Dalton Transactions

Towards Elucidating Structure of Ligand-protected Nanoclusters

Journal:	*Dalton Transactions*
Manuscript ID	DT-FRO-04-2020-001418.R1
Article Type:	Frontier
Date Submitted by the Author:	27-May-2020
Complete List of Authors:	Cowan, Michael; University of Pittsburgh, Chemical and Petroleum Engineering Mpourmpakis, Giannis; University of Pittsburgh, Chemical and Petroleum Engineering

Towards Elucidating Structure of Ligand-protected Nanoclusters

Michael J. Cowan and Giannis Mpourmpakis*

Department of Chemical and Petroleum Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA

*Corresponding author. Email: gmpourmp@pitt.edu

ABSTRACT

Ligand-protected metal nanoclusters (NCs) are organic-inorganic nanostructures, exhibiting high stability at specific “magic size” compositions and tunable properties that make them promising candidates for a wide range of nanotechnology-based applications. Synthesis and characterization of these nanostructures has been achieved with atomic precision, offering great opportunities to study the origin of new physicochemical property emergence at the nanoscale using theory and computation. In this *Frontiers* article, we highlight the recent advances in the field of ligand-protected metal NCs, focusing on stability theories on monometallic and heterometal doped NCs, and NC structure prediction. Furthermore, we discuss current challenges on predicting previously undiscovered NCs and propose future steps to advance the field through applying first principles calculations, machine learning, and data-science-based approaches.

INTRODUCTION

Ligand-protected metal nanoclusters (NCs) are a unique class of small (tens to a few hundred metal atoms), atomically precise nanomaterials that have attracted tremendous interest in recent years.[1] Due to their size regime, NCs can exhibit molecular-like photophysical[2, 3] and solubility[4] properties, which differentiate them from larger metal nanoparticles.[5] Furthermore, with a number of promising physicochemical properties, NCs have found use in many applications over a broad range of fields. For instance, the luminescent properties of NCs have opened avenues for their use in detecting biomolecules[6], as well as imaging cancer cells[7, 8] and bacteria.[9] Additionally, NCs have emerged as efficient and selective catalysts owing to their high surface-to-volume ratio and discrete electronic states (i.e. molecular-like highest occupied – lowest unoccupied molecular orbital (HOMO-LUMO) gap rather than metallic character).[10] NCs can catalyze a variety of reactions, including among others the hydrogenation of nitrobenzaldehyde,[11] photocatalytic degradation of organic pollutants,[12] and the electrocatalytic reduction of CO_2.[13-16] With the ever-growing possibilities for practical nanotechnological applications, research interest for NCs will continue to expand.

