle}}} \quad \text{(B-4)}$$

$$+ \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}}[1+\cos(n_i\phi_i+\delta_i)]}_{U_{\text{dihedral}}} + \underbrace{\sum_i\sum_{j\neq i} 4\,\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i\sum_{j\neq i}\frac{q_i q_j}{\epsilon r_{ij}}}_{U_{\text{nonbond}}}$$

The bond Stretching U_{bond}, as presented on Equation B-4, is modeled as a harmonic potential where k_i^{bond} is the force constant for the bond, $\vec{r_i}$ is the instantaneous bond length, and $\vec{r_0}$ the equilibrium bond length. The energy increases as the bond length deviates from its equilibrium value. The Angle Bending U_{angle} involves the bending of angles between bonds, k_i^{angle} is the force constant for bending, θ_i is the actual angle, and θ_0 is the equilibrium angle. This component penalizes deviations from the equilibrium angle. For torsional interactions $U_{dihedral}$, the energy is associated with the rotation about bond axes, where k_i^{dihe} represents the torsion force constant, n_i is the multiplicity of the dihedral angle, ϕ_i is the dihedral angle, and is the phase shift. For the intermolecular interactions, the first term is the van der Waals Forces, modeled using the Lennard-Jones potential, which includes attractive and repulsive components. The parameters ϵ_{ij} and σ_{ij} represent the depth of the potential well and the finite distance at which the inter-particle potential is zero, respectively. The other term is the Coulombic Forces. The electrostatic interaction between charged particles, where q_j are the charges, r_{ij} is the distance between particles, and is the dielectric constant of the medium, adjusting the strength of the interaction. Other terms are also taken into account, like the improper term helping to get the correct geometry. It is important to set these parameters accurately to ensure the "real" behaviors and predict new phenomena under experimental conditions.

Different force fields have been parametrized and optimized, and their performance is dependent on the approximation made (from QM) or experimental data that are used. AMBER (Assisted Model Building with Energy Refinement) is predominantly used for biomolecular simulations, such as proteins and nucleic acids.[11] It has a good balance between accuracy and computational efficiency and is known for its well-tested parameter sets for biomolecules and its ability to simulate biologically relevant macromolecules effectively. Therefore, AMBER, with its robust parameter sets and efficient algorithms, remains a popular choice for researchers focusing

on biological molecules, where fast and reliable simulations are crucial. Nevertheless, its performance can be less optimal for systems involving novel drug-like compounds as it is not necessarily well-represented in its parameters set.[12] Other force fields exist with their different applications. CHARMM (Chemistry at HARvard Macromolecular Mechanics), which can be used for a simulation involving complex biological systems and interactions, including membrane-protein interactions.[13] OPLS-AA (Optimized Potentials for Liquid Simulations-All Atoms) is suitable for ionic and organic liquids.[14,15] AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) is a more recent force field that uses multipole electrostatics rather than just point charges and includes polarizability. It offers enhanced accuracy in modeling electrostatic interactions due to its use of distributed multipole analysis, making it suitable for complex environments where polarization effects are significant. It increases the accuracy of simulations involving complex molecular interactions and specific solvent effects. In AMOEBA, polarizability is implemented through a dynamic response model, where the induced dipoles are calculated based on the local electric field. This approach allows AMOEBA to adaptively model the changing electron distribution around atoms as they interact with their environment, enhancing the accuracy of simulating molecular behavior under various conditions.[16] The inclusion of polarizability in AMOEBA enables the simulation to capture changes in electron distribution that occur during chemical reactions. This feature is fundamental for accurately predicting reaction pathways and interaction energies, particularly in systems where polarization effects drive the chemical dynamics, such as the binding of ligands to metal centers like in our case with thiolate-SAMs. However, this force field is not reactive and cannot reproduce the formation and the breaking of the bonds that can sometimes happen for complex reactivity. Nevertheless, it is widely used and gives accurate predictions for biological systems. Nowadays, AMOEBA+, integrates advancements such as machine learning potentials, which enhance its predictive power and efficiency.[17] By learning from a vast dataset of molecular interactions, AMOEBA+ is designed to optimize the accuracy of its predictions while managing computational demands more effectively.

As computational capabilities continue to advance, the development of more sophisticated and computationally efficient force fields will likely revolutionize our ability to simulate and understand complex molecular phenomena. Ongoing research into hybridization promises to bridge the gap between accuracy and efficiency, giving the next generation of molecular dynamics simulations.

Periodic Boundary Conditions Having discussed molecular dynamics (MD), ensembles, thermostats, and force fields, we now turn our attention to an essential concept in MD simulations—periodic boundary conditions (PBC). The use of MD allows us to simulate thousands of atoms, but for the study of a surface to understand, for example, the stability of our thiolates-SAM on gold or silver, we cannot have an infinite space it would require too many resources.[18]In practical terms, PBC are essential for simulating bulk properties and extensive systems without the computational overhead of simulating an infinite number of atoms. By replicating a small section or 'box' of atoms across an infinite virtual lattice, PBC allows us to model extensive systems with finite computational resources. To counter this, PBC are introduced to replicate a specific box in all directions, as presented in Figure B-3. When a particle crosses one boundary of the simulation box, it re-enters from the opposite boundary.

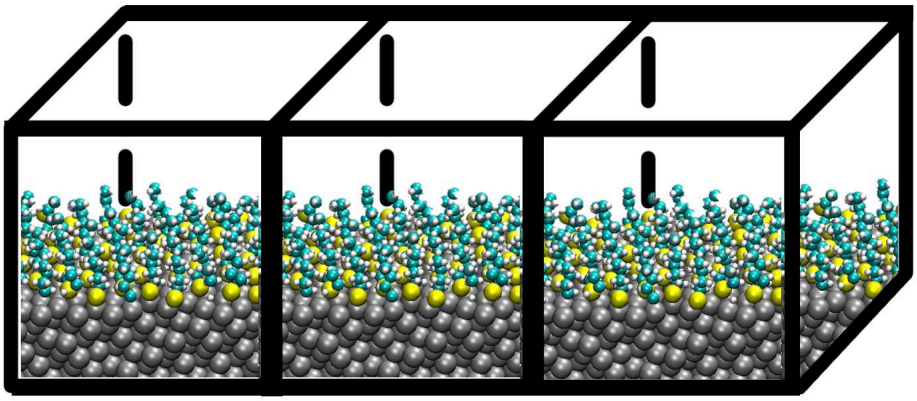


Figure B-3: Scheme of the periodic boundary conditions with at its center a silver surface capped with thiolates-SAMs

This not only maintains the density of the system but also ensures continuous and realistic interaction dynamics among particles. Particles interact only with the closest periodic image of the other particles. The use of PBC can sometimes introduce artifacts, particularly in systems with significant surface effects or where the size of the box is comparable to the range of interaction forces. Care must be taken to choose a box size large enough to minimize such effects, ensuring that the system properties are not unduly influenced by the artificial periodicity. Moreover, to ensure all the physics into infinite periodic images of the long-range interactions between particles, the Ewald summation is also used.[19]

For studies focusing on surface phenomena, such as the stability of thiolate-SAMs on gold or silver, PBC allows us to explore surface interactions under realistic bulk conditions. This approach is critical for accurately assessing how SAMs organize and bind to surfaces. While PBC are a powerful tool for simulating endless space, they must be applied with precociousness. The choice of boundary conditions, box size, and the method of interactions between periodic images must be carefully considered to avoid physical problems, such as phase behavior and diffusion characteristics.

Molecular Dynamics set-up To initiate a molecular dynamics (MD) simulation, the first step involves selecting an appropriate interaction model, such as classical force fields like AMBER, or polarizable force fields such as AMOEBA or ReaxFF. Then, the coordinates of each atom are defined within a simulation box, configured to enforce periodic boundary conditions, effectively modeling infinite spatial continuity. Using software like TINKER for AMOEBA-based simulations, the MD process begins by initializing atomic positions and velocities, often assigning velocities based on the Maxwell-Boltzmann distribution to match the desired initial temperature.

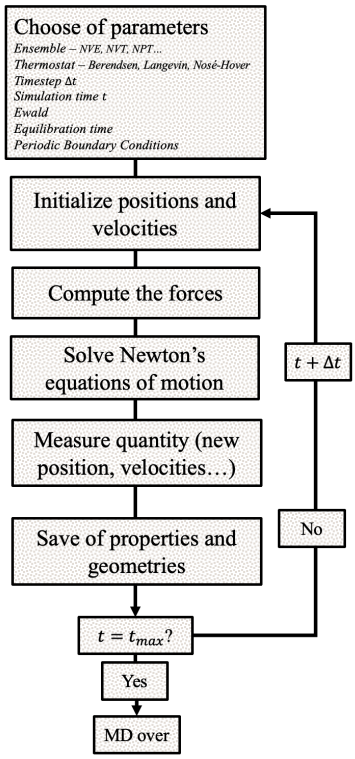


Figure B-4: Schematic of the different steps involved in a Molecular Dynamics simulation from the initial step to the end of the simulation.

The simulation parameters, including the choice of ensemble (e.g., NVT for constant temperature or NVE for constant energy), timestep, and total simulation time, are carefully selected to ensure accurate dynamics. During the simulation, forces on each atom are calculated at every timestep, using the chosen force fields to solve Newton's equations of motion. This iterative process updates the positions and velocities of atoms, driving the system towards equilibration. Velocity rescaling is applied at each timestep to maintain the system at the target temperature, particularly in thermostatted ensembles like NVT, using algorithms such as the Berendsen thermostat. Several steps are required to extract the geometry during the simulation. From that, it can be calculated different properties to understand how the system is reacting to different conditions. When reaching the simulation time (tmax), the simulation ends. The pathway is presented in Figure B-4

MD is then giving a good prediction for a large system, and its parametrization should be carefully chosen. As said before, to see reactivity, one needs to choose a reactive/polarization force fields that captures more precisely physical interactions. Therefore, this is mandatory if we want to capture the formation and breaking of bonds during the capping of thiolates-SAMs on gold and silver nanoparticles.

B.1.b) ReaxFF

The Reactive Force Field, called ReaxFF, is a computational tool enabling the perform accurate simulations of chemical reactions involving complex molecules and materials. Developed initially by Adri van Duin et al., it was designed to bridge the gap between quantum mechanics (QM) and classical force fields, offering a balance between computational efficiency and the ability to capture the nuances of chemical reactivity.[20] ReaxFF distinguishes itself from traditional molecular dynamics simulations by its unique ability to handle bond formation and dissociation dynamically. This capability makes it exceptionally well-suited for studying a wide range of chemical processes, such as combustion[21], catalysis[22], and material degradation[23], where bond-breaking and bond-making are essential features. Its development was motivated by the need for a more robust and versatile model that could keep pace with the increasing complexity of the systems being studied, especially those that could not be adequately described by conventional force fields. Classical force fields, with their fixed bonding scenarios, are typically limited to simulating molecular dynamics without involving chemical reactions. In contrast, quantum

mechanical methods, while highly accurate in predicting chemical phenomena, are often computationally intensive and not feasible for large systems or long simulation times.

ReaxFF combines the efficiency of classical force fields with a reactivity component; it allows the simulation of large-scale systems over practical timescales while still capturing critical chemical reaction events. This capability opens up new ways for exploring complex material behaviors and chemical processes in a more detailed and physically realistic manner.

In practice, ReaxFF has enabled researchers and scientists to conduct simulations that were previously beyond reach. For instance, in Materials Sciences, ReaxFF has been applied to investigate the stability, mechanical properties, and reactive properties of polymers[24], metals[25] , and semiconductors[26] under various conditions.

The broad applicability and adaptability of ReaxFF stem from its parameterization, which can be tailored to fit specific chemical environments based on experimental data or high-level quantum calculations. New parametrizations using the ReaxFF processed have been made for different materials.[27,28] This flexibility ensures that ReaxFF remains a powerful tool where realistic modeling of chemical phenomena is crucial for the development of new technologies and materials. Nevertheless, one should be prudent for transferability.

Theoretical Foundations of ReaxFF To delve deeper into the theoretical foundations of the Reactive Force Field (ReaxFF), it is imperative to understand the formulation of its force fields and their differentiation from non-reactive force fields. The core of ReaxFF's capability lies in its force field formulation, which is distinctively built to handle the dynamic nature of chemical reactions. The force field used for ReaxFF is a complex function of the atomic coordinates and extends beyond mere pairwise interactions to include many-body terms. This approach is crucial in accurately describing bond-breaking/making events during a simulation. It is divided into two main blocks: the first, which is based on bond order, and the other on non-bonded interactions. Only the general equations are presented here to get an overview of the philosophy behind ReaxFF without going too deep into mathematical expressions. At the heart of the ReaxFF simulation is the bond order calculation. Unlike classical force fields that assume a fixed bond order, ReaxFF dynamically calculates bond orders based on the distances between atoms and their hybridization states. The bond order in ReaxFF influences the strength of interactions, such as bond stretching,

bending, and torsion, and thus allows for the description of bond formation and breaking without the need to explicitly specify these events a priori.

1. **Energy Contributions** The energy in a ReaxFF simulation is composed of several contributions, as presented in Equation B-5, reflecting the complex nature of reactive systems. These include terms for bond energy, lone pair energy, angles, torsions, over/under coordination penalties, van der Waals interactions, and Coulomb interactions. The balance and interplay among these energy terms allow ReaxFF to adapt dynamically to the changing molecular structure in a reaction pathway.

$$E_{tot} = E_{Bond} + E_{over} + E_{under} + E_{angle} + E_{torsion} + E_{VDWaals} + E_{Coulomb} + E_{specific} \quad \text{(B-5)}$$

The bond order energy is calculated from a bond order term BO'_{ij} which is the foundation of this method. BO'_{ij} is determined through three exponentials as presented in Equation B-6, each defined between a distance r_{ij} range depending on the type of bonds and dynamically adjusts during simulations.

$$BO'_{ij} = \exp\left[p_{bo,1}\cdot\left[\frac{r_{ij}}{r_0^{\sigma}}\right]^{p_{bo,2}}\right] + \exp\left[p_{bo,3}\cdot\left[\frac{r_{ij}}{r_0^{\pi}}\right]^{p_{bo,4}}\right] + \exp\left[p_{bo,5}\cdot\left[\frac{r_{ij}}{r_0^{\pi\pi}}\right]^{p_{bo,6}}\right] \quad \text{(B-6)}$$

The σ bond with $p_{bo,1}$ and $p_{bo,2}$ from a distance between 1.5 Å and 2.5 Å. The first π bond $p_{bo,3}$ and $p_{bo,4}$ from 1.2 Å to 1.75 Å and the second π bond $p_{bo,5}$ and $p_{bo,6}$ between 1 Å and 1.4 Å. All those constant terms are updated during the simulation.

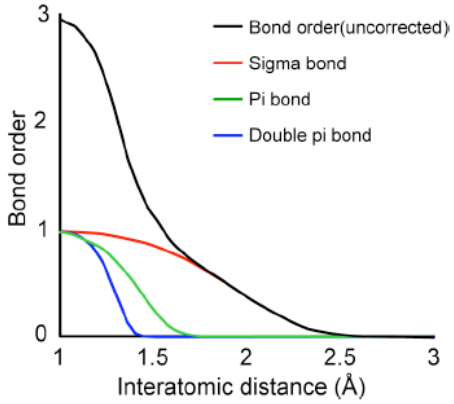


Figure B-5: Evolution of the bond order BO'_{ij} for carbon-carbon interactions concerning the distance r_{ij} adapted from Adri Van Duin et al. paper (2011)[29]

As we can see in Figure B-5, the bond order evolves depending on the distance of the C-C bond with a maximum bond order of 3 at short interaction. It involves an exponential decay term to reflect the weakening of bonding interactions with increasing distance. It allows us to accurately model the longer range and partially bonded configurations of transition states.

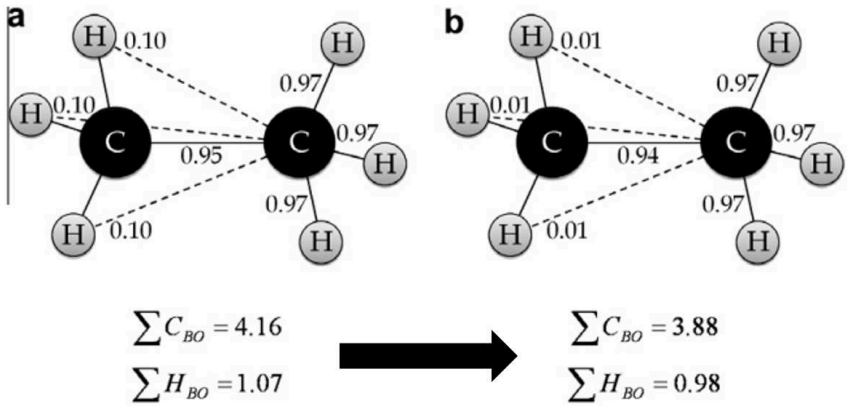


Figure B-6: Effect of coordinates correction on bond order of ethane molecule adapted from Adri van Duin et al. paper (2011) [29]

Nevertheless, BO'_{ij} needs some corrections as it over or underestimates its value where all atoms exceed their valences. For example, for ethane in Figure B-6a, without the correction, the bond order for both carbons and hydrogens is overestimated. This means that the actual bond order calculation requires some corrections terms, as the values for those atoms should not exceed 4 and 1. A correction factor based on the valence theory is then applied to the bond order to not exceed the bond order degree for each atom, which capped the specific values for each atom, as shown in Figure B-6b. This new formulation is then defined in Equation B-7.

$$BO_{ij} = BO'_{ij}.f_1(\Delta'_i,\Delta'_j).f_4(\Delta'_i,BO'_{ij}).f_5(\Delta'_i,BO'_{ij}) \quad \text{(B-7)}$$

The functional form of each function is described in the paper of Adri van Duin et al.[29] The important term here, Δ'_i, corresponds to the degree of deviation of the sum of the uncorrected bond orders around an atomic center.

However, small over/under coordination can remain to ensure the correction of two new constants appeared: the p_{over} and p_{under}. The over-coordinate term creates a penalty on the system, ensuring that the E_{over} is going to zero for the undercoordinated system, thus $\Delta_i < 0$ as

described in Equation B-8. For an undercoordinated atom, the resonance of the π-electron needs to be taken into account. E_{under} defined in Equation B-9 is then used only if the bonds between atoms i and j partly have π-bond features.

$$E_{over} = p_{over}.\Delta_i.\left(\frac{1}{1+\exp(\lambda_6.\Delta_i)}\right) \quad \text{(B-8)}$$

$$E_{under} = -p_{under}.\frac{1-\exp(\lambda_7.\Delta_i)}{1+\exp(-\lambda_8.\Delta_i)}.f_6(BO_{ij,\pi}.\Delta_j) \quad \text{(B-9)}$$

Other terms require a reformulation to obtain a dependence on each energy term to the BO. Thus, the valence angle term goes to zero as the BO goes to zero, taking into account the effect of over/under coordination where the penalty is applied to the central atom. It depends on an equilibrium term Θ_o and the valence angle Θ_{ijk}. The equilibrium term is linked to the sum of π-bond orders around the central atom j. Thus, it changes the equilibrium angle based on the geometry of the central atom j and its neighbors. Other terms are included, like the torsion angles and the conjugated energy, which are important and described also by the BO.

2. **The non-bonded interactions** van der Waals (VDW) and Coulomb Interactions are also taken into account to capture repulsive and attractive forces for each system, which is crucial to predict accurately the interactions, especially in complex molecular systems. For the VDW forces, a distance-corrected Morse potential is used, including a shielding interaction, defined in Equation B-10.

$$E_{VdWaals} = D_{ij} \cdot \left\{\exp\left[\alpha_{ij} \cdot \left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}}\right)\right] - 2 \cdot \exp\left[\frac{1}{2} \cdot \alpha_{ij} \cdot \left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}}\right)\right]\right\} \quad \text{(B-10)}$$

In ReaxFF, the shielding adjusts the strength of the interaction potential based on the local environment of particles. It allows the avoid excessively high repulsion between bonded atoms and atoms sharing a valence angle.

For the Coulomb interaction, atomic charges are calculated with the Electron Equilibration Method (EEM)[30] method, which uses the electronegativity χ_n and hardness $\gamma_{n,j}$ parameters for each element in the system, optimized by quantum mechanical data. The colombic energy is then determined by Equation B-11. Other methods exist, such as the Charge Equilibration (Qeq), which uses ionization potentials and electron affinities to determine the charges.[31] Its extension with the

Extended Qeq (EQeq) that relies on multiple ionization energies and electron affinities has been shown to improve the charge predictions for high-oxidation state transition metals.[32]

$$E_{Coulomb} = C \cdot \frac{q_i \cdot q_j}{\left[r_{ij}^3 + \left(\frac{1}{\gamma_{ij}}\right)^3\right]^{1/3}} \quad \text{(B-11)}$$

Where γ_{ij} is one of the coulomb parameters that can be optimized for example to reproduce the Qeq orbital overlap correction.

The objectives of ReaxFF are to predict the heat of formations with errors of about 4 kcal/mol, bond lengths within 0.01 Å, and angles within 2° from their literature values. The use of these potential energy functions allows ReaxFF to simulate molecular dynamics by solving Newton's equations of motion for atoms, where the forces are derived from the potential energy gradients concerning atomic positions. Through these calculations, ReaxFF can dynamically adjust molecular structures, simulate reaction pathways, and predict products in chemical reactions—capabilities that are important in understanding and designing chemical processes and materials. It is adding a new step that dynamically adjusts bonding scenarios based on the atomic positions and energies, allowing for the simulation of complex reactions.

The next section will further elucidate the process of parameterization for gold and silver in ReaxFF, highlighting its importance in tailoring force fields for specific chemical environments.

For the consideration of Au/S/C/H, the force field parameters were calibrated using a combination of experimental and DFT calculations by Järvi et al.[29] First, an optimization of elemental sulfur interactions was performed, followed by S-C, S-H, and Au-C interactions, as an optimization of the Au-S parameters was finally done. The DFT functional used here was the Perdew-Burke-Ernzerhof (PBE), described in more detail in section C. They found that it is the best functional in terms of accurate prediction of experimental data. Binding and dissociation energies were then compared to DFT calculation for each interaction. A four-layer slab was used, with the last layer fixed with different specific parameters. Overall, the authors calculated a cohesive energy for Au of about 2.98 eV, which agrees well with experimental and DFT results. By plotting the interatomic potential energy for S-S, S-C, and S-H, as presented in Figure B-7,

they observed that while they were able to mostly match DFT results, the S-S bond interactions remained underestimated.

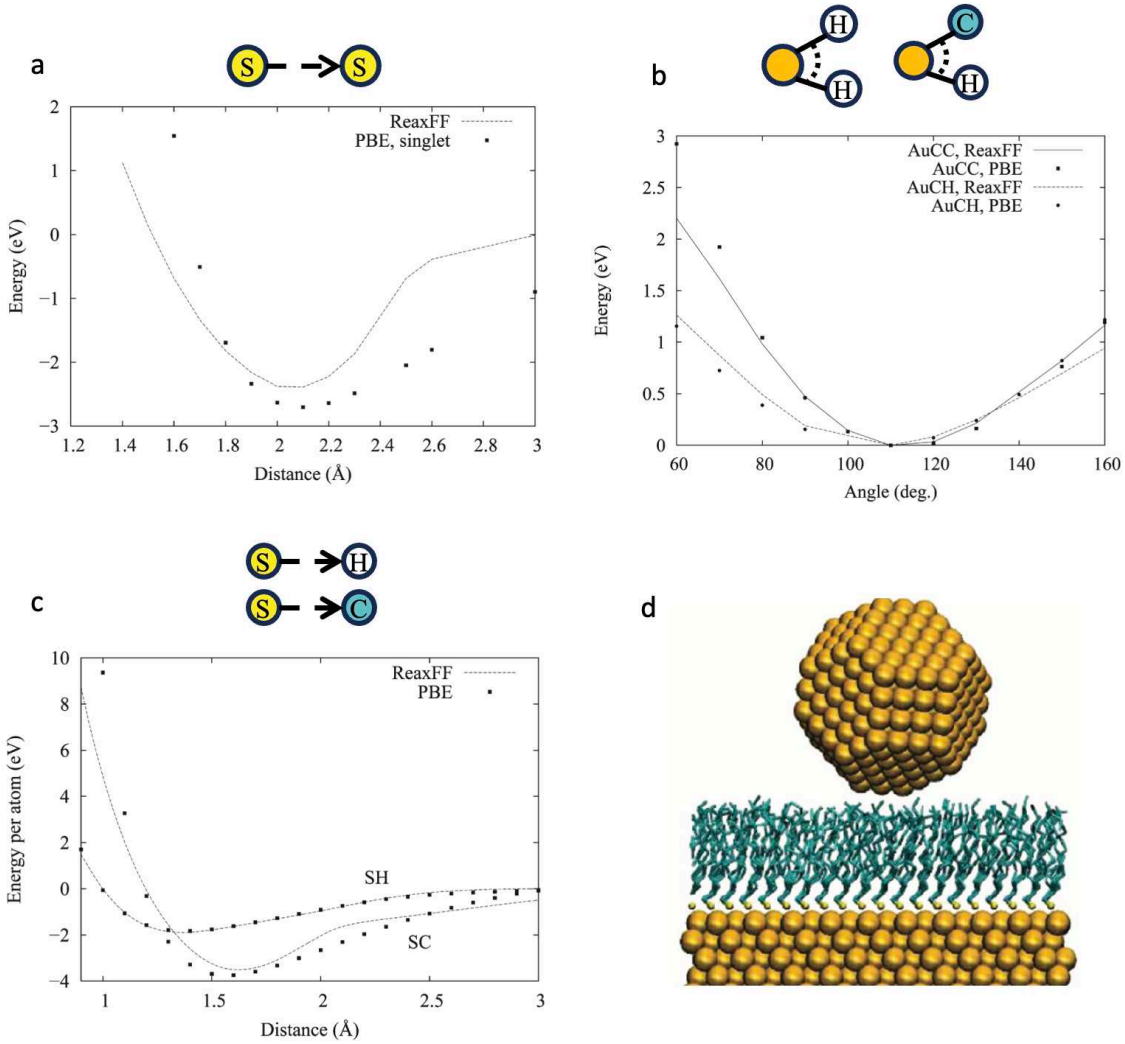


Figure B-7: The ReaxFF validation for the Au/S/C/H potential. (a) is the interatomic potential energy for S-S, and (b) is the distortion energy of AuCH and AuHH. (c) is the interatomic potential energy for S-H and S-C (d), is the simulation of a gold nanoparticle with a thiolates-SAM gold surface.

The potential's effectiveness was confirmed through its application in studying the interactions of gold clusters on thiol-monolayered gold surfaces. This simulation demonstrated the potential's ability to accurately model the reactivity between gold and the surface sulfur, particularly noting the formation of S-Au-S bonds. Consequently, these results validate the potential capability to simulate complex interfaces that involve thiol groups and gold, highlighting its utility in understanding and predicting the behavior of such nanoscale systems.

The drawback of the parametrization done by van Duin et al. is that it does not accurately reproduce the geometry of gold-thiolate systems, particularly for the S-Au-S staple motifs.

In practice, it is crucial to accurately describe this specific bonding observed experimentally in gold-sulfur systems. Bae et al. addressed this by reoptimizing both the Au-S and Au-Au bond parameters, as well as the S-Au-S angle bending parameters.[33] They employed DFT with a polarized triple-zeta (TZP) basis set with a frozen core configuration for each element involved—gold atoms were treated with a [$1s^2$–$4f^{14}$] frozen core, sulfur atoms with a [$1s^2$–$2p^6$] frozen core and carbon atoms with a [1s2] frozen core. Additionally, the zeroth-order regular approximation (ZORA) was used to account for scalar relativistic effects, which is crucial for simulating the behavior of heavy elements like gold. Upon testing, it was found that the original ReaxFF parametrization inaccurately predicted the S-Au-S angle, deviating from the expected 180° and not aligning with DFT calculations presented in Figure B-8.

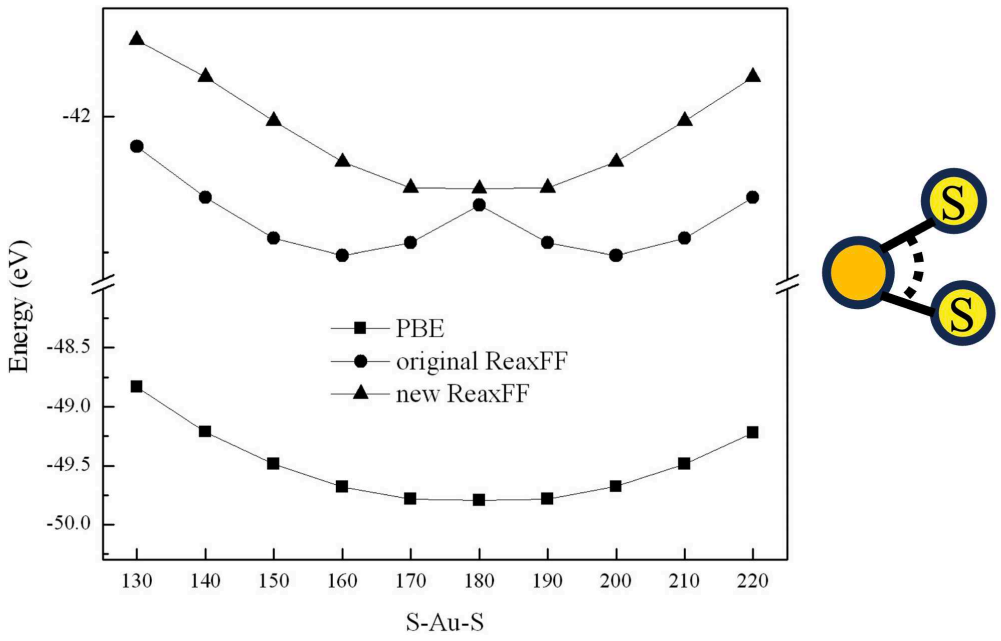


Figure B-8: Angle bending energy for S-Au-S. A comparison between the original ReaxFF (Van Duin et al.), the new ReaxFF (Bae et al.), and DFT calculation. Adapted from Bae et al. (2013)

The reparameterization aimed to refine this aspect, allowing for a more accurate representation of both small and large thiolate gold clusters. The new ReaxFF parameters improved the description of angle bending while maintaining bond accuracy compared to PBE-fitted models.

Further advancements were made by extending the simulation to dynamic and reactive environments with detailed pathway analysis, offering a broader understanding of functionalization processes under realistic conditions. This approach helps in designing experimental strategies for gold functionalization in practical applications involving proteins.[34] By incorporating parameters for nitrogen and oxygen, which are present in cysteine, the simulation successfully mimicked actual biological conditions. Additionally, the restructuring process and the formation of a staple were observed during prolonged simulations when the adsorbate was surrounded by water molecules, presented in Figure B-9.

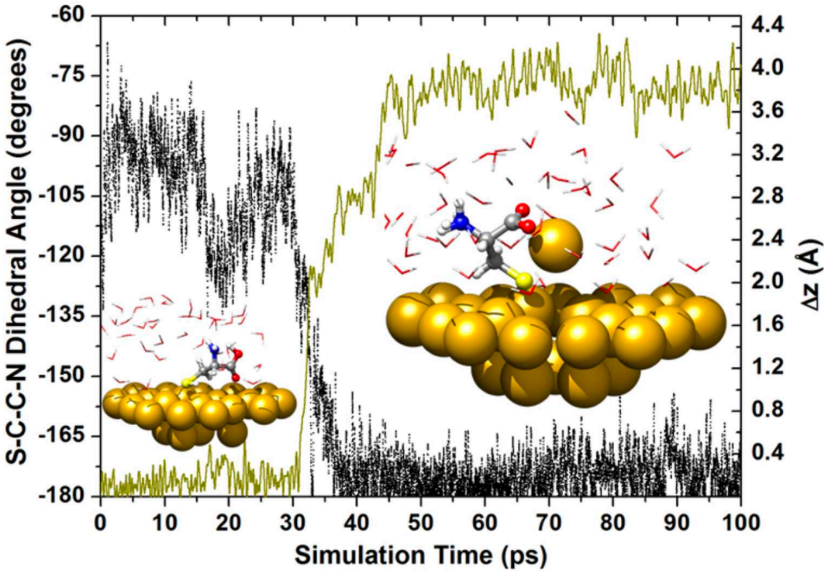


Figure B-9: Evolution of the torsion angle of the nitrogen group (in yellow) during a simulation of a gold surface surrounded by 330 water molecules and one cysteine. The evolution of the distance between an anchor gold atom and its initial position along the z-axis is also shown (in black). Adapted from Monti et al. (2016)[34]

We just saw how an initial ReaxFF potential (for gold) could be reoptimized to be more accurate in the description of specific bonds (Au-S) to capture complex reactions or to match with new ligands containing new atoms (cysteine). Now, concerning silver, Dulong et al.[27] developed a new reactive force field (ReaxFF) for silver clusters capped with thiolate self-assembled monolayers (SAMs), enhancing the existing parameters by comparing them with those developed

by Lloyd et al. for Ag/Zn/O systems.[35] Initially, a comprehensive training set comprising theoretical data—including geometries, relative energies, potential energy surfaces, and atomic charges—was constructed to inform the parameterization process. The objective was to minimize the squared weighted errors (SSE) across these data points, normalizing each error by the standard deviation to fine-tune the ReaxFF predictions. The validation of the newly optimized parameters involved using a probability function, with simulations run across eight replicas to ensure statistical robustness. The theoretical data were calculated with DFT calculations employing the PBE functional, with silver modeled using the modLAN2DZ basis set, which incorporates pseudopotentials to enhance the accuracy. For sulfur, carbon, and hydrogen, the 6-311++G (2d, dp) basis set was utilized. Post-parametrization, the new ReaxFF model was compared with the ReaxFF(lit) from Lloyd et al., as depicted in Figure B-10, specifically analyzing the maximum deviations in Ag-Ag bonds across various cluster sizes. This comparison revealed that deviations decrease with larger clusters (a minimum of 20 silver atoms), affirming that the new model provides a more accurate description of Ag-Ag bonding in larger clusters compared to the previous model.

