JYU DISSERTATIONS 91

Emmi Pohjolainen

Atomistic Simulation View on Gold Nanocluster Functionalities via Ligand Shell Dynamics

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF MATHEMATICS
AND SCIENCE

JYU DISSERTATIONS 91

Emmi Pohjolainen

Atomistic Simulation View on Gold Nanocluster Functionalities via Ligand Shell Dynamics

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Ylistönrinteen salissa FYS1
kesäkuun 14. päivänä 2019 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in Ylistönrinne, auditorium FYS1, on June 14, 2019 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2019

Editors
Timo Sajavaara
Department of Physics, University of Jyväskylä
Ville Korkiakangas
Open Science Centre, University of Jyväskylä

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7781-8

ISBN 978-951-39-7781-8 (PDF)
URN:ISBN:978-951-39-7781-8
ISSN 2489-9003

ABSTRACT

Pohjolainen, Emmi
Atomistic Simulation View on Gold Nanocluster Functionalities via Ligand Shell Dynamics
Jyväskylä: University of Jyväskylä, 2019, 86 p.(+included articles)
(JYU Dissertations
ISSN 2489-9003; 91)
ISBN 978-951-39-7781-8 (PDF)

Gold nanoclusters have been traditionally extensively studied by experimental and DFT methods. The past results have revealed their character as an interesting species at the limit of bulk and atomic, with properties tunable for numerous potential applications. Especially, biological and medical applications have raised interest with gold nanoclusters acting as potential labels for imaging, and as drug carriers. To understand in detail their behavior in solvent and protein environments at atomic scale, simulations complementing the experimental data are needed. While such size and time scales are for the most parts beyond reach for DFT methods, such systems are well within capabilities of all-atom MD simulations. In this thesis, the force field parameters for performing such simulations are developed, and applied for different types of gold nanocluster systems to study their functionalities. MD simulations were successfully applied in revealing ligand shell conformations to explain structural functionalities of gold nanoclusters with unknown ligand shell structure. Considering larger systems, MD simulations were applied in investigating solvent and protonation conditions for gold nanocluster self-assembly and superstructure formation. Furthering complexity, simulations with gold nanoclusters and a full enterovirus were performed to study atomic scale interactions between the two. Also free energy calculations were performed to shed light on binding affinities of various drug-based molecules to the virus, and implications of adding the gold nanoclusters to the picture were concluded. Altogether, the results show that MD simulations fill a niche in investigating such systems alongside experimental and DFT methods.

Keywords: Gold nanocluster, Molecular dynamics, Simulations, Force field

Author

Emmi Pohjolainen
Department of Physics
Nanoscience Center
University of Jyväskylä
Jyväskylä, Finland

Supervisors

Academy Professor Hannu Häkkinen
Departments of Physics and Chemistry
Nanoscience Center
University of Jyväskylä
Jyväskylä, Finland

Professor Gerrit Groenhof
Department of Chemistry
Nanoscience Center
University of Jyväskylä
Jyväskylä, Finland

Reviewers

Assistant Professor Zoe Cournia
Biomedical Research Foundation
Academy of Athens
Athens, Greece

Researcher Giulia Rossi
Physics Department
University of Genoa
Genoa, Italy

Opponent

Professor Jaakko Akola
Department of Physics
Norwegian University of Science and Technology
Trondheim, Norway
and
Laboratory of Physics
Faculty of Natural Sciences
Tampere University
Tampere, Finland

PREFACE

The work presented in this thesis was carried out at the Nanoscience Center, Department of Physics at the University of Jyväskylä during 2014-2019. I would like to acknowledge NGS-NANO graduate school for the initial and Vilho, Yrjö and Kalle Väisälä Foundation for two years of financial support. CSC Finnish IT center for Science, Finnish Grid Infrastructure and PRACE are acknowledged for computational resources.

First, I would like to express my gratitude to my supervisors Academy Professor Hannu Häkkinen and Professor Gerrit Groenhof. Little did I know what I poured myself into, when Hannu first introduced me to this project in starting my master's thesis work. Thank you for believing in me letting me continue this far, and thank you both for all the invaluable guidance and opportunities during these years.

With all the simulation work first and foremost being a collaboration with the experimentalists, I would like to thank Professor Mika Pettersson, senior Lecturer Tanja Lahtinen, Lecturer Varpu Marjomäki, Dr. Satu Mustalahti, Dr. Nonappa and Dr. Tiia-Riikka Tero and everyone part of the gold-virus consortium meetings for extremely useful discussions.

I would also like to give warm thanks to the reviewers of this thesis, Assistant Professor Zoe Cournia and Researcher Giulia Rossi for taking the time to offer their comments.

Being part of two groups with different expertise has been a massive privilege. I would like to thank all the past and present members of both groups for great discussions, guidance, company and friendship; Elli and Noora, you have been such great company in conference trips whether it was hustle and bustle in China, gazillion stairs in Switzerland or tracking down Max Planck's grave in Germany - after dark, naturally.

Especially I'd like to thank Dr. Sami Malola; I doubt there was ever a challenge with gold nanocluster computations he was not able to solve and present in a beautiful picture. In lack of ideas on how to continue and what to improve the advice of Dr. Dmitry Morozov were always plenty and welcome. Also special thanks to Dr. Andre Clayborne for encouragement, Dr. Xi Chen and Ms. Saara Lautala, I really enjoyed working with all of you. Mutual suffering is the best of its kind! In general, thank you everyone at the Nanoscience Center for contributing to a great working environment and community.

Finally, I would like to thank my family and friends, life outside science. Especially my parents Sonja and Kalle, and my sister Tiina for their (sometimes tough) love and support, it was all beyond essential for me to complete this work. I always had a lively home to come to and reload at. Bed was always made, and good company guaranteed for a tired Helsinki-conference-goer at Tiina's. Lastly, thank you Marko for your love, endless (IT and other) support, (somewhat forced) interest in my work, patience and humor, every day.

Jyväskylä, April 2019

Emmi Pohjolainen

LIST OF INCLUDED ARTICLES

I Emmi Pohjolainen, Xi Chen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *Journal of Chemical Theory and Computation*, 12 (3), pp. 1342–1350, **2016**.

II Tiia-Riikka Tero, Sami Malola, Benedek Koncz, Emmi Pohjolainen, Saara Lautala, Satu Mustalahti, Perttu Permi, Gerrit Groenhof, Mika Pettersson, Hannu Häkkinen. Dynamic Stabilization of the Ligand-Metal Interface in Atomically Precise Gold Nanoclusters Au_{68} and Au_{144} Protected by *m*eta-Mercaptobenzoic Acid. *ACS Nano*, 11 (12), pp. 11872–11879, **2017**.

III Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**.

The author performed most of the molecular dynamics simulations, performed the analyses and wrote the first drafts for the publications **I** and **III**.

For the paper **II** the author performed or devised most of the molecular dynamics simulations and result analyses, and participated in writing the manuscript.

An unpublished work on simulating gold nanocluster self-assemblies is presented as the part of the results of this thesis.

The author has contributed to the following publications not included in this thesis:

A.I Emmi Pohjolainen, Hannu Häkkinen, Andre Clayborne. The Role of the Anchor Atom in the Ligand of the Monolayer-Protected $Au_{25}(XR)_{18}^{-}$ Nanocluster. *The Journal of Physical Chemistry C*, 119 (17), pp. 9587–9594, **2015**.

ABBREVIATIONS

MD	molecular dynamics
DFT	density functional theory
*p*MBA	*para*-mercaptobenzoic acid
*m*MBA	*meta*-mercaptobenzoic acid
PET	phenylethanethiol
NMR	nuclear magnetic resonance
IR	infrared
EM	electron microscopy
TEM	transmission electron microscopy
EV1	echovirus 1
PBC	periodic boundary conditions
LJ	Lennard-Jones
VDW	van der Waals
RMSD	root mean square deviation
MSD	mean square displacement
RESP	restrained electrostatic potential
TI	thermodynamic integration
FEP	free energy perturbation
R	generic organic group
UV-VIS	ultraviolet - visible spectroscopy
AI	artificial intelligence

CONTENTS

1 INTRODUCTION

At the heart of any computational method is, on one hand being able to reproduce the known, experimental results. That is, confirmation that the method used to describe the system is sufficiently accurate to reproduce what can be considered as known. Naturally, since such a method aims to describe how nature works, basis for such method should come from what is known about laws of physics.

On the other hand, computations aim to provide an explanation to the experimental results, at an accuracy or scale unreachable to the experiments themselves. That is, advance our understanding of nature. At best, the computational results may be utilized as predictive means to, e.g., guide materials design and, more generally, interpret various types of experimental data.

Inherently, with any computational method come the limitations, most often sacrificing accuracy over the time and size scales. Thus the system and its properties to be studied determine the computational method chosen for the task at hand. While the most cut-throat division could be seen to be quantum vs. classical mechanics, both sides contain numerous methods designed to describe different systems and their properties at different levels of theory. Currently, the most common choice of method for describing dynamics of large (bio)molecules at relevant timescales, is classical molecular dynamics.

Gold nanoclusters, atomically precise nanoscale objects, have raised interest as nanomaterials due to their distinctive properties. From basic research point of view, they are interesting objects at the limits of what could be considered bulk and atomic [1]. Moreover, ligand-protected, atomically precise clusters provide tunability for different applications by control of the ligand shell chemistry. This property has motivated the usage of gold nanoclusters in life-science applications.

Numerous computational studies of gold nanoclusters have been performed using DFT or quantum chemistry methods, which have greatly advanced the understanding of these systems. However, in going towards understanding their role in biological and as larger scale systems in solvent environment, dynamics at time scales way beyond such methods are needed. On the other hand, such simulations should provide atomistic scale understanding on structures and dynamics of the systems, that is, reach beyond experiments. This is an ideal challenge for molecular dynamics simulations, method whose utilization in gold nanocluster systems has been scarce compared to DFT methods.

Gold may be considered as atypical in MD simulations, that traditionally have triumphed in protein simulations. The current parameterization procedures, while not straightforward, enable force field generation and simulations of such unaccustomed species. The testing and validation of such parameters for ones purpose, becomes then crucial. Essentially, this is enabled by numerous gold nanocluster crystal structures and vast complementing experimental data to compare and refer to.

In the work presented in this thesis, a generalized force field parameterization for thiolate-protected gold nanoclusters is presented and validated against both experimental and DFT data, work **I**. The force field applicability is then demonstrated in various types simulations to shed light on:
(1) ligand shell structure of gold nanoclusters with unknown experimental structure (**II**)
(2) stability of self-assembled superstructures of gold nanoclusters (unpublished work) and
(3) binding affinities of gold nanoclusters to virus (**III**).

2 GOLD NANOCLUSTERS

Gold nanoclusters are nanoscale objects composed from few to some hundreds of gold atoms at maximum. As such, they can be considered as objects between bulk (inert) and atomic (reactive) gold, a property contributed to gold's several oxidation states at nanoscale [2] and that has made them an interesting research topic in understanding size-specific and finite-size effects of materials. Especially, properties and stability of gold nanoclusters are extremely size-dependent, up to the point it is often referred that every atom counts [3,4]. Also an interesting way to tune the properties in atomic scale includes replacing some of the gold atoms with other metals, such as silver, i.e., producing heterometallic clusters [5].

While bare gold nanoclusters are prone to aggregation, it is conventional to passivate them by, e.g., ligand layer composed of organic molecules or polymers [3,4]. Essentially then, it is the ligand layer that determines the interactions of the gold nanoclusters with the environment. This contributes, for example, to solubility, stability, optical, fluorescent and sensing properties. For example water solubility, dictated by the ligand shell, is desired for biocompatibility. Thus functionalizing the gold nanoclusters for specific applications via ligand shell tuning is of interest [6]. Common functional groups for passivation include thiolates, phosphines, selenolates and tellurolates [7].

For the scope of this thesis, homometallic gold nanoclusters passivated or protected by thiolates are considered.

2.1 Thiolate protected gold nanoclusters

Synthesis and characterization of such atomically precise gold nanoclusters have been reported for several sized clusters [3,4]. Some of the most used thiolates to that end include phenylethane thiol (PET) and *para* and *meta* mercaptobenzoic acids [8], shown in figure 1.

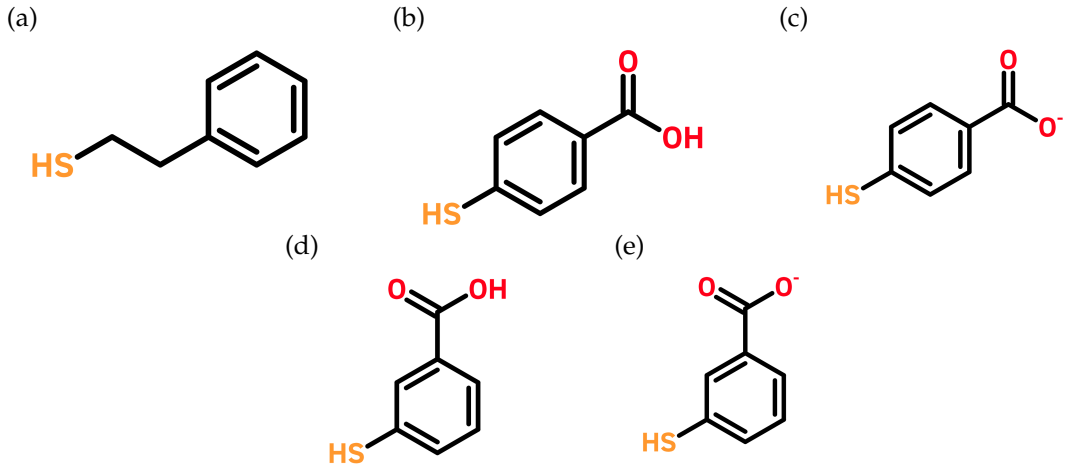


FIGURE 1 Common thiolate ligands of gold nanoclusters: (a) phenylethane thiol; (b) and (c) *para* mercaptobenzoic acid in protonated and deprotonated forms; (d) and (e) *meta* mercaptobenzoic acid in protonated and deprotonated forms.

Fundamentally, the gold-sulfur bond strength is similar to that of gold-gold bonds, indicating that bonding to thiolate enables modifying gold-gold bonding network [2] . Furthermore, the type of the ligand and its stericity along with inter-ligand interactions are likely to guide the formation of gold nanoclusters [9].

It has been observed that compositions of such thiolate protected gold nanoclusters follow certain relations in terms of the number of gold atoms in the cluster and the number of protecting thiolates. Moreover, certain numbers, referred to as "magic", of gold atoms in the cluster compositions appear independent of the ligand type, suggesting intrinsic stability of specific sized clusters [3,4].

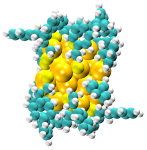
However, the ligand effects are manifested, for example, as observed unexpected bonding motifs and cluster compositions with bulky ligands compared to the smaller ones [10]. Also, more subtle changes in the ligand type have similar effects [11,12].

2.1.1 Structural and electronic stability

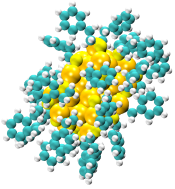
The first single-particle crystallization and structure determination up to molecular scale of $Au_{102}pMBA_{44}$ gold nanocluster [13], figure 2(c) provided the atomic resolution of the gold-sulfur interface as the "staple" motif, found to be generally conserved among different gold nanoclusters, as suggested by DFT [14].

The determination up to molecular scale of nine thiolate protected gold nanocluster structures, along with DFT calculations of several different thiolate protected gold nanocluster structures, have revealed the remarkable stability of these clusters due to both geometry and electronic structure [3,4]. Some example structures are presented in figure 2.

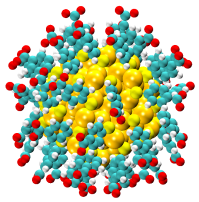
(a) $Au_{25}PET_{18}^{-}$, diameter ≈ 2.5 nm



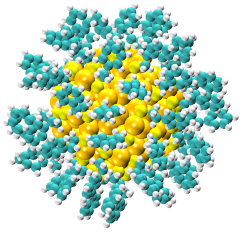
(b) $Au_{38}PET_{24}$, diameter ≈ 2.8 nm



(c) $Au_{102}pMBA_{44}$, diameter ≈ 2.9 nm



(d) $Au_{144}PET_{60}$, diameter ≈ 3.4 nm



(e) $Au_{144}pMBA_{60}$, diameter ≈ 3.3 nm

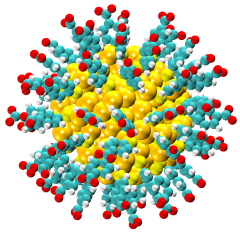


FIGURE 2 Examples of thiolate-protected gold nanocluster structures (used as reference in force field parameterization of work **I**). (a) $Au_{25}PET_{18}^{-}$ [15–18]; (b) $Au_{38}PET_{24}$ [19, 20]; (c) $Au_{102}pMBA_{44}$ [13]; (d) $Au_{144}PET_{60}$ [21]; (e) $Au_{144}pMBA_{60}$ (theoretical model based on $Au_{144}PET_{60}$ structure). Atom colors: C=cyan,O=red, H=white, S=yellow, Au=gold. Adapted with permission from Emmi Pohjolainen, Xi Chen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *Journal of Chemical Theory and Computation*, 12 (3), pp. 1342–1350, **2016**. Copyright 2016 American Chemical Society.

Closer look into such structures have revealed, in terms of geometry that the core of the gold nanocluster is often highly symmetric. However, not all the gold atoms of the cluster are considered part of this core, but so-called surface gold atoms participate in protecting the core by forming $RS(AuSR)_n$ units, often called staples. These surface gold atoms are bound to sulfurs in the thiolates, presented schematically in figure 3. This scheme is often referred to as "divide and protect", with part of the Au atoms in the core and part in the protective layer. This scheme also carries an electronic consideration [9] explained later. [2–4]

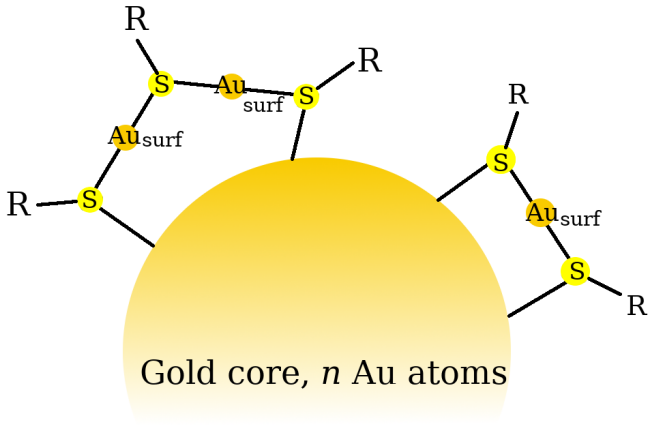


FIGURE 3 Divide and protect -scheme with separated n core Au atoms and surface Au_{surf} atoms (part of the staples) in two representative staple units; left: V-shaped staple unit with $RS{-}Au_{surf}{-}SR{-}Au_{surf}{-}SR$ and right: rectangular staple unit with $RS{-}Au_{surf}{-}SR$.

The staple units obtained from experimental atomistic structure determinations have revealed units of different lengths, i.e., containing one (rectangular staple) or two (V-shaped staple) gold atoms and two or three sulfur atoms [10]. In general, the V-shaped (two Au atoms) or even longer motifs are common with smaller clusters with higher surface curvature [4, 10]. On the other hand, also so-called simple bridging motif with direct bonding of thiolate to Au is possible [10].

To this end, structures presented in figure 2 may be considered in another way, assigning the core and the surface (staple) gold atoms more clearly: for example, $Au_{25}PET_{18}^-$ consists of an icosahedral 13-atom core, with the remaining 12 surface Au atoms as parts of the six V-shaped staple units forming the protective shell. Similarly, $Au_{144}PET_{60}$ consists of 114 atom core with 30 rectangular staples in the protective shell. In cases of $Au_{102}pMBA_{44}$ and $Au_{38}PET_{24}$ both V-shaped and rectangular staples are observed [2].

This division of gold atoms into core and surface atoms is both geometric and electronic, as revealed by the DFT calculations [9]: the Au atoms in the core are neutral (metallic) while the surface Au atoms are oxidized, due to electron attraction of the thiolates. This constitutes as the formal electron count of the protected gold nanocluster to correspond to a shell closing number of so-called superatomic orbitals of a spherically symmetric potential (given the core of the cluster is symmetric). This provides and electronic explanation to the stability of "magic" sized clusters [2, 22, 23].

2.1.2 Protonation and solubility

Both experiments and DFT computations have revealed numerous interesting properties of gold nanoclusters including structural, optical and electronic, and thus numerous potential applications have been suggested [1]. In general, the metal core is responsible for the properties including electronic and optical properties, while the ligand shell governs the chemical nature of the gold nanocluster, in terms of solubility and reactivity [24].

For the purposes of this thesis the focus is on structures, interactions and dynamics of the ligand shell with respect to its environment, a factor often missing in the DFT calculations, and the gold nanocluster functionalities thereof.

Considering the common ligands presented in figure 1, while the PET-protected gold nanoclusters are organosoluble, the potential water solubility of the MBA-protected clusters are of special interest in biological applications [8, 24].

The MBA ligands exist in different protonation states, protonated (neutral, figure 1 (b),(d)) and deprotonated (negatively charged figure 1 (c),(e)). The $Au_{102}p\text{MBA}_{44}$ nanocluster, has been shown to be water soluble down until pH 5, with minimum 7/44 ligands deprotonated. In the fully protonated (neutral) form the cluster is soluble in short-chain alcohols, such as methanol. Moreover, at pH 6.2, half of the ligands are expected to deprotonated [24]. In addition to the titration experiment, constant pH molecular dynamics simulations have been used to predict the charge distribution, i.e., positions of the protonated and deprotonated ligands on the cluster surface [24].

In addition, the protonation and the charge state of the gold nanoclusters control their affinity to charged or hydrophobic species, thus the charge state is of paramount importance considering, e.g., biological applications [24]. Especially, to this end the effects of nanoparticle surface chemistry on interactions with membranes have been extensively studied [25–33].

In addition to titration methods, solution NMR has also been used in characterizing gold nanoclusters behavior in solvents, namely in terms of hydrodynamic size and diffusion coefficients [34–36].

2.2 Selected Applications

In accordance with the extraordinary properties of gold nanoclusters, several potential applications have been envisioned, ranging from sensing, bioimaging and cancer treatment [6] to catalysis [37] and electronic applications [6].

In the following, selected potential applications in keeping with properties discussed in the last section, and working as motivations for the simulations presented in this thesis, are summarized in more detail.

The effect of even a small change in the ligand on the structure of gold nanoclusters is evident from reports of the structures of *m*MBA protected ligands (figure 1) of Au_{68}*m*MBA_{32} [11, 38] and Au_{144}*m*$MBA_{\approx 40}$ [11, 12]. While the structure of Au_{102}*p*MBA_{44} exhibits symmetrical positions of gold atoms, figure 4 (a), the reported electron microscopy structures of the *m*MBA protected clusters with 68 [38] and 144 [12] gold atoms show clearly more random positions, figure 4 (b) and (c).

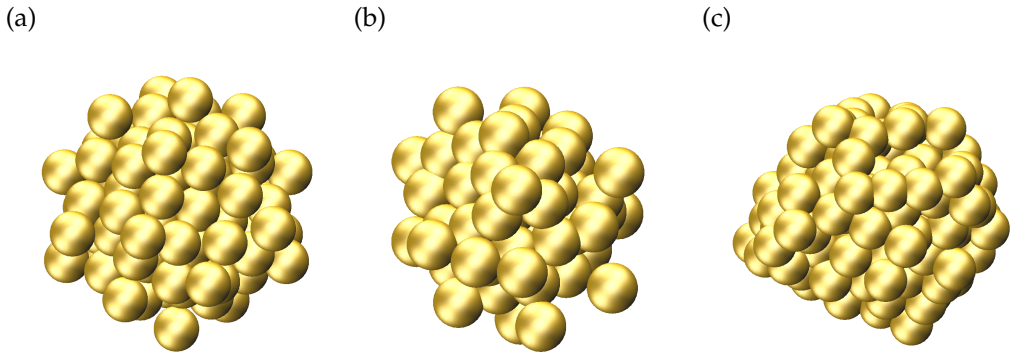


FIGURE 4 Positions of gold atoms of (a) Au_{102}*p*MBA_{44} [13], (b) Au_{68}*m*MBA_{32} [38] and (c) Au_{144}*m*$MBA_{\approx 40}$.

In addition, it was shown that the *m*MBA protected clusters showed increased binding to proteins with exposed sulfurs and thiol-modified DNA compared to the *p*MBA protected clusters [11]. This property could be useful in bioimaging applications, and has been successfully shown in tracking an intracellular transaction [39].

The importance of the protonation state of the *p*MBA protected gold nanoclusters for the cluster solubility was discussed in a previous section. Yet the usability of this property was truly manifested in an observation of formation of ordered superstructures, self-assembled from Au_{102}*p*MBA_{44} clusters upon change of solvent conditions from water to methanol [40].

The structures produced in this manner included 2D sheets ordered in hexagonal packing, forming flakes of some sheets thick and 3D spherical shells of one cluster thick and $\approx$ 200 nm in diameter. The 3D shells were also able to form superstructures from superstructures, i.e., multiple shells bonded to each other [40].

Potential applications of such colloidal include functional materials that could have tunable plasmonic and mechanical properties, making them, e.g., potential drug carriers [40].

Gold nanoclusters have also been utilized as labels in virus imaging in two schemes: (1) covalent linking of functionalized $Au_{102}p$MBA$_{44}$ clusters to cysteine residues on the virus surface [41] and (2) Linking of functionalized $Au_{102}p$MBA$_{44}$ nanoclusters to non-covalently bound pocket factor located inside the hydrophobic pocket of the virus [42].

The advantage of using gold is the provided contrast over the protein material of the virus when imaged with TEM [41, 42]. Of especial advantage of using atomically precise gold nanoclusters is the site-specificity and robustness of the covalent and the non-covalent binding [41, 42]. For the remainder of this section, the non-covalent scheme (2) is considered and schematically presented in figure 5.

Echovirus 1 is an icosahedral virus, composed of 60 protomer units. Each protomer unit has a cavity called hydrophobic pocket, that surround the five-fold symmetry axis in the full structure. Each or some of the pockets are occupied by a pocket factor, palmitic acid in the natural state, bound in place by non-covalent interactions (figure 5, 1.). The pocket factors have a role in stability of the virus, while their removal is connected to the genome release. Thus, the pocket is also a viable drug target, in terms of replacement of the natural pocket factor with a drug molecule (figure 5, 1.-2.) stabilizing the virus capsid to prevent its opening. Generally, the uncoating mechanisms of viruses are, including the role of the pocket factors, not well understood, making the pocket a potential label target as well. Using the known working drug molecule Pleconaril as the basis, a new molecule was synthesized to enable both better binding to the pocket and the ability to connect, with an appropriate length alkyl chain (figure 5, 3.-4.), to a gold nanocluster (figure 5, 5.) that would reach out from the pocket on the cluster surface. Indeed, this binding scheme was proven viable in numerous types of experiments [42].

Especially it was concluded that the gold nanocluster binding does not prevent the infectivity of the virus, while increase in the stability against temperature increase was observed. Furthermore, the binding process was deemed dynamic, possibly due to virus breathing, and not all the pockets were simultaneously occupied by the gold nanocluster bound probes [42].

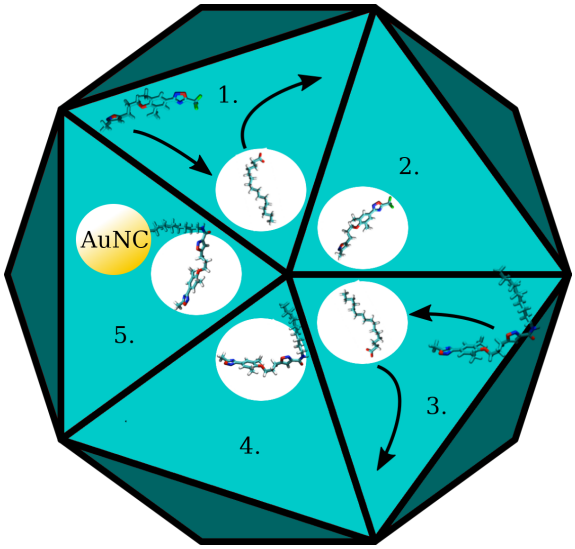


FIGURE 5 Schematic of non-covalent linking of gold nanoclusters to hydrophobic pockets of EV1: Natural pocket factor (palmitic acid) inside the hydrophobic pocket (symmetrically located around 5-fold symmetry axis) may be replaced by 1.-2. Drug molecule Pleconaril;l 3.-4. Pleconaril-based Kirtan1 molecule with a tail connectable to gold nanocluster (AuNC), 5.

3 MOLECULAR DYNAMICS SIMULATIONS

3.1 An atomistic scale laboratory on a computer

Molecular dynamics (MD) simulations may be described as experiments performed on a computer. That is, the system may be built up, and the conditions may be chosen to correspond to an experimental set-up. Furthermore, the conditions or systems potentially challenging to be built experimentally are in principle straightforward to set-up computationally. In the scope of this thesis MD simulations are considered at classical level and atomistic scale.

Once the system and the conditions are set-up, the simulation evolves in time according to Newtonian mechanics and to the model, i.e., potentials, selected to describe the interactions between the atoms. As a result, a trajectory of each atom or molecule in the system is obtained as a "molecular movie". Such an atomistic accuracy, and the time scale resolution of an MD simulation are beyond reach for experimental methods. Inherently however, the quality of the system description determines the quality of the data extracted from it [43].

As introduced previously, the purpose of the simulations can be seen to bridge the gap between theory and experiment [43]. The connection of the simulation to the theory naturally comes in as the rules according to which the simulation proceeds. The connection between the microscale simulation laboratory and the macroscale experimental laboratory, is not as intuitive.

The experiment most often contains many more molecules and takes a much longer time than the simulation. Thus the connection between the experimentally measured macroscopic properties and the simulated, microscopic properties, needs to be established. This crucial connection arises from statistical mechanics. From this point of view, the quality of the simulation is determined by the sufficiency of the statistical sampling [43].

MD simulations have proven to be an extremely useful tool in studying proteins [44–46], even up to size scale of full viruses [47, 48].

3.2 Statistical mechanics for molecular dynamics

The partition function $Z = \sum_i e^{\frac{-E_i}{k_B T}}$, sum over all system states, formally connects the thermal averages (quantities that are measured) to the microscopic properties of the system. Essentially then, in order to access the thermal averages, partition function Z needs to be obtained. Direct obtaining of Z is however an enormous task, due to numerous number of states some of which are physically irrelevant. Thus, one might envision partial sampling of Z. [43, 49]

Generally, the negative natural logarithm of the partition function defines the thermodynamic potential of an ensemble (e.g, Helmholtz free energy F in NVT ensemble, Gibbs free energy G in NPT ensemble), that has minimum at equilibrium. Thus, the form of the partition function depends on the ensemble. [43]

Classically, the microscopic state of the system of N molecules can be expressed in terms of positions $\overline{r} = (\overline{r}_1, \overline{r}_2, ..., \overline{r}_N)$ and momenta $\overline{p} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_N)$. This spans a $6N$ - dimensional phase space formed by positions and momenta. [43] In essence, the phase space average is the classical equivalent of the ensemble average. In a given ensemble, the distribution of points in the phase space at a given time is determined by the partition function Z of that ensemble. [50]

While experimental measurements are considered to be in fact time averages, natural way of thinking computing of corresponding quantities is also via time averaging. That is, the points Γ in phase space evolve in time $\Gamma(t)$. The time average of general thermodynamic property A is then [50]:

$$\begin{aligned} A_{\text{obs}} = A_{\text{macroscopic}} = \langle A \rangle_{\text{time}} = \langle A(\Gamma(t)) \rangle_{\text{time}} &= \lim_{t_{\text{obs}} \to \infty} \frac{1}{t_{\text{obs}}} \int_0^{t_{\text{obs}}} \mathrm{d}t A(\Gamma(t)) \\ &= \frac{1}{\tau_{\text{obs}}} \sum_{\tau=1}^{\tau_{\text{obs}}} A(\Gamma(\tau)) \\ &= \sum_{\Gamma} A(\Gamma)\rho(\Gamma) \end{aligned} \tag{1}$$

In practice, the time t_{obs} is finite and in simulations the time is replaced by discrete time steps τ, hence the switch from integral to sum on the second to last row of equation 1. Implied is that, in order to evolve the system in time, we need to know its initial state ($\Gamma(0)$). Furthermore, the underlying assumption is that the time-average of the equation 1 is independent of this initial state. If this holds, averaging over time is equal to averaging over the ensemble members, i.e., the last row of equation 1, where ρ is the probability distribution function of the ensemble, thus encompassing the partition function Z. [50]

The equality of averaging over time and ensemble members is called the ergodic hypothesis. These two averaging schemes represent the two ways of sampling the partition function: Molecular dynamics (MD) and Monte Carlo (MC), [50] respectively. The remainder of this thesis focuses on MD only.

3.3 Dynamics

Accessing the thermodynamical averages by averaging over the time, implies one needs a description for time evolution (dynamics) of the system in phase space. In classical MD, Newton's equations of motion are used for the task. [43,50]

Considering a classical system with electron motions averaged out based on Born-Oppenheimer approximation, the Hamiltonian can be written as a sum of kinetic $K(\overline{p})$ and potential $V(\overline{r})$ energies:

$$H(\overline{r},\overline{p}) = K(\overline{p}) + V(\overline{r}) \tag{2}$$

While the kinetic part $K = \sum_{i=1}^{N} \sum_{\alpha} \frac{p_{i\alpha}^2}{2m_i}$ is a simple sum over molecules N and momentum components α, the potential part V contains the intermolecular interactions. That is, V is a function containing sums over single atoms, atom pairs, triplets, quadruplets etc. Often, only single atom and effective pair potentials are included. [43,50]

According to Boltzmann distribution, weight of a conformation at temperature T follows $e^{-\frac{V(\overline{r})}{k_B T}}$. Considering this with respect to sampling of phase space as in equation 1, the situation may seem hopeless. In MD however the sampling is aided by the fact that the kinetic energy K enables crossing of barriers in order of $k_B T$. [51]

The time evolution of each atom's (index i) position and velocity according to Newton's equations of motion can be solved as:

$$\begin{aligned} \frac{\mathrm{d}\overline{v}_i(t)}{\mathrm{d}t} &= \frac{\overline{F}_i(\overline{r}_i(t))}{m_i} \\ \frac{\mathrm{d}\overline{r}_i}{\mathrm{d}t} &= \overline{v}_i(t) \end{aligned} \tag{3}$$

The forces $\overline{F}_i$ acting on each atom need to be computed from the gradient of V (the interaction potential) as

$$\overline{F}_i = -\frac{\partial V}{\partial \overline{r}_i} \tag{4}$$

In practice, the integration of equations in 3 is performed at small time intervals, Δt, called the timestep of the simulation; often in order of femtoseconds. The timestep ought to be chosen small enough to describe the highest frequency motions in the system. In order to increase the time step and thus simulation efficiency, it is customary to constrain certain bonds lengths or angles. This operation decreases the fastest frequencies and thus allows for the time step increase. [51]

3.4 Force fields

The interaction potential V holds within the approximations applied to describe the system at sufficient accuracy. For example, in classical MD V includes the electronic effects in an effective way. In general, V may be cast in various different forms, with parameters obtained in numerous different ways. The different interaction potentials are called force fields. [51]

Force fields come in different flavors, that often suit specific purposes. That is, for example, range of types molecules or conditions the force field is applicable for. More fundamentally, the force field defines the functional form, i.e., which interactions are included and how. The philosophy behind obtaining the parameters associated with the functional form of the force field may also differ, from fitting to experimental or higher-level computational data. As a result, these parameters are generally not interchangeable between different force fields. [51]

The functional form of a classical force field for molecules contains terms for bonded and non-bonded interactions. Often, the bonded interactions contain terms up to four-body interactions: bonds, angles and proper and improper dihedral angles. Non-bonded interactions are effective and pairwise consisting of van der Waals and electrostatic terms:

$$
\begin{aligned}
V = & \sum_{\text{bonds}} k_b(b-b_0)^2 + \sum_{\text{angles}} \frac{1}{2}k_a(a-a_0)^2 \\
& + \sum_{\text{dihedrals}} k_d(1+\cos{(nd-\delta)}) + \sum_{\text{impropers}} \frac{1}{2}k_i(i-i_0)^2 \\
& + \sum_{i,j>i} \left(\frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^{6}} \right) + \sum_{i,j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}
\end{aligned}
\tag{5}
$$

Bond, angle and improper dihedral potentials are harmonic potentials with force constants k_a, k_b, k_i and equilibrium or reference values a_0, b_0, i_0. These interactions exist between bonded atoms and are two, three or four -body, respectively. The improper dihedral potential ensures the planarity of, e.g., ring structures. The proper dihedral potential, a four-body interaction, describes the transitions of torsions (e.g., from *cis* to *trans*) and follows a sinusoidal form. [51]

The last row of equation 5 is the non-bonded part, consisting of 6-12 Lennard-Jones potential and Coulomb parts. The parameters C_{12} and C_6 encompass element-specific parameters, depth of the potential and the minimum distance, that are combined with so-called established combination rules. Coulomb potential contains partial charges, q_i and q_j, of an atom pair. ϵ_0 is the vacuum permittivity. Both these potentials are between pairs of atoms, with interatomic distance r_{ij}, thus describing the effect of van der Waals and electrostatic interactions effectively (not including triplets, quadruplets etc.). [43, 51]

In practice, summation over all pairs is truncated by a cut-off distance between the atoms to reduce computational cost. The Coulomb interaction is however long-ranged in nature, i.e., longer range than the typical cut-off. Inclusion of this long-range part is crucial, and various methods, such as Ewald summation, have been developed for the purpose and not discussed here in detail; merely mention that using Ewald summation requires an overall neutral simulation box to avoid artifacts. [51]

3.4.1 Amber force fields

One of the widely used standard force-fields for protein and nucleic acid simulations is Amber [52]. It follows the form of equation 5, and both quantum chemistry calculations and experimental data of amino acids or peptide fragments have been used for parameter fitting [53].

Reference bond and angle values are based on crystal structures, and force constants adapted to reproduce normal modes of different peptide fragments. Torsion parameters are adapted to match experimental or computed torsional barriers, or experimental conformational populations. Partial charges of atoms are obtained from (restricted) fits to computed quantum electrostatic potential. Lennard-Jones parameters are based on reproducing the densities and heats of vaporization in liquid simulations. These parameters have been improved over the course of time. [53]

While Amber force fields were originally parameterized for proteins and nucleic acids, the development of GAFF [54] (Generalized Amber Force Field) enables Amber-compatible parameterization of standard organic molecules to complement the simulations, e.g., with ligands and drug molecules [53].

3.5 System and environment

In this section, certain system and simulation set-up specifics relevant for the work presented in this thesis are summarized.

In the framework of the previous sections, MD is here applied considering all atoms (nuclear coordinates) as point masses carrying partial charges, interacting through equation 5 - like function. That is, the bonds are treated with various harmonic terms and non-bonded interactions are Lennard-Jones and Coulombic type. The parameters of equation 5 follow Amber force field philosophy. The description is classical and chemical reactions or bond breaking cannot occur. Dynamics follow classical Newton's equations of motion.

3.5.1 Solvent

One of the conditions crucial in experiments and thus of importance in simulations as well, is the solvent environment. For biomolecules, the solvent is usually water with some concentration of ions.

The solvent may be presented implicitly or explicitly. In the works of this thesis, the latter approach is applied. There are numerous different explicit water models [44], e.g., TIP3P [55], parameterized to reproduce correct water density and heat of vaporization in simulations.

Parameters for numerous solvents in addition to water, for different force fields have been parameterized and readily available, for example in Virtualchemistry.org [56, 57]. In general, ion parameters are also available for any force field.

3.5.2 Boundaries

The simulation box necessarily has boundaries. While the system size or simulation box is preferably kept as small as possible, the nature of the box boundary may have significant effects on the system. Essentially, the simulation is prone to suffer from surface effects and thus not appropriately describe a macroscopic bulk system. [43, 44, 50, 51]

The problem is often circumvented by using periodic boundary conditions (PBC), meaning that the simulation box is infinitely repeated through space and thus no wall or surface atoms between systems exists. [43, 44, 50, 51]

Naturally when applying PBC the size of the central box must be chosen according to the range of the interactions, to prevent each atom from interacting with its periodic image. In practice also so-called minimum image convention is applied: pair-interactions need only be counted once, i.e., between nearest neighbors, not between all periodic images. [43, 44, 50, 51]

3.5.3 Ensemble

In the spirit of mimicking the experimental conditions in simulations, one could envision perhaps the most fundamental conditions to be controlled, or kept constant, are temperature and pressure.

Several established methods to achieve MD simulations with constant temperature and / or pressure (NVT or NPT ensemble), exist and not presented here in detail.

Generally in the simplest of cases, for temperature control the velocities of atoms may be scaled (i.e., kinetic energy corresponds to the target temperature). For pressure control, the size of the simulation box may be scaled. Formally such methods do not yield known ensembles, yet the average temperature and pressure, while fluctuating, are at target values. [43, 44, 50, 51]

3.6 Limitations and applicability of MD

The applicability or correctness of MD simulations are limited by two factors: Accuracy of force field and the extent of sampling. [51]

Force field challenges relate to its functional form and parameters to that end (equation 5). First, systems or phenomena for which the classical description is sensible can be studied. The same applies for the parameters that, in accordance with the force field, are suitable and fitted with specific types of properties and conditions in mind, and are regularly updated and improved. [53]

Issues in sampling can be understood in terms of equation 1, where formally $t \to \infty$. Since t is of course finite, ergodicity is not guaranteed. Nevertheless, improved sampling requires longer simulation time, and naturally the phase space is bigger for more complex systems. In practice, one aims for sufficient simulation length that generates a representative ensemble of the system. [43, 44, 50, 51]

Also in equation 1 it is assumed that an equilibrium ensemble is sampled. Therefore it is customary to allow for a pre-production equilibration period to stabilize the system and exclude initial large fluctuations. Again depending on the system and the property in question, this may take long time, and is not necessarily easily realized. [44]

3.7 Accessible properties

Once the simulation trajectory is obtained, useful data need to be extracted from it: one wants to tie the simulation properties to the experimental ones. That is, to ensure simulations are applicable in describing the system and properties in question.

Perhaps the most obvious way of assessing the simulation trajectory, is investigating the time evolution, averages, drifts and fluctuations, of a generic property A, that could relate to the structure or energetics of the system. Ideally, A has, be it simple or complex, connection to an experimentally measurable property. Alternatively, comparisons can be made to data obtained from a higher-level computation.

Implied by equation 1, thermodynamic properties, such as internal and kinetic energies, temperatures and pressures, may be estimated as time averages of the simulations. Furthermore, properties dependent on fluctuations in thermodynamic quantities are accessible. Pair distributions of averaged atomic distances may be used to describe structures and time correlations may be connected to experimental transport coefficients and spectra. [43, 51, 58, 59]

Examples of experimental techniques that may be connected to simulation data include IR or Raman spectroscopy (normal mode analyses), X-ray scattering (atomic positions, mobilities, occupancies), NMR (orientations, distances, rate constants). [59]

While many biological processes are still out of reach for simulation times currently available, experimental information may be utilized in simulations not only by means of structures, but as guidance to steer or enforce the simulation to a certain direction or conformation. [44]

The following subsections summarize the type of properties relevant for the works of this thesis.

3.7.1 Structures and interactions

Considerations of simple distance based interactions, such as hydrogen bonds that are easily obtained from simulations, may be used for arguing for structure (in)stability and reveal interaction patterns impossible to probe in detail in experiments. In the same spirit, several structural characteristics relating to, e.g., system size and surface area are available at truly atomistic scale.

Also dynamical behavior, that is relevant for the simulation time scale, may be extracted and compared to experimental time scales or mobilities, or computed barriers from a higher level of theory.