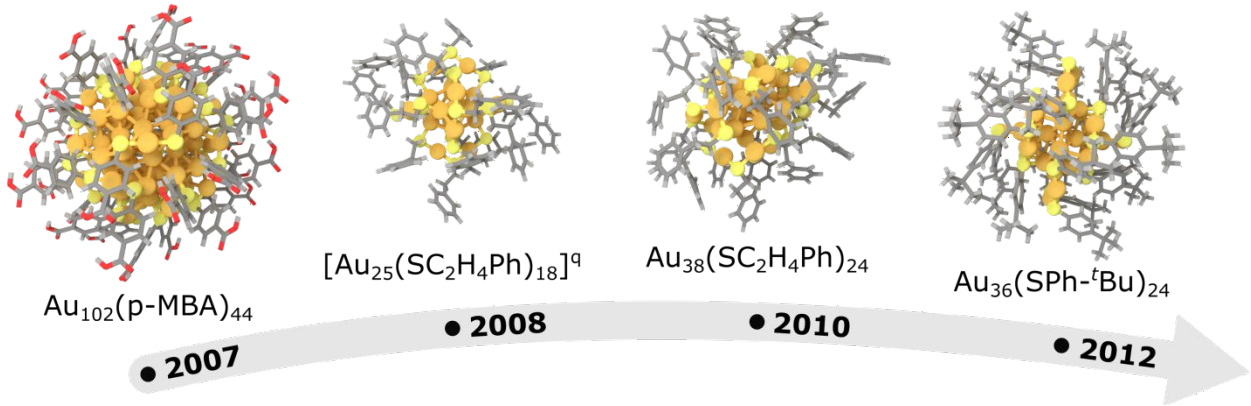


Figure 1. Timeline scheme of the first thiolate-protected Au NC structures determined experimentally. Both the anionic (q = -1) and neutral (q = 0) $Au_{25}(SC_2H_4)_{18}$ structures were determined in 2008. Gold and yellow balls represent Au and S atoms, respectively. Red, gray, and light gray sticks represent O, C, and H atoms of the organic ligands, respectively.

One of the most popular synthesis methods of ligand-protected NCs was introduced by Brust *et al.* in 1994, where Au salts were reduced in the presence of organic thiols and a strong reducing agent ($NaBH_4$).[17] The synthesis of thiolate-protected Au nanoparticles was achieved exhibiting a distribution of diameters ranging 1-3 nm, but structural characterization with atomic precision was still lacking. Over the next decade, improvements were made to the synthesis process such that Au NCs with exact molecular weights were discovered using separation methods and mass spectrometry.[18-20] True atomic precision, however, was not achieved until 2007, where Jadzinsky and Calero *et al.* were able to characterize the exact atomic positions of Au_{102}(p-MBA)$_{44}$ (p-MBA = para-mercaptobenzoic acid) within an unprecedented 1.1 Å resolution through single crystal X-ray diffraction.[21] With this pioneering work, the discovery of additional NC structures followed soon after, including both the anionic[22, 23] and neutral[24] $Au_{25}(SC_2H_4Ph)_{18}$ in 2008, the

$Au_{38}(SC_2H_4Ph)_{24}$ in 2010,[25] and the $Au_{36}(SPh\text{-}^tBu)_{24}$ in 2012 (Figure 1).[26] Today, the number of known experimentally determined Au NCs has exploded.[27]

Ligand-protected NCs are also referred to as "magic size" NCs, since they exhibit high stability at specific compositions (i.e. at specific *n* and *m* of $Au_n(SR)_m$ structures).[28] This emergence of magic sizes was primarily due to the development of a size focusing synthesis method.[29] According to the size focusing method, after the initial formation of polydisperse Au NCs, the solution is exposed to "harsh" conditions, such as excess thiol concentration and elevated temperatures. This leads to only the most stable NC surviving in a monodisperse environment, thus, "focusing" the NC distribution to a single NC structure. Additional methods to synthesize new magic sizes have also been developed. These include ligand-based approaches to control NC size[30] and NC-NC transformations through ligand exchange.[31, 32] Furthermore, significant research has been done to introduce heterometals into Au NCs, forming a new class of alloy NC derivatives with distinct properties.[33 34] Alloy NCs can be synthesized by heterometal doping Au NCs to form analogues of their monometallic counterparts,[35-37] or even entirely new structures.[38-40] With the continuous advancement of synthesis and post-synthetic manipulation methods, there is practically no end in sight to the discovery of new, atomically precise, ligand-protected NCs.

The experimental discovery of ligand-protected NCs and determination of their structure with atomic precision has undoubtedly generated tremendous interest in the field and as such, a number of important questions emerge: What makes ligand-protected NCs stable at very specific, magic sizes? How does heterometal doping affect overall NC stability? How can we predict mono- and bimetallic NCs of magic size that have yet been experimentally determined? Herein, we summarize the current state of NC research on understanding NC stability, as well as expanding NC materials space through heterometal doping. We also focus on the current progress and challenges of NC structure prediction and propose potential future steps.