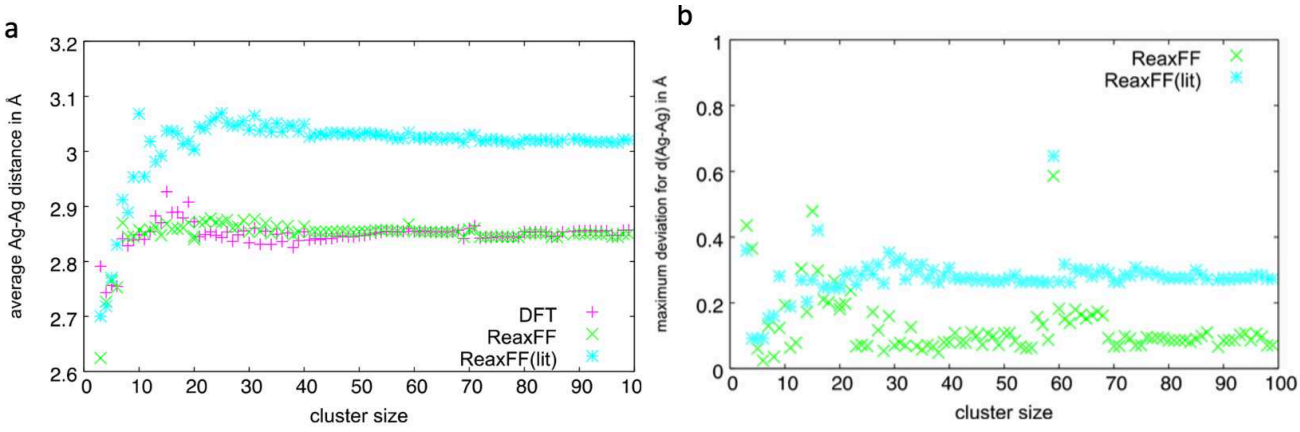
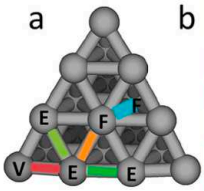


Figure B-10: (a) Average distance and (b) distance deviation of the Ag-Ag bond for different cluster sizes. A comparison between the force fields of Lloyd et al.[35] and Dulong et al. Adapted from Dulong et al. (2020).[27]

Optimization of the cage-like Ag20 cluster was performed to determine if the interatomic distances align with those predicted by density functional theory (DFT). This cluster features diverse atomic sites, including facets, edges, and vertices. It should mimic the behavior of larger clusters yet possess a face-centered cubic (fcc) structure that does not demand excessive computational resources to treat it using DFT.

However, a significant limitation of the ReaxFF model is its uniform treatment of atoms of the same element within the cluster. This illustrates the model's inherent constraints and highlights its inability to differentiate individual atoms based on their positions. While Lloyd's initial potential tends to overestimate distances, Dulong's revised ReaxFF model more accurately averages these distances. It achieves a mean value of 2.85 Å for silver, as presented in Figure B-11.



	Ag_{20} DFT	Ag_{20} ReaxFF(lit)	Ag_{20} ReaxFF(new)	Au_{20} DFT	Au_{20} ReaxFF
VE	2.77	2.81	2.82	2.71	2.877
EE_{edge}	2.77	3.02	2.85	2.67	2.879
EE_{fac}	2.91	3.03	2.85	2.94	2.885
FE	2.81	3.03	2.85	2.81	2.882
FF	3.04	3.00	2.85	3.15	2.880

Figure B-11: Comparison of the distances for the Ag_{20} pyramid-like structure optimized by DFT, ReaxFF from Lloyd et al.[35] , and ReaxFF from Dulong et al. The Au_{20} pyramid-like structure is also optimized by DFT and ReaxFF from Järvi et al.[29]

Comparatively, for gold, distances derived from *Järvi*'s potential averaged 2.88 Å, as presented in Figure B-11, which is not consistent with DFT calculations. *(This inability to differentiate between atoms of the same type, a fundamental aspect of the ReaxFF approach, is a primary motivation for the development of new potentials discussed in Section C of the project.)* To validate the revised potential, energies were analyzed, comparing results from DFT calculations with both Lloyd's original and Dulong's new ReaxFF models. This included evaluations from cluster energies to dissociation energies for Ag-Ag and Ag_{19}-Ag interactions. Notably, the new model significantly improved the accuracy of energy estimations for small clusters, with the minimum energy error for cluster dissociations reduced to 1.2 kcal/mol, compared to the 15 kcal/mol observed with the previous potential and presented in Figure B-12.

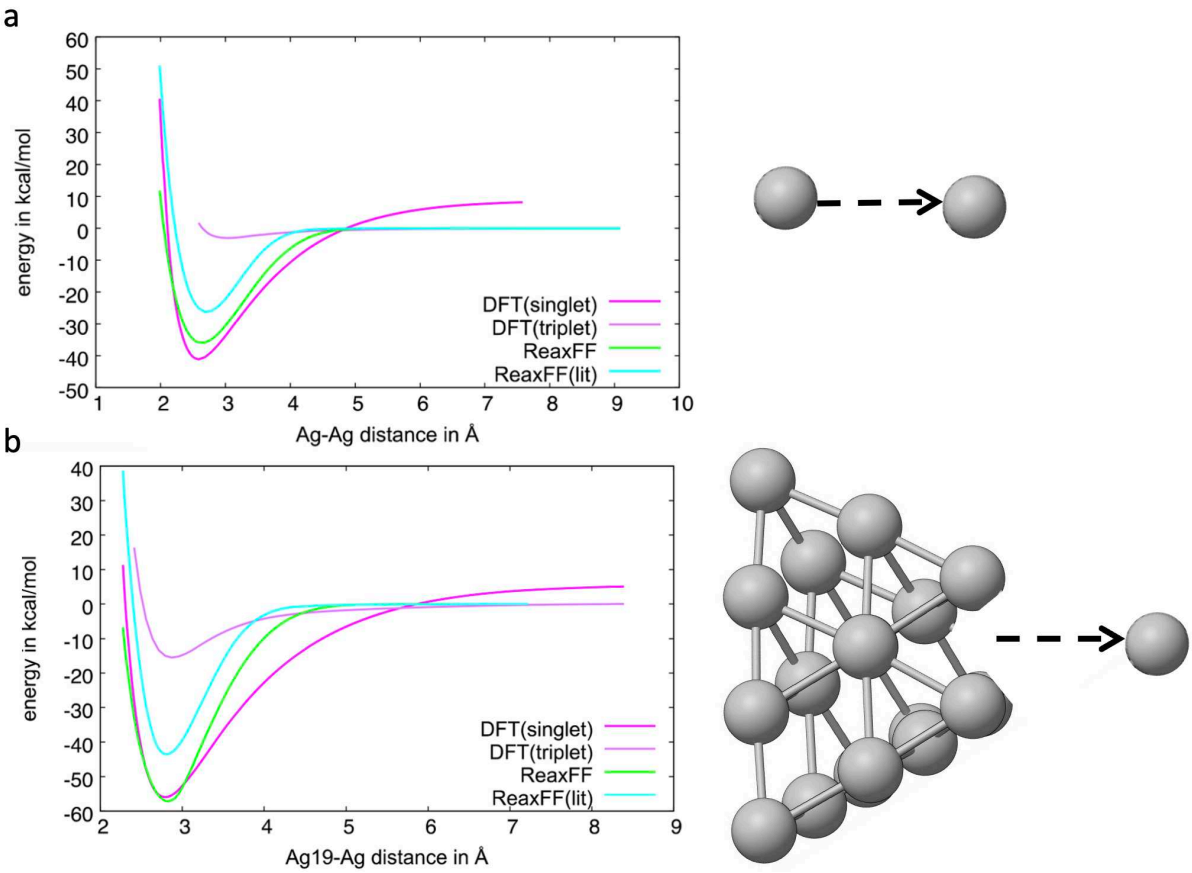
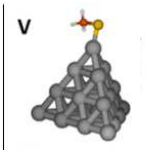
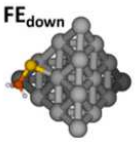
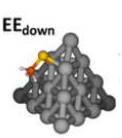


Figure B-12: Potential energy surface of the dissociation for the (a) dimer of silver Ag_2 and (b) Ag_{20} pyramid-like structure. Adapted from Dulong et al. (2020)[27]

The parameterization of the Ag-S system was conducted through the analysis of five distinct MeS-Ag_{20} isomers, each representing different sulfidation adsorption sites on the silver cluster.

The ReaxFF potential developed by Dulong et al. aligns closely with DFT calculations, exhibiting energy discrepancies within the chemical accuracy threshold of less than 1 kcal/mol. This is detailed in Table 1, which presents the optimized geometry energies for Ag_{20}-SCH_3 isomers, comparing the results from DFT with the potential from Dulong et al.

Table B-1: Optimized geometry energies for different Ag_{20}-SCH_3 isomers. Comparing DFT with potential from Dulong et al.

					
ReaxFF	11.9	0.0	3.3	11.1	11.1
DFT	11.6	0.0	3.1	11.8	10.6

Furthermore, the energy hierarchy across various adsorption sites is maintained, as shown in Table B-1. The corresponding potential energy surface for the dissociation of the Ag-S bond closely matches DFT calculations, further validating the accuracy of the ReaxFF model developed by Dulong et al., as illustrated in Figure B-13.

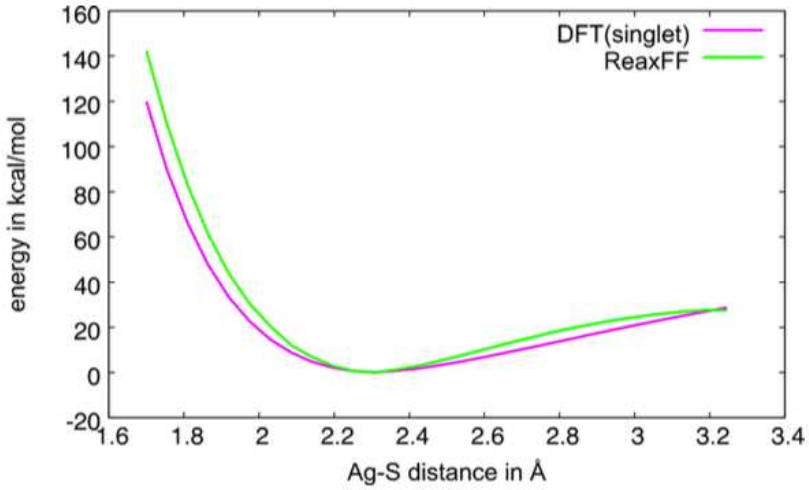


Figure B-13: Potential energy of the dissociation of Ag-S bonds. A comparison between ReaxFF was made by Dulong et al. and the DFT. Adapted from Dulong et al. (2020)[27]

The primary goal of this research is to employ these parametrized force fields to explore the reactivity of silver surfaces and nanoparticles under various conditions. This includes studies on complex reactivity phenomena such as sulfidation, core-shell interactions, and the formation of defects during simulations, ensuring that the model can accurately predict these intricate chemical processes.

B.2) Computational strategy

Our goal is to see if the present parametrization of ReaxFF potential can predict the correct reactivity on silver surfaces and nanoparticles when capped with alkylthiolates. Two different strategies have been used depending on the metal that we studied. For gold, three different FF exist, the first was created in 2011 by van Duin et al. and allows one to perform calculations on a system containing gold, sulfur, carbon, and hydrogen.[29] As explained before, it has specifically been designed to study the cluster capping by SAMs. Then, the approach by Bae et al. is an improvement of the previous iteration of the force field.[33] In this FF, the angle bending parameters have been improved as they saw that the PES for S-Au-S bending was not fitting DFT. Aikens changed the parameters to better describe those partial energies. The last potential, created by Monti et al., has been parametrized to simulate the mechanism of interaction of a cysteine with gold in water solution.[34] In this study, they also took into account the formation of staples and accurately parametrized their potential to reproduce such a feature. From those three FF, we tried to see which one would simulate the best the SAMs of thiolates on gold surfaces. For silver, the Dulong et al. potential is used.[27]

B.2.a) NATOMOS

NATOMOS (Nanoscale ATOMistic Optimized Simulations) is a program improved during the thesis, which is designed to provide comprehensive insights into the structural dynamics and adsorption characteristics of nanocrystals. Moreover, it allows one to create different structures using the SAMmaker module.

1. Creation of the different systems

The SAMmaker tool developed in this thesis generates the positions of metallic atoms and ligands for $\sqrt{3}$x$\sqrt{3}$R30°an $\sqrt{7}$x$\sqrt{7}$R19.1°self-assembled monolayers (SAMs). Initially, SAMmaker calculates the positions of the metallic atoms within an fcc (face-centered cubic) slab, with the highest layers corresponding to the (111) plane. The lines of the metallic atoms forming the (111) plane point in the x direction, as illustrated in Figure B-14.

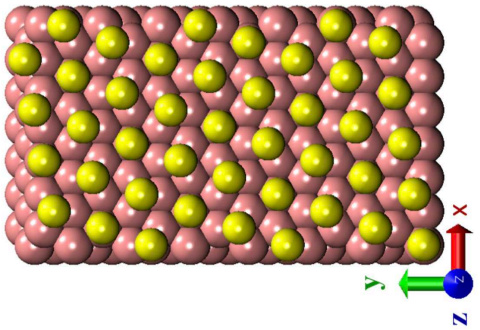


Figure B-14: $\sqrt{7x7}R19.1°$ Metallic Surfaces in red-capped with sulfur in yellow created by SAMmaker

The user can then define the positions of the ligands for the first lines in the x direction. The sulfur-headgroups ligands for the different assemblies are colored in yellow in Figure B-14. These ligands are then repeated in the y direction, taking into account a possible shift in the x direction. Users can customize several parameters, including the length of the alkyl chain (ranging from methane to butane), the restructured conformation, the number of gold layers constituting the surface, and the size of the simulation box, as depicted in Figure B-15. The program automatically calculates the appropriate box size to facilitate easy parameterization of periodic boundary conditions (PBC).

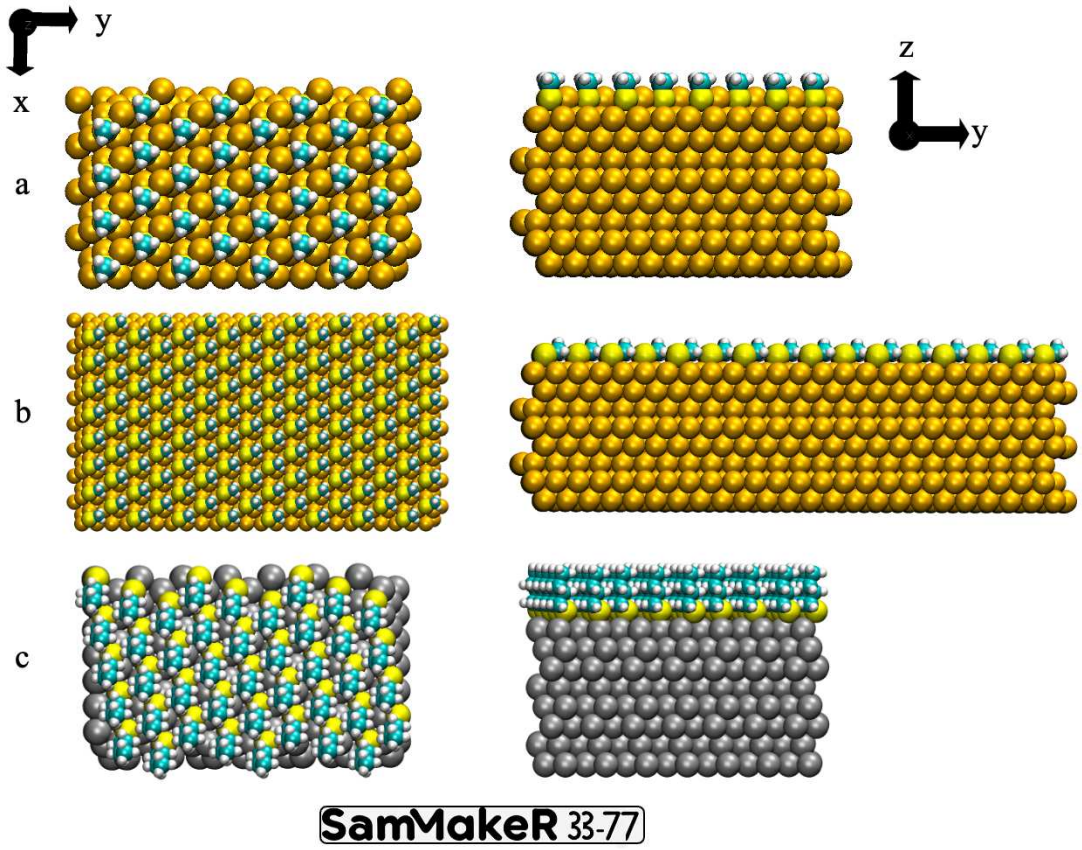


Figure B-15: Visualization of different assemblies that can be created by the use of SAMmaker. (a) is the surface of gold and SMe-Au-SMe staples. (b) the $\sqrt{3x3}R30°$assembly for unrestructured gold. (c) the $\sqrt{7x7}R19.1°$ for silver.

Another particularly useful capability of SAMmaker involves the capability of converting specific surface configurations into nanoparticles ranging from 2 to 10 nm in diameter. This transformation allowed for a comparative analysis between different nanoparticles and their corresponding surfaces. By doing so, we could investigate how nanoparticle size influences the stability and reactivity of thiolate-SAMs.

Additional features:

•**Customizable Alkyl Chain Length:** Users can select between methane and butane for the alkyl chain length, allowing for tailored simulations based on specific research requirements. Other ligands can be added depending on the file provided.

•**Restructured Conformation:** The tool provides an option to specify the restructured conformation, enabling more accurate modeling of realistic surface interactions.

•**Adjustable Surface Layers:** The number of gold and silver layers constituting the surface can be adjusted, providing flexibility in modeling different surface thicknesses.

•**Automatic Box Size Calculation:** SAMmaker automatically determines the correct size of the simulation box, ensuring proper parametrization of PBC and eliminating the need for manual adjustments.

2. **Analysis of the dynamics**

Reading Nanocrystal Data (readNC function): This function is the entry point of the simulation, where it reads the atomic coordinates from input files, setting the stage for subsequent analysis. This step is crucial as it establishes the initial conditions of the nanocrystals, defining how atoms are spatially arranged. The function ensures that the coordinates of gold or silver and ligand atoms are accurately parsed, providing the foundation for identifying structural features such as neighbors, edges, vertices, and facets.

Analyzing Nanocrystals (anaNC0 function): Here, the program calculates the neighbors of each gold atom based on a set distance threshold. This is essential for classifying atoms into structural categories such as vertices, edges, and facets, which are pivotal for understanding the nanocrystal's morphology and its potential reactive sites.

Ligand Adsorption Analysis (analig function): This function identifies and categorizes adsorbed ligands by their relative positions to the gold atoms. Quantifying ligands at different adsorption sites provides insights into the binding efficiency and site-specific characteristics of the nanocrystal, which are crucial for understanding the ligand's protective and functional roles, as discussed in the research articles.[36]

Staple Finder (staplefinder function): This function searches for staple motifs, which are indicative of strong ligand binding and restructuration of the nanocrystal.

SAM Finder (SAMfinder function): It identifies and analyzes self-assembled monolayers (SAMs) on the nanocrystal surfaces. The formation and analysis of SAMs are important for evaluating how modifications in the ligand or surface can affect overall nanocrystal behavior.

Output Generation: Finally, the program generates output files that can be used for further analysis or visualization. This includes detailed statistics about the nanocrystal and ligand configurations, which are crucial for publishing results or further computational analysis.

The program effectively integrates complex functionalities to simulate and analyze gold nanocrystals with ligands.

B.2.b) Data treatment and analysis

To validate the ReaxFF potential developed by Dulong et al. for accurately predicting the reactivity between thiolates-SAMs and silver surfaces, we employed several computational tools. Initially, SAMmaker was used to construct the thiolate-SAM silver surface assemblies, notably the 77 and 33 configurations. Periodic boundary conditions (PBC) were applied to maintain manageable surface sizes and prevent interactions with adjacent simulation cells. A space was intentionally left at the top of each surface to avoid unintended interactions with the surface below, as presented in Figure B-16.

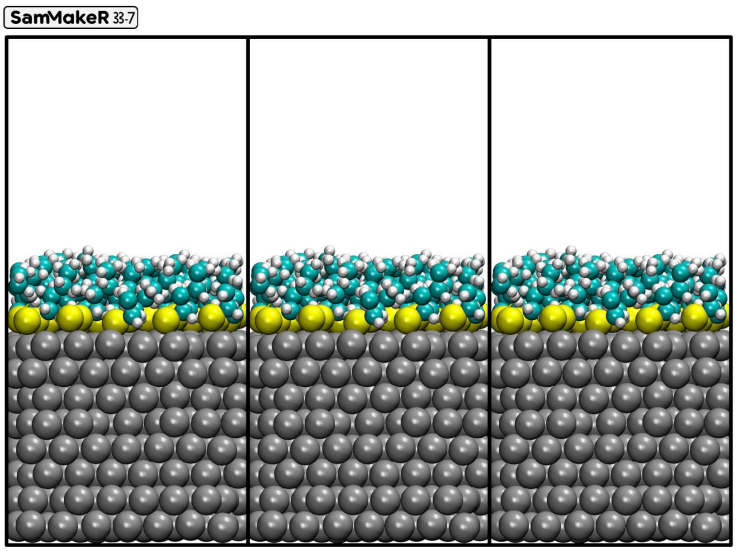


Figure B-16: Silver surfaces capped with thiolates-SAMs created by SAMmaker.

Following the assembly creation, NATOMOS was utilized to assess the stability of the SAM during heating processes and thermal equilibration at 300K. This analysis also extended to exploring various properties to detect any reactive effects that might manifest under these conditions.

References

(1) Spreiter, Q.; Walter, M. Classical Molecular Dynamics Simulation with the Velocity Verlet Algorithm at Strong External Magnetic Fields. *Journal of Computational Physics* **1999**, *152* (1), 102–119. https://doi.org/10.1006/jcph.1999.6237.

(2) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *The Journal of Chemical Physics* **1984**, *81* (8), 3684–3690. https://doi.org/10.1063/1.448118.

(3) Basconi, J. E.; Shirts, M. R. Effects of Temperature Control Algorithms on Transport Properties and Kinetics in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (7), 2887–2899. https://doi.org/10.1021/ct400109a.

(4) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *The Journal of Chemical Physics* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.

(5) Ruhle, V. Berendsen and Nose-Hoover Thermostats.

(6) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover Chains: The Canonical Ensemble via Continuous Dynamics. *The Journal of Chemical Physics* **1992**, *97* (4), 2635–2643. https://doi.org/10.1063/1.463940.

(7) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, Second edition.; Oxford University Press: Oxford, United Kingdom, 2017.

(8) *Langevin thermostat for rigid body dynamics | The Journal of Chemical Physics | AIP Publishing*. https://pubs.aip.org/aip/jcp/article/130/23/234101/925280/Langevin-thermostat-for-rigid-body-dynamics (accessed 2024-06-06).

(9) Vasseur, L.; Barbault, F.; Monari, A. Interaction between Yersinia Pestis Ail Outer Membrane Protein and the C-Terminal Domain of Human Vitronectin. *J. Phys. Chem. B* **2024**, *128* (16), 3929–3936. https://doi.org/10.1021/acs.jpcb.4c00965.

(10) Gopan, N.; Sathian, S. P. A Langevin Dynamics Study of Nanojets. *Journal of Molecular Liquids* **2014**, *200*, 246–258. https://doi.org/10.1016/j.molliq.2014.10.009.

(11) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(12) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Structure, Function, and Bioinformatics* **2006**, *65* (3), 712–725. https://doi.org/10.1002/prot.21123.

(13) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry* **2010**, *31* (4), 671–690. https://doi.org/10.1002/jcc.21367.

(14) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225–11236. https://doi.org/10.1021/ja9621760.

(15) Sambasivarao, S. V.; Acevedo, O. Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. *J. Chem. Theory Comput.* **2009**, *5* (4), 1038–1050. https://doi.org/10.1021/ct900009a.

(16) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A. Jr.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549–2564. https://doi.org/10.1021/jp910674d.

(17) Wang, Y.; Inizan, T. J.; Liu, C.; Piquemal, J.-P.; Ren, P. Incorporating Neural Networks into the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2024**, *128* (10), 2381–2388. https://doi.org/10.1021/acs.jpcb.3c08166.

(18) Wang, Y.; Solano Canchaya, J. G.; Dong, W.; Alcamí, M.; Busnengo, H. F.; Martín, F. Chain-Length and Temperature Dependence of Self-Assembled Monolayers of Alkylthiolates on Au(111) and Ag(111) Surfaces. *J. Phys. Chem. A* **2014**, *118* (23), 4138–4146. https://doi.org/10.1021/jp412285v.

(19) TOUKMAJI, A. Y.; BOARD, J. A. Ewald Summation Techniques in Perspective : A Survey. *Comput. phys. commun* **1996**, *95* (2–3), 73–92.

(20) van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396.

(21) Agrawalla, S.; van Duin, A. C. T. Development and Application of a ReaxFF Reactive Force Field for Hydrogen Combustion. *J. Phys. Chem. A* **2011**, *115* (6), 960–972. https://doi.org/10.1021/jp108325e.

(22) Shin, Y. K.; Kwak, H.; Vasenkov, A. V.; Sengupta, D.; van Duin, A. C. T. Development of a ReaxFF Reactive Force Field for Fe/Cr/O/S and Application to Oxidation of Butane over a Pyrite-Covered Cr2O3 Catalyst. *ACS Catal.* **2015**, *5* (12), 7226–7236. https://doi.org/10.1021/acscatal.5b01766.

(23) Lotfi, R.; van Duin, A. C. T.; Biswas, M. M. Molecular Dynamics Simulations of Perfluoropolyether Lubricant Degradation in the Presence of Oxygen, Water, and Oxide Nanoparticles Using a ReaxFF Reactive Force Field. *J. Phys. Chem. C* **2018**, *122* (5), 2684–2695. https://doi.org/10.1021/acs.jpcc.7b09660.

(24) Chenoweth, K.; Cheung, S.; van Duin, A. C. T.; Goddard, W. A.; Kober, E. M. Simulations on the Thermal Decomposition of a Poly(Dimethylsiloxane) Polymer Using the ReaxFF Reactive Force Field. *J. Am. Chem. Soc.* **2005**, *127* (19), 7192–7202. https://doi.org/10.1021/ja050980t.

(25) Lahouari, A.; Piquemal, J.-P.; Richardi, J. ReaxFF Simulations of Self-Assembled Monolayers on Silver Surfaces and Nanocrystals. *J. Phys. Chem. C* **2024**, *128* (3), 1193–1201. https://doi.org/10.1021/acs.jpcc.3c07098.

(26) Psofogiannakis, G.; van Duin, A. C. T. Development of a ReaxFF Reactive Force Field for Si/Ge/H Systems and Application to Atomic Hydrogen Bombardment of Si, Ge, and SiGe (100) Surfaces. *Surface Science* **2016**, *646*, 253–260. https://doi.org/10.1016/j.susc.2015.08.019.

(27) Dulong, C.; Madebene, B.; Monti, S.; Richardi, J. Optimization of a New Reactive Force Field for Silver-Based Materials. *J. Chem. Theory Comput.* **2020**, *16* (11), 7089.

(28) Mohammadtabar, K.; Guerrero, E.; Romero Garcia, S.; Shin, Y. K.; van Duin, A. C. T.; Strubbe, D. A.; Martini, A. Development and Application of a ReaxFF Reactive Force Field for Ni-Doped MoS2. *J. Phys. Chem. C* **2023**, *127* (25), 12171–12183. https://doi.org/10.1021/acs.jpcc.3c00668.

(29) Järvi, T. T.; Van Duin, A. C. T.; Nordlund, K.; Goddard, W. A. Development of Interatomic ReaxFF Potentials for Au–S–C–H Systems. *J. Phys. Chem. A* **2011**, *115* (37), 10315–10322. https://doi.org/10.1021/jp201496x.

(30) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108*, 4315.

(31) Rappe, A. K.; Goddard, W. A. I. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95* (8), 3358–3363. https://doi.org/10.1021/j100161a070.

(32) Martin-Noble, G. C.; Reilley, D.; Rivas, L. M.; Smith, M. D.; Schrier, J. EQeq+C: An Empirical Bond-Order-Corrected Extended Charge Equilibration Method. *J. Chem. Theory Comput.* **2015**, *11* (7), 3364–3374. https://doi.org/10.1021/acs.jctc.5b00037.

(33) Bae, G.-T.; Aikens, C. M. Improved ReaxFF Force Field Parameters for Au–S–C–H Systems. *J. Phys. Chem. A* **2013**, *117* (40), 10438–10446. https://doi.org/10.1021/jp405992m.

(34) Monti, S.; Carravetta, V.; Ågren, H. Simulation of Gold Functionalization with Cysteine by Reactive Molecular Dynamics. *J. Phys. Chem. Lett.* **2016**, *7* (2), 272–276. https://doi.org/10.1021/acs.jpclett.5b02769.

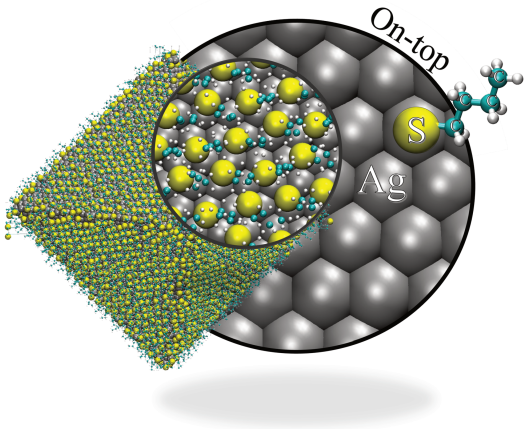
(35) Lloyd, A.; Cornil, D.; van Duin, A. C. T.; van Duin, D.; Smith, R.; Kenny, S. D.; Cornil, J.; Beljonne, D. Development of a ReaxFF Potential for Ag/Zn/O and Application to Ag Deposition on ZnO. *Surface Science* **2016**, *645*, 67–73. https://doi.org/10.1016/j.susc.2015.11.009.

(36) Richardi, J.; Fadigas, M. ReaxFF Molecular Dynamics Simulations of Large Gold Nanocrystals. *J. Chem. Theory Comput.* **2022**, *18* (4), 2521.

B.3) ReaxFF Simulations of Self-Assembled Monolayers on Silver Surfaces and Nanocrystals

Author's Note: The section that follows is adapted from an article originally published in the Journal of Physical Chemistry C (https://pubs.acs.org/doi/full/10.1021/acs.jpcc.3c07098).To clearly distinguish between the original content and some new additions, italic have been used. This formatting is intended to help differentiate between the unchanged text from the article and the newly added content.

Abstract: The self−assembled monolayers of alkanethiolates on Ag (111) surfaces and nanoparticles are studied using molecular dynamics. Reactive force fields allow simulations of very large systems, such as nanoparticles of 10 nm. Stable assemblies are obtained as experimentally observed for these systems. Only nanoparticles smaller than 4 nm show a spontaneous restructuring of the metallic core. The preferred adsorption site is found to be in an on−top position, which is in good agreement with recent X−ray absorption near edge structure experiments. Moreover, similar distances between the sulfur headgroups are found on the facets and edges.



B.3.a) Introduction

Silver nanoparticles (NPs) are characterized by unique optical, electrical, thermal, and biological properties.[1,2] In particular, they are employed as antibacterial agents in tissues, food products, and health care.[3] Thus, silver nanoparticles have exceptional antimicrobial activity. Moreover, their surface plasmon resonance gives them enhanced absorption and scattering of light. Therefore, nanosilver is nowadays one of the most widely used nanomaterials. For many applications, silver NPs are protected by ligands, such as thiolates or amines. These ligands form self-assembled monolayers (SAMs) on the NP surface. SAMs are usually used to avoid

nanoparticle's coalescence. But they are also useful to modify the properties of NPs, such as their solvability in specific solvents. Alkanethiolates are, for example, used to tune the antibacterial properties of silver NPs.[3] Their presence modulates the Ag+ ion release and helps to avoid nondesirable side effects. Recently, the synthesis of atomically defined thiolate-protected nanoclusters was achieved, opening up many possible applications.[4,5]

Our study focuses on the alkanethiolates SAM, which form on Ag(111) surfaces and NPs. Experimentally, a $\sqrt{7}x\sqrt{7}R19.1°$ SAM (denoted as 77 SAM in the following) is observed.[6,7] This geometry is denoted by Wood's notation.[8] It means that the thiolate head groups form a hexagonal assembly that is extended by $\sqrt{7}$ and turned by $19.1°$ with respect to the layer of Ag atoms on the Ag(111) surface. Moreover, experimental and theoretical studies have also shown that, due to the presence of thiolates, the silver surface can be restructured.[4,5] In this case, a silver atom appears between three sulfur atoms of the thiolates due to bonds breaking on the surface. For Au (111) surfaces, two superstructures are found, $c(2x4)$ and $\sqrt{3}x\sqrt{3}R30°$ SAMs (denoted as 33). The former is usually caused by the formation of staples, which corresponds to a restructuration of the surface and is found in the majority.[9,10] For this first study, however, we focus on non-restructured silver facets.

Silver NPs protected with thiolates have been widely studied by experiments in the last years[11–17] using techniques like EXAFS (extended X-ray absorption fine structure), XANES (X-ray absorption near edge structure), or XPS (synchrotron-based X-ray photoelectron spectroscopy). These experiments led to two central questions. First, in some experiments, the presence of silver sulfide was observed between the Ag core and the surrounding thiolates.[11,12,17] However, other experiments found only a small amount of silver sulfide.[13–17] The eventual origin of silver sulfides has not been elucidated until now.[17] Second, very recent experiments indicate an on-top adsorption site for the thiolates. This is in contradiction to recent theoretical density functional theory (DFT) studies, which predict a preference for the bridge position on flat silver surfaces and NPs.[18,19]

The aim of this paper is to apply the recently developed ReaxFF[20] force field (FF) in order to perform molecular dynamics simulations of thiolate SAMs on silver. Recently, several aspects of silver NPs have been studied using FF simulations[21–25] and DFT calculations.[20] The FF studies handle aspects that are not in close relation with thiolate-coated NPs, as investigated here. These

aspects include the interaction between silver NPs and proteins,[22] the interaction with amines,[23] and the formation of hollow silver NPs under laser irradiation.[24] The DFT study investigates the adsorption of methanethiolates on both silver nanoparticles and (111) surfaces. The results revealed that bridge sites are the most thermodynamically stable adsorption locations, emphasizing their significance. Additionally, it was observed that smaller nanoparticles display increased reactivity, accompanied by significant adsorption energies, further highlighting their role as reactive platforms.

However, to the best of our knowledge, there is only one study that focused on thiolates Ag(111) surfaces.[25] These simulations do not succeed in reproducing the 77 SAM on silver that was experimentally found. Moreover, no simulation of silver NPs with thiolates has been published. The use of ReaxFF may improve this situation for three reasons.

First, classical force fields using, for example, Lennard-Jones parameters naturally lead to hollow sites as preferred adsorption sites (see, for example, the study in ref [26]). This is because the hollow sites enable more contact between the adsorbing atom and the metallic atoms. Therefore, such FFs have difficulties simulating 77 SAMs characterized by the adsorption of bridges and on-top sites. Within the ReaxFF, the Ag−S interaction is handled as a bond, and angle-dependent interactions enable stabilization of the adsorption of bridge and on-top sites compared to the hollow sites. Thus, we will show here that the new FF is able to reproduce the 77 SAM on Ag (111) planes in agreement with the experiment.

