Accelerating Peptoid-Surface Simulations: Peptoid Binding on Au(111) and Au(100)

Tai-Yu Pan

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington
2018

Reading Committee:
Jim Pfaendtner
David Beck

Program Authorized to Offer Degree:
Chemical Engineering
In nature, proteins and peptides are found to have specific binding on inorganic surface, especially on gold surface. However, the main challenge for synthesized proteins and peptides is their poor stability. Thus, sequence defined peptoids were found to mimic the natural behavior. F. Yan et al. recently discovered a method for designing peptoid that was able to change the morphology of the gold nanoparticles to star-shaped. However, no matter proteins, peptides or peptoids, the binding mechanism remains unknown due to difficult observation in conformational change from laboratory experiment. In our study, we used molecular dynamics simulation to investigate in atomic level for systems of peptoid on Au(111) and Au(100) surfaces, which are the most common structure for gold in nature. Combined with enhanced sampling method, we were able to obtain the binding free energy. In addition, to sample the strong binding systems, we introduced the two-step metadynamics method to accelerate the simulations.
TABLE OF CONTENTS

List of Figures ...................................................................................................................................................... ii

List of Tables ........................................................................................................................................................ iii

Chapter 1. Introduction ........................................................................................................................................ 1

1.1 Molecular Dynamics ........................................................................................................................................ 1

1.2 Peptoid on Au(111) and Au(100) Surface .................................................................................................... 2

Chapter 2. Methods ............................................................................................................................................... 5

2.1 Classical MD Simulations ............................................................................................................................... 5

2.2 Two-Step Metadynamics ............................................................................................................................... 8

Chapter 3. Results and Discussion .................................................................................................................... 13

3.1 Convergence of Two-Step MetaD ................................................................................................................ 13

3.2 Comparison of Au(100) and Au(111) ......................................................................................................... 18

Chapter 4. Conclusion .......................................................................................................................................... 22

Bibliography .......................................................................................................................................................... 23

Appendix A ............................................................................................................................................................ 26
LIST OF FIGURES

Figure 1.1. Peptoid-induced formation of gold nanostructures ........................................ 4
Figure 2.1. Arrangement of virtual sites on Au(111) and Au(100) surface ....................... 6
Figure 2.2. CV vs simulation time on Au(100)/Au(111) for standard MetaD. ............... 11
Figure 3.1. CV vs simulation time on Au(100)/Au(111) for WTMetaD......................... 15
Figure 3.2. Gaussian height vs simulation time on Au(100)/Au(111) for WTMetaD...... 16
Figure 3.3. Reweighted free energy surface on Au(100)/Au(111) surface.................... 17
Figure 3.4. Conformation of the peptoid in local minimum on Au(100) and Au(111) ... 20
Figure 3.5. Normalized water density profile for blank tests and production runs ........ 21
LIST OF TABLES

Table 2.1. GoIP-Charmm LJ parameters for Au(111) surface .............................................. 6
Table 2.2. GoIP-Charmm LJ parameters for Au(100) surface .............................................. 7
Table 2.3. Summary of system setup ......................................................................................... 8
Table 2.4. Metadynamics parameters ....................................................................................... 12
ACKNOWLEDGEMENTS

I am proud of being one of the members in PRG group. First, I would like to thank my thesis advisor Professor Jim Pfaendtner. Though Jim was on sabbatical in the second year, he still answered to my questions and problems in research anytime anywhere. He was so energetic and always encouraged me whenever I felt frustrated. He was supportive to my decision after graduation and gave me useful suggestion. I have learned a lot under his instruction. I would also like to thank all the PRG group members: Kejia, Khushmeen, Luke, Sarah, Luiz, Chris, Wesley, Kayla, Josh, Coco, and Arushi. They were awesome lab mates and always provided solutions to my problems. Especially Arushi, she was the best mentor during the period and there wasn’t any error message that she couldn’t solve. I would also like to thank Professor David Beck as second reader of this thesis. He brought me into the world of data science and aroused my interest in computer science. Lastly, simulations were able to complete by using the HYAK super computer at University of Washington.
DEDICATION