3.7.2 Diffusion coefficient

One experimentally measurable transport coefficient is the diffusion coefficient D. In the hydrodynamic model D is related to mean square displacement (MSD) of atoms, straightforward to obtain from an MD simulation, by Einstein diffusion relation [58]:

$$D = \lim_{t\to\infty} \frac{\overbrace{\left\langle (\overline{r}_i(t) - \overline{r}_i(0))^2 \right\rangle}^{\text{MSD}}}{6t} \tag{6}$$

Diffusion coefficients are prone to suffer from effects of the simulation box periodicity. In a periodic system the diffusing particle is affected by hydrodynamic (self)-interactions stemming from the surrounding boxes. Thus, the diffusion coefficients calculated from MD simulations are strongly box-size dependent. The most straightforward correction attempt to this is performing simulations in different box sizes, followed by an extrapolation to the infinite dilution. [60]

3.7.3 Free energy of binding

In simulations often free energies of solvation or binding are considered. Binding free energies are of particular interest in the field of drug design, with the ability to rank between the optimal binding between different molecules and investigate the detailed contributing factors or interactions. In the following, basic concept of free energy along with methods for its computation are presented. [44]

3.7.3.1 Basic concepts

Free energy and entropy are fundamental concepts in thermodynamics considering stability and equilibrium, related to the negative logarithm of the partition function as described in section 3.2. Ultimately, relative stability between systems under given conditions is determined by the free energy difference [61]. Absolute free energies however are not accessible via statistical averages and thus impossible to compute for complex systems [51, 61]. Free energy differences on the other hand are obtainable through statistical averages, and thus computable from simulations [50].

Direct determining free energy differences between states A and B, proportional to $\log \frac{Z_B}{Z_A}$, is however also challenging, due to insufficient, and more importantly lesser sampling of some state(s) of interest compared to the other(s). Several different established methods for free energy estimation, even for complex and non-equilibrium systems have been developed and presented in the following. [61]

3.7.3.2 Transitions between states

The time scale of standard simulations of large molecules rarely allow witnessing transitions between states of interest (say, folded and unfolded protein, bound and unbound ligand) spontaneously. Such challenges may be overcome by use of alchemical or steered MD, enforcing the transition between states, either slowly (reversibly) or fast (irreversibly). [44]

One could consider transition between states A and B via alchemical or steered reaction coordinate route. Alchemical transition refers to changing the nature of the system's atom(s) in terms of the force field parameters and thus the interaction potential and the Hamiltonian. In the reaction coordinate case transition between states is via, e.g., change in distance between atoms. [44, 62]

Different techniques are available for extracting the free energy differences from such transitions. Techniques include for example thermodynamic integration (TI), free energy perturbation (FEP) and non-equilibrium methods based on Jarzynski's equality. [62]

3.7.3.3 Equilibrium techniques

TI makes use of the fact that, for a fixed value of reaction coordinate or alchemical parameter λ, derivative of the free energy is obtained as a statistical average, provided the system is at equilibrium. In practice, derivatives are computed for a sequence of λ values ranging slowly or reversibly from 0 to 1 (system A to B). The full free energy difference profile of the $A \rightarrow B$ transition is then recovered by integrating these derivatives. [62]

In FEP method the relation of the free energy difference to the exponential difference in Hamiltonian's of the two states H_A and H_B is utilized. One may consider the H_B as being perturbed with term V_{BA}: $H_B = H_A + V_{BA}$. Now V_{BA} may be computed from the ensemble of system A for both states A and B (or vice versa). In practice, if the phase space overlap of the two states is low, for sufficient overlap or small enough V_{BA} for trustworthy sampling, the transformation needs to be stepwise, i.e., the perturbation is proportional to the number of steps n as $\frac{V_{BA}}{n}$. In addition, an estimator called Bennett acceptance ratio is often applied to reduce the number of the needed transformation steps; analogous are different histogram methods not discussed here in detail. [62]

3.7.3.4 Non-Equilibrium techniques

Considering TI now in terms of irreversible transformation, the system is driven out of equilibrium and the work performed is larger than the free energy difference between the states, constituting of both free energy difference and heat dissipation [44, 61–64]:

$$\langle W(t) \rangle \geq \Delta G \tag{7}$$

Jarzynski's equality [65] states the relationship between the exponential average of the work and free energy difference [44, 61–64]:

$$\left\langle e^{-\beta W(t)} \right\rangle = e^{-\beta \Delta G} \tag{8}$$

The average entails the ensemble average over initial conditions and the work relates to the change in Hamiltonian in transition from state A to B. [44,61–64]

More generally, free energy difference also has a lower boundary [44,61–64]:

$$-\langle W(t)\rangle \leq \Delta G \leq \langle W(t)\rangle \tag{9}$$

Suggesting the relation between free energy and work performed in forward (probability density p_{forward}) and backward (probability density p_{backward}) transitions. [44,61–64] Indeed, this relation was shown by Crooks [66,67]:

$$\frac{p_{\text{forward}}(w = W(t))}{p_{\text{backward}}(w = -W(t))} = e^{\beta(w-\Delta G)} \tag{10}$$

Conceptually these equalities state the connection to the equilibration free energy of the arbitrarily fast transition from equilibrated state A to the state B not at equilibrium, provided there are enough realizations for the exponential to converge in equation 8. As the transformation rate increases, sampling of small work values that weigh most in the exponential average, decreases, thus one needs more realizations for converging the exponential average. Thus, whether it is computationally cheaper to perform many enough fast realizations of the transition or resort to methods requiring equilibrium and slow transforming to ensure overlap between the states, is ambiguous. [44,61–64]

Direct application of Jarzynski's equality is known to suffer from convergence issues. Different schemes for treating the issues of the sampling bias have been suggested, simplest of which (assuming the work distribution is gaussian) are based on simple cumulant expansions [63]. For non-gaussian work distributions, resampling and block averaging schemes have been suggested [68,69].

Jarzynski's equality and Crooks relation have been applied and validated in numerous types of systems, both experimentally (single molecule stretching) and computationally [70]. Perhaps the most famous demonstration is the equality between the experimental free energies obtained from the work profiles of reversible and several irreversible (non-equilibrium) stretching events of an RNA molecule [71,72].

3.8 Software

In order to perform simulations described above, in addition extensive computer resources, an efficient algorithm machinery is needed.

One of the widely used and most versatile readily available MD packages to efficiently set-up, perform and analyze MD simulations is Gromacs [73–78]. It is compatible with standard force fields and different types of dynamics and energy minimization schemes are available, along with steered simulations and free energy calculations, among many other features. Specifications for ensembles, boundaries, solvent models, cut-off schemes, constraints etc. can be made straightforwardly and there are several options. Running in parallel enables treating large systems for relevant time scales.

3.9 Molecular dynamics simulations of gold nanoclusters: a literature review

The aim of this section is to provide a short literature review on generally covered topics in relatively recent MD simulations, including atomistic, coarse-grain and QM/MM, of gold nanoclusters. While DFT is and has been often the method of choice in computational investigations of gold nanoclusters, the following results demonstrate the potential and applicability of MD simulations in many respects of gold nanocluster research, complementing and adding to the experimental data and also providing predictive prospects. In addition to structural and environmental investigations, different free energy methods have been actively employed to study, e.g., aggregation and membrane intrusion.

In terms of structures and properties, MD simulations have been applied in shedding light on effects caused by, e.g., ligand shell coverage density [79] and ligand size or composition [80–82]. In addition, the effects of gold nanocluster size and ligation on interactions between the nanoclusters [83] and between gold nanoclusters and solvents or ions [84–87] have been studied with MD simulations.

Self-assembling [88] packing characteristics, along with aggregation behavior [89–91] dependent on ligand and core properties have also been studied utilizing MD simulations. Such aggregation properties can be further applied in numerous sensing applications, as suggested by different simulations [92–95] as well. Considering biological applications more closely, gold nanocluster aggregation can be utilized in detecting target DNA, with the aggregation behavior then naturally of paramount importance, which can be studied in detail by MD simulations [96].

In terms of biological and medical applications, interactions of gold nanoclusters with proteins and membranes is crucial. Such interactions are governed by the surface properties of the nanoclusters. That is, appropriately functionalized gold nanoparticles could be utilized in specific binding to e.g., receptors, once the interactions are well understood. To that end, numerous studies have been performed investigating in atomic detail the effects of nanoparticle properties on its interactions and reactions with different proteins and thus the protein structure [97–102], along with binding of different peptides on nanocluster surfaces [103, 104].

Interactions between nanocluster and membranes is a widely studied topic, due to its implications for using nanoparticles in physiological environments or cells. To this end, MD simulations have revealed important and interesting properties showing dependence of interactions and binding, fusion and intrusion on, e.g., nanocluster size, charge and surface properties along with external effects. On many occasions, different free energy methods are applied in detailed studies of membrane penetration [25–29, 31–33, 105–120].

4 RESULTS AND DISCUSSION

In this chapter the main results of the papers **I-III** along with unpublished results for gold nanocluster self-assemblies are presented. Several experiments of gold nanocluster systems motivated us to advance the knowledge from these experiments by simulation methods. Especially, the aim was to investigate gold nanocluster dynamics in solvent environments (aggregation behavior) and in a biolabeling application (virus or protein environment). Furthermore, simulations would enable fast exploring of numerous structures in complementing experimental structure solving of new gold nanoclusters. In order to achieve dynamics simulations in solvent and biological environments, along with numerous structures to scan through, methods beyond DFT need to be applied. For this task we have chosen MD simulations.

The foundation for performing MD simulations for gold nanoclusters is laid by parameterizing and validating the force field for gold nanoclusters, work **I**, discussed in chapter 4.1. The applicability and generalizability of the force field is then demonstrated in work **II** (chapter 4.2), alongside experimental and other computational methods, by shedding light on structures of experimentally unsolved ligand shells of new gold nanoclusters. Simulation system complexity is further increased in simulating gold nanocluster self-assemblies (unpublished results, chapter 4.3) and interactions with virus (paper **III**, chapter 4.4).

4.1 Generalized gold nanocluster force-field: parameterization and validation (paper I)

While few force field parameterizations for gold nanoclusters are available, the usual philosophy is the applicability to one type of cluster only. Due to increasing number of gold nanocluster structures emerging, we aimed at generating a force field that would be easily transferable to different sized thiolate-protected gold nanoclusters. Due to our interest in simulating gold nanoclusters in combination with proteins and in different solvent environments, along with relatively straightforward parameterization procedure, the parameters were generated to be compatible with Amber force fields.

As reference systems we utilized four different experimentally solved and one computationally produced gold nanocluster structures, with four different symmetric core sizes from 25-144 gold atoms, with two types of thiolate ligands PET and *p*MBA. The structures are presented in figure 2. To this end, no core-specific bonds are introduced between gold atoms and all gold atoms carry zero partial charges. That is, the gold atoms in the cluster are held together by Lennard-Jones potential only. We have utilized Lennard-Jones parameters developed by Heinz *et al.* [121] for gold surfaces. The strength of the LJ potential between gold atoms using these parameters was proved sufficient to keep the gold core symmetry stable under 300 K temperature and 1 bar pressure in the 100 ns timescales of our test simulations.

The connection between gold and sulfur in the reference systems comes in the form of rectangular or V-shaped units often called staples, figure 3. Since the staples are a conserved form in several different gold nanocluster structures, we selected a set of parameters for the bonded interactions (bonds, angles, torsions) that could be transferable between the same staple type. In practice, the reference values for these parameters are taken as averages from structures of different gold nanoclusters, those shown in figure 2, while force constants were adopted from Heikkilä *et al.* [86] and Banerjee *et al.* [122].

Finally, one needs parameters for the thiolate ligands, that in our cases were standard organic molecules (R), for which Amber force field compatible atom types and most of the bond, angle and torsion parameters were readily available. Partial charges needed to be evaluated, for which we used RESP, a procedure recommended for Amber compatible parameterization. Charge fitting was done on a staple-like unit, with charges on gold atoms restricted to zero. The schematic of the parameter construction presented in figure 6.

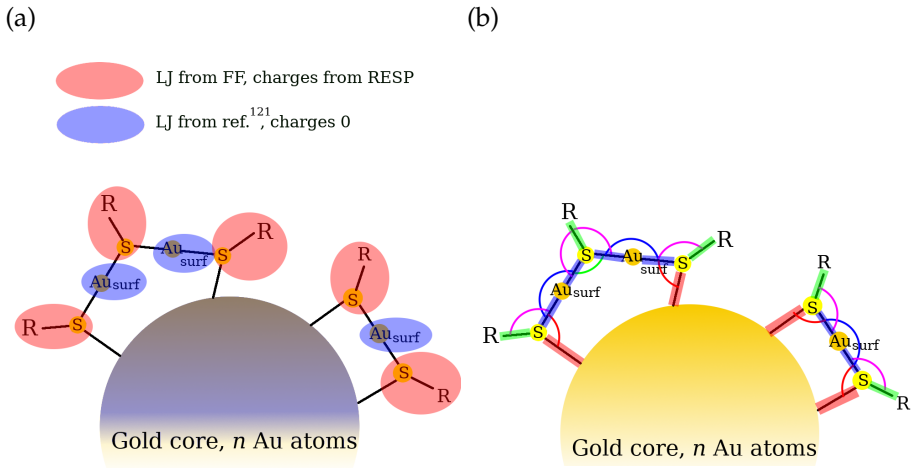


FIGURE 6 Schematic representation of (a) non-bonded and (b) bonded parameters of gold nanoclusters. The R group bonded parameters obtained from Amber force field. In (b) the colored arches show the defined angles and rectangles the bonds. Torsion angles are not drawn.

For validation of the force field we have compared results obtained from the simulations to (1) available experimental results and (2) DFT, i.e., higher level computational results. As the very first measure we estimated the force field applicability based on structural stability of the structures.

Simple investigation of the MD trajectories revealed that in all the studied cases symmetry and gold core structure remained intact, while especially the V-shaped staples showed fluctuations. While comparisons of MD to MD/DFT simulations revealed no clear differences between the sampled geometries, due to the difference in time scales of the two methods the physicality of the V-shaped staple fluctuations could not be concluded.

The classical MD trajectories for all the studied clusters also revealed flipping of ligands around the sulfur atoms in the staple units. DFT calculations were used to estimate the energy barriers for the flipping, implying that flipping in room temperature could be expected.

In terms of validating against experimental results, we turn to diffusion coefficients and crystal structure simulations. Diffusion coefficients were computed for three different PET protected clusters in organic solvents, for which experimental data was available. Six different box sizes were used for each cluster in order to extrapolate to infinite dilution and thus eliminate the periodicity effects. The results are summarized in table 1. While computed values are consistently smaller compared to the experimental ones, the trend is consistent with the experimental data, suggesting the force field captures the trends in cluster behavior in solvent environment.

TABLE 1 Computed and experimental diffusion coefficients. The computed result given as an extrapolate to infinite box size. Adapted with permission from Emmi Pohjolainen, Xi Chen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *Journal of Chemical Theory and Computation*, 12 (3), pp. 1342–1350, **2016**. Copyright 2016 American Chemical Society.

System	D (10^{-5} cm^2/s)	$D_{\text{experimental}}$ (10^{-5} cm^2/s)
$Au_{25}PET_{18}^-$	0.373 ± 0.09	0.484 ± 0.04
$Au_{38}PET_{24}$	0.274 ± 0.07	0.374 ± 0.03
$Au_{144}PET_{60}$	0.251 ± 0.04	0.356 ± 0.03

Finally, we attempted mimicking the experimental crystallization conditions of one of the gold nanocluster structures, and simulated a periodic unit cell with 72 clusters. Simulations in the crystallization mimicking conditions stabilized while pure water solvent destabilized the crystal structure, shown in figure 7. This further implies the force field describes the inter-cluster and cluster-solvent interactions in a meaningful way.

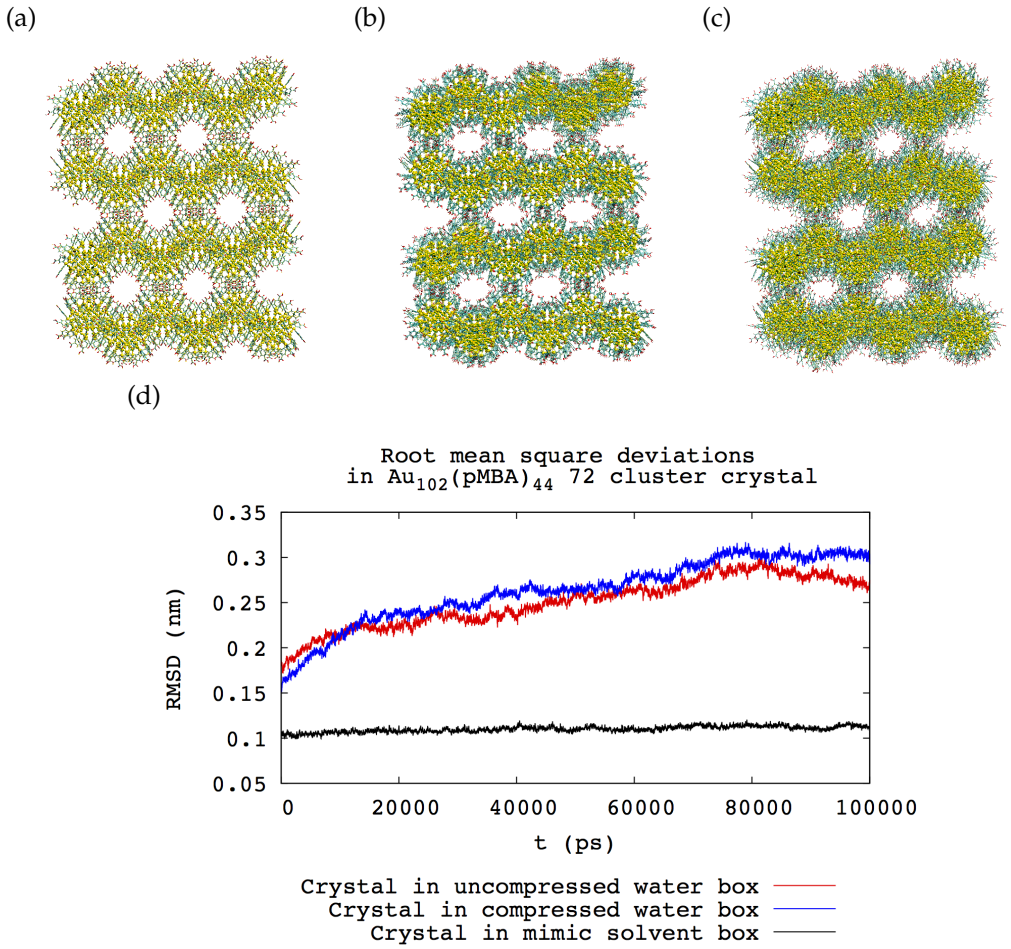


FIGURE 7 Original crystal structure in 72-cluster unit cell (a) and end structures of 100 ns simulations with (b) solvent condition mimicking reported [13] crystallization conditions and (c) pure water solvent. (d) shows the RMSD during simulations in mimic and water solvents (solvent molecules not shown for clarity). Adapted with permission from Emmi Pohjolainen, Xi Chen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *Journal of Chemical Theory and Computation*, 12 (3), pp. 1342–1350, **2016**. Copyright 2016 American Chemical Society.

Summarizing, we have demonstrated the applicability of the force field in describing to sufficient accuracy intra-cluster and inter-cluster interactions, along with cluster-solvent interactions. This enables investigating gold nanoclusters by simulations at size and time scales, numerous clusters in solvent and protein environments, that are beyond reach of DFT methods. In the following sections, we further demonstrate the applicability of this force field in gold nanocluster systems of increasing complexity.

In addition to the works presented here, the force field parameters presented have been utilized in works of others as well. Such works include interpreting NMR spectra with help of MD simulations [36], aiding in gold nanocluster multimer characterization [123], revealing interaction patterns in gold nanocluster sensing applications [92–95], and in determining protonation patterns on the ligand shell from constant pH MD simulations [24]. These force field parameters have also been reported being used in producing Amber compatible parameters in NanoModeler program [124].

4.2 Ligand shell structure prediction of *m*MBA -protected gold nanoclusters (paper II)

The force field parameterization was performed with respect to gold nanoclusters for which complete experimental structures, including the gold core and the ligand shell were available. In the water-soluble cluster structure used in the parameterization (Au_{102}*p*MBA_{44}), the type of the ligand is *para*-mercaptobenzoic acid (*p*MBA). Moreover, the clusters referenced in the force field parameterization had a symmetric or close-packed gold core.

Subsequent works have reported synthesis and characterization of three different sized stable gold nanoclusters with *m*MBA ligand [11, 12, 38]. The difference to *p*MBA is the position of the carboxyl acid group in the ring. While the structures of the gold cores for Au_{68} and Au_{144} sized clusters were solved, shown in figure 4, the ligand shell structure (and also the ligand count for the Au_{144} cluster) were undetermined. Interestingly, the structures of the gold cores of these clusters were clearly more asymmetric compared to those with *p*MBA ligand.

Reactivity of *m*MBA protected gold nanoclusters to thiol-modified DNA and proteins with exposed cysteines, and thus sulfur was also reported [11]. In terms of interactions with the environment and thus e.g., such life-science applications, the ligand layer is naturally of utter importance. To this end, various experimental and computational techniques were employed in this work to shed light on the structure and dynamics of the ligand layers of these two *meta*-mercaptobenzoic (*m*MBA) acid protected clusters.

The ligand count for Au_{68} was previously determined to be 32 by mass spectrometry [38], and supported by the NMR studies of this work. For the Au_{144} cluster the ligand count based on mass spectrometry was previously estimated at around 40 (unresolved by NMR). Based on NMR spectra both *m*MBA protected clusters Au_{68} and Au_{144} show no symmetry in the ligand shell.

The estimated ligand count in Au_{144}*m*$MBA_{\approx 40}$ is significantly lower compared to known structure of $Au_{144}PET_{60}$. Differences are also seen in the NMR and UV-VIS spectra of these two clusters with the same core size.

All this indicates the change of ligand from *para*-mercaptobenzoic acid to *meta*-mercaptobenzoic acid ligand, while still resulting in stable clusters, significantly affects the structure and properties of the gold nanoclusters. With an aim to shed light on the ligand shell structure and dynamics, we turned to MD simulations.

MD simulations enabled us to simulate numerous models of Au_{68}*m*MBA_{32} and Au_{144}*m*$MBA_{40\text{-}53}$ nanoclusters, in different solvent conditions and with varying protonation patterns. The model structures, i.e., the ligand positions (and number), were generated with an algorithm based on known existing gold nanocluster structures [125]. Few of the simulated models are presented in figure 8.

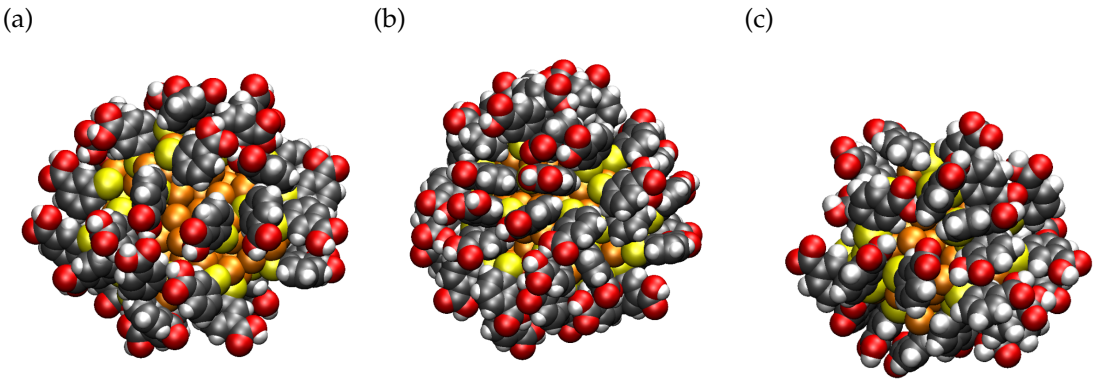


FIGURE 8 Model structures for MD simulations. (a) and (b) show Au_{144}*m*MBA_{40} and Au_{144}*m*MBA_{53}, respectively. (c) shows one of the models for Au_{68}*m*MBA_{32}. Atom colors: C=gray,O=red, H=white, S=yellow, Au=gold. Reprinted with permission from Tiia-Riikka Tero, Sami Malola, Benedek Koncz, Emmi Pohjolainen, Saara Lautala, Satu Mustalahti, Perttu Permi, Gerrit Groenhof, Mika Pettersson, Hannu Häkkinen. Dynamic Stabilization of the Ligand-Metal Interface in Atomically Precise Gold Nanoclusters Au_{68} and Au_{144} Protected by meta-Mercaptobenzoic Acid. *ACS Nano*, 11 (12), pp. 11872–11879, **2017**. Copyright 2017 American Chemical Society.

For simulations the *m*MBA was parameterized similar to *p*MBA in work **I**. Due to asymmetricity of the gold cores the gold atom positions were restrained to those presented in the experimental structures. Only Au-S bonds were adopted from the force field bonded interactions of work **I**, since no information on conserved staple-like units for these clusters was available.

Several different types of weak interactions or conformations between ligands and between ligands and gold were recognized from the simulation data of 50 ns runs, presented in figure 9. These included hydrogen bonds and aromatic contacts between the ligands, along with so called π-Au interaction and O$=$C$-$OH$\cdots$Au interaction. While especially the large number of aromatic contacts indicate efficient packing of ligands in all systems, of special interest are the interactions between ligands and gold. In the frame of the force field description, with no charges on Au atoms, the apparent interactions between ligands and gold need to stem from solvent effects and are thus referred to as conformations. The nature of these conformations as weak interactions between ligands and gold was subsequently confirmed by DFT calculations and experimental data.

In simulations conformations where the aromatic ring of the *m*MBA group lies flat on the surface of the gold cluster were observed, referred to as π-Au interaction, figure 9(e). According to DFT computations the flat lying conformation was energetically more stable than the up-standing conformation. Two arrangements of the carboxyl acid group, syn and $anti$ (figure 9(a), (b), were also observed in simulations. While the syn conformation is generally more dominant species in ambient conditions, the solvent environment is known to alter this balance. Moreover, for both syn and $anti$ conformations dynamic or switching conformations were observed with the OH group occasionally lying close (0.23-0.27 nm) of a surface Au atom, referred to as O$=$C$-$OH$\cdots$Au conformation and a hydrogen bond-like interaction (figure 9(c), (d)), recently also confirmed by others [126]. Similar conformations were not observed for reference systems containing *p*MBA ligands.

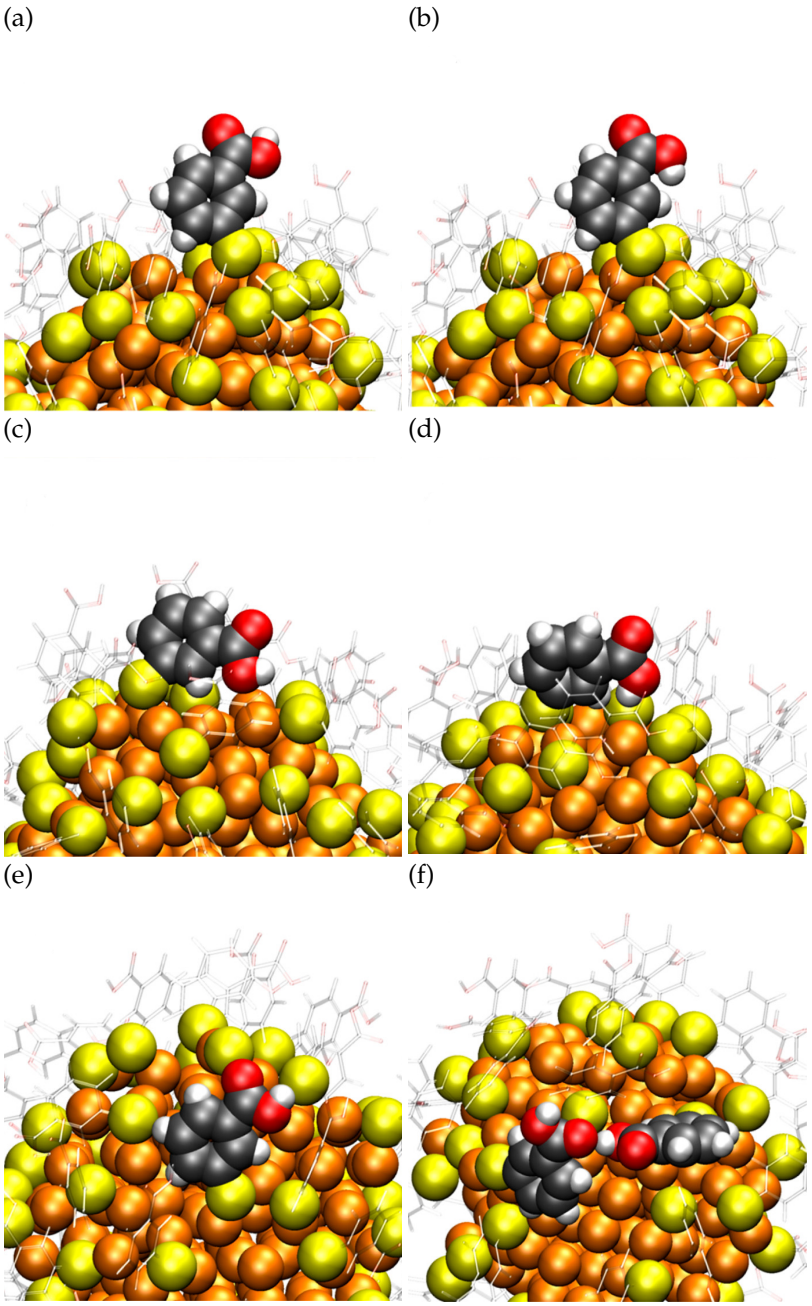


FIGURE 9 Ligand conformations observed in simulations. (a) and (b) show *syn* and *anti* conformations of the COOH group, respectively. In (c) and (d) these conformations lie close to the cluster surface. (e) shows a ligand in *syn* conformation lying flat on the surface and (f) demonstrated the interligand hydrogen bond. Only one ligand is shown for clarity. Atom colors: C=gray,O=red, H=white, S=yellow, Au=gold. Reprinted with permission from Tiia-Riikka Tero, Sami Malola, Benedek Koncz, Emmi Pohjolainen, Saara Lautala, Satu Mustalahti, Perttu Permi, Gerrit Groenhof, Mika Pettersson, Hannu Häkkinen. Dynamic Stabilization of the Ligand-Metal Interface in Atomically Precise Gold Nanoclusters Au_{68} and Au_{144} Protected by meta-Mercaptobenzoic Acid. *ACS Nano*, 11 (12), pp. 11872–11879, **2017**. Copyright 2017 American Chemical Society.

The appearance of syn and $anti$ conformations of mMBA in these clusters could be confirmed by the experimental (figure 10) and DFT IR spectroscopy data of these systems, where for both $Au_{68}mMBA_{32}$ and $Au_{144}mMBA_{\approx 40}$ clusters three peaks are observed in the carboxyl acid region. In contrast, both free mMBA and $Au_{102}pMBA_{44}$ show only two peaks. DFT computed frequencies of different geometries were used in assigning the peaks. The two lowest frequency peaks of $Au_{68}mMBA_{32}$ and $Au_{144}mMBA_{\approx 40}$ clusters qualitatively matched between the free mMBA and $Au_{102}pMBA_{44}$ cluster, corresponding to syn and $anti$ conformations of the carboxyl acid groups. The highest frequency peaks of $Au_{68}mMBA_{32}$ and $Au_{144}mMBA_{\approx 40}$ clusters based on DFT computations were assigned to the $anti$ conformation where there is a close contact between Au and the OH group, i.e., the O$=\!=$C$-$OH$\cdots$Au interaction.

In addition to investigating the interactions, simulation structures were utilized in DFT computation of the UV-VIS spectrum for $Au_{144}mMBA_{40}$ and $Au_{144}mMBA_{53}$ clusters. Supporting the lower ligand count, the computed spectrum for $Au_{144}mMBA_{40}$ shows a better agreement with the experimental spectrum, figure 10 (b).

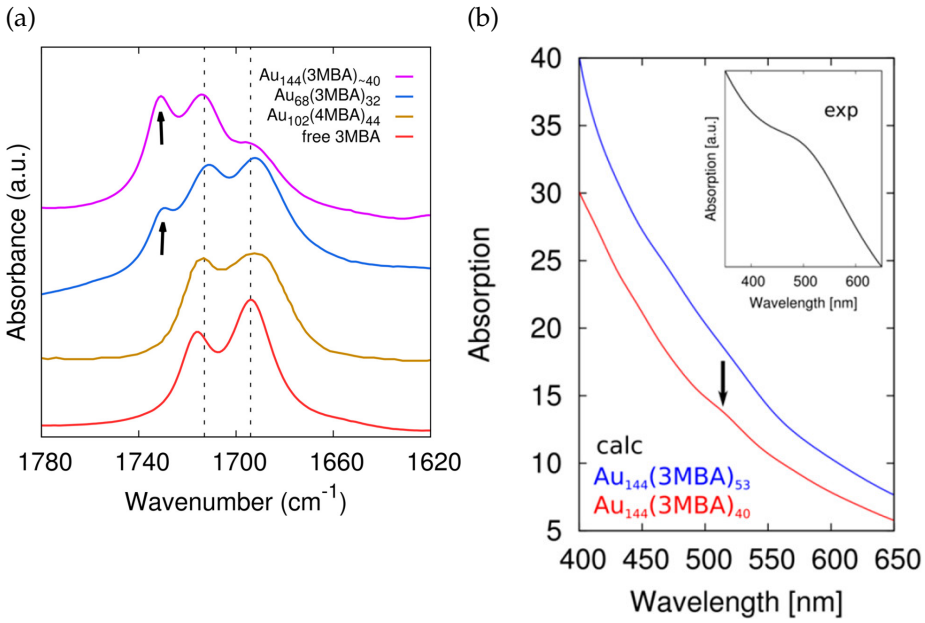


FIGURE 10 3MBA and 4MBA in the legends refer *m*MBA and *p*MBA, respectively. Experimental (a) IR spectra for *m*MBA protected gold nanoclusters (purple Au_{144}*m*$MBA_{\approx 40}$. The arrow indicates the peak interpreted as the O$=$C$-$OH$\cdots$Au interaction present for *m*MBA protected clusters only; (b) DFT computed UV-VIS absorption spectra for *m*MBA protected gold nanoclusters with structures obtained from MD simulation snapshots: Au_{144}*m*MBA_{40} (red) and Au_{144}*m*MBA_{53} (blue). Inset shows the experimental spectrum of Au_{144}*m*$MBA_{\approx 40}$. Reprinted with permission from Tiia-Riikka Tero, Sami Malola, Benedek Koncz, Emmi Pohjolainen, Saara Lautala, Satu Mustalahti, Perttu Permi, Gerrit Groenhof, Mika Pettersson, Hannu Häkkinen. Dynamic Stabilization of the Ligand-Metal Interface in Atomically Precise Gold Nanoclusters Au_{68} and Au_{144} Protected by meta-Mercaptobenzoic Acid. *ACS Nano*, 11 (12), pp. 11872–11879, **2017**. Copyright 2017 American Chemical Society.

The dynamic nature of the weak ligand-Au interactions described above, could explain the reactivity of *m*MBA protected gold nanoclusters towards thiol-modified DNA or proteins with exposed cysteines [11], namely by exposing / providing gold sites for these molecules to react.

The results of this work demonstrate the importance of MD simulations in complementing experimental information, by revealing atomistic scale information on the conformations and further interactions and dynamics of the ligand shell. While finding the right composition and conformations of the ligand shell is an impossible task with enormous number of possibilities, the use of MD simulations enables investigating several of them. Especially, MD simulations provided structures for higher-scale DFT computations producing spectral data directly comparable to the experiments.

In general, the detailed MD results show that the interactions and the observed conformations are dependent on the number and the positioning of ligands, the protonation state and the pattern of the ligand shell and the solvent conditions. While the actual statistics are not relevant for the results of this work, we point out the ability to harvest more detailed data from the simulations. This also implies the effects of protonation state and pattern along with solvent conditions, potentially important in future simulations.

In summary, the MD simulations contributed in shedding light on the stability of the *m*MBA protected gold nanoclusters in spite of the low ligand count or protection. Thus, we have demonstrated the applicability and generalizability of the force field in work **I** for systems it was not originally parameterized for.

4.3 Self-assembled superstructures of $Au_{102}pMBA_{44}$ gold nanoclusters

The results presented in this section are unpublished and not peer-reviewed.

4.3.1 Motivation

This work was motivated by the work of Nonappa *et al.* [40] reporting the self-assembly of $Au_{102}pMBA_{44}$ gold nanoclusters into hexagonally close-packed, 2D flakes composed of few monolayers of gold nanoclusters, with lateral dimensions in order of hundreds of nanometers. Such 2D sheets in turn bent to form 3-dimensional, one-cluster-thick shell-like structures of ≈ 200 nm in diameter. This self-assembly is controlled with solvent conditions by gradually changing from water to methanol by dialysis.

The *p*MBA ligand of the $Au_{102}pMBA_{44}$ gold nanoclusters allows for the surface patterning by varying protonated and deprotonated ligands, and thus promote certain inhomogeneous hydrogen bonding network, likely responsible for the superstructure packing. However, experimental information on the protonation state, let alone the pattern, of the gold nanoclusters in the superstructures is not available.

Our aim for this work is to explain in atomistic detail the origins of stability of both the 2D and 3D $Au_{102}pMBA_{44}$ gold nanocluster superstructures in terms of interactions and solvent conditions. This information may be utilized in design of new self-assembled nanomaterials.

4.3.2 System descriptions

To investigate the conditions leading to stable superstructures in atomistic detail, we have performed MD simulations in various conditions. While simulation time-scales do not allow superstructure formation from a random configuration of gold nanoclusters as such, simulations were initiated from ordered structures mimicking the experimentally observed superstructures.

To this end two types of superstructures of $Au_{102}pMBA_{44}$ were constructed for the simulations: (1) An infinite 2D-sheet with the packing parameters obtained from experimental data and (2) spherical shell of 100 gold nanoclusters of ≈ 17 nm in diameter (smaller compared to experimental structures of up to ≈ 200 nm in diameter). The structures are presented in figure 11.

(a) (b)

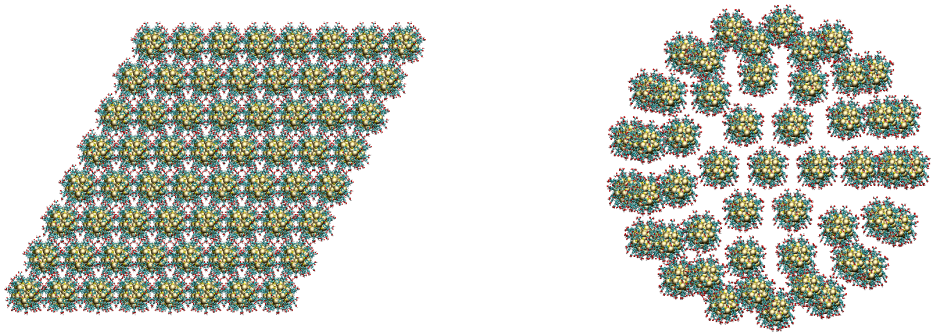


FIGURE 11 Simulated superstructures of Au_{102}*p*MBA_{44}. (a) 2D hexagonally packed sheet of 64 gold nanoclusters in the unit cell; (b) spherical shell of 100 gold nanoclusters.

In terms of investigating the effect of solvent conditions for the superstructure stability, the 2D sheet systems were simulated in (a) pure water; (b) pure methanol and (c) an even mixture of water and methanol. Similar solvent conditions ((d)-(h)) were used for spherical shell systems as well, along with mixtures of (g) water / methanol initially outside / inside the shell and (h) water / methanol initially inside / outside the shell. In all systems NaCl is added to concentration of 0.1 M + neutralizing cations when needed (max. cation concentration 1.0 M in systems with all clusters deprotonated). In solvent conditions (g) and (h) ions are initially divided according to their solubility in water and methanol, i.e., 90% initially in the water. Different solvent conditions are presented in figure 12.

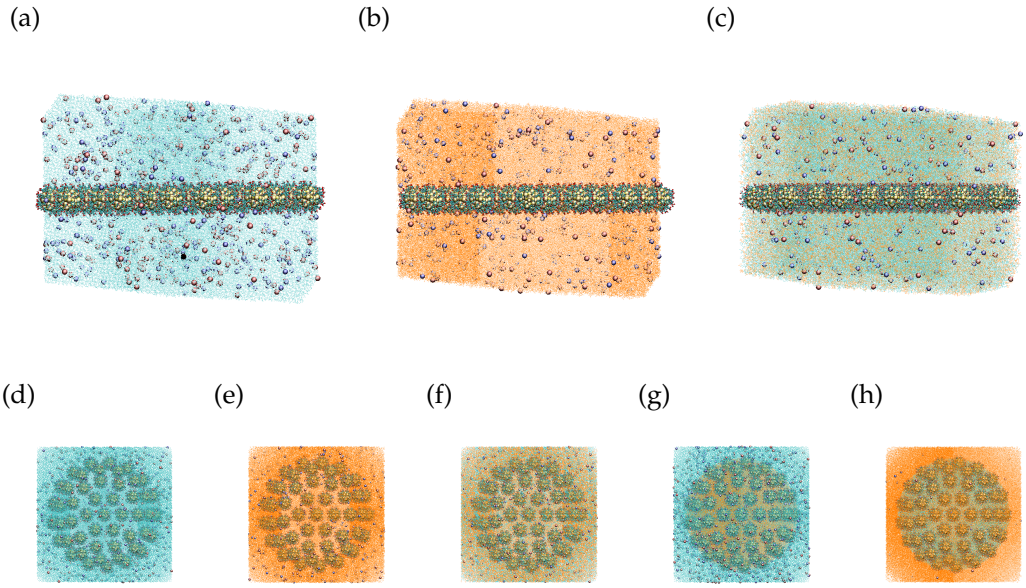


FIGURE 12 Solvent conditions of superstructure simulations, water shown in cyan and methanol in orange colors. Ions, Na and Cl are shown in blue and red, respectively.

The titratable group *p*MBA in the Au_{102}*p*MBA_{44} cluster ligand shell enables numerous possibilities for the cluster charge or the protonation state, depending on the pH. According to recent titration experiment [24] of the Au_{102}*p*MBA_{44} cluster in water at pH=7, the charge state of -22 e has been suggested, i.e., half of the 44 ligands protonated / deprotonated. In the lack of any additional knowledge on the charge state or the pH of the experiments mimicked here, we have performed the simulations in the -22 e state along with 0 e state (fully protonated) and the -44 e charge state (fully deprotonated). The simulated protonation patterns are presented in figure 13.

In addition to the overall charge, the 44 ligand positions allow for patterning of the cluster surface with charges, determined by how the protonated and deprotonated ligands are positioned. Thus in considering the -22 e charge state, we have constructed eight different patterns, figure 13 (c)-(j). Patterns (c)-(f) are constructed to have a belt of neutral or negatively charged ligands in the middle, with top and bottom then correspondingly negatively charged or neutral. (g) and (h) have all neutral or negatively charged ligands on same side of the cluster. In patterns (c)-(h) the divisions are made both vertically and horizontally with respect to the C_2 symmetry axis of the gold nanocluster. Pattern (i) is a pattern suggested by a constant pH MD simulation [24] and pattern (j) is randomly constructed.