Second, the silver surface may be restructured. This can only be simulated with reactive FFs, allowing for the breaking of the Ag−Ag bonds; this would be the case for ReaxFF. Moreover, ReaxFF enables us to study if the thiolate is decomposed during adsorption, which may explain the formation of silver sulfide.

Third, with the help of massively parallel computing, ReaxFF can investigate NPs made of a hundred thousand atoms in contrast to DFT methods. Thus, the largest system studied here is the butanethiolate-coated NP of 10 nm, made of 57 826 atoms. Moreover, the evolution with temperature can also be studied, which implies simulation lengths up to 4 ns. Such ReaxFF has already been applied for several metallic NPs made of cobalt, copper, and gold.[27–31] In particular, for gold NPs, it predicts staple formation and on-top adsorption sites in full agreement with experiments.[31]

Here, the ReaxFF is used to calculate the properties of silver surfaces and NPs (with diameters ranging from 2 to 10 nm) coated with thiolates. To study the influence of chain length, methanethiolates and butanethiolates are used. While methanethiolates have been widely studied on flat silver surfaces, for NPs, usually longer chains with at least butane have been used. For the sake of comparison, we nevertheless investigated methanethiolate on NPs. In this first study, we start from non-restructured surfaces to see if the restructuring spontaneously happens, as previously observed for gold NPs with thiolates. We also ignored the presence of the solvent. It is expected that the solvent plays a minor role in the adsorption properties, which are mainly governed by the strong Ag−S interactions. However, it might be nonnegligible for the energy difference between 77 and 33 SAMs, which should be investigated in the future.

The study focuses on nanocrystals of octahedral shape, which have been chosen since their facets are made of Ag (111) planes. This enables a direct comparison with the results obtained for the Ag (111) surfaces. Moreover, NPs of octahedral shape are often observed in experiments.[32]

The structure of this article is as follows: First, we present the ReaxFF approach and the simulation protocol. We will then introduce the new analysis tools, SAMmaker and SAMfinder, that were specifically developed for this study. They allow the construction and finding of SAMs on metallic surfaces and NPs. Then, the stability of the thiolate SAMs on the Ag (111) surfaces is investigated. Finally, we turn to the predicted properties of the SAMs on NPs.

B.3.b) Methods

B.3.b.1) The ReaxFF model

The reactive force field method proposed by van Duin et al.[33] enables the formation and breaking of bonds during classical molecular dynamics simulations. The formalism is described in detail in the paper by Chenoweth et al.[34] Here, we use the recently developed FF for AgSCH.[20] It is able to reproduce the DFT values for the average distances between silver atoms in NPs with a precision of 0.05 Å. It also correctly describes the energetical and geometric properties of thiolates on an Ag20 pyramid arrangement.[20]

The charges of the atoms are obtained through the usual electronegativity equalization method.[35,36]

B.3.b.2) Simulation Protocol

Molecular dynamics simulations are carried out in the NVT ensemble using a time step of 0.25 fs. The temperature is controlled using the Berendsen thermostat with a damping constant of 5 fs. The reliability of the Berendsen thermostat has been verified through simulations using the NVT Nose−Hoover thermostat, yielding identical results within the statistical margin. The parallel program LAMMPS is used for all simulations.[37,38] First, an equilibration is performed by increasing the temperature from 0 to 300 K. The simulation is then continued at 300 K to obtain better-converged properties. To test correct equilibration and convergence, three runs are performed:

- Equilibration: 0.3ns, convergence at 300K: 0.3 ns
- Equilibration: 1 ns, convergence at 300K: 1 ns
- Equilibration: 3 ns, convergence at 300K: 1 ns

To obtain standard errors for all properties, three independent simulations are carried out, starting from different initial atomic positions for the SAM position with a shift of 0.05 Å. It has been found that in the case of NPs, an alkane chain strictly perpendicular to the silver surface has to be avoided since this leads to a strong reduction of the SAM frequency of about 30%. Indeed, the alkane chains go together and move the S atoms from their initial positions corresponding to the chosen SAM.

B.3.b.3) Set up Tools

A new Python tool called SAMmaker was developed to set up SAMs on flat surfaces and NPs. It has been included in the Python library NATOMOS. *How this tool works is explained in the section 3a.*

We observed that it is very important to have a compact initial layer of alkane without strong repulsion to obtain stable SAMs. Therefore, all ligands are slightly declined and oriented in the same directions. Moreover, orientations with very small repulsions were searched for. Figure 1 shows typical snapshots obtained after equilibration for 77 and 33 SAM of thiolates on the Ag (111) surface.

To set up the NPs, the SAMs obtained with SAMmaker are transferred to the NP facets using the NATOMOS tool NTM_setup, which has been described in detail in a previous paper[31]

(and in section 3b and 3c). It is interesting to note that there are differences in the stability of SAMs for NPs compared to surfaces. Thus, starting from an orientation of the methanethiolate with the S−C bond perpendicular to the plane led to a well-defined SAM on the surface.

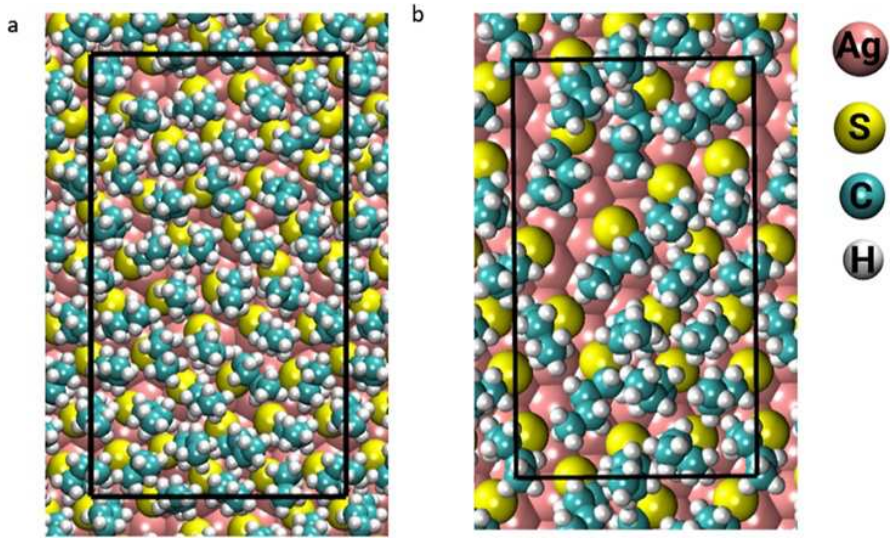


Figure 1: Snapshots of a SAM of methanethiolates on Ag (111) at 300 K after equilibration of 0.3 ns. (a) 77 SAM and (b) 33 SAM. The simulation box for each system is represented in black

However, when the same orientation is used for the NPs, the SAM becomes unstable during the simulation. This shows the importance of well-defining the initial orientation to obtain a large degree of SAM.

B.3.b.4) Analysis Tools

The SAMs were analyzed with the help of the Python tool NTM_ana. It was described in detail in a previous paper.[31] It determines the following properties studied in this paper:

• the number of neighbors for the silver atoms,

• the locations and types of the adsorption sites occupied by the ligands,

• the average distances of neighboring ligand head groups.

Please note that two atoms are counted as neighbors whenever the distance is smaller than the position of the minimum after the first peak of the pair distribution function (minimum for Ag−Ag: 3.75, Ag−S: 2.92, and S−S: 5.5 Å).

Here, a new function, SAMfinder, has been included in Ntm_ana. It allows the determination of the percentage of ligands participating in a SAM characterized as 33 or 77, for example. For this purpose, a percolation method is used. All properties obtained by the analysis are shown in Tables S1 to S4 in the Supporting Information.

B.3.c) Results and Discussion

B.3.c.1) Thiolate SAMs on Ag (111) Surfaces

Simulations on two different thiolate SAMs on an Ag (111) plane were carried out (see Figure 1). On one hand, the 77 SAM experimentally observed on non-restructured Ag (111) surfaces is studied. On the other hand, we also investigated the 33 SAM for silver, usually observed on an Au (111) when no staple formation occurs. The second SAM was chosen for the following reasons: The Ag (111) and Au (111) surfaces have the same geometry, leading to the same distances between the ligands in a given SAM. Moreover, staple formation, such as for Au (111), is not observed on Ag (111). Therefore, the 33 SAM without staples should also be possible for steric reasons on Ag (111). A successful force field should, however, predict that the 77 SAM is more stable than the 33 SAM, in agreement with experiments.

We now study the simulations starting from 77 and 33 SAMs as a function of the temperature. Figure 1 shows snapshots of both SAMs obtained at 300 K. First, the stability of the SAMs during the simulation is studied. Then, their energies and entropies are compared, and finally, their properties, such as the occupation of adsorption sites, are discussed. The temperature is continuously increasing. Every 50 K, the simulation is stopped and continued at a fixed temperature for 30k, 100k, and 300k steps corresponding to the three equilibrations (0.3, 1, and 3 ns). This is done to ensure that the results are well-converged. Only the last third of these runs at fixed temperatures are used to calculate the SAM frequencies and energies.

B.3.c.1.1. Stability of the SAMs on Silver Surfaces

Figure 2 plots the percentage of thiolates belonging to 77 and 33 SAM for methane and butanethiolate obtained for the longest simulation run (3 ns), as determined with the module SAMfinder.

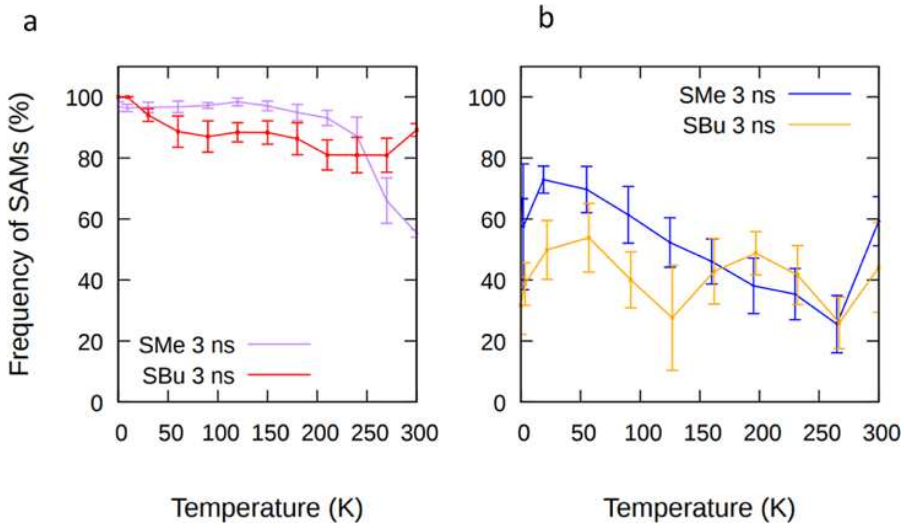


Figure 2: Frequencies of 77 (a) and 33 (b) SAMs of methane and butanethiolates on the Ag (111) surface for an equilibration step of 3 ns corresponding to an increase in the temperature.

The results for both 77 and 33 SAMs at different simulation times are shown in the F*igure S1*. Please note that parts a and b show the frequencies obtained from simulations starting from 77 and 33 SAM, respectively.

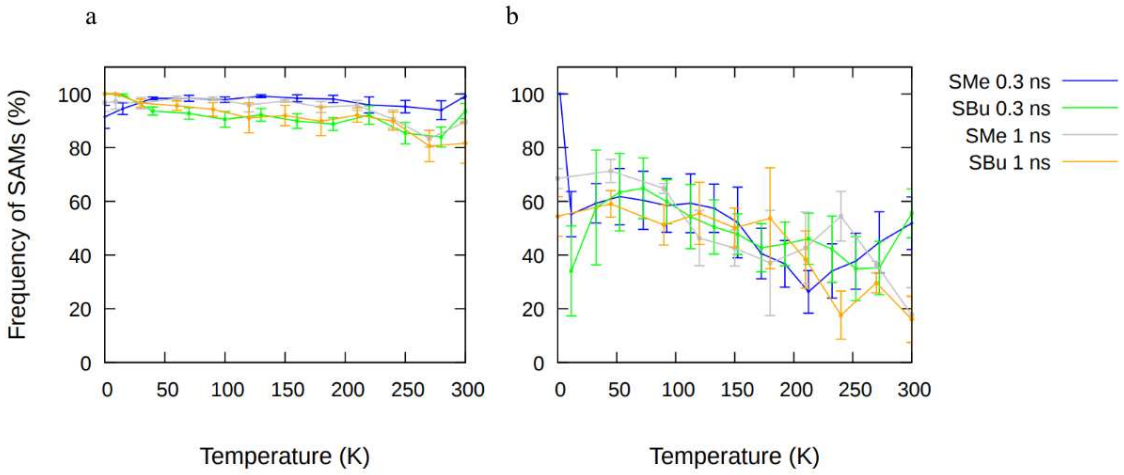


Figure S 1: Frequencies of 77 (a) and 33 (b) SAMs of methane and butane

In the case of butanethiolate, the 77 SAM occupied nearly the total surface up to 300 K, independent of the simulation time. This is also found for methanethiolate until 250 K. However, at higher temperatures, the percentage of methanethiolates in 77 SAMs decreases. This becomes more pronounced when the simulation time used for the equilibrations increases (see Figure S1). To investigate the stability of the 77 SAM at 300 K, the simulation run is continued for 1 ns at that

temperature (see Figure S2). We observed a complete disappearance of the 77 SAM for methanethiolate in contrast to butanethiolate, which rests stable.

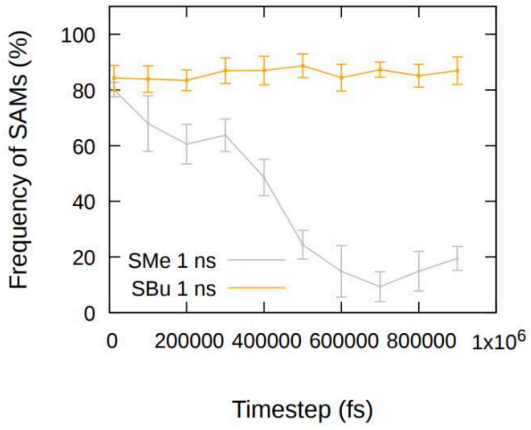


Figure S 2: Frequencies of 77 SAMs of methane and butane thiolates on Ag

Studying the snapshots at 300 K (see Figure S3), we found a desorption of the methanethiolates, which is not observed in experiments. Thus, it shows the limit of our force field. Since we want to study stable SAMs, we only considered simulations before this happens (the first 10 000 steps at 300 K).

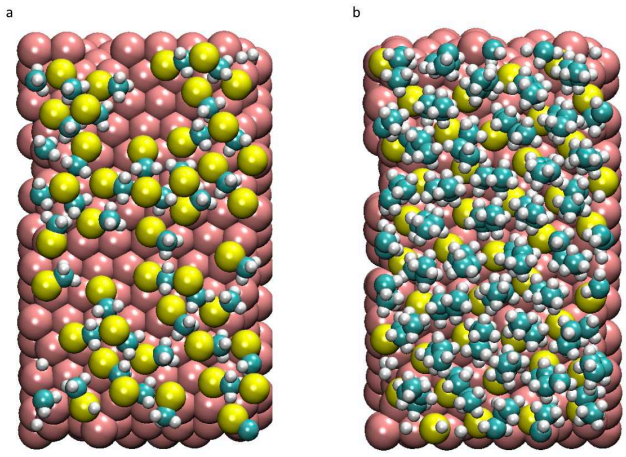


Figure S 3: Snapshot of both the methane (a) and butane (b) thiolates SAMs on Ag (111) surface after 4 ns.

Figure 2 also shows the frequency of 33 SAMs starting from a perfect 33 SAM, where the frequency of SAM rapidly decreases to about 20% for the methane and stays stable for the butane with some fluctuation during the simulation. Note that the results are the same within the statistical

error for the 1 and 3 ns simulations. This indicates that the results are well-converged. The only exception is the simulation for methanethiolate at temperatures larger than 200 K, which has been explained by a very slow desorption of the thiolates. To be sure that a similar process does not happen for butanethiolate, we have continued the simulation at any temperature for 3 ns. The results are stable within the statistical errors.

B.3.c.1.2. Energy and Entropy of the SAMs

In order to compare the stability of both SAMs, we first calculate their binding energies per thiolates, defined as

$$\Delta E_{bind} = \frac{1}{n_{Thio}}\left(E^{M\text{-}SAM} - E^{Metal} - n_{Thio}E^{Thiolate}\right)$$

where $E^{M\text{-}SAM}$ corresponds to the energy of the system, which corresponds to the SAMs in contact with the metal layer. E^{Metal} and $E^{Thiolate}$ are, respectively, the energies of the metal layer and of the single thiolate molecule. n_{Thio}corresponds to the number of thiolates present on the surface. We show in the following the differences in binding energy between 77 and 33 SAMs. By calculating this difference, one can easily see that the calculation of $E^{Thiolate}$ is not necessary, which allows us to decrease the possible errors that can emerge with the simulation of a small system using ReaxFF. Please note that the binding energy difference is negative, implies that the 77 SAM is energetically preferred.

Figure 3 displays the evolution of the binding energies as a function of the temperature for methane and butanethiolates on silver for 3 ns.

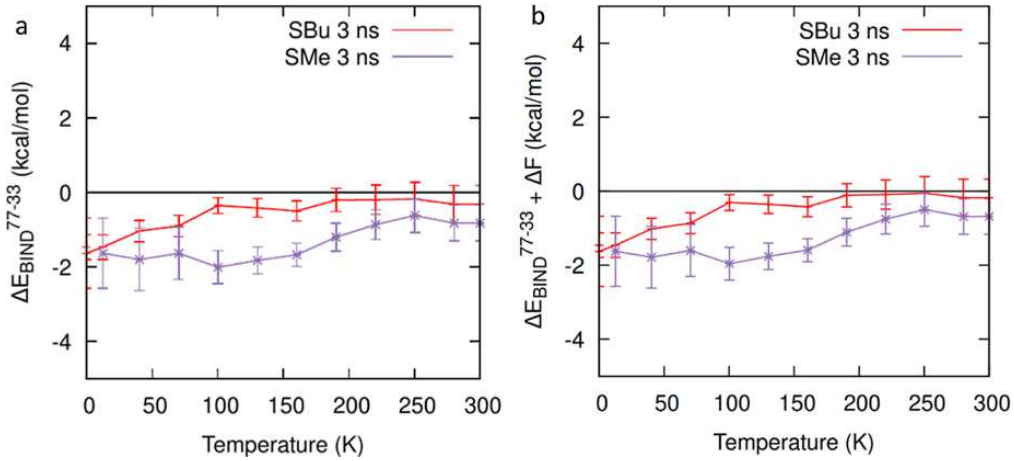


Figure 3: a) Is the difference of binding energies between the 77 and the 33 SAMs and (b) the entropy is added as a function of the temperature for both butane and methanethiolates on the Ag (111) surface. Simulation length: 3 ns.

In Figure S4, the results for 0.3 and 1 ns are shown. It is obvious that sufficiently long simulation runs must be carried out to obtain stable results.

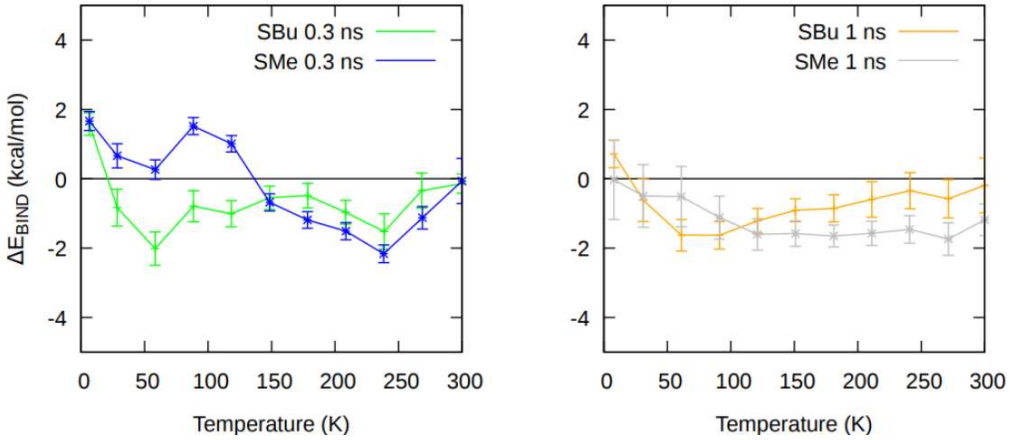


Figure S 4*:* Difference of binding energies between the 77 and the 33 SAMs as a function of the temperature for both butane and methane thiolates on the Ag (111) surface. Simulation length: (a) 0.3 ns and (b) 1 ns

In the case of methanethiolate, Figure 3 shows that the binding energy difference varies between −2 and −1 kcal mol^{-1}. This indicates that 77 SAM is energetically favored. However, for butanethiolate, the binding energy difference is between −2 and 0 kcal mol^{-1}. Hence, the energetic stabilization of the 77 conformations is not proved in this case.

In order to determine which SAM is the most stable, we need to consider the influence of entropy. It can be estimated from statistical mechanics using the partition function of translation. *The free energy of a ligand in a SAM related to entropy can be obtained from equation:*

$$F = -kT \ln q_t$$

where qt is the partition function of translation for the ligand with mass m defined as

$$q_t = \frac{V}{h^3}\sqrt{2\pi mkT}$$

V is the volume in which the ligand can move for the SAM. The difference in free energy between both SAMs can then be calculated from

$$\Delta F = -kT \ln\left(\frac{V_{77}}{V_{33}}\right)$$

Using the relation $S = -\left(\frac{\partial F}{\partial T}\right)$*, it can be seen this ΔF is directly related to the entropy. It is now assumed that the ligand movement, perpendicular to the metallic surface, does not depend on the type of SAM, but on the metal-ligand interaction. Then the ratio of volumes can be approximated by V77/V33 ≃ S77/S33 where S77 and S33 are the surface per ligand for both SAMs. It can be calculated for the 33 and 77 SAMs as 20.37 Å2 and 15.84 Å*^{2}. The ligand molecule can move within the space left by the neighboring ligands. For the 33 SAM, this space is larger due to a larger surface per thiolate than that for the 77 SAM. It is, therefore, expected that the 33 SAM is entropically favored. In Figure 3b, the estimation of the entropy difference is added to show how it may influence the stability of the SAMs. The evaluation of the free energy part due to entropy at 300 K gave 0.149 kcal mol^{-1} between the 77 and the 33 SAM. This shows that the entropy favors 33 SAMs, as expected. However, in the case of methanethiolate, the binding energy is much larger than the estimated free energy due to entropy. This means that the 77 SAM is the more stable assembly for our force field, which is in good agreement with the experiment. As discussed above, this conclusion cannot be made in the case of butanethiolate.

B.3.c.1.3) The Properties of the 77 SAMs

We now look at the properties of the 77 SAMs, which will be important for the comparison with the results obtained in the following section for the NPs.

It is important to note that during all simulations, no spontaneous decomposition of the ligands is observed. We first study the number of Ag neighbors around the silver atom. The expected numbers of 12 and 9 neighbors for the bulk and surface atoms, respectively, were obtained as shown in Figure S5.

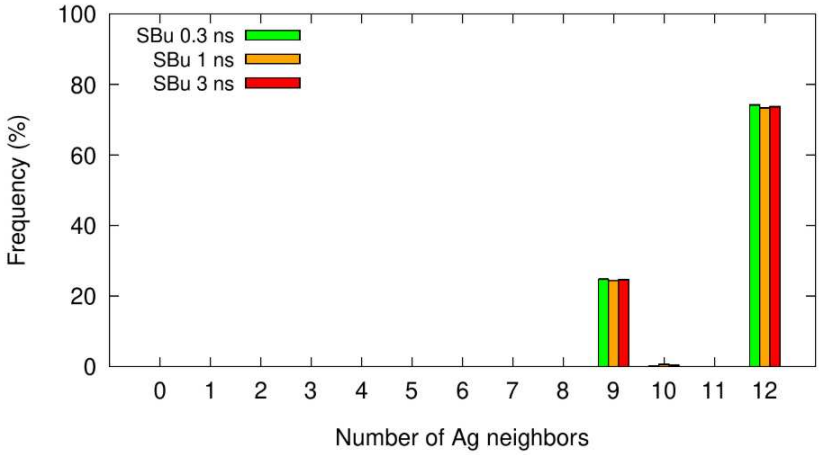


Figure S 5: Frequencies of atoms as a function of the number of Ag neighbors

No spontaneous restructuring of the metallic layer is observed. DFT theory and experiments, however, have shown such a restructuration.[6,7] This may mean that such a restructuring is too slow to take place in our simulations. In future simulations starting from a restructured surface, it is planned to study their stability and properties in comparison with the non-restructured layer investigated here.

We now turn to the occupation frequencies of adsorption sites shown in Figure S6 for butanethiolates. A comparison of the results obtained after 1 and 3 ns shows agreement within the statistical accuracy, which indicates that the results are well converged.

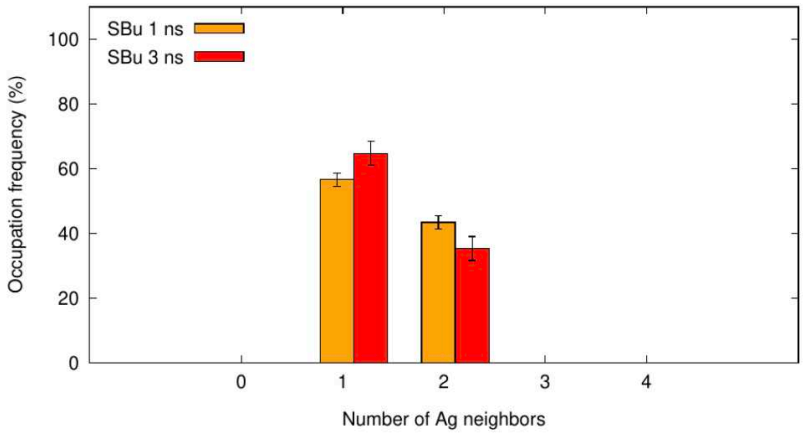


Figure S 6: Occupation frequencies of adsorption sites on the facet for butane thiolates on Ag (111) surface at different simulation lengths, where respectively 1, 2 and 3 Ag neighbors correspond to on-top, bridge and hollow sites.

A similar occupation of the on-top (52%) and bridge (48%) sites is found where the S atom is in contact with one or two silver atoms, respectively. These sites are typically occupied in 77 SAMs when studied by DFT.[6] Please note that we found that these sites are usually occupied in a way so that the adsorbing atom is shifted from the ideal on-top or bridge site.

Finally, the average distance between the sulfur head groups of the ligands was determined (Figure S7). A value of 4.4 Å is found, which corresponds to the one expected for the 77 SAM.

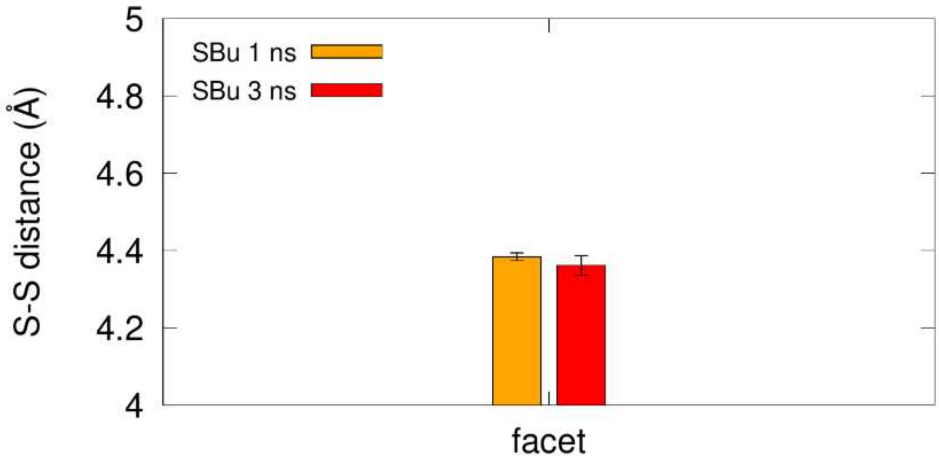


Figure S 7: Average distance between sulfur atoms neighbors of butane thiolates on a 77 SAM at different simulation lengths 1 ns and 3 ns. Result at the end of the equilibration at 300 K

B.3.c.2) Thiolate SAMs on Silver Nanoparticles

Figure 4 shows the typical snapshot obtained for simulations of 2 and 10 nm NP with butanethiolate at 300 K (1 ns).

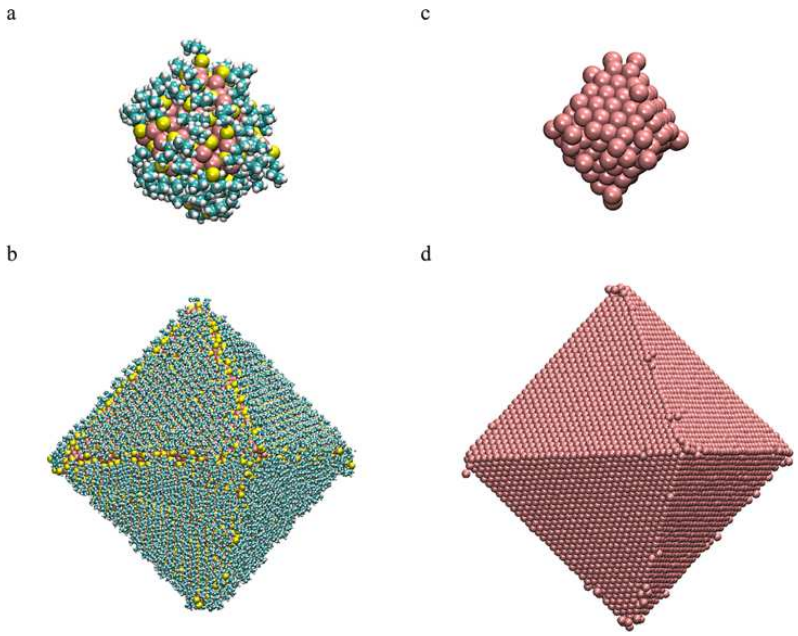


Figure 4: Snapshots of the simulation of butanethiolate-covered silver nanoparticles of 2 nm (a) and 10 nm (b) at the end of the convergence step for 1 ns at 300 K. The ligands are removed to show any reconstruction of the metallic core in (c) and (d) snapshots.

Figure S8 presents snapshots of 4 nm. The NPs of 2 and 4 nm are also shown without the ligands to make any restructuration visible. No spontaneous decomposition of the thiolates was observed. Therefore, the formation of silver sulfide observed in some experiments[11,12,15] cannot be explained by these experiments. However, we cannot exclude that a restructured silver layer may lead to this kind of decomposition, which should be studied in the future. Also, other ligands such as allyl thiolate (for which silver sulfide has been often observed[12,17]) should be tested.

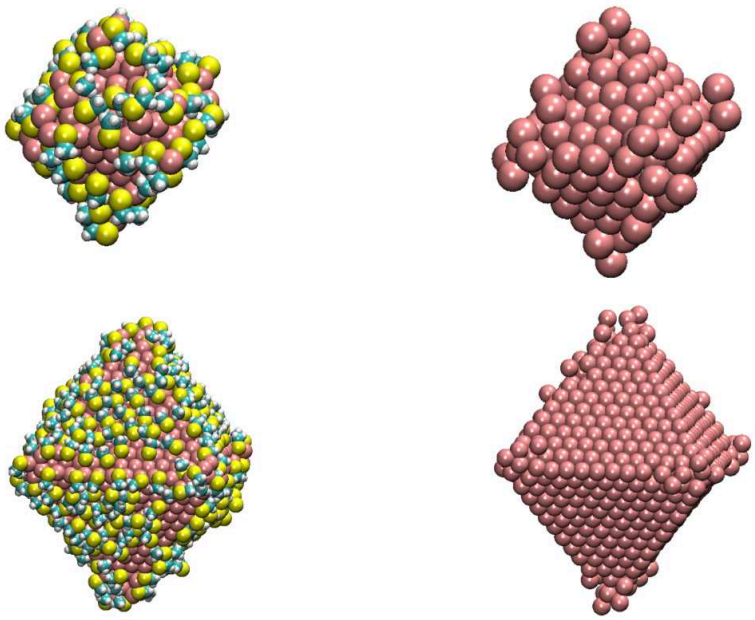


Figure S 8: Snapshots of the simulation of methane thiolate covered silver nanoparticles of 2 nm (a) and 4 nm (c) at the end of the convergence step for 1 ns at 300 K. The ligands are removed to show any reconstruction of the metallic core on b and d snapshots.

We will now first study whether the 77 SAM is also stable on NPs. In the following, the properties, such as the occupation of adsorption sites and the distance between the sulfur atoms, are discussed for the NPs. The sulfur atoms on the edges in Figure 4 may give the impression of being separated from the alkane chains. This is not the case. They are oriented only to the facets of the NPs.

B.3.c.2.1) Is the 77 SAM Stable on Nanoparticles?

We now study the frequency at which 77 SAMs are found on the NPs for butanethiolates. For 4 nm, a butane frequency of about 78% is found at low temperatures, as shown in Figure 5. This is significantly lower than the frequencies between 90 and 100% observed on the flat surfaces (Figure 2).

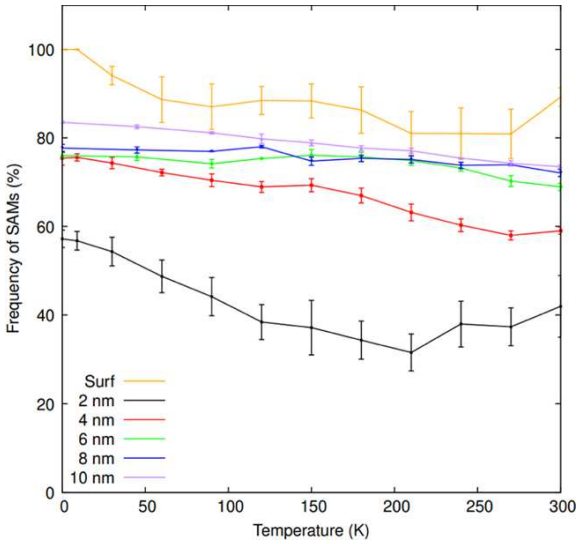


Figure 5: Frequencies of 77 SAMs on NPs for butanethiolates depend on the temperature. Comparison between surfaces and NPs of different sizes from 2 to 10 nm; equilibration time of 1 ns.

To understand this difference, we studied where the SAMs are located (see Figure 6). Therefore, all sulfur atoms belonging to SAMs are colored yellow. They are found principally in the centers of the facets, and the SAM formation is perturbed by the edges.