To all my family, who is the most important to me and supported me for higher education opportunity
Chapter 1. INTRODUCTION

1.1 MOLECULAR DYNAMICS

Molecular dynamics (MD) simulation has become an important tool for the study of protein and peptoid systems. It can provide an accurate, atomic-level description which may not be found in conventional experiments. MD simulation includes the numerical, step by step, solution of the Newton’s equations of motion, which can be written as:

\[ m_i \frac{\partial^2 r_i}{\partial t^2} = F_i \quad \text{and} \quad F_i = -\frac{\partial V}{\partial r_i} \]  

(1.1)

By solving the classical equations of motion, we can predict the position of atoms for each time step, and then form a molecular trajectory. The key to the solution is the force, which is derived from the potential energy \( V \). Potential Energy can be subdivided into two parts: non-bonded and bonded interactions. Non-bonded interactions include the Lennard-Jones potential and Coulomb interaction, while bonded interactions consist of bond, angle, and torsion potential. The Lennard-Jones (LJ) potential between two atoms equals:

\[ V_{LJ}(r_{ij}) = 4\epsilon_{ij} \left( \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \]  

(1.2)

The calculation process can take huge computational resources. Thanks to the use of high performance computing (HPC), MD simulation nowadays can be applied to more complex systems and larger time scales.

In our study, we use MD simulations to discover the mechanism of peptoid binding on Au(111) and Au(100) surface. To calculate the binding energy, the peptoid has to be given a bias potential to escape from the local minimum energy state. Metadynamics (MetaD) is a commonly used enhanced sampling method.[1]–[3] An external history-dependent bias potential is
constructed in the space of selected collective variables (CVs). However, in the strong binding case, the simulation can take over several microseconds, which exceeds a month with available computational resources. Excitingly, we developed a two-step MetaD method to accelerate the process. This approach is more efficient and can be applied to a larger free energy system.

1.2 Peptoid on Au(111) and Au(100) Surface

In nature, proteins and peptides are found to have specific binding on inorganic surface, especially on gold surface. Many research has been done to use proteins and peptides to assemble functional materials for wide range of applications. Though some researchers tried to use synthetic proteins and peptides, they still faced obstacles due to the difficulty in predicting structures of proteins and peptides and their poor stabilities against thermal and chemical degradation. Some researchers investigated the sequence-defined polymers to mimic the behavior of nature proteins and peptides for higher stability. C. Chen et al. demonstrated that peptoids, namely poly-N-substituted glycines, were able to self-assemble into 2D networks comprising hexagonally patterned nanoribbons on mica surfaces, and—in the case of lipid-like peptoids—into 2D membrane-mimetic materials.

Since the morphology-dependent physical and chemical properties of plasmonic nanomaterials are significant for applications in sensing, photonics and catalysis, some researchers used proteins and peptides to control mineral formation. Recently, F. Yan et al. discovered a method for designing peptoid that was able to control the morphology of the gold nanoparticles. They successfully changed spherical nanoparticles to hedgehog-shaped. As shown in Figure 1.1, the peptoid, containing three Nce and six N4-Chpe groups (Figure 1.1A), induced formation of five-star shaped gold nanostructures (Figure 1.1B and C) when incubated with HAuCl₄ and HEPES, while changing from HAuCl₄ to AgNO₃ results in forming nanoribbons
comprising uniform, discrete, spherical silver nanoparticles (Figure 1D and E).[16] In this case, the peptoid controlled the nucleation of plasmonic nanoparticles and their assembly into superstructures. However, the underlying mechanism remains unknown.

The vision of our study is to achieve a predictive understanding of principles underlying peptoid-controlled plasmonic nanocrystal formation and develop rules for designing peptoids that lead to the formation of plasmonic nanomaterials for which the morphology is programmable and the resulting function is predictable. An atomic level of observation is necessary to discover the rules governing bio-controlled formation of plasmonic nanomaterials. We used molecular dynamics simulation to investigate the underlying interactions on the interface. We reproduced the system from experiments from F. Yan et al.[16] Furthermore, the gold structure of Au(100) and Au(111) are mostly found in nature. We took both in consideration and made a comparison. Combined with our two-step metadynamics enhanced sampling method, we were able to obtain the free energy surface and to illustrate the interaction mechanism.
Figure 1.1. Peptoid-induced formation of gold nanostructures. (A) Structures of peptoid Nce$_3$(N$_4$-Clpe)$_6$. (B) and (C) TEM images of gold nanostars. (D) and (E) TEM images of nanoribbons of silver nanoparticles. This figure was taken in whole from reference [16].
Chapter 2. METHODS

2.1 CLASSICAL MD SIMULATIONS

The Classical Simulations were performed on the GROMACS 5.1.2[17], [18] with the force filed GoIP-CHARMM[19], [20]. This force field contains terms to describe the dynamic polarization of gold atoms, chemisorbing species, and the interaction between sp² hybridized carbon atoms and gold. Both Au(111)/(100) slabs were constructed as a size around 6.1 nm × 6.1nm with atomic 6 layers by using a lattice parameter of 2.93 Å. The virtual sites were placed on the surface of gold slabs, as shown in Figure 2.1. This geometry arrangement and the L-J parameters was optimized by a combination of experimental and first principles data. LJ parameters for Au(111) and Au(100) was given in Table 2.1. The peptoid (Figure 1.1A) was put above gold surface in a box of height 6.1 nm filled with TIP3P water. An energy minimization was conducted using the steepest descent algorithm with van der Waals, neighbor list and coulomb cutoffs at 1nm. The minimized configuration was equilibrated at 300 K in the NVT ensemble, using Velocity-rescaling temperature coupling (\(\tau = 0.1\) ps) for 5ns. Covalent bonds with hydrogen atoms were constrained using LINCS[21] to simulate with a time step of 2 fs. Blank tests, containing only water molecules and gold surface, were done by same NVT configuration with a time step of 1 fs. All gold atoms were frozen during all simulations to maintain the surface structure. The specifics of the simulations are tabulated in Table 2.2.
Figure 2.1. Arrangement of virtual sites on (a) the Au(111) surface and (b) the Au(100) surface. Real gold atoms are shown in gold, virtual sites in blue. This figure was taken in whole from reference [20].

Table 2.1. GoIP-Charmm LJ parameters for AUI and AUB on Au(100) and Au(111) surface$^a$

<table>
<thead>
<tr>
<th></th>
<th>Au(100)</th>
<th>Au(111)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\varepsilon$ (kJ/mol) $\sigma$ (Å)</td>
<td>$\varepsilon$ (kJ/mol) $\sigma$ (Å)</td>
</tr>
<tr>
<td>Au–Au</td>
<td>2.10</td>
<td>4.00</td>
</tr>
<tr>
<td>Au–N</td>
<td>5.40</td>
<td>3.00</td>
</tr>
<tr>
<td>Au–O(carbonyl)</td>
<td>0.60</td>
<td>3.35</td>
</tr>
<tr>
<td>Au–O(water)</td>
<td>1.50</td>
<td>3.275</td>
</tr>
<tr>
<td>Au–H</td>
<td>0.60</td>
<td>2.725</td>
</tr>
</tbody>
</table>

$^a$Parameters for surface virtual sites (AUI) and bulk gold atoms (AUB) are given, while no LJ terms are assigned to real gold surface atoms (AUS).
Table 2.2. Summary of system setup

<table>
<thead>
<tr>
<th></th>
<th>Au(100)</th>
<th>Au(111)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total atoms</td>
<td>21253</td>
<td>21796</td>
</tr>
<tr>
<td>Simulation box size</td>
<td>$6.15 \times 6.15 \times 6.15$ nm$^3$</td>
<td>$6.15 \times 6.09 \times 6.1$ nm$^3$</td>
</tr>
<tr>
<td>Number of gold layers</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
2.2 TWO-STEP Metadynamics