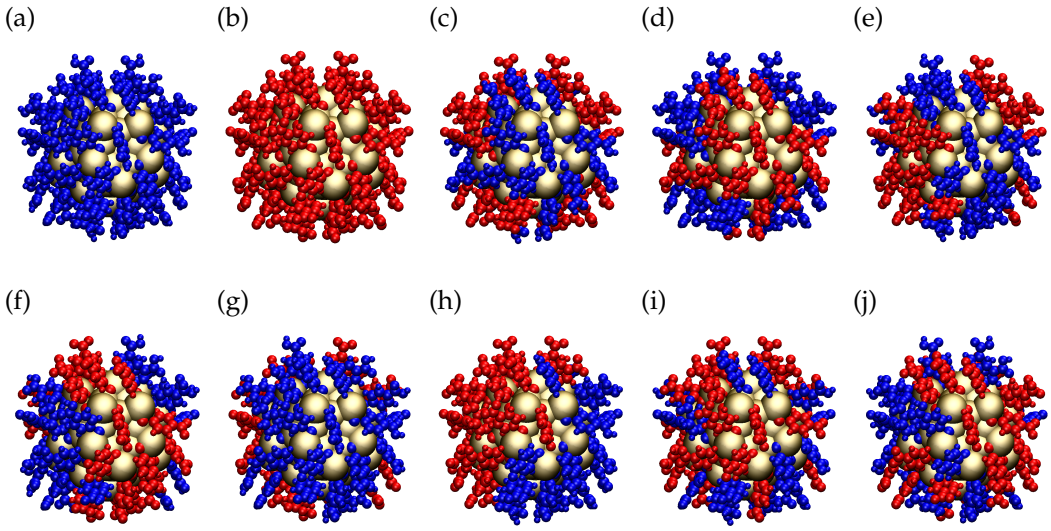


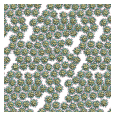
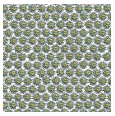
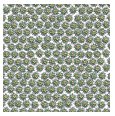
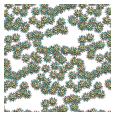
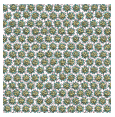
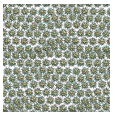
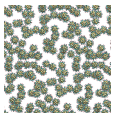
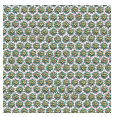
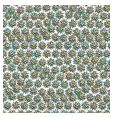
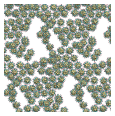
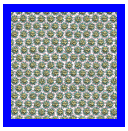
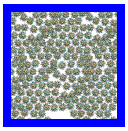
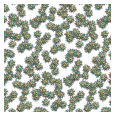
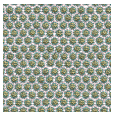
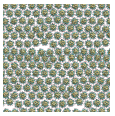
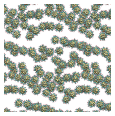
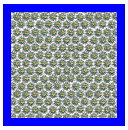
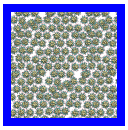
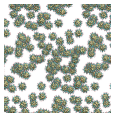
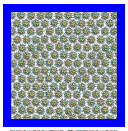
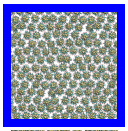
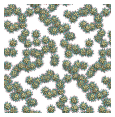
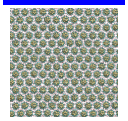
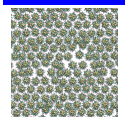
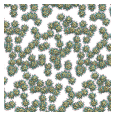
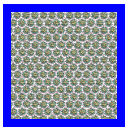
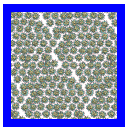
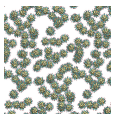
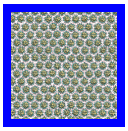
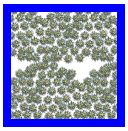
FIGURE 13 Protonation patterns of Au_{102}*p*MBA_{44}. Protonated (neutral) ligands shown in blue, deprotonated (charge -1e per ligand) shown in red. (a) and (b) show fully protonated and deprotonated clusters, charges 0 and -44e, respectively; (c)-(j) have half of the ligands protonated and deprotonated in different patterns, and the charge -22e for each.

Considering the solvent and protonation pattern conditions, MD simulations of 30-100 ns were performed for 30 2D sheet systems and 30-200 ns for 50 spherical shell systems. The specific description of the MD simulations and the analyses are presented in appendix 1. The main results are presented in the following section.

4.3.3 Simulation results - 2D sheets

The end structures of 2D sheets are presented in table 2. The structures from 100 ns simulations are framed in blue.

TABLE 2 End structures after 30 - 100 ns simulation; 100 ns end structures framed in blue, 30 ns end structures unframed. The different protonation states in rows (a)-(j) and the solvent conditions in columns (a)-(c). 2-3 periodic images shown.

Identifier	(a)	(b)	(c)
(a)			
(b)			
(c)			
(d)			
(e)			
(f)			
(g)			
(h)			
(i)			
(j)			

Considering the different solvent conditions (columns), visually the structure is best retained in methanol (b), while in water (a) generally all systems fall apart. Retaining of structure to varying degrees is observed in mixture (c).

In terms of protonation states (rows) ranking by visual means remains ambiguous, while differences are observed. For example in mixed solvent state (c) protonation state (j) clearly distorts more compared to other protonation states. Average distances between neighboring clusters in water and methanol are presented in figure 14.

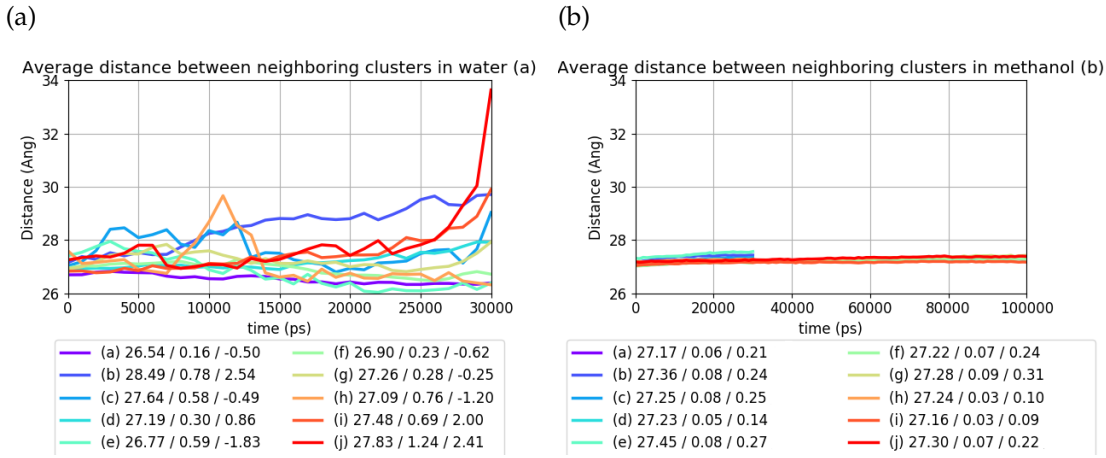


FIGURE 14 Average distance between neighboring cluster pairs (Å) of different protonation states (a)-(j) in water (a) and in methanol (b). Average / standard deviation / drift are reported in the legends.

In agreement with the visual inspection, the neighbor cluster distances in water (decreased stability) show larger deviations and drifts compared to methanol (increased stability). Considering the closeness to the experimental reference value of 27 Å, the differences in averages between protonation states are still minute. For example in methanol largest difference between states (e) 27.45 Å and (i) 27.16 Å is only 0.29 Å.

Since evidently differences in stability of the 2D sheets are observed depending on the solvent condition and the protonation pattern, we take a detailed look into interactions between clusters to account for these differences. To this end, we focus on hydrogen bonds and ionic interactions. Namely, the superstructure formed of fully deprotonated clusters (protonation state (b)), unable to form inter-cluster hydrogen bonds, appears stable in both methanol and in mixture (figure 2, row (b), columns (b), (c)). In addition, the superstructure formed of fully protonated clusters (protonation state (a)) and thus neutral, not forming ionic interactions, appears stable in methanol and mixture (figure 2, row (a), columns (b), (c)).

Average number of hydrogen bonds and chelating cations per cluster are presented in figure 15. The cut-off for hydrogen bond length and angle in the analysis are 0.35 nm and 30°, respectively. Cation is considered as chelating if located within 0.30 nm of two carboxyl acid group carbons simultaneously. Analyses considering also cations close to the ligands (within 0.3 nm) were performed without the chelating condition, which resulted on average each deprotonated group compensated by a close cation for all of the studied systems.

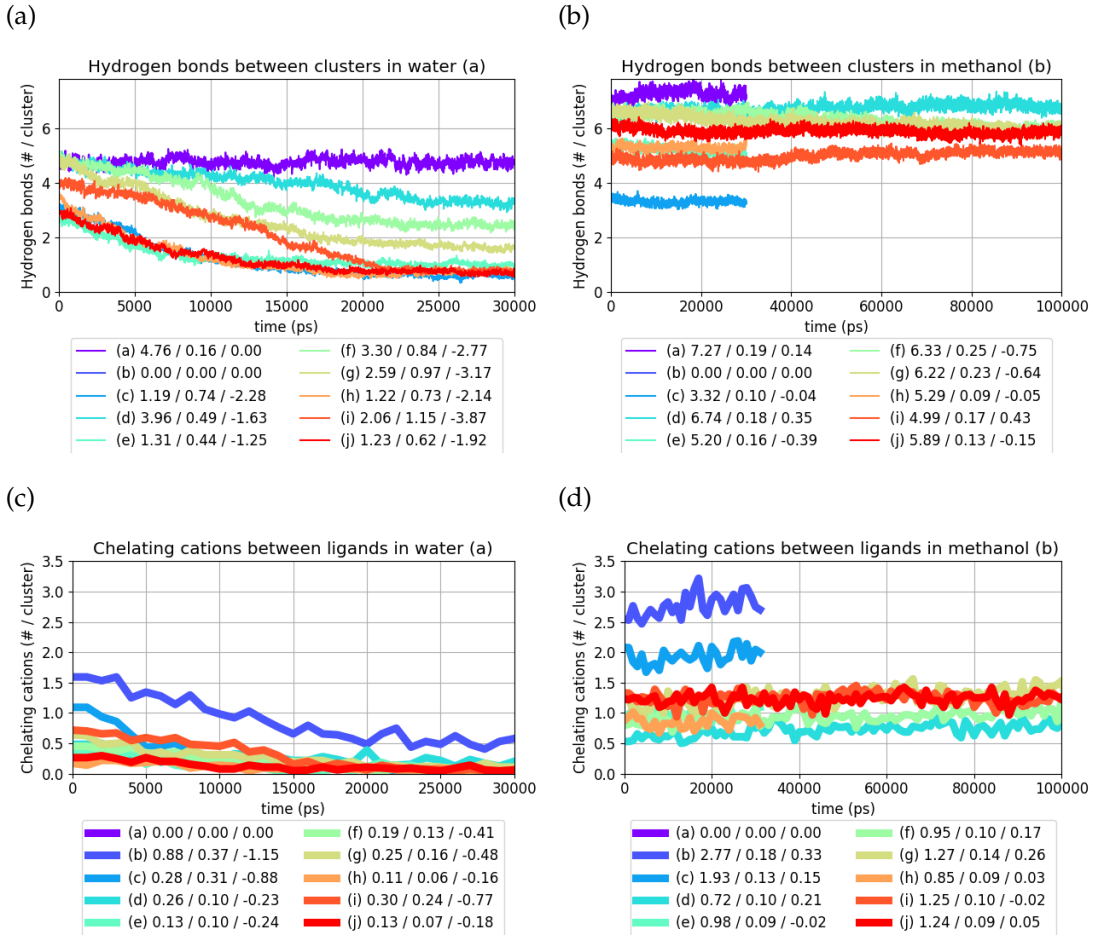


FIGURE 15 Average number of hydrogen bonds between gold clusters (/cluster) of different protonation states (a)-(j) in water (a) and in methanol (b). The average number of chelating cations (/cluster) of different protonation states (a)-(j) in water (c) and in methanol (d). Average / standard deviation / drift are reported in the legends.

From figure 15 it can be seen that on average fewer hydrogen bonds and fewer chelating cations between ligands are observed in water compared to methanol. This is likely due to poorer solubility of nanoclusters in methanol, and the poor solubility of NaCl in methanol. This implies that, since stability in water was deemed worse than in methanol based on visual observation, these interactions reflect the superstructure stability. Furthermore, differences between protonation states appear, so that the smaller number of inter-cluster hydrogen bonds may be in turn compensated by a larger number of chelating cations. This effect could be clearly seen for example for protonation state (c) in methanol with on average smaller number of inter-cluster hydrogen bonds and larger number of chelating cations compared to the other protonation states.

In order to obtain consistent ranking between the different protonation states among different solvents, a simple scoring function was fitted to find weights w_{hbonds} and w_{ions} for number of hydrogen bonds Hb and chelating cations Ic according to $R = w_{\text{hbonds}} * Hb + w_{\text{ions}} * Ic$. R is defined as the goodness or stability score based on the closeness of average neighbor distance to the experimental reference value and percentage of neighbor cluster distances remaining within 0.2 nm of this reference value during the simulation. From the fitted weights ($w_{\text{hbonds}} = 2.15$, $w_{\text{ions}} = 4.91$) the new stability scores are calculated and normalized to give maximum (best) score of 10. The resulting scores are presented in table 3.

TABLE 3 Stability scores of different protonation states (rows (a)-(j)) in different solvents columns (a)-(c). The last row shows the average $\pm$ the standard deviation over different protonation states for each solvent conditions and last column shows the average $\pm$ the standard deviation for each protonation state over different solvent conditions.

Identifier	(a)	(b)	(c)	average
(a)	5.2	8.0	6.5	6.6 ± 1.1
(b)	2.2	6.9	4.3	4.5 ± 1.9
(c)	2.0	8.5	5.5	5.3 ± 2.7
(d)	5.0	9.2	5.4	6.5 ± 1.9
(e)	1.8	8.2	5.9	5.3 ± 2.7
(f)	4.1	9.3	5.8	6.4 ± 2.2
(g)	3.5	10.0	5.9	6.5 ± 2.7
(h)	1.6	7.9	5.3	4.9 ± 2.6
(i)	3.0	8.6	5.6	5.7 ± 2.3
(j)	1.7	9.6	4.7	5.3 ± 3.3
average	3.0 ± 1.3	8.6 ± 0.9	5.5 ± 0.6	5.7 ± 0.7

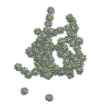
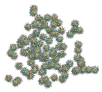
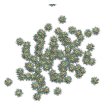
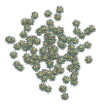
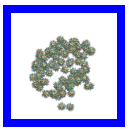
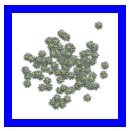
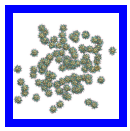
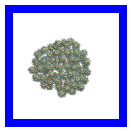
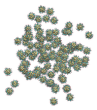
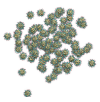
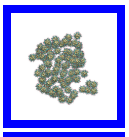
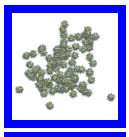
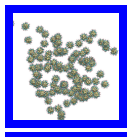
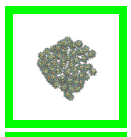
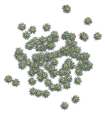
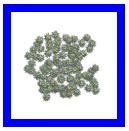
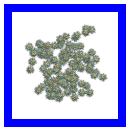
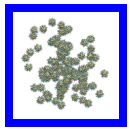
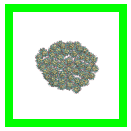
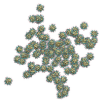
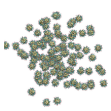
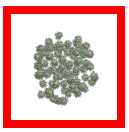
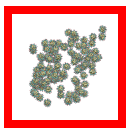
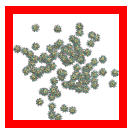
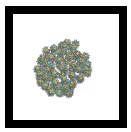
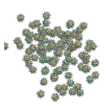
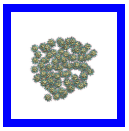
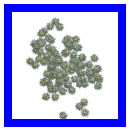
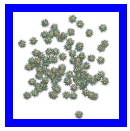
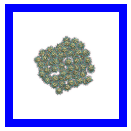
The averages over protonation states for different solvent conditions in table 3 reflect the visually observed (figure 2) stabilities, from best to worst, in methanol (b), mixture (c) and water (a). Considering the averages over solvent states for different protonation patterns, the most stable ones based on this scoring are (a), (d), (f) and (g), and the least stable ones are (b) and (h). However, while the trend in solvent conditions is clear, the differences between protonation states considering the deviations are much smaller.

Based on this scoring we find that the most stabilizing solvent condition, protonation pattern combination is the full methanol and protonation pattern (g). This corresponds to a protonation state of all the neutral ligands on one side and all the deprotonated ligands on the other side of the cluster, as divided horizontally with respect to the C_2 symmetry axis, and is thus amphiphilic or Janus-like in nature. Interestingly, the protonation pattern (h) exhibiting low stability in terms of the scores, has also a protonation pattern with all the neutral ligands on one side and all the deprotonated ligands on the other side of the cluster, but divided vertically with respect to the C_2 symmetry axis.

4.3.4 Simulation results - spherical shells

The end structures of spherical shells are presented in table 4. The structures after >30 ns simulations are framed in different colors.

TABLE 4 End structures after 30 - 200 ns simulation; 50 ns, 100 ns, 150 ns and 200 ns end structures framed in red, blue, black and green, respectively, 30 ns end structures unframed. The different protonation states shown in rows (a)-(j) and the solvent conditions in columns (d)-(h).

Identifier	(d)	(e)	(f)	(g)	(h)
(a)					
(b)					
(c)					
(d)					
(e)					
(f)					
(g)					
(h)					
(i)					
(j)					

Considering the different solvent conditions (columns), similarly to 2D sheets, water (column (d)) appears visually the most destabilizing. To be noted is that, due to smaller size of the spherical shell compared to the experimental one, higher surface tension in the simulated structures could be expected due to the higher curvature. This may affect the stability of the shell structures. The overall most stabilizing condition, independent of the simulation time scale, appears to be the mixture (h) (i.e., water initially inside the shell), though squeezing of the spherical shell is observed. While in mixtures (f) and (g) structure retaining to varying degrees is observed for some systems, the longer simulations (framed structures in figure 4) consistently fall apart in these solvent conditions. Retaining of the shell-like structure is observed in methanol (e) to some extent, with no apparent squeezing.

As for the 2D sheet structures, ranking the protonation states (rows) by mere visual means is ambiguous. To consider structures in more detail, average superstructure gyration radii in water, methanol and mixture (h) are presented in figure 16.

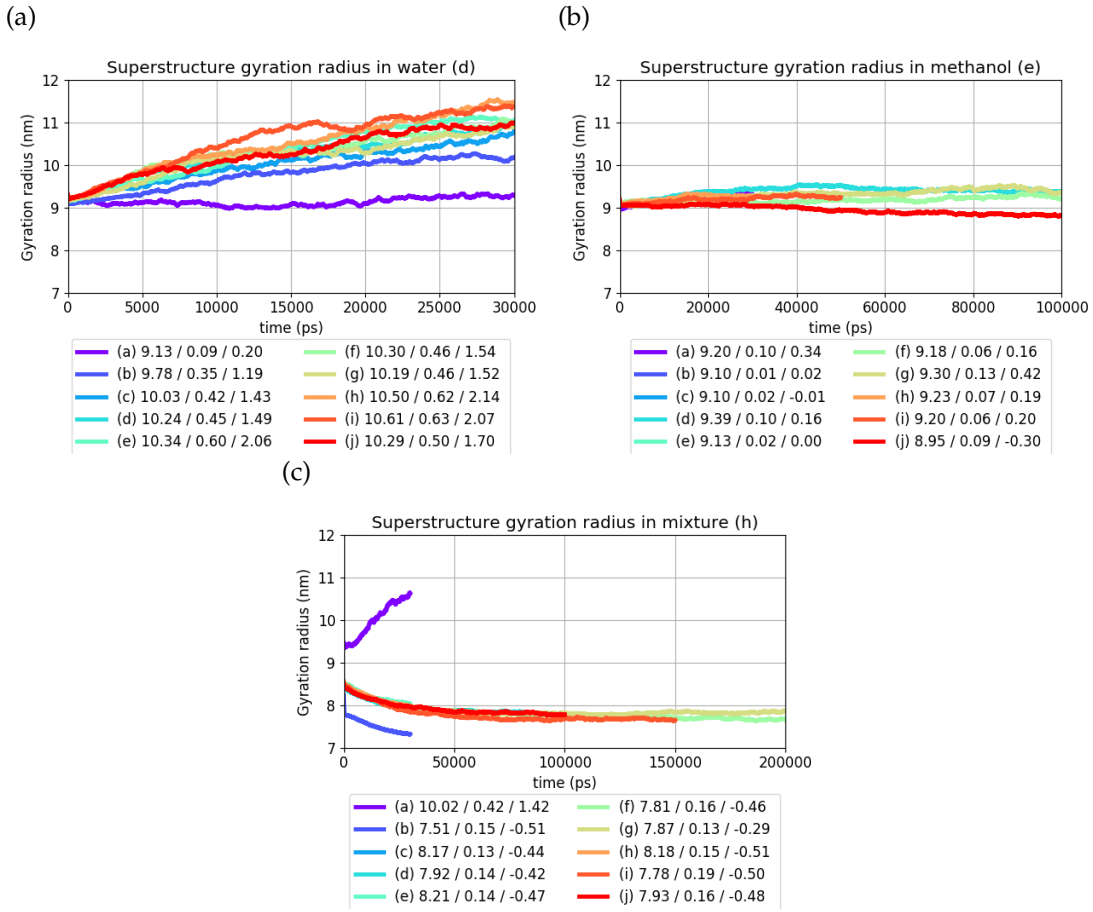


FIGURE 16 Superstructure gyration radius of different protonation states (a)-(j) in solvents (d), (e) and (h), figures (a)-(e). Average / standard deviation / drift are reported in the legends.

In agreement with the visual inspection, from figure 16 it can be seen that the gyration radii in water (decreased stability) are, for the most protonation states, larger in terms of averages, deviations and drifts compared to methanol and mixture (h). In methanol all protonation states show similar values in gyration radii, while in mixture (h) the averages are smaller and the extreme protonation states (a) (visually unstable in figure 4) and (b) show the largest and smallest gyration radii, respectively.

Similarly to 2D sheets, a detailed look into interactions between clusters is given in terms of hydrogen bonds and chelating cations between ligands. Average number of hydrogen bonds per cluster are presented in figure 17 for solvent conditions in water, methanol and mixture (h).

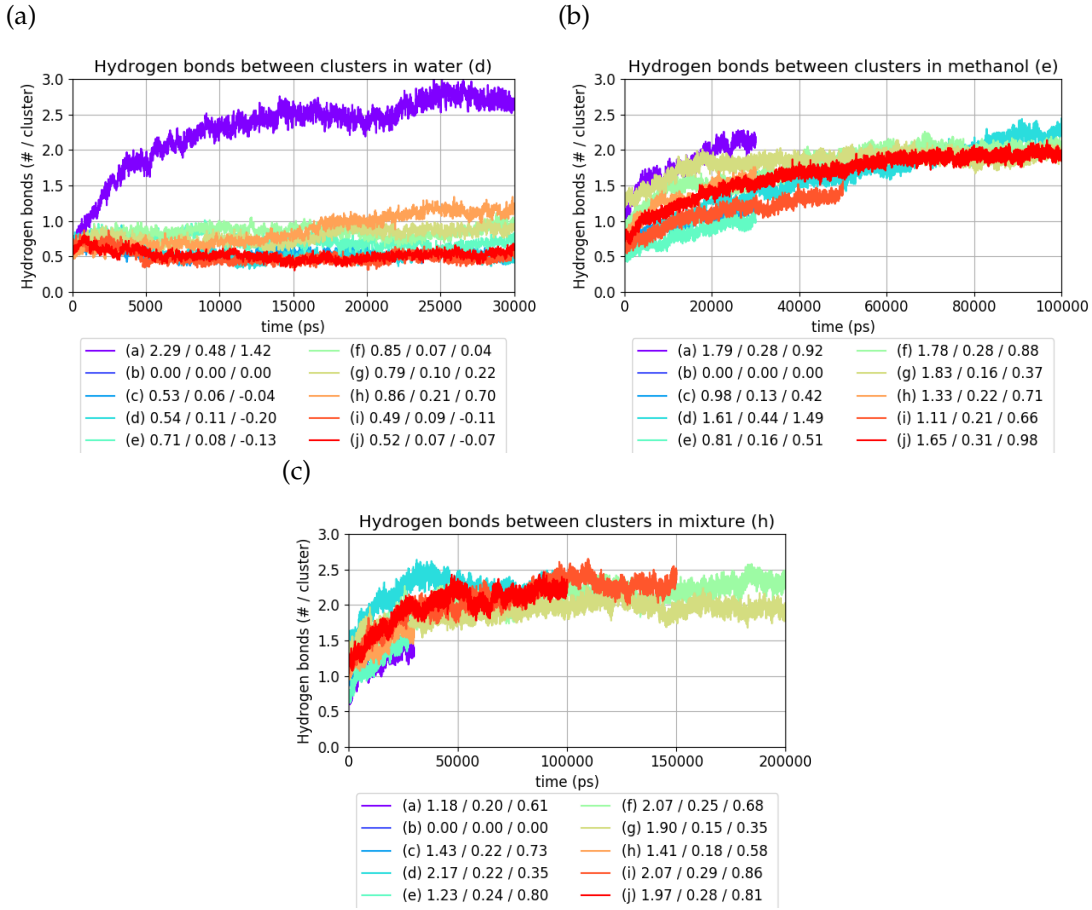


FIGURE 17 Average number of hydrogen bonds between gold clusters (/cluster) of different protonation states (a)-(j) in solvents (d), (e) and (h), figures (a)-(c). Average / standard deviation / drift are reported in the legends.

From figure 17 it can be seen that most often there are fewer hydrogen bonds between clusters in water compared to methanol and mixture (h), implying poorer solubility of nanoclusters in methanol compared to water. On the other hand protonation state (a) in water shows the most inter-cluster hydrogen bonds, while visually the structure is unstable (figure 4)). This implies that hydrogen bonds do not determine the stability of the shell-like superstructures alone. Furthermore, on average more hydrogen bonds are observed in mixture (h) compared to methanol, which reflects the observed squeezing and stability.

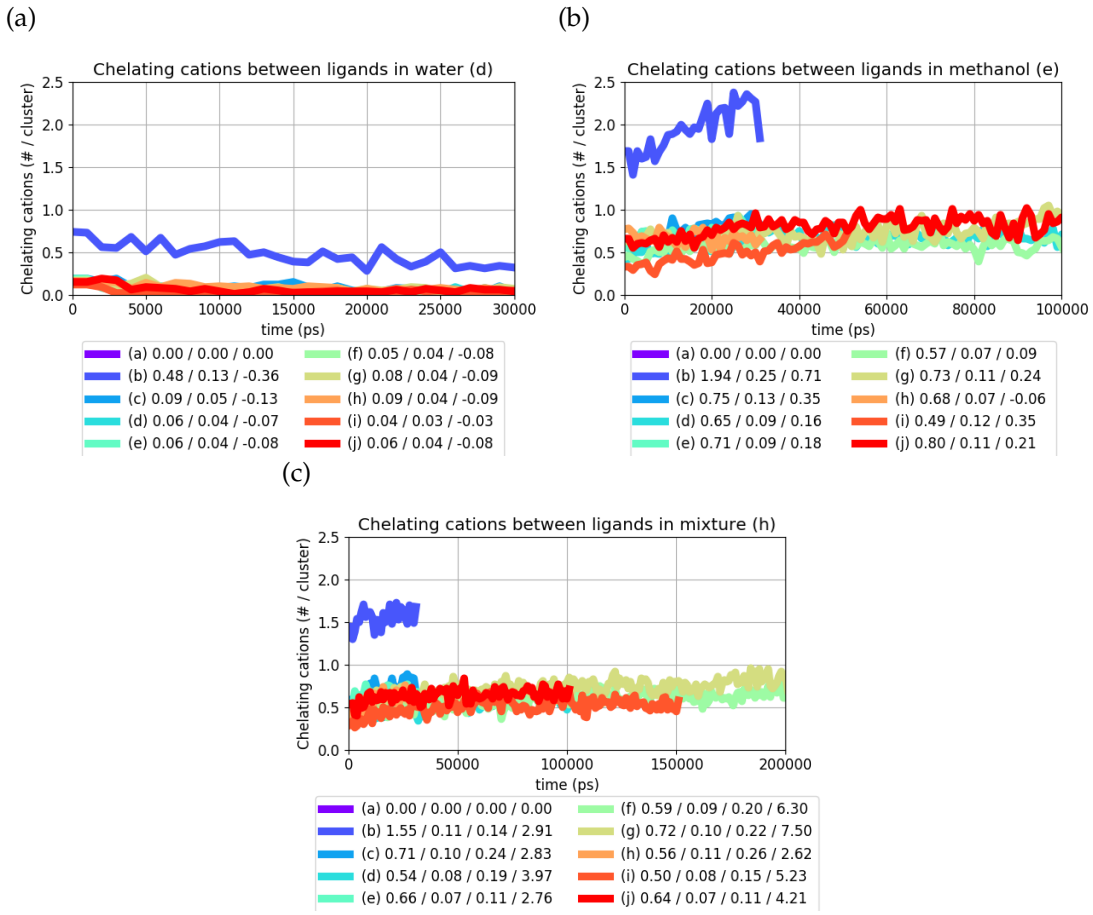


FIGURE 18 Average number of chelating cations (/cluster) of different protonation states (a)-(j) in solvents (d), (e) and (h), figures (a)-(c). Average / standard deviation / drift are reported in the legends.

From figure 18 it can be seen, similarly to case of hydrogen bonds, that there are fewer chelating cations in water compared to methanol and mixture (h), again likely due to the poor solubility of NaCl in methanol. This implies that the chelating cations also play role in the superstructure stability. The protonation state (b) (all deprotonated, charge -44e / cluster) in all solvents has the most chelating cations compared to other protonation states, explained by the high negative charge per cluster. On average, similar number of chelating cations are observed in mixture (h) and methanol. Analyses considering close cations without chelating condition were also performed, and similarly to the 2D sheets on average each deprotonated group was compensated by a close cation for all of the studied systems.

Ranking between the different protonation states was obtained similarly to 2D sheets with a simple scoring function fitted to find weights w_{hbonds} and w_{ions} for number of hydrogen bonds Hb and chelating cations Ic according to $R = w_{\text{hbonds}} * Hb + w_{\text{ions}} * Ic$. R is defined as the goodness or stability score based on the gyration radius (larger gyration radius, i.e., more scattered clusters gives a lower score). From the fitted weights ($w_{\text{hbonds}} = 0.67$, $w_{\text{ions}} = 0.81$) the new stability scores are calculated and normalized to give maximum (best) score of 10. The resulting scores are presented in table 5.

TABLE 5 Stability scores of different protonation states (rows (a)-(j)) in different solvents columns (d)-(h). The last row shows the average $\pm$ the standard deviation over different protonation states for each solvent conditions and last column shows the average $\pm$ the standard deviation for each protonation state over different solvent conditions.

Identifier	(d)	(e)	(f)	(g)	(h)	average
(a)	8.1	6.3	4.2	6.3	4.2	5.8 ± 1.5
(b)	2.1	8.3	5.4	3.4	6.7	5.2 ± 2.2
(c)	2.3	6.7	4.9	3.3	8.1	5.1 ± 2.1
(d)	2.2	8.5	5.3	3.1	10.0	5.8 ± 3.0
(e)	2.8	5.9	5.4	3.6	7.2	5.0 ± 1.6
(f)	3.2	8.8	5.9	3.7	9.9	6.3 ± 2.7
(g)	3.1	9.6	5.6	3.7	9.8	6.4 ± 2.9
(h)	3.4	7.6	4.8	4.0	7.4	5.4 ± 1.7
(i)	1.9	6.0	4.8	2.5	9.5	4.9 ± 2.7
(j)	2.1	9.3	6.1	4.4	9.7	6.3 ± 2.9
average	3.1 ± 1.7	7.7 ± 1.3	5.2 ± 0.5	3.8 ± 1.0	8.3 ± 1.8	5.6 ± 0.6

The averages over protonation states for different solvent conditions in table 5 reflect the visually observed (figure 4) stabilities, with water and mixture (g) showing low scores, mixture (f) an intermediate score and methanol and mixture (h) having high scores. Considering the averages for different protonation patterns, the most stable ones based on the scoring are (f), (g) and (j), and the least stable ones are (e) and (i). As for the 2D sheets, the trend in solvent conditions is clear but the differences between protonation states considering the deviations are much smaller.

The most stabilizing solvent condition, protonation pattern combination is found for mixture (h) and the protonation state (d). This corresponds to a protonation state of all the neutral ligands forming a horizontal belt in the middle of the cluster (with respect to C_2 symmetry axis) and the deprotonated ligands divided on top and bottom of the cluster. The lowest stability based on this scoring is found for protonation pattern (i), which is the pattern predicted by a constant pH simulation.

4.3.5 Conclusions

We have performed molecular dynamics simulations of two types of different experimentally determined superstructures of $Au_{102}p$MBA$_{44}$ gold nanoclusters, 2D sheets and spherical shells, in various solvent and protonation conditions to systematically investigate their stability.

Visual investigation reveals clear differences in stabilities of superstructures in different solvent conditions, and more subtle differences between protonation states and patterns. In agreement with the experimental results of Nonappa *et al.* [40], we find none of the superstructures stable in pure water, while the 2D sheets show excellent retaining of the hexagonally packed superstructure in methanol. While also spherical shells retain some resemblance to the initial shell-like structure in pure methanol, systems in mixture where water lies initially inside the shells appear, while squeezed, the best retained of all the solvent conditions. While this initial dividing of water and methanol inside and outside the spherical shell is not completely retained during the simulation, i.e., mixing occurs, the breaking of superstructure with reverted division, i.e., water initially outside the shell suggests such shell structures may enclose water.

The evident interactions for $Au_{102}p$MBA$_{44}$ systems consist of hydrogen bonds, yet we also find chelating ionic interactions to play role in the superstructure stability. Thus, one may envision the most efficient ligand layer patterning for the superstructure stability being a balance between hydrogen bonds and ionic interactions. The effect of the solvent then comes in as the preference for forming solvent-ligand hydrogen bonds over ligand-ligand hydrogen bonds in water and vice versa in methanol. Similarly, due to poor solubility of NaCl in methanol ions prefer to lay closer to gold nanoclusters enabling increased number of chelating interactions.

Adopting such simplistic view of the two contributing interactions, we have scored the different protonation states, for which determining the stability merely visually is not feasible. To this end, we find this scoring captures the evident stability differences of the solvent conditions, i.e., least score for water (unstable systems) and the best score for methanol (for 2D sheets) or mixture (h) (spherical shells). Considering the protonation patterns, for both superstructure types we find states (f) and (g) showing on average high scores. State (f) corresponds to a protonation state of all the negatively ligands forming a vertical belt in the middle of the cluster (with respect to C_2 symmetry axis) and the neutral ligands divided on top and bottom of the cluster. State (g) corresponds to a protonation state of all the neutral ligands on one side and all the deprotonated ligands on the other side of the cluster, as divided horizontally with respect to the C_2 symmetry axis.

Based on the scoring results, it appears certain amphiphilic nature of the ligand shell improves the stability of the superstructure, whereas more patched ligand shells result in lower stability. In addition, it seems to matter how the ligands are divided with respect to the symmetry axis. The differences in averages between protonation states are small, and longer simulations along with non-homogeneous initial cluster orientations would be needed to fully conclude the stability of certain protonation states over the others. Further insight into the stability of the superstructures could be obtained from computing free energies of different configurations.

In summary, in spite of the limitations described above, we are able suggest trends contributing to the stability of $Au_{102}p$MBA$_{44}$ nanocluster superstructures, in terms of both the solvent conditions and the protonation patterns. In both 2D and 3D cases the clusters with amphiphilic surfaces seem to increase the stability. 2D sheets appear most stable in pure methanol, while the 3D shell structure is further stabilized by enclosing water. This information could be useful in the future design of metal-cluster self-assemblies and superstructures thereof.

4.4 $Au_{102}pMBA_{44}$ gold nanoclusters at the bio-interface: Binding to an enterovirus (paper III)

Tuning of the ligand shell of a gold nanoclusters enables making them bio-compatible or water soluble, such as the $Au_{102}pMBA_{44}$ cluster. Bio-compatibility of these clusters can be utilized in applications including imaging and sensing. The work presented in this section was motivated by the use of $Au_{102}pMBA_{44}$ nanoclusters as contrast agents in virus imaging with TEM. [41,42] Experimentally, the gold nanoclusters are known to connect to viruses covalently through sulfur linkages between a functionalized gold nanoclusters and exposed cysteines. The non-covalent connection occurs via replacement of a fatty acid molecule inside a hydrophobic pocket on the virus surface, by a drug-like molecule, Kirtan1, linked to the gold nanocluster. The aim of this work was to study the interactions and dynamics between different pocket factors, presented in figure 19, and viruses at atomistic scale.

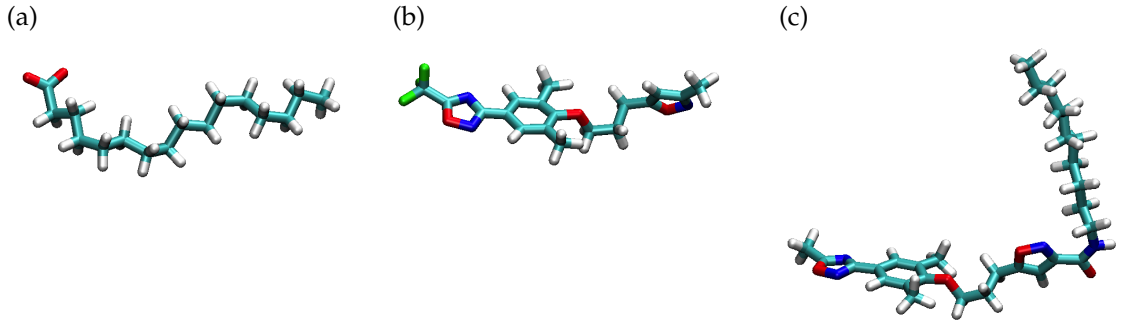


FIGURE 19 Pocket factors of echovirus 1 (EV1). (a) The natural pocket factor palmitic acid; (b) Pleconaril - a known drug molecule against EV1 and (c) Kirtan1 with a hydrophobic tail connectable to the gold nanocluster. Atom colors: C=cyan, O=red, H=white, N=blue, F=green. Adapted with permission from Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**. Copyright 2017 American Chemical Society.

In order to compare between different pocket factors, we estimated their binding affinities, i.e., binding free energies to the hydrophobic pockets of echovirus 1. Due to the complexity of the system, along with differences in the structures of the investigated pocket factors, the equilibrium free energy methods and alchemical methods were out of question. Therefore we resorted to non-equilibrium free energy methods, where steered MD simulations were used for simultaneously pulling all 60 pocket factors out of the pockets and recording the force profiles of the processes. Works performed are then obtained through integration of the force profiles. These works are averaged to obtain the free energy estimate from Jarzynski's equality as described in section 3. While the virus is constantly moving, pulling could only be performed in the forward direction (out from the pocket), thus Crooks theorem could not be utilized here.

Echovirus 1 is an icosahedral virus that contains 60 symmetrically positioned hydrophobic pockets. With each containing a pocket factor, each pulling simulation generates 60 work values for the exponential averaging. We performed 25 pulling simulations of 4.5 ns with pulling velocity of 1 nm/ns, starting from different snapshots of un-steered virus simulations, generating in total of 1500 work values, for each of the five pocket factor systems. An example pulling procedure for palmitic acid is presented in figure 20.

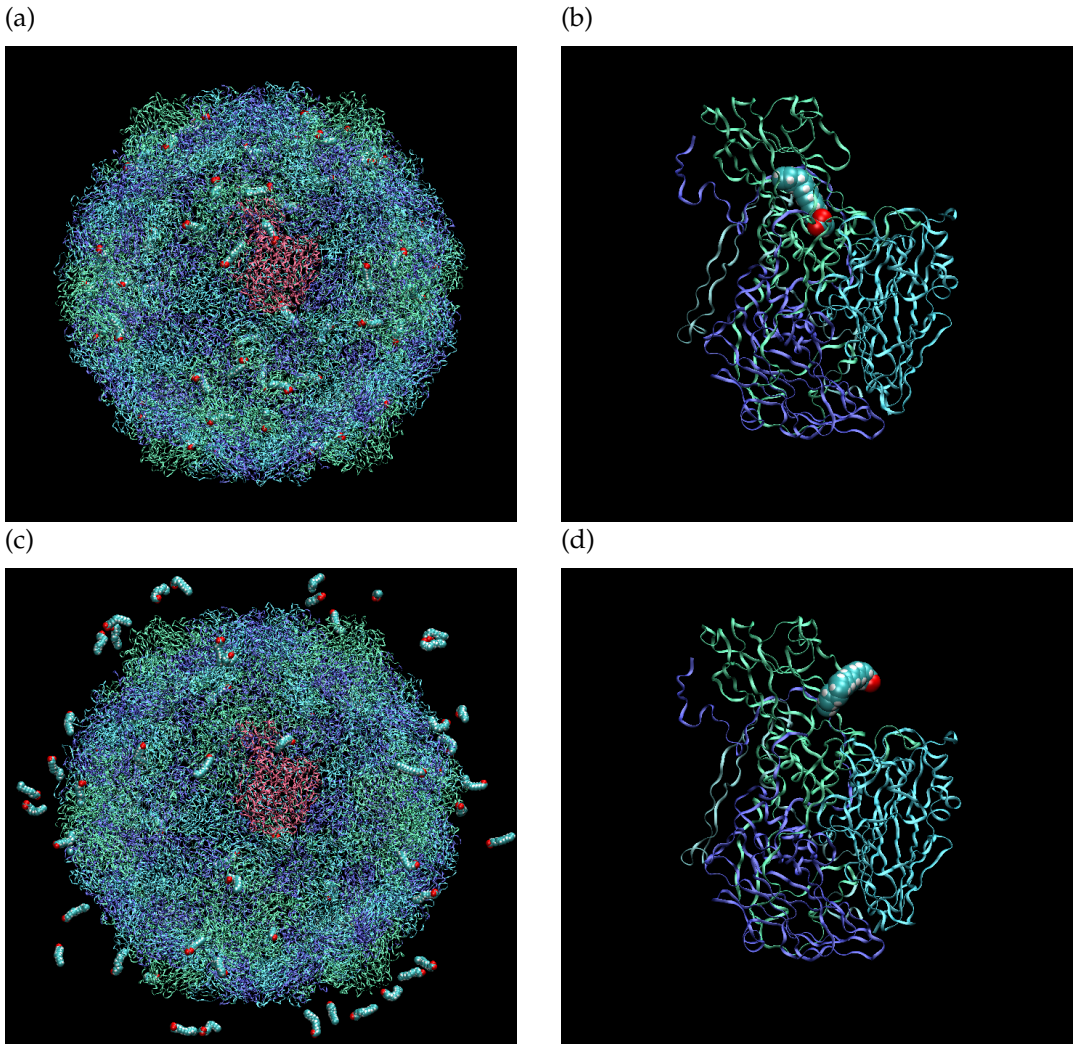


FIGURE 20 Pulling procedure visualized. (a) shows the full virus capsid with 60 hydrophobic pockets occupied a palmitic acid molecule. (b) shows the close-up of one protomer unit shown in red in (a). (c) and (d) show the structures at the end of one pulling simulation. Virus protein shown in ribbon and pocket factors as VDW spheres. Reprinted with permission from Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**. Copyright 2017 American Chemical Society.

Similar pulling end states, i.e., externalized pocket factors, from one pulling simulations and full virus capsid for the remaining pocket factors are presented in figure 21.

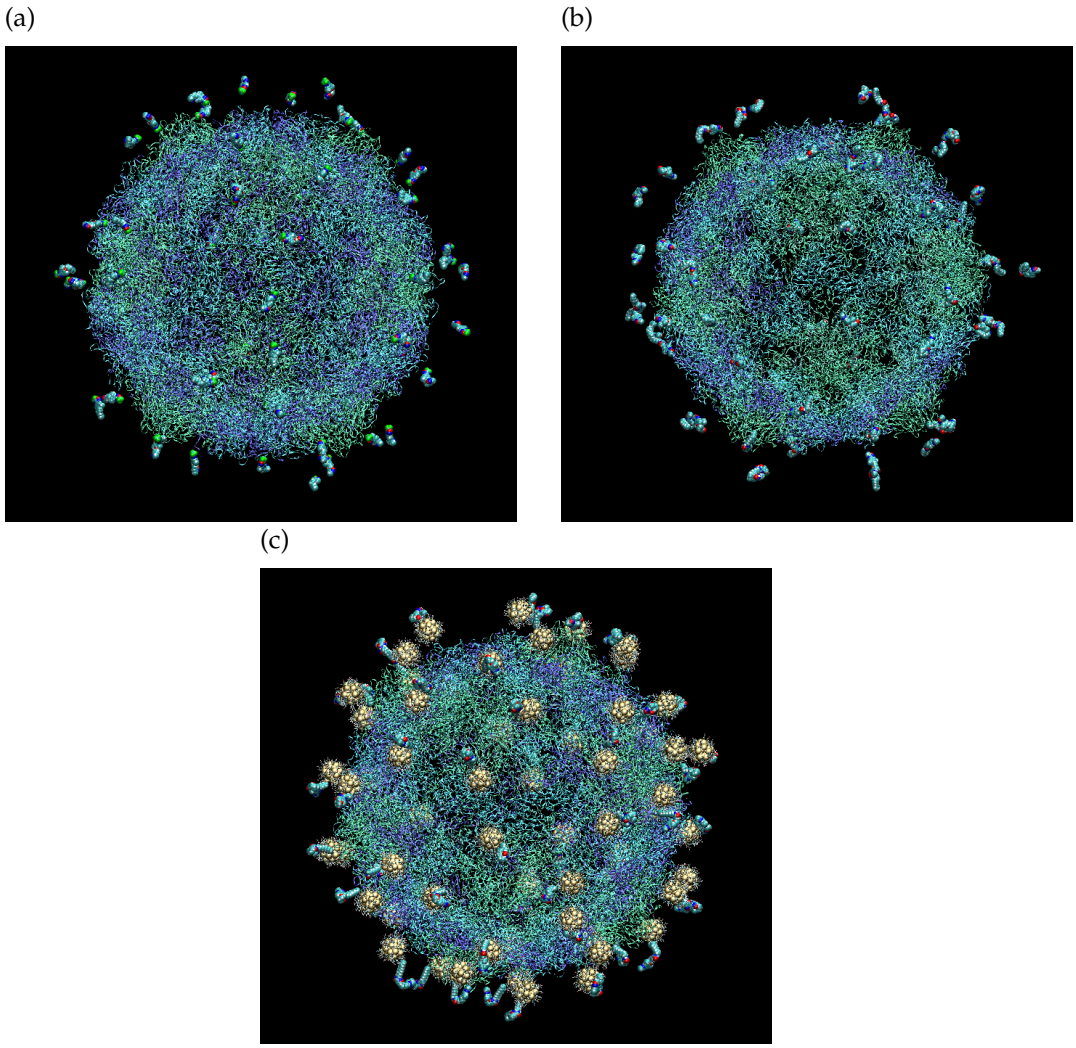


FIGURE 21 Pulling end states for systems with pocket factor (a) Pleconaril, (b) Kirtan1, (c) Kirtan1 + half deprotonated gold nanocluster. Reprinted with permission from Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**. Copyright 2017 American Chemical Society.