UNDERSTANDING THE ORIGIN OF MAGIC SIZE STABILITY

Experimental advances in NC synthesis and characterization have provided nanostructures with atomic level precision to theory. In addition, developments in theory and increase in computational power have enabled the investigation of large systems, such as ligand-protected NCs, with accurate, first principles methods.[4, 41] As a result, the combination of both ends has led to the detailed elucidation of NC properties, such as stability. Many metal NC structural rules and stability theories have been developed due to the expanding number of experimentally synthesized NCs. However, the first structural rule came as a prediction that was later confirmed through the determination of the $Au_{102}(p\text{-}MBA)_{44}$. The "divide and protect" theory,[42] first introduced by Häkkinen *et al.* in 2006, states that the NC structure consists of two distinct sections: i) a highly symmetric core made solely of metal atoms, which is protected by ii) a shell of ligand-metal motifs. For thiolate-protected NCs, the protecting groups form as $L\text{-}(M\text{-}L)_x$ units (M = metal, L = thiolate ligand).[42] These units range in sizes, known as different x-Mers,[5] and include x = 0 (i.e. a bridging thiol[43] or a μ_3-coordinated sulfide group[44-46]). Additionally, protecting ring motifs have been observed (e.g. octameric ring in the $Au_{20}(SPh\text{-}^tBu)_{16}$[47]) where metal-ligand

units form a complete loop around the NC core. It should be noted that $(M\text{-}L)_x$ ring structures can form as prenucleation species during NC[48] and nanoparticle[49] synthesis and their presence may affect particle dispersity.[50] These $L\text{-}(M\text{-}L)_x$ protecting groups became known as "staple" motifs due to the staple-like appearance of dimers ($L\text{-}(M\text{-}L)_2$) around the core of the first determined NC, the $Au_{102}(p\text{-}MBA)_{44}$.[21, 51] As an example of the divide and protect theory, Figure 2 demonstrates how the $Au_{25}(SR)_{18}$ NC can be deconstructed into a 13-atom icosahedron core and a protecting shell of six dimeric staple motifs. Remarkably, this simple yet powerful theory universally captures the structural makeup of all thiolate-protected Au NCs.

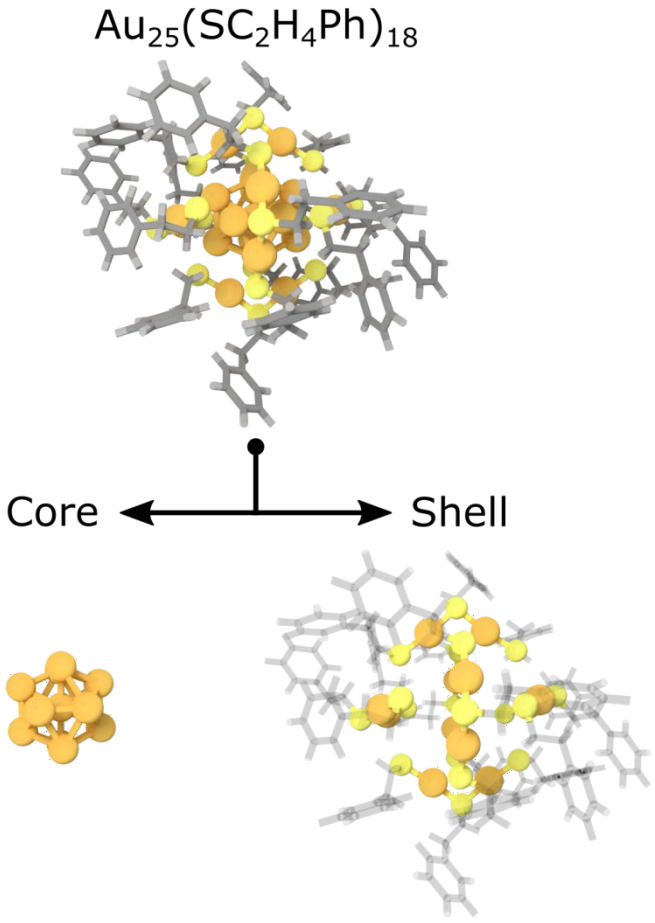


Figure 2. Illustration of the divide and protect theory[42] using the $Au_{25}(SC_2H_4Ph)_{18}$ structure[24] (top). The NC can be decomposed into a Au_{13} icosahedron core (left) protected by a shell of six dimeric staple motifs, RS-Au-SR-Au-SR (right). Gold and yellow balls correspond to Au and S atoms, respectively. R groups (C and H atoms, represented as sticks) are faded in the shell image for clarity.