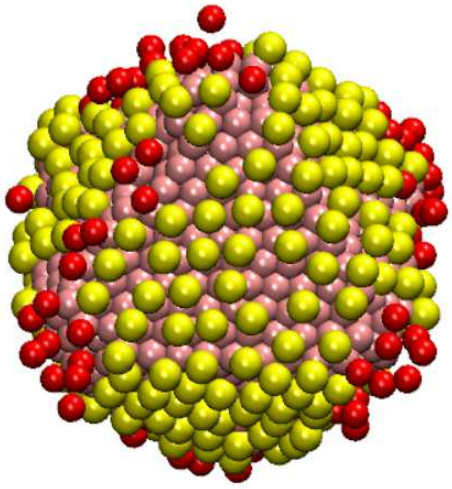


Figure 6: Snapshots of the simulation of 77 butanethiolate SAMs on Ag (111) nanoparticles of 4 nm at the end of the convergence step for 1 ns at 300 K. Atoms belonging to the SAM are colored in yellow.

By increasing the temperature, the SAM frequency decreases. A frequency of 77 SAMs of 60% is observed at 300 K. Let us look at the evolution of the 77 SAM frequency with particle size (see Figure 5). From 2 to 10 nm, the SAM frequency decreases at 300 K for butanethiolate.

Nevertheless, depending on the NP diameter, the decrease is more or less pronounced. The more we increase the NP diameter, the more the coverage of the facets grows, and the SAM formation is then less perturbed by the edges. Figure S9 illustrates the stability profiles of 77 SAMs for both methane and butane thiolates on a 4 nm Ag (111) nanoparticle, observed over different simulation durations. It is evident that as the simulation length increases to 1 ns, particularly around 200 K, the instability of methane thiolates becomes more pronounced.

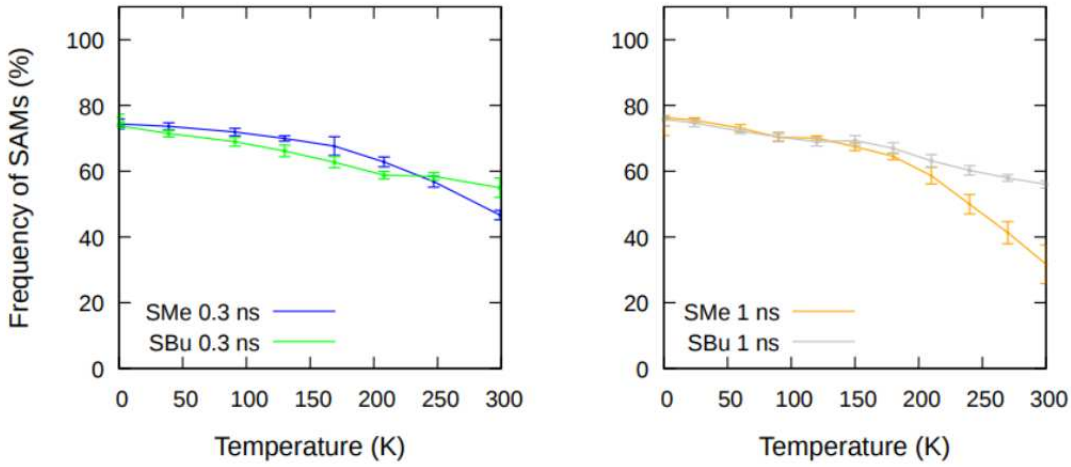


Figure S 9: Frequencies of 77 SAM for both methane and butane thiolates SAMs on a 4 nm Ag (111) NP for different for 0.3 ns, 1 ns and 3 ns simulation length during the equilibration step

B.3.c.2.2) The Structure of the Metallic Core

We now turn to the number of neighbors observed for the silver atoms. For a perfect octahedron, 7 and 9 neighbors are found on the edges and facets, respectively. A change in these numbers would indicate a metallic restructuring due to the ligands, for example, on the edges. This local restructuring should be distinguished from the global one observed on silver surfaces in experiments. As explained in Section 3. Such general restructuring is impossible to see during ReaxFF simulations. Such a phenomenon was widely observed in the ReaxFF simulation for gold NPs, where the metallic atoms were extracted on the facets and, in particular, at the edges due to staple formation.[28,31] Figure 7 shows the frequencies of silver atoms as a function of their number of neighbors for different NP sizes with butanethiolate SAMs.

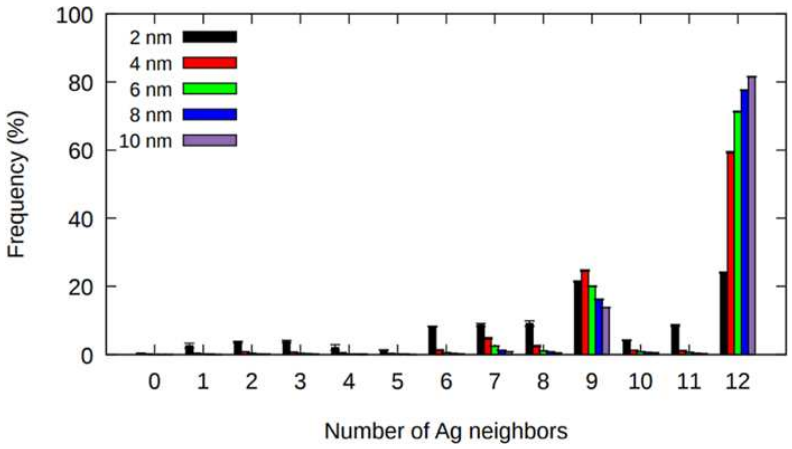


Figure 7: Frequencies of silver atoms as a function of the number of neighbors for NPs from 2 to 10 nm for butanethiolate (equilibration time of 1 ns and convergence time of 1 ns). The property was taken at the end of the convergence at 300 K.

In Figure S10, the neighbor frequencies of 4 nm NPs are plotted for butanethiolate at different simulation times. The frequencies of neighbors agree for the three runs of different lengths, showing that these results are well converged.

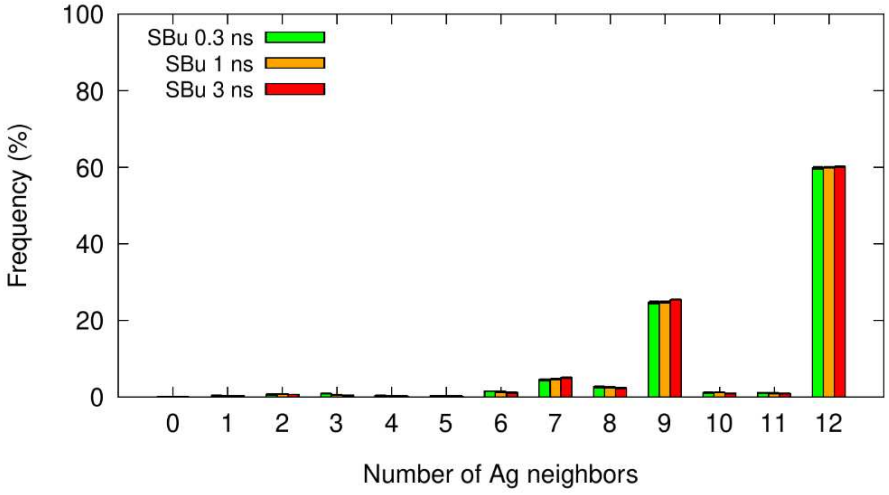


Figure S 10: Frequencies of silver atoms as a function of their number of neighbors for 4 nm for butane thiolate at different equilibration times from 0.3 ns to 3 ns. Property taken at the end of the convergence at 300K.

By focusing on Figure 7, it can be seen that the neighbor frequencies for the NPs larger than 2 nm are similar to those of a perfect octahedron, with peaks for 7, 9, and 12 neighbors. Only at the edges do some atoms change their positions leading to atoms with 6 or 8 neighbors. These

results are very different from the ones found for gold NPs previously discussed.[31] This shows that no surface restructuring spontaneously happens for large silver NPs.

However, for the 2 nm NP, the behavior is markedly different (see black lines in Figure 5). A large fraction of atoms changes the number of neighbors, even on the facets. This change in morphology is confirmed in the snapshots shown in Figure S8.

B.3.c.2.3) Location and Occupation of Adsorption Sites

We first examine where the ligands are located on the NP surface. Ligands are counted for the vertices or edges if they are in contact with at least one silver atom on the vertices or edges, respectively. The remaining ligands are attributed to the facets. For example, in the case of 4 nm particles with butanethiolate, 55.7, 39.4, and 4.9% of the ligands are adsorbed on the facets, edges, and vertices, respectively. These frequencies do not significantly change during the simulation and correspond to the frequencies initially set up. Similar behavior was observed for all NP sizes and ligands. In Table S2, all frequencies giving the locations of ligands are collected.

We now turn to the question of which adsorption sites the ligands occupy. In Figure 8, the frequencies of head groups as a function of their number of silver neighbors are shown for butanethiolate and different NP sizes. On a perfect Ag (111) surface, the headgroup of a ligand occupying the on-top, bridge, or hollow site has 1, 2, or 3 neighbors, respectively. For NPs larger than 4 nm, this is also true here since the degree of silver atoms extracted by ligands is very low, as previously shown, and the NP surface is well described by the Ag (111) plane. The frequencies of neighbors agree for the three runs with different simulation lengths (see Figure S11).

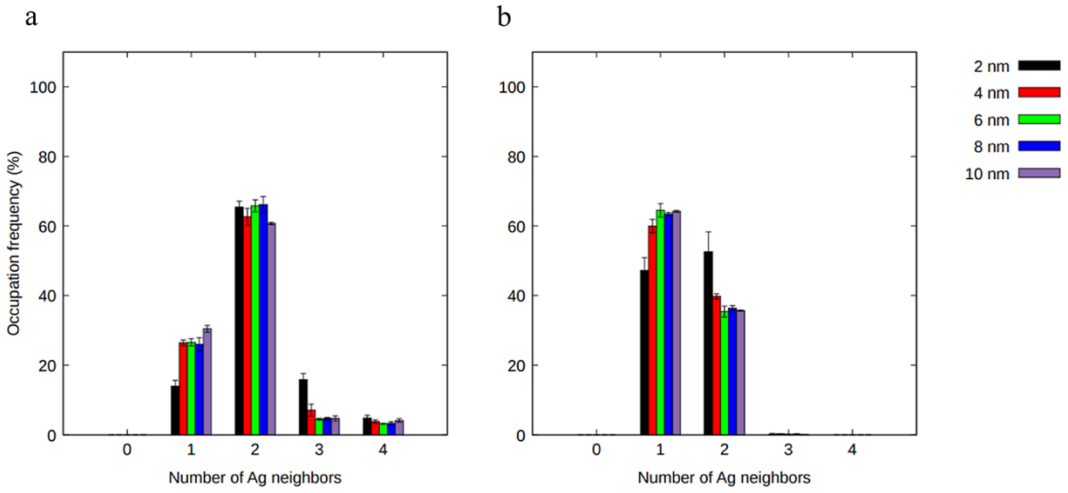


Figure 8: Occupation frequencies of S atoms in contact with a given number of silver atoms on the (a) edges and (b) facets for butanethiolates on the Ag (111) nanostructure from 2 to 10 nm with a simulation length of 1 ns at the end of the convergence at 300 K.

Similar to the surface results discussed in Section 3.1, the ligands on the NP facets occupy only on-top and bridge sites, with a preference for the on-top site. Also, the NP size has little influence on the occupation frequencies, except for the smallest NP of 2 nm. On the NP's edges, the results are quite different where the head groups are mainly in contact with two atoms, which indicates a preference for the bridge site.

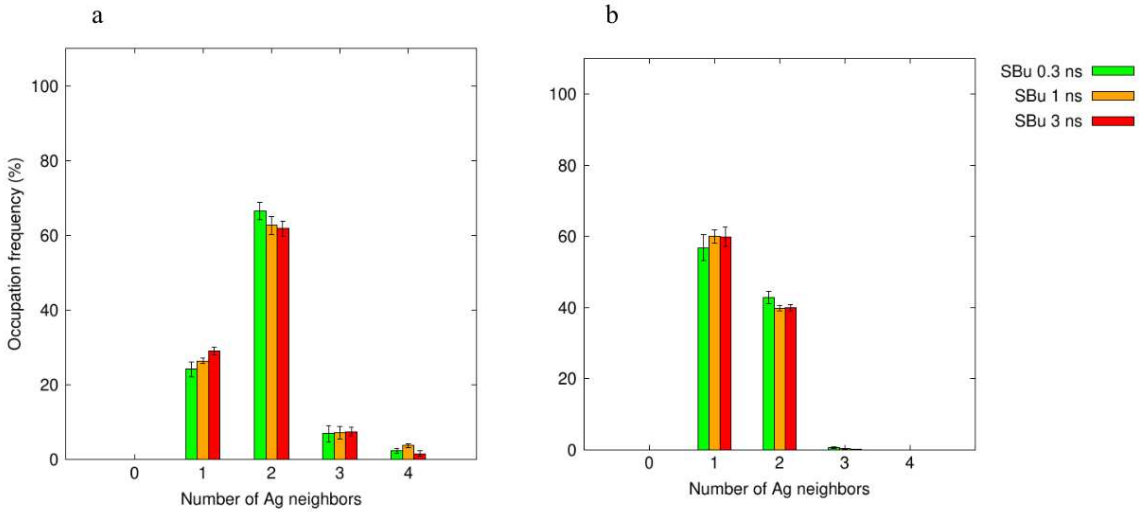


Figure S 11: Occupation frequencies of adsorption sites on the (a) edges and (b) facets for butane thiolates on an Ag (111) nanoparticle of 4 nm, where respectively 1, 2 and 3 Ag neighbors correspond to top, bridge and hollow sites. Results for equilibration times 0.3, 1 and 3 ns at the end of the convergence at 300 K.

The occupation of the on-top sites found on the facets and edges is in good agreement with recent S-edge XANES experiments,[17] which clearly show an occupation of this site. Here, we show that this occupation is related to a dense assembly such as the 77 SAM, which can be realized only by occupying both on-top and bridge sites. This explains the occupation of the on-top sites even when the bridge sites are preferred, as shown in DFT calculations.[18,19] In contrast, less dense assemblies such as the 33 SAM would imply only an occupation of the bridge site that is not consistent with experiments.

In the gold NP ReaxFF simulations, an on-top site preference was also observed on the facets. In good agreement with experiments for gold NPs.[31] A large fraction of ligands was also in contact with three or even four gold atoms, which is not the case for the simulation obtained here for silver NPs. Please note that, in particular, for the synthesis of NPs, first thiols are physisorbed, which leads to the formation of thiolate with a strong bond between the ligand and the silver atom. What happened with the H atoms of the thiols is still an open question; it may be adsorbed by the metal. In our work, the H atoms are ignored since they are much smaller than the thiolates and may easily fit in the voids between the ligand head groups. We have carried out simulations for the 77 SAM placed on the silver surface H atoms between sulfur atoms. The frequency of SAMs is reduced by about 20%, which shows that these hydrogens may perturb the assembly.

Finally, we study the average distances between neighboring sulfur head groups of the ligands. The S−S distances are shown in Figure 9 for butanethiolate and different NP sizes. The S−S distance is close to 4.4 Å for both the edge and the facet, as expected for a 77 SAM. For NPs larger than 4 nm, this distance does not evolve with the NP size or alkane chain length of the ligand. There is also no significant difference between the edge and the facet. In a previous simulation of gold NPs using nonreactive force fields,[39,40] the S−S distances on the edges were found to be about 0.2 Å smaller than those on the facets, which implies a different assembly on the edges. These results could not be confirmed by recent ReaxFF simulations of gold NPs.

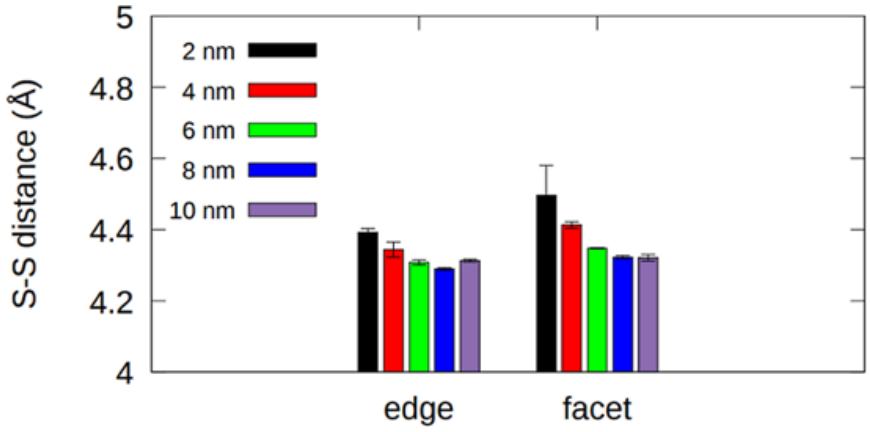


Figure 9: Average S−S distance between first neighbors on the edges and facets of silver NPs from 2 to 10 nm for butanethiolates at the end of convergence at 300 K for an equilibration length of 1 ns.

Here, we show that, also for silver, the S−S distance on the edges is not reduced. For smaller NPs of 2 nm, the average S−S distance increases to 4.5 nm, and the distances on the edges are slightly reduced with respect to the facets.

Conclusions

We have shown here that the ReaxFF approach is able to yield a stable 77 SAM typical of butanethiolates on the Ag (111) surface, demonstrating its effectiveness in simulating and predicting surface adsorption behaviors with high accuracy. It is interesting to note that ReaxFF used in a recent paper to study gold NPs has shown completely different behaviors with the formation of staples on the Au (111) surfaces. The formation of staples is in good agreement with the experiment and the DFT calculation.[19] This shows that ReaxFF yields different properties for both materials in contrast to other force fields. In the recent paper, it was also shown that ReaxFF was capable of reproducing the formation of staples on Au (111) surfaces. Thus, it is possible with ReaxFF to distinguish both materials, which is usually quite difficult for FFs since both metals are characterized by the same geometry. However, we observed that the 77 SAM of methanethiolate is not stable for very long simulations. In addition, the SAM for butanethiolate is not energetically favored with respect to other SAMs. This shows the model's current parametrization limits. Moreover, the reconstructed surface usually observed for silver is made of silver adatoms on an Ag (111) surface. This corresponds to an important restructuring with respect to our systems, which is impossible to see during the ReaxFF simulations. But our simulations do not even show any perturbation of the silver layer in contact with the thiolates. This indicates the good stability of these unrestructured SAMs. It would be, however, interesting in the future to study already restructured silver surfaces with thiolates; the idea would be to compare the binding energies of these systems with the unrestructured ones obtained here. This would allow us to see which one is energetically preferred and if restructuration can be experimentally expected.

The simulations here also show that dense assemblies such as the 77 SAMs found on the Ag (111) surface should appear on silver NPs. To the best of our knowledge, there are no available experiments studying this question. During the simulations, no decomposition of the alkanethiolate was found. The origins of the silver sulfide observed in some experiments are an open question.[11,12,17] In the future, this question should also be investigated for restructured surfaces and other ligands.

Also, no spontaneous restructuring of the NP surfaces was found, except for the small NPs inferior to 4 nm. A global restructuring of the silver surfaces, as observed in experiments, is not attainable

with ReaxFF simulations. In the future, we plan to perform simulations starting from a restructured surface and compare its binding energies with those found here.

Two adsorption sites on silver NP facets were found: one on- top and another at the bridge site, where the former is preferred: this is in good agreement with recent experiments.

Regarding the distance between S-headgroup ligands, we observed no difference between the facets and the edges.

B.4) ReaxFF Simulations of Self-Assembled Monolayers on Gold Surfaces and Clusters

The same approaches used for silver have been tested on gold. Three different potentials were used (previously described in section B.3): Agren et al., Aikens et al., and VanDuin et al. The binding energy was then plotted to observe how the assembly evolved during the simulation, as depicted in Figure B-17. We found that for the Aiken and Van Duin potentials, the simulations did not converge, stopping at 25 K and 150 K, respectively. Furthermore, direct visualization of the surface for both potentials revealed total deconstruction, with some atoms leaving the surface. However, with Agren's potential, the binding energy evolved up to 300 K, and the surface appeared stable.

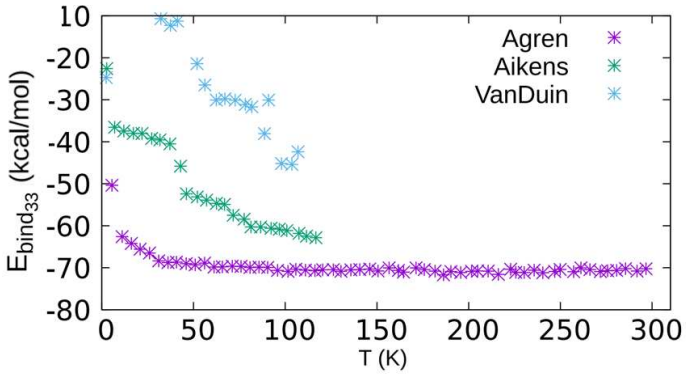
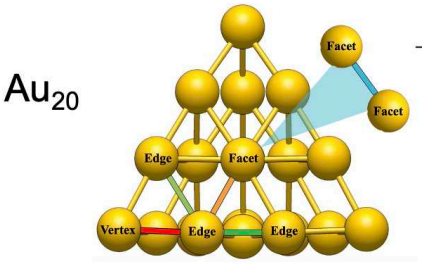


Figure B-17: Binding energy for 33 methanethiolates-SAM on Au (111) from 0 K to 300 K using three different potentials.

We wanted to see then if the Agren Force Fields was able to discriminate between gold atoms in the popular pyramid-like structure. By optimizing with Agren's potential, we performed a radial analysis on an Ag_{20} pyramid like structure and we saw that compared to DFT the distances

are equivalents. Meaning that we lack of accuracy when regarding the precise interactions of Au-Au bonds.



Distance in Å	ReaxFF	DFT
Vertex-edge	2.88	2.76
Edge-facet	2.88	2.86
Facet-Facet	2.88	3.09
Edge-Edge	2.88	2.71
Edge-Edge facet	2.88	3.00

Figure 15: Distances of Au-Au atoms for an optimized structure from ReaxFF and DFT.

From that a question emerged for silver . Is the current force fields made by Dulong et al., is capable of predicting accurate distances on small and large clusters ?

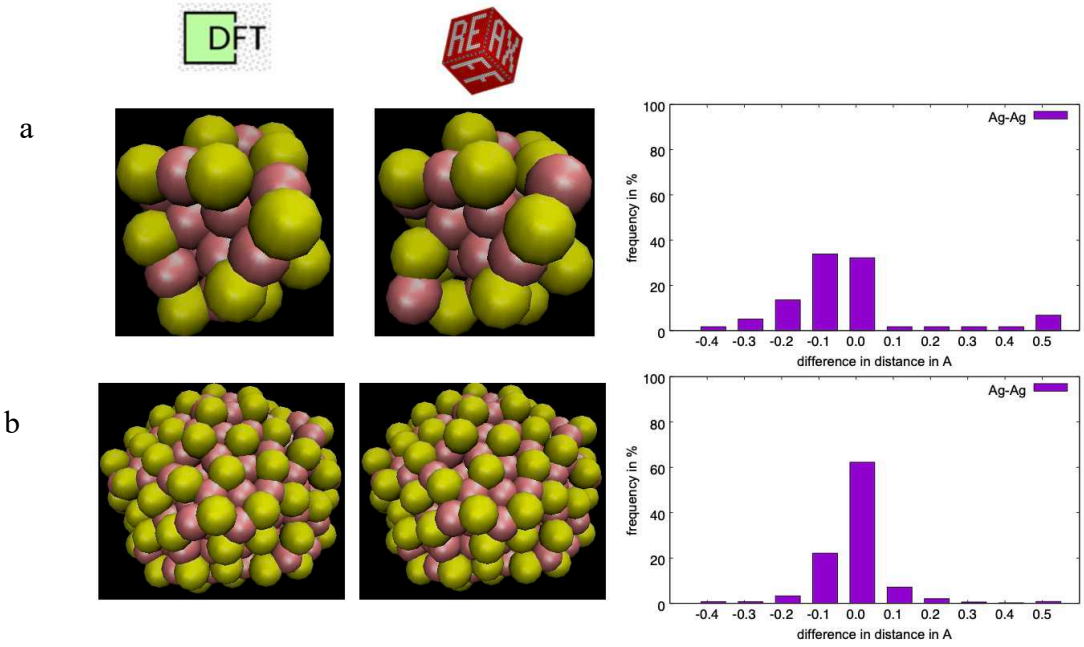


Figure B-18: Comparison of Ag-Ag distances between DFT and ReaxFF for (a) Ag25 and (b) Ag374 protected with akylthiolates. For the sake of comprehension, only silver and sulfur atoms are represented by the colors red and yellow, respectively.

As shown in Figure B-18b, for large clusters, the difference in distances between DFT and ReaxFF is quite small, with a high frequency at 0 Å. The deviation is more noticeable for Ag25,

depicted in Figure B-18a, where we have a large proportion, which is shifted to 0.1 Å, indicating less accuracy in describing the expected distances.

This implies that the current parameterization of all the potentials created for gold or silver do not accurately predict the stability of different assemblies between the 33 and the 77, nor does it accurately describe each Ag-Ag distance. Therefore, new approaches are needed to achieve accurate descriptions for both small and large clusters and different surfaces.

References

(1) Morones, J. R.; Elechiguerra, J. L.; Camacho, A.; Holt, K.; Kouri, J. B.; Ramirez, J. T.; Yacaman, M. J. The Bactericidal Effect of Silver Nanoparticles. *Nanotechnology* **2005**, *16*, 2346.

(2) Jain, P. K.; Huang, X.; El-Sayed, I. H.; El-Sayed, M. A. Noble Metals on the Nanoscale: Optical and Photothermal Properties and Some Applications in Imaging, Sensing, Biology, and Medicine. *Acc. Chem. Res.* **2008**, *41*, 1578.

(3) Le Ouay, B.; Stellacci, F. Antibacterial Activity of Silver Nanoparticles: A Surface Science Insight. *Nano Today* **2015**, *10*, 339.

(4) Alhilaly, M. J.; Bootharaju, M. S.; Joshi, C. P.; Besong, T. M.; Emwas, A.-H.; Juarez-Mosqueda, R.; Kaappa, S.; Malola, S.; Adil, K.; Shkurenko, A. [Ag 67 (SPhMe 2) 32 (PPh 3) 8] 3+ : Synthesis, Total Structure, and Optical Properties of a Large Box-Shaped Silver Nanocluster. *J. Am. Chem. Soc.* **2016**, *138* (44), 14727.

(5) Tasaka, Y.; Nakamura, K.; Malola, S.; Hirata, K.; Kim, K.; Koyasu, K.; Häkkinen, H.; Tsukuda, T. Electron Binding in a Superatom with a Repulsive Coulomb Barrier: The Case of $[Ag_{44}(SC_6H_3F_2)_{30}]^{4-}$ in the Gas Phase. *J. Phys. Chem. Lett.* **2020**, *11* (8), 3069–3074. https://doi.org/10.1021/acs.jpclett.0c00786.

(6) Torres, D.; Carro, P.; Salvarezza, R. C.; Illas, F. Evidence for the Formation of Different Energetically Similar Atomic Structures in Ag (111) - (7 × 7)-R 19.1° - CH 3 S. *Phys. Rev. Lett.* **2006**, *97* (22), 226103.

(7) Woodruff, D. P. The Interface Structure of N-Alkylthiolate Self-Assembled Monolayers on Coinage Metal Surfaces. *Phys. Chem. Chem. Phys.* **2008**, *10*, 7211.

(8) Unertl, W. N. *Handbook of Surface Science*; 1996; Vol. 1.

(9) Grönbeck, H.; Häkkinen, H.; Whetten, R. L. Gold–Thiolate Complexes Form a Unique c(4 × 2) Structure on Au(111). *J. Phys. Chem. C* **2008**, *112*, 15940.

(10) Voznyy, O.; Dubowski, J. J.; Yates, J. T.; Maksymovych, P. The Role of Gold Adatoms and Stereochemistry in Self-Assembly of Methylthiolate on Au(111). *J. Am. Chem. Soc.* **2009**, *131* (36), 12989.

(11) Padmos, J. D.; Zhang, P. Surface Structure of Organosulfur Stabilized Silver Nanoparticles Studied with X-Ray Absorption Spectroscopy. *J. Phys. Chem. C* **2012**, *116* (43), 23094.

(12) Battocchio, C.; Meneghini, C.; Fratoddi, I.; Venditti, I.; Russo, M. V.; Aquilanti, G.; Maurizio, C.; Bondino, F.; Matassa, R.; Rossi, M. Silver Nanoparticles Stabilized with Thiols: A Close Look at the Local Chemistry and Chemical Structure. *J. Phys. Chem. C* **2012**, *116* (36), 19571.

(13) Farrag, M.; Thämer, M.; Tschurl, M.; Bürgi, T.; Heiz, U. Preparation and Spectroscopic Properties of Monolayer-Protected Silver Nanoclusters. *J. Phys. Chem. C* **2012**, *116* (14), 8034.

(14) Maya Girón, J. V.; Zelaya, E.; Rubert, A.; Benítez, G.; Carro, P.; Salvarezza, R. C.; Vela, M. E. Surface Chemistry of 4-Mercaptobenzoic Acid Self-Assembled on Ag(111) and Ag Nanoparticles. *J. Phys. Chem. C* **2013**, *117* (47), 24967.

(15) Azcárate, J. C.; Corthey, G.; Pensa, E.; Vericat, C.; Fonticelli, M. H.; Salvarezza, R. C.; Carro, P. Understanding the Surface Chemistry of Thiolate-Protected Metallic Nanoparticles. *J. Phys. Chem. Lett.* **2013**, *4* (18), 3127.

(16) Vericat, C.; Vela, M. E.; Corthey, G.; Pensa, E.; Cortés, E.; Fonticelli, M. H.; Ibañez, F.; Benitez, G. E.; Carro, P.; Salvarezza, R. C. Self-Assembled Monolayers of Thiolates on Metals: A Review Article on Sulfur-Metal Chemistry and Surface Structures. *RSC Adv.* **2014**, *4* (53), 27730.

(17) Marchioni, M.; Battocchio, C.; Joly, Y.; Gateau, C.; Nappini, S.; Pis, I.; Delangle, P.; Michaud-Soret, I.; Deniaud, A.; Veronesi, G. Thiolate-Capped Silver Nanoparticles: Discerning Direct Grafting from Sulfidation at the Metal-Ligand Interface by Interrogating the Sulfur Atom. *J. Phys. Chem. C* **2020**, *124* (24), 13467.

(18) Li, A.; Piquemal, J. P.; Richardi, J.; Calatayud, M. Butanethiol Adsorption and Dissociation on Ag (111): A Periodic DFT Study. *Surf. Sci.* **2016**, *646*, 247.

(19) Becerril, D.; Noguez, C. Adsorption of a Methylthio Radical on Silver Nanoparticles: Size Dependence. *J. Phys. Chem. C* **2015**, *119* (20), 10824.

(20) Dulong, C.; Madebene, B.; Monti, S.; Richardi, J. Optimization of a New Reactive Force Field for Silver-Based Materials. *J. Chem. Theory Comput.* **2020**, *16* (11), 7089.

(21) Hazarika, Z.; Jha, A. N. Computational Analysis of the Silver Nanoparticle-Human Serum Albumin Complex. *ACS Omega* **2020**, *5* (1), 170.

(22) Yoneya, M.; Sugisawa, S. Simulation of Colloidal Silver Nanoparticle Formation from a Precursor Complex. *J. Phys. Chem. C* **2019**, *123* (17), 11257.

(23) Jiang, C.; Mo, Y.; Wang, H.; Li, R.; Huang, M.; Jiang, S. Molecular Dynamics Simulation of the Production of Hollow Silver Nanoparticles under Ultrafast Laser Irradiation. *Comput. Mater. Sci.* **2021**, *196*, 110545.

(24) Gu, M.; Liu, T.; Xiao, X.; Li, G.; Liao, W. Simulation and Experimental Study of the Multisized Silver Nanoparticles Sintering Process Based on Molecular Dynamics. *Nanomaterials* **2022**, *12* (6), 1030.

(25) Alexiadis, O.; Harmandaris, V. A.; Mavrantzas, V. G.; Site, L. D. Atomistic Simulation of Alkanethiol Self-Assembled Monolayers on Different Metal Surfaces via a Quantum, First-Principles Parametrization of the Sulfur–Metal Interaction. *J. Phys. Chem. C* **2007**, *111*, 6380.

(26) Djebaili, T.; Abel, S.; Marchi, M.; Richardi, J. Influence of Force-Field Parameters on the Atomistic Simulations of Metallic Surfaces and Nanoparticles. *Phys. Chem. C* **2017**, *121*, 27758.

(27) Bae, G.-T.; Aikens, C. M. Improved ReaxFF Force Field Parameters for Au–S–C–H Systems. *J. Phys. Chem. A* **2013**, *117* (40), 10438–10446. https://doi.org/10.1021/jp405992m.

(28) Monti, S.; Carravetta, V.; Ågren, H. Simulation of Gold Functionalization with Cysteine by Reactive Molecular Dynamics. *J. Phys. Chem. Lett.* **2016**, *7* (2), 272–276. https://doi.org/10.1021/acs.jpclett.5b02769.

(29) Zhang, X.-Q.; Iype, E.; Nedea, S. V.; Jansen, A. P. J.; Szyja, B. M.; Hensen, E. J. M.; van Santen, R. A. Site Stability on Cobalt Nanoparticles: A Molecular Dynamics ReaxFF Reactive Force Field Study. *Phys. Chem. C* **2014**, *118*, 6882.

(30) Keith, J. A.; Fantauzzi, D.; Jacob, T.; Van Duin, A. C. Reactive Forcefield for Simulating Gold Surfaces and Nanoparticles. *Phys. Rev. B* **2010**, *81*, 235404.

(31) Richardi, J.; Fadigas, M. ReaxFF Molecular Dynamics Simulations of Large Gold Nanocrystals. *J. Chem. Theory Comput.* **2022**, *18* (4), 2521.

(32) Xia, Y.; Xiong, Y.; Lim, B.; Skrabalak, S. E. Shape-Controlled Synthesis of Metal Nanocrystals: Simple Chemistry Meets Complex Physics? *Angew. Chem., Int. Ed.* **2009**, *48*, 60.

(33) van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396.

(34) Chenoweth, K.; van Duin, A. C.; Goddard, W. A. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A* **2008**, *112*, 1040.

(35) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108*, 4315.

(36) Janssens, G. O.; Baekelandt, B. G.; Toufar, H.; Mortier, W. J.; Schoonheydt, R. A. Comparison of Cluster and Infinite Crystal Calculations on Zeolites with the Electronegativity Equalization Method (EEM). *J. Phys. Chem.* **1995**, *99*, 3251.

(37) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1.

(38) Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. Parallel Reactive Molecular Dynamics: Numerical Methods and Algorithmic Techniques. *Parallel Comput.* **2012**, *38*, 245.