While no atomistic scale structural changes in proximity of the hydrophobic pocket were investigated in detail, conformations of the pocket factors along with simple RMSD check of the hydrophobic pocket proximate proteins did not reveal any large scale changes induced by pulling.

The resulting pulling work distributions, shown in figure 22, have significant standard deviation and skew, thus instead of simple cumulant estimators of Jarzynski's equality, we resorted to the scheme based on block averaging and extrapolation, proposed by Zuckerman *et al* [68, 69] for non-gaussian work data to estimate the binding free energies.

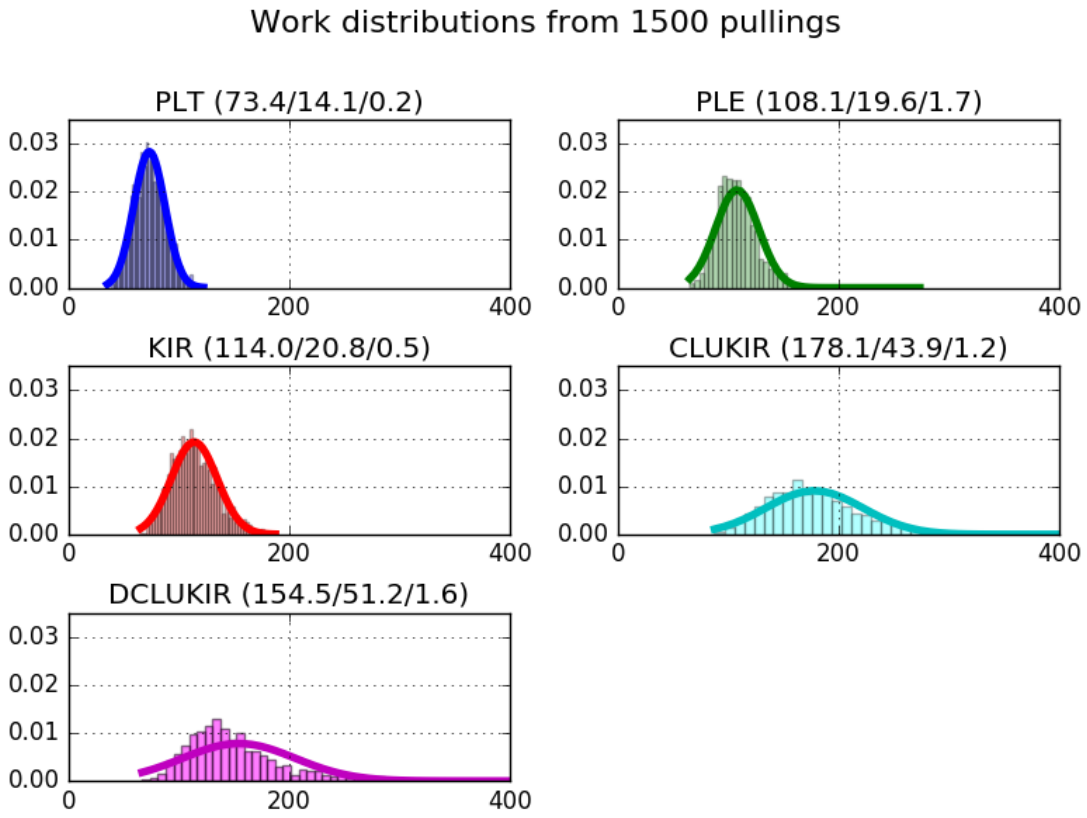


FIGURE 22 Work distributions of 25 pulling simulations or 1500 pullings for pocket factor systems: PLT=palmitic acid, PLE=Pleconaril, KIR=Kirtan1, CLUKIR=Kirtan1+half deprotonated gold nanocluster, DCLUKIR=Kirtan1+fully deprotonated gold nanocluster. Values in parentheses () show mean / standard deviation / skew in kJ/mol. Reprinted with permission from Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**. Copyright 2017 American Chemical Society.

The final estimated binding affinities, presented in table 6, are clearly overestimated, by comparison to one the highest binding affinities between biotin and avidin (≈ -85.4 kJ/mol) appearing in nature. That is, replacing the natural pocket factor with a binding affinity as high as reported here, would be an extremely rare event. On the other hand, this pocket factor is a known drug target and thus replaceable by other molecules.

TABLE 6 Binding affinity, ΔG, estimates for different pocket factors. Adapted with permission from Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**. Copyright 2017 American Chemical Society.

pocket factor	$\Delta G\pm$ error (kJ/mol)
palmitic acid	-168.1 $\pm$ 12
Pleconaril	-299.9 $\pm$ 18
Kirtan1	-295.3 $\pm$ 13
Kirtan1 + half deprotonated gold nanocluster	-392.3 $\pm$ 14
Kirtan1 + fully deprotonated gold nanocluster	-293.6 $\pm$ 19

We investigated several factors potentially causing the overestimation, of which the pulling velocity had the clearest and the most significant effect: Decreasing the pulling velocity decreases the binding affinity estimate. Nevertheless, the trend remained the same among different pulling velocities, thus the qualitative result is unaffected by the compromise made by choosing a pulling velocity fast enough to enable efficient statistics.

Important consideration in applying Jarzynski's equality is the sufficient sampling of work values or converging the exponential average as discussed in chapter 3. In the frame of the block averaging / extrapolation scheme utilized here, we observed the convergence of the results between using 1000 work values and the full data set of 1500 work values. Still, the relatively small data set may affect the overestimation of the binding affinities.

The complexity of the system opens possibilities for example to cooperative effects (pulling all factors at once may not be feasible), dependency on the starting structure (bias due to limited number of initial configurations), and dependency on the pulling distance (i.e., determining the point of detaching). While in our investigations these factors do not seem to matter in terms of the qualitative results, due to relatively small statistics the effects of them cannot be fully excluded.

Qualitatively the computed results are nevertheless reasonable and in line with the experiments, in the sense that: (1) The natural pocket factor has the lowest estimated affinity, and can therefore be replaced by those exhibiting larger estimated binding affinities, such as the known drug-molecule and its derivatives. (2) The two molecules having similar structures inside the pocket, i.e., the drug molecule and its derivative, have similar estimated binding affinities.

Based on the estimates, the binding affinity of the functionalized gold nanocluster depends on the protonation state of the gold nanocluster. In the fully deprotonated state, the affinity is close to that of the drug-molecule and its derivative without the gold cluster. In any case, the affinity is not decreased by the connection of the gold nanocluster. In the case of half protonated / half deprotonated gold nanocluster the affinity increases significantly, likely due to interactions between the ligand shell and the virus surface proteins and / or the poorer solubility in water compared to the fully deprotonated cluster.

In addition to binding affinities, simulations reveal in more detail the factors that contribute to these binding affinities, i.e., interactions that break during the detaching. By considering a very simple additive model of hydrogen bonds, hydrophobic and aromatic interactions as contributing to the binding affinity, it is evident that all these are important.

To summarize, the simulations confirm, in line with the experiments, that the non-covalent method of linking gold nanoclusters to virus surface at specific locations is possible. Furthermore, similar simulations could be utilized in estimating qualitative binding affinities of different types of new molecules to the hydrophobic pocket. The design of new molecules for similar applications could be at first stages guided by considering the interactions revealed by the atomistic scale investigations provided by simulations.

5 SUMMARY AND OUTLOOK

Force field parameterization for performing MD simulations with gold nanoclusters was developed and applied in numerous simulations, by us and other groups, in investigating the functionalities of gold nanoclusters for different purposes.

Considering the works presented in this thesis, we have demonstrated the applicability and usefulness of MD simulations, alongside experiments and DFT computations, in investigating the ligand shell functionality in cases of (1) interactions within one cluster (work **II**), (2) interactions between cluster and solvent and between clusters themselves (self-assembly) and (3) interactions between clusters and virus protein (work **III**). Qualitative agreement with available experimental data confirms, in cases of these works, the sufficiency of accurate description of interactions with ligand shell and the environment, while the gold is inert in terms of electrostatics. This description allows for performing computationally relatively cheap simulations, with a force field that is easily transferable to different kind of gold nanoclusters.

While the structure of the ligand shell (or the positions gold atoms) may not always be resolved experimentally, in future different types of AI methods [125], together with MD simulations may provide bonding or interaction motifs for structure predictions.

Implications of importance of the solvent conditions, especially ions, could be further investigated by computations and experiments. This could reveal more details about the self-assemblies of gold nanoclusters [40], and further their utilization as reversibly forming / deforming objects for, e.g., drug carrier applications. While simulation times for full virus simulations are still short, disabling witnessing any large conformational changes, advances in virus imaging techniques may in future provide new structures for simulations to more accurately investigate, e.g., the implication of gold nanoclusters to their structure and function.

Clear indications of the importance of protonation state and pattern of the gold nanocluster ligand shell were concluded in all the works presented here. While testing all possible states and patterns in simulations is impossible, in the future simulating at constant pH [127], i.e., allowing dynamic protonation / deprotonation during the simulation, could be important and reveal interesting details on the protonation dynamics of these clusters under different conditions. Inherently, the importance of the protonation pattern relates to the effects of charges in the system. Therefore, for certain applications inclusion of both gold partial charges and the charge polarizability would be essential.

Recently, fusing of stable gold nanoclusters into another stable cluster has been observed [128]. Describing such dynamics with simulations could reveal the mechanisms of this, yet treating of the system at QM/MM [85,101] level or at least with reactive force fields [104] would then be required.

While parameterizing organic ligands is relatively straightforward, parameters for heterometallic along with clusters fully composed of other metal than gold, would likely be of interest in future simulations.

REFERENCES

[1] T. Tsukuda and H. Häkkinen, editors. *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*. Elsevier, **2015**.

[2] H. Häkkinen. The gold-sulfur interface at the nanoscale. *Nat Chem.*, 4: 443–55, **2012**.

[3] T. Tsukuda and H. Häkkinen. Chapter 1 - Introduction. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 1 – 7. Elsevier, **2015**.

[4] S. Takano and T. Tsukuda. Chapter 2 - Controlled Synthesis: Size Control. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 9 – 38. Elsevier, **2015**.

[5] Y. Wang, H. Yang, and N. Zheng. Chapter 4 - Structural Engineering of Heterometallic Nanoclusters. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 73 – 102. Elsevier, **2015**.

[6] N. Goswami, J. Li, and J. Xie. Chapter 11 - Functionalization and Application. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 297 – 345. Elsevier, **2015**.

[7] Y. Negishi, Y. Niihori, and W. Kurashige. Chapter 3 - Controlled Synthesis: Composition and Interface Control. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 39 – 71. Elsevier, **2015**.

[8] K. Sokołowska, S. Malola, M. Lahtinen, V. Saarnio, P. Permi, K. Koskinen, M. Jalasvuori, H. Häkkinen, L. Lehtovaara, and T. Lahtinen. Towards Controlled Synthesis of Water-Soluble Gold Nanoclusters: Synthesis and Analysis. *J. Phys. Chem. C*, 123(4): 2602–2612, **2019**.

[9] H. Häkkinen, M. Walter, and H. Grönbeck. Divide and Protect: Capping Gold Nanoclusters with Molecular Gold-Thiolate Rings. *J. Phys. Chem. B*, 110(20): 9927–9931, **2006**.

[10] Q. Tang and D. Jiang. Chapter 7 - Structure Prediction by Density Functional Theory Calculations. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 161 – 187. Elsevier, **2015**.

[11] M. Azubel and R. D. Kornberg. Synthesis of Water-Soluble, Thiolate-Protected Gold Nanoparticles Uniform in Size. *Nano Lett.*, 16(5): 3348–3351, **2016**.

[12] M. Azubel, A. L. Koh, K. Koyasu, T. Tsukuda, and R. D. Kornberg. Structure Determination of a Water-Soluble 144-Gold Atom Particle at Atomic Resolution by Aberration-Corrected Electron Microscopy. *ACS Nano*, 11(12): 11866–11871, **2017**.

[13] P. D. Jadzinsky, G. Calero, C. J. Ackerson, D. A. Bushnell, and R. D. Kornberg. Structure of a Thiol Monolayer-Protected Gold Nanoparticle at 1.1 Å Resolution. *Science*, 318(5849): 430–433, **2007**.

[14] D. Jiang, M. L. Tiago, W. Luo, and S. Dai. The "Staple" Motif: A Key to Stability of Thiolate-Protected Gold Nanoclusters. *J. Am. Chem. Soc.*, 130(9): 2777–2779, **2008**.

[15] Y. Negishi, K. Nobusada, and T. Tsukuda. Glutathione-Protected Gold Clusters Revisited: Bridging the Gap between Gold(I)-Thiolate Complexes and Thiolate-Protected Gold Nanocrystals. *J. Am. Chem. Soc.*, 127(14): 5261–5270, **2005**.

[16] M. W. Heaven, A. Dass, P. S. White, K. M. Holt, and R. W. Murray. Crystal Structure of the Gold Nanoparticle [N(C8H17)4][Au25(SCH2CH2Ph)18]. *J. Am. Chem. Soc.*, 130(12): 3754–3755, **2008**.

[17] M. Zhu, C. M. Aikens, F. J. Hollander, G. C. Schatz, and R. Jin. Correlating the Crystal Structure of A Thiol-Protected Au25 Cluster and Optical Properties. *J. Am. Chem. Soc.*, 130(18): 5883–5885, **2008**.

[18] M. Zhu, W. T. Eckenhoff, T. Pintauer, and R. Jin. Conversion of Anionic [Au25(SCH2CH2Ph)18]- Cluster to Charge Neutral Cluster via Air Oxidation. *J. Phys. Chem. C*, 112(37): 14221–14224, **2008**.

[19] N. K. Chaki, Y. Negishi, H. Tsunoyama, Y. Shichibu, and T. Tsukuda. Ubiquitous 8 and 29 kDa Gold:Alkanethiolate Cluster Compounds: Mass-Spectrometric Determination of Molecular Formulas and Structural Implications. *J. Am. Chem. Soc.*, 130(27): 8608–8610, **2008**.

[20] H. Qian, W. T. Eckenhoff, Y. Zhu, T. Pintauer, and R. Jin. Total Structure Determination of Thiolate-Protected Au38 Nanoparticles. *J. Am. Chem. Soc.*, 132(24): 8280–8281, **2010**.

[21] H. Qian and R. Jin. Controlling Nanoparticles with Atomic Precision: The Case of Au144(SCH2CH2Ph)60. *Nano Lett.*, 9(12): 4083–4087, **2009**.

[22] H. Häkkinen. Chapter 8 - Electronic Structure: Shell Structure and the Superatom Concept. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 189 – 222. Elsevier, **2015**.

[23] M. Walter, J. Akola, O. Lopez-Acevedo, P. D. Jadzinsky, G. Calero, C. J. Ackerson, R. L. Whetten, H. Grönbeck, and H. Häkkinen. A unified view of ligand-protected gold clusters as superatom complexes. *PNAS*, 105(27): 9157–9162, **2008**.

[24] J. Koivisto, X. Chen, S. Donnini, T. Lahtinen, H. Häkkinen, G. Groenhof, and M. Pettersson. Acid–Base Properties and Surface Charge Distribution of the Water-Soluble Au102(pMBA)44 Nanocluster. *J. Phys. Chem. C*, 120(18): 10041–10050, **2016**.

[25] G. Rossi and L. Monticelli. Gold nanoparticles in model biological membranes: A computational perspective. *Biochim Biophys Acta Biomembr.*, 1858(10): 2380 – 2389, **2016**. Biosimulations of lipid membranes coupled to experiments.

[26] G. Rossi and L. Monticelli. Simulating the interaction of lipid membranes with polymer and ligand-coated nanoparticles. *Adv. Phys.: X*, 1(2): 276–296, **2016**.

[27] E. Heikkilä, H. Martinez-Seara, A. A. Gurtovenko, I. Vattulainen, and J. Akola. Atomistic simulations of anionic Au144(SR)60 nanoparticles interacting with asymmetric model lipid membranes. *Biochim Biophys Acta Biomembr.*, 1838(11): 2852 – 2860, **2014**.

[28] X. Quan, C. Peng, D. Zhao, L. Li, J. Fan, and J. Zhou. Molecular Understanding of the Penetration of Functionalized Gold Nanoparticles into Asymmetric Membranes. *Langmuir*, 33(1): 361–371, **2017**.

[29] P. Gkeka, P. Angelikopoulos, L. Sarkisov, and Z. Cournia. Membrane Partitioning of Anionic, Ligand-Coated Nanoparticles Is Accompanied by Ligand Snorkeling, Local Disordering, and Cholesterol Depletion. *PLOS Comput. Biol.*, 10: 1–10, **2014**.

[30] L. Becucci, R. Guidelli, F. Polo, and F. Maran. Interaction of Mixed-Ligand Monolayer-Protected Au144 Clusters with Biomimetic Membranes as a Function of the Transmembrane Potential. *Langmuir*, 30(27): 8141–8151, **2014**.

[31] F. Simonelli, D. Bochicchio, R. Ferrando, and G. Rossi. Monolayer-Protected Anionic Au Nanoparticles Walk into Lipid Membranes Step by Step. *J. Phys. Chem. Lett.*, 6(16): 3175–3179, **2015**.

[32] E. Heikkilä, H. Martinez-Seara, A. A. Gurtovenko, M. Javanainen, H. Häkkinen, I. Vattulainen, and J. Akola. Cationic Au Nanoparticle Binding with Plasma Membrane-like Lipid Bilayers: Potential Mechanism for Spontaneous Permeation to Cells Revealed by Atomistic Simulations. *J. Phys. Chem. C*, 118(20): 11131–11141, **2014**.

[33] R. Gupta and B. Rai. Effect of Size and Surface Charge of Gold Nanoparticles on their Skin Permeability: A Molecular Dynamics Study. *Sci. Rep.*, 7: 45292, **2017**.

[34] K. Salorinne, T. Lahtinen, J. Koivisto, E. Kalenius, M. Nissinen, M. Pettersson, and H. Häkkinen. Nondestructive Size Determination of Thiol-Stabilized Gold Nanoclusters in Solution by Diffusion Ordered NMR Spectroscopy. *Anal. Chem.*, 85(7): 3489–3492, **2013**.

[35] K. Salorinne, T. Lahtinen, S. Malola, J. Koivisto, and H. Häkkinen. Solvation chemistry of water-soluble thiol-protected gold nanocluster Au102 from DOSY NMR spectroscopy and DFT calculations. *Nanoscale*, 6: 7823–7826, **2014**.

[36] K. Salorinne, S. Malola, O. Andrea Wong, C. D Rithner, x. Chen, C. Ackerson, and H. Häkkinen. Conformation and dynamics of the ligand shell of a water-soluble Au 102 nanoparticle. *Nat. Commun.*, 7: 10401, **2016**.

[37] Y. Chen and R. Jin. Chapter 10 - Atomically Precise Gold Nanoclusters Catalyzed Chemical Transformations. In T. Tsukuda and H. Häkkinen, editors, *Protected Metal Clusters: From Fundamentals to Applications*, volume 9 of *Frontiers of Nanoscience*, pages 263 – 296. Elsevier, **2015**.

[38] M. Azubel, J. Koivisto, S. Malola, D. Bushnell, G. L. Hura, A. L. Koh, H. Tsunoyama, T. Tsukuda, M. Pettersson, H. Häkkinen, and R. D. Kornberg. Electron microscopy of gold nanoparticles at atomic resolution. *Science*, 345(6199): 909–912, **2014**.

[39] M. Azubel, S. D. Carter, J. Weiszmann, J. Zhang, G. J. Jensen, Y. Li, and R. D. Kornberg. FGF21 trafficking in intact human cells revealed by cryo-electron tomography with gold nanoparticles. *Elife*, 8, **2019**.

[40] Nonappa, T. Lahtinen, J. S. Haataja, T.-R. Tero, H. Häkkinen, and O. Ikkala. Template-Free Supracolloidal Self-Assembly of Atomically Precise Gold Nanoclusters: From 2D Colloidal Crystals to Spherical Capsids. *Angew. Chem. Int. Ed.*, 55(52): 16035–16038, **2016**.

[41] V. Marjomäki, T. Lahtinen, M. Martikainen, J. Koivisto, S. Malola, K. Salorinne, M. Pettersson, and H. Häkkinen. Site-specific targeting of enterovirus capsid by functionalized monodisperse gold nanoclusters. *PNAS*, 111(4): 1277–1281, **2014**.

[42] M. Martikainen, K. Salorinne, T. Lahtinen, S. Malola, P. Permi, H. Häkkinen, and V. Marjomäki. Hydrophobic pocket targeting probe for enteroviruses. *Nanoscale*, 7, **2015**.

[43] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, Oxford University Press, New York, **1991**.

[44] S. A. Adcock and J. A. McCammon. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.*, 106(5): 1589–1615, **2006**.

[45] J. R. Perilla, B. C. Goh, C. K. Cassidy, B. Liu, R. C. Bernardi, T. Rudack, H. Yu, Z. Wu, and K. Schulten. Molecular dynamics simulations of large macromolecular complexes. *Curr Opin Struct Biol.*, 31: 64 – 74, **2015**. Theory and simulation/Macromolecular machines and assemblies.

[46] J. L. Klepeis, K. Lindorff-Larsen, R. O. Dror, and D. E. Shaw. Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struct Biol.*, 19(2): 120 – 127, **2009**. Theory and simulation / Macromolecular assemblages.

[47] T. Reddy and M. S. Sansom. Computational virology: From the inside out. *Biochim Biophys Acta Biomembr.*, 1858(7, Part B): 1610 – 1618, **2016**. New approaches for bridging computation and experiment on membrane proteins.

[48] J. A. Hadden and J. R. Perilla. All-atom virus simulations. *Curr Opin Virol.*, 31: 82 – 91, **2018**. Virus structure and expression • Viral evolution.

[49] R. Bowley and M. Sánchez. *Introductory statistical mechanics*. Clarendon Press, 2nd ed edition, **1999**.

[50] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, Academic Press, San Diego, **2002**.

[51] W. F. van Gunsteren and H. J. C. Berendsen. Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *Angew. Chem. Int. Ed.*, 29(9): 992–1023, **1990**.

[52] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 78(8): 1950–1958, **2010**.

[53] J. Ponder and D. Case. Force fields for protein simulations. Protein simulations. *Adv. Prot. Chem*, 66: 27–85, **2003**.

[54] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9): 1157–1174, **2004**.

[55] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2): 926–935, **1983**.

[56] C. Caleman, P. J. van Maaren, M. Hong, J. S. Hub, L. T. Costa, and D. van der Spoel. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.*, 8(1): 61–74, **2012**.

[57] D. van der Spoel, P. J. van Maaren, and C. Caleman. GROMACS molecule liquid database. *Bioinformatics*, 28(5): 752–753, **2012**.

[58] J. M. Haile. *Molecular Dynamics Simulation: Elementary Methods*. John Wiley Sons, Inc., New York, NY, USA, 1st edition, **1992**.

[59] O. Becker, A. MacKerell, B. Roux, and M. Watanabe. *Computational Biochemistry and Biophysics*. Taylor & Francis, **2001**.

[60] I.-C. Yeh and G. Hummer. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B*, 108(40): 15873–15879, **2004**.

[61] D. A. Kofke. Free energy methods in molecular simulation. *Fluid Ph. Equilibria*, 228-229: 41 – 48, **2005**. PPEPPD 2004 Proceedings.

[62] T. Lelièvre, G. Stoltz, and M. Rousset. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, **2010**.

[63] C. Dellago and G. Hummer. Computing Equilibrium Free Energies Using Non-Equilibrium Molecular Dynamics. *Entropy*, 16: 41–61, **2013**.

[64] H. Xiong, A. Crespo López, M. Marti, D. Estrin, and A. E. Roitberg. Free Energy Calculations with Non-Equilibrium Methods: Applications of the Jarzynski Relationship. *Theor Chem Acc*, 116: 338–346, **2006**.

[65] C. Jarzynski. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.*, 78: 2690–2693, **1997**.

[66] G. E. Crooks. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.*, 90: 1481–1487, **1998**.

[67] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60: 2721–2726, **1999**.

[68] D. M. Zuckerman and T. B. Woolf. Overcoming finite-sampling errors in fast-switching free-energy estimates: extrapolative analysis of a molecular system. *Chem. Phys. Lett.*, 351(5): 445 – 453, **2002**.

[69] F. M. Ytreberg and D. M. Zuckerman. Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems. *J. Comput. Chem.*, 25(14): 1749–1759, **2004**.

[70] G. Hummer and A. Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *PNAS*, 98(7): 3658–3661, **2001**.

[71] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante. Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality. *Science*, 296(5574): 1832–1835, **2002**.

[72] D. Collin, F. Ritort, C. Jarzynski, S. B Smith, I. J. Tinoco, and C. Bustamante. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature*, 437: 231–4, **2005**.

[73] H. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.*, 91(1): 43 – 56, **1995**.

[74] E. Lindahl, B. Hess, and D. van der Spoel. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model*, 7: 306–317, **2001**.

[75] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. GROMACS: Fast, flexible, and free. *J. Comput. Chem.*, 26(16): 1701–1718, **2005**.

[76] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, 4(3): 435–447, **2008**.

[77] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7): 845–854, **2013**.

[78] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2: 19 – 25, **2015**.

[79] J. M. Devi. Simulation studies on structural and thermal properties of alkane thiol capped gold nanoparticles. *J. Mol. Graph.*, 74: 359 – 365, **2017**.

[80] A. K. Giri and E. Spohr. Influence of Chain Length and Branching on the Structure of Functionalized Gold Nanoparticles. *J. Phys. Chem. C*, 122(46): 26739–26747, **2018**.

[81] A. K. Chew and R. C. Van Lehn. Effect of Core Morphology on the Structural Asymmetry of Alkanethiol Monolayer-Protected Gold Nanoparticles. *J. Phys. Chem. C*, 122(45): 26288–26297, **2018**.

[82] R. C. Van Lehn and A. Alexander-Katz. Structure of Mixed-Monolayer-Protected Nanoparticles in Aqueous Salt Solution from Atomistic Molecular Dynamics Simulations. *J. Phys. Chem. C*, 117(39): 20104–20115, **2013**.

[83] O. Villarreal, R. A. Rodriguez, L. Yu, and T. Wambo. Molecular dynamics simulations on the effect of size and shape on the interactions between negative Au18(SR)14, Au102(SR)44 and Au144(SR)60 nanoparticles in physiological saline. *Colloids Surf. A*, 503, **2016**.

[84] S. A. Alsharif, L. Y. Chen, A. Tlahuice-Flores, R. L. Whetten, and M. J. Yacaman. Interaction between functionalized gold nanoparticles in physiological saline. *Phys. Chem. Chem. Phys.*, 16: 3909–3913, **2014**.

[85] D. M. Chevrier, L. Raich, C. Rovira, A. Das, Z. Luo, Q. Yao, A. Chatt, J. Xie, R. Jin, J. Akola, and P. Zhang. Molecular-Scale Ligand Effects in Small Gold–Thiolate Nanoclusters. *J. Am. Chem. Soc.*, 140(45): 15430–15436, **2018**.

[86] E. Heikkilä, A. A. Gurtovenko, H. Martinez-Seara, H. Häkkinen, I. Vattulainen, and J. Akola. Atomistic Simulations of Functional Au144(SR)60 Gold Nanoparticles in Aqueous Environment. *J. Phys. Chem. C*, 116(17): 9805–9815, **2012**.

[87] O. A. Perfilieva, D. V. Pyshnyi, and A. A. Lomzov. Molecular Dynamics Simulation of Polarizable Gold Nanoparticles Interacting with Sodium Citrate. *J. Chem. Theory Comput.*, 15(2): 1278–1292, **2019**.

[88] A. Kyrychenko, G. V. Karpushina, S. I. Bogatyrenko, A. P. Kryshtal, and A. O. Doroshenko. Preparation, structure, and a coarse-grained molecular dynamics model for dodecanethiol-stabilized gold nanoparticles. *Comput. Theor. Chem.*, 977(1): 34 – 39, **2011**.

[89] J. M. Devi. Aggregation of thiol coated gold nanoparticles: A simulation study on the effect of polymer coverage density and solvent. *Comput. Mater. Sci.*, 86: 174 – 179, **2014**.

[90] T. Kister, D. Monego, P. Mulvaney, A. Widmer-Cooper, and T. Kraus. Colloidal Stability of Apolar Nanoparticles: The Role of Particle Size and Ligand Shell Structure. *ACS Nano*, 12(6): 5969–5977, **2018**.

[91] R. C. Van Lehn and A. Alexander-Katz. Ligand-Mediated Short-Range Attraction Drives Aggregation of Charged Monolayer-Protected Gold Nanoparticles. *Langmuir*, 29(28): 8788–8798, **2013**.

[92] W. Phanchai, U. Srikulwong, A. Chompoosor, C. Sakonsinsiri, and T. Puangmali. Insight into the Molecular Mechanisms of AuNP-Based Aptasensor for Colorimetric Detection: A Molecular Dynamics Approach. *Langmuir*, 34(21): 6161–6169, **2018**.

[93] P. Toomjeen, W. Phanchai, C. Choodet, A. Chompoosor, R. Thanan, C. Sakonsinsiri, and T. Puangmali. Designing an Aptasensor Based on Cysteamine-Capped AuNPs for 8-Oxo-dG Detection: A Molecular Dynamics Approach and Experimental Validation. *J. Phys. Chem. B*, 123(5): 1129–1138, **2019**.

[94] M. Khavani, M. Izadyar, and m. R. Housaindokht. A combined MD/QM study on the sensing mechanism of Pb2+ by glutathione functionalized gold nanoparticles. *J. Mol. Liq.*, 280, **2019**.

[95] M. Khavani, M. Izadyar, and M. R. Housaindokht. Modeling of the Functionalized Gold Nanoparticle Aggregation in the Presence of Dopamine: A Joint MD/QM Study. *J. Phys. Chem. C*, 122(45): 26130–26141, **2018**.

[96] S. W. Shin, S. Y. Ahn, S. Yoon, H. S. Wee, J. W. Bae, J. H. Lee, W. B. Lee, and S. H. Um. Differences in DNA Probe-Mediated Aggregation Behavior of Gold Nanomaterials Based on Their Geometric Appearance. *Langmuir*, 34(49): 14869–14874, **2018**.

[97] G. Brancolini, L. Bellucci, M. C. Maschio, R. D. Felice, and S. Corni. The interaction of peptides and proteins with nanostructures surfaces: a challenge for nanoscience. *Curr. Opin. Colloid Interface Sci.*, 41: 86 – 94, **2019**.

[98] L. Riccardi, L. Gabrielli, X. Sun, F. De Biasi, F. Rastrelli, F. Mancin, and M. De Vivo. Nanoparticle-Based Receptors Mimic Protein-Ligand Recognition. *Chem*, 3: 92–109, **2017**.

[99] F. Simonelli, G. Rossi, and L. Monticelli. Role of Ligand Conformation on Nanoparticle–Protein Interactions. *J. Phys. Chem. B*, 123(8): 1764–1769, **2019**.

[100] F. Tavanti, A. Pedone, and M. C. Menziani. Competitive Binding of Proteins to Gold Nanoparticles Disclosed by Molecular Dynamics Simulations. *J. Phys. Chem. C*, 119(38): 22172–22180, **2015**.

[101] V. Rojas-Cervellera, L. Raich, J. Akola, and C. Rovira. The molecular mechanism of the ligand exchange reaction of an antibody against a glutathione-coated gold cluster. *Nanoscale*, 9: 3121–3127, **2017**.

[102] M. Tang, N. S. Gandhi, K. Burrage, and Y. Gu. Interaction of gold nanosurfaces/nanoparticles with collagen-like peptides. *Phys. Chem. Chem. Phys.*, 21: 3701–3711, **2019**.

[103] Q. Shao and C. K. Hall. Binding Preferences of Amino Acids for Gold Nanoparticles: A Molecular Simulation Study. *Langmuir*, 32(31): 7888–7896, **2016**.

[104] M. Samieegohar, F. Sha, A. Z. Clayborne, and T. Wei. ReaxFF MD Simulations of Peptide-Grafted Gold Nanoparticles. *Langmuir*, 35(14): 5029–5036, **2019**.

[105] P. U. Atukorale, Z. P. Guven, A. Bekdemir, R. P. Carney, R. C. Van Lehn, D. S. Yun, P. H. Jacob Silva, D. Demurtas, Y.-S. Yang, A. Alexander-Katz, F. Stellacci, and D. J. Irvine. Structure–Property Relationships of Amphiphilic Nanoparticles That Penetrate or Fuse Lipid Membranes. *Bioconjug Chem.*, 29(4): 1131–1140, **2018**.

[106] R. Gupta and B. Rai. Penetration of Gold Nanoparticles through Human Skin: Unraveling Its Mechanisms at the Molecular Scale. *J. Phys. Chem. B*, 120(29): 7133–7142, **2016**.

[107] T. Lunnoo, J. Assawakhajornsak, and T. Puangmali. In Silico Study of Gold Nanoparticle Uptake into a Mammalian Cell: Interplay of Size, Shape, Surface Charge, and Aggregation. *J. Phys. Chem. C*, 123(6): 3801–3810, **2019**.

[108] E. S. Melby, C. Allen, I. U. Foreman-Ortiz, E. R. Caudill, T. R. Kuech, A. M. Vartanian, X. Zhang, C. J. Murphy, R. Hernandez, and J. A. Pedersen. Peripheral Membrane Proteins Facilitate Nanoparticle Binding at Lipid Bilayer Interfaces. *Langmuir*, 34(36): 10793–10805, **2018**.

[109] A. Torchi, F. Simonelli, R. Ferrando, and G. Rossi. Local Enhancement of Lipid Membrane Permeability Induced by Irradiated Gold Nanoparticles. *ACS Nano*, 11(12): 12553–12561, **2017**.

[110] R. C. Van Lehn and A. Alexander-Katz. Pathway for insertion of amphiphilic nanoparticles into defect-free lipid bilayers from atomistic molecular dynamics simulations. *Soft Matter*, 11: 3165–3175, **2015**.

[111] P. Angelikopoulos, L. Sarkisov, Z. Cournia, and P. Gkeka. Self-assembly of anionic, ligand-coated nanoparticles in lipid membranes. *Nanoscale*, 9: 1040–1048, **2017**.

[112] X. Chen, D. P. Tieleman, and Q. Liang. Modulating interactions between ligand-coated nanoparticles and phase-separated lipid bilayers by varying the ligand density and the surface charge. *Nanoscale*, 10: 2481–2491, **2018**.

[113] J.-e. Park, M. Seo, E. Jang, H. Kim, J. S. Kim, and S.-J. Park. Vesicle-like assemblies of ligand-stabilized nanoparticles with controllable membrane composition and properties. *Nanoscale*, 11: 1837–1846, **2019**.

[114] Z. Shen, W. Baker, H. Ye, and Y. Li. pH-Dependent aggregation and pH-independent cell membrane adhesion of monolayer-protected mixed charged gold nanoparticles. *Nanoscale*, **2019**.

[115] R. C. Van Lehn and A. Alexander-Katz. Energy landscape for the insertion of amphiphilic nanoparticles into lipid membranes: A computational study. *PLoS One*, 14(1): 1–19, **2019**.

[116] R. C. Van Lehn and A. Alexander-Katz. Fusion of Ligand-Coated Nanoparticles with Lipid Bilayers: Effect of Ligand Flexibility. *J. Phys. Chem. A*, 118(31): 5848–5856, **2014**.

[117] R. C. Van Lehn and A. Alexander-Katz. Membrane-Embedded Nanoparticles Induce Lipid Rearrangements Similar to Those Exhibited by Biological Membrane Proteins. *J. Phys. Chem. B*, 118(44): 12586–12598, **2014**.

[118] R. C Van Lehn, M. Ricci, P. H.J. Silva, P. Andreozzi, J. Reguera, K. Voitchovsky, F. Stellacci, and A. Alexander-Katz. Lipid tail protrusions mediate the insertion of nanoparticles into model cell membranes. *Nat. Commun.*, 5: 4482, **2014**.

[119] R. C. Van Lehn, P. U. Atukorale, R. P. Carney, Y.-S. Yang, F. Stellacci, D. J. Irvine, and A. Alexander-Katz. Effect of Particle Diameter and Surface Composition on the Spontaneous Fusion of Monolayer-Protected Gold Nanoparticles with Lipid Bilayers. ***N**ano Lett.*, 13(9): 4060–4067, **2013**.

[120] J. Lin and A. Alexander-Katz. Cell Membranes Open "Doors" for Cationic Nanoparticles/Biomolecules: Insights into Uptake Kinetics. *ACS Nano*, 7(12): 10799–10808, **2013**.

[121] H. Heinz, R. A. Vaia, B. L. Farmer, and R. R. Naik. Accurate Simulation of Surfaces and Interfaces of Face-Centered Cubic Metals Using 12-6 and 9-6 Lennard-Jones Potentials. *J. Phys. Chem. C*, 112(44): 17281–17290, **2008**.

[122] S. Banerjee, J. Montgomery, and J. Gascón. A QM/MM approach for the study of monolayer-protected gold clusters. *J. Mater. Sci.*, **2012**.

[123] T. Lahtinen, E. Hulkko, K. Sokołowska, T.-R. Tero, V. Saarnio, J. Lindgren, M. Pettersson, H. Häkkinen, and L. Lehtovaara. Covalently linked multimers of gold nanoclusters Au102(p-MBA)44 and Au 250(p–MBA)n. *Nanoscale*, 8: 18665–18674, **2016**.

[124] S. Franco-Ulloa, L. Riccardi, F. Rimembrana, M. Pini, and M. De Vivo. NanoModeler: A Webserver for Molecular Simulations and Engineering of Nanoparticles. *J. Chem. Theory Comput.*, 15(3): 2022–2032, **2019**.

[125] S. Malola, P. Nieminen, J. Hämäläinen, T. Kärkkäinen, and H. Häkkinen. A General Method for Structure Prediction of Metal-Ligand Interfaces of Hybrid Nanoparticles, **2019**. Preprint in ChemRxiv (March 1).

[126] M. Rigoulet, S. Massou, E. D. Sosa Carrizo, S. Mallet-Ladeira, A. Amgoune, K. Miqueu, and D. Bourissou. Evidence for genuine hydrogen bonding in gold(I) complexes. *PNAS*, 116(1): 46–51, **2019**.

[127] S. Donnini, F. Tegeler, G. Groenhof, and H. Grubmüller. Constant pH Molecular Dynamics in Explicit Solvent with λ-Dynamics. *J. Chem. Theory Comput.*, 7(6): 1962–1978, **2011**.

[128] T. Dainese, S. Antonello, S. Bogialli, W. Fei, A. Venzo, and F. Maran. Gold Fusion: From Au25(SR)18 to Au38(SR)24, the Most Unexpected Transformation of a Very Stable Nanocluster. *ACS Nano*, 12(7): 7057–7066, **2018**.

APPENDIX 1 SIMULATIONS AND ANALYSES AU_{102}*P*MBA_{44} NANOCLUSTER SUPERSTRUCTURES

APPENDIX 1.1 Simulations

All simulations were performed using Gromacs 5.1.4 [78] simulation package and Amber99sb-ildn [52] force field with parameters for thiolate protected nanoclusters from work **I**.

2D sheet structures were formed by placing one 8x8 sheet to hexagonal simulation box to generate infinite sheet structure in xy direction when using periodic boundaries, while in the z direction the distance between sheets is 25 nm. For the spherical shells dodecahedric simulation boxes were used with distance to solvent 1.2 nm.

The solvent was added to the box with gmx solvate tool: TIP3P water and methanol with GAFF (generalized Amber force field) parameters obtained from Virtual chemistry [56,57] were used. Ions (NaCl) were added to a concentration of 0.1 M + neutralizing cations. Depending on the protonation state of the nanoclusters, the total ion concentration varies between 0.1-1.0 M. In systems where solvents, water and methanol, are initially separated to restricted areas (figure 12 g, h) the whole system was first solvated in water which was then removed from the volume then resolvated with methanol. Ions were added to reflect their solubility in each solvent, i.e., 90% of the ions were added to water volume, and the remaining to methanol volume.

Each system was first energy minimized with steepest descents. For solvent relaxation 1 ns NVT run with position restraints on solute atoms excluding the hydrogens was performed. In order to avoid bias from placing the clusters initially in identical orientations, the systems were equilibrated at 300 K NVT for 10 ns, with translational motion of each cluster restricted while rotational motion was allowed. This equilibration was then followed by 30 ns to 200 ns production run at 300 K, 1 bar NPT with no restraints.

In all NVT and NPT simulations we used leap frog integrator with 2 fs time step, periodic boundary conditions, a 1.0 nm Lennard-Jones cut-off with dispersion correction for energy and pressure, PME electrostatics with a 1.0 nm cut-off and 0.12 nm grid spacing, the velocity-rescale thermostat with a reference temperature of 300 K and coupling time constant of 0.1 ps, and the Berendsen barostat with a reference pressure of 1 bar and coupling time constant of 1 ps. To avoid breaking of the sheets semi-isotropic pressure coupling was used with 2D sheet systems, with 0 compressibility in the xy direction. In systems with pure methanol we used methanol compressibility as reported in Virtualchemistry of 1.09 1/GPa. For pure water and water-methanol mixtures we used compressibility of water 0.45 1/GPa. All covalent bond lengths were constrained with the LINCS algorithm.

APPENDIX 1.2 Stability analyses

To determine the retaining of the superstructure, in addition to visual inspection, distances between neighboring gold nanoclusters in 2D sheets were tracked and compared to the experimentally determined distance (2.7 nm) between the clusters in the hexagonally packed sheet. The structures of spherical shells were analyzed using radii of gyration determined using gmx gyrate tool with default settings.

The non-covalent interactions between gold nanoclusters are the basis of superstructure stability. Firstly, the acidic *p*MBA ligands can form hydrogen bonds, thus given the optimal solvent condition, orientation and the protonation pattern of the clusters, the stable superstructure could be expected to form. Hydrogen bonds between *p*MBA ligands in production NPT trajectory of 30-200 ns were analyzed using Gromacs gmx hbond tool with default cut-off angle and radius of 30°and 0.35 nm, respectively.

Secondly, *p*MBA ligand in deprotonated form carries a negative charge, thus we have investigated ionic interactions between nanoclusters and cations. Especially, we have looked at so-called chelating cation interactions, where the cation lies close (within 3 Å) to carboxyl acid group carbon of two ligands simultaneously.

For both type of interactions the number of contacts averaged per cluster are presented. The average, standard deviation and drift are reported, with drift calculated as the difference between last and first point of a linear fit to the data in question.

For consistent ranking of the system stabilities, we performed a simple fit to find weights w_{hbonds} and w_{ions} for number of hydrogen bonds Hb and chelating cations Ic according to $R = w_{\text{hbonds}} * Hb + w_{\text{ions}} * Ic$. R is defined as the goodness or stability score based on the closeness of average neighbor distance to the experimental reference value, and the percentage of neighbor cluster distances remaining within 0.2 nm of this reference value during the simulation in 2D sheets. In spherical shell systems score is based on gyration radius. From the fitted weights the new stability scores are calculated and compared between the systems.