A second structure-based rule, developed by Dass in 2012, captures the composition constraints of magic size NCs.[52] The "nano-scaling law" describes the number of metal atoms and number of ligands as analogues to the volume (V) and surface area (SA) of NCs, respectively. Through analysis of the known magic sizes, the NCs were found to follow the relationship $SA \propto aV^{2/3}$ (a = scaling factor), which is a known scaling law of primitive geometric shapes. In other words,

$Au_n(SR)_m$ NCs follow $n \propto am^{2/3}$, which rationalizes the specific *n* and *m* pairs that emerge in the magic sizes. Subsequent work revealed that the law holds across the entire size regime of NCs and that the type of ligand affects the scaling factor.[53-55]

Although structural rules can capture geometric trends in metal NCs, they do not provide a means of rationalizing the stability of these nanoscale systems. As a result, all stable NCs generally follow the divide and protect theory and nano-scaling law, but a theoretically predicted NC that adheres to these rules may not be stable. Thus, there has been extensive work on rationalizing the stability of magic size NCs through both their geometric (atomic positions) and electronic (electron configuration) structures. The earliest method utilized an electron counting approach. Drawing from the jellium model, the superatom theory describes that stable NCs have a closed electronic valence shell and a relatively large HOMO-LUMO gap.[56, 57] This theory, which was introduced in 2008,[56] was able to capture the stability of many magic size NCs that were experimentally discovered later. However, the continuous synthesis of stable NCs revealed cases that the superatom theory could not capture, thus limiting its application as a universal stability model.[58, 59] Although extensions to the theory have been reported, like the superatom network[60] and united cluster model,[59] none of these methods alone can rationalize the entire population of magic size NCs.[61] Therefore, these theories, although instrumental in introducing fundamental concepts to the field, lack universal predictive power of the NC structural and electronic stability.

In 2016, Xu *et al.* developed the grand unified model, or GUM, which uses two-electron elementary blocks, namely the triangular Au_3 and tetrahedral Au_4, to rationalize the core structures in stable Au NCs.[62] These blocks are formed from a combination of three different Au "flavors" (1e, 0.5e, and 0e valence states) which are determined by core-shell bonding based on divide and protect theory.[42] GUM revealed that known NC cores can be assembled by combining elementary blocks. Moreover, these NC cores follow duet or octet electron counting rules, which rationalizes their high stability. Notably, GUM has been applied to 70+ ligand-protected Au NCs and can be leveraged as a guide to experimentalists towards new NCs.[62, 63] However, similar to the superatom theory, GUM does not explicitly capture the complete electronic structure of metal NCs, which could allow for "false positive" predictions when exploring new structures (e.g. it does not capture subtle ligand effects which can dictate the size of stable NCs[30]). It is also currently constrained to Au systems. Nevertheless, GUM is an important discovery that helps in rationalizing stability across the complete Au NC size regime.