(39) Djebaili, T.; Richardi, J.; Abel, S.; Marchi, M. Atomistic Simulations of the Surface Coverage of Large Gold Nanocrystals. *J. Phys. Chem. C* **2013**, *117*, 17791.

(40) Djebaili, T.; Richardi, J.; Abel, S.; Marchi, M. Atomistic Simulations of Self-Assembled Monolayers on Octahedral and Cubic Gold Nanocrystals. *J. Phys. Chem. C* **2015**, *119*, 21146.

C) Study of gold nanoparticles using machine learning potential

C.1) Introduction

As shown in the previous chapters, molecular dynamics simulations with reactive force fields do not accurately describe the phenomena that can happen when thiolates react with the gold (or silver) surfaces. Even if they are parametrized using ab initio reference calculations and they demand less computational resources than full quantum calculations, they appear not to be suited for these systems. The main reasons for these limitations are based on the definition of the methods themselves. Indeed, the mathematical functions used as functional forms for force fields often include drastic physics simplifications, leading to inaccuracies in describing complex molecular systems.[1,2] To be more specific, ReaxFF struggles with model systems with complex interactions. For gold, it accurately predicts atomic diffusion, but when interactions with the thiolates are added, it struggles with surface reconstruction effects, as shown by the test performed using the three known potentials in the previous chapter. These limitations could also explain the non-sulfidation of the silver nanoparticles. Therefore, new methods should be used since accuracy has always been a more general problem when dealing with force fields. In that context, the emergence of deep learning approaches can serve as a new strategy to provide a more efficient way to describe these specific systems.

Artificial intelligence is one of the most popular tools and can be used for a lot of things, including writing, creating images, videos, music, driving, and even talking. Among all its available tools, machine learning (ML) can be used for science, as illustrated by Figure C-1. In comparison to classical programming, ML approaches are trained based on statistical structures, and thanks to the large amount of available data that exists across all domains, one can easily train its own ML potential.

It is then possible to create a flexible model which can capture complex interactions. In practice, we will use Neural Network interatomic potentials whose parameters will be learned from ab initio calculations to predict complex interactions. For example, they should be well suited for the study of our gold systems that exhibit reactivity effects, such as for the formation of staples. Different potentials have already been created based on small gold clusters[3] and gold-protected clusters with sulfur[4] , but our goal here is to create a general model that can predict from small clusters to large nanoparticles capped or not with thiolates-SAMs. Over the years, different types

of ML models have been introduced. Among them, we can cite the kernel method in the 1990's based on decision boundaries between two sets of data that belong to different categories. It gives a linear regression in a non-linear space defined by a similarity Kernel over data points (or descriptors of data points).[5] Such an approach is efficient in lattice dynamics simulations for calculating thermodynamics properties reproducing quantum data with high accuracy. Nevertheless, it involves high computational complexity with larger datasets, and the kernel selection requires a lot of trial and error, which can affect the efficiency of the learning.[6] This is one of the reasons why people are also using Neural-Network (NN) based ML methods. They are built as follows:

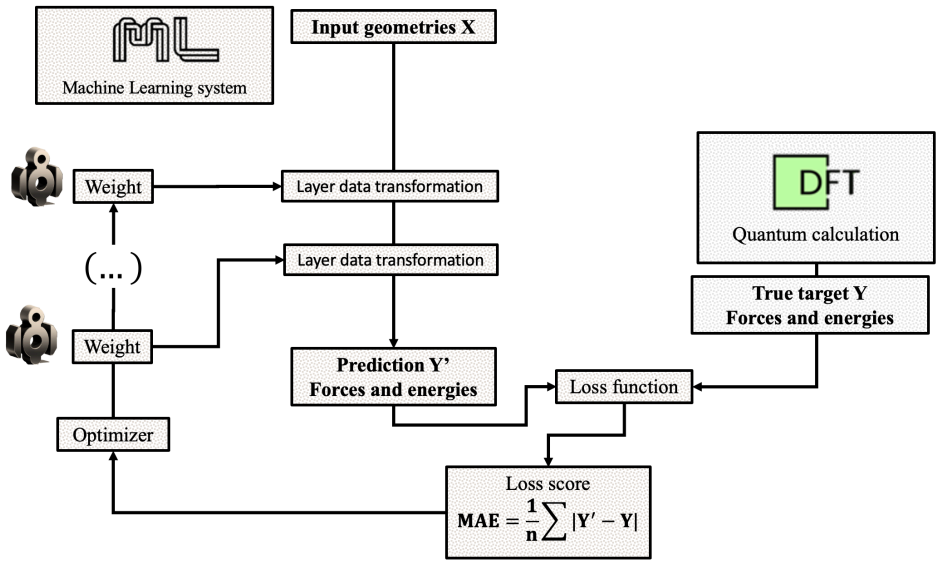


Figure C-16: Schematic of the machine learning process to create the desired potential depending on reference DFT calculations.

So, as represented in Figure C-1, the objective of the ML approach is to fit the model to the true target Y, which can originate from DFT calculations, for example, including energies and forces. Indeed, it is going to use the input geometries to compute its prediction by successive data transformations and then, from the loss function, consider if the weight of each layer needs to be optimized. In the next section, we are going to focus more on the Neural Network "black box," which allows us to translate the data to predict the right values.

Neural network architecture NN can be used to process a lot of data following the procedure represented in Figure C-2. Each layer is composed of a specific number of neurons, which can

vary depending on the complexity of the task. A neuron (X) is a fundamental computational unit that receives inputs from external sources or other neurons. Each neuron has weights (W) and a bias (b) that are optimized during the training. Then take place the activation function(f) which is used to reach the specification problem that the neuron is trying to solve. It is used to introduce a non-linearity (log), as represented in Figure C-2. The schematic of Figure C-2 displays how a single neuron is defined.

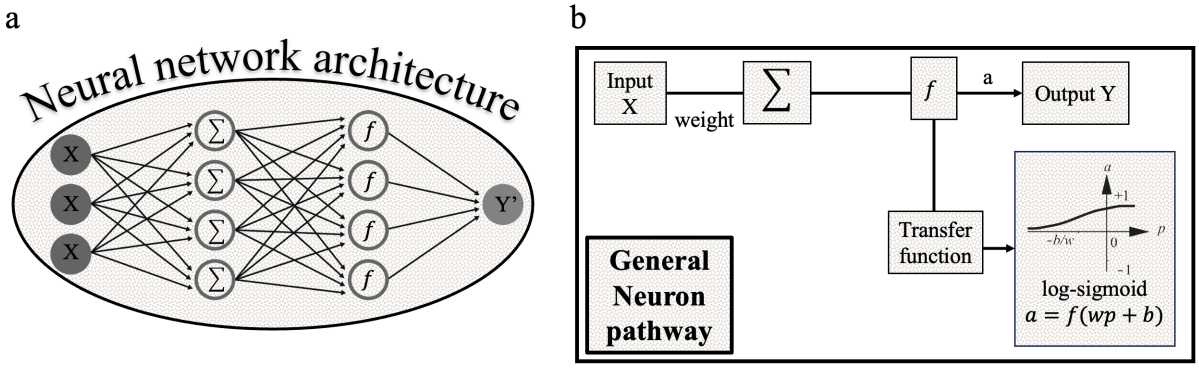


Figure C-2: (a) Simple Neural Network architecture for an arbitrary number of inputs. (b) General neuron pathway with arbitrarily chosen log-loss functions. Adapted from Neural Network Design Book from Martin T. Haga[7]

The input X_i goes into the neuron, and then a summation is operated as defined in Equation C-1 to gather all the information from the input as the weight W_i of each neuron adds the bias $b.$,

$$z = \sum_{i=1}^{n} W_i X_i + b \tag{C-1}$$

Where z represents the total input to the neuron from all incoming connections, where n is the number of inputs to the neuron. Then, it passes through a transfer function, which processes the scalar neuron output. Each neuron in the layer is connected to all inputs, and the output of one layer can be used as input for the next layer. Therefore, more than one layer can be used to capture complex patterns present within the data. In practice, a single-layer NN can only learn basic interactions and one to keep in mind the accuracy problem of FF potentials. However, higher layers can build upon the first layer, which determines only "simple features" to be refined and determined in more detail using the different features and their links between the different layers. Deeper networks, therefore, increase their "expressiveness," which means that sufficient data and

computational resources are provided. From an overall point of view, it improves learning and allows for the generalization of unseen data. Figure C-3 shows a schematic of a multiple-layer NN.

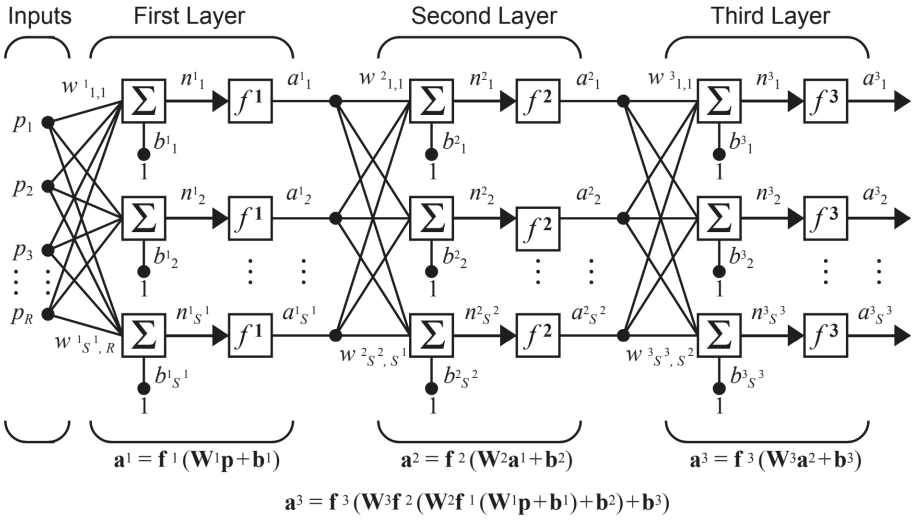


Figure C-3: Schematic of a three-layer neural network. Each one of them contains basic neural network parameters (activation function, weights, bias, etc..). Adapted from Neural Network Design Book from Martin T. Haga[7]

C.2) Machine Learning for Interatomic Potential

Neural network-based ML has significantly increased the development of interatomic potentials. Indeed, it enables atomistic simulations at ab initio level accuracy without requiring too much resources. So far, different methods have been developed:

- Behler-Parinello introduces a neural network approach for representing high-dimensional potential energy surfaces (PES) that significantly speeds up calculations while maintaining ab initio accuracy.[8] Their method involves a generalized neural network structure where the total energy is a sum of atomic contributions learned through DFT-based training data. The network uses symmetry functions to describe atomic environments invariantly, allowing accurate predictions across varying system sizes and conditions.
- SE(3)-Equivariant Graph Neural Networks, which handle geometric tensors, provide a better description of atomic environments.[9] It can learn from less data (compared to other NN approaches) while achieving high accuracy. Some models have already been created based on this method, like Neural Equivariant

Interatomic Potentials (NequIP)[9] or the Equivariant Hierarchy-based Graph Networks (EGHNs)[10].

In this Thesis, since Graph Neural Networks (GNN) have been used, we are going to define them in the next section and explain how they can be handled with ease in the context of the Force-field-enhanced Neural Network Potentials (FeNNol) Python module. Within FeNNol, the first module that has been created is the module including the Force-Field-Enhanced Neural Network InteraXtions (FENNIX) framework.

Overview of the FENNIX Framework

FeNNix is a framework that can be used to build a force-field-enhanced ML model.[11] It is a hybrid model combining machine learning and classical force fields to predict local energy contributions. The inclusion of long-range electrostatics and dispersion effects is also one of the main motivations for the creation of the framework. The ML model uses equivariant neural networks to predict the local properties of molecules. It is based on the Allegro model, which makes it strictly local and allows for favorable computational scaling with the size of the system.[12] Moreover, the equivariance of Allegro allows for the prediction of tensorial properties such as atomic multipoles. More information concerning Allegro can be found in the article by Muesaelian et al.[12], where the authors explain in more detail its architecture, which utilizes tensor products of learned equivariant representations to handle many-body potentials. They were able to obtain superior accuracy and scalability across various benchmarks as Allegro allowed them to predict out-of-distribution data by recovering the kinetic properties of complex materials. It also provides an accurate method for simulating atomistic dynamics with large systems and for long timescales.

The workflow of the Allegro model can be given as follows. First off, each atom at its environment and neighboring $N(i)$ as defined in equation C-2:

$$N(i) = \left\{ j \text{ s.t.} \left\| \overrightarrow{R_{ij}} \right\| < r_c \right\} \qquad \text{(C-2)}$$

Where j corresponds to the indices of atoms that are considered neighbors of atom i such that (s.t.) the distance between atom i to atom j, presented by $\left\| \overrightarrow{R_{ij}} \right\|$, must be less than the cutoff radius r_c. The cutoff defines the neighborhood around atom i within which other atoms are considered to be neighbors. Then, we need to model the interactions between pairs of atoms.

1. **Two-body (2B) Feature using a Multi-Layer Perceptron (MLP)**

   To model these interactions, we employ a two-body descriptor $\left(x_{ij}, V_{ij}^{nlp}\right)$ for each pair of atoms. The first component of this descriptor, x_{ij}^{2B} is derived from a multilayer perceptron (MLP). An MLP is a type of neural network that includes one input layer, one or more hidden layers, and an output layer. Each layer is fully connected to the next, allowing the network to capture complex relationships in the data.

- First, a chemical species encoding is performed. Each atom's chemical species is encoded using a one-hot encoding method, denoted as $I(Z_i)$ for atom i and $I\left(Z_j\right)$ for atom j. One-hot encoding transforms categorical data into a binary vector. For example, if there are five possible types, and if atom i is of type 1, it would be represented as [1, 0, 0, 0, 0]. This study, which was based on the periodic table, represented each element using sine and cosine functions, reflecting the similarities and periodic trends among elements. The row and column indices of each element in the periodic table are coded separately. As shown in Figure C-4, the first part corresponds to the row and the other to the column. For the row, we then have a difference between the hydrogen and the other atoms that are from the same period (2nd) in the periodic table. In comparison, for the column, as shown on the right in Figure C-4, we have the same values for silver and gold because they are from the same group (11th). This type of coding allows good generalization and similarity inference in the ML model.

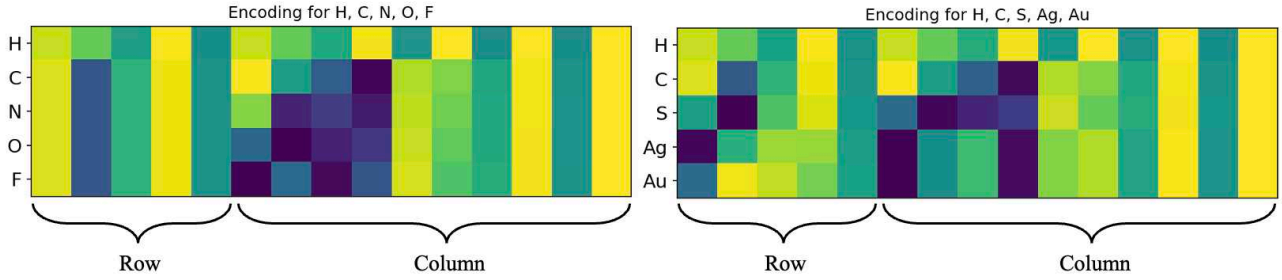


Figure C-4: Positional encoding of different chemical species. (a) is one simple example with atoms in the same row (b) is the chemical species that are studied in this thesis. Adapted from Plé et al. (**2023**)[11]

In this thesis, the encoding of molecular systems is fundamentally based on their electronic structures, using a model known as SpookyNet. This model uses a detailed encoding scheme that incorporates several key aspects of each element within the system. Firstly, it uses the

atomic number, which provides a fundamental identifier that is intrinsic to the chemical properties and behavior of the element. In addition, the model encodes the total number of electrons, which is crucial for representing the electronic state of the element and its potential interactions within molecules. Furthermore, SpookyNet explicitly includes the spin state of the electrons, an essential quantum mechanical property that influences magnetic interactions and the overall stability of molecular structures. This inclusion is particularly significant as it allows SpookyNet to account for the diversity of molecular states, enabling it to predict properties across different electronic configurations with greater accuracy.

- Then, when it is done, we need to represent the interatomic distances. The distance between atoms i and j, R_{ij}, is represented using a Bessel basis function, noted $B(R_{ij})$. This function projects the interatomic distance onto a basis that enhances the model's ability to discern subtle variations of such quantity. The Bessel functions are particularly useful for representing wave-like properties in space and are defined as $B(R_{ij}) = [B_1(R_{ij}), \dots, B_{Nbasis}(R_{ij})]$.
- The cutoff function, $f_c(R_{ij})$, ensures that the influence of distant atoms diminishes smoothly as their separation approaches a predefined cutoff distance, r_c. This function is crucial for maintaining computational efficiency (i.e., it reduces the number of operations to be performed) while focusing on locally relevant interactions.

By using an MLP to process these inputs, we can generate a sophisticated description of atomic interactions, which are then utilized in molecular dynamics simulations to predict material properties more accurately. This method allows for flexibility in handling various atomic species and interaction distances, enhancing the predictive capabilities of computational models. x_{ij}^{2B} is then defined in equation C-3 as follows:

$$x_{ij}^{2B} = \mathrm{MLP}_{2B}[I(Z_i) \parallel I(Z_j) \parallel B(R_{ij})] f_c(R_{ij}) \qquad \text{(C-3)}$$

Where it is a concatenation ‖ (joining operation) between the electronic information from atom i ($I(Z_i)$) and j ($I(Z_j)$) merging both information into our two-body (2B) descriptor, then the projected radial basis $B(R_{ij})$ composed of Bessel basis function is concatenated. From that, we get a descriptor, which is composed of all the chemical environments between atoms i and j.

$f_c(R_{ij})$ corresponds to the function which is going smoothly towards zero when approaching the distance cutoff r_c.

1. **The local environment interaction**

   The second component of the two-body descriptor $V_{ij}^{nlp,2B}$ uses the two-body scalar embedding where the two-body equivariant features are projected onto a basis of real spherical harmonics Y_{ij}^{lp} and defined as:

$$V_{ij}^{nlp,2B} = \left[\mathrm{MLP}_{\mathrm{embed}}^{2B}\left(x_{ij}^{2B}\right)\right]^{nlp} Y_{ij}^{lp} \qquad \text{(C-4)}$$

It is composed of indices n, l, and p that maintain the equivariance concerning the symmetries. The use of equivariance is important as it demonstrates accurate descriptions of structural and kinetic properties for complex materials. It is defined as a tensor labeled with a channel index n, a rotational index l, and a parity index p. The rotational index can have different values. $l = 0$ corresponds to scalar/invariant quantities, $l = 1$ to vector-like objects and is $l \geq 2$ to objects encompassing higher-order tensors. The tensor product is calculated using the $e3nn$ model,[13] given the equivariant multiplication operation for two representations. They have the particularity to satisfy two conditions :(i) its binarity and (ii) and its equivariance. To give an example, let's consider the interaction between atoms i and j. Atoms i have their environment $N(i)$ defined by the cutoff radius and all its chemical information is stored in a vector called $\mathrm{I}(\mathrm{Z_i})$. It is seeing the atom j from a distance $\left\|\overrightarrow{R_{IJ}}\right\|$ which also has its chemical information stored in the vector $\mathrm{I}(\mathrm{Z_j})$. The embedding is then processed by first creating a global equivariant environment neighborhood composed of the two descriptors previously described and of a new feature, L, which is defined as the layer index. The equation becomes:

$$\Gamma_i^{nlp,(L)} = \sum_{j \in N(i)} \left[\mathrm{MLP}_{\mathrm{embed}}^{(L)}\left(x_{ij}^{(L-1)}\right)\right]^{nlp} Y_{ij}^{lp} \qquad \text{(C-5)}$$

The interaction is then processed using a tensor product creating a "latent space" composed of all possible combinations of rotational l and parity indices p that are allowed by symmetry. m

is a multiplicity index that must be integrated as it helps distinguish between between $(l_1, p_1) \wedge (l_2, p_2)$ possible combinations that produce (l, p). The latent space is defined as:

$$L_{ij}^{nmlp,(L)} = \left(\Gamma_i^{nl_1p_1,(L)} \otimes V_{ij}^{nl_2p_2,(L)}\right)^{nmlp} \tag{C-6}$$

Where a tensor product $\otimes$ is applied between the local $\Gamma_i^{nl_1p_1,(L)}$ environment of the atom i and the $V_{ij}^{nl_2p_2,(L)}$ interaction descriptor between atoms i and j. It contains all the combinations of rotational l and parity p that are allowed by symmetry. Then, to get the new pairwise embedding, the scalar product is redefined by:

$$x_{ij}^{(L)} = \alpha x_{ij}^{(L-1)} + \sqrt{1-\alpha^2} f_c(R_{ij}) \times \mathrm{MLP}_{\mathrm{latent}}^{(L)}\left[x_{ij}^{(L-1)} \parallel L_{ij}^{nm01,(L)}\right]_{n,m} \tag{C-7}$$

Where the α coefficient is a value ranging between 0 and 1 and allowing to easily propagate scalar information from the $L-1$ layer to the L layer assuring communication between them. It can either be changed for each layer or set to a specific value. Then, the equivariant features are obtained by a linear combination of the element of the latent space with the same indices l and p from all channels and multiplicities.

$$V_{ij}^{nlp,(L)} = \sum_{n',m} w_{n',m}^{nlp,(L)} L_{ij}^{n'mlp,(L)} \tag{C-8}$$

Where $w_{n',m}^{nlp,(L)}$ is the weights that are optimized in the training procedure.

2. **The energy contributions**

The contributions of the total energy are defined by the product of the three different energies:

$$E^{\mathrm{OP1}} = \sum_i E_i^{\mathrm{NN}} + \sum_{i,j<i} E_{ij}^{\mathrm{C}} + \sum_{i,j<i} E_{ij}^{\mathrm{D}} \tag{C-9}$$

Where E_i^{NN} is the neural network contribution determined by processing the many-body descriptors previously described. It provides a local energy contribution to the short-range interactions, which remains limited by the rc cutoff radius. The neural network also provides atomic partial charges q_i^{NN} and atomic volumes v_i^{NN}. Then E_{ij}^{C} is the electrostatic interaction

computed via a Coulomb potential. It depends on the number of valence electrons for both atoms i and j, respectively N_i and N_j and their partial charges, q_i and q_j , predicted by the neural network.

$$E_{ij}^{\mathrm{C}} = \frac{1}{R_{ij}}\left[N_i N_j + N_j(q_i - N_i) f_\alpha\left(\frac{R_{ij}}{r_i^{\mathrm{vdw}}}\right) + N_i\left(q_j - N_j\right) f_\alpha\left(\frac{R_{ij}}{r_j^{\mathrm{vdw}}}\right) + (q_i - N_i)\left(q_j - N_j\right) f_\beta\left(\frac{R_{ij}}{r_i^{\mathrm{vdw}}}\right) f_\beta\left(\frac{R_{ij}}{r_j^{\mathrm{vdw}}}\right)\right] \quad \text{(C-10)}$$

Where f_α and f_β are damping functions with α and β adjustable parameters that are assumed to be universal. $r_{i,j}^{\mathrm{vdw}}$ are the environment-dependent van der Waals (vDW) radii of each atom i and j. They are defined as the ratio between the atomic volume predicted by the neural network v_i^{NN} and the volume of the atom without any interaction (v_i^{free}) which is defined as:

$$r_i^{vdw} = \left(\frac{v_i^{\mathrm{NN}}}{v_i^{\mathrm{free}}}\right)^{\frac{1}{3}} r_i^{vdw,free} \quad \text{(C-11)}$$

The final energetic contribution is the dispersion interaction E_{ij}^{D} uses of pairwise Tkatchenko-Scheffler[14] model:

$$E_{ij}^{\mathrm{D}} = \frac{-2C_{6,i}C_{6,j}}{\left(C_{6,i}\frac{\alpha_j}{\alpha_i} + C_{6,j}\frac{\alpha_i}{\alpha_j}\right)R_{ij}^6}\sigma_{ij}\left(R_{ij}\right) \quad \text{(C-12)}$$

where C_6 is a coefficient derived from the electron density using accurate free-atom reference data, which is used to determine the vdW interactions.[11] α_i, $\alpha_j$and $\sigma_{ij}\left(R_{ij}\right)$ are the isolated atom polarizabilities and the sigmoid damping function respectively:

$$C_{6,i} = \left(\frac{v_i^{\mathrm{NN}}}{v_i^{\mathrm{free}}}\right)^2 C_{6,i}^{\mathrm{free}}, \alpha_i = \left(\frac{v_i^{\mathrm{NN}}}{v_i^{\mathrm{free}}}\right)\alpha_i^{\mathrm{free}} \quad \text{(C-13)}$$

$$\sigma_{ij}(R_{ij}) = \left[1 + e^{-\gamma\left(\frac{R_{ij}}{sr_i^{\text{vdw}} + r_j^{\text{vdw}}} - 1\right)}\right]^{-1} \quad \text{(C-14)}$$

Where γ and s are adjustable parameters that are assumed universal.

To summarize, the FeNNix framework utilizes the Allegro model to process chemical embeddings, applying symmetry operations equivariantly across all elements. It allows for accurate local environments and keeps track consistently of chemical properties. Allegro's innovative approach not only ensures that symmetry operations are applied uniformly but also integrates a mechanism to adjust the weights of these operations during the neural network training phase, enhancing the model's adaptability and accuracy. Once the embeddings are processed, the framework employs a local many-body descriptor to calculate the interaction energies between atoms. The FeNNix model can then determine the different energetics contributions from short- to long-range within molecular systems. It is a robust physical model to predict and analyze molecular energies with high precision. The schematic of the procedure is presented in Figure C-5.

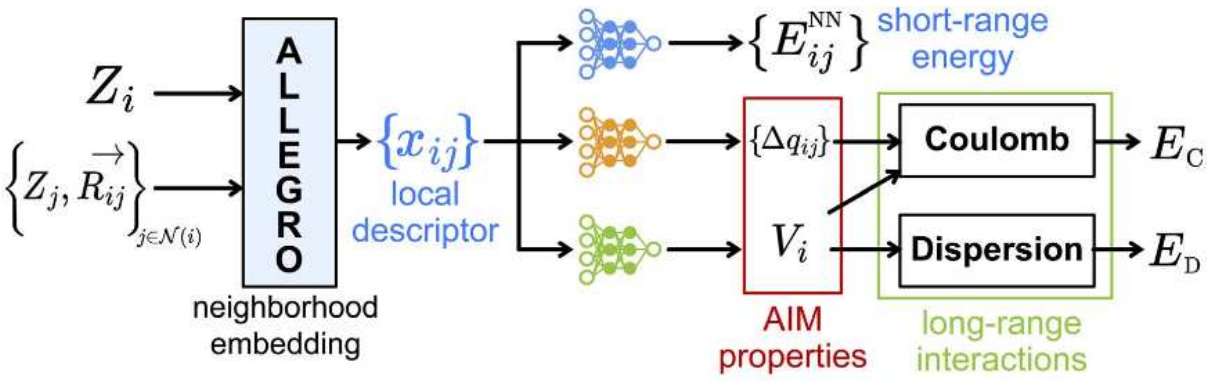


Figure C-5: Procedure describing the FeNNix method. Adapted from Plé et al. (**2023**)[11]

Training modules

FeNNix's training process is designed to mirror DFT-derived total energies E^{DFT} and forces F_{ij}^{DFT}. It forms the basis for a loss function that quantifies differences between the model's predictions and the DFT benchmarks. Such quantification is essential for adjusting the model's weights, thus tuning its predictions. Specifically, the loss function—typically a mean squared error for both energies and forces—guides systematic weight adjustments to enhance accuracy and reliability in predicting chemical properties. The training comprises four stages, starting with this foundational alignment defined as

$$L^{(1)} = \lambda_E (E^{\mathrm{DFT}} - E^{\mathrm{NN}})^2 \qquad \text{(C-15)}$$

$$+\lambda_F \sum_{j=1}^{3} \sum_{i=1}^{N_{\mathrm{at}}} \left(F_{ij}^{\mathrm{DFT}} - F_{ij}^{\mathrm{NN}}\right)^2$$

$$+\lambda_q \sum_{i=1}^{N_{\mathrm{at}}} \left(q_i^{\mathrm{DFT}} - q_i^{\mathrm{NN}}\right)^2$$

$$+\lambda_v \sum_{i=1}^{N_{\mathrm{at}}} \left(\frac{v_i^{DFT}}{v_i^{\mathrm{free}}} - \frac{v_i^{\mathrm{NN}}}{v_i^{\mathrm{free}}}\right)^2$$

The parameters λ can be set to get a specific trend within the learning. For example, one could set the parameters as $\lambda_E = \lambda_F$ which will give the same path during the learning. Alternatively, one could set $\lambda_E < \lambda_F$ to favor the learning of the forces over energies. This last parameterization accelerates the training procedure at the beginning. This loss is then averaged over the configurations in a batch. This process favors the multi-output network and provides models with better generalization capability and flexibility. At this stage, all the parameters within the embedding are trained with a starting learning rate that we are calling here τ_{ini}. Furthermore, τ_{ini} is reduced when the error on the training set is converging, and this first stage stops when the learning rate reaches 10^{-4}.

In the second stage, the embedding is frozen for the multi-layered preceptors for charges and volumes. Therefore, the learning focuses only on short-range interactions. Then, by keeping the previous weight for λ_E and λ_F the second stage begins, ensuring that the volumes and charges

are not modified during its process. It stops when the errors between the validation set and the train set drop to the same order.

FeNNol framework

Developed by Plé et al.[15], FeNNol is an open-source Python library designed to enhance atomistic machine learning models. Leveraging the FeNNix model it simplifies the creation, training, and execution of molecular dynamics (MD) simulations across various systems. Built on the Jax framework, FeNNol supports GPU acceleration, making it highly efficient for complex model processing.[15]

1. **Chemical and radial encoding:** To perform the learning, we need to parametrize first the encoding of the chemical environment. Different chemical encodings can be used with FeNNol. The one-hot encoding described before, one based on the electronic occupancy called TeaNet[16], electronic structure encoding called SpookyNet[17], position encoding presented previously in Figure C-4, and 4D encoding based on the Stowe-Janet-Scerri periodic table[18]. A combination of those encodings can also be done and optimized during the process.

   The radial encoding can be done using simple Gaussian basis, Bessel basis, inverse-distance Gaussian, Fourier expansion, and Bernstein polynomials.
2. **Physics Modules** Some modules can be used in FeNNol. They encompass, for example, the Coulomb module computing the electrostatic interactions between distributed charges. For the computations of long-range electrostatic interaction, the Ewald summation is used. The use of this technique effectively handles the computational challenges of the slow convergence of electrostatic interaction summation. It gives rapid convergence, flexibility, and accuracy.[19] Different modules can also be used to handle the dispersion/exchange interactions, charge equilibration, and repulsive pair potential.

With this foundation, one can build their own FeNNol potential. A key aspect of this method is the introduction of a new embedding named CRATE (Configurable Resource from Atomic Environment), specifically developed to combine chemical and geometric information. It incorporates species embeddings along with radial and angular information. The embedding process also utilizes radial basis and angle embeddings and tensorial information, features that can be interconnected across different layers using message-passing. This method iteratively refines

an embedding vector for each atom, x_i, within the system across multiple layers by incorporating data from the local environment.

Initially, two distinct vectors are utilized: r_i; which corresponds to information from a previous layer, and s_i which combines local resources. The transformation process is governed by the following equations:

$$r_i = W_r^{(L)} x_i + b_r^{(L)} \quad \text{(C-16)}$$

$$s_i = W_s^{(L)} x_i + b_s^{(L)} \quad \text{(C-17)}$$

Where $W_r^{(L)}$ and $W_s^{(L)}$ are weight matrices at the L-th layer of the NN. These matrices are responsible for transforming the input embedding vector x_i into the new vector r_i containing information from the previous layer and s_i combining local resources. The bias vectors are $b_r^{(L)}$ and $b_s^{(L)}$ are added to the linear transformations to shift the output vectors r_i and s_i. Local geometric resources are then defined in Equation C-18 by a vector Ri:

$$R_i = r_i \parallel \left(\parallel_{g \in \mathrm{res}(L)} g\left(x_i, \{s_j\}\right) \right) \quad \text{(C-18)}$$

This vector R_i contains all the local resources from atoms i and j and the information of the previous layer of atoms i (r_i). It is an iterative concatenation $\parallel_{g \in \mathrm{res}(L)}$, meaning that we concatenate all local resources computed for the layer L s_j, with the embedding x_i using a function g. Each function g processes its inputs to compute a set of features that describe the geometric relationship between atom i and its neighbors $g\left(x_i, \{s_j\}\right)$. It is then mixed through a Multi-Layer Perceptron (MLP) depending on the layer L to update the embedding.

$$x_i{}^{(L+1)} = \sigma\left(F^{(L)}\right) \odot x_i{}^{(L)} + u\left(MLP^{(L)}(R_i)\right) \quad \text{(C-19)}$$

Where $\sigma\left(F^{(L)}\right)$ is a trainable "forget gate". With σ, the sigmoid function is used as a weight to control the flow of information as it ranges between (0,1). $F^{(L)}$ represents a vector of trainable parameters.

- A value close to 0 for $\sigma\left(F^{(L)}\right)$ means *forget a lot of the previous information,*

- A value close to 1 retains *most of the information*.

The element-wise multiplication $\odot$, applied the "gate" values to the embedding x_i, which filters the embedding based on the relevance decided by the forget gate. This first part of Equation C-19 maintains a balance between old and new data, ensuring the model's adaptability. On the right-hand side of Equation C-19, only the activation function u is used to enable complex transformations.

Regarding the radial resources g_i^{rad}, they are based on a radial basis $B(r_{ij})$ which can be defined with Bessel functions, for example. As in FeNNix, it is multiplied by a function $f(r_{ij})$ that tends to zero when atoms are outside the cutoff radius. It is defined as in equation C-20.

$$g_i^{\text{rad}} = \sum_j s_j \otimes B(r_{ij}) f(r_{ij}) \tag{C-20}$$

Where here we have a tensor product $\otimes$ between the local resources of the neighbors s_j of atom i with the radial basis.

Concerning the angular resources, it is a combination of triplets of atoms. To build it, it is first reduced to a chemical-radial basis D_{ij} and D_{ik}, which are considered as two edges. (In a Graph Neural Network, one can consider each atom as a node and each bond as an edge).

$$[D_{ij}]_c = \tilde{f}(r_{ij}) \sum_{ab} [\tilde{B}(r_{ij})]_a [s_j]_b \left[W_{\text{ang}}^{(L)}\right]_{abc} \tag{C-21}$$

Where $\tilde{f}(r_{ij})$ is the switching function which goes to 0 depending on the cutoff chosen by the user. $\tilde{B}(r_{ij})$ is the radial basis for the triplet graph, and the $W_{\text{ang}}^{(L)}$ is the trainable tensor. Then, from that, the angles between the two edges are projected into an angular basis Θ_{ijk} and we get our angular scalar g_i^{ang} which is defined in Equation C-22:

$$g_i^{\text{ang}} = \sum_{\{jk\}} \Theta_{ijk} \otimes (D_{ij} \odot D_{ik}) \tag{C-22}$$

Where we have a tensor product between the angular basis and result of the element-wise multiplication $\odot$ between the two edges D_{ij} and D_{ik}. From this, we get the angular resources,

which depend on a cutoff that can be different from the one used for the radial resources of Equation C-20.