ORIGINAL PAPERS

I

A UNIFIED AMBER-COMPATIBLE MOLECULAR MECHANICS FORCE FIELD FOR THIOLATE-PROTECTED GOLD NANOCLUSTERS

Reprinted with permission from

Emmi Pohjolainen, Xi Chen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen.
Journal of Chemical Theory and Computation, 12 (3), pp. 1342–1350, **2016**.

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

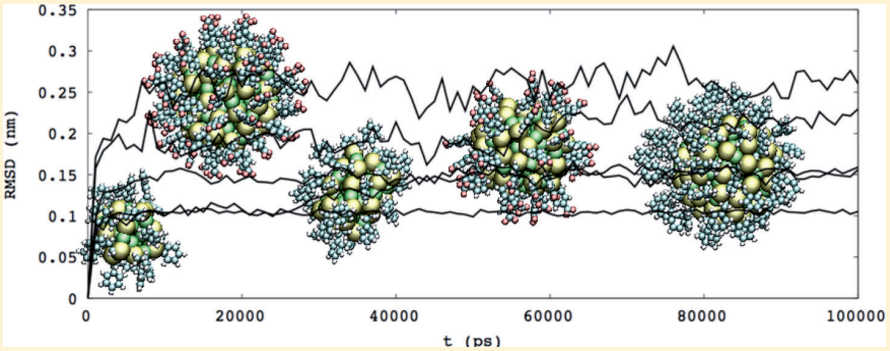
A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters

Emmi Pohjolainen,[†] Xi Chen,[‡] Sami Malola,[†] Gerrit Groenhof,*[,‡] and Hannu Häkkinen*[,†,‡]

[†]Department of Physics, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland, FI-40014

[‡]Department of Chemistry, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland, FI-40014

Supporting Information



ABSTRACT: We present transferable AMBER-compatible force field parameters for thiolate-protected gold nanoclusters. Five different sized clusters containing both organo-soluble and water-soluble thiolate ligands served as test systems in MD simulations, and parameters were validated against DFT and experimental results. The cluster geometries remain intact during the MD simulations in various solvents, and structural fluctuations and energetics showed agreement with DFT calculations. Experimental diffusion coefficients and crystal structures were also reproduced with sufficient accuracy. The presented parameter set contains the minimum number of cluster-specific parameters enabling the use of these parameters for several different gold nanoclusters. The parameterization of ligands can also be extended to different types of ligands.

INTRODUCTION

Thiolate-protected gold nanoclusters $Au_n(SR)_m$ have been extensively studied during the past years, both experimentally and theoretically.[1−4] Clusters studied so far range in size and nature of the ligand and exhibit interesting electronic,[5,6] optical,[7] or catalytic properties[8,9] that could make them promising building blocks for nanotechnology applications. Because the ligand R-group can easily be made biocompatible, these thiolate-protected clusters are also increasingly used in biological research. For example, the clusters have been conjugated with biomolecules to act as contrast agents in electron microscopy imaging[10] or as molecular rulers.[11,12]

Whereas density functional theory (DFT) has been widely used to investigate the electronic, optical, and even catalytic properties of gold nanoclusters, the system size and time scales that must be reached to investigate the interactions between gold clusters and biomolecules precludes DFT and necessitates classical molecular dynamics simulation in combination with a molecular mechanics force field instead. To date, several molecular dynamics simulations have been performed on systems containing gold, either as a surface[13,14] or nanoparticle.[15−29] However, as most of these studies were focused on a specific gold cluster, the force field parameters are often not transferable to clusters with different sizes or ligands.

A notable exception are reactive force fields for gold nanoparticles, such as REAX-FF,[30−33] which are readily transferred between clusters. However, due to the high complexity of the potential functions, large number of parameters, and lack of compatibility with the available biomolecular force fields, reactive force fields are not yet available in mainstream molecular dynamics programs. Moreover, the small time step with which the equations of motion have to be integrated in reactive force field simulations creates a severe bottleneck for simulating large biological systems. Alternatively, the QM/MM approach,[34] in which the gold cluster is described at the QM level, is transferable[35] as well, but the computational effort required for the QM calculation restricts accessing the biologically relevant time scales in the simulation.

To systematically investigate the effect of size and composition of thiolate-protected gold clusters on their interplay with biomolecules, such as proteins, membranes, or nucleic acids, we need a transferable force field that is compatible with an existing biomolecular force field. Since the AMBER force field[36] has been specifically developed for a

Received: November 5, 2015
Published: February 4, 2016

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

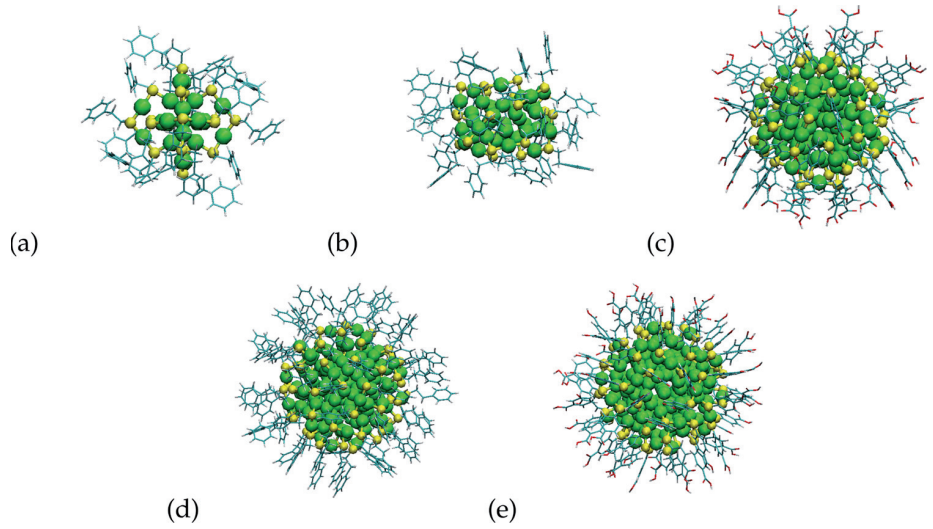


Figure 1. Clusters studied in this work. Coloring convention: green = gold; yellow = sulfur; cyan = carbon; white = hydrogen; red = oxygen. *p*MBA = para-mercaptobenzoic acid and PET = phenyl-ethanethiol. (a) $Au_{25}(PET)_{18}^{-}$, (b) $Au_{38}(PET)_{24}$, (c) $Au_{102}(pMBA)_{44}$, (d) $Au_{144}(PET)_{60}$, and (e) $Au_{144}(pMBA)_{60}$.

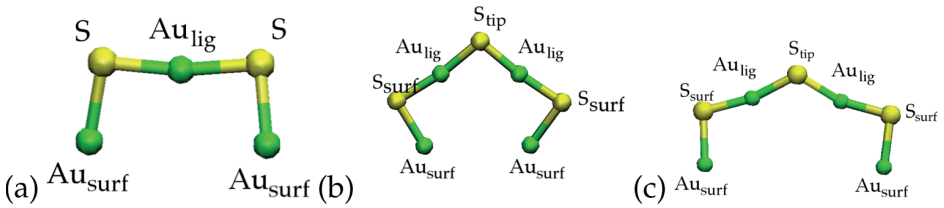


Figure 2. Different ligand unit formations related to clusters studied in this work with naming conventions and colors: gold (green) and sulfur (yellow). (a) Rectangular staple formation present in Au_{38}, Au_{102}, and Au_{144} clusters. The V-shape staple formation in (b) appears in Au_{25} and Au_{38} clusters and in (c) in Au_{102} clusters (together with (a)). The alkylthiol groups are omitted.

wide range of biomolecules, we have optimized a set of generic parameters to model the energetics of thiolate-protected gold clusters with the AMBER potential functions. To facilitate transferability, we avoided bonding interactions between gold atoms in the core and rely on noncovalent interaction between gold atoms instead to maintain the core structure. To verify the validity of the new force field, we have performed MD simulations on five representative thiolate-protected gold nanoclusters (Figure 1) and compared the results of these simulations to the available experimental data and DFT computations.

METHODS

In this work, we derive general force field parameters for thiolate-protected gold nanoclusters based on a representative set of such clusters. The geometries of these clusters were obtained from DFT calculations. All MD simulations were performed with Gromacs, versions 4.5 and 4.6.[37]

Parameterization. The shape and anatomy of a gold cluster is controlled by the number of gold atoms, chemical nature of the ligands, and number of these ligands. The core consists of gold atoms arranged in shells. This arrangement depends on the number of gold atoms. For example, the Au_{25} cluster has one gold atom in the center around which 12 gold atoms form an icosahedron. In Au_{38}, the core is less symmetric and is composed of two merged icosahedra. Larger clusters contain more gold layers with decahedral (Au_{102}) and icosahedral (Au_{144}) centers and have a more globular overall shape.

In previous force fields, constraints were employed to fix the distances between the gold atoms of the core and maintain its geometry.[16,19] Such an approach, however, requires knowledge about the interatomic distances, which vary with cluster size. Therefore, force fields based on constraints between the core atoms are not transferable and cannot be used to predict the geometry of new gold clusters.

Instead, transferability requires that the force field functions and parameters for the gold core do not depend on the details of the cluster. To fulfill that requirement, we used the noncovalent Lennard-Jones interatomic potential with parameters for metallic gold,[38] as in ref 19, 21, and 39. In addition, no partial charges are assigned to the gold atoms. Although this choice was obviously motivated by our requirement that the force field parameters are transferable to other clusters, it was justified by Bader analyses on various thiolate-protected gold nanoclusters, which suggest that the core atoms are indeed almost neutral in most clusters.[5,35,40]

The core is separated from the ligands by a sulfur−gold interface layer that protects the core. The sulfur atoms connect the ligands to the core by forming so-called staples (Figure 2), in which the thiolates bind covalently with one gold atom at the surface (Au_{surf}) and another at the interface (Au_{lig}). To describe these covalent interactions with a minimal set of parameters, valid for a wide range of gold clusters, the equilibrium bond lengths and angles were defined as the average of these values

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

in the DFT-optimized structures of the five clusters, shown in Figure 1. However, because the $Au_{lig}-S_{tip}-Au_{lig}$ angle in the V-shaped staples ranges from 100° in the smaller $Au_{25}(PET)_{18}$ and $Au_{38}(PET)_{18}$ clusters to 119.2° in the larger $Au_{102}(PET)_{44}$ cluster (Figure 2), it was not possible to use a single equilibrium value for this angle in all clusters. Therefore, we introduced two values instead, depending on the shape of the staple. Nevertheless, because the choice for which angle value to use can be determined either *a priori* or *a posteriori*, having two values for this angle does not affect transferrability.

The harmonic force constant for the Au_x-S bonds was adopted from ref 35. This force constant is approximately an order of magnitude smaller than the usual angle force constants in AMBER. The harmonic force constants for the Au_x-S-Au_x and $S-Au_{lig}-S$ angles were taken from ref 19, which are comparable to those in AMBER. Nonbonded 1−4 interactions involving the atoms of a staple were excluded and implicitly accounted for by the torsion potential.

In the V-shaped staples of the $Au_{102}(PET)_{44}$ cluster, it was necessary to increase the $S-Au_{lig}-S$ angle force constant in order to maintain the strained staple shape open and stable in simulations. This force constant was also taken from ref 19 and is about three times larger than the typical angle force constants in AMBER. For angles involving gold and ligand atoms (Au_x-S-C_x), we used the same force constants as Banerjee et al.,[35] which are somewhat smaller than those in AMBER.

Finally, the ligand layer consists of organic molecules that contain atom types for which AMBER parameters are already available. However, for some combinations of these atoms, bonded or angle parameters were missing. Missing equilibrium values were determined by averaging these values over the DFT optimized structures, whereas missing force constants were selected based on similarity with existing interactions. The partial charges on the ligand atoms were optimized with Ambertools12,[41] following the RESP charge fitting procedure recommended for AMBER.[42] Charges were derived for a model system consisting of two gold atoms connected to two ligands via a sulfur atom, that is, $Au_2(PET)_2$ and $Au_2(pMBA)_2$. Similar model systems have been used previously in gold−thiolate parameter derivations.[35,43] The geometries of $Au_2(PET)_2$ and $Au_2(pMBA)_2$ were optimized using Gaussian 09,[44] using the B3LYP functional in combination with LANL2DZ basis set and the W06 density fitting basis.[45] The electron density of the optimized structure was used to calculate the electrostatic potential at points selected according to Merz−Singh−Kollman scheme[46,47] in Gaussian 09. The radius of Au was set to 0.17 nm. Atomic charges were fitted to this potential in a two-stage RESP procedure, constraining the charges of Au to zero. For the *p*MBA ligand the partial charges were fitted for both protonated and deprotonated structures. For the charged $Au_{25}PET_{18}^-$ cluster, the negative charge was distributed over all ligands by scaling all charges. The optimized charges were compared to the default AMBER charges of similar residue parts, that is, phenylalanine, with no significant discrepancies observed. Although we only considered the *p*MBA and PET ligands (Figure 3), the procedure to parameterize the ligand is sufficiently general to be applied to other ligands as well. Therefore, parameters for new ligands can be easily obtained without modifying any of the core or interface parameters.

Force Field Validation. *Molecular Dynamics Simulations.* The new force field was tested by performing 100 ns MD simulations of the five clusters (Figure 1) in various solvents. In all simulations, the leapfrog Verlet integrator was used with a

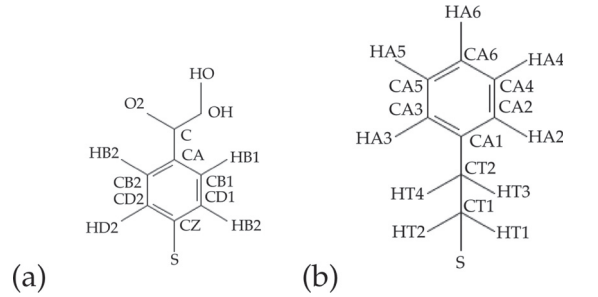


Figure 3. Alkylthiol ligands and naming conventions used in this work: (a) *p*MBA = para-mercaptobenzoic acid and (b) PET = phenylethanethiol.

time step of 2 fs, periodic boundary conditions, 1.0 nm Lennard-Jones cutoff with dispersion correction for energy and pressure, PME electrostatics with a 1.0 nm cutoff and 0.12 nm grid spacing,[48] velocity-rescale thermostat with a reference temperature of 298.15 K and coupling time constant of 0.1 ps,[49] and Berendsen barostat with a reference pressure of 1 bar and coupling time constant of 1 ps.[50] All bond lengths were constrained with the LINCS algorithm.[51] The $Au_{25}(PET)_{18}$ and $Au_{38}(PET)_{24}$ clusters were simulated in chloroform and the $Au_{144}(PET)_{60}$ cluster in dichloromethane. Chloroform and dichloromethane (GAFF[52]) parameters and structures were obtained from the Virtual chemistry Web site (virtualchemistry.org).[53,54] Both neutral and anionic forms of the $Au_{25}(PET)_{18}^-$ cluster were simulated. Clusters $Au_{102}(pMBA)_{44}$ and $Au_{144}(pMBA)_{60}$ were simulated in TIP3P water, with all *p*MBA groups deprotonated. The SETTLE algorithm was used to constrain the internal degrees of freedom of the water molecules.[55] Sodium and chloride ions were added to neutralize the system at a concentration of 0.1 M. Prior to the production simulations, each system was energy minimized and equilibrated under NVT (100 ps) and NPT (100 ps) conditions. The deprotonated $Au_{102}(pMBA)_{44}$ and $Au_{144}(pMBA)_{60}$ systems were equilibrated for 1 ns under NPT conditions with a 1 fs time step.

Density Functional Theory Computations. All the model cluster structures $Au_{25}(PET)_{18}^-$, $Au_{38}(PET)_{24}$, $Au_{102}(pMBA)_{44}$, $Au_{144}(PET)_{60}$, and $Au_{144}(pMBA)_{60}$ were relaxed to the energy minimum configuration using density functional theory (DFT) with PBE functional as implemented in the software package GPAW that uses real-space grids and projector-augmented waves.[56] For the relaxation of the smaller clusters, $Au_{25}(PET)_{18}^-$ and $Au_{38}(PET)_{24}$, the criterion for the maximum residual forces acting on atoms was 0.05 eV/Å, whereas for the larger clusters, $Au_{102}(pMBA)_{44}$, $Au_{144}(PET)_{60}$, and $Au_{144}(pMBA)_{60}$, the criterion was 0.1 eV/Å. A grid spacing of 0.2 Å was used in all calculations.

Activation barriers for flipping the chirality of the sulfur atoms were computed for a model system containing a single staple with three thiolate ligands (PET or *p*MBA) and four gold atoms (Figure 4). The reaction coordinate was defined as the torsion angle between the Au_1-S_2 and S_3-C_4 bond vectors. This angle was varied in 23 and 18 discrete steps of 5.7°and 6.5° for the *p*MBA and PET model systems, respectively, followed by a constrained geometry optimization keeping both torsion angle and the two dangling gold atoms fixed. Afterward, the potential energy profiles were re-evaluated at the force field level to verify the validity of the latter.

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

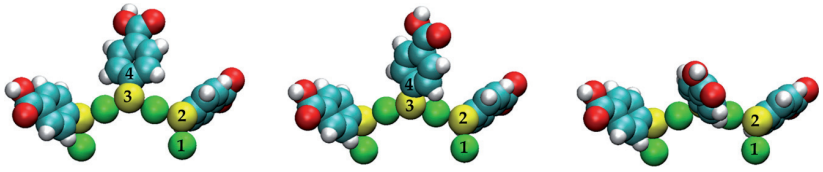


Figure 4. Model system and reaction coordinate for computing the energy profile for flipping the chirality of the central sulfur atoms (three points from the flipping path shown). The reaction coordinate is defined as the torsion angle between the bond vectors Au_1-S_2 and S_3-C_4. For the rest of the ligand, both PET and *p*MBA are used, but only *p*MBA is shown here.

Further validation based on DFT was performed by comparing structural fluctuations between the force field and DFT MD simulations of $Au_{25}PET_{18}^-$ and $Au_{25}(SH)_{18}^-$ at 300 K. The force field simulations were carried out in vacuum, without cut-offs for Lennard-Jones or electrostatic interactions, using a stochastic dynamics integrator with a 1 fs time step, reference temperature of 300 K, and friction coefficient of 0.01 ps^{-1}. Bonds involving hydrogen atoms were constrained. The classical DFT-based molecular dynamics simulations were performed with the LDA functional and the LCAO double-ζ basis set using a Langevin integrator, with a friction of 0.004 ps^{-1} and a 1 fs time step. The $Au_{25}PET_{18}^-$ cluster was thermalized slowly to 300 K and subsequently simulated for 0.5 ps, whereas the 10 ps molecular dynamics trajectory of $Au_{25}(SH)_{18}^-$ was taken from previous work.[57]

Diffusion Coefficients. Diffusion coefficients for free diffusion of $Au_{25}(PET)_{18}$ and $Au_{38}(PET)_{24}$ in chloroform and $Au_{144}(PET)_{60}$ in dichloromethane were determined from 20 independent NPT simulations of 10 ns at 289.15 K and 1 bar. Initial conditions were generated by randomly assigning velocities from a Maxwell–Boltzmann distribution to the atoms of the starting structure that was created by placing the cluster in the center of a cubic periodic box filled with solvent. As before, the force field parameters for the chloroform and dichloromethane solvents were obtained from virtualchemistry.org.[53,54] To account for finite size effects, the simulations were repeated with six different box sizes.[58] Prior to the production runs, the simulation boxes were equilibrated for 2.1 ns.

Diffusion coefficients were calculated from the Einstein diffusion equation (in three dimensions):

$$D_{\text{cluster}} = \lim_{t\to\infty} \frac{\langle \|\mathbf{r}_i(t) - \mathbf{r}_i(0)\|^2 \rangle}{6t} \quad (1)$$

where $\langle \|\mathbf{r}_i(t) - \mathbf{r}_i(0)\|^2 \rangle$ denotes the mean square displacement (*MSD*). The Stokes–Einstein equation relates the hydrodynamic radius r of a spherical particle in homogeneous solution to its diffusion coefficient by

$$r = \frac{k_B T}{6\pi D_{\text{cluster}} \eta} \quad (2)$$

where η is the viscosity of the solvent.

Following Wang and Hou, we averaged the mean-square-displacements (MSDs) over the 20 independent 10 ns simulations for each cluster.[59] The diffusion coefficient was obtained by linear fitting the MSD from 1 to 9 ns in eq 1. As suggested in ref 58, a weighted linear fit was performed to the diffusion coefficients as a function of the inverse box length. The weights were obtained from the standard deviations of diffusion coefficients in each box. The diffusion coefficient at infinite dilution was estimated from the intercept of this fit.

Crystal Simulations. Further validation was performed by simulating a crystal of $Au_{102}(pMBA)_{44}$[60] under the experimental crystallization conditions and comparing the structural dynamics to the X-ray data. For this purpose, a periodic unit cell with 72 neutral (protonated) $Au_{102}(pMBA)_{44}$ clusters was created and solvated in a mixture of water/methanol/NaCl/NaOAc, which corresponds approximately to the crystallization conditions reported in experiment.[60] The water was described with the TIP3P model.[61] The methanol (GAFF) parameters were obtained from virtualchemistry.org, and sodium acetate was parameterized with Ambertools12, using the default settings. After energy minimization, the system was equilibrated for 1 ns at 200 K and fixed volume with position restraints on the heavy atoms of the clusters. After that, position restraints were removed, and the system was further equilibrated for 1 ns at 298.15 K and constant volume and subsequently for another ns at 298.15 K and 1 bar pressure. After these equilibration steps, the system was simulated for 100 ns at 298.15 K and 1 bar pressure. To investigate the influence of the crystallization liquid, we also performed a simulations with pure water, modeled as TIP3P water molecules.[61]

RESULTS AND DISCUSSION

Parameters. Table 1 lists the AMBER atom types for thiolate-protected gold nanoclusters, Lennard-Jones parameters σ and ϵ, and partial charges. The Lennard-Jones parameters for gold were obtained from ref 38, whereas all other atom types are standard types found in AMBER. Because the atom types determine the force constants and equilibrium values for bonded interactions, we only report the bonded parameters that we have introduced ourselves (Table 2). These force constants and equilibrium values describe interactions that involve known atom types but that were not yet included in the default AMBER force field.

For the clearly more open V-angle of the $Au_{102}(pMBA)_{44}$ cluster compared to similar staple shapes of the $Au_{25}(PET)_{18}$ and $Au_{38}(PET)_{18}$) clusters, we recommend using the larger angle and force constant to avoid strong structural deformations. For simulating new clusters with a known geometry of the staples, obtained from either experiment or theory, we suggest to take the value that is closest to the reference value. If, in contrast, the geometry is not known and the purpose of the force field computations is to determine the structure, we suggest choosing the smaller angle/force constant combination first and update these parameters if needed.

Force Field Validation. In all simulations, the cluster symmetries remained intact during 100 ns, suggesting the validity of using only a Lennard-Jones potential to describe the interactions between the gold atoms in the core of the cluster. Next, we inspected the conformational dynamics of the staples. In contrast to the dynamics of the gold atoms in the core, rather large fluctuations of the staple geometries were observed

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342–1350

Table 1. Nonbonded Parameters for All Atom Types Used in Simulations[a]

atom name	atom type	σ (nm)	ϵ (kJ/mol)	q (e)
Au	Au	0.2629	22.1330	0
*p*MBA				
C	C	0.339967	0.359824	0.6641 (0.7245)
CZ	CA	_ ″ _	_ ″ _	0.0689 (0.0827)
CD1	CA	_ ″ _	_ ″ _	−0.0557 (−0.0707)
CB1	CA	_ ″ _	_ ″ _	−0.1604 (−0.1985)
CA	CA	_ ″ _	_ ″ _	−0.0029 (0.0463)
CB2	CA	_ ″ _	_ ″ _	−0.1604 (−0.1985)
CD2	CA	_ ″ _	_ ″ _	−0.0557 (−0.0707)
HD1	HA	0.259964	0.06276	0.1191 (0.1000)
HB1	HA	_ ″ _	_ ″ _	0.1596 (0.1458)
HD2	HA	_ ″ _	_ ″ _	0.1191 (0.1000)
HB2	HA	_ ″ _	_ ″ _	0.1596 (0.1458)
HO	HO	0.000000	0.00000	0.4619 (−)
O2	O	0.295992	0.87864	−0.5237 (−0.7282)
OH	OH	0.306647	0.880314	−0.6417 (−0.7282)
S	S	0.356359	1.04600	−0.1518 (−0.3504)
PET				
CA1	CA	0.339967	0.359824	0.0298 (0.0267)
CA2	CA	_ ″ _	_ ″ _	−0.1582 (−0.1613)
CA3	CA	_ ″ _	_ ″ _	−0.1582 (−0.1613)
CA4	CA	_ ″ _	_ ″ _	−0.1799 (−0.1829)
CA5	CA	_ ″ _	_ ″ _	−0.1799 (−0.1829)
CA6	CA	_ ″ _	_ ″ _	−0.0932 (−0.0964)
CT1	CT	0.339967	0.45773	−0.0026 (−0.0057)
CT2	CT	_ ″ _	_ ″ _	−0.1101 (−0.1132)
HT3	H1	0.247135	0.0656888	0.0773 (0.0742)
HT4	H1	_ ″ _	_ ″ _	0.0773 (0.0742)
HA2	HA	0.259964	0.06276	0.1468 (0.1437)
HA3	HA	_ ″ _	_ ″ _	0.1468 (0.1437)
HA4	HA	_ ″ _	_ ″ _	0.1469 (0.1438)
HA5	HA	_ ″ _	_ ″ _	0.1469 (0.1438)
HA6	HA	_ ″ _	_ ″ _	0.1287 (0.1257)
HT1	HC	0.264953	0.065888	0.0903 (0.0872)
HT2	HC	_ ″ _	_ ″ _	0.0903 (0.0872)
S	S	0.356359	1.04600	−0.1990 (−0.2021)

[a]Partial charges for deprotonated *p*MBA and PET ligands to obtain the negatively charged $Au_{25}(PET)_{18}^-$ cluster are presented in parentheses.

Table 2. Bonded Parameters[a]

bond	k_b (kJ mol^{-1} nm^{-2})	d_{eq} (nm)
Au_{surf}–S	62730[35]	0.241
Au_{lig}–S	62730[35]	0.233
CA–S	198321.6*	0.175
CT–S	99113.0[35]	0.184
angle	k_θ (kJ mol^{-1} rad^{-2})	θ_{eq} (deg)
Au_{surf}–S–Au_{lig}	460.240[19]	91.3
S–Au_{lig}–S	460.240[19]	172.4
Au_{lig}–S–Au_{lig}	460.240/1460.240[19]	100.0/119.2
Au_{surf}–S–C	146.370[35]	111.6
Au_{lig}–S–C	146.370[35]	106.8
CA–C–OH	585.76*	112.0
CA–C–O	585.76*	126.0
CA–CA–S	418.40*	120.0
S–CT–HC	418.40*	107.0
H1–CT–CA	418.40*	109.0
dihedral	(kJ/mol)	ψ
X–X–CA–S	4.60240*	180.00*
C–CA–CA–CA	4.60240*	180.00*

[a]Parameters with known atom types but not by default included in AMBER are marked with ∗.

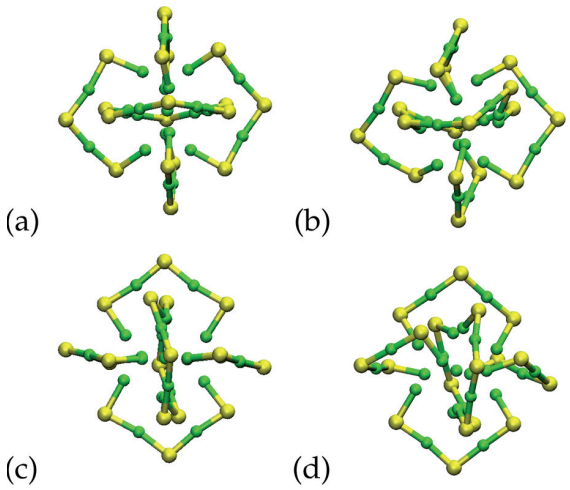


Figure 5. Staple distortion in the simulation of $Au_{25}(PET)_{18}^-$. Panels (a) and (c) show the original DFT geometry, and panels (b) and (d) show the structure averaged from the 100 ns MD simulation.

in our trajectories. In particular, the V-shaped staples of the $Au_{25}(PET)_{18}^-$ cluster deviate significantly from the DFT optimum geometry, as illustrated in Figure 5.

In Figure 5, it is shown that the staples exhibit certain folding or tilting from the top of the staple toward the cluster core. Similar folding was also seen in vacuum MD simulations (next section). Because this tilting or folding may not be physically realistic, we recommend using these parameters with the smallest $Au_{25}(PET)_{18}^-$ cluster with caution. Despite the folding, the cluster core retains its shape during the simulations, and the diffusion coefficient is in line with the experimental one (see below).

To verify whether the fluctuations in staple geometries are realistic, we performed DFT molecular dynamics simulations and compared the staple dynamics to that observed in the force field simulations.

MD/DFT Simulations. The dynamics and structural fluctuations of the $Au_{25}(PET)_{18}^-$ cluster were studied in a picosecond time scale in vacuum with DFT/MD simulations and compared to classical MD simulations. Structural fluctuations were studied by examining distributions of atomic distances and angles in different simulations, presented in Figures 6 and 7.

The distributions of distances and angles generally broaden with increasing simulation time due to staple deformation. The main differences between the DFT/MD and force field MD simulations are the averages of these distributions. However, as the averages depend strongly on the simulation time in the force field MD simulation, we believe that the observed differences between the Amber and DFT ensembles are due to insufficient sampling of the latter rather than inaccuracies of the force field. The overlap of the distributions furthermore suggest that our force field parameters permit sampling DFT geometries in classical MD simulations.

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

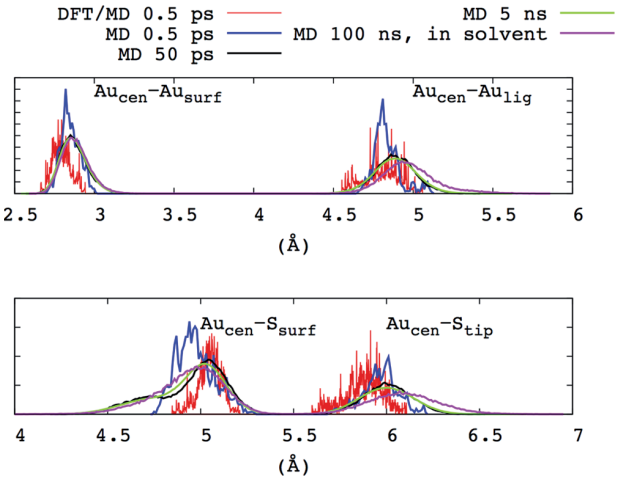


Figure 6. Distribution of atom distances from the central Au atom.

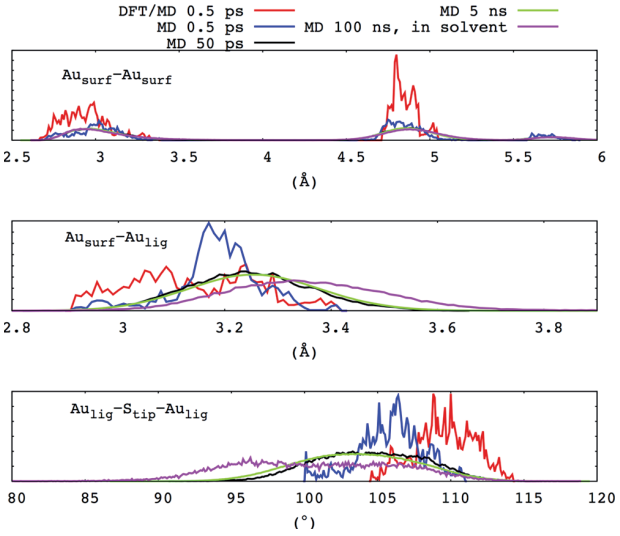


Figure 7. Bond and angle distributions.

Comparing the distance and angle distributions to longer DFT/MD simulation (10 ps) for $Au_{25}(SH)_{18}^-$ presented in ref 57, there appears to be no drastic differences in the observed distance and angle ranges. Especially, the ranges of the reported distances and angle appearing on the longer time scale MD simulations do not exceed those observed already in 10 ps DFT/MD. Mean square displacements of atom positions reported in ref 57 also qualitatively agree to those observed in this work for similar time scales. The peak positions or averages for bond distances qualitatively agree as well between MD and DFT/MD. The only clear difference appears again in the S_{surf}–S_{tip}–S_{surf} angle, which on average for MD appears smaller compared to DFT/MD.

In conclusion, the structural fluctuations in the smallest cluster do not drastically differ between force field MD simulations and DFT/MD simulations. However, the 0.5 ps simulation at DFT/MD was probably too short to capture the distortion or folding of the staples that was observed in the longer force field MD simulations. Therefore, at this point, it is difficult to say whether the distortions observed in the smallest cluster are realistic or not. Nevertheless, the range of angles and distances are in agreement with the DFT/MD simulations. Despite distortions of the staples, the force field maintains the overall shape of the cluster, suggesting that the intracluster interactions are sufficiently well described for the purposes of this work.

Ligand Flipping. Also changes in the chirality of sulfur atoms or flipping of the ligands over the staples were observed in the simulations. Examples of ligand flips in 100 ns simulations are presented in Figure 8. Flips occur when the sign of the dihedral

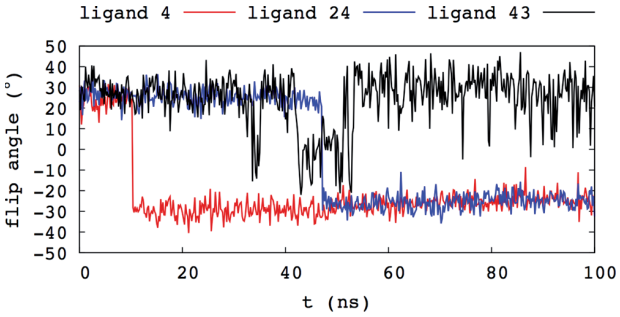


Figure 8. Ligand flips in 100 ns simulations for three example ligands in the $Au_{102}(p\text{MBA})_{44}$ cluster. The change of sign in the angle indicates flipping of the ligand.

angle C–Au–Au–S changes (Figure 8). Figure 8 shows the flip angle versus time for three example ligands of the $Au_{102}(p\text{MBA})_{44}$ cluster. Flips were observed in simulations for all the five clusters studied.

Figure 8 shows ligand flips for three example ligands of the $Au_{102}(p\text{MBA})_{44}$ cluster. Similar flips are also observed for other ligands and in simulations of the other four clusters as well. In order to validate whether the flips are realistic, we calculated DFT energy profiles associated with these flips.

A potential energy profile for flipping the central ligand in one isolated staple unit of $Au_{102}(p\text{MBA})_{44}$ and $Au_{25}(\text{PET})_{18}$ was computed with DFT. For comparison, we computed the MM energies along these pathways as well. Both profiles are shown in in Figure 9.

Constrained relaxation DFT predicts barriers of around 0.3 and 0.6 eV for the central ligand flipping in the Au_{102} and Au_{25} V-shaped staples, respectively. Corresponding barriers in MD are around 0.35 and 0.45 eV. The barriers in MD are therefore of the same order of magnitude compared to DFT. Due to these relatively low barriers, flipping at room temperature is anticipated for all these clusters and was indeed observed in the MD simulations.

Diffusion Coefficients. Diffusion coefficients of three clusters for which experimental data is available[62] are listed in Table 3 for various box sizes. An error-weighted linear fit of the diffusion coefficient as a function of $1/L$ was extrapolated to infinite box size to correct for finite size effects. The extrapolated diffusion coefficients are listed in bold. Experimental results taken from ref 62 are presented for comparison.

The experimentally observed trend that the diffusion coefficient increases with decreasing cluster size is also observed in our simulations. In general, the calculated diffusion coefficients are smaller than the experimental ones. Nevertheless, the diffusion coefficients with the error estimates from the simulations are within the error estimates of the experimental diffusion coefficients.

The discrepancies between calculated and experimental diffusion coefficients may have various origins in addition to

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342–1350

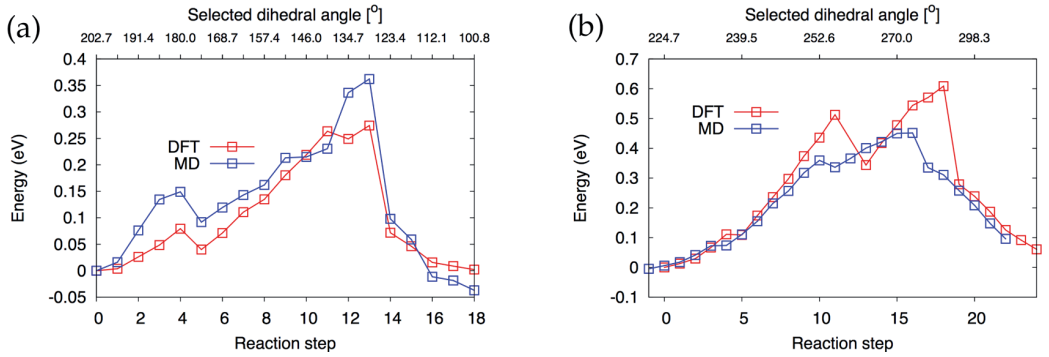


Figure 9. Ligand flipping potentials in isolated unit structures from clusters (a) Au_{102} and (b) Au_{25}.

Table 3. Diffusion Coefficients Calculated from Einstein Relation between 1 and 9 ns and Averaged over 20 Individual 10 ns Simulations ± Standard Deviation[a]

system	L (nm)	$D_{cluster}$ (10^{-5} cm_2/s)	$D_{experimental}$ (10^{-5} cm^2/s)
$Au_{25}(PET)_{18}$	2.27	0.014 ± 0.009	
	3.23	0.077 ± 0.079	
	4.13	0.177 ± 0.126	
	5.12	0.333 ± 0.449	
	6.06	0.373 ± 0.306	
	7.17	0.686 ± 0.595	
	∞	**0.373 ± 0.09**	**0.484 ± 0.04**
$Au_{38}(PET)_{24}$	2.55	0.021 ± 0.011	
	3.45	0.059 ± 0.045	
	4.23	0.144 ± 0.125	
	5.23	0.270 ± 0.266	
	6.36	0.233 ± 0.221	
	7.38	0.377 ± 0.308	
	∞	**0.274 ± 0.07**	**0.374 ± 0.03**
$Au_{144}(PET)_{60}$	3.44	0.012 ± 0.010	
	4.71	0.065 ± 0.060	
	5.83	0.118 ± 0.110	
	7.17	0.131 ± 0.085	
	8.34	0.127 ± 0.136	
	9.62	0.355 ± 0.264	
	∞	**0.251 ± 0.04**	**0.356 ± 0.03**

[a]Final result for infinite box size estimated from the intercept of the weighted 6 point linear fit to results for different box lengths L. Experimental diffusion coefficients[62] are presented for comparison ± the reported experimental standard deviation at maximum of 8%.

inaccuracies in the force field. First, the convergence of the calculated diffusion coefficients is rather poor. Second, the diffusion coefficient is strongly dependent on solvent viscosity. Viscosity in turn can depend on the deuteration of the solvent. In the experiments, deuterated solvents ($CDCl_3$ and CD_2Cl_2) were used, whereas we used parameters for the protonated solvents ($CHCl_3$ and CH_2Cl_2).

Despite these differences, the overall agreement between the diffusion of the cluster in the simulations and experiment is encouraging and suggests the relative hydrodynamic radii of the clusters are in line with experiment. As the hydrodynamic radius depends strongly on the fluctuations of the staples, we can also consider these fluctuations to be realistic and part of the normal thermal fluctuations of the gold clusters in solution.

Crystal Simulation. As a final test for the validity of our force field parameters, we performed a 100 ns MD simulation of a crystal of $Au_{102}(p\text{MBA})_{44}$, for which the experimental X-ray structure is available.[60] We performed simulations of 72 asymmetric units in a periodic unit cell. The empty cavities between the clusters were filled with either pure water or a mixture of water, methanol, NaCl, and NaOAc, which closely resembles the solvent used during crystallization.[60] Whereas in simulations with water as a crystal solvent, the agreement with the experiment quickly deteriorates (Supporting Information); the crystal is stable when the mixture of solvents is used (Figure 10). The small structural deviations between the experimental structure and the averaged simulated structure shown in Figure 10 could be due to using a single protonation state of the *p*MBA ligands, which affects intermolecular hydrogen bonding. Alternatively, the solvent may have not been fully equilibrated

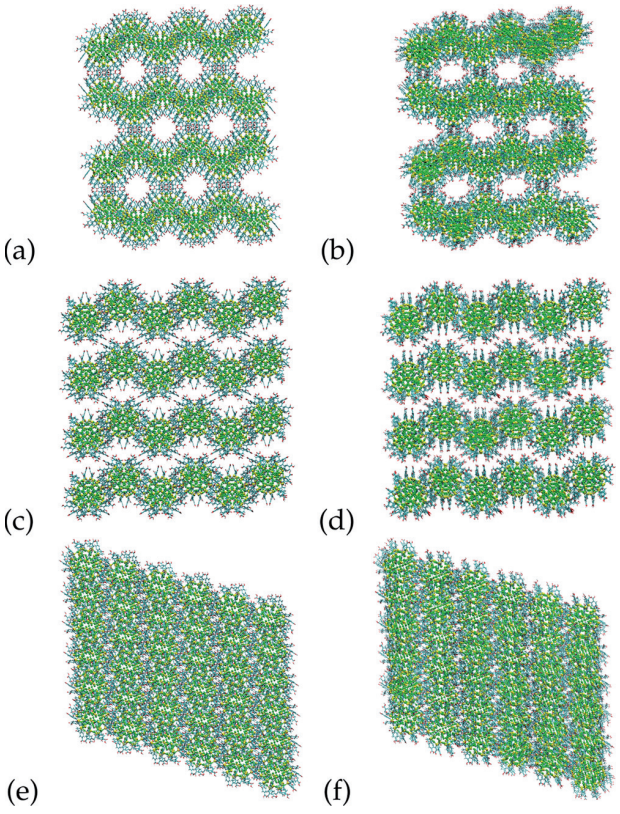


Figure 10. Simulation of 72 cluster crystal of $Au_{102}(p\text{MBA})_{44}$. Panels (a), (c), and (e) show the experimental crystal structure, and panels (b), (d), and (f) show the averaged structure from the 100 ns production simulation.

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

during the simulation. Nevertheless, the overall agreement with the X-ray data further supports the view that our force field is sufficiently accurate for classical MD simulations, and we are currently performing MD simulations of various systems containing gold clusters.

In the Supporting Information, we provide the force field files with our parameters for the Gromacs simulation program. The contents of these files are easily converted to other programs that support AMBER force fields.

CONCLUSION AND OUTLOOK

We have optimized AMBER force field parameters for thiolate-proteced gold clusters. For transferability, we have optimized one parameter set for a wide range of clusters, rather than multiple sets for specific clusters. Whereas the latter strategy may provide the most accurate description of the cluster that has been parameterized, the parameterization would need to be repeated for every new cluster considered, which can be time consuming. Also, geometries must be available beforehand, which precludes the use of the force field to predict the geometry of new clusters that are too large for a DFT treatment. Instead, we have optimized the parameters by considering five representative thiolate-protected gold clusters simultaneously. Furthermore, we use a noncovalent Lennard-Jones potential to describe the interactions between gold atoms in the core, which is essential for a transferable force field.

We have validated our force field parameters by comparing the results of classical MD simulations of gold clusters to the results of both DFT calculations and the available experimental data. The energetics and structural fluctuations show nice agreement to DFT data, and also experimentally determined diffusion constants and crystal structures could be reproduced with sufficient accuracy in MD simulations.

Because AMBER is a force field with parameters for a wide variety of biological compounds, including nucleic acids, proteins, sugars, and phospholipids, our work opens the possibility to systematically investigate the effect of thiolate-protected gold nanoclusters on the structure and stability of biomolecules. This is particularly relevant now that gold clusters are increasingly being used as contrast agents or molecular rulers in bioimaging applications. Understanding the interactions between biological matter and gold clusters will also be essential for designing new strategies to conjugate these clusters with biomolecules. Atomistic simulations can provide these insights and will therefore play an important role in driving this field forward.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b01053.