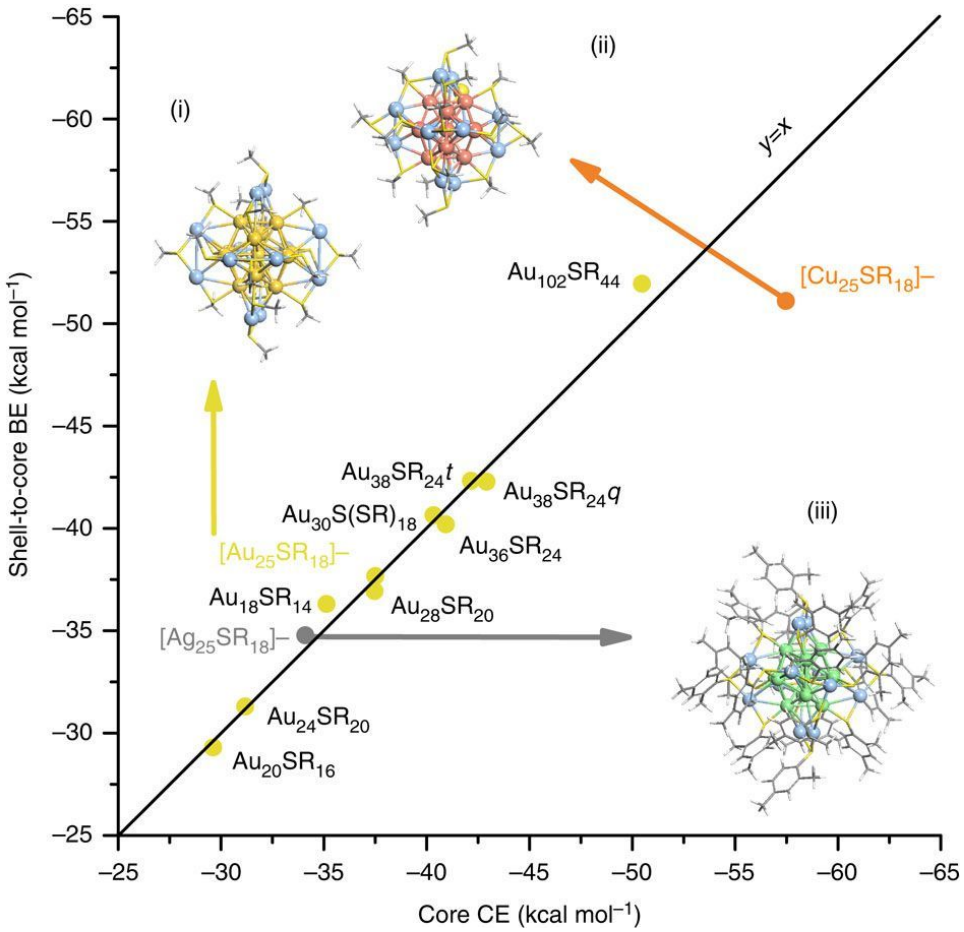


Figure 3. Thermodynamic stability model applied to a sample of thiolate-protected metal NCs. Both core CE and shell-to-core CE are calculated with DFT. Solid black line indicates perfect parity. Inlaid structures represent $[M_{25}SR)_{18}]^{-}$, where i) M = Au, ii) M = Cu, and iii) M = Ag. Balls indicate metal atoms and solid lines the ligands. Different colored balls indicate metal atoms in the core vs. shell regions of the NC. Adapted with permission from ref. 64. Copyright © 2017, Springer Nature.

Recently, Taylor and Mpourmpakis developed the Thermodynamic Stability Model (TSM).[64] The TSM is the first model that not only captures the complete geometric and electronic NC structure (i.e. exact atomic positions and electron configuration, respectively), but also incorporates fundamental thermodynamics to rationalize NC stability (see Table 1).[64] The model leverages the divide and protect theory[42] of a core-shell structure and builds on chemical potential contributions between the core and the shell region of the NCs. Given the harsh conditions undergone during size focusing synthesis, a stable NC achieves chemical equilibrium between its core metal kernel and shell of staple motifs. This equilibrium of chemical potentials can be approximated by two electronic properties, i.e. the core cohesive energy (CE) and shell-to-core binding energy (BE), which can both be calculated using density functional theory (DFT). Core CE is the average bond strength between core metal atoms in the presence of the protecting shell motifs, whereas shell-to-core BE is the binding strength of the shell motifs to the core. When applied to a range of metal NCs (including Au, Ag, and Cu systems), the TSM reveals a fine energy