The equivariant resources are then defined by three different components (equivariant tensors are defined with hat notation). First, an equivariant neighborhood basis from a chemical radial basis $\hat{\rho}_i^{(L)}$. It depends on two parameters. The interaction feature $b_{ij}^{(L)}$ defined in Equation C-23, which is a tensor product between the embedding of atom j and the radial basis.

$$b_{ij}^{(L)} = W_e^{(L)} \left(s_j \otimes B(r_{ij}) f(r_{ij}) \right) \tag{C-23}$$

Where $W_e^{(L)}$ is a trainable weight matrix specific to the layer L that converts the radial chemical basis to a user-selected $N_{channels}$ equivariant channels. It allows the model to handle complex transformations, e.g. to ensure rotational invariance. The number of channels is kept small because of the computational cost.

Second, the equivariant embedding of the interaction between atoms i and j at layer L.

$$\hat{V}_{ij}^{(L)} = \begin{cases} \hat{V}_j^{(L-1)} \; if \; L > 1 \; and \; message - passing \\ else \; \hat{Y}_{ij} \end{cases} \tag{C-24}$$

Where $\hat{V}_j^{(L-1)}$ is in the context of message passing the equivariant embedding (and $L \neq 0$) from the previous layer L-1 for atom j. The information from atom j is directly used to update the interaction for the current layer L. $\hat{Y}_{ij}$ depends on a user-defined maximum degree l_{max} and is the tensor representing the direction coefficient between atom i and j in the real spherical harmonics.

Then, from that, we can define, as expressed in Equation C-25, the chemical-radial basis $\hat{\rho}_i^{(l)}$

$$\hat{\rho}_i^{(L)} = \hat{\rho}_i^{(L-1)} + \sum_j b_{ij}^{(L)} \otimes \hat{V}_{ij}^{(L)} \tag{C-25}$$

After forming the neighborhood basis $\hat{\rho}_i^{(L)}$ we initialize the equivariant embedding $V_i^{(0)}$ for the first layer L=0 as:

$$\left[V_i^{(0)}\right]_{n\lambda m} = \sum_{n'} \left[W_v^{(0)}\right]_{nn'\lambda} \left[\hat{Y}_{ij}\right]_{n'\lambda m} \tag{C-26}$$

Where n, n', m and λ are indices used to navigate and manipulate tensor components. $\left[\mathrm{W}_{\mathrm{v}}^{(0)}\right]_{\mathrm{nn'\lambda}}$ indicates the weight matrix that maps input tensor components from index n' to n. $\left[\hat{\mathrm{Y}}_{\mathrm{ij}}\right]_{\mathrm{n'm}}$ interaction tensor between atoms i and j capturing features indexed by n', λ, and m.

Then, after forming the neighborhood basis of Equation C-25 and the initialize embedding of equation C-26, a series of tensor products is done to update $\hat{V}_i$:

$$\left[\hat{L}_i^{(L,t)}\right]_{n\lambda m} = \sum_{p=(\lambda_1,\lambda_2\to\lambda)} \left[W_{\text{path}}^{(L,t)}\right]_p \left(\sum_{m_1,m_2} C_{m_1 m_2 m}^{\lambda_1\lambda_2\lambda} \left[\hat{\rho}_i^{(L)}\right]_{n\lambda_1 m_1} \left[\hat{V}_i^{(L-1,t)}\right]_{n\lambda_2 m_2} \right) \tag{C-27}$$

$$\left[\hat{V}_i^{(L,t)}\right]_{n\lambda m} = \sum_{n'} \left[W_v^{(L,t)}\right]_{nn'} \left[\hat{V}_i^{(L-1,t)} + L_i^{(L,t)}\right]_{n'\lambda m} \tag{C-28}$$

Where $C_{m_1 m_2 m}^{\lambda_1\lambda_2\lambda}$ are Clebsch–Gordan coefficients for coupling different angular momentum states, allowing the model to preserve the symmetries of the system, such as rotationally equivariant representations. $W_{\text{path}}^{(L,t)}$ is the path weight which are trainable parameters that weigh different tensors products paths. It allows the control of how symmetric components interact and influence the updated embedding, ensuring the stability of the iterative tensor products. Equation C-28 is just the update of the equivariant embedding $\hat{V}_i$ followed by channel mixing.

Finally, the equivariant resources g_i^{E3} at layer L is just a concatenation of the scalar irreps from the different $L_i^{(L,t)}$ and is defined in Equation C-29:

$$g_i^{\mathrm{E3}} = \|_{t,n} \left[L_i^{(L,t)}\right]_{n00} \tag{C-29}$$

More information about long-range resources is also possible with FeNNol and all mathematical functions about it are expressed in the paper by Plé et al.[15] In this thesis we did not use long-range resources. The main reason for using CRATE is its efficiency, as it requires less resources than the Allegro method.

FeNNol training procedure

1. **The dataset and its evaluation:** First, the dataset should be created. It should be composed of the cartesian coordinates, the energies, and the forces in a pickle format file. The batch size should be set up appropriately. This parameter determines the number of geometries processed during a single learning iteration.
2. **Optimization parameters:** We used the Adabelief[20] optimizer to train our models, which is a stochastic gradient descent with learning rate adaptability. It adjusts the learning rate based on the 'belief' in the observed gradient's direction, determined by how closely the observed gradients follow the predicted one. If the actual gradient significantly deviated from its exponential moving average, it indicated low reliability. Consequently, AdaBelief takes smaller steps. Conversely, if it is too close to the predicted, larger steps are taken. The algorithm is explained and compared to the Adam optimizer in the paper from Zhuang et al.[20] Some parameters during the learning can also be frozen to do fine-tuning or to perform transfer learning.
3. **Loss and stages:** Different losses can also be defined – mean squared error, mean absolute error, or log hyperbolic cosine, for example. It can be set for either the whole system or per atom. Different stages can be performed. For example, if one wants to do a transfer learning[21] between different methods like DFT and ReaxFF, one can set a first learning on ReaxFF and then adjust them with DFT with another learning called transfer-learning. The number of epochs and learning rate can then be reparametrized for the second learning.

MD simulations The FeNNol module allows users to select various ensembles and thermostats, including Langevin, Bussi-Parinello, or Nose-Hoover, and to specify simulation times. Users can also choose between periodic and non-periodic boundary conditions. It is also based on GPU, which speeds up the computation. This makes it particularly advantageous for conducting complex simulations that require extensive computational resources.

Additionally, the inherent flexibility and precision of the FeNNol framework make it particularly suited for developing new potential models in material science, such as those for gold and its interactions with thiolates.

One of the primary motivations for this project is the development of an accurate potential to predict complex reactivities, such as the restructuring of thiolate self-assembled monolayers on gold surfaces.

To initiate the machine learning process, it was essential to assemble a comprehensive ab initio reference dataset. This dataset is fundamental for the training mechanism, providing the necessary data points to teach the model how to accurately predict and simulate complex chemical reactions and interactions. The next chapter will detail the procedures followed to gather and refine this data, highlighting the sources, selection criteria, and preprocessing steps that ensure the dataset's relevance and utility for training the FeNNol model.

C.3) Preparation of the training set

C.3.a) Quantum calculation

To get accurate reference data for the data set, quantum methods such as DFT, CCSD(T), or MP2 are the best choices. They provide a full electronic analysis for each atom in a system. Based on the resolution of the time-independent Schrödinger equation using the electronic molecular Hamiltonian, $H\Psi(x,t)=E\Psi(x,t)$. They can give information about total energies, forces, charges, and dispersion energy. These approaches use the Born-Oppenheimer approximation, which relates to the fact that electrons move much faster than nuclei, allowing the equation to be separated into electronic and nuclear parts.

To approximate the wavefunction, different methods can be used, and they are all based on the Self Consistent Field (SCF) procedure, which is an iterative method approximating the wave function / electronic density.[22] Offering access to a set of orbitals.

The Hartree-Fock (HF) method This method is based on a molecular orbital approximation where the total wavefunction of the system is the product of one-electron wavefunctions. Each one is defined as an orbital and is determined according to a basis set. This approximation considers the independence of every individual electron. It is a variational method that considers that the energy of the HF wavefunction is always greater than the exact value.[23] The objective of the SCF using the HF method is then to determine the best coefficient that minimizes the energy. It includes the exchange energy, which accounts for the antisymmetry of the electronic wavefunction. Arising from the Pauli exclusion principle, it reflects the repulsion and correlated motion of the electrons.

It gives an accurate description of the electron distribution within a system.[24] While Hartree-Fock provides a useful approximation of the wave function, it fails to account for electron correlation, leading to significant errors in systems where these interactions are strong.

Full configuration interaction (FCI) Is the most accurate methods that consider all possible electron configurations for a given number of electrons and orbitals. In this method, the wave function is defined as a linear combination of determinants. It belongs to the so-called post-Hartree-Fock class of methods, as the initial molecular orbitals originate from an HF calculation.[25,26] FCI constructs the wavefunction as a linear combination of all possible electrons and fully captures electron correlation, providing accurate energy estimation. Even though this method provides a very accurate electronic description, it is computationally costly, which makes it impossible to use for large systems on present computers. Indeed, its scaling is O(N!), where N is the number of orbitals that are used. The factorial growth comes from the detailed consideration of electron interaction across all configurations, *one of the main arguments from which arises the use of possible quantum computing*. Different methods have been developed to approximate the results of FCI without having extreme computational expenses, such as selected CI (sCI), which uses only a subset of the most important configurations based on the contributions to the overall wave function, reducing the required computational resources.

Couple Cluster Singles and Doubles (CCSD) It's a post-non-variational Hartree-Fock method that includes all single and double electron excitation from a reference determinant. Its name comes from the exponential ansatz of the cluster operator. It can also include a perturbative treatment of triple excitations CCSD(T), which is considered one of the most accurate methods for QM (or golden standard). Even though CCSD is less computationally intensive than FCI, it still captures a significant proportion of the correlation energy. The method is scaling at O(N^6).[27]

Density Functional Theory DFT is a class of methods that is widely used as it offers a good balance between accuracy and computational efficiency. DFT is used in both academic research and industrial applications, ranging from the development of new materials and drugs to the study of catalytic processes. It approximates the electron density rather than the wavefunction: [28]

1. **Hohenberg and Kohn theorem:** DFT is based on the Hohenberg-Kohn theorem, which states that the energy of the ground state of any interacting particle system is a function of the electron density.[29] Moreover, it is a variational theorem where the energy is always greater or equal to

the exact energy. This means that the true ground state density minimizes the energy functional, providing a way to calculate properties by minimizing $E[\rho] concerning\ \rho$. From these statements, the Kohn-Sham method is used, offering a practical way to approximate this function.

2. **The Kohn-Sham (KS) model:** This approach is a method that implements the principle of DFT. The total energy functional is decomposed into several terms. The kinetic energy of non-interacting electron $T_s[\rho]$ which defines a fictitious system of non-interacting electron which has the same density as the interacting one. $V_{ext}[\rho]$ is the external potential energy present here due to the external potential acting on the electrons. The Hartree energy, $E_H[\rho]$ which is the classical electrostatic energy. Finally, the $E_{xc}[\rho]$ which is the exchange-correlation energy that accounts for the remaining non-classical interactions.[30] To find the ground state density, the KS equations should then be resolved. They are derived by minimizing the total energy functional concerning the density under the constraints that density comes from orbitals that are solutions to a single-particle Schrödinger equation. This method reduces the many-body problem to a set of single-particle problems, thus decreasing computational costs. It gives a good balance between accuracy and cost. Different functionals have been proposed to describe $E_{xc}[\rho]$.
3. **The functionals:** Different functionals are used to describe the exchange-correlation energy. In this study, we are focusing on gold, to choose the best functional which provides good accuracy without requiring excessive computational effort, one should follow the so-called Jacob's ladder of DFT.[31] It sorts the DFT functionals from the very simple ones to the most complex ones. The functional with the lowest accuracy is the local density approximation (LDA) where $E_{xc}[\rho]$ depends only on the density at one point. Then, the generalized gradient approximations (GGA) use both the density and its gradient at each point, greatly reducing the bond dissociation energy error and improving transition-state barriers. Next is the meta-GGA, which depends on the $T_s[\rho]$. Next, the hybrid functionals mix some amount of exact exchange within a GGA approach. It allows to mimic static correlation effects and produces a highly accurate functional coming of course with its cost.[32] The choice of the right functional is then really important to capture all the electronic properties of a system. A benchmark study has been conducted on various functionals for neutral gold clusters, ranging from GGA to hybrid

ones. The Perdew-Burke-Ernzerhof (PBE) functional was found to give a minimal error, making it highly suitable for our study.[3]

4. **Basis sets:** They are used to describe the electron density distribution. Different types of basis sets are used in DFT calculations. Slater-type orbitals (STOs) have the advantage of closely mimicking the actual shape and behavior of atomic orbitals with a correct structure, leading to a more accurate physical representation. They decrease exponentially, which matches the behavior of actual atomic orbitals. However, the major drawback of STOs in computational chemistry is their computational complexity, especially when it comes to evaluating multicenter integrals required for molecular systems. For hydrogen-like atoms, they provide exact analytical solutions. Gaussian-type orbitals (GTO) provide a practical application for a wide range of chemical systems, from small organic molecules to large biological complexes and materials. Their use simplifies the calculation of molecular integrals as analytical expressions exist for multicentred integrals, making them more tractable than STOs. They are also computationally efficient and more flexible, where they can be combined to approximate a variety of orbital shapes and sizes with strong capabilities to model the electron cusp problem. Then, plane wave basis sets also exist and are suitable for the study of periodic systems in condensed matter physics and materials science. They are uniformed across space and can be optimized by adjusting a cutoff energy, which determines the maximum kinetic energy of the plane waves included in the basis set. Choosing the correct basis set, such as GTOs for molecular systems and plane waves for periodic systems, is crucial, as it affects the accuracy of the results and the computational resources required.
5. **Pseudopotentials and dispersion:** To reduce computational requirements, pseudopotentials can be used. They replace the core electrons with an effective core potential (ECP), thus focusing on the valence electrons, which are the most responsible for chemical properties. Moreover, to accurately describe the physical and chemical properties of a system, one can also consider dispersion interactions. It arises from the correlated movements of electrons between atoms and molecules, which are not captured by conventional DFT due to their local treatment of electron correlation. These involve adding a correction term (i.e. dispersion-corrected DFT) to the energy calculated by DFT to account for dispersion. It Is then renamed DFT-d2 or d3, depending on the empirical term calculated. For example, the use of Grimme's dispersion.[33]

Overall, one should carefully combine the choices of the right functional, basis set, pseudopotential, and dispersion correction to get a satisfactory description of a given studied system. For gold systems, a vast literature of DFT studies exists. It has been found that the best functional to describe the bond between two gold atoms is the Perdew-Burke-Ernzerhof mixed with a d3 correction, also known as PBE-d3, which is a GGA functional. [34-35] This functional gives total energy properties in good agreement with the experiment with a low mean average error. Concerning the basis set, the Stuttgart basis sets are used, which are also called Stuttgart relativistic small core (RSC) effective core potential (ECP). RSC basis can handle heavy atoms without requiring too much resources and by still staying accurate. It replaces the core electrons of an atom with a potential that effectively accounts for their shielding effects on the valence electrons. It also takes into account relativistic effects, which are crucial for the description of gold. Thus, there is a spin-orbit coupling happening for gold, which is defined as a relativistic interaction between the spin and the orbit motion of electrons. This relativistic effect leads to a contraction of the s and p orbitals and, conversely, a potential expansion of the d and f orbitals. The gold Stuttgart ECP incorporates these relativist effects into its formulation, making it more accurate and realistic than non-relativistic basis sets. Using such an approach strengthens the description of all the gold systems that need to be computer-using DFT to obtain reference energies and forces.

C.3.b) Dataset preparation

To perform the learning discussed in section C-3, one should have accurate ab initio reference computations. Here, with the use of FeNNol, we want to learn about energies and forces. Nevertheless, to do the calculation, we need to find different equilibrium and non-equilibrium systems to get a flexible potential. First, the goal was to find the different systems that were already existing in the literature. We took our reference structures from the Muller group[36] , which consists of a large dataset of 63 015 low-energy atomically precise nanoclusters that are available for a lot of elements. Concerning gold clusters, structures from 3 to 55 atoms are gathered with both local and global minimum. As we are limited in terms of computational resources, we've decided to limit our calculation to a maximum number of 25 gold atoms. Another point that was considered is the switch in terms of geometry. It is known that from 12 to 15 gold atoms, the clusters go from a 2D to a 3D structure. We ensured ourselves to gather more data in this range for the potential to be able to differentiate between those. Therefore, we tried to gather as many geometries as we could from different studies and databases. In total, we were able to gather several minimum

structures that were equilibrated. Nevertheless, for the potential to be flexible, we needed to include non-equilibrium structures.

Molecular dynamics simulations Using MD simulations via LAMMPS combined with ReaxFF, we were able to create non-equilibrium structures. We used the Monti et al. force fields[37]with a Berendsen thermostat, a timestep of 2 fs, and heating of the simulation box until 1300 K. To ensure that the system is beyond equilibrium, a small deviation of 0.25 Å was applied at each timestep during the simulation. A filter had also been applied at the end of each simulation to remove redundant geometries. A radial analysis was performed by comparing each pair of atoms during the simulation. If the geometry from the timestep t is too close to the previous one at the timestep $t-1$ then it is removed from the dataset. To give an example, at a low temperature under 600 K, the different systems did not move enough as such temperature appears pretty low for such stable systems even with the use of deviation. Nevertheless, when reaching a higher temperature and arriving around the boiling point of gold, which is around 1337 K, the clusters started to move a lot, and we were able to get non-equilibrium systems. Moreover, we wanted to create a potential that can capture clusters for both chemistry at the surface and on the clusters. Nevertheless, we were not able to create reference data with a large number of atoms, so since it is stable with ReaxFF, we used it for the surface by simulating two different sizes of nanoparticles. The main drawback of using ReaxFF as a reference is that it does not accurately distinguish between each gold atom and treats them equally. This means that nanocrystals cannot see the differences between atoms located on edges and facets. Therefore, it does not accurately predict specific interactions, which may account for the lack of prediction for complex interactions. However, it can simulate a stable gold surface. Our idea was to combine ReaxFF's surface stability for gold systems with accurate DFT descriptions of gold clusters. This should allow us to obtain a flexible potential capable of capturing both surface types and cluster chemistry.

The process that uses the DFT reference data is presented in Figure C-6. First, global and local minimum are gathered from different databases and studies. Then, MD is applied on each structure using ReaxFF, and the filter is applied. DFT calculation is then done on each non-redundant structure using the Gaussian software. The PBE function is used with the Stuttgart basis set and Grimme dispersion d3 correction. Note that we are using unrestricted functional, an important point since it allows for paired electrons to occupy different spatial orbitals. Therefore,

for an odd number of electrons, it is more flexible and describes more accurately the spin polarization. It ensures a good treatment of open-shell systems, which forces equal spin distributions. Moreover, gold atoms exhibit significant relativistic effects and spin-orbit coupling due to their high atomic number. Unrestricted PBE can better account for these effects because it allows for different spatial distributions of spin-up and spin-down electrons. This is crucial for accurately modeling the electronic structure of gold clusters.

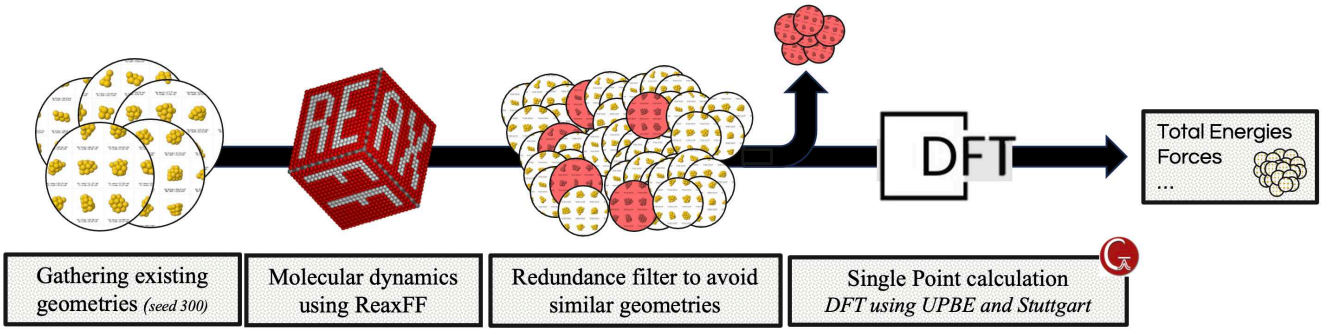


Figure C-6: Scheme of the dataset preparation.

From that, we got specific energies and forces for each system, and we prepared the dataset as presented in Figure C-7. Systems from 3 to 25 atoms were used, having a total of 15 946 structures for the training along with 1500 structures for the validation. The cluster containing 22 gold atoms has been excluded due to a problem of convergence.

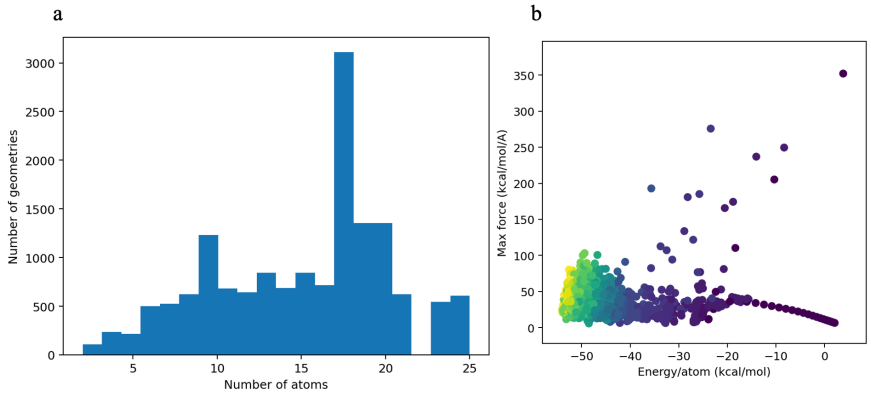


Figure C-7: DFT dataset properties with (a) the number of systems that are used in comparison to the number of atoms and (b) the relationship Between Energy per Atom and Maximum Force.

Concerning the ReaxFF simulation, the dataset properties are presented in Figure C-8. We did a simple simulation on two nanoparticles of two different sizes, i.e., 2 and 4 nm, which are

called ns2 and ns4, respectively. We wanted to obtain non-equilibrium structures, so, as for the DFT dataset, an increase until 1300 K was done, along with a shift of each atom of about 0.25 Å.

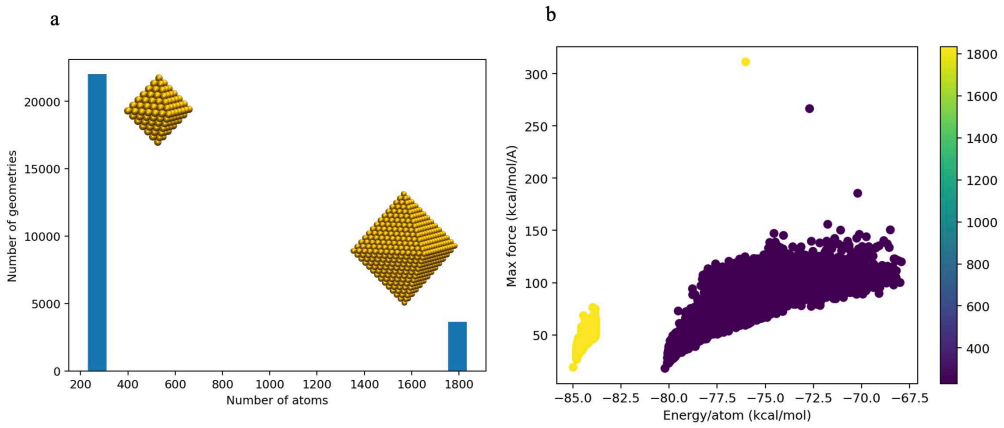


Figure C-8: ReaxFF dataset properties with (a) the number of systems that are used in comparison to the number of atoms and (b) the relationship Between Energy per Atom and Maximum Force

The filter is also used in this part to avoid redundancies. In total, for the ReaxFF dataset, we have 21 812 training geometries along with 3849 validation geometries. Then, using both datasets, the learning process can begin.

C.4) Development of a machine learning potential

In the present context, all learning processes are based on accurately capturing the energies and forces. In this implementation, we utilized the CRATE module for the embedding process, which is specifically parametrized to handle the complexity of these interactions. The CRATE module is configured with a dimensionality of 512 and consists of 2 layers. For the radial basis functions, the model employs 10 Bessel functions, which are particularly effective in modeling the radial dependencies of atomic interactions due to their smooth, oscillatory nature.

The embedding of atomic species is informed by their electronic structure, providing a nuanced representation that aligns closely with physical chemistry principles. This approach ensures that the embeddings reflect the inherent properties of each element, enhancing the model's accuracy and predictive capabilities.

For the computation of energies, the model leverages a dedicated Neural Network module, which is structured with [256, 128, 64, 32, 1] neurons across its layers. We interpret the output neuron as an atomic contribution to the potential energy of the system.

The different parameters used for defining the potential include a cutoff radius fixed at 5.3 Å, beyond which interactions between atoms are not considered. Two message-passing are used, giving a receptive field of 10.6 Å. L_{max} is equal to 2, $N_{channels}$ to 16 and two atom-centred channel-wise tensor products are used. The learning rate is initially set at 10^{-4} and gradually reduced to 10^{-5}, facilitating stable convergence over training iterations. This careful management of the learning rate helps prevent overfitting and ensures that the model reliably learns the underlying patterns in the data.

To have a flexible potential that can simulate both the surfaces and the clusters, a transfer learning procedure is performed. First, a learning procedure based on ReaxFF simulations is done. As this method, with the use of the fore fields made by Monti et al.[37], gives stable gold surfaces, we will use it to have a good starting point for the simulation of the surfaces. Then, from that, the potential is refined using the DFT dataset. To ensure that the potential is working accurately, some tests are then performed:

- First, a comparison between the radial distribution function of each potential is done. The objective is to see if the surface stability can be reproduced with our different potentials.
- An energetic analysis is also performed by comparing the dissociation of one gold atom from the Au_{20} cluster, which is a pyramid-like structure. A comparison between our potentials, the DFT, and ReaxFF is performed.
- A radial analysis is also done on the optimized Au_{20} cluster. To see if our potential can discriminate between each gold atom that is on the cluster and if it is close to the DFT calculation.

C.4.a) ReaxFF-based potential

As explained before, the FeNNol python module is used. It allows us to easily create machine learning potentials and then use them directly to perform MD simulations. The parameters that were used for the first step of the transfer learning on the ReaxFF dataset is a batch size of 64 with a max epoch of 2k. The batch size, as explained in the previous section, is the structures that are considered at each epoch. As it is a GNN, we had to reduce the number of batch size as a lot of atoms are considered. Thus, it increases a lot the computation requirements. For 2 and 4 nm, we respectively considered a total of 230 and 1 834 atoms. During the learning, every pair of atoms is considered and needs to be analyzed to create a graph of the different layers.

By launching the learning with all the geometries, we saw that the learning was pretty slow, with a preprocessing time reaching 7.8 minutes, and considering that we had a total of 2k epochs, it would have taken 10 days to reach the end. Therefore, we considered having parallel learning with only ns2 to get fast learning and be able to change some parameters. We have then three different learning steps. The first is considering the same weight for both energies and forces where $\omega_e = \omega_f = 1$. Followed by a change in terms of weight, with $\omega_e = 2 \wedge \omega_f = 1$. We wanted to see how forcing the learning to go towards the energy impacts the accuracy of the learning. Both learnings are based only on the reference of the ns2 systems. The last is the slowest, with the same weight for both energies and forces, $\omega_e = \omega_f = 1$ but this time it considers both ns2 and ns4 geometries. The reason behind this comparison is to see the different patterns depending on the number of geometries and weights, as presented in Figure C-9. By using the same weight for the energies and the forces as depicted in green in Figure C-9, we can see that the learning for the energies is going down quickly, reaching 0.641 kcal/mol per atom after only 15 epochs. Concerning the forces, we are going down also, but it took more epochs for the training to reach its plateau of 0.119 kcal/mol/Å. Moreover, for the energies, the errors are still considered pretty high as it is counted "per atom." To counter this, we wanted to make the learning for the energy stronger by changing their respective weights with 2 for energies and 1 for forces represented in red.

For the energies, it appears that it does not change at all the error as it is reaching its plateau as fast as with the previous learning with a value of 0.629 kcal/mol. Nevertheless, for the forces when using $\omega_e = 2$ and $\omega_f = 1$ (*the red curve*), there is a large influence where two different plateaux were reached. One after 100 epoch with a value of 0.778 kcal/mol/Å, which stays stuck until 1000 epoch.

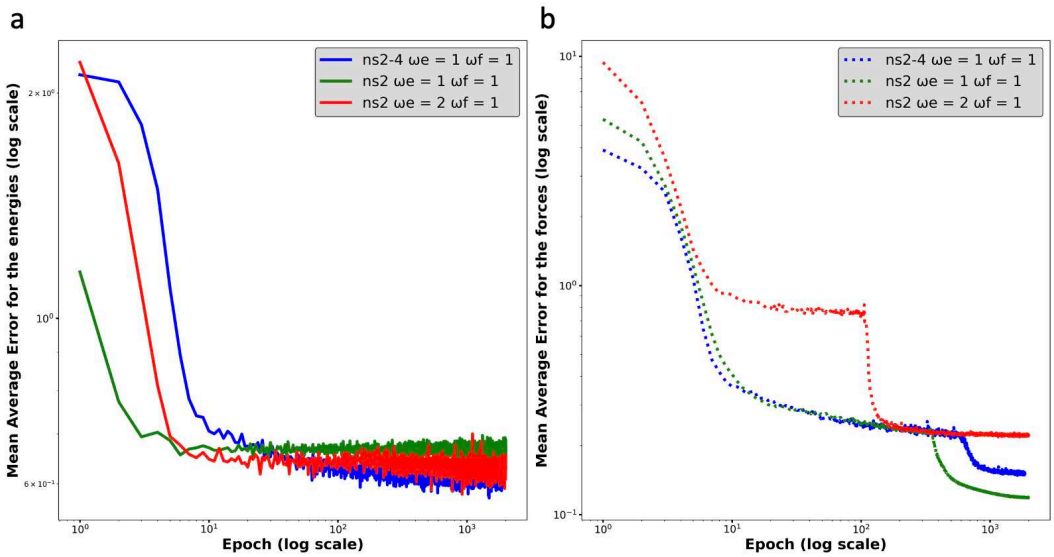


Figure C-9: log plot of the mean average error along with the number of epochs for first learning from the ReaxFF dataset. (a) is the MAE of the energies and (b) of the forces. The green and red curves correspond to the learning only composed of the ns2 systems; the red uses different weights for the learning for the forces and the energies, and the blue one considers all the systems ns2 and ns4.

Then, the learning is going down again, reaching another plateau at the same level as the first learning with a value of 0.211 kcal/mol/Å. Finally, by comparing it with the learning with all the geometries ns2 and ns4, we can see that for the energies, we are getting lower than the two others with a value of 0.575 kcal/mol. Concerning the forces, we are reaching 0.150 kcal/mol. Nevertheless, the values for the energies are still pretty high, but to see if we can still have stable surfaces, another analysis is done.

Molecular Dynamics is then performed with the use of the three potentials. The temperature is set at 300 K using the NVT ensemble along with the Langevin thermostat. The system studied is the Au (111) surface with 8 layers, using 1 ns dynamic with a timestep of 2 fs. By looking directly at the simulation as presented in Figure C-10, it appears that we have good stability on the surface with no differences between each learning.

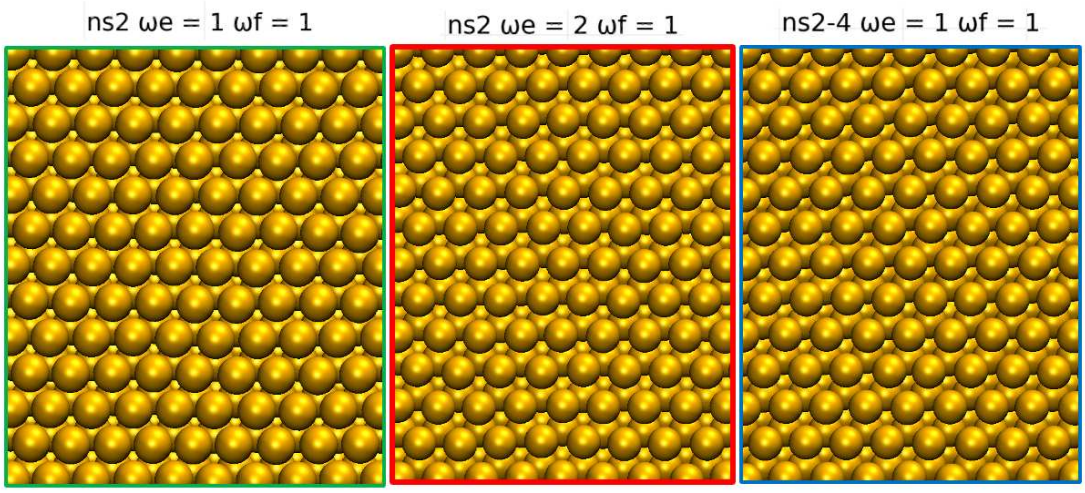


Figure C-10: Visualization of the ending of each MD simulation made with the different potentials. The green corresponds to the learning of only the ns2 systems; the red has a different weight for the learning of the forces and the energies, and the blue one considers all the systems ns2 and ns4

Nevertheless, to ensure this stability, the radial distribution function is plotted to see how the distances between atoms are evolving depending on the potential, as shown in Figure C-11. It provides a statistical description of each atom around its center. It allows us to see any differences in terms of description for each atom's environment.

There are no differences in terms of stability for all three MD simulations. There is a high intensity at 2.88 Å, which corresponds to the equilibrate distances found in MD simulation using ReaxFF. It means that the first potential is already capable of getting a stable gold surface by mimicking the normal behavior based on ReaxFF energies and forces. There is no real influence on the different weights for the learning. We've then picked the potential with the same weights containing ns2 and ns4 structures.