Additional data on MD simulation of the Au_{102} crystal in water. (PDF)

Description of the parameter files of the Gromacs force field. (ZIP)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: gerrit.x.groenhof@jyu.fi (G.G.).

*E-mail:hannu.j.hakkinen@jyu.fi (H.H.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The computations were done at the Nanoscience Center of University of Jyväskylä and at the CSC − the Finnish IT Center for Science. This work was funded by the Academy of Finland and the NGS-NANO graduate school.

REFERENCES

(1) Häkkinen, H. *Nat. Chem.* **2012**, *4*, 443−455.

(2) Tsukuda, T. *Bull. Chem. Soc. Jpn.* **2012**, *85*, 151−168.

(3) Sardar, R.; Funston, A. M.; Mulvaney, P.; Murray, R. W. *Langmuir* **2009**, *25*, 13840−13851.

(4) Jin, R. *Nanoscale* **2010**, *2*, 343−362.

(5) Walter, M.; Akola, J.; Lopez-Acevedo, O.; Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Whetten, R. L.; Grönbeck, H.; Häkkinen, H. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9157−9162.

(6) Lopez-Acevedo, O.; Clayborne, P. A.; Häkkinen, H. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *84*, 035434.

(7) Cui, M.; Zhao, Y.; Song, Q. *TrAC, Trends Anal. Chem.* **2014**, *57*, 73−82.

(8) Lopez-Acevedo, O.; Kacprzak, K. A.; Akola, J.; Häkkinen, H. *Nat. Chem.* **2010**, *2*, 329−334.

(9) Chen, X.; Häkkinen, H. *J. Am. Chem. Soc.* **2013**, *135*, 12944−12947.

(10) Marjomäki, V.; Lahtinen, T.; Martikainen, M.; Koivisto, J.; Malola, S.; Salorinne, K.; Pettersson, M.; Häkkinen, H. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 1277−1281.

(11) Mathew-Fenn, R. S.; Das, R.; Silverman, J. A.; Walker, P. A.; Harbury, P. A. B. *PLoS One* **2008**, *3*, e3229.

(12) Reinhard, B. M.; Siu, M.; Agarwal, H.; Alivisatos, A. P.; Liphardt, J. *Nano Lett.* **2005**, *5*, 2246−2252.

(13) Verde, A. V.; Acres, J. M.; Maranas, J. K. *Biomacromolecules* **2009**, *10*, 2118−2128.

(14) Kokh, D. B.; Corni, S.; Winn, P. J.; Hoefling, M.; Gottschalk, K. E.; Wade, R. C. *J. Chem. Theory Comput.* **2010**, *6*, 1753−1768.

(15) Alsharif, S. A.; Chen, L. Y.; Tlahuice-Flores, A.; Whetten, R. L.; Yacaman, M. J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 3909−3913.

(16) Van Lehn, R. C.; Alexander-Katz, A. *J. Phys. Chem. C* **2013**, *117*, 20104−20115.

(17) Van Lehn, R. C.; Alexander-Katz, A. *J. Phys. Chem. B* **2014**, *118*, 12586−12598.

(18) Van Lehn, R. C.; Ricci, M.; Silva, P. H. J.; Andreozzi, P.; Reguera, J.; Voïtchovsky, K.; Stellacci, F.; Alexander-Katz, A. *Nat. Commun.* **2014**, *5*, 4482.

(19) Heikkilä, E.; Gurtovenko, A. A.; Martinez-Seara, H.; Häkkinen, H.; Vattulainen, I.; Akola, J. *J. Phys. Chem. C* **2012**, *116*, 9805−9815.

(20) Heikkilä, E.; Martinez-Seara, H.; Gurtovenko, A. A.; Vattulainen, I.; Akola, J. *Biochim. Biophys. Acta, Biomembr.* **2014**, *1838*, 2852−2860.

(21) Krupa, P.; Mozolewska, M. A.; Rasulev, B.; Czaplewski, C.; Leszczynski, J. *Task Quarterly* **2014**, *18*, 337−341.

(22) Lee, K. H.; Ytreberg, F. M. *Entropy* **2012**, *14*, 630−641.

(23) Lal, M.; Plummer, M.; Purton, J.; Smith, W. *Proc. R. Soc. London, Ser. A* **2011**, *467*, 1986−2003.

(24) Antúnez-García, J.; Mejía-Rosales, S.; Pérez-Tijerina, E.; Montejano-Carrizales, J. M.; José-Yacamán, M. *Materials* **2011**, *4*, 368−379.

(25) Böyükata, M. *Phys. E (Amsterdam, Neth.)* **2006**, *33*, 182−190.

(26) Erkoç, Ş. *Phys. E (Amsterdam, Neth.)* **2000**, *8*, 210−218.

(27) Mariscal, M. M.; Olmos-Asar, J. A.; Gutierrez-Wing, C.; Mayoral, A.; Yacaman, M. J. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11785−11790.

(28) Legenski, N.; Zhou, C.; Zhang, Q.; Han, B.; Wu, J.; Chen, L.; Cheng, H.; Forrey, R. C. *J. Comput. Chem.* **2011**, *32*, 1711−1720.

(29) Brancolini, G.; Toroz, D.; Corni, S. *Nanoscale* **2014**, *6*, 7903−7911.

(30) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *J. Phys. Chem. A* **2001**, *105*, 9396−9409.

(31) Keith, J. A.; Fantauzzi, D.; Jacob, T.; van Duin, A. C. T. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *81*, 235404.

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

(32) Järvi, T. T.; van Duin, A. C. T.; Nordlund, K.; Goddard, W. A. *J. Phys. Chem. A* **2011**, *115*, 10315−10322.
(33) Bae, G.-T.; Aikens, C. M. *J. Phys. Chem. A* **2013**, *117*, 10438−10446.
(34) Groenhof, G. *Methods Mol. Biol.* **2013**, *926*, 43−66.
(35) Banerjee, S.; Montgomery, J.; Gascón, J. *J. Mater. Sci.* **2012**, *47*, 7686−7692.
(36) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(37) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(38) Heinz, H.; Vaia, R. A.; Farmer, B. L.; Naik, R. R. *J. Phys. Chem. C* **2008**, *112*, 17281−17290.
(39) Ramezani, F.; Habibi, M.; Rafii-Tabar, H.; Amanlou, M. *DARU* **2015**, *23*, 1−5.
(40) Malola, S.; Häkkinen, H. *J. Phys. Chem. Lett.* **2011**, *2*, 2316−2321.
(41) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California, San Francisco, 2012.
(42) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(43) Leng, Y.; Krstić, P. S.; Wells, J. C.; Cummings, P. T.; Dean, D. J. *J. Chem. Phys.* **2005**, *122*, 244721.
(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, P. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01 ; Gaussian, Inc.: Wallingford, CT, 2009.
(45) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057−1065.
(46) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129−145.
(47) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.
(48) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(49) Bussi, G.; Parrinello, M. *Comput. Phys. Commun.* **2008**, *179*, 26−29.
(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(51) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(52) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(53) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. *J. Chem. Theory Comput.* **2012**, *8*, 61−74.
(54) van der Spoel, D.; van Maaren, P. J.; Caleman, C. *Bioinformatics* **2012**, *28*, 752−753.
(55) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(56) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dułak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; Kristoffersen, H. H.; Kuisma, M.; Larsen, A. H.; Lehtovaara, L.; Ljungberg, M.; Lopez-Acevedo, O.; Moses, P. G.; Ojanen, J.; Olsen, T.; Petzold, V.; Romero, N. A.; Stausholm-Møller, J.; Strange, M.; Tritsaris, G. A.; Vanin, M.; Walter, M.; Hammer, B.; Häkkinen, H.; Madsen, G. K. H.; Nieminen, R. M.; Nørskov, J. K.; Puska, M.; Rantala, T. T.; Schiøtz, J.; Thygesen, K. S.; Jacobsen, K. W. *J. Phys.: Condens. Matter* **2010**, *22*, 253202.
(57) Mäkinen, V.; Häkkinen, H. *Eur. Phys. J. D* **2012**, *66*, 310.
(58) Yeh, I.-C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873−15879.
(59) Wang, J.; Hou, T. *J. Comput. Chem.* **2011**, *32*, 3505−3519.
(60) Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Bushnell, D. A.; Kornberg, R. D. *Science* **2007**, *318*, 430−433.
(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(62) Salorinne, K.; Lahtinen, T.; Koivisto, J.; Kalenius, E.; Nissinen, M.; Pettersson, M.; Häkkinen, H. *Anal. Chem.* **2013**, *85*, 3489−3492.

DOI: 10.1021/acs.jctc.5b01053
J. Chem. Theory Comput. 2016, 12, 1342−1350

II

DYNAMIC STABILIZATION OF THE LIGAND-METAL INTERFACE IN ATOMICALLY PRECISE GOLD NANOCLUSTERS AU_{68} AND AU_{144} PROTECTED BY META-MERCAPTOBENZOIC ACID

Reprinted with permission from

Tiia-Riikka Tero, Sami Malola, Benedek Koncz, Emmi Pohjolainen, Saara Lautala, Satu Mustalahti, Perttu Permi, Gerrit Groenhof, Mika Pettersson, Hannu Häkkinen. *ACS Nano*, 11 (12), pp. 11872–11879, **2017**.

ACSNANO

www.acsnano.org

ARTICLE

Dynamic Stabilization of the Ligand–Metal Interface in Atomically Precise Gold Nanoclusters Au_{68} and Au_{144} Protected by *meta*-Mercaptobenzoic Acid

Tiia-Riikka Tero,[†,#] Sami Malola,[‡,#] Benedek Koncz,[†,#] Emmi Pohjolainen,[‡,#] Saara Lautala,[‡] Satu Mustalahti,[†] Perttu Permi,[†,§] Gerrit Groenhof,[†] Mika Pettersson,[†] and Hannu Häkkinen*[,†,‡]

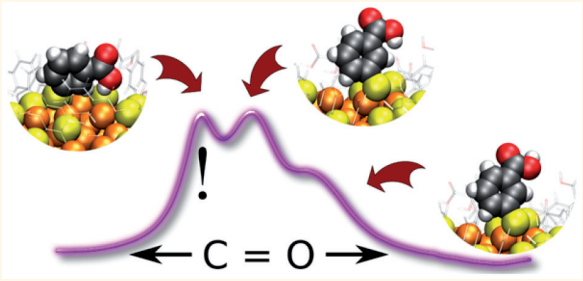
[†]Department of Chemistry, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland

[‡]Department of Physics, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland

[§]Department of Biology and Environmental Science, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland

Supporting Information

ABSTRACT: Ligand-stabilized, atomically precise gold nanoclusters with a metal core of a uniform size of just 1–3 nm constitute an interesting class of nanomaterials with versatile possibilities for applications due to their size-dependent properties and modifiable ligand layers. The key to extending the usability of the clusters in applications is to understand the chemical bonding in the ligand layer as a function of cluster size and ligand structure. Previously, it has been shown that monodispersed gold nanoclusters, stabilized by *meta*-mercaptobenzoic acid (*m*-MBA or 3-MBA) ligands and with sizes of 68–144 gold atoms, show ambient stability. Here we show that a combination of nuclear magnetic resonance spectroscopy, UV–vis absorption, infrared spectroscopy, molecular dynamics simulations, and density functional theory calculations reveals a distinct chemistry in the ligand layer, absent in other known thiol-stabilized gold nanoclusters. Our results imply a low-symmetry C_1 ligand layer of 3-MBA around the gold core of Au_{68} and Au_{144} and suggest that 3-MBA protects the metal core not only by the covalent S–Au bond formation but also *via* weak π–Au and O=C–OH···Au interactions. The π–Au and –OH···Au interactions have a strength of the order of a hydrogen bond and thus are dynamic in water at ambient temperature. The –OH···Au interaction was identified by a distinct carbonyl stretch frequency that is distinct for 3-MBA-protected gold clusters, but is missing in the previously studied $Au_{102}(p\text{-MBA})_{44}$ cluster. These thiol–gold interactions can be used to explain a remarkably low ligand density on the surface of the metal core of these clusters. Our results lay a foundation to understand functionalization of atomically precise ligand-stabilized gold nanoclusters *via* a route where weak ligand–metal interfacial interactions are sacrificed for covalent bonding.



KEYWORDS: *gold nanocluster, thiols, molecular dynamics, NMR spectroscopy, vibrational spectroscopy, carboxylic acid*

Manipulation of gold nanoparticles with atomic and molecular precision, in forms of 1–3 nm atomically precise clusters covered by a monolayer of organic molecules, has been made possible over the past decade by breakthroughs in synthetic chemistry and characterization techniques.[1,2] Properties of these versatile building blocks of nanomaterials can be tuned by controlling the size of the metal core and the chemical nature of the protecting molecular layer. Water-soluble clusters are relevant in life science and nanomedicine, for instance, for imaging, detection, and theranostics. Since the clusters interact with the environment *via* the protecting ligand layer, characterizing the structure and dynamics of that layer in solution phase is key to understand the interaction to the environment.[3] This is an important challenge to be solved for engineering practical applications.

Among several water-soluble thiols that can stabilize nanometer-scale gold clusters,[4] *meta*-mercaptobenzoic acid (3-MBA) was recently shown to produce three distinct-size clusters that are readily reactive to thiol-modified DNA and proteins with an exposed cysteine group.[5] The atomic structure of the gold core of the smallest cluster, Au_{68}, was previously

Received: November 2, 2017
Accepted: November 14, 2017
Published: November 14, 2017

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872–11879

solved *via* single-particle 3D reconstruction of transmission electron microscopy (TEM) data.[6] The companion paper of the current work describes the atomic structure of the gold core of the largest cluster Au_{144} (ref 7). Here, we concentrate on the structure and dynamics of the protecting ligand layer of Au_{68} and Au_{144} clusters in solution and demonstrate dynamic interactions at the ligand–metal interface that contribute both to chemical protection of the gold core and opening active metal sites for conjugation reactions. These interactions are suggested from molecular dynamics (MD) simulations and density functional theory computations (DFT) and supported by infrared (IR) spectroscopy data. This observation can explain the documented reactivity of 3-MBA-protected gold nanoclusters and demonstrates a potential route for bioconjugation where covalent bonds to gold can irreversibly be added at the ligand–metal interface. The existence of gold sites at the ligand–metal interface, that are protected by weak dynamic interactions and may be available, for example, for catalytic reactions at ambient conditions in water, creates an unexplored paradigm for the ligand–metal chemistry at the thiolate–gold interface.[8]

RESULTS

3-MBA-protected Au_{144} and Au_{68} clusters were produced from a synthesis described earlier by using thiol-to-gold ratios of 7:1 and 2:1, respectively.[5] We characterized the samples by nuclear magnetic resonance (NMR) spectroscopy and diffusion ordered spectroscopy (DOSY-NMR) as well as UV–vis and infrared (IR) spectroscopies.

The ligand count of the $Au_{68}(3\text{-MBA})_x$ cluster was previously determined to be $x \sim 32$.[6] ESI-MS, XPS, and TGA analyses reported in the companion paper indicate that the ligand count x in $Au_{144}(3\text{-MBA})_x$ may be as low as 40.[7] This is in stark contrast to known organo-soluble thiol-protected gold nanoclusters of a similar size, such as the $Au_{144}(SCH_2CH_2Ph)_{60}$.[9,10] Our ^{1}H NMR data of $Au_{144}(3\text{-MBA})_{\sim 40}$ in D_2O showed one broad feature without resonance fine structure in the aromatic region (5–9.5 ppm, Figure 1 and Figure S1, Supporting Information), and no significant correlations were seen in 2D spectra (Figures S2 and S3). This differs qualitatively from the previously published ^{1}H NMR data of $Au_{144}(SCH_2CH_2Ph)_{60}$[11] and $Au_{102}(p\text{-MBA})_{44}$[3,12] as well as from $Au_{68}(3\text{-MBA})_{32}$,

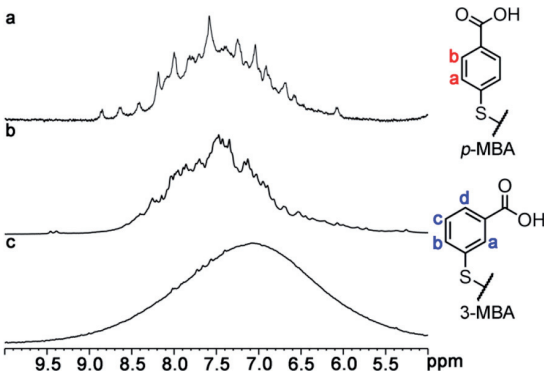


Figure 1. One-dimensional ^{1}H NMR spectra. (a) Aromatic region of ^{1}H NMR spectrum of $Au_{102}(p\text{-MBA})_{44}$ in D_2O-NaOH. (b) The same for $Au_{68}(3\text{-MBA})_{30-32}$ and (c) $Au_{144}(3\text{-MBA})_n$ in D_2O. Chemical structures of *p*- and 3-mercaptobenzoic acid ligands with proton labels are shown on the right. The data for (a) is taken from ref 12.

measured in this work (Figure 1 and Figure S4), where a large number of resonance signals are visible in the 1D spectrum and their 2D correlations give detailed information on the local symmetry environments in the ligand layer. Previously, we have been able to make a successful assignment of the signals to the specific ligands in the protecting layer of $Au_{102}(p\text{-MBA})_{44}$.[12]

UV–vis absorption of $Au_{144}(3\text{-MBA})_{\sim 40}$ showed a weak plasmonic shoulder around 500 nm (Figure 2), which is absent

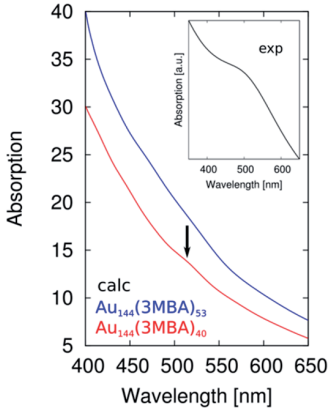


Figure 2. UV–vis absorption spectra. Theoretical spectra obtained from LR-TDDFT computations for $Au_{144}(3\text{-MBA})_n$ models for x = 40 (red) and x = 53 (blue), compared to the experimental spectrum of $Au_{144}(3\text{-MBA})_{\sim 40}$ (inset). The arrow shows a weak plasmonic band at 510 nm computed for $Au_{144}(3\text{-MBA})_{40}$, agreeing qualitatively well with the band around 500 nm in the experiment.

in organo-soluble $Au_{144}(SCH_2CH_2Ph)_{60}$.[9,13] This points to unexpected, significant differences in the response of the metal electrons to light in these two systems that have exactly the same number of gold atoms.

The experimental results described above show unambiguously that the 3-MBA ligand induces very different properties of $Au_{144}(3\text{-MBA})_{\sim 40}$ from the ones of other gold nanoclusters of similar size. Particularly, the location of the acid group at the 3-position in the aromatic ring (Figure 1) creates a possibility for weak interactions within the ligand layer that are absent in *p*-MBA protected clusters or in organo-soluble clusters. This motivated us to consider computational models (Figure 3) to study the structure and dynamics of the 3-MBA ligand layer of Au_{144} and Au_{68}. We built hundreds of structural models (Supporting Information, Section B) of $Au_{68}(3\text{-MBA})_{32}$ and $Au_{144}(3\text{-MBA})_x$ with x = 40, 42, 44, 47, 49, 51, 53, and simulated the dynamics of the ligand layer in water and in methanol, with varying degrees of protonation between 0 and 50%, at 300 K by using GROMACS MD software.[14] The coordinates of the gold atoms in the metal cores of the models for $Au_{68}(3\text{-MBA})_{32}$ and $Au_{144}(3\text{-MBA})_x$ were taken from refs 6 and 7, respectively, in addition, the coordinates of the metal core of $Au_{144}(3\text{-MBA})_x$ are given in Section H, Supporting Information.

Visual inspection of MD trajectories revealed several weak interactions in the ligand layer and at the ligand–gold interface, such as formation of interligand hydrogen bonds, interligand π stacking (aromatic contacts), π–Au interaction where the aromatic ring lies "flat" on the gold core, and hydrogen-bonding-like O=C–OH···Au interaction when the hydroxyl group is rotated toward the gold core. This interaction was

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872–11879

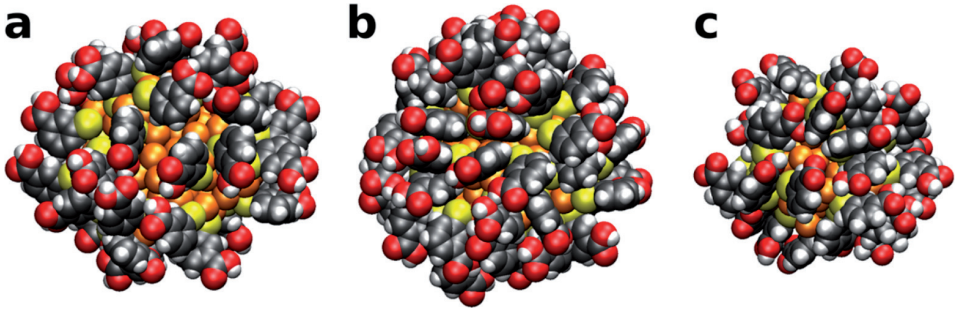


Figure 3. Structural models for 3-MBA-protected Au_{68} and Au_{144} clusters. (a) $Au_{144}(3\text{-MBA})_{40}$, (b) $Au_{144}(3\text{-MBA})_{53}$, (c) $Au_{68}(3\text{-MBA})_{32}$. Au, golden; S, yellow; C, dark gray; O, red; H, white.

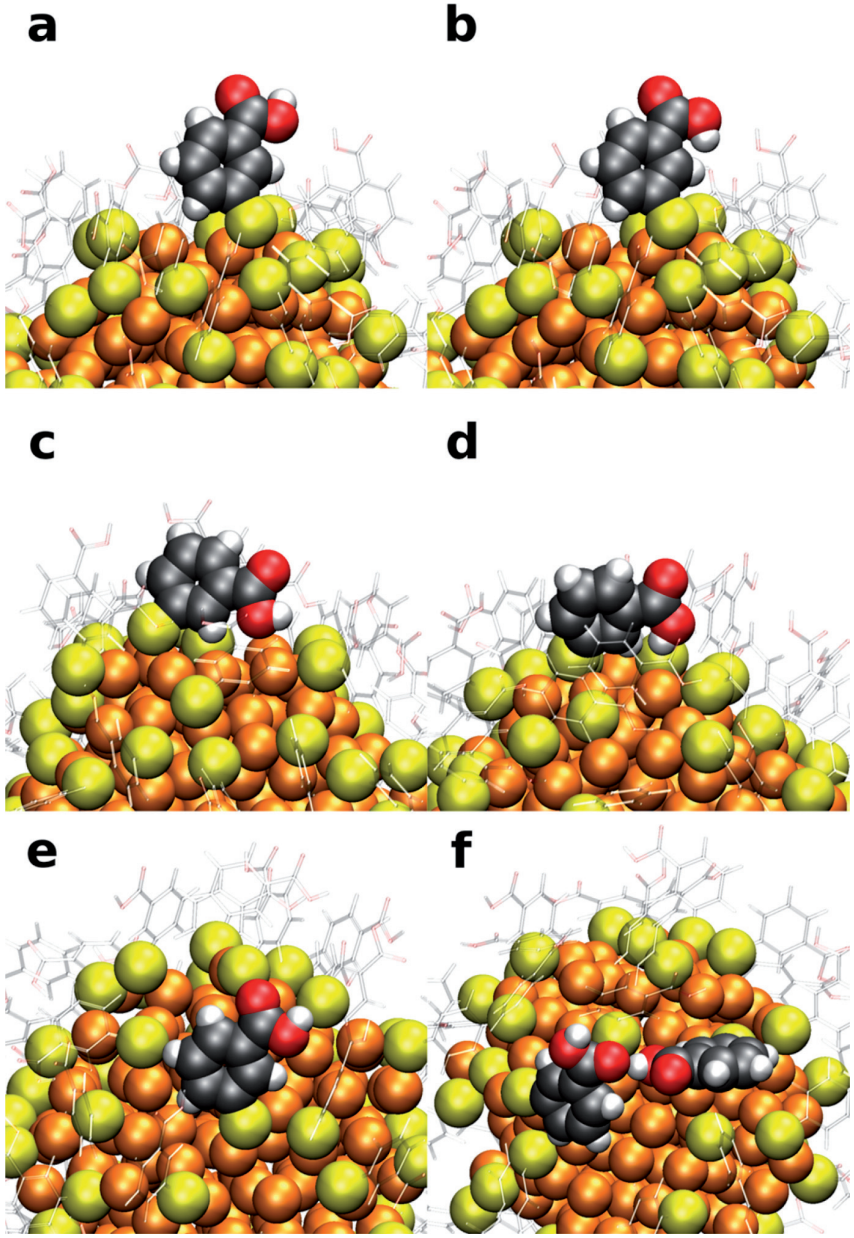


Figure 4. Interligand and ligand−metal interactions. Various conformations of 3-MBA and weak interactions in the ligand layer, found in the MD simulations. (a) *syn* pointing out, (b) *anti* pointing out, (c) *syn* close to gold core, (d) *anti* close to gold core, (e) *syn* lying flat on the core, and (f) interligand hydrogen bond. DFT calculations show that structures in a, c, and d are practically isoenergetic (within 25 meV).

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

dynamic, that is, was continuously formed and broken during the course of the runs. Figure 4 visualizes snapshots of the types of the weak interactions in the ligand layer. Statistical analysis of weak interactions over 50 ns trajectories is shown in Table S6.

Although two 3-MBA ligands can in principle form a hydrogen-bond pair, the analysis revealed only 0 to 1.9 hydrogen bonds (Figure 4f) in the ligand layer for $Au_{68}(3\text{-MBA})_{32}$ and 0.3 to 5.4 bonds in the ligand layer for $Au_{144}(3\text{-MBA})_{40-53}$. The higher deprotonation state was found to promote the higher number of bonds. The low number of the interligand hydrogen bonds could be explained by hydrogen bonding with the solvent molecules and by a rather strong tendency for a competing interaction, namely interligand π stacking. On average, we found 1.4−1.5 aromatic contacts per ligand in $Au_{68}(3\text{-MBA})_{32}$ and 1.3−1.8 contacts per ligand in $Au_{144}(3\text{-MBA})_{40-53}$, with the higher number of contacts in clusters with the higher number of ligands. These numbers indicate a good degree of ligand packing, taking into account that the maximum is two contacts per ligand if all the aromatic rings in the ligand layer would be "sandwiched" between two neighboring ligands.

The analysis revealed also several other interesting and unexpected interactions at the ligand−metal interface (see Table S6 and Figure S4). It is well-known that carboxylic acids can adopt two spatial arrangements in the preferable planar geometry. In these conformations, the hydroxyl group is pointing either parallel or antiparallel to the carbonyl group (*syn* and *anti* conformations, respectively). In the absence of any strong external interactions, *syn* conformer (Figure 4a) has the lower energy and usually it will be the dominating species at room temperature in the gas phase, however, explicit solvent−solute interactions can reduce the *syn*/*anti* energy difference dramatically in solution. Our MD simulations showed that 3-MBA molecules in the ligand layer frequently transformed to *anti* conformation (Figure 4b) as well as to distinct *syn*/*anti* conformations where the OH group was rotated toward a close-lying Au atom at the metal core surface (Figure 4c,d). This gave rather close OH···Au interactions of 0.23−0.27 nm. DFT computations based on locally relaxed structures from MD simulations showed that configurations in Figure 4a,c,d are in practice isoenergetic within 0.025 eV (which is less than the inherent DFT inaccuracy of determining energies) and about 0.28 eV more favorable than the *anti* configuration in Figure 4b (note, however, that no explicit solvent was considered in the DFT calculations, thus the potential stabilization effect of hydrogen bonding to solvent for the configuration shown in Figure 4b is neglected). In all MD runs we also observed π−Au interactions where the aromatic ring was "lying down" on the core surface (Figure 4e). DFT computations showed that the π−Au interaction stabilizes the "lying down" 3-MBA ligand by 0.15 eV as compared to a "standing up" configuration.

To seek experimental signals of these weak interactions, we turned to vibrational spectroscopy (Figure 5). It is well-known that the carbonyl stretch frequency is a sensitive measure for the local environment of the O═C−OH moiety. IR data measured for $Au_{144}(3\text{-MBA})_{\sim 40}$ in deuterated methanol (MeOD) showed three peaks in the carbonyl region at 1695, 1714, and 1731 cm^{-1}, while the free 3-MBA and previously studied $Au_{102}(p\text{-MBA})_{44}$ cluster showed only two peaks. These two peaks belong to the *syn* and *anti* conformers of solvated molecules as explained in the Supporting Information (section C and Figure S9). For $Au_{68}(3\text{-MBA})_{32}$ we observed the same three peaks as for the Au_{144} cluster at 1692, 1711, and 1732

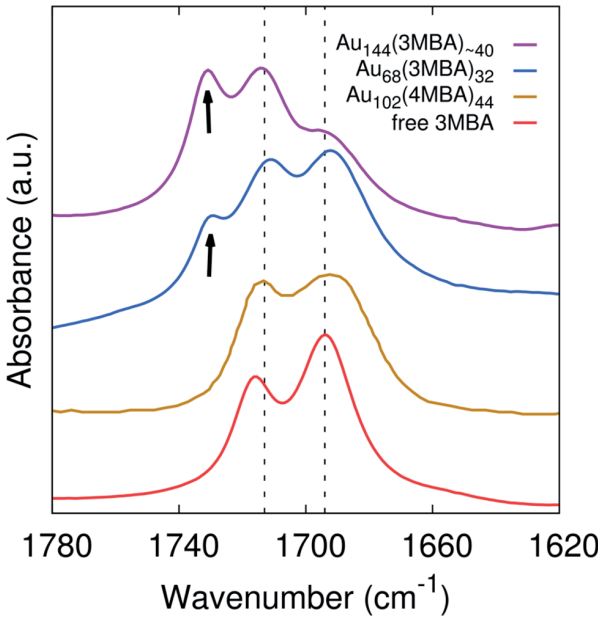


Figure 5. IR spectra of gold clusters. $Au_{144}(3\text{-MBA})_{\sim 40}$ (purple), $Au_{68}(3\text{-MBA})_{32}$ (blue), $Au_{102}(p\text{-MBA})_{44}$ (brown), free 3-MBA (red). The result for $Au_{102}(p\text{-MBA})_{44}$ is taken from ref 15. The arrow indicates the distinct peak around 1730 cm^{-1}, observable only in 3-MBA-passivated clusters, and interpreted as the signal of the O═C−OH···Au interaction. All spectra were measured in MeOD.

cm^{-1}. This result implied three distinct configurations of the O═C−OH group in the ligand layer. The two lowest frequencies matched with the signals from the free 3-MBA ligand and with the previously studied $Au_{102}(p\text{-MBA})_{44}$ cluster in MeOD[15] (Figure 5 and Table 1). Temperature dependence

Table 1. IR-Active C═O Modes[a]

system	*syn*	*anti*	*anti*/interface
Exp 3-MBA in CCl_4	1744	–	–
Exp 3-MBA in MeOD	1693	1715	–
Exp $Au_{144}(3\text{-MBA})_{\sim 40}$	1695	1714	1731
Exp $Au_{68}(3\text{-MBA})_{32}$	1692	1711	1732
Calc free 3-MBA	1727	1748	–
Calc $Au_{144}(3\text{-MBA})_{53}$	–	1754	1761

[a]Exp = measured. Calc = calculated. Unit is cm^{-1}. "Anti/interface" refers to the −OH···Au interaction shown in Figure 4d.

of the data showed that the relative intensities of the peaks changed with the temperature, indicating dynamic equilibrium between them (Figures S10 and S11). Repeating the same measurements for $Au_{68}(3\text{-MBA})_{32}$ in water under various pH conditions, we observed a broad carbonyl band centered around 1685 cm^{-1} at lower pH values where most of the ligands are still protonated, but could not resolve individual features within the band (Figure S12).

We were able to assign the three IR signals measured in MeOD (Figure 5) with help of DFT computations. The computed C═O stretch frequencies of *syn* and *anti* forms of a free 3-MBA are 1727 and 1748 cm^{-1}, respectively. For the free 3-MBA, we observed frequencies of 1693 and 1715 cm^{-1} in MeOD at room temperature, yielding the *anti*−*syn* difference as 22 cm^{-1}. This matched perfectly the computed *anti*−*syn*

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

frequency difference of 21 cm^{-1} and allowed us to assign the peaks at 1695 cm^{-1} ($Au_{144}(3\text{-MBA})_{\sim 40}$) and at 1692 cm^{-1} ($Au_{68}(3\text{-MBA})_{32}$) to the *syn* form (Figure 4a) and peaks at 1714 cm^{-1} ($Au_{144}(3\text{-MBA})_{\sim 40}$) and at 1711 cm^{-1} ($Au_{68}(3\text{-MBA})_{32}$) to the *anti* form (Figure 4b).

Interestingly, DFT computations done on a locally relaxed configuration of the O═C−OH···Au interaction of 3-MBA in the *anti* form (Figure 4d) gave the C═O stretch of 1761 cm^{-1}, which is 7 cm^{-1} higher than the computed value for the "normal" *anti* configuration in Figure 4b and 13 cm^{-1} higher than the computed value for the free *anti* 3-MBA. We thus assigned the highest frequency observed for both $Au_{144}(3\text{-MBA})_{\sim 40}$ and $Au_{68}(3\text{-MBA})_{32}$ to the O═C−OH···Au interaction visualized in Figure 4d. This interaction at the ligand−metal interface has not been reported before for any thiolate-protected gold nanocluster.

Self-assembling monolayers (SAMs) of thiols on flat gold surfaces have been studied intensively from early 1980's as a convenient system for modifying physical and chemical properties of gold surfaces for applications.[8] A vast majority of the work has been done by using organosoluble thiols, which cannot have such interactions with gold as what is discussed here for 3-MBA. However, in a few cases, MBA-SAM systems have been studied.[16,17] Ref 16 is particularly interesting since it presents a systematic FTIR investigation of SAMs on gold consisting of 2-MBA, 3-MBA, and 4-MBA thiols. The freshly prepared 3-MBA SAM shows C═O stretch frequencies of both free (monomer) and associated (dimerized) molecules.[16] Upon aging in H_2O and CO_2−free air, the monomer peak shifts to lower frequencies (from 1750 to 1739 cm^{-1}) indicating rearrangement of the ligand layer and possibly interaction with the gold surface. It can be noted that the lower frequency is close to our measured peak at 1731 cm^{-1}, assigned here to the thiol−gold interaction.

DISCUSSION

The results reported in this work demonstrate several properties of 3-MBA-protected gold nanoclusters that set them apart from previously characterized organo-soluble or water-soluble gold clusters stabilized by other thiol ligands. Below, the distinct properties of $Au_{68}(3\text{-MBA})_{32}$ and $Au_{144}(3\text{-MBA})_{\sim 40}$ are summarized, and some of their implications are discussed.

C_1-Symmetric Ligand Layer and Low Thiol Coverage. Our complementary analysis of $^1H-^1H$ and $^1H - ^{13}C$ correlations using 2D TOCSY, NOESY, and ^{13}C-HSQC experiments indicated 126 signals from the ligand layer of $Au_{68}(3\text{-MBA})_{32}$. Since each 3-MBA contributes by 4 distinct signals, this result implies 31−32 symmetry-unique ligand environments (Tables S1−S5 and Figures S5−S8). This count is remarkably close to the previously reported molecular composition of this cluster based on mass spectrometry[6] and shows unambiguously that the symmetry of the ligand layer is C_1. ^1H NMR spectrum for $Au_{144}(3\text{-MBA})_{\sim 40}$ showed just one broad, unresolved feature in the aromatic region which made the more detailed analysis impossible (Figure 1). Our previously published ^1H NMR spectrum of compositionally precisely determined, organo-soluble $Au_{144}(SCH_2CH_2Ph)_{60}$ nanocluster[11] was highly structured, which shows that while the ligand layer of $Au_{144}(SCH_2CH_2Ph)_{60}$ must have a higher symmetry, although still nondetermined, the most likely symmetry of the ligand layer of $Au_{144}(3\text{-MBA})_{\sim 40}$ is C_1 just as in $Au_{68}(3\text{-MBA})_{32}$. In such a case, there are about 160 signals produced by $Au_{144}(3\text{-MBA})_{\sim 40}$ that merge together in the aromatic region to produce the observed broad feature in Figure 1c.

The C_1-symmetric ligand layer implies chirality. While we have not tried to separate the possible enantiomers of $Au_{144}(3\text{-MBA})_{\sim 40}$ or $Au_{68}(3\text{-MBA})_{32}$ clusters, we have calculated the theoretical circular dichroism (CD) spectrum of the model cluster $Au_{144}(3\text{-MBA})_{40}$. The spectrum (Figure S13) has strong signals in the visible and near-IR range, which confirms the chiral nature of the model cluster.

The ligand count of $Au_{144}(3\text{-MBA})_{\sim 40}$ is remarkably low as compared to its similar-size organo-soluble counterparts $Au_{144}(SCH_2CH_2Ph)_{60}$,[9,13] $Au_{133}(SPh^tBu)_{52}$,[18]) and $Au_{130}(p\text{-MBT})_{50}$ (p-MBT = para methylbenzene thiol.[19] Previous work[20,21] has shown that bulky organosoluble ligands can lead to gold cluster structures and gold:thiol ratios that are significantly different from sterically small ligands. However, our work implies that water-soluble thiol ligands such as 3-MBA can have distinct effects on the structure of the thiol−gold interface in case, where the ligand has an additional functionality (weak interaction to gold) to the conventional gold−sulfur bond. Here, our results can be contrasted to a recent work[22] that reports the total structure (from X-ray crystallography) of the 4-MBA protected gold cluster $Au_{146}(4\text{-MBA})_{57}$. This cluster, despite having almost the same count of gold atoms as $Au_{144}(3\text{-MBA})_{\sim 40}$, has a very different number of thiols, different metal core, metal−thiol interface, and ligand layer.[22] The relatively low ligand density at the thiol−gold interface found in this work for $Au_{144}(3\text{-MBA})_{\sim 40}$ and $Au_{68}(3\text{-MBA})_{32}$ was found to lead to a situation where, in addition to S−Au bonds, some of the thiols interact with the gold also by intermolecular interactions *via* O═C−OH···Au and π−Au configurations. These interactions, while known to exist in small gold complexes,[23,24] have not been previously discussed in the context of large clusters.

The low ligand count in $Au_{144}(3\text{-MBA})_{\sim 40}$ is also supported by comparison of the computed and measured UV−vis absorption spectrum (Figure 2); the computed spectrum for the atomistic model of $Au_{144}(3\text{-MBA})_{40}$ agrees qualitatively with the experimental data, indicating the experimentally observed weak plasmonic band, while the spectrum of the $Au_{144}(3\text{-MBA})_{53}$ model does not. Note that it has been well established previously that organo-soluble $Au_{144}(SCH_2CH_2Ph)_{60}$ clusters do not show the plasmon peak in the experimental or simulated UV−vis spectrum,[13] which makes a clear contrast to $Au_{144}(3\text{-MBA})_{\sim 40}$. From a theoretical viewpoint, one can understand the tendency to support the plasmon as the number of thiols is decreased and number of gold atoms remains the same, as the density of free electrons in the same volume increases.[25,26] The ultimate reasons why 3-MBA thiols stabilize low-symmetry gold nanoclusters with low-density ligand layer are still to be found. Reference MD simulations for $Au_{102}(p\text{-MBA})_{44-x}$ and $Au_{144}(p\text{-MBA})_{60-y}$ models with incomplete ligand layers (x = 12, y = 19) indicated that π−Au interactions with "lying down" p-MBA ligands are possible as well, but the −OH···Au interaction is not sterically possible (Table S6). We speculate here that the initial 3-MBA/Au^I polymer phase, from which the clusters nucleate by reduction, already contains weak thiol−gold interactions in addition to the S−Au covalent bonds, which lead to the lower ligand density in the protecting monolayer.

Reactivity to Thiol-Modified DNA and Proteins. The existence of the weakly protected gold sites at the thiol−gold

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

interface gives a basis for understanding the reported reactivity of 3-MBA-protected gold nanoclusters toward thiol-modified DNA and proteins with an exposed cysteine,[5] thought previously to occur *via* thiol-to-thiol exchange reaction. Azubel and Kornberg reported recently[5] (Figure S14) that while reacting 4-MBA (*p*-MBA) protected gold clusters to thiol-modified oligonucleotides, one sees the original unreacted gold band and a few higher bands where presumably thiol-to-thiol ligand exchange reaction produces 1, 2, ... additions of the new molecule to the cluster. This is an equilibrium reaction, and since the binding energies of 4-MBA and thiol-oligos are presumably close to each other, one sees all the time also the unreacted gold cluster band. On the other hand, in the reaction for 3-MBA protected gold, the original band disappears completely, indicating that the binding of thiol-oligos is irreversible (equilibrium fully in the product side) and all clusters bind at least one thiol-oligo. In the light of our results, this can now be understood since (a) the weak interactions appearing at the interface of 3-MBA/gold are only of the order of 0.1−0.2 eV, and thiol binding is of the order of 2 eV, and (b) the weak interactions are dynamic in water at room temperature, and the gold sites protected by these interactions will be all exposed to incoming thiol-oligo, inducing the irreversible binding. Our simulations predict that there is only a low number (on average 3−4) of gold sites per cluster that are protected by the weak interactions, which agrees very well with the observed maximum number of thiol-oligos bound per cluster (Figure S14).[5]

Catalytic Sites at the Thiol−Gold Interface. Previously, several concepts have been discussed in relation to usage of ligand-stabilized gold nanoclusters as catalysts.[2,27,28] Our work defines a paradigm where utilizing the few weakly protected gold sites of Au_{68}(3-MBA)$_{32}$, Au_{144}(3-MBA)$_{\sim 40}$, and other nanoclusters, that could be stabilized by ligands bonding to gold *via* dynamic weak interactions, opens avenues for engineering size-controlled, ambient-stable, catalytic gold nanocatalysts that combine the concepts of heterogeneous and homogeneous catalysis. In that respect, the suggested π−Au interaction is particularly relevant, since lifting the aromatic ring at the interface exposes up to three Au sites for bonding of reactant molecules.

CONCLUSIONS

In summary, our experimental and computational characterization of Au_{68}(3-MBA)$_{32}$ and Au_{144}(3-MBA)$_{\sim 40}$ shows that gold nanoclusters stabilized by 3-MBA thiols have several distinct characteristics as compared to many of the previously studied organo-soluble or water-soluble gold clusters. The protecting ligand layer has only C_1 symmetry which imposes an asymmetric metal core. The acidic group in the *meta* position of the aromatic ring in 3-MBA gives rise to weak interactions within the ligand layer and at the ligand−gold interface. These interactions are missing in organo-soluble ligand layers and also in the previously well-studied *p*-MBA-stabilized clusters such as Au_{102}(*p*-MBA)$_{44}$. These interactions may also be indirectly responsible for the unusually low ligand coverage in the Au_{144}(3-MBA)$_{\sim 40}$. Several currently unknown compositions and sizes of 3-MBA-protected gold nanoclusters will undoubtedly be found by variations of the known syntheses, which will open unexplored possibilities for applications of these materials in biolabeling, catalyzing biochemical reactions, imaging, detection, and theranostics.

METHODS

NMR Measurements. For structural characterization of Au_{68}(3-MBA)$_{32}$ and Au_{144}(3-MBA)$_{\sim 40}$ clusters, a set of 2D correlation experiments, that is, TOCSY and NOESY for ^{1}H−^{1}H through bond and through space connectivities, respectively, and ^{13}C-HSQC for ^{1}H−^{13}C one-bond connectivities, were employed along with 1D ^{1}H NMR and DOSY NMR. TOCSY and NOESY spectra were collected using mixing times of 20 and 200 ms, respectively. All NMR spectra were measured in D_2O at 303 K using Bruker Avance III HD NMR spectrometer, operating at 800 MHz of ^{1}H frequency, and equipped with a cryogenically cooled TCI ^{1}H, ^{13}C, ^{15}N probe head.