balance between the core CE and the shell-to-core BE for experimentally synthesized (stable) NCs, as shown by the parity plot in Figure 3. Importantly, the TSM is not constrained to electron counting rules, thus holding predictive power to test on any theoretical candidate structures. For example, although the $[Cu_{25}(SR)_{18}]^-$ follows the superatom theory, it has not been synthesized (due to oxidation affinity of Cu),[65] which can be rationalized by its lack of energy balance under the TSM. Specifically, the Cu NC is well-below the parity line in Figure 3, indicating that the average bond strength of the core Cu atoms is larger than the binding strength of the shell dimer motifs onto the core. Due to chemical potential approximation for each energetic property, this imbalance can also be described as the core and shell of the NC failing to achieve chemical equilibrium, and therefore stability. Another example is the $Au_{36}(SR)_{24}$ NC, which has been experimentally synthesized[26] and TSM predicts it to be a stable NC (energy balance in Figure 3), but is predicted as an unstable NC by the original superatom theory. Further efforts have revealed the TSM's capability to capture stability across the entire size regime of metal NCs (up to the $Au_{279}(SR)_{84}$).[55] Additionally, the TSM has been applied to metal NCs under electrocatalytic conditions, where NCs become catalytically active after partial ligand loss.[14, 15] Specifically, recent results using the TSM revealed the robustness of the $[Au_{25}(SR)_{18}]^-$, which maintains an energy balance (i.e. thermodynamic stability) after partial ligand removal.[66] It is important to note that ligand effects play a large role on NC stability. Some specific examples include interligand interactions and their competition with solvent effects, ligand effects on metal-sulfur bonds, and steric hindrance between ligands.[54] Importantly, the TSM captures ligand effects through its calculated energetic properties, including interfacial (core-shell) strain induced from dramatic changes in ligand structure. However, subtle ligand effects, caused by weak interactions or slight changes in ligand structure (e.g. methyl position in a methylbenzenethiolate ligand) are currently not captured by the model. Thus, further work is required to incorporate these subtle ligand effects into the TSM to achieve a truly universal prediction framework. In addition, other electronic stability criteria have been investigated in literature, including Mulliken electronegativity and chemical hardness. Although the TSM is rooted in a derivation from chemical equilibrium, more work should be done to investigate the relationships between the model and other energetic-based stability criteria. Nevertheless, the promising results with the TSM reveals its potential as a general method to both rationalize and predict stability of thiolate-protected metal NCs.

Table 1. Comparison of NC stability theories. The "~" symbol indicates the theory partially captures structure and electronics through accounting for number of atoms and electron counting rules. (GUM: Grand Unified Model. TSM: Thermodynamic Stability Model.

THEORY	CAPTURES STRUCTURE	CAPTURES ELECTRONICS	CAPTURES THERMODYNAMICS
SUPERATOM[56]	~	~	No
GUM[62]	Yes	~	No
TSM[64]	Yes	Yes	Yes

OPENING THE MATERIALS SPACE THROUGH HETEROMETAL DOPING

Since the initial structure determinations of Au NCs, there has been a great deal of interest in exploring the synthesis and properties of alloy NCs. Shortly after the structure of $[Au_{25}(SC_2H_4Ph)_{18}]^-$ was confirmed, Jiang and Dai conducted a theoretical study on $M@Au_{24}(SC_2H_4Ph)_{18}]^q$ NCs using DFT.[67] Heteroatoms M were selected based on the superatom theory (i.e. structure must maintain an $8e^-$ valence shell closure) as well as limiting the charge of the system to $-2 \leq q \leq +2$ (NC charge originates from the presence of counterions during synthesis). The study identified sixteen heteroatoms that were suitable candidates to dope the NC, as each theoretical structure maintained structural integrity and exhibited a large HOMO-LUMO gap. Additionally, the work revealed that doping can manipulate the electronic properties of the NC (e.g. d-block metal doping led to HOMO-LUMO gaps between 1.0 and 1.5 eV), opening avenues to property tuning.[67]

Currently, there are many examples reported in literature that successfully convert Au NCs into new alloy structures through heterometal doping. For instance, the $[Au_{25}(SR)_{18}]^-$, considered the most widely studied NC,[68] has been successfully doped with a number of different heterometals.[36, 38, 69] The metal exchange approach, where a monometallic NC reacts with a heterometal-ligand complex to yield an alloy NC, has been applied to many other magic size NCs as well.[37, 70-72] Additionally, alloy NCs have been created by co-reducing different metal salts during initial NC synthesis. Yang *et al.* co-reduced $AgBF_4$ and $ClAuPPh_3$ to form $Au_{12}Ag_{32}(SR)_{30}$.[73] Through a similar approach, Song *et al*. reported a new $Au_{52}Cu_{72}(SR)_{55}$ nanoalloy