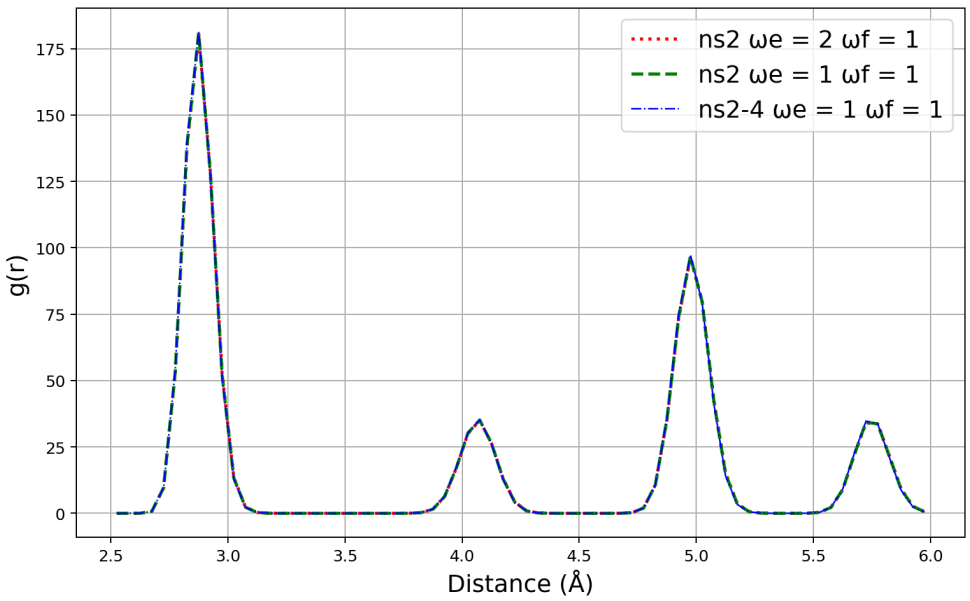


Figure C-11: radial distribution function for the three different MD simulations at 300K, performed using different learning parameters. The green curve corresponds to the learning limited to the ns2 systems, the red has a different weight for the learning for the forces and the energies, and the blue one considers all the systems ns2 and ns4.

C.4.b) Refinement with DFT calculations

As explained previously, we are using the ReaxFF potentials while performing a final transfer learning on DFT reference data. The dataset preparation for the DFT is explained in section 4b. Different learning procedures were performed to see the influence of the number of epochs to get a good yield between parameters from ReaxFF and DFT. 1k to 5k epochs are then used, as presented in Figure C-12. The goal is not to lose too much information from the previous potential and to gain enough from the clusters. The risk is, therefore, to overfit on clusters DFT dataset and to obtain unstable MD simulations on the Au (111) surface. In such a case, the surface would lose its stability and suffer from an agglomeration effect. This indicates that the stability of the surface obtained from the ReaxFF pre-learning is compromised. The opposite is also true: if the learning process on the DFT reference data is insufficient, the resulting precision for small clusters and the discrimination of each gold atom will also be inadequate.

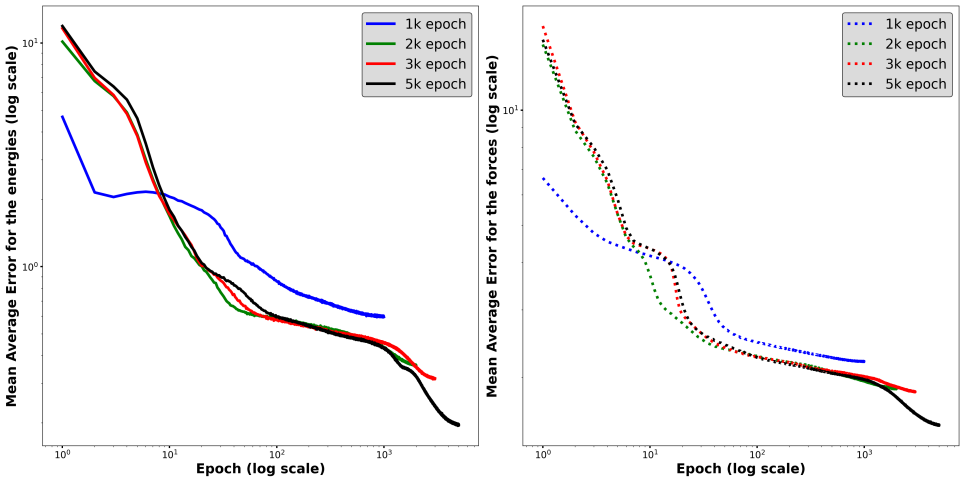


Figure C-12: comparison between different transfer learning procedures from ReaxFF potential with ns2 using different numbers of epochs.

Comparing the different epochs (as presented in Table C-1), one can see that the more we increase the number of epochs, the more the mean average error decreases. By reaching the 5k epoch, we can go down to 0.196 kcal/mol (per atom) and 1.493 kcal/mol/Å for the forces and the energies, respectively.

Table C-2: Mean average errors for energies and forces for learning with different numbers of epochs

Number of epochs	MAE on energies *(kcal/mol)*	MAE on forces *(kcal/mol/Å)*
1000	0.598	2.196
2000	0.361	1.862
3000	0.314	1.828
5000	0.196	1.493

Nevertheless, by looking at the radial distribution function, it can be seen that when reaching 1k, we can reproduce the right distances for the surface for the equilibrium peak. But when looking to 2k and more, we have a shift on the left for the distances. Furthermore, this shift in distances can be directly seen by looking at the snapshots as presented in Figure C-13. It appears that it tends to cluster-like structures. The most stable conformations are those with the number of epochs corresponding to 1k and 3k, the others being overfitted. Nevertheless, for the 3k epochs case, it can be seen that the function appears smoother, which means that the system is melting,

which does not represent the wanted stability for the solid gold surface, which is not supposed to melt at 300 K.

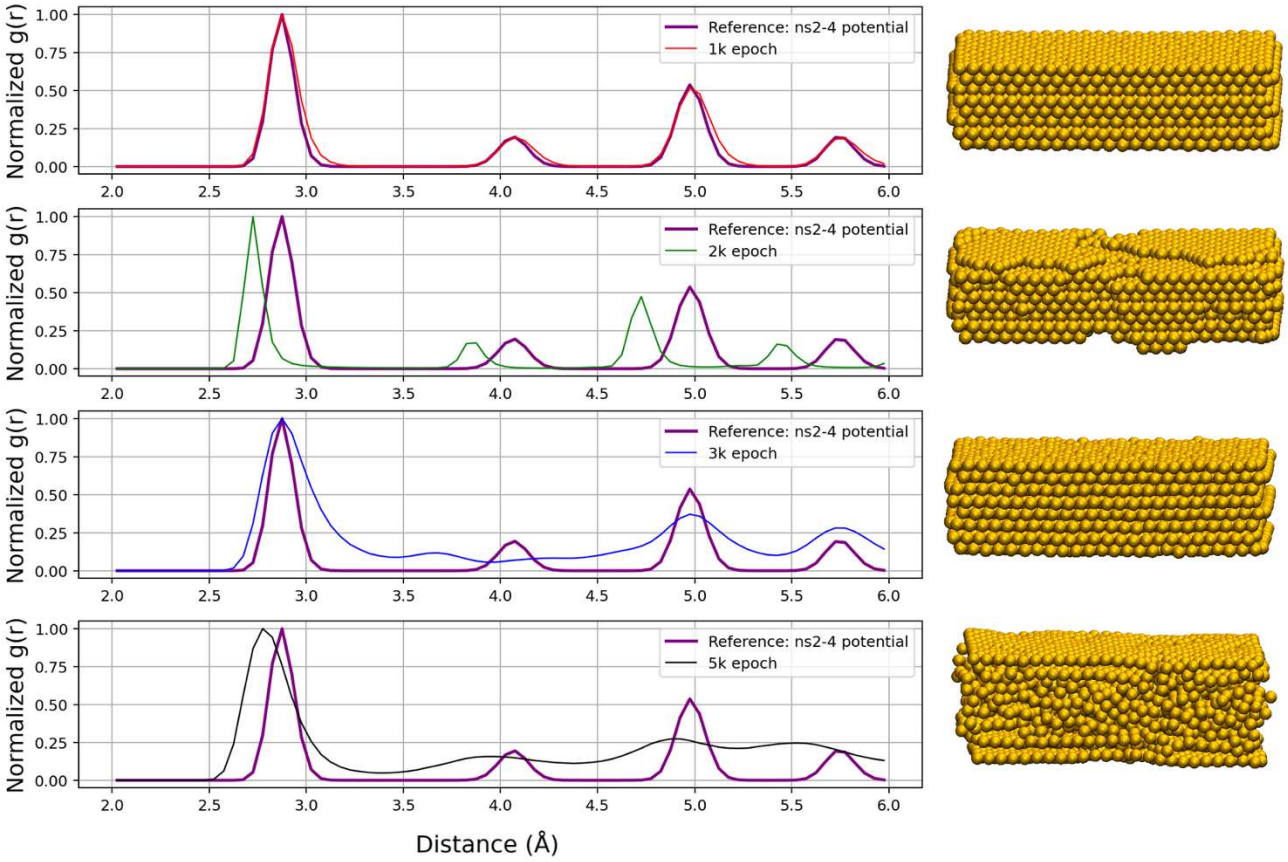


Figure C-13: Comparison between radial distribution functions at 300 K. Different numbers of epochs are presented in red 1k, green 2k, blue 3k, and black 5k and compared to a reference the black line, which corresponds to the distribution from the previous model. The end of the simulation is also presented on the right-hand side of the figure.

So, to ensure that there is no stability for 2k and 3k, we did a test using different parameters. We performed the learning this time per atom for the energy instead of the whole system, and we used the same analysis. First, by plotting (see Figure C-14) both learning, we are getting an MAE for 2k at 0.061 kcal/mol per atom and 0.920 kcal/mol/Å for energies and forces, respectively. For 3k, 0.057 kcal/mol per atom and 0.910 kcal/mol/Å are found. It means that using the learning per atom approach allows us to get a better accuracy even on the forces that are finally going down to the chemical accuracy.

Both presented learning procedures, then giving more or less the same MAE, but as before, to ensure the stability of the surface, we did perform MD simulations on it again at 300K. By plotting the distribution radial function, we can see, as presented in Figure C-15, that going up to 3k does not give accurate distances. Contrary to the previous 2k model, it seems to give a more convenient model for the surface.

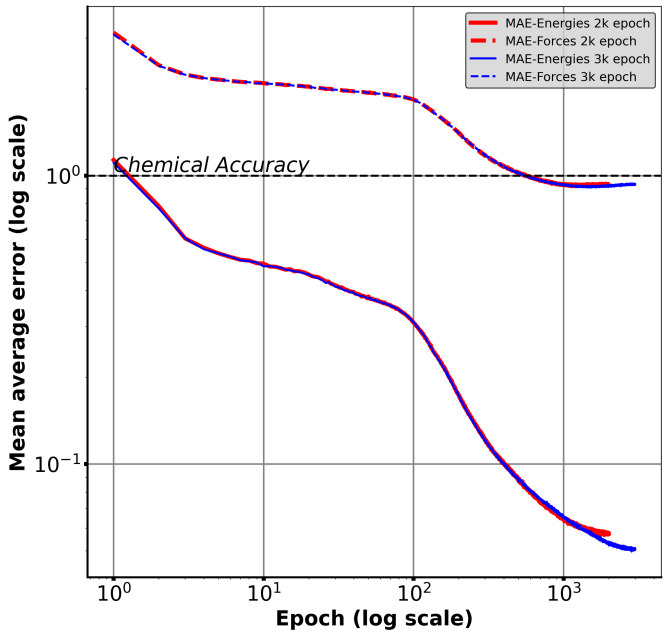


Figure C-14: Mean average errors for forces and energies for ns2-ns4 transfer learning with two different numbers of epochs. In red and in blue, the number of epochs is respectively 2k and 3k.

Nevertheless, by looking directly at the structure, one should notice that there is a problem of stability within the layers with a loss of signal at high distances. Finally, as discussed before, we have too smooth functions between peaks, meaning that the surface of gold is not as stable enough as it should be.

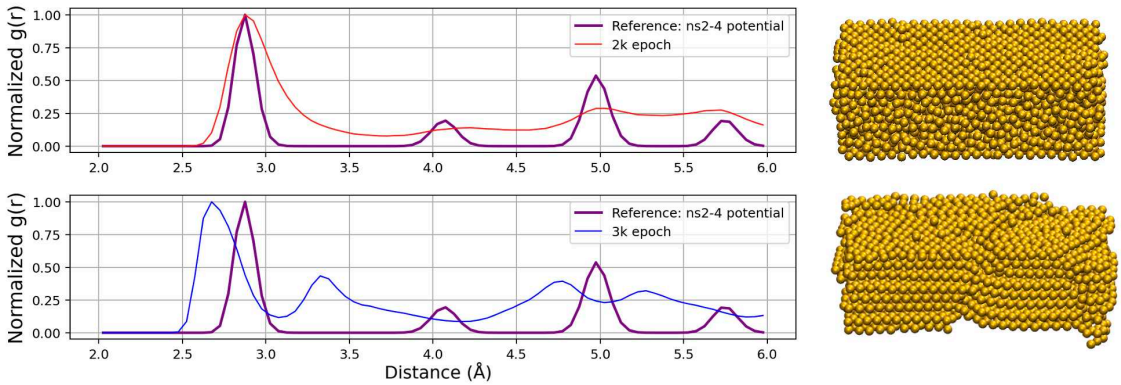


Figure C-15: radial distribution function for the transfer learning potential containing ns2 and ns4 with different number of epochs. The red line is 2k epoch, and the blue line is 3k. The reference is in purple and corresponds to the potential from the previous potential only using ReaxFF geometries.

This means that going up to 2k does not give a stable surface. Moreover, we saw that doing the learning per atom for energies can give better accuracy for the potential. From that, we narrowed our number of epochs from 1000 to 1500, which should avoid overfitting the DFT dataset, and we used the learning per atom for the energies. Doing that, our new potentials should be able to fit better with the DFT reference and keep the stability of the surface.

Three different potentials were then analyzed using 1000, 1250, and 1500 epochs, respectively. Their respective MAE for energies and forces are given in Table C-2. As expected, the value of the errors for the forces is going lower when one is fitting the energies per atom. Meaning that we are now reaching more accurate potential.

Table C-3: Mean average error of the energies (per atom) and forces for different numbers of epochs.

Number of epochs	MAE Energies *(per atom) (kcal/mol)*	MAE Forces *(kcal/mol/Å)*
1000	0.070	0.973
1250	0.066	0.940
1500	0.064	0.927

By comparing DFT energies and forces data to the results of the models, we observe a good linear regression for all three learnings, with MAE going down to chemical accuracy. To ensure the quality of each potential, we then validate them by looking at the distances of the Au_{20} pyramid-like structure. If we can get better distances than those obtained using ReaxFF and close to DFT reference, let's say with a deviation of 0.01 Å, this first test would show that the potential may be accurate for gold systems from small clusters to surfaces. Thus, it could be able to accurately model nanoparticles.

To study further the quality of the potential, we took one vertex atom of the pyramid and increased its distance from the cluster. At each point, we calculate using DFT single points to get reference data. To have good accuracy, we changed the spin from singlet to triplet at long-range (around 5 Å). We did the same single-point calculations with ReaxFF and our FeNNol potentials. As we can see in Figure C-16, for the dissociation energy, we are fitting much better than ReaxFF, and we are close to the DFT reference. The equilibrium distance is also pretty close to DFT. By looking at the different profiles, the 1250 epoch model seems to have the most appropriate potential.

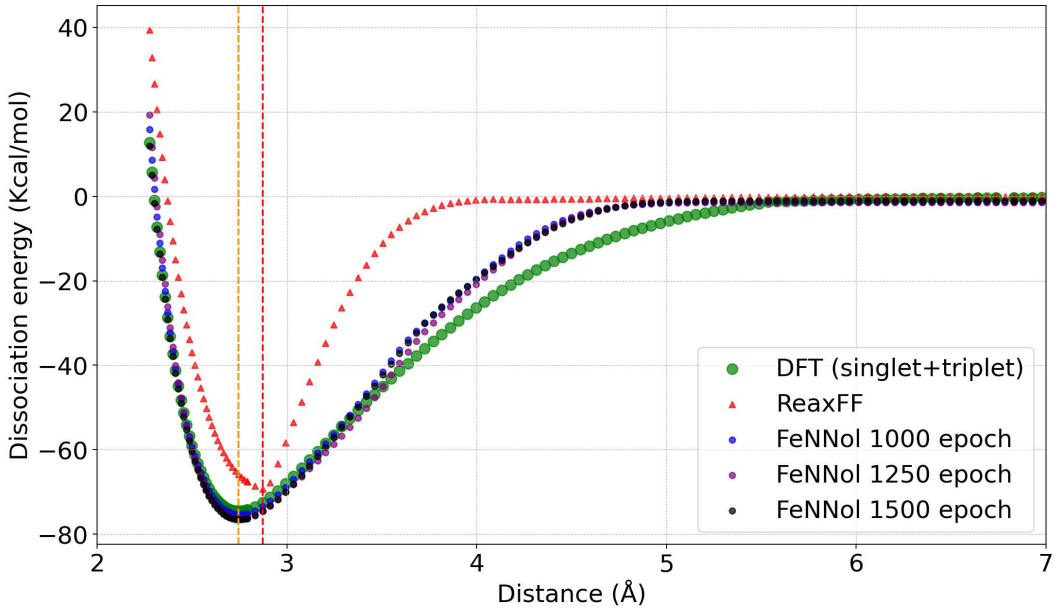
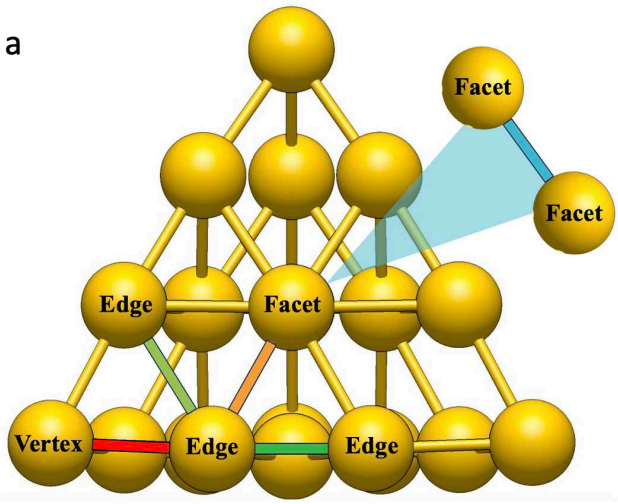


Figure C-16: Potential energy surface of the dissociation of one vertex gold atom from Au_{20} pyramid-like structure with different potentials having different numbers of epochs along with DFT and ReaxFF profiles. The dashed lines correspond to the equilibrium distance for DFT and FeNNol in orange and for ReaxFF in red.

In particular, the FeNNol potentials are in better agreement between 3 and 4 Å, which should allow a better agreement for cluster geometries where the distances may strongly vary. Moreover, FeNNol is capturing the equilibrium distance with no deviation at 2.74 Å.

To verify that, we also optimized the structure with our potentials and investigated the different distances on the Au_{20} cluster, as discussed in the previous section, since it depends on where the gold atom is placed, we have different distances. We considered then all the atoms on the vertex, on the edges, and the facets. In disagreement with DFT, ReaxFF predicts all distances as being identical, with an average distance of 2.88 Å. However, on the ML side, we can see we have different distances for each bond that matches with DFT with a deviation limited to about 0.01 Å. We are then assured that our model is discriminating each pair of atoms and can give a good PES dissociation profile.

a



b

	ReaxFF	1000 epoch	1250 epoch	1500 epoch	DFT
Vertex-edge	2.88	2.75	2.75	2.75	2.76
Edge-facet	2.88	2.85	2.85	2.85	2.86
Facet-Facet	2.88	3.08	3.06	3.14	3.09
Edge-Edge	2.88	2.70	2.71	2.70	2.71
Edge-Edge facet	2.88	2.99	2.99	2.98	3.00

Figure C-17: (a) schematic of the Au20 pyramid-like structure (b) distances for each pair of atoms

Moreover, the 1000 epoch potential is giving the most accurate distances compared to DFT with a difference of ± 0.01Å. In conclusion, we are finally able to obtain the potential to reproduce DFT results for clusters while maintaining a stable conformation for surfaces. Molecular dynamics can, therefore, easily be used with this potential to simulate different conditions with a high accuracy.

C.4.c) Cluster potential

We succeeded in getting a flexible potential that can compute both surfaces and clusters. Nevertheless, even if the distances for the Au_{20} pyramid-like structure are accurate, we wanted to see if we could enhance the description for the clusters by only using the DFT reference data. We then focused only on the clusters, and we trained a new potential without using the ReaxFF dataset. The training from more than 16k different DFT-based clusters is done, and the results are displayed in Figure C-18. First, it appears that the model learns more from the forces than from the energies, but there is a cross-section at the end where the energies are going down to the chemical accuracy. Different numbers of epochs have been used, and 3k seems to be enough to get accurate energies with an MAE of about 0.612 kcal/mol for energies (*not per atom*) and 1.120 kcal/mol/Å for forces. By comparing with the DFT calculations, we see that we have a good agreement between DFT and our model.

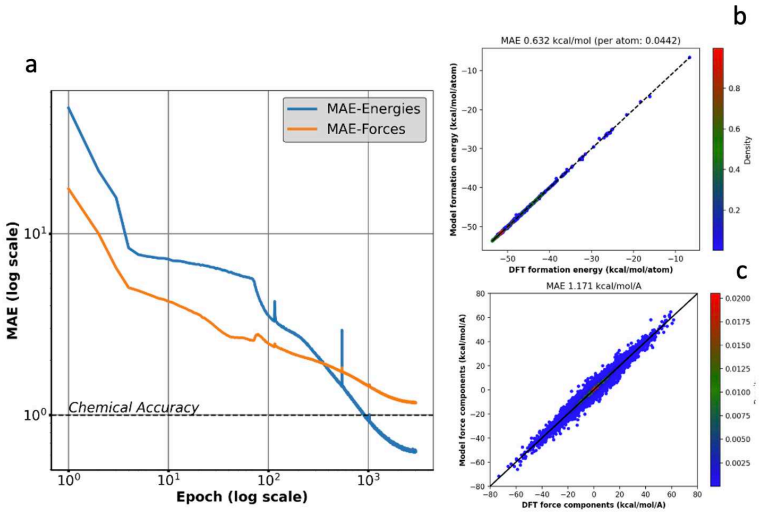


Figure C-18: (a) learning profile showing the MAE compared to the number of epochs. (b) Comparison between formation energy of the model and DFT. (c) Comparison between forces from the model and DFT.

Therefore, here also, we can reproduce DFT accuracy using our model. To validate the model, we did an MD simulation on non-encountered structures using different geometries from small to large clusters that contain more than 25 atoms.

We compared the radial distribution of each pair of atoms to see if our model could differentiate between each gold atom. As previously, we studied the Au_{20} clusters. ReaxFF does not predict the same profile as the DFT for short and long-range interaction, and its equilibrium distance is around 2.88 Å. On the opposite, our model follows the dissociation energy profile until 3.55 Å, which is the distance cutoff for the training as depicted in Figure C-19. Moreover, the equilibrium distance is the same at 2.74 Å. It means that for this cluster, the model reproduced well the DFT energetic. The next validation step is to see if it is possible to differentiate specifically the different distances depending on where the gold atom is located.

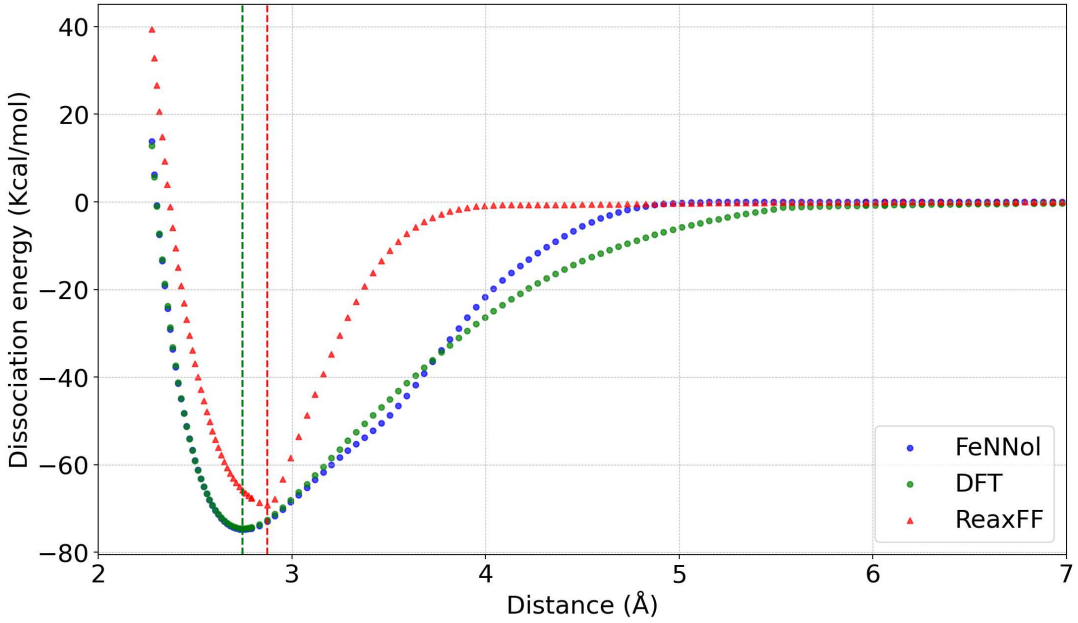


Figure C-19: Potential energy surface of the dissociation of one vertex gold atom from Au_{20} pyramid-like structure comparing FeNNol potential, DFT, and ReaxFF. The dashed lines correspond to the equilibrium distance for DFT and FeNNol in green and for ReaxFF in red.

The gold atom in a cage-like cluster can occupy three distinct sites: the vertex, the edge, and the facet, as depicted in Table C-3. A thorough description of these sites is crucial for studying the reactivity of such clusters. Additionally, in regular bulk-like clusters, which exhibit surface reactivity, these sites must also be accurately represented in the model. To ensure our model's capability to match with DFT, geometry optimization of the Au_{20} cluster was performed using all

three methods. DFT highlights the significance of atomic distances within the cluster, showing variations between different sites (see Table C-3). As anticipated, ReaxFF predicts all gold atoms to have the same distance, which contrasts with our observations. Our model aligns well with DFT, particularly for the equilibrium bond lengths. The FeNNol potential accurately matches the distances for facet-facet and edge-edge gold atoms, with other site predictions showing an error margin of approximately ± 0.01 Å. The ability of the FeNNol potential to distinguish between gold atoms is a critical feature, as it enables the detection of different reactivities on the cluster's surface, a capability not shared by ReaxFF.

Table C-4: Comparison between different methods based on the distances for each pair of atoms for the optimized Au_{20} pyramid-like structure.

	ReaxFF	FeNNol	DFT
Vertex-edge	2.88	2.75	2.76
Edge-facet	2.88	2.85	2.86
Facet-Facet	2.88	3.09	3.09
Edge-Edge	2.88	2.71	2.71
Edge-Edge facet	2.88	2.99	3.00

We carried out another validation test using Au_7 and Au_{27} to see if the model was capable of getting accurate distances for a small cluster and, more specifically, for the larger cluster that had not been seen during the training set (up to 25 atoms).

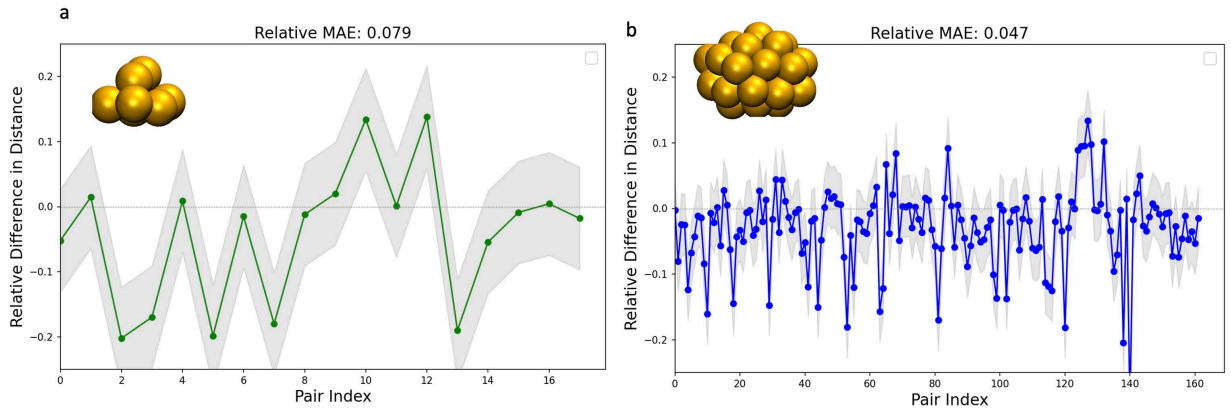


Figure C-20: Difference in terms of distances between DFT and FeNNol between each pair of gold atoms (a) cluster composed of 7 gold atoms (b) cluster with 27 gold atoms.

By looking at the relative difference of distances between each pair of atoms, we can see that we are close to the DFT results as depicted in Figure C-20, with a mean absolute error of 0.047 Å for the Au_{27} and 0.079 Å for the Au_7.

Then, we did a test on the formation energy for clusters from 2 to 20 gold atoms to see if our model could fit with DFT compared to ReaxFF. Geometry optimization was performed, and the formation energies are depicted in Figure C-21.

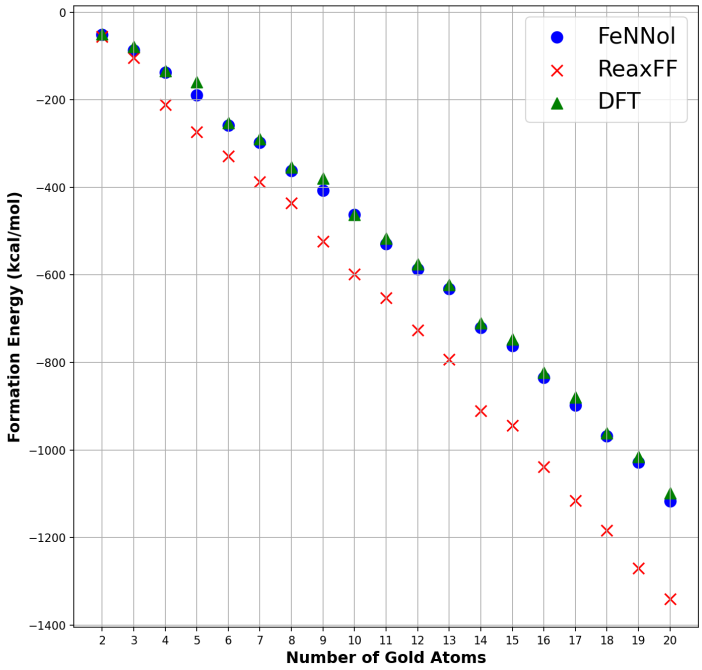


Figure C-21: Comparison of the formation energy for different optimized gold clusters between FeNNol, ReaxFF in red, and DFT in green.

These results demonstrate that we are then able to get a proper potential that fits perfectly well with the DFT energy. The optimization with FeNNol was pretty fast, and when one wants to do MD as it uses GPU, it enhances the speed.

Conclusion

Overall, we employed the FeNNol module alongside classical and quantum mechanical methods to develop a flexible machine-learning potential capable of accurately simulating gold nanoparticle systems. Our investigations revealed the effectiveness of integrating molecular dynamics (MD) simulations with density functional theory (DFT) calculations to refine our understanding of cluster dynamics and surface interactions. By focusing on a range of cluster sizes, we optimized the balance between computational demand and simulation accuracy. Our findings demonstrate that the FeNNol potential, when calibrated against both ReaxFF and DFT datasets, provides a robust tool for predicting the energetics and structural dynamics of gold clusters. Notably, our approach allowed for the differentiation between gold atoms within a structure, a crucial factor for advancing our understanding of cluster reactivity and stability. Radial distribution function analyses further supported the stability and accuracy of the potentials across different learning parameters.

Further analysis will be performed to determine if our potential can predict the effect of higher temperatures up to the boiling point. The next step is to include thiolates to see if we can predict staple formations. We aim to gather a sufficient number of structures to capture all the different complex reactivities. The use of a large database should speed up the learning of our model, and we know that some researchers are currently gathering a lot of crystal structures.[38] In parallel, more detailed simulations will be included to enhance the accuracy of the potential. For example, we plan to include CCSD simulations to improve the description of the bonding between Au-S, S-S, and S-C. Learning over charges will also be included to enhance the prediction accuracy of our model. Moreover, we want to include solvent effects. Our objective is to develop a highly accurate potential for gold capped with thiolates-SAMs.

In the future, this potential could be used with more ligands, enabling applications such as studying the nanomaterial/biomolecule interface.

This process has been fully automated, and the same process will be applied to silver. As explained in section B, this method could predict the sulfidation of the nanoparticle. Another idea is to use active learning, which could enhance the accuracy of our different potentials. Copper is also one of our objectives to predict for all three metals. However, it is more complicated than gold and silver due to its multiple oxidation states and lower stability.

References

(1) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106* (5), 1589–1615. https://doi.org/10.1021/cr040426m.

(2) Shamsi, Z.; Cheng, K. J.; Shukla, D. Reinforcement Learning Based Adaptive Sampling: REAPing Rewards by Exploring Protein Conformational Landscapes. *J. Phys. Chem. B* **2018**, *122* (35), 8386–8395. https://doi.org/10.1021/acs.jpcb.8b06521.

(3) Baek, H.; Moon, J.; Kim, J. Benchmark Study of Density Functional Theory for Neutral Gold Clusters, Au_n (n = 2–8). *J. Phys. Chem. A* **2017**, *121* (12), 2410–2419. https://doi.org/10.1021/acs.jpca.6b11868.

(4) McCandler, C. A.; Pihlajamäki, A.; Malola, S.; Häkkinen, H.; Persson, K. A. Gold–Thiolate Nanocluster Dynamics and Intercluster Reactions Enabled by a Machine Learned Interatomic Potential. *ACS Nano* **2024**, *18* (29), 19014–19023. https://doi.org/10.1021/acsnano.4c03094.

(5) Mishin, Y. Machine-Learning Interatomic Potentials for Materials Science. *Acta Materialia* **2021**, *214*, 116980. https://doi.org/10.1016/j.actamat.2021.116980.

(6) Ladygin, V. V.; Korotaev, P. Yu.; Yanilkin, A. V.; Shapeev, A. V. Lattice Dynamics Simulation Using Machine Learning Interatomic Potentials. *Computational Materials Science* **2020**, *172*, 109333. https://doi.org/10.1016/j.commatsci.2019.109333.

(7) Hagan, M. T.; Demuth, H. B.; Beale, M. H.; Jesús, O. D. *Neural Network Design*; Martin Hagan, 2014.