UV−vis and IR Spectroscopy. The FT-IR measurements were done with a Nicolet Magna IR 760 spectrometer using a homemade flow-cell cuvette equipped with CaF_2 windows. The optical path was set to 25 μm for the pure ligand measured in MeOD and 500 μm for other measurements. Temperature of the sample was controlled with a MiraCool thermostat, by circulating isopropanol in the aluminum cuvette-holder. The system was equilibrated for 30 min at every temperature. All spectra were measured with 2 cm^{-1} resolution and processed with Happ−Genzel apodization before Fourier transform. 256 or 512 scans were averaged. All spectra presented have the solvent spectrum subtracted. The UV−vis measurements were done with a PerkinElmer Lambda 850 spectrometer in a 1 mm quartz cuvette.

Molecular Dynamics Simulations. The initial configurations for the ligand layers of Au_{68}(3-MBA)$_{32}$ and Au_{144}(3-MBA)$_{40-53}$ were created by using an in-house algorithm (see details on Section B, Supporting Information). The gold core was fixed to the experimental TEM structure.[6,7] The ligand layer was further relaxed and dynamically simulated by using the GROMACS 5.0.4[14] MD simulation package with an AMBER force field for thiol−gold interactions, parametrized in our previous work.[29] Each of the structural models was simulated with varying degrees of the protonation state. After initial equilibration phases (see technical details in the Supporting Information, pages 21−22), the statistics were collected over 50 ns. Au_{102}(*p*-MBA)$_{44}$[1] and Au_{144}(*p*-MBA)$_{60}$ (model based on ref 10) clusters with full and defective ligand shells were also simulated for reference.

DFT Calculations. Selected configurations from the GROMACS MD simulations were further studied by DFT calculations. We used the code package GPAW[30] which employs the projector augmented wave method (PAW) in the real-space grid. Scalar relativistic effects for gold were included in the PAW setup. Total energy calculation and structure optimization were carried out by using the Perdew−Burke−Ernzerhof (PBE) exchange−correlation functional[31] and a real space grid spacing of 0.2 Å. The structure optimization was carried out until the residual forces were below 0.1 eV/Å. During the optimization, all Au atoms were kept fixed with the positions determined from the TEM experiments.[6,7] The optical absorption spectra for the optimized structures were calculated by using the linear response time-dependent DFT (LR-TDDFT) as implemented in GPAW.[32] The PBE functional was used for the exchange−correlation kernel, and the spacing in the real space grid was 0.25 Å. The plotted continuous absorption spectra in Figure 2 of the main text are sums of the individual absorption lines broadened with 0.05 eV Gaussians. IR vibration analysis of selected ligands was done on structures optimized by a tighter criterion with 0.05 eV/Å residual force.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsnano.7b07787.

Figures S1−S14, Tables S1−S6, list of Au coordinates of Au_{144}(3-MBA)$_{\sim 40}$ from ref 7, additional NMR analysis of Au_{68}, details of the computational methods, analysis of MD simulations, analysis of IR data (PDF)

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

AUTHOR INFORMATION

Corresponding Author

*E-mail: hannu.j.hakkinen@jyu.fi.

ORCID

Satu Mustalahti: 0000-0001-9685-0015
Mika Pettersson: 0000-0002-6880-2283
Hannu Häkkinen: 0000-0002-8558-5436

Author Contributions

#These authors contributed equally. T.-R.T. and P.P. measured and interpreted the NMR data. B.K. and S.Mu. measured and analyzed the IR and UV−vis data. E.P. and S.L. performed the MD simulations. S.Ma. devised the algorithm for creating initial conformations of the ligand layer and performed the DFT calculations on the electronic structure and IR frequencies. G.G., M.P., and H.H. advised the work and contributed to the interpretation of data. H.H. compiled the initial manuscript draft which was commented on by all authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants 258806, 266492, 290077, and 294217 from the Academy of Finland. H.H. acknowledges the Finnish Academy Professorship. The computations were done at the Nanoscience Center (NSC) in University of Jyväskylä and at the Finnish national supercomputer center CSC. We thank M. Azubel and R. D. Kornberg for sending samples of 3-MBA-protected Au_{68} and Au_{144} clusters for analysis as well as for numerous discussions while preparing the joint submission of papers and T. Tsukuda for discussions on the mass spectrometry data of Au_{144}.

REFERENCES

(1) Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Bushnell, D. A.; Kornberg, R. D. Structure of a Thiol Monolayer-Protected Gold Nanoparticle at 1.1 Angstrom Resolution. *Science* **2007**, *318*, 430−433.

(2) Tsukuda, T.; Häkkinen, H. *Protected Metal Clusters: From Fundamentals to Applications*; Elsevier: Amsterdam, 2015.

(3) Salorinne, K.; Malola, S.; Wong, O. A.; Rithner, C. D.; Chen, X.; Ackerson, C. J.; Häkkinen, H. Conformation and Dynamics of the Ligand Shell of a Water-Soluble Au102 Nanoparticle. *Nat. Commun.* **2016**, *7*, 10401.

(4) Ackerson, C. J.; Jadzinsky, P. D.; Kornberg, R. D. Thiolate Ligands for Synthesis of Water-Soluble Gold Clusters. *J. Am. Chem. Soc.* **2005**, *127* (2005), 6550−6551.

(5) Azubel, M.; Kornberg, R. D. Synthesis of Water-Soluble, Thiolate-Protected Gold Nanoparticles Uniform in Size. *Nano Lett.* **2016**, *16*, 3348−3351.

(6) Azubel, M.; Koivisto, J.; Malola, S.; Bushnell, D.; Hura, G. L.; Koh, A. L.; Tsunoyama, H.; Tsukuda, T.; Pettersson, M.; Häkkinen, H.; Kornberg, R. D. Electron Microscopy of Gold Nanoparticles at Atomic Resolution. *Science* **2014**, *345*, 909−912.

(7) Azubel, M.; Koh, A. L.; Koyasu, K.; Tsukuda, T.; Kornberg, R. D. Structure Determination of a Water-Soluble 144-Gold Atom Particle at Atomic Resolution by Aberration-Corrected Electron Microscopy. *ACS Nano* **2017**, DOI: 10.1021/acsnano.7b06051.

(8) Häkkinen, H. The Gold−Sulfur Interface at the Nanoscale. *Nat. Chem.* **2012**, *4*, 443−455.

(9) Qian, H.; Jin, R. Controlling Nanoparticles with Atomic Precision: The Case of Au 144(SCH2CH2Ph)60. *Nano Lett.* **2009**, *9*, 4083−4087.

(10) Lopez-Acevedo, O.; Akola, J.; Whetten, R. L.; Gronbeck, H.; Häkkinen, H. Structure and Bonding in the Ubiquitous Icosahedral Metallic Gold Cluster Au-144(SR)(60). *J. Phys. Chem. C* **2009**, *113*, 5035−5038.

(11) Salorinne, K.; Lahtinen, T.; Koivisto, J.; Kalenius, E.; Nissinen, M.; Pettersson, M.; Häkkinen, H. Nondestructive Size Determination of Thiol-Stabilized Gold Nanoclusters in Solution by Diffusion Ordered NMR Spectroscopy. *Anal. Chem.* **2013**, *85*, 3489−3492.

(12) Salorinne, K.; Lahtinen, T.; Malola, S.; Koivisto, J.; Häkkinen, H. Solvation Chemistry of Water-Soluble Thiol-Protected Gold Nanocluster Au102 from DOSY NMR Spectroscopy and DFT Calculations. *Nanoscale* **2014**, *6*, 7823−7826.

(13) Negishi, Y.; Nakazaki, T.; Malola, S.; Takano, S.; Niihori, W.; Kurshige, W.; Yamazoe, S.; Tsukuda, T.; Häkkinen, H. A Critical Size for Emergence of Nonbulk Electronic and Geometric Structures in Dodecanethiolate-Protected Au Clusters. *J. Am. Chem. Soc.* **2015**, *137*, 1206−1212.

(14) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(15) Koivisto, J.; Chen, X.; Donnini, S.; Lahtinen, T.; Häkkinen, H.; Groenhof, G.; Pettersson, M. Acid-Base Properties and Surface Charge Distribution of the Water-Soluble Au102(*p*-MBA)44 Nanocluster. *J. Phys. Chem. C* **2016**, *120*, 10041.

(16) Wells, M.; Dermody, D. L.; Yang, H. C.; Kim, T.; Crooks, R. M.; Ricco, A. J. Interactions Between Organized, Surface-Confined Monolayers and Vapor-Phase Probe Molecule. 9. Structure/Reactivity Relationship Between Three Surface-Confined Isomers of Mercaptobenzoic Acid and Vapor-Phase Decylamine. *Langmuir* **1996**, *12*, 1989−1996.

(17) Lee, J. R. I.; Willey, T. M.; Nilsson, J.; Terminello, L. J.; De Yoreo, J. J.; van Buuren, T. Effect of Ring Substitution Position on the Structural Conformation of Mercaptobenzoic Acid Self-Assembled Monolayers on Au(111). *Langmuir* **2006**, *22*, 11134−11141.

(18) Dass, A.; Theivendran, S.; Nimmala, P. R.; Kumara, C.; Jupally, V. R.; Fortunelli, A.; Sementa, L.; Barcaro, G.; Zuo, X.; Noll, B. C. Au 133(SPh- tBu)52 Nanomolecules: X-ray Crystallography, Optical, Electrochemical, and Theoretical Analysis. *J. Am. Chem. Soc.* **2015**, *137*, 4610−4613.

(19) Chen, Y.; Zeng, C.; Liu, C.; Kirschbaum, K.; Gayathri, C.; Gil, R. R.; Rosi, N. L.; Jin, R. Crystal Structure of Barrel-Shaped Chiral Au130(*p*-MBT)50 Nanocluster. *J. Am. Chem. Soc.* **2015**, *137*, 10076−10079.

(20) Krommenhoek, P. J.; Wang, J.; Hentz, N.; Johnston-Peck, A. C.; Kozek, K. A.; Kalyuzhny, G.; Tracy, J. B. Bulky Adamantanethiolate and Cyclohexanethiolate Ligands Favor Smaller Gold Nanoparticles with Altered Discrete Sizes. *ACS Nano* **2012**, *6*, 4903−4911.

(21) Nishigaki, J.-I.; Tsunoyama, R.; Tsunoyama, H.; Ichikuni, N.; Yamazoe, S.; Negishi, Y.; Ito, M.; Matsuo, T.; Tamao, K.; Tsukuda, T. A New Binding Motif of Sterically Demanding Thiolates on a Gold Cluster. *J. Am. Chem. Soc.* **2012**, *134*, 14295−14297.

(22) Vergara, S.; Lukes, D. A.; Martynowycz, M. W.; Santiago, U.; Plascencia-Villa, G.; Weiss, S. C.; de la Cruz, M. J.; Black, D. M.; Alvarez, M. M.; Lopez-Lozano, X.; Barnes, C. O.; Lin, G.; Weissker, H. C.; Whetten, R. L.; Gonen, T.; Yacaman, M. J.; Calero, G. MicroED Structure of Au_{146}(p-MBA)$_{57}$ at Subatomic Resolution Reveals a Twinned FCC Cluster. *J. Phys. Chem. Lett.* **2017**, *8*, 5523−5530.

(23) Schmidbaur, H.; Raubenheimer, H. G.; Dobrzanska, L. The Gold-Hydrogen Bond, Au-H, and the Hydrogen Bond to Gold, Au···H-X. *Chem. Soc. Rev.* **2014**, *43*, 345−380.

(24) Tiekink, E. R. T.; Zukerman-Schpector, J. Gold-pi Aryl Interactions as Supramolecular Synthons. *CrystEngComm* **2009**, *11*, 1176−1186.

(25) Walter, M.; Akola, J.; Lopez-Acevedo, O.; Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Whetten, R. L.; Grönbeck, H.; Häkkinen, H. A Unified View of Ligand-Protected Gold Clusters as Superatom Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9157−9162.

(26) Malola, S.; Lehtovaara, L.; Enkovaara, J.; Häkkinen, H. Birth of the Localized Surface Plasmon Resonance in Monolayer-Protected Gold Nanoclusters. *ACS Nano* **2013**, *7*, 10263−10270.

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

(27) Lopez-Acevedo, O.; Kacprzak, K. A.; Akola, J.; Häkkinen, H. Quantum Size Effects in Ambient CO Oxidation Catalysed by Ligand-Protected Gold Clusters. *Nat. Chem.* **2010**, *2*, 329−334.
(28) Chen, X.; Häkkinen, H. Protected but Accessible: Oxygen Activation by a Calixarene-Stabilized Undecagold Cluster. *J. Am. Chem. Soc.* **2013**, *135*, 12944−12947.
(29) Pohjolainen, E.; Chen, X.; Malola, S.; Groenhof, G.; Häkkinen, H. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *J. Chem. Theory Comput.* **2016**, *12*, 1342−1350.
(30) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dulak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; Lehtovaara, L.; Ljungberg, M.; Lopez-Acevedo, O.; Moses, P. G.; Ojanen, J.; Olsen, T.; Petzold, V.; Romero, N. A.; Stausholm-Møller, J.; Strange, M.; et al. Electronic Structure Calculations with GPAW: a Real-Space Implementation of the Projector Augmented-Wave Method. *J. Phys.: Condens. Matter* **2010**, *22*, 253202.
(31) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(32) Walter, M.; Häkkinen, H.; Lehtovaara, L.; Puska, M.; Enkovaara, J.; Rostgaards, C.; Mortensen, J. J. Time-Dependent Density-Functional Theory in the Projector Augmented-Wave Method. *J. Chem. Phys.* **2008**, *128*, 244101.

DOI: 10.1021/acsnano.7b07787
ACS Nano 2017, 11, 11872−11879

III

EXPLORING STRATEGIES FOR LABELING VIRUSES WITH GOLD NANOCLUSTERS THROUGH NON-EQUILIBRIUM MOLECULAR DYNAMICS SIMULATIONS

Reprinted with permission from

Emmi Pohjolainen, Sami Malola, Gerrit Groenhof, Hannu Häkkinen. *Bioconjugate Chemistry*, 28 (9), pp 2327–2339, **2017**.

Bioconjugate Chemistry

Article

pubs.acs.org/bc

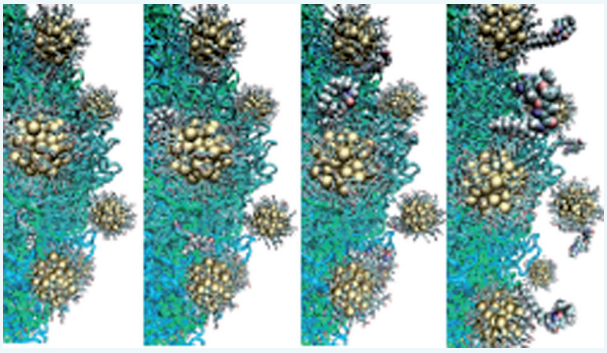
Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations

Emmi Pohjolainen,[†] Sami Malola,[†] Gerrit Groenhof,[*,‡] and Hannu Häkkinen[*,†,‡]

[†]Department of Physics and [‡]Department of Chemistry, Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland FI-40014

Supporting Information

ABSTRACT: Biocompatible gold nanoclusters can be utilized as contrast agents in virus imaging. The labeling of viruses can be achieved noncovalently but site-specifically by linking the cluster to the hydrophobic pocket of a virus via a lipid-like pocket factor. We have estimated the binding affinities of three different pocket factors of echovirus 1 (EV1) in molecular dynamics simulations combined with non-equilibrium free-energy calculations. We have also studied the effects on binding affinities with a pocket factor linked to the $Au_{102}pMBA_{44}$ nanocluster in different protonation states. Although the absolute binding affinities are over-estimated for all the systems, the trend is in agreement with recent experiments.[3] Our results suggest that the natural pocket factor (palmitic acid) can be replaced by molecules pleconaril (drug) and its derivative Kirtan1 that have higher estimated binding affinities. Our results also suggest that including the gold nanocluster does not decrease the affinity of the pocket factor to the virus, but the affinity is sensitive to the protonation state of the nanocluster, i.e., to pH conditions. The methodology introduced in this work helps in the design of optimal strategies for gold−virus bioconjugation for virus detection and manipulation.



INTRODUCTION

Recent years have seen many breakthroughs in the synthesis and characterization of atom-precise, ligand-stabilized gold nanoclusters in the size range of 1 to 3 nm.[1] The ability to produce biocompatible thiolate-protected gold nanoclusters via ligand tuning has promoted their use in biological applications (for example, as potential imaging agents,[2−5] molecular rulers[6−8] and in biological sensing and medical analysis).[9−11] Recent studies have demonstrated the potential use of thiolate-protected gold nanoclusters as contrast agents in virus imaging with TEM[2,3] by either covalent[2] or noncovalent[3] linking of gold nanoclusters to the virus surface. While the covalent linking relies on the formation of sulfur linkages between the gold nanocluster functionalized by a suitable linker molecule and accessible cysteine residues of the protein,[2] the noncovalent linking is achieved by the replacement of the natural pocket factor (e.g., palmitic acid in EV1) in the hydrophobic pocket of the virus by a pleconaril (PLE) drug-like molecule Kirtan1 with a linker tail that extends to the outside of the hydrophobic pocket and connects covalently to the gold nanocluster.[3] Although the noncovalent binding was deemed successful based on the transmission electron microscopy (TEM) imaging,[3] no molecule-scale structural information is available from the experiments. This has motivated us to study the interactions between the viruses, hydrophobic pocket factors, and the gold nanoclusters via large-scale all-atom molecular dynamics simulations consisting of the full EV1 virus in water and composed of about 3.2 million atoms.

Molecular dynamics (MD) simulations are commonly used to investigate large biological systems. In the field of gold nanoclusters and their effects in biological systems, MD simulations have been utilized, for example, in studying the interplay between gold nanoclusters and membranes,[12−18] RNA,[19] DNA,[20] and proteins.[20,21] Ligand binding can also be addressed in MD simulations, and we use that here to explore the binding affinity of various pocket factors.

For large complex systems, equilibrium free-energy approaches are often impractical due to the long equilibration times associated with the changes under investigation. The non-equilibrium approaches based on Jarzynski equality enable the obtaining of the free-energy estimate from the distribution of work values of many repeated non-equilibrium processes via exponential averaging.[22] Steered MD simulations combined with free-energy calculations can shed light on many types of properties and processes of biomolecules, such as elastic properties,[23] and, additionally, for viruses,[24,25] protein folding and conformations,[26−29] ion transport,[30−32] and ligand-binding and -dissociation processes.[33−38] In the field of gold nano-

Received: June 27, 2017
Revised: August 11, 2017
Published: August 14, 2017

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

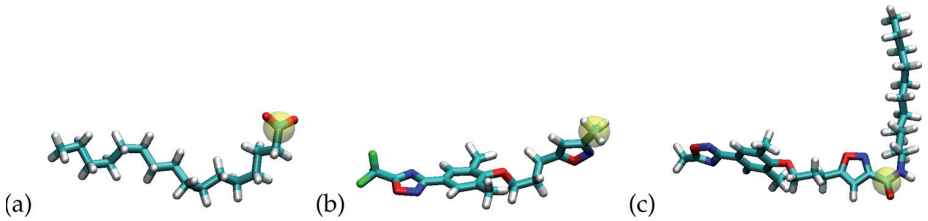


Figure 1. Pocket factors of EV1. (a) Natural pocket factor palmitic acid (PLT), (b) drug molecule pleconaril (PLE), and (c) pleconaril-derivative Kirtan1 molecule with a linker tail extending outside the pocket (KIR). The pull group carbon is shown highlighted in yellow. Coloring convention: cyan, carbon; red, oxygen; blue, nitrogen; green, fluorine; and white, hydrogen.

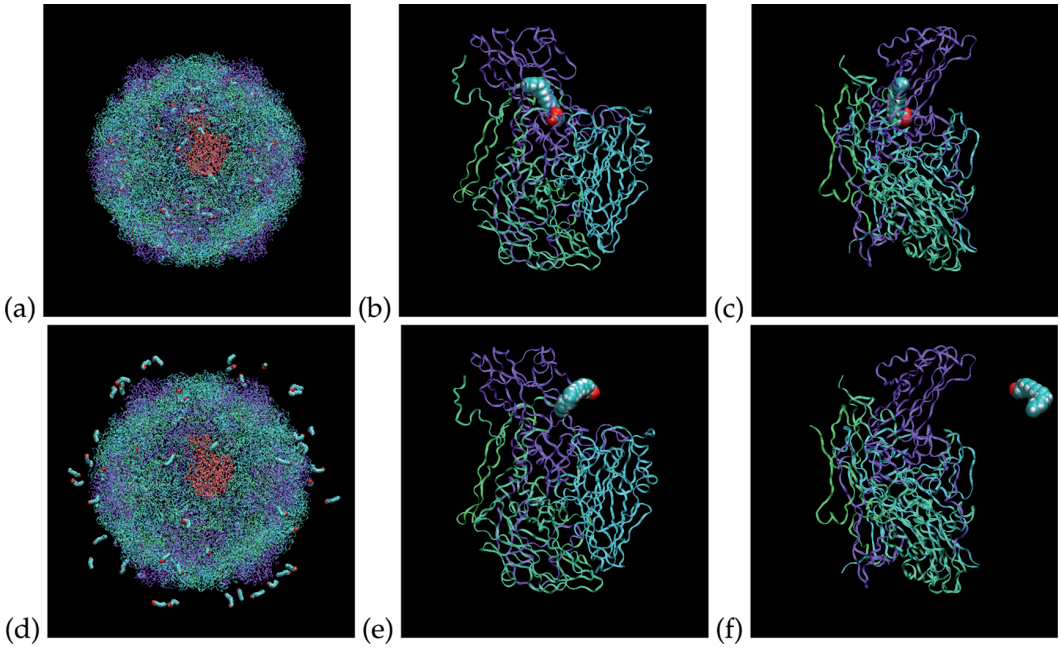


Figure 2. (a) Full virus with PLT pocket factor (drawn in VDW spheres) at the beginning of the pulling simulation, with 1/60 unit highlighted in red and close-ups in panels b and in c with 90° rotation around a vertical axis. (d–f) The full virus and 1/60 unit close-ups at the end of one pulling simulation.

clusters, free-energy calculations have been utilized in studying their interactions with membranes[18] and RNA.[19]

In practice, the direct Jarzynski estimate is biased due to the finite sample size, and corrections need to be applied.[39,40] Simplest corrections include using cumulant expansions of the work distributions.[39] However, such simple corrections are not suitable for highly non-Gaussian distributions.[41] However, an alternative option is to utilize the Crooks fluctuation theorem[42] by performing non-equilibrium simulations in both the forward and the reverse directions of the process being studied to reduce the bias of the estimate.[43]

In this work, we studied echovirus 1 (EV1) with three different pocket factors shown in Figure 1 (the natural pocket factor palmitic acid, a drug molecule (pleconaril), and a pleconaril-derivative Kirtan1 molecule) by means of all atom MD simulations with a complete model of the EV1 protein capsid in water. We have investigated the interactions and estimated the binding affinities of the different pocket factors using free-energy calculations. Because the complexity and size, on the one hand, and the lack of similarity between the pocket factors on the other hand, preclude equilibrium free-energy approaches, such as thermodynamic integration, free-energy perturbation, or umbrella sampling, we have used steered center of mass pulling MD simulations and a non-equilibrium Jarzynski equality-based, bias-corrected analysis method, as suggested by Zuckerman et al.[40,41] instead for obtaining the binding affinity estimates. Due to the complexity of the system, the simulations were only performed in one direction, namely from the inside of the pocket into the water. Because we simultaneously pulled out all 60 pocket factors in our production simulations, we obtained statistics from 60 pullings per each simulation, benefiting the extensive statistics that such non-equilibrium methods call for.

In addition, we have studied the effects of different simulation parameters and initial conditions on the free-energy estimates. Lastly, we have studied the effects of adding the gold nanoclusters by linking a $Au_{102}pMBA_{44}$ gold nanocluster to a Kirtan1 molecule to study in detail the effects on the binding affinity and interactions and, hence, assess the applicability of such functionalized gold nanoclusters in virus-imaging applications.[2,3]

RESULTS AND DISCUSSION

Pulling Simulations. A total of 25 4.5 ns pulling simulations were performed for each of the 5 systems. Figure 2 shows snapshots at the beginning and at the end of one

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327–2339

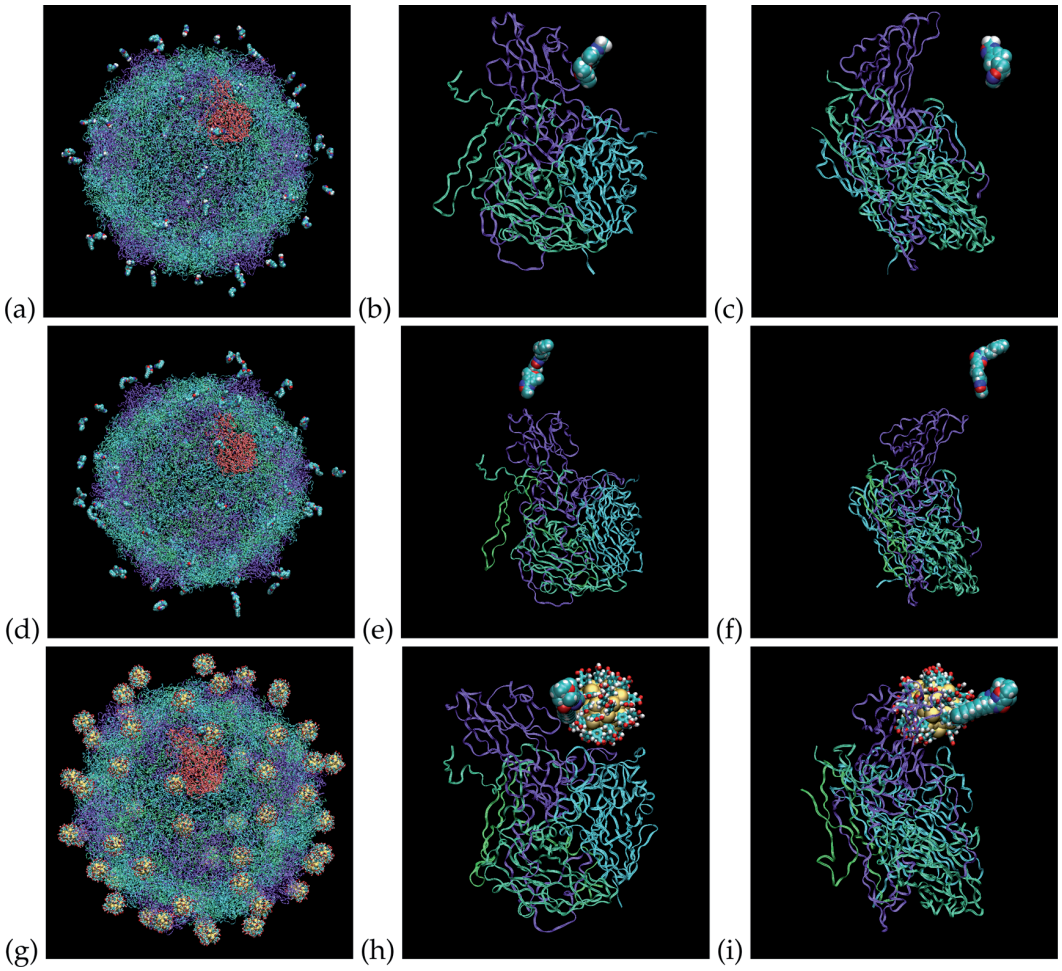


Figure 3. (a–c) Full virus and 1/60 unit close-ups with PLE, (d–f) with KIR, and (g–i) with CLUKIR pocket factor at the end of one pulling simulation.

representative pulling simulation for the natural pocket factor palmitic acid (PLT) system with the full virus and a close-up of one of the 1/60 protomer units. The end states from one representative pulling simulation of the full virus and 1/60 unit for the PLE, KIR, and Kirtan1 linked to Au102pMBA44 gold nanocluster in a half protonated and -deprotonated state (CLUKIR) systems are presented in Figure 3.

From Figures 2 and 3, it can be seen that all the pocket factors exit the pockets. From Figure 3g–i, the trend can be seen of the gold nanoclusters tending to remain close to the virus surface while the pocket-factor part slips past the cluster and exits the pocket. This could affect the binding affinity estimate if the gold nanocluster would significantly affect the water solubility of Kirtan1. To address this, we performed extended pulling simulations in which the gold nanocluster completely detached from the virus surface.

The root mean square deviations (RMSD) of the hydrophobic pocket in the extended reference and pulling trajectories are presented in Figure 4. The black vertical lines denote the time at which the pulling simulation was started. The left side of the black line shows the RMSD of the reference trajectory, and the right side shows the 4.5 ns extension of the reference trajectory and pulling trajectory in green. The scale of the time axis is widened for the extended and pulling part for clarity.

From Figure 4, it can be seen that for all systems the RMSDs observed in the pulling simulations do not significantly exceed the values observed in the extended reference trajectory. Furthermore, no visual rupture of the protein or the pocket factors were observed in the pulling simulations. We also compared pocket-factor conformations in pulling and in free simulations (Supporting Information), and no significant differences were observed. Therefore, we conclude on the basis of this structural analysis that pulling does not seem to induce unphysical conformational changes in the protein.

Free-Energy Calculations. The work distributions from the 25 pulling simulations (i.e., 1500 pulling events) with mean, standard deviation, and skewness indicated (in kJ/mol) are presented in Figure 5. From this, it can be seen that the work distributions show large standard deviations and positive skew. It is therefore clear that analysis methods requiring near-Gaussian distributed data are not applicable. Instead, we resort to the analyses methods suggested by Zuckerman et al.,[40,41] suitable for skewed and non-Gaussian data. Comparing the averages of the work distributions, the following trend is observed:

$$\langle W_{\mathrm{PLT}} \rangle < \langle W_{\mathrm{PLE}} \rangle < \langle W_{\mathrm{KIR}} \rangle < \langle W_{\mathrm{DCLUKIR}} \rangle < \langle W_{\mathrm{CLUKIR}} \rangle$$

The convergence of the block averaged direct Jarzynski estimate (average and standard deviation from 500 repetitions

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327–2339

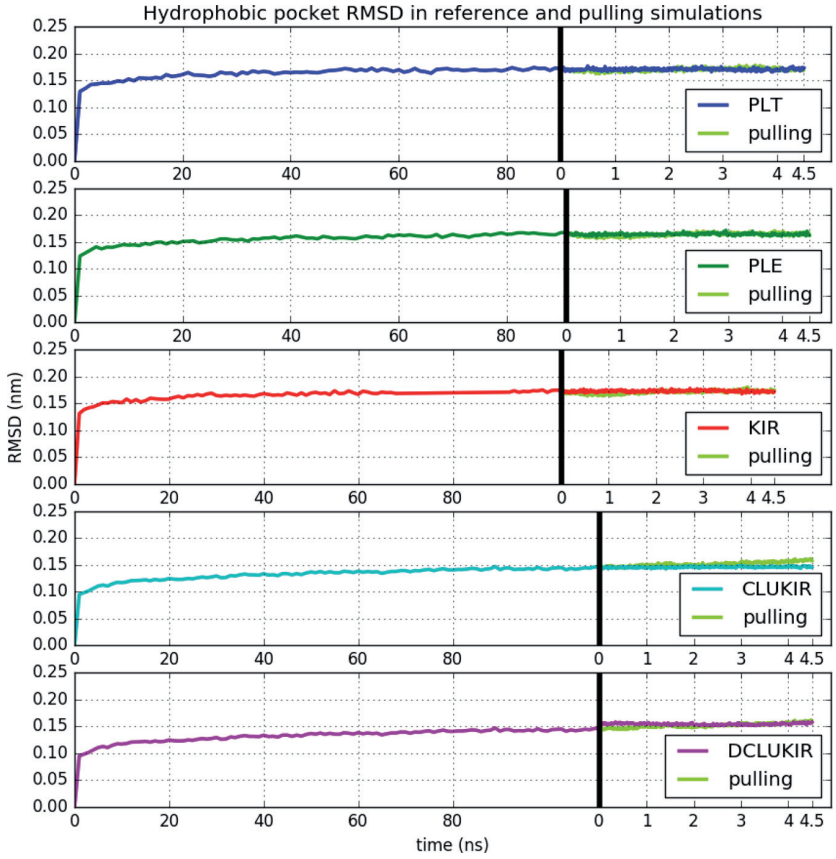


Figure 4. Root mean square deviations of the hydrophobic pocket in extended reference and pulling trajectories.

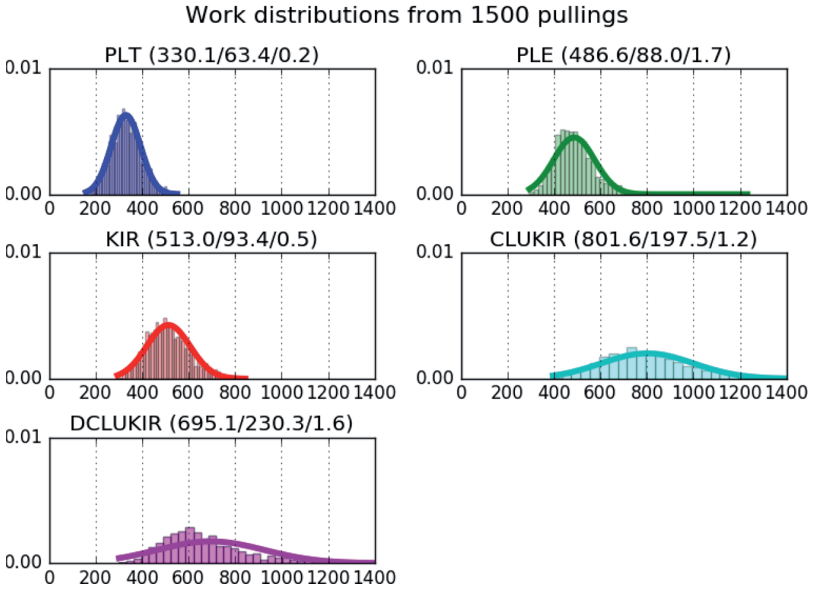


Figure 5. Work distributions from 25 pulling simulations with indicated mean, standard deviation, and skewness in kJ/mol.

of each N, with $n = 1 \cdots N$ as described earlier) as a function of total number of work values N included in the analysis is illustrated in Figure 6. From there, it can be seen that the change in the estimate in between for example the $N = 1000$ and 1500 estimates is relatively small for all of the systems. Therefore, the full data set of $N = 1500$ work values is expected to be sufficient to give at least qualitatively relevant results. The mean and standard deviation of the work distributions and the binding affinity estimates of the block-averaged direct Jarzynski estimates (JE) and the results with linear (LE), quadratic (QE), and cumulative integral (CI) extrapolations are presented in Table 1.

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

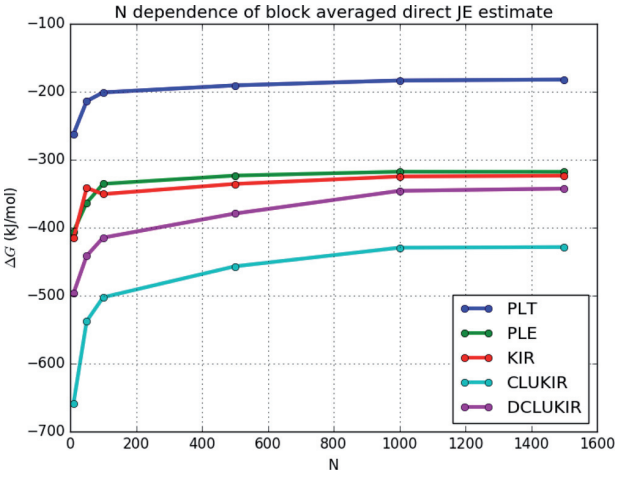


Figure 6. Block-averaged direct Jarzynski estimates with different total number of work values N included in the analysis (N = 1500 is the maximum). Shown is the average from 500 repetitions.

As the final estimate, we report the average plus or minus the combined error obtained from the LE, QE, and CI estimates, from lowest to highest binding affinity: $\Delta G_{PLT} = -168.1 \pm 12$ kJ/mol, $\Delta G_{DCLUKIR} = -293.6 \pm 19$ kJ/mol, $\Delta G_{KIR} = -295.3 \pm 13$ kJ/mol, $\Delta G_{PLE} = -299.9 \pm 18$ kJ/mol, and $\Delta G_{CLUKIR} = -392.3 \pm 14$ kJ/mol. The natural pocket factor palmitic acid shows the lowest binding affinity, and Kirtan1 bonded to the half-protonated and -deprotonated gold nanocluster shows the highest binding affinity. Pleconaril, Kirtan1, and Kirtan1 bonded to the fully deprotonated gold nanocluster show similar affinities.

Among the highest binding affinities in nature is the one of avidin−biotin complex, with a dissociation constant of 10^{-15} M[44] or a binding affinity of −85.4 kJ/mol. Compared to this value, the absolute binding affinity estimates obtained here are high, and the replacement of the natural pocket factor with such a high binding affinity would be implausible. The overestimation of the binding affinities could originate from different factors discussed in the next section, such as pulling velocity and pulling distance. The general trend in the calculated estimates is nevertheless in agreement with the experimental result that palmitic acid can be replaced by pleconaril or Kirtan1 molecules, which, in this analysis, show higher affinity estimates compared to that of palmitic acid. Furthermore, pleconaril and Kirtan1 exhibit similar affinities, which could be expected based on the similarity of the core structures of these molecules. This could also suggest the linker tail does not notably affect the binding affinity.

It is to be noted that in the experiments,[3] a 5000-fold excess of Kirtan1 was used, which contributes to the binding affinity with an entropic term of $-k_BT \log(5000) \approx -21$ kJ/mol. In comparison to the estimated binding affinity values as well as error estimates, this contribution is relatively small and does not change the interpretation of the simulation results.

Based on these results, the addition of the gold nanocluster in a half-protonated and -deprotonated state appears to overall enhance binding of Kirtan1 to the virus. The affinity estimate decreases with the fully deprotonated state of the cluster, and this estimate is close to the estimate of pure Kirtan1, possibly because the deprotonated gold nanocluster enhances the solubility.

Interactions. Protein−ligand interactions were analyzed for a representative pulling trajectory and listed in Table 2. The contact counts are shown as the difference between the initial state (average of the first 300 ps of the pulling simulation) and the final state (average of last 300 ps of the pulling simulation) so that the negative numbers denote broken and the positive numbers newly formed contacts. All counts were divided by 60 to obtain an estimate for the interactions of one pocket factor (plus a gold nanocluster in CLUKIR and DCLUKIR systems). HB denotes hydrogen bonds, HP denotes hydrophobic bonds, R denotes ring contacts, and pf denotes pocket factors. The contacts for the linker part of Kirtan1 molecule are presented in parentheses.

The negative numbers quantifying the contacts between the protein and the pocket factors or the *p*MBA groups indicate overall breaking of interactions between these parts of the systems occurs during the pulling simulation for all the studied systems. Similarly for all the systems hydrogen bonds between the solvent and the pocket factors or the *p*MBA groups are increasingly formed as indicated by the positive numbers.

To build a rapid scoring function to estimate the binding affinity trends, the relative strengths of the interactions were estimated simply by assigning weights HB_w, HP_w, and R_w to best fit to $HB_w \times HB + HP_w \times HP + R_w \times R = \Delta G_i$. The values giving the best fit are HB_w = 31.25, HP_w = 16.75, and R_w = 25.25. In general, it can be estimated that not merely hydrogen bonds but also both hydrophobic and ring contacts are likely to contribute to the binding affinities of these molecules.

Snapshot Dependency. The mean and standard deviation of the work distributions and the free energy estimate obtained from the LE, QE, and CI analyses for different snapshots, and their average are presented in Table 3. The total number of work values was 60 for each snapshot (1 simulation per snapshot).

Based on the estimates and their errors from different snapshots, in Table 3, there seems no systematic dependency of the work values or binding affinity estimates on the reference simulation starting snapshot for any of the systems. Therefore, no bias is expected to be caused by whether the pulling was started from a later or an earlier snapshot of the reference simulation, indicating sufficient equilibration of the reference simulation for the purposes of this work.

Table 1. Work and Binding-Affinity Results in kJ/mol from N = 1500 Work Values with the Reported Affinity Estimate Averages and Standard Deviations Obtained from 500 Repetitions of the Block Averaging Plus Extrapolation Analyses

system	W (mean/standard deviation)	JE	LE	QE	CI
PLT	330.1/63.4	−182.2/0.1	−172.9/2.2	−167.5/0.2	−164.0/10.9
PLE	486.6/88.0	−317.9/0.1	−307.7/2.3	−303.0/0.2	−289.0/17.9
KIR	513.0/93.4	−323.5/0.2	−285.9/9.2	−308.2/0.3	−291.6/8.2
CLUKIR	801.6/197.5	−428.5/0.2	−391.5/7.8	−399.8/0.3	−385.5/11.3
DCLUKIR	695.1/230.3	−342.5/0.3	−269.9/17.0	−311.8/0.5	−299.0/8.2

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

Table 2. Counted Interactions from One Representative Pulling Trajectory for Each System[a]

interaction type	PLT	PLE	KIR	CLUKIR	DCLUKIR
HB: protein−pf	−3.9	−0.2	−0.6	−0.8	−1.0
HB: protein−*p*MBA	–	–	–	−3.4	−3.5
HB: pf−*p*MBA	–	–	–	0	0
HB: solvent−pf	4.3	2.8	4.8	4.7	5.3
HB: solvent−*p*MBA	–	–	–	4.9	6.2
HP: protein−pf	−10.4	−11.9	−11.4 (−4.1)	−10.6 (−3.6)	−9.7 (−3.3)
HP: protein−*p*MBA	–	–	–	−10.0	−5.1
HP: pf−*p*MBA	–	–	–	0.8 (1.2)	0.1 (0.9)
R: protein−pf	–	−7.2	−6.8	−7.2	−7.4
R: protein−*p*MBA	–	–	–	−2.0	−1.6
R: pf−*p*MBA	–	–	–	1.7	0.3

[a]HB, hydrogen bonds; HP, hydrophobic contacts; and R, ring contacts. A negative sign denotes overall broken interactions, whereas no negative sign denotes overall formed interactions.

Table 3. Work (Distributions Presented in Figure S1) and Binding Affinity Results in kJ/mol from N = 60 Work Values for Pulling Simulations Started from Different Snapshots[a]

system	starting snapshot (ns)	W (mean/standard deviation)	ΔG (kJ/mol)
PLT	29	338.1/70.5	−187.2 ± 35
PLT	52	329.1/46.6	−221.4 ± 29
PLT	77	329.1/67.7	−175.2 ± 78
PLT	114	332.8/72.3	−131.0 ± 135
<PLT>	average	332.3/65.2	−150.6 ± 41
PLT	**full**	**330.1/63.4**	**−168.1 ± 12**
PLE	29	504.1/151.0	−335.1 ± 58
PLE	57	485.9/71.7	−357.3 ± 36
PLE	77	492.3/80.5	−299.7 ± 103
PLE	91	504.8/123.8	−310.9 ± 41
<PLE>	average	496.8/111.8	−320.3 ± 32
PLE	**full**	**486.6/88.0**	**-299.9 ± 18**
KIR	29	516.8/77.9	−310.3 ± 99
KIR	57	502.9/85.7	−331.3 ± 56
KIR	92	501.3/98.6	−342.4 ± 61
KIR	103	519.6/92.7	−340.5 ± 44
<KIR>	average	510.1/89.5	−336.5 ± 32
KIR	**full**	**513.0/93.4**	**−295.3 ± 13**
CLUKIR	26	786.1/170.0	−444.1 ± 65
CLUKIR	57	768.8/199.5	−424.6 ± 75
CLUKIR	82	801.5/200.3	−401.9 ± 122
CLUKIR	111	828.8/260.2	−483.7 ± 62
<CLUKIR>	average	796.3/211.2	−428.7 ± 41
CLUKIR	**full**	**801.6/197.5**	**−392.3 ± 14**
DCLUKIR	26	696.8/256.1	−367.5 ± 69
DCLUKIR	57	663.2/202.3	−293.1 ± 129
DCLUKIR	82	684.8/221.0	−363.9 ± 102
DCLUKIR	111	691.1/210.6	−366.6 ± 82
<DCLUKIR>	average	684.0/223.8	−308.4 ± 64
DCLUKIR	**full**	**695.1/230.3**	**−293.6 ± 19**

[a]The estimate from the full data set is shown in bold for comparison.

While the exact time point of the reference simulation snapshot used as a starting structure does not have systematic effect on the binding affinity estimate, in comparing the averaged snapshot results to the full result, differences of up to ∼41.2 kJ/mol can be observed. These discrepancies in the estimates could indicate limitations in the comparability of these results due to an insufficient amount of data obtained from one simulation (or 60 work values from distinguishable pocket factors) only. This is also indicated by the large errors of ΔG estimates obtained for averaged snapshot results.