Metadynamics was implemented using the PLUMED[22] library in GROMACS. As introduced in Chapter 1, it is used to encourage the system to explore the high free energy area by adding Gaussian hills. The total bias potential is sum of these Gaussians which are deposited along the trajectory, as shown below:

\[ V_{bias}(s, t) = \sum_{k\tau < t} W(k\tau) \exp\left(-\sum_{i=1}^d \frac{(s_i - s_i(q(k\tau)))^2}{2\sigma_i^2}\right) \]  

(2.1)

where \( s \) is the CVs, \( \tau \) the Gaussian deposition stride, \( \sigma_i \) the width of the Gaussian for the \( i \)-th CV, and \( W(k\tau) \) the height of the Gaussian.

In standard MetaD, Gaussians with constant height are added along the entire simulation. However, it is difficult to find a perfect value of Gaussian height. The simulation will take longer by using smaller value, while the system will be hard to converge by using larger one. This problem can be solved by using well-tempered metadynamics (WTMetaD)[4], [23], [24]. In WTMetaD, the Gaussian height is decreased as:

\[ W(k\tau) = W_0 \exp\left(\frac{-V(s(q(k\tau)), k\tau)}{k_B \Delta T}\right) \]  

(2.2)

where \( W_0 \) is an initial Gaussian height, \( \Delta T \) an input parameter with the dimension of a temperature, and \( k_B \) the Boltzmann constant. With this approach, the bias potential can smoothly converge in the limit time scale. Usually, the system is expected to converge in hundreds of nanoseconds by WTMetaD. However, the Gaussian height is decreased by an exponential term, which means after a period of time, the height will be very small. In the case of strong binding system, it will take over several microseconds and use a lot of computational resources to escape from local minimum.

In our study, we introduced a two-step MetaD, which combined the both benefits from standard MetaD and WTMetaD. The CV selected for biasing with metadynamics were the radius
of gyration of the peptoid (CV1) and the z-distance between the gold surface and the center of mass of the peptoid (CV2). Production simulations were performed with multiple walkers[25] metadynamics, with initial structures generated by using steered MD to position four starting structures in disparate regions of the CV space, and it was carried out in the NVT ensemble using the cut-off mentioned in section 2.1. First, we ran a standard MetaD. All CVs had an initial Gaussian height of 0.05 kJ/mol and hills were deposited every 4 ps. The COLVAR file generated by PLUMED was monitored during the simulation. As shown in Figure 2.2, we terminated the simulations when at least one walker came off the gold surface (local minima) and fluctuated for about 50 ns. And then, the utility sum_hills in PLUMED was used to sum the Gaussians deposited during the simulation and stored in the HILLS file. The fes.dat file was generated, which contains the free energy estimate and the derivatives of all CVs along the grid of CV space. From standard MetaD, we were able to obtain an estimate of the underlying free energy surface quickly, and used it as a static bias potential for additional MetaD simulations. We alternated the signs for columns of free energy and derivatives in order to use it as external bias for step two. This file should be included in argument EXTERNAL for following WTMetaD. WTMetaD started with an initial Gaussian height of 1.0 kJ/mol for all CVs and bias factor of 50. Hills were deposited every 1 ps. The simulations were considered converged if the free energy profile of the CV of interest did not change significantly during the last 25% of the production run. The MetaD parameters are tabulated in Table 2.3.

Reweighting method was used to reconstruct the free energy surface (FES) as a function of CVs by following equation[26]

$$ F(s) = -\frac{1}{\beta} \log P(s) \quad (2.3) $$
where $\beta = 1/k_B T$, $T$ is the temperature, $P(s)$ is the probability distribution of the collective variables. Reweighting is typically used to rebuild the full Boltzmann distribution due to the distortion of unbiased variables. In our case, we used bias potential from step 1 as an external bias for step 2. Thus, reweighting was needed to include all bias potential and reconstruct the FES. D. Branduardi et al. introduced an alternative and simpler algorithm which can be applied to WTMetaD with multiple walkers.[27], [28] In the approximate method, we took the final bias potential from all simulations and re-processed the trajectory from the start. This gave the relative weights of each frames so that we could reconstruct the $P(s)$ by histogram and substitute into Eq. 2.3.
Figure 2.2. Collective variable along the simulation time on Au(100)/Au(111) respectively for standard MetaD.
<table>
<thead>
<tr>
<th></th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>standard MetaD</td>
<td>WTMetaD</td>
</tr>
<tr>
<td>biased variables</td>
<td>radius of gyration, distance from surface</td>
<td></td>
</tr>
<tr>
<td>sigma</td>
<td>0.02 nm, 0.025 nm</td>
<td>0.02 nm, 0.025 nm</td>
</tr>
<tr>
<td>bias factor</td>
<td>---</td>
<td>50</td>
</tr>
<tr>
<td>hill height</td>
<td>0.05 kJ/mol</td>
<td>1.00 kJ/mol</td>
</tr>
<tr>
<td>hill deposition rate</td>
<td>2000 steps</td>
<td>500 steps</td>
</tr>
<tr>
<td>number of walkers</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
Chapter 3. RESULTS AND DISCUSSION

3.1 CONVERGENCE OF TWO-STEP META\!D

To determine the convergence of the system, we first looked at the CV versus simulation time. In step 1, as shown in Figure 2.2, the peptoid was stuck on both Au(100) and Au(111) surface for most of simulation time. In this period, Gaussian hills were continually added to fill up the energy basin. At 200 ns, walker 4 was able to escape from the local minimum and explore a new region of the CV space. This was a good sign to terminate the simulation and turn into the next step because there was enough bias potential to push the peptoid away from the surface. Step 2, we used WTMeta\!D to smoothly converge the bias potential. Figure 3.1 shows the CV plot for WTMeta\!D simulation. All four walkers with significant fluctuation were able to come on and off the gold surface repeatedly, which means the system diffused in the entire CVs space.

Second, we checked the Gaussian height along the simulation time. As mentioned in Chapter 2.2, the Gaussians with constant height were added in standard Meta\!D simulation. We obtained an estimate of underlying free energy surface. WTMeta\!D was needed to again sample the system with static bias from step 1 and add new bias with decreased Gaussian height as Equation 2.2. As shown in Figure 3.2, Gaussian height was reduced as an exponential profile along the entire simulation time in step 2.

Third, the most important step, we reconstructed the free energy surface by reweighting. The simulations were considered converged if the free energy profile of the CV of interest did not change significantly during the last 25% of the production run. As shown in Figure 3.3, we took 50%, 75%, 90%, 100% of simulation time frames to rebuild the free energy profile. On Au(100) surface, the difference between 75% of frames (orange) and 100% of frames (red) was smaller
than 1%. On Au(111) surface, four lines were overlapped by each other. Both systems were considered to converge to the underlying free energy landscape.

In the case of strong binding system, it is difficult to converge in limit time scale. In general, it takes over several microseconds with significant computational resources in our experiences. In our approach, we successfully converged the system in 550 ns by two-step metadynamics with four multiple walkers. The system was able to escape from the local minimum and explore the entire CVs space. This method at least reduced 50% of computational cost in our case and can be applied to even some weaker binding systems.
Figure 3.1. Collective variable along the simulation time on Au(100)/Au(111) respectively for WTMetaD.
Figure 3.2. Gaussian height along the simulation time on Au(100)/Au(111) respectively for WTMetaD.
Figure 3.3. Reweighted free energy surface on Au(100)/Au(111) surface. 50%, 75%, 90%, 100% of frames were taken respectively.
3.2 COMPARISON OF AU(100) AND AU(111)

Many research has been done in gold binding peptides with simulations. L. B. Wright et al. found the adsorption of peptides to Au(111) surface is stronger than to Au(100).[29] Braun et al. studied the interaction of a gold binding peptide on Au(111) and Au(211) surface and made comparison on interaction energies and water population.[30]–[33] Heinz et al. investigated binding peptides on different surfaces including Au(100) and Au(111). [34] M. Hoefling et al. calculated the potential of mean force between 20 amino acids and Au(111) surface.[35]

In our study, as shown in Figure 3.3, the estimated peptoid binding energy on Au(100) and Au(111) surfaces were 96.84 kJ/mol and 110.14 kJ/mol respectively. In general, the free energy of adsorption approaches or exceeds 40 kJ/mol in protein-inorganic pairs is considered to be strong.[36] In our system, the binding energies are 2 to 3 times to this value. It also indicates that the system was difficult to converge in limit simulation time scale without two-step metadynamics method. The reason leading the significant binding energy is the side chains on the peptoid. M. Hoefling et al. showed that the amino acids (Phe, Trp, Tyr) with aromatic groups have large adsorption energy in range of 40 to 45 kJ/mol.[35] The peptoid in our system contains 6 phenyl chloride groups, which have very high aromaticity. It may cause \( \pi \)-electron mediated effects with the gold surface. In addition, aromatic rings enable the maximum contact to the plain surface. As shown in Figure 3.4, six rings lied flatly on both Au(100)/Au(111) surface. By looking into the trajectories, we discovered that this structure was stable and appeared repeatedly whenever the peptoid touched the surface.

There is a 15 kJ/mol difference in favor of binding on Au(111) surface. Many researchers also found this preference over Au(100) surface.[29], [37] with peptides. To illustrate the underlying mechanism, we first looked at the slightly difference in conformation of the peptoid
between Au(100) and Au(111) surface. Though the six rings were able to convert to 2D structure on both surfaces, the rest of the residues, 3 Nce groups, acted differently during the simulation. As shown in Figure 3.4(a) and (c), Nce groups on Au(100) tend to be away from the surface while groups on Au(111) have stronger affinity to the surface. This can also be proved by reweighted free energy profile shown in Figure 3.3. The FES on Au(111) surface has a clear point at lowest free energy, while it shows a range on Au(100), and it is due to the calculation of center of mass of the peptoid. Nce group, which contains carboxylic group, is considered to be hydrophilic due to hydrogen bond. Thus, we further looked into the water profiles of both systems. Figure 3.3 shows the distance near 0.5 nm from surface has the local minimum, so we focus on the water profiles within this range. The blank tests contained only water molecules and gold surface. There is no surprise to see both systems have a slight drop in water density when including the peptoid due to hydrophobicity of 6 aromatic rings, as shown in Figure 3.5. Furthermore, Au(111) has greater water affinity than Au(100). Combined the information above, we suggested that the higher water adsorption on Au(111) enables the Nce groups to come closer to the surface, and further enhance the binding energy of the peptoid.

The other important factor to affect the binding energy is the geometry of the surface and the peptoid. As shown in Figure 2.1, the virtual sites in purple have actual LJ potential parameters for surface interaction. This arrangement is calculated by a combination of experimental and first-principles data.[20] The geometry of the virtual sites in Au(111) is a normal hexagon with a side length of 1.69 Å, while it’s a square on Au(100) with a side length of 2.93 Å. The aromatic ring on the peptoid is also a hexagon with an average bond length of 1.40 Å. The similar geometry of the Au(111) and the aromatic rings is able to maximize the contact area, which leads to the greater binding affinity over Au(100).
Figure 3.4. Conformation of the peptoid in local minimum on (a)(b) Au(100) and (c)(d) Au(111) surface.
Figure 3.5. Normalized water density profile for (a)(b) blank tests and (c)(d) production runs respectively.
Chapter 4. CONCLUSION

In our study, we successfully introduced two-step metadynamics, combined with standard and well-tempered metadynamics, to converge the systems with high binding energy in 550 ns simulation time. We reconstructed the free energy surface by reweighting with affordable computational resources in limit simulation time scale. Our approach is more efficient and can be applied to wide range of systems.

We further demonstrated that the physical and chemical properties of the side chain is the main effect on binding affinity. The aromatic ring enables the full contact with the plain surface. The conformational difference of the peptoid will be a minor factor, which affected by intramolecular and intermolecular interactions, and water profiles on the surface. We also reported a clear preference of adsorption on Au(111) over Au(100) due to the geometry of the gold surface.
BIBLIOGRAPHY