(8) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401. https://doi.org/10.1103/PhysRevLett.98.146401.

(9) Batzner, S.; Smidt, T.; Sun, L.; Mailoa, J.; Kornbluth, M.; Molinari, N.; Kozinsky, B. SE(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials. March 1, 2021. https://doi.org/10.21203/rs.3.rs-244137/v1.

(10) Han, J.; Huang, W.; Xu, T.; Rong, Y. Equivariant Graph Hierarchy-Based Neural Networks. arXiv October 15, 2022. https://doi.org/10.48550/arXiv.2202.10643.

(11) Plé, T.; Lagardère, L.; Piquemal, J.-P. Force-Field-Enhanced Neural Network Interactions: From Local Equivariant Embedding to Atom-in-Molecule Properties and Long-Range Effects. *Chem. Sci.* **2023**, *14* (44), 12554–12569. https://doi.org/10.1039/D3SC02581K.

(12) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B. Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics. *Nat Commun* **2023**, *14* (1), 579. https://doi.org/10.1038/s41467-023-36329-y.

(13) Geiger, M.; Smidt, T. E3nn: Euclidean Neural Networks. arXiv July 18, 2022. http://arxiv.org/abs/2207.09453 (accessed 2024-06-19).

(14) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102* (7), 073005. https://doi.org/10.1103/PhysRevLett.102.073005.

(15) Plé, T.; Adjoua, O.; Lagardère, L.; Piquemal, J.-P. FeNNol: An Efficient and Flexible Library for Building Force-Field-Enhanced Neural Network Potentials. arXiv May 6, 2024. https://doi.org/10.48550/arXiv.2405.01491.

(16) Takamoto, S.; Izumi, S.; Li, J. TeaNet: Universal Neural Network Interatomic Potential Inspired by Iterative Electronic Relaxations. *Computational Materials Science* **2022**, *207*, 111280. https://doi.org/10.1016/j.commatsci.2022.111280.

(17) Unke, O. T.; Chmiela, S.; Gastegger, M.; Schütt, K. T.; Sauceda, H. E.; Müller, K.-R. SpookyNet: Learning Force Fields with Electronic Degrees of Freedom and Nonlocal Effects. *Nat Commun* **2021**, *12* (1), 7273. https://doi.org/10.1038/s41467-021-27504-0.

(18) Lalvani, H. 4D-Cubic Lattice of Chemical Elements. *Found Chem* **2020**, *22* (2), 147–194. https://doi.org/10.1007/s10698-019-09350-7.

(19) Shi, W.; He, B.; Pu, B.; Ren, Y.; Avdeev, M.; Shi, S. Software for Evaluating Long-Range Electrostatic Interactions Based on the Ewald Summation and Its Application to Electrochemical Energy Storage Materials. *J. Phys. Chem. A* **2022**, *126* (31), 5222–5230. https://doi.org/10.1021/acs.jpca.2c02591.

(20) Zhuang, J.; Tang, T.; Ding, Y.; Tatikonda, S. C.; Dvornek, N.; Papademetris, X.; Duncan, J. AdaBelief Optimizer: Adapting Stepsizes by the Belief in Observed Gradients. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc., 2020; Vol. 33, pp 18795–18806.

(21) Zaverkin, V.; Holzmüller, D.; Bonfirraro, L.; Kästner, J. Transfer Learning for Chemically Accurate Interatomic Neural Network Potentials. *Phys. Chem. Chem. Phys.* **2023**, *25* (7), 5383–5396. https://doi.org/10.1039/D2CP05793J.

(22) Blinder, S. M. Basic Concepts of Self-Consistent-Field Theory. *Am. J. Phys.* **1965**, *33* (6), 431–443. https://doi.org/10.1119/1.1971665.

(23) Laaksonen, L.; Pyykkö, P.; Sundholm, D. Fully Numerical Hartree-Fock Methods for Molecules. *Computer Physics Reports* **1986**, *4* (5), 313–344. https://doi.org/10.1016/0167-7977(86)90021-3.

(24) Bartlett, R. J.; Stanton, J. F. Applications of Post-Hartree—Fock Methods: A Tutorial. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley, 1994; Vol. 5, pp 65–169. https://doi.org/10.1002/9780470125823.ch2.

(25) Knowles, P. J.; Handy, N. C. A Determinant Based Full Configuration Interaction Program. *Computer Physics Communications* **1989**, *54* (1), 75–83. https://doi.org/10.1016/0010-4655(89)90033-7.

(26) Wu, X.; Su, P. Chapter 1 - Very Brief Introduction to Quantum Chemistry. In *Quantum Chemistry in the Age of Machine Learning*; Dral, P. O., Ed.; Elsevier, 2023; pp 3–25. https://doi.org/10.1016/B978-0-323-90049-2.00006-8.

(27) Lipkowitz, K. B.; Boyd, D. B. *Reviews in Computational Chemistry, Volume 14*; John Wiley & Sons, 2009.

(28) Sholl, D. S.; Steckel, J. A. DENSITY FUNCTIONAL THEORY. 253.

(29) Eschrig, H. *The Fundamentals of Density Functional Theory*; Ebeling, W., Pilkuhn, M., Wilhelmi, B., Series Eds.; TEUBNER-TEXTE zur Physik; Vieweg+Teubner Verlag: Wiesbaden, 1996; Vol. 32. https://doi.org/10.1007/978-3-322-97620-8.

(30) Engel, E.; Dreizler, R. M. *Density Functional Theory: An Advanced Course*; Theoretical and Mathematical Physics; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011. https://doi.org/10.1007/978-3-642-14090-7.

(31) Zhang, I. Y.; Xu, X. On the Top Rung of Jacob's Ladder of Density Functional Theory: Toward Resolving the Dilemma of SIE and NCE. *WIREs Computational Molecular Science* **2021**, *11* (1), e1490. https://doi.org/10.1002/wcms.1490.

(32) Hofmann, M.; Schaefer, H. F. Computational Chemistry. In *Encyclopedia of Physical Science and Technology (Third Edition)*; Meyers, R. A., Ed.; Academic Press: New York, 2003; pp 487–506. https://doi.org/10.1016/B0-12-227410-5/00129-0.

(33) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006**, *27* (15), 1787–1799. https://doi.org/10.1002/jcc.20495.

(34) Kepp, K. P. Benchmarking Density Functionals for Chemical Bonds of Gold. *J. Phys. Chem. A* **2017**, *121* (9), 2022–2034. https://doi.org/10.1021/acs.jpca.6b12086.

(35) *Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional | The Journal of Chemical Physics | AIP Publishing*. https://pubs.aip.org/aip/jcp/article/110/11/5029/183578/Assessment-of-the-Perdew-Burke-Ernzerhof-exchange (accessed 2024-06-24).

(36) *Quantum Cluster Database*. https://muellergroup.jhu.edu/qcd/ (accessed 2024-07-23).

(37) Monti, S.; Carravetta, V.; Ågren, H. Simulation of Gold Functionalization with Cysteine by Reactive Molecular Dynamics. *J. Phys. Chem. Lett.* **2016**, *7* (2), 272–276. https://doi.org/10.1021/acs.jpclett.5b02769.

(38) Pihlajamäki, A.; Malola, S.; Häkkinen, H. Creating a Catalog of the Crystal Structures of Monolayer-Protected Clusters. May 24, 2024. https://doi.org/10.26434/chemrxiv-2024-kxjpb.

General conclusion

Throughout this study, we demonstrated that the current ReaxFF parametrization for the AgSCH potential, developed by Dulong et al., accurately predicts stable interactions between thiolates-SAMs. However, it was found that for butane thiolates, the energetically favored conformation (77) was not distinct from less stable assemblies (33), indicating that ReaxFF may not precisely calculate the correct energy for each conformation. Additionally, no spontaneous formation of silver sulfides was observed, suggesting a limitation in the model's ability to predict nanoparticle surface restructuring, although small systems like ns2 and ns4 showed some destabilization.

For larger nanoparticles, this destabilization was less pronounced due to a higher proportion of facets relative to edges. Experimental preferences for the on-top site and sulfur-head group distances in the 77 assembly were accurately predicted, confirming the potential's reliability in simulating butanethiolates-SAM on silver. However, adjustments are necessary to improve energy descriptions across different systems.

Regarding methanethiolates-SAMs, the stability predictions were inaccurate, potentially due to the same parametrization issues. Preliminary tests on gold using three different potentials did not allow for a detailed analysis due to these parametric constraints.

To address these shortcomings, we employed machine learning with the FeNNol module package, facilitating the creation of tailored potentials for various systems. This involved preparing a dataset from theoretical studies and databases, leading to the generation of 15,430 non-redundant structures. These were subjected to single-point DFT calculations to obtain reference energies and forces, which informed the machine learning model.

The model achieved chemical accuracy within 0.066 kcal/mol for energies and 0.940 kcal/mol/Å for forces. By comparing the radial distribution function of the surfaces to stable reference surfaces, we confirmed that using 1000 epochs during training led to stable surface predictions. The potential also closely matched DFT distances in an Au_{20} pyramid-like structure within 0.01 Å, surpassing the accuracy of ReaxFF, especially in differentiating gold atoms based on their positional context.

In conclusion, while the existing ReaxFF parametrizations provided a foundation, the integration of machine learning significantly enhanced the accuracy and applicability of our simulations. In the coming months, further tests are planned to determine if our model can effectively differentiate between various isomers. Additionally, we aim to incorporate thiolates into our simulations. This step is crucial as it will enable us to refine the current gold potential into a more flexible model capable of capturing effects such as surface restructuring. Once this method is validated, a similar approach will be applied to silver to address more complex reactivities, such as sulfidation. These advancements will not only enhance the predictive power of our simulations but also broaden their applicability in understanding and designing nanomaterial interfaces.

Résumé en français

Les nanoparticules (NPs) se sont imposées comme un pilier de la nanotechnologie en raison de leurs propriétés physiques, chimiques et biologiques distinctives. Ces particules, souvent composées de métaux de transition, varient de groupes de deux à des milliers d'atomes. Cette thèse explore la simulation de ces nanoparticules, en se concentrant spécifiquement sur les métaux précieux tels que l'or, l'argent et le cuivre. Ces métaux ont de nombreuses applications, comme l'argent en tant qu'agent antibactérien ou l'or en tant qu'agent anticancéreux. Pour l'argent, les ions se propagent dans la bactérie en interagissant directement avec son ADN et sa mitochondrie. Néanmoins, une trop forte concentration d'ions dans le corps peut entraîner des effets indésirables. Pour limiter et contrôler ces effets, on peut utiliser le thiolate. En raison de sa forte interaction covalente avec l'argent (ou l'or) et le soufre, cela favorise la stabilité de l'interaction. Cette interaction se manifeste sous la forme d'une monocouche auto-assemblée où le ligand forme une couche dense sur toute la surface. Différentes structures ont été trouvées en fonction du métal étudié. Pour l'or, deux structures existent : une avec une simple adsorption sur la surface à différents sites appelée la $\sqrt{3x3}R30°$ (33). Cette appellation provient de la notation de Wood's qui permet de définir l'assemblement en fonction de l'endroit où le substrat s'adsorbe. De plus, une autre structure a été trouvée, accompagnée cette fois d'une restructuration de la surface. Certains atomes d'or sortent de la surface en cassant leur liaison pour se placer entre les atomes de soufre. Cette structure est très stable et favorisée par les défauts présents sur la surface d'une nanoparticule, par exemple. Quant à l'argent, en fonction de la concentration de thiol en interaction, cela peut soit s'adsorber normalement et créer la structure $\sqrt{7x7}R19.1°$ (77), soit se restructurer et former ce que l'on appelle un effet de core@shell Ag_2S autour de la nanoparticule. Cette étude est centrée sur ces différents systèmes. L'objectif est de déterminer si les méthodes actuelles permettent de prédire ces phénomènes de restructuration et de stabilité de cette monocouche auto-assemblée.

La dynamique moléculaire est utilisée dans un premier temps car nous voulons étudier des systèmes assez larges qui ne sont pas modélisables avec des méthodes de chimie quantique. Nous avons utilisé un champ de force réactif appelé ReaxFF. Ce champ de force permet, pendant une simulation, la cassure et la formation de liaisons, ce qui permet d'observer la restructuration de la surface lorsque les thiolates sont déposés sur la surface. Dans un premier temps, un logiciel appelé SAM-Maker a été créé pour permettre à un utilisateur de créer les différentes structures. Il fait

maintenant partie du package NATOMOS, un outil d'analyse et de création de structures métalliques qui sera intégré à TINKER. Ainsi, nous avons créé deux structures d'argent : celle qui a été trouvée expérimentalement, la 77, et celle pour l'or, 33, en utilisant du méthanethiolate et du butanethiolate. Cela permettra de voir si le potentiel actuel est capable de prédire énergétiquement que la structure stable pour l'argent est bien la 77. Nous voulions aussi étudier l'impact de la longueur de la chaîne sur la stabilité de notre monocouche.

Nous avons utilisé le potentiel AgSCH créé par Dulong et al. Nous avons réalisé des dynamiques en chauffant une boîte de simulation de 0 à 300K pour observer la stabilité de la simulation pendant cet état de chauffe. Différents temps de simulation ont été pris en compte : 0.3, 1 et 3 ns avec un pas de temps de 0.25 fs. Une étape de vérification de la stabilité a aussi été introduite tous les 50 K pour vérifier la stabilité de notre dynamique. À la fin de la chauffe, une simulation supplémentaire est également effectuée pour vérifier que le système est stable à 300 K. Nous avons démontré ici que l'approche ReaxFF est capable de produire une monocouche stable de 77, typique des butanethiolates sur la surface Ag (111), démontrant son efficacité pour simuler et prédire les comportements d'adsorption de surface. Il est intéressant de noter que ReaxFF, utilisé dans un article récent pour étudier les NP d'or, a montré des comportements complètement différents avec la formation d'agrafes sur les surfaces Au (111). La formation de ces agrafes correspond bien à l'expérience et au calcul DFT. Cela montre que ReaxFF produit des propriétés différentes pour les deux matériaux, en contraste avec d'autres champs de force. Dans un article récent, il a également été démontré que ReaxFF était capable de reproduire la formation d'agrafes sur les surfaces Au (111). Ainsi, il est possible avec ReaxFF de distinguer les deux matériaux, ce qui est généralement assez difficile pour les champs de force puisque les deux métaux sont caractérisés par la même géométrie. Cependant, nous avons observé que la monocouche de 77 de méthanothiolate n'est pas stable pour des simulations très longues. De plus, la monocouche de butanethiolate n'est pas énergétiquement favorisée par rapport aux autres SAMs. Cela montre les limites de la paramétrisation actuelle du modèle. De plus, la surface reconstruite habituellement observée pour l'argent est constituée d'adatomes d'argent sur une surface Ag(111). Nos simulations ne montrent même aucune perturbation de la couche d'argent en contact avec les thiolates, ce qui indique la bonne stabilité de ces monocouches non restructurées. Il serait cependant intéressant à l'avenir d'étudier des surfaces d'argent déjà restructurées avec des thiolates ; l'idée serait de comparer les énergies de liaison de ces systèmes avec celles obtenues ici pour les

systèmes non restructurés. Cela nous permettrait de voir lequel est énergétiquement stable et si une restructuration peut être expérimentée. Les simulations montrent également que des assemblages denses tels que les SAMs de 77 trouvés sur la surface Ag (111) devraient apparaître sur les NP d'argent. À notre connaissance, il n'y a pas d'expériences disponibles étudiant cette question. Lors des simulations, aucune décomposition de l'alkanothiolate n'a été trouvée. Les origines du sulfure d'argent observé dans certaines expériences restent une question ouverte. À l'avenir, cette question devrait également être étudiée pour des surfaces restructurées et d'autres ligands. Aucune restructuration spontanée des surfaces des NP n'a été trouvée, sauf pour les petits NP inférieurs à 4 nm. Une restructuration globale des surfaces d'argent, comme observée dans les expériences, n'est pas réalisable avec les simulations ReaxFF. À l'avenir, nous prévoyons d'effectuer des simulations à partir d'une surface restructurée et de comparer ses énergies de liaison avec celles trouvées ici. Deux sites d'adsorption sur les facettes des NP d'argent ont été trouvés : un en position sur le dessus et un autre au site de pont, le premier étant préféré ; cela est en bon accord avec les expériences récentes. Concernant la distance entre les ligands du groupe S, nous avons obtenu la même valeur que d'autre études expérimentales et n'avons observé aucune différence entre les facettes et les bords.

La même méthode a donc été envisagée pour l'or avec l'idée qu'il serait possible de prédire la stabilité de la restructuration. Néanmoins, même en ayant utilisé les trois potentiels existants, il nous a été impossible d'obtenir de bons résultats montrant ce phénomène. C'est pourquoi nous avons envisagé l'utilisation des algorithmes d'apprentissage pour créer un nouveau potentiel capable de prédire avec plus de précision les différents systèmes étudiés.

FeNNol est un programme Python qui permet à un utilisateur de facilement créer son propre potentiel à partir de données qu'il aura créées au préalable. La partie la plus difficile de cette étude a été de réunir les différentes structures entre les différentes publications et banques de données. Au total, 300 structures ont été réunies, avec des clusters allant de 3 à 25 atomes d'or. Pour augmenter ce nombre et obtenir des structures non équilibrées, de la dynamique moléculaire a été effectuée. L'utilité d'avoir des structures hors équilibre permet au potentiel créé d'être beaucoup plus flexible en termes de réactivité qui pourrait se produire durant la dynamique. À partir des 300 structures, les dynamiques ont donc été faites en chauffant le système. Puis, un filtre est appliqué pour assurer qu'il n'y a pas de structure redondante. Ce filtre est basé sur une analyse

radiale de toutes les paires d'atomes, comparant les différentes géométries de chaque simulation. Au total, 15 946 structures uniques ont été créées. À partir de cela, des calculs DFT ont été faits en utilisant la fonctionnelle PBE et la base Stuttgart qui prend en compte les pseudo-potentiels et les effets relativistes. Depuis ces calculs, nous avons extrait les énergies et les forces de référence pour chaque système, qui serviront de référence pour le potentiel. Dans un premier temps, pour réussir à obtenir un potentiel flexible entre les clusters et les surfaces, un préapprentissage est effectué. Nous voulions des énergies et des références de surfaces d'une méthode qui permet de simuler des surfaces stables. Pour cela, nous avons utilisé ce que nous connaissions de ReaxFF, qui est capable de générer des surfaces d'or pur très stables. Deux tailles de nanoparticules ont donc été utilisées : 2 nm et 4 nm, donnant un total de 20 000 structures de référence. Pour s'assurer que les structures n'étaient pas trop similaires, la dynamique sur ces systèmes a été effectuée à très haute température, approchant la température de fusion de l'or. Le premier apprentissage basé sur les valeurs de ReaxFF est donc effectué. En analysant la fonction de distribution radiale, nous avons pu constater que le potentiel reproduisait bien les distances observées pour une surface parfaite, ce qui signifie que le potentiel est capable de reproduire ReaxFF. Néanmoins, cela ne suffit pas pour avoir un potentiel flexible sur les clusters, car ReaxFF n'est pas capable d'avoir une bonne précision pour les petits systèmes comme les très petits clusters. Pour cela, nous avons affiné notre modèle avec les références DFT calculées au préalable. Différents tests de paramètres ont été effectués et, après quelques optimisations, nous avons pu trouver le meilleur modèle. Il est capable d'avoir une surface stable et de décrire avec précision DFT les clusters. Pour s'en assurer, une analyse énergétique a été réalisée. Nous avons étudié le cluster Au20, connu pour avoir une forme pyramidale. Un atome d'or est donc arraché de ce cluster et l'énergie de dissociation est calculée pour chaque distance. Notre modèle est bien en accord avec la DFT sans aucun décalage pour la distance d'équilibre. N'ayant pas pris en considération les effets à longue portée pendant l'apprentissage, nous avons encore un peu de mal à décrire ce qui se passe à plus de 3.75Å. Cependant, cela reste très précis autour de l'équilibre. De plus, pour s'assurer de sa précision, nous avons optimisé le cluster et comparé les distances avec le même cluster pyramidal optimisé par DFT et par ReaxFF. En regardant chaque distance associée, nous avons constaté que, contrairement à ReaxFF, nous arrivons à discriminer chaque atome d'or en fonction de sa position avec une précision de 0.01 Å. D'autres tests sont à venir pour tester différents isomères et voir si notre modèle arrive à prédire leurs énergies précisément. Nous avons également créé un autre modèle

plus précis mais utilisable seulement pour les systèmes comprenant moins de 220 atomes d'or, car au-delà une chimie de surface peut apparaître. Ce potentiel est basé uniquement sur les références DFT et est capable de déterminer l'énergie de chaque cluster en accord avec la DFT. De plus, la courbe de dissociation est d'autant plus précise. Il est aussi capable de prédire des clusters qu'il n'a pas vus, avec plus de 25 atomes d'or, et d'avoir les bonnes distances et énergies.

En conclusion, cette thèse a démontré l'efficacité des simulations basées sur le champ de force ReaxFF pour prédire la stabilité des monocouches auto-assemblées à base butanethiolates sur l'argent. L'utilisation d'algorithmes d'apprentissage a permis de créer des potentiels plus précis pour l'or, ouvrant la voie à des simulations encore plus fiables et détaillées. Ces avancées contribuent à une meilleure compréhension des interactions à l'échelle nano et offrent des perspectives intéressantes pour les applications futures des nanoparticules dans divers domaines.

Supplementary information

Table S1. Average distances between neighboring S head groups. The results are given for different silver nanocrystal diameters capped by butane thiolate. The values on the vertex, edges and facets are shown.

2 nm / SBu	
total	4.44 ± 0.01
vertex	4.39 ± 0.01
edge	4.39 ± 0.01
4 nm / SBu	
total	4.43 ± 0.01
vertex	4.37 ± 0.04
edge	4.34 ± 0.02
6 nm / SBu	
total	4.35 ± 0.00
vertex	4.34 ± 0.07
edge	4.22 ± 0.01
8 nm / SBu	
total	4.31 ± 0.00
vertex	4.32 ± 0.06
edge	4.23 ± 0.00
10 nm / SBu	
total	4.29 ± 0.01
vertex	4.41 ± 0.10
edge	4.26 ± 0.01

Table S2. Relative frequencies of the S head group as a function of their adsorption position for silver nanocrystals with butane thiolate.

2 nm / SBu		
vertex	edge	facet
16.99 ± 0.63	65.77 ± 1.06	17.26 ± 0.88
4 nm / SBu		
vertex	edge	facet
4.93 ± 0.23	39.37 ± 1.13	55.67 ± 0.96
6 nm / SBu		
vertex	edge	facet
2.20 ± 0.06	26.27 ± 0.44	71.57 ± 0.49
8 nm / SBu		
vertex	edge	facet
1.13 ± 0.09	21.53 ± 0.33	77.33 ± 0.41
10 nm / SBu		
vertex	edge	facet
0.67 ± 0.07	18.23 ± 0.18	81.07 ± 0.23

Table S3. Frequencies of silver atoms with a given number of neighbors for silver nanocrystals with butane thiolate. First line denoted by "total" gives the absolute number. All other values are given in %.

2 nm / SBu							
Neihbors	0	1	2	3	4	5	6
Frequency	0.44 ± 0.11	2.14 ± 0.41	4.59 ± 0.43	3.92 ± 0.20	1.69 ± 0.17	1.78 ± 0.29	6.47 ± 0.63
Neighbors	7	8	9	10	11	12	
Frequency	10.51 ± 0.79	8.37 ± 0.54	21.60 ± 0.44	3.70 ± 0.48	8.47 ± 0.35	23.99 ± 0.39	

4 nm / SBu							
Neihbors	0	1	2	3	4	5	6
Frequency	0.10 ± 0.00	0.30 ± 0.00	0.80 ± 0.10	0.53 ± 0.09	0.37 ± 0.03	0.30 ± 0.06	1.43 ± 0.18
Neighbors	7	8	9	10	11	12	
Frequency	4.73 ± 0.13	2.57 ± 0.15	24.87 ± 0.27	1.23 ± 0.09	1.03 ± 0.12	59.97 ± 0.17	

6 nm / SBu							
Neihbors	0	1	2	3	4	5	6
Frequency	0.00 ± 0.00	0.10 ± 0.00	0.20 ± 0.00	0.13 ± 0.03	0.03 ± 0.03	0.10 ± 0.00	0.40 ± 0.00
Neighbors	7	8	9	10	11	12	
Frequency	2.80 ± 0.00	0.73 ± 0.03	20.50 ± 0.06	0.83 ± 0.03	0.50 ± 0.06	71.73 ± 0.09	

8 nm / SBu							
Neihbors	0	1	2	3	4	5	6
Frequency	0.00 ± 0.00	0.10 ± 0.00	0.17 ± 0.03	0.10 ± 0.00	0.03 ± 0.03	0.03 ± 0.03	0.27 ± 0.03
Neighbors	7	8	9	10	11	12	
Frequency	1.50 ± 0.00	0.60 ± 0.00	16.47 ± 0.03	0.63 ± 0.03	0.23 ± 0.03	78.03 ± 0.07	

10 nm / SBu							
Neihbors	0	1	2	3	4	5	6
Frequency	0.00 ± 0.00	0.00 ± 0.00	0.10 ± 0.00	0.10 ± 0.00	0.03 ± 0.03	0.00 ± 0.00	0.17 ± 0.03
Neighbors	7	8	9	10	11	12	
Frequency	1.03 ± 0.03	0.40 ± 0.00	13.87 ± 0.03	0.50 ± 0.00	0.20 ± 0.00	81.87 ± 0.03	

Table S4. Frequencies of the S atoms in contact with a given number of silver atoms for silver nanocrystals with butane thiolate.

2 nm / SBu					
Neighbors	0	1	2	3	4
vertex	0.00 ± 0.00	12.51 ± 3.19	70.32 ± 4.78	15.87 ± 2.89	1.30 ± 0.71
edge	0.00 ± 0.00	13.98 ± 1.67	65.36 ± 1.78	15.82 ± 1.78	4.85 ± 0.76
facet	0.00 ± 0.00	47.19 ± 3.65	52.58 ± 5.68	0.23 ± 0.23	0.00 ± 0.00
centre	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

4 nm / SBu					
Neighbors	0	1	2	3	4
vertex	0.00 ± 0.00	15.71 ± 3.47	76.12 ± 3.06	6.12 ± 0.00	2.04 ± 2.04
edge	0.00 ± 0.00	26.44 ± 0.79	62.62 ± 2.44	7.12 ± 1.70	3.81 ± 0.53
facet	0.00 ± 0.00	59.94 ± 1.94	39.75 ± 0.75	0.31 ± 0.05	0.00 ± 0.00
centre	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

6 nm / SBu					
Neighbors	0	1	2	3	4
vertex	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	3.18 ± 1.36	96.82 ± 5.45
edge	0.00 ± 0.00	23.37 ± 0.69	56.69 ± 1.03	6.98 ± 0.57	12.96 ± 1.33
facet	0.00 ± 0.00	64.09 ± 1.91	35.67 ± 1.50	0.24 ± 0.13	0.00 ± 0.00
centre	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

8 nm / SBu					
Neighbors	0	1	2	3	4
vertex	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	100.00 ± 7.96
edge	0.00 ± 0.00	10.37 ± 0.33	24.19 ± 0.47	3.26 ± 0.93	62.19 ± 1.07
facet	0.00 ± 0.00	47.03 ± 0.66	27.91 ± 0.84	0.99 ± 0.16	24.07 ± 0.37
centre	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

10 nm / SBu					
Neighbors	0	1	2	3	4
vertex	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	100.00 ±10.45
edge	0.00 ± 0.00	2.19 ± 0.33	6.01 ± 0.93	2.73 ± 0.33	89.07 ± 0.33
facet	0.00 ± 0.00	24.49 ± 0.27	15.04 ± 0.18	1.48 ± 0.00	59.00 ± 0.37
centre	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00

#**********SAM_maker**********
#1: import SAMmaker function
#2: Launch the function by only specifying your periodicity and the number of layers that you need (ex: SAMmaker (2,8))
note that you need the geometry of your ligand which is a typical xyz file like the following;
5

S	0.00910	0.02410	-0.00000
C	-0.00001	-0.00001	1.82399
H	1.02220	0.18723	2.28677
H	-0.66575	0.71105	2.37279
H	-0.32292	-1.02249	2.22332

#3: Some questions will appear for you to specify the desire system
#==> example:
#==> <What kind of metal do you want for your surface? (Au or Ag) -> Ag
#==> <What conformation do you want? (33, 77) -> 77
#==> <Do you want the restructured surface? (y,n) --> n
#==> <Which ligand do you want on your restructured surface? (m or b) ---> m
#Then you'll get your xyz file with the desire surface.

#**********NTM_setup**********
#From your previous surface from SAM_maker

#step 1
#Specify $1 = program $2 = surface from SAM maker $3 = name of the output $4 = the mode $5 = number of atoms $6 = xyz file of the ligand used $7-$9 = simulation box
#It is creating a csam file which is used in order to create the NP
#exemple:
python3 ntm_setup060722.py Ag_SCH3_77_m_Unrestructured.xyz data.surf 5 400 Methane_Thiolate.xyz 20.09000 34.79690 300.0

#step 2
#Specify $1 = program $2 = the NP without ligands $3 = name of the output $4 = the mode $5 = number of atoms $6 = xyz file of the ligand used $7-$9 = simulation box
#exemple
python3 ../ntm_setup060722.py ../surface_xyz/nc6octa.xyz data.surf 2 4000 Methane_Thiolate.xyz 106 106 106

#**********NTM_ana**********
#specify $1 = program $2 = simu xyz at each timestep $3 = the NP without ligands $4 = the step that you want to analyze $5-$7= simulation box $8= the metal that you have
#exemple
python3.11 ntm_ana090623.py toto_reax.xyz ns6Ag287_77N.xyz 0 106 106 106 silver >> anaAg.dat
#The anaAg.dat is containing all the information for the specific simulation step

GUIDLINE FOR FENNOL

First you need to have miniconda3 (https://docs.conda.io/projects/miniconda/en/latest/install) on your system and activate the jax_username environment (source jax_username). You should have this (jaxnff) [username@clustername]

For the learning :

You should have a file .yaml structured as follow

```
device: cuda:1 #learning with gpu you can change that to device: cpu

output_directory: run_dir #directory where the info is stored

log_file: train.log #file where the train info is stored

print_timings: True


training:

  dspath: /home/adamML/ML_training/gold_clusters/small_clusters/gold_clusters.pkl #the pickles that are read which contains the reference energies and the validations energies (pickles are format from python environment)

  batch_size: 128 #The number of geometries that are tested

  max_epochs: 1000 #The total number of steps that the learning is taking

  nbatch_per_epoch: 200 #Number of steps that the NN is seeing the batch so for 128 geometries the NN is seeing them 128x200

  nbatch_per_validation: 50 #The number of steps that the NN is testing the learning with the validation

  ema_decay: 0.99 #The decay of the loss function in order to average along the loss and not step onto minimum local

  default_loss_type: mse #mean square error can be change to log for example

  # default_status: frozen


  weight_decay: 1.e-4 #The decay that is used for the different specific targets

  decay_targets: [energy,embedding/species_linear_0,embedding/Wa_0,embedding/dxi_0,embedding/species_linear_1,embedding/Wa_1,embedding/dxi_1] #The targets

  # decay_targets: [refine_embedding_1,energy]

  energy_terms: [energy,repulsion] #The terms that are used for the energy learning
```

```
  print_stages_params: True
  # print_model: True
  frozen: [repulsion] #The repulsion is frozen which means that the NN is not changing it during the process. The repulsion comes from reference values


  # coordinates_ref_key: coordinates_ref


  stages:
    train_spice1:
      init_lr: 1.e-4 #The convergence criterion ?
      lr: 1.e-4
      final_lr: 1.e-5
      # end_event: [rmse_f, 5.]
      loss:
        e:
          key: total_energy
          # ref: e_formation_dft
          ref: formation_energy
          # remove_ref_sys: True
          weight: 1.e-1 #Weigh of the learning
          unit: kcalpermol
        f:
          key: forces
          ref: forces
          weight: 1. #Weigh of the learning in this situation the NN is learning more from the forces
          unit: kcalpermol
          threshold: 3.


model:
```

```
cutoff: 5.3 #cutoff for the graph it won't see things far from 5.3 A
energy_terms: [energy,repulsion]
preprocessing:
  graph:
    switch_params:
      switch_type: polynomial
      p: 3.
      trainable: True
  graph_filter_angle:
    module_name: GRAPH_FILTER
    cutoff: 3.5
    graph_key: graph
    graph_out: graph_angle
    remove_hydrogens: True
    switch_params:
      switch_type: polynomial
      p: 20.
      trainable: True
  angle_extension:
    module_name: GRAPH_ANGULAR_EXTENSION
    graph_key: graph_angle

modules:
  embedding:
    module_name: MACARON
    dim: 512
    dim_src: 64
    dim_dst: 32
    dim_angle: [8,32]
```

```
    nmax_angle: 4
    nlayers: 2
    radial_basis:
      dim: 10
      basis: spooky
      trainable: True
    radial_basis_angle:
      dim: 8
      basis: bessel
      trainable: True
    species_encoding:
      # dim: 64
      encoding: electronic_structure
      trainable: False
      zmax: 86
    embedding_key: embedding
    graph_angle_key: graph_angle
    message_passing: True
    mixing_hidden: [512,512]
    activation_mixing: tssr2
    angle_combine_pairs: True

  energy:
    module_name: NEURAL_NET
    neurons: [256,128,64,32,1]
    activation: swish
    input_key: embedding
    squeeze: True
```

```
  repulsion:
    module_name: repulsion_zbl
    trainable: True
```

To launch > fennol_train nameofthefile.yaml

To launch a dynamic:

Need to have the file “input.fnl” constructs as follow:

```
device cpu (or put cuda:0 to use gpu)
model_file latest_model.fnx #model that will be use for the simulation from the learning (author’s note: use the best_model.fnx instead of the final one)
print_variance yes

xyz_input{
  file Au6.xyz    #Geo that will be use
  indexed no      # Index for each atom
  box_info no     # second line that is read to get the info of the box simu
}
estimate_pressure no
nblist_stride = 20    #The number of step where the program is writing the geometry
nblist_skin = 0.2     #the ??

nbeads 32

nsteps = 1000000 #number of steps that is process
dt[fs] = 0.5 #timestep
gamma[THz] = 20. #property for langevin thermostat
temperature = 300.
```

```
tdump[ps] = 1 #dumps for the temperature ?
nprint = 100 #step where the info is printed

thermostat LGV

qtb{
  tseg[ps]=0.25
  omegacut[cm1]=15000.
  skipseg = 5
  startsave = 50
  agamma  = 1.e-1
}
```

Then the command to launch the input file is

fennol_md input.fnl > log.out

The geometries at every nblist_stride are stored in the .arc file and all the informations about the energies, the temperature are stored in the log.out.

The .arc file can be directly read with VMD.