Dependency on Pulling Velocity. The mean and standard deviation of the work distributions (denoted W_i, where i is the pulling velocity in nm/ns) and the free-energy estimate obtained from the LE, QE, and CI analyses similarly to previous sections (denoted ΔG_i) for pulling simulations with different pulling velocities are presented in Table 4. The total number of work values was 60 for each velocity (one simulation per velocity). The estimates from full data (W_{full} and ΔG_{full}) with 1 nm/ns pulling velocity are shown for comparison.

Considering the work data in Table 4, a consistent trend of decreasing average W with decreasing pulling velocity can be observed for all the systems. The free-energy estimates ΔG_i also decrease with decreasing pulling velocity, with the exception of the PLT system between $\Delta G_{0.25}$ and $\Delta G_{0.5}$. This is probably due to the small number of work values in the distributions and, hence, statistical inaccuracy. Comparing the average ΔG_1 for one simulation data and average ΔG_{full}, similar to the previous section, significant differences of up to 45.3 kJ/mol can be observed. The trends stay nevertheless the same among different velocities, such that the PLT system shows the lowest and the CLUKIR system the highest affinity, and PLE and KIR systems show similar affinities. However, with the lower pulling velocities, a clear narrowing in the difference between PLE, KIR, and CLUKIR systems can be observed. Considering the error estimates and the fact that, as noted also in the previous section, analyses based on data from only 1 simulation or 60 work values can produce results in discrepancy with those from the full data set, this narrowing in the difference may be an artifact from the analyses with insufficient statistics. Figure 7 shows the average and standard deviation (as error bars) of the pulling work values of 60 ligands in one pulling simulation with different pulling velocities for the PLE system.

Although there is a systematic effect of pulling velocity on the binding-affinity estimates, the general trends in binding affinities between the systems stay the same among different pulling velocities. In addition, the 1 nm/ns pulling velocity, a compromise between accuracy and statistical efficiency, should still fall within the activated process regime with no significant frictional effects. This can also be seen in Figure 7 with a visual change of regime around the pulling velocity of 2 nm/ns. Therefore, the analyses using a 1 nm/ns pulling velocity are expected to give over-estimated but qualitatively reliable results.

Cooperative Effects. The mean and standard deviation of the work distributions (denoted W_i, where i denotes the pulled

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

Table 4. Work (Distributions Presented in Figure S2) and Binding Affinity Results in kJ/mol from $N = 60$ Work Values for Pulling Simulations Performed with Different Pulling Velocities

system	W_{10}	W_1	$W_{0.5}$	$W_{0.25}$	W_{full}
PLT	547.9/85.8	327.8/59.9	288.2/59.1	253.0/58.0	330.1/63.4
PLE	771.1/122.8	504.8/123.8	456.2/94.1	400.2/101.6	486.6/88.0
KIR	834.6/120.3	519.6/92.7	463.6/89.7	427.2/101.9	513.0/93.4
CLUKIR	1699.7/396.1	808.9/194.3	680.1/170.3	575.4/136.4	801.6/197.5
system	ΔG_{10}	ΔG_1	$\Delta G_{0.5}$	$\Delta G_{0.25}$	ΔG_{full}
PLT	−331.9 ± 60	−157.7 ± 77	−118.8 ± 76	−144.8 ± 19	−168.1 ± 12
PLE	−515.4 ± 86	−311.6 ± 36	−287.1 ± 83	−248.6 ± 38	−299.9 ± 18
KIR	−565.7 ± 94	−340.6 ± 48	−275.1 ± 90	−259.5 ± 29	−295.3 ± 13
CLUKIR	−838.1 ± 576	−404.2 ± 78	−388.6 ± 46	−272.5 ± 66	−392.3 ± 14

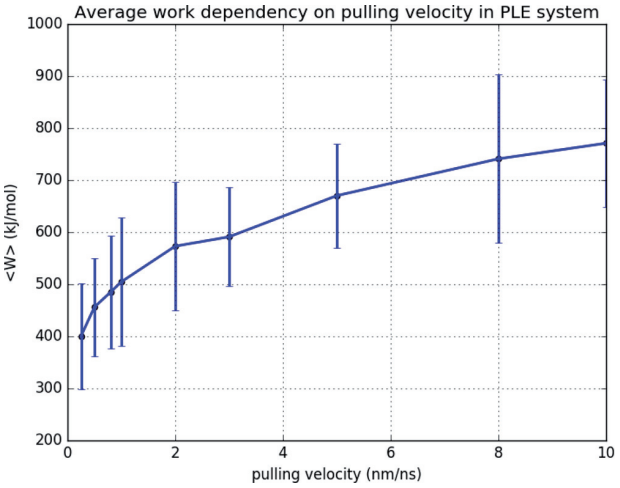


Figure 7. Velocity dependence of average work in one pulling simulation of 60 ligands of the PLE system.

group(s)) and direct Jarzynski free-energy estimate (denoted JE_i) for pulling simulations with different simultaneously pulled groups are presented in Table 5. The results extracted for the corresponding groups from a simulation with all groups pulled simultaneously are shown in parentheses.

After the work distributions and JE estimates from different subgroup pullings between the systems are compared, there appears to be no clear systematic trend in the estimates being larger or smaller for specific groups. Similarly, there is no consistent increase or decrease observed for specific groups in comparison with the results extracted from the simulation with all groups pulled simultaneously.

On average, however, there is a tendency for group 1(21) (that is: singly pulled pocket factor number 21) to exhibit the largest and penta1 and penta12 the smallest average work values and JE estimates. Comparing the estimates extracted from the simulation with all groups pulled, the differences between these subgroups are not as evident. This could suggest that there may be some dependency on how many and which groups are being pulled simultaneously. However, taking into account that the differences between subgroup pullings and the corresponding results from the simulation with all groups pulled are mostly within the standard deviations of the work distributions, in combination with a smaller set of work values for the JE estimate, it is difficult to conclude whether the observed effects originate from cooperativity or from statistical deviations between different simulations. Although we cannot exclude the possibility of the cooperative effects, the possible effect is unlikely to affect the binding affinity estimates in the large-scale statistics.

Extended Pulling. The mean and standard deviation of the work distributions (denoted W_i, where i is the pulling time in ns) and the free-energy estimate obtained from the LE, QE, and CI analyses, similar to previous sections for full systems (denoted ΔG_i) or the direct Jarzynski estimate for the pentamer systems (denoted JE_i) for extended pulling simulations, are presented in Table 6. The estimates obtained from the full data (W_{full}, ΔG_{full}, and JE_{full}) are shown for comparison.

Considering pulling in the full system and taking into account the large error estimates, no significant differences in W or ΔG estimates can be observed between 4.5 and 6 ns pulling for any of the systems. However, possibly due to statistical inaccuracy resulting from smaller set of work values, the ΔG estimates of KIR and CLUKIR are closer to each other compared to ΔG_{full} results, while DCLUKIR exhibits

Table 5. Work (Distributions Presented in Figure S3) and Direct JE Results in kJ/mol for Pulling Simulations Performed with Different Simultaneously Pulled Groups

system	W_{penta1}	$W_{penta12}$	$W_{random12}$	$W_{1(21)}$	$W_{1(59)}$
PLT	361.0/50.4 (345.0/69.7)	325.6/56.6 (327.0/49.6)	333.7/61.2 (326.0/70.3)	364.4 (272.8)	350.8 (333.6)
PLE	476.3/93.5 (483.6/83.8)	478.0/64.0 (504.7/66.3)	493.4/73.7 (503.0/81.5)	511.3 (500.2)	486.7 (494.1)
KIR	463.7/80.4 (493.5/34.6)	497.6/69.3 (499.0/62.1)	475.1/72.1 (490.0/70.7)	422.4 (500.9)	518.6 (461.8)
CLUKIR	832.1/112.4 (818.8/134.4)	823.7/185.5 (852.4/223.0)	891.6/165.1 (906.7/152.8)	899.7 (809.4)	515.8 (615.8)
system	JE_{penta1}	$JE_{penta12}$	$JE_{random12}$	$JE_{1(21)}$	$JE_{1(59)}$
PLT	−302.0 (−280.2)	−233.5 (−259.7)	−237.8 (−204.7)	−364.4 (−272.8)	−350.8 (−333.6)
PLE	−374.2 (−374.7)	−368.0 (−385.2)	−378.5 (−327.3)	−511.3 (−500.2)	−486.7 (−494.1)
KIR	−374.1 (−443.3)	−372.9 (−402.0)	−389.4 (−364.9)	−422.4 (−500.9)	−518.6 (−461.8)
CLUKIR	−665.3 (−669.6)	−580.7 (−610.6)	−691.4 (−705.2)	−899.7 (−809.4)	−515.8 (−615.8)

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

Table 6. Work (Distributions Presented in Figure S4) and Direct JE Results in kJ/mol for Extended Pulling Simulations of Kirtan1-Containing Systems for the Full Virus and 1 Pentamer (1/12) Unit

system	$W_{4.5}$	W_6	W_{full}	$\Delta G_{4.5}$	ΔG_6	ΔG_{full}
KIR	522.1/92.9	527.8/92.1	513.0/93.4	−358.8 ± 40	−371.6 ± 35	−295.3 ± 13
CLUKIR	808.9/194.3	882.8/211.3	801.6/197.5	−401.4 ± 88	−386.7 ± 168	−392.3 ± 14
DCLUKIR	692.4/224.8	760.3/262.2	695.1/230.3	−317.5 ± 145	−308.5 ± 161	−293.6 ± 19

system	$W_{4.5}$	W_{15}	W_{full}	$JE_{4.5}$	JE_{15}	JE_{full}
KIR (pentamer)	569.0/151.1	453.8/83.9	513.0/93.4	−421.9	−340.5	−312.2
CLUKIR (pentamer)	1036.8/119.6	891.7/55.3	801.6/197.5	−914.0	−835.0	−412.9
DCLUKIR (pentamer)	430.7/123.3	554.4/161.8	695.1/230.3	−342.2	−383.0	−318.9

consistently the smallest ΔG estimates. Again, significant differences in the ΔG estimate in comparing the results from the full data set of 1500 work values versus a single pulling set of 60 work values can be observed.

For pentamer systems, the extension of the pulling distance results in a decrease of the average work and JE estimate for KIR and CLUKIR systems, while for DCLUKIR, the effect is opposite. This could be due to insufficient statistics. In pentamer pullings, the CLUKIR system exhibits clearly larger work values and JE estimates compared with KIR and DCLUKIR. While large error estimates and small statistics make it difficult to draw definite conclusions, it appears that, while the affinity of Kirtan molecule to the virus is affected by the gold nanocluster's protonation state, the gold nanocluster (in either protonation state) does not significantly decrease the affinity of Kirtan molecule to the virus, even when full detachment of the clusters from the virus surface is allowed.

CONCLUSIONS

We have estimated the binding affinities of five different pocket factor systems to the hydrophobic pocket of EV1 virus capsid utilizing MD simulations with non-equilibrium free-energy calculations and analysis methods based on the Jarzynski equality with block averaging and extrapolation schemes that account for the bias associated with the direct estimate. The binding affinity estimates were obtained by extracting work values from steered MD simulations of simultaneously pulling all of the 60 pocket factors out from the pockets that they occupied. A total of 25 such pulling simulations for each of the 5 systems were performed, constituting a total of 1500 work values per system for the free-energy estimations.

While experimental results for absolute binding affinities of these systems are not available by comparing these results to one the highest binding affinities in the nature of an avidin−biotin complex (with binding affinity of approximately −85.4 kJ/mol),[44] the over-estimation is evident. This could in part be explained by the relatively high pulling velocity used in the simulations. Lowering of pulling velocity lowers the affinity estimates with no observed convergence. While the pulling velocity is high, no structural deformation of the protein or pocket factors in terms of RMSD or visual inspection were observed in the pulling simulations. Experiments suggest[3] that the exchange of the pocket factor to be a dynamic process, with the virus retaining its functionality and, hence, overall structure upon the exchange. In this sense, our simulations do not contradict the experiments.

In addition, the pulling distance has effect on binding affinity estimates of systems containing the Kirtan1 molecule. The effects of both pulling velocity and distance could be in part due to insufficient statistics related to the analyses of these effects.

Nevertheless, the qualitative trends are reproduced consistently for different simulation parameters, such that the natural pocket factor palmitic acid shows the lowest affinity. The known drug molecules pleconaril and its derivative Kirtan1, sharing similar core structures, show similar affinities that are both higher than that of palmitic acid. Conceptually a molecule with lower binding affinity can be replaced by one with higher binding affinity. In this sense, our results are in qualitative agreement with recent experiments[3] showing that palmitic acid is replaceable by pleconaril or Kirtan1 molecules. Furthermore, the experiments[3] show decreased melting temperatures of virus capsids with hydrophobic pockets occupied by palmitic acid, in comparison to capsids with pockets occupied by pleconaril or pleconaril-based molecules while retaining infectivity. This indicates the higher stability of capsids containing drug-based molecules in the hydrophobic pockets and, hence, higher binding affinities. Our calculations are also in agreement with this trend. Therefore, despite the over-estimation, the simulations give qualitatively relevant results, and similar simulation procedures could be utilized in estimating the binding affinities of new molecules in this virus system.

After the gold nanoclusters were attached to Kirtan1 for the purposes of imaging, the highest binding affinity is consistently exhibited by Kirtan1 conjugated to half-protonated and -deprotonated gold nanoclusters, while with the fully deprotonated nanoclusters, the affinity is more similar to that of pure Kirtan1. This suggests that the affinity is sensitive to the protonation state of the gold nanocluster, i.e., pH conditions.

Comparing the binding affinity estimate trends to the numbers of interactions, in addition to hydrogen bond interactions, hydrophobic and ring contacts are all expected to play a crucial role in determining the binding affinities of these molecules. The simple model of correlating numbers of interactions and binding-affinity estimates could be taken into consideration when designing molecules for purposes similar to those described here.

Conjugating Kirtan1 to the gold nanocluster in either of the protonation states does not seem to significantly decrease its affinity to the virus. Our results therefore suggest that using noncovalent binding to virus pockets schemes are a viable route to labeling viruses with gold nanoparticles. The consistency of the trends among various conditions suggest that simulations could be used to predict binding affinity of new pocket factors with respect to the natural pocket factor palmitic acid. This could be utilized in the design of new pocket factors.

In this work, we have only considered the labeling of the virus with gold nanoclusters in terms of binding affinities and large-scale conformational changes, not focusing on any possible atomistic scale structural changes that may arise as a result of this labeling. A more-detailed study of these possible

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

structural changes and interactions would be important in further assessing the effects of the gold nanoclusters on the structure and functions of the virus.

METHODS

In this work, we performed all-atom MD simulations for five systems consisting of a full model of EV1 capsid (without the genome) with all of its 60 hydrophobic pockets occupied by one of the three different pocket factors: (1) natural pocket factor palmitic acid, (2) pleconaril, and (3) Kirtan1. In addition, two systems with (4) Kirtan1 linked to $Au_{102}pMBA_{44}$ gold nanocluster in a half-protonated and -deprotonated state and (5) in a fully deprotonated state were simulated. These simulations mimic the experiments in ref 3. We have estimated the binding affinities of the different pocket factors from the work distributions of simultaneously pulling out the 60 pocket factors from the pockets using steered molecular dynamics simulations and Jarzynski equality with block averaging and an extrapolation analysis scheme suggested by Zuckerman and co-workers.[40,41]

Construction of the Virus−Pocket Factor Systems. The structure of the virus protein was obtained from the Protein databank with a PDB ID 1EV1 protomer structure for echovirus 1[45] containing proteins VP1−VP4 as well as two different fatty acids: myristic acid bonded to the VP4 chain and palmitic acid located in the hydrophobic pocket of VP1. The missing amino acid residues from the VP4 chain were added to the structure using Modeler 9.11 loop builder[46,47] and manual alignment.

The force-field parameters for myristic and palmitic acids were obtained by using suitable available parameters of the Amber99sb-ildn[48] force field. The partial charges assuming deprotonated acids were optimized following the RESP charge fitting procedure recommended for Amber.[49] The geometry optimization and ESP calculations according to the Merz−Singh−Kollman scheme[50,51] were performed with Gaussian09[52] at a HF/6-31F* level of theory. Atomic charges were fitted to the obtained potential in a two-stage RESP fit procedure with Ambertools12[53] and are included in the Supporting Information.

The full structure with 60 protomer units was constructed with visual molecular dynamics (VMD)[54] based on PDB BIOMT records and the available mono2poly script. In constructing the Gromacs topology, protonations for all amino acid residues except those of histidines were assigned on the basis of their intrinsic pK_a value and a pH of 7. Histidines were protonated based on visualizing the possibilities of hydrogen bond formations in one protomer unit. Replacement of the natural pocket factor palmitic acid by pleconaril and Kirtan1 was done by fitting the structures onto the palmitic acid position in the original crystal structure using VMD[54] and the Molefacture tool.

Due to the higher complexity of pleconaril and Kirtan1 structures (Figure 1) compared to palmitic acid, the force-field parameters were not obtained from the readily available parameters but derived using an AmberTools12[53] Antechamber program to obtain the generalized Amber force field parameters.[55] The Gromacs topologies were obtained using ACPYPE code (emulating amb2gmx).[56] The charge derivations were performed similarly to palmitic and myristic acids and are also included in the Supporting Information.

Linking a $Au_{102}pMBA_{44}$ cluster to the Kirtan1 molecule was done manually using VMD[54] by connecting the terminal C of the linker tail of Kirtan1 to the terminal O atom of the *p*MBA group. The *p*MBA that connects to the linker was selected to correspond to one of the reported *p*MBA groups most likely to undergo ligand exchange reactions in the $Au_{102}pMBA_{44}$ nanocluster.[57] The force field bonded parameters for the aliphatic linker were selected on the basis of standard Amber alkane parameters of similar interactions. To enable linking, one H atom bonded to the terminal C atom of Kirtan1 was removed. The charge of the removed H atom was evenly distributed to the remaining atoms to conserve the overall neutral charge of the molecule. The force-field parameters for $Au_{102}pMBA_{44}$ were described in a previous paper.[58]

Reference Molecular Dynamics Simulations. Prior to the steered pulling simulations, 90−120 ns molecular dynamics simulations were performed (later referred as the reference simulations) for all five systems to test the simulation setup and to obtain an ensemble of statistically independent starting configurations for the pulling simulations. Before the reference simulations, energy minimizations and short equilibrations were performed. For the systems without the gold nanoclusters, the equilibration procedure consisted of 50 ps NVT (NVT = constant number of atoms, volume and temperature) at 200 K plus 100 ps NVT at 300 K plus 100 ps NPT (NPT = constant number of atoms, pressure and temperature) at 300 K at 1 bar, with position restraints on the protein and pocket factors. For the systems with gold nanoclusters, we used first 1 ns of NPT (300 K and 1 bar) annealing with position restraints, followed by 1 ns of unrestrained NPT at 300 K and 1 bar. Each system was simulated in dodecahedral simulation box solvated with TIP3P water[59] and sodium and chloride ions,[60] which were added to neutralize each system at a concentration of 0.1−0.15 M.

For the reference simulations, we used Gromacs versions 4.5 and 4.6[61,62] and the Amber99sb-ildn force field[48] with added parameters for thiolate-protected gold nanoclusters,[58] myristic and palmitic acids, pleconaril and Kirtan1. In the simulations, we used a leapfrog integrator with a 2 fs time step, periodic boundary conditions, a 1.0 nm Lennard−Jones cutoff with dispersion correction for energy and pressure, PME (PME = particle mesh Ewald method) electrostatics with a 1.0 nm cutoff and 0.12 nm grid spacing,[63] a velocity-rescale thermostat with a reference temperature of 300 K and a coupling time constant of 0.1 ps,[64] and the Berendsen barostat with a reference pressure of 1 bar and coupling time constant of 1 ps.[65] For improved performance, the lengths of covalent bonds containing hydrogens were constrained with the LINCS algorithm[66] for the systems not containing gold nanoclusters. For the gold-nanocluster-containing systems, constraints were applied on all bond lengths.

Pulling Simulations. In the pulling simulations, we have the center-of-mass pulling code in Gromacs 5.0.4[67] with a harmonic potential with a force constant of 10 000 kJ/mol nm ^{2} and a pulling velocity of 1 nm/ns. Comparable force constants and pulling velocities have been used in previous works.[27,68,69] The carbon atom on which the pulling force was exerted (indicated in Figure 1) was selected on the basis of initial tests. A total of three different pulling groups, carbons at the ends of the molecule and the center of mass of the full molecule, were tested for the pulling of palmitic acid. The selection of the pulling group was made based on the smallest average work obtained from one pulling simulation of 60 pocket factors. The pulling direction for each pocket factor was taken as the vector connecting the center of mass of the pocket factor and virus

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

measured from the averaged structure of the reference simulation. For CLUKIR and DCLUKIR systems, the same directions were used as for KIR system.

Each pulling simulation started from a snapshot taken from the reference simulation with random velocities assigned to the atoms. All 60 pocket factors were pulled out simultaneously during a 4.5 ns simulation. Visual inspection confirmed externalization of all pocket factors. Simulation settings were otherwise identical to those of the reference simulations, but constraints were applied on hydrogen bond lengths in all systems. Forces obtained from the simulations were integrated to obtain the work for pulling each pocket factor from its initial position inside the pocket to the final position outside the pocket. A total of 25 pulling simulations were performed for each system, constituting a total 1500 work values for the free-energy estimations.

We also analyzed the effect of pulling out the pocket factor on the conformational dynamics of the virus by comparing the reference simulation of the virus to the situation after the pocket factors have been pulled out. To investigate the effect of pulling on the structure of the hydrophobic pocket and its proximity, the RMSD of one representative pulling trajectory and of the reference trajectory extended by 4.5 ns (length of pulling trajectory) without pulling were compared for all systems. Protein residues located within a 1.5 nm spherical volume of the pocket factor at the starting configuration were included in the analysis.

To better understand the origin of the binding affinities and build a simple scoring function for later use, we investigated the interactions in one representative pulling trajectory for each system in more detail by simple distance criteria. The hydrogen bonds were analyzed using Gromacs' g_hbond tool with a 0.35 nm and 30° cutoff radius and angle. Hydrophobic contacts were counted as the number of hydrophobic protein residues having an atom closer than 0.7 nm to the center of mass of a pocket factor. Aromatic ring contacts were counted similarly, but only atoms in the rings were included, and instead of centers of mass of the full pocket factors, centers of mass of the rings were used in the distance criterion (for PLE, KIR, CLUKIR, and DCLUKIR systems). For the CLUKIR and DCLUKIR systems, the interactions of the *p*MBA groups were also considered. To illustrate the effects of these interactions in the most simple manner, their relative strengths were estimated based on counted contacts by assigning weights HB_w, HP_w, and R_w to best fit to $HB_w \times HB + HP_w \times HP + R_w \times R = \Delta G_i$, where HB, HP, and R denote the (overall broken and formed) number of hydrogen bonds, hydrophobic contacts, and ring contacts, respectively, and ΔG_i is the binding affinity estimate of each system.

Free-Energy Estimation. The free energy or binding affinity ΔG can be estimated from the non-equilibrium pulling work values W with the Jarzynski equality:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta G} \quad (1)$$

where $\beta = \frac{1}{k_B T}$

Because direct estimation is biased, corrections are needed. Due to the non-Gaussian nature of our work distributions and technical difficulties with using forward−reverse simulations, we utilized block averaging and extrapolation analyses methods as introduced by Zuckerman and co-workers in refs 40 and 41. The full set of $N = 1500$ work values was resampled by taking $n = 1 \cdots N$ unique work values and calculating ΔG from eq 1. The procedure was repeated for each n a total of N/n times. The averaged direct Jarzynski estimate ΔG is then obtained as a smooth function of work values in the block.

To extrapolate beyond $n = 1 \cdots N$ work values, the conversion $\chi = n^{-\tau}$ is performed. Extrapolations can then be applied to the χ, average ΔG data to obtain the ΔG estimate for large number of work values. Following the procedure of Zuckerman et al., τ was fixed to 0.5 for linear and quadratic fits, and three alternative extrapolated estimates were obtained by:

(1) performing a linear fit to the 1/5 tail of the ΔG, χ data and
(2) performing a quadratic fit to the full ΔG, χ data.

Estimate (3) was obtained by utilizing cumulative integral (CI) extrapolation, as presented in ref 41. With the optimal value of τ, the CI values converge for the small χ (1/5) tail, and the best estimate is the cumulative integral value at minimum χ, corresponding to value of the optimized τ for which the slope of the 1/5 tail is closest to 0. The block averaging and extrapolations were repeated 500 times to obtain averages and standard deviations for each estimate.

Factors Affecting the Free-Energy Estimate. To estimate the effects of different simulation parameters and initial conditions, we performed pulling simulations using various pulling velocities, pull groups, distances, number of pocket factors being pulled simultaneously, and protonation state of the gold nanocluster. In addition, we also tested a possible dependency on the starting snapshot and inspected the end structures to confirm that the pulling had not disrupted the protein or the pocket factors. The statistics of this section are often from one pulling simulation only, which for the full system equals $N = 60$ pulling work values for the analyses. In case of data sets with $N < 60$, block averaging and extrapolations were not performed, but only the work distributions and direct Jarzynski estimates are compared.

To verify that there is no initial bias originating from the reference simulations, we studied the dependency of the binding affinity estimate on the reference trajectory snapshot, from which the pulling was started, by comparing the free-energy estimates from pulling simulations initiated from 4 independent reference simulation snapshots between 26 and 114 ns for all 5 systems.

Dependency of binding affinity on pulling velocity can be considered to originate from friction (linear dependence on pulling velocity, dominating with pulling velocities of above approximately 1 m/s (=1 nm/ns)) and activated processes (logarithmic dependence on pulling velocity, dominating for slower velocities).[70] To address the effect of pulling velocity on binding affinity estimates, additional pulling simulations corresponding to approximately 4.5 nm pulling were performed with pulling velocities of 10, 0.5, and 0.25 nm/ns for the PLT, KIR, and CLUKIR systems. For the PLE system additional pulling simulations with velocities of 0.8, 2.0, 3.0, 5.0, and 8.0 nm/ns were performed. The results are compared to the estimate with 1 nm/ns velocity started from the same snapshot of the reference trajectory.

By simultaneously pulling all 60 pocket factors, we neglect cooperative effects, if any. Whether such cooperative effects exist, was tested for PLT, PLE, KIR, and CLUKIR systems by performing pulling simulations with only a subset of the pocket factors being pulled out simultaneously:

(1) all pocket factors of one pentamer unit (5 simultaneous pullings, $N = 5$, denoted penta1),

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

(2) one symmetrically equivalent pocket factor of each pentamer unit in the whole virus capsid (N = 12, denoted penta12),
(3) random pocket factor from each pentamer unit (N = 12, denoted random12), and
(4) one pocket factor (N = 1, denoted 1(group number)) at a time, for two different groups.

Free-energy estimates were compared with the free-energy estimates when all 60 pocket factors were pulled out simultaneously. Because the number of pullings was significantly smaller, only the averages and standard deviations of the work values and the direct Jarzynski estimate are compared.

For CLUKIR and DCLUKIR systems, the pulling time of 4.5 ns results in the externalization of the pocket factors, but the gold nanoclusters remain close to the virus surface. This might cancel the effects of the gold nanoclusters on the solubility of the Kirtan1 molecule. To estimate the effect of the gold nanoparticle on the overall solubility of the linker−gold nanoparticle construct, the pulling was extended to 6 ns for systems KIR, CLUKIR, and DCLUKIR. Further extensions require a larger simulation box. To avoid the steep increase in computational costs, these simulations were performed on a subsystem containing one isolated pentamer unit, 1/12 of the full virus. A series of 15 ns pulling simulations for the KIR, CLUKIR, and DCLUKIR subsystems were performed to achieve the clear detaching of gold clusters from the virus surface. Because the number of work values is lower than 60, only the averages and standard deviations of the work values and direct Jarzynski estimate are compared.

In the first set of pulling simulations, the protonation state of the gold nanocluster *p*MBA groups was selected to be half protonated and half deprotonated (with a charge of −22e per cluster), reasonable for physiological pH conditions[71] and the protonation−deprotonation pattern obtained from constant pH MD simulations.[71] To study the effect of difference in the protonation state, we performed a 120 ns reference simulation and 25 pulling simulations using clusters with all *p*MBA groups deprotonated (DCLUKIR, charge of −44e per cluster).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.bioconjchem.7b00367.

Conformation analysis of pocket factors in pulling and free simulations. (PDF)
Gromacs topologies (force-field parameters) for myristic acid. (TXT)
Gromacs topologies for palmitic acid. (TXT)
Gromacs topologies for pleconaril. (TXT)
Gromacs topologies for Kirtan1. (TXT)
Gromacs topologies for Kirtan1 connected to gold nanoclusters in the CLUKIR protonation state. (TXT)
Gromacs topologies for Kirtan1 connected to gold nanoclusters in the DCLUKIR protonation state. (TXT)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: gerrit.x.groenhof@jyu.fi.
*E-mail: hannu.j.hakkinen@jyu.fi.

ORCID

Hannu Häkkinen: 0000-0002-8558-5436

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grant nos. 258806, 294217, and 292352 from the Academy of Finland. E.P. acknowledges a postgraduate study grant from the Vilho, Yrjö, and Kalle Väisälä Foundation. The computations were done at the CSC, the Finnish IT Center for Science (Grand Challenge project EVLIGANDS), at the Nanoscience Center of the University of Jyväskylä by utilizing the Finnish Grid and Cloud Infrastructure (FGCI; persistent identifier urn:nbn:fi:research-infras-2016072533), and at the HLRS High-Performance Computing Center, Stuttgart. We thank V. Marjomäki and M. Martikainen for numerous discussions on the experimental aspects of gold−virus binding and A. Ruokonen for preliminary modeling work on hydrophobic pocket factors.

■ ABBREVIATIONS

EV1, echovirus 1; MD, molecular dynamics; *p*MBA, para-mercaptobenzoic acid; TEM, transmission electron microscopy; PLT, palmitic acid; PLE, pleconaril; KIR, Kirtan1; CLUKIR, Kirtan1 connected to 50% deprotonated $Au_{102}pMBA_{44}$ gold nanocluster; DCLUKIR, Kirtan1 connected to 100% deprotonated $Au_{102}pMBA_{44}$ gold nanocluster; JE, direct Jarzynski estimate; LE, linear extrapolation; QE, quadratic extrapolation; CI, cumulative integral extrapolation; HB, hydrogen bonds; HP, hydrophobic contacts; R, ring contacts; penta1, one full pentameric subunit of EV1 with total of 5 pocket factors; penta12, one symmetrically positioned unit of all pentameric subunits of EV1 with total of 12 pocket factors; random12, one randomly positioned unit of all pentameric subunits of EV1 with total of 12 pocket factors; 1(21), one pocket factor number 21; 1(59), one pocket factor number 59

■ REFERENCES

(1) Tsukuda, T., and Häkkinen, H. (2015) *Protected Metal Clusters: From Fundamentals to Applications*, Elsevier, Amsterdam, The Netherlands.
(2) Marjomäki, V., Lahtinen, T., Martikainen, M., Koivisto, J., Malola, S., Salorinne, K., Pettersson, M., and Häkkinen, H. (2014) Site-specific targeting of enterovirus capsid by functionalized monodisperse gold nanoclusters. *Proc. Natl. Acad. Sci. U. S. A. 111*, 1277−1281.
(3) Martikainen, M., Salorinne, K., Lahtinen, T., Malola, S., Permi, P., Häkkinen, H., and Marjomäki, V. (2015) Hydrophobic pocket targeting probes for enteroviruses. *Nanoscale 7*, 17457−17467.
(4) Zhang, W., Ye, J., Zhang, Y., Li, Q., Dong, X., Jiang, H., and Wang, X. (2015) One-step facile synthesis of fluorescent gold nanoclusters for rapid bio-imaging of cancer cells and small animals. *RSC Adv. 5*, 63821−63826.
(5) Chattoraj, S., Amin, M. A., Mohapatra, S., Ghosh, S., and Bhattacharyya, K. (2016) Cancer Cell Imaging Using in Situ Generated Gold Nanoclusters. *ChemPhysChem 17*, 61−68.
(6) Sönnichsen, C., Reinhard, B. M., Liphardt, J., and Alivisatos, A. P. (2005) A molecular ruler based on plasmon coupling of single gold and silver nanoparticles. *Nat. Biotechnol. 23*, 741−45.
(7) Mathew-Fenn, R. S., Das, R., Silverman, J. A., Walker, P. A., and Harbury, P. A. B. (2008) A Molecular Ruler for Measuring Quantitative Distance Distributions. *PLoS One 3*, e3229.
(8) Reinhard, B. M., Siu, M., Agarwal, H., Alivisatos, A. P., and Liphardt, J. (2005) Calibration of Dynamic Molecular Rulers Based on

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

Plasmon Coupling between Gold Nanoparticles. *Nano Lett. 5*, 2246−2252.
(9) Si, Y., Sun, Z., Zhang, N., Qi, W., Li, S., Chen, L., and Wang, H. (2014) Ultrasensitive Electroanalysis of Low-Level Free MicroRNAs in Blood by Maximum Signal Amplification of Catalytic Silver Deposition Using Alkaline Phosphatase-Incorporated Gold Nanoclusters. *Anal. Chem. 86*, 10406−10414.
(10) Zhang, N., Si, Y., Sun, Z., Chen, L., Li, R., Qiao, Y., and Wang, H. (2014) Rapid, Selective, and Ultrasensitive Fluorimetric Analysis of Mercury and Copper Levels in Blood Using Bimetallic Gold-Silver Nanoclusters with "Silver Effect"-Enhanced Red Fluorescence. *Anal. Chem. 86*, 11714−11721.
(11) Zhang, L., Li, S., Dong, M., Jiang, Y., Li, R., Zhang, S., Lv, X., Chen, L., and Wang, H. (2017) Reconstituting redox active centers of heme-containing proteins with biomineralized gold toward peroxidase mimics with strong intrinsic catalysis and electrocatalysis for H2O2 detection. *Biosens. Bioelectron. 87*, 1036−1043.
(12) Van Lehn, R. C., Ricci, M., Silva, P. H. J., Andreozzi, P., Reguera, J., Voïtchovsky, K., Stellacci, F., and Alexander-Katz, A. (2014) Lipid tail protrusions mediate the insertion of nanoparticles into model cell membranes. *Nat. Commun. 5*, 4482.
(13) Van Lehn, R. C., and Alexander-Katz, A. (2015) Pathway for insertion of amphiphilic nanoparticles into defect-free lipid bilayers from atomistic molecular dynamics simulations. *Soft Matter 11*, 3165−3175.
(14) Van Lehn, R. C., and Alexander-Katz, A. (2014) Membrane-Embedded Nanoparticles Induce Lipid Rearrangements Similar to Those Exhibited by Biological Membrane Proteins. *J. Phys. Chem. B 118*, 12586−12598.
(15) Heikkilä, E., Martinez-Seara, H., Gurtovenko, A. A., Vattulainen, I., and Akola, J. (2014) Atomistic simulations of anionic Au144(SR)60 nanoparticles interacting with asymmetric model lipid membranes. *Biochim. Biophys. Acta, Biomembr. 1838*, 2852−2860.
(16) Rossi, G., and Monticelli, L. (2016) Gold nanoparticles in model biological membranes: A computational perspective. *Biochim. Biophys. Acta, Biomembr. 1858*, 2380−2389.
(17) Rossi, G., and Monticelli, L. (2016) Simulating the interaction of lipid membranes with polymer and ligand-coated nanoparticles. *Adv. Phys.: X 1*, 276−296.
(18) Simonelli, F., Bochicchio, D., Ferrando, R., and Rossi, G. (2015) Monolayer-Protected Anionic Au Nanoparticles Walk into Lipid Membranes Step by Step. *J. Phys. Chem. Lett. 6*, 3175−3179.
(19) Mudedla, S. K., Azhagiya Singam, E. R., Balamurugan, K., and Subramanian, V. (2015) Influence of the size and charge of gold nanoclusters on complexation with siRNA: a molecular dynamics simulation study. *Phys. Chem. Chem. Phys. 17*, 30307−30317.
(20) Krupa, P., Mozolewska, M. A., Rasulev, B., Czaplewski, C., and Leszczynski, J. (2014) Towards Mechanisms of Nanotoxicity-Interaction of Gold Nanoparticles with Proteins and DNA. *Task Quarterly 18*, 337−341.
(21) Brancolini, G., Toroz, D., and Corni, S. (2014) Can small hydrophobic gold nanoparticles inhibit [small beta]2-microglobulin fibrillation? *Nanoscale 6*, 7903−7911.
(22) Jarzynski, C. (1997) Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett. 78*, 2690−2693.
(23) Maruthamuthu, V., Schulten, K., and Leckband, D. (2009) Elasticity and Rupture of a Multi-Domain Neural Cell Adhesion Molecule Complex. *Biophys. J. 96*, 3005−3014.
(24) Zink, M., and Grubmüller, H. (2010) Primary Changes of the Mechanical Properties of Southern Bean Mosaic Virus upon Calcium Removal. *Biophys. J. 98*, 687−695.
(25) Arkhipov, A., Roos, W. H., Wuite, G. J., and Schulten, K. (2009) Elucidating the Mechanism behind Irreversible Deformation of Viral Capsids. *Biophys. J. 97*, 2061−2069.
(26) Wang, Q., Xue, T., Song, C., Wang, Y., and Chen, G. (2016) Study on the Application of the Combination of TMD Simulation and Umbrella Sampling in PMF Calculation for Molecular Conformational Transitions. *Int. J. Mol. Sci. 17*, 692.
(27) Park, S., Khalili-Araghi, F., Tajkhorshid, E., and Schulten, K. (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys. 119*, 3559−3566.
(28) Chen, W., Lou, J., Hsin, J., Schulten, K., Harvey, S. C., and Zhu, C. (2011) Molecular Dynamics Simulations of Forced Unbending of Integrin $\alpha_V \beta_3$. *PLoS Comput. Biol. 7*, 1−13.
(29) Kappel, C., and Grubmüller, H. (2011) Velocity-Dependent Mechanical Unfolding of Bacteriorhodopsin Is Governed by a Dynamic Interaction Network. *Biophys. J. 100*, 1109−1119.
(30) Åqvist, J., and Luzhkov, V. (2000) Ion permeation mechanism of the potassium channel. *Nature 404*, 881−884.
(31) De Fabritiis, G., Coveney, P. V., and Villà-Freixa, J. (2008) Energetics of K+ permeability through Gramicidin A by forward-reverse steered molecular dynamics. *Proteins: Struct., Funct., Genet. 73*, 185−194.
(32) Giorgino, T., and De Fabritiis, G. (2011) A High-Throughput Steered Molecular Dynamics Study on the Free Energy Profile of Ion Permeation through Gramicidin A. *J. Chem. Theory Comput. 7*, 1943−1950.
(33) Goh, B. C., Wu, H., Rynkiewicz, M. J., Schulten, K., Seaton, B. A., and McCormack, F. X. (2016) Elucidation of Lipid Binding Sites on Lung Surfactant Protein A Using X-ray Crystallography, Mutagenesis, and Molecular Dynamics Simulations. *Biochemistry 55*, 3692−3701.
(34) Ytreberg, F. M. (2009) Absolute FKBP binding affinities obtained via nonequilibrium unbinding simulations. *J. Chem. Phys. 130*, 164906.
(35) Bu, L., Beckham, G. T., Shirts, M. R., Nimlos, M. R., Adney, W. S., Himmel, M. E., and Crowley, M. F. (2011) Probing Carbohydrate Product Expulsion from a Processive Cellulase with Multiple Absolute Binding Free Energy Methods. *J. Biol. Chem. 286*, 18161−18169.
(36) Alvarez, H. A., McCarthy, A. N., and Grigera, J. R. (2012) A Molecular Dynamics Approach to Ligand-Receptor Interaction in the Aspirin-Human Serum Albumin Complex. *J. Biophys. 2012*, 642745.
(37) Zhang, J.-L., Zheng, Q.-C., Li, Z.-Q., and Zhang, H.-X. (2012) Molecular Dynamics Simulations Suggest Ligand's Binding to Nicotinamidase/Pyrazinamidase. *PLoS One 7*, 1−12.
(38) Lemkul, J. A., and Bevan, D. R. (2010) Assessing the Stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics. *J. Phys. Chem. B 114*, 1652−1660.
(39) Hummer, G. (2001) Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys. 114*, 7330−7337.
(40) Zuckerman, D. M., and Woolf, T. B. (2002) Overcoming finite-sampling errors in fast-switching free-energy estimates: extrapolative analysis of a molecular system. *Chem. Phys. Lett. 351*, 445−453.
(41) Ytreberg, F. M., and Zuckerman, D. M. (2004) Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems. *J. Comput. Chem. 25*, 1749−1759.
(42) Crooks, G. E. (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top. 60*, 2721−2726.
(43) Lu, N., Kofke, D. A., and Woolf, T. B. (2004) Improving the efficiency and reliability of free energy perturbation calculations using overlap sampling methods. *J. Comput. Chem. 25*, 28−40.
(44) Green, N. M. (1975) Avidin. *Adv. Protein Chem. 29*, 85−133.
(45) Filman, D., Wien, M., Cunningham, J., Bergelson, J., and Hogle, J. (1998) Structure determination of echovirus 1. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 1261−1272.
(46) Šali, A., and Blundell, T. L. (1993) Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol. 234*, 779−815.
(47) Fiser, A., Do, R. K. G., and Šali, A. (2000) Modeling of loops in protein structures. *Protein Sci. 9*, 1753−1773.
(48) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet. 78*, 1950−1958.

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339

(49) Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem. 97*, 10269−10280.
(50) Singh, U. C., and Kollman, P. A. (1984) An approach to computing electrostatic charges for molecules. *J. Comput. Chem. 5*, 129−145.
(51) Besler, B. H., Merz, K. M., and Kollman, P. A. (1990) Atomic charges derived from semiempirical methods. *J. Comput. Chem. 11*, 431−439.
(52) Frisch, M. J. et al. (2009) *Gaussian09*, Revision D.01, Gaussian Inc., Wallingford CT.
(53) Case, D. A. et al. (2012) *AMBER 12*; University of California, San Francisco, CA.
(54) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J. Mol. Graphics 14*, 33−38.
(55) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157−1174.
(56) Sousa da Silva, A. W., and Vranken, W. F. (2012) ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes 5*, 1−8.
(57) Heinecke, C. L., Ni, T. W., Malola, S., Mäkinen, V., Wong, O. A., Häkkinen, H., and Ackerson, C. J. (2012) Structural and Theoretical Basis for Ligand Exchange on Thiolate Monolayer Protected Gold Nanoclusters. *J. Am. Chem. Soc. 134*, 13316−13322.
(58) Pohjolainen, E., Chen, X., Malola, S., Groenhof, G., and Häkkinen, H. (2016) A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *J. Chem. Theory Comput. 12*, 1342−1350.
(59) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926.
(60) Aqvist, J. (1990) Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem. 94*, 8021−8024.
(61) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput. 4*, 435−447.
(62) Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics 29*, 845−854.
(63) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(64) Bussi, G., and Parrinello, M. (2008) Stochastic thermostats: comparison of local and global schemes. *Comput. Phys. Commun. 179*, 26−29.
(65) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(66) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.
(67) Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1−2*, 19−25.
(68) Marklund, E. G., Mahmutovic, A., Berg, O. G., Hammar, P., van der Spoel, D., Fange, D., and Elf, J. (2013) Transcription-factor binding and sliding on DNA studied using micro- and macroscopic models. *Proc. Natl. Acad. Sci. U. S. A. 110*, 19796−19801.
(69) Aprodu, I., Redaelli, A., and Soncini, M. (2008) Actomyosin Interaction: Mechanical and Energetic Properties in Different Nucleotide Binding States. *Int. J. Mol. Sci. 9*, 1927−1943.
(70) Heymann, B., and Grubmüller, H. (1999) AN02/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chem. Phys. Lett. 303*, 1−9.
(71) Koivisto, J., Chen, X., Donnini, S., Lahtinen, T., Häkkinen, H., Groenhof, G., and Pettersson, M. (2016) Acid-Base Properties and Surface Charge Distribution of the Water-Soluble Au102(pMBA)44 Nanocluster. *J. Phys. Chem. C 120*, 10041−10050.

DOI: 10.1021/acs.bioconjchem.7b00367
Bioconjugate Chem. 2017, 28, 2327−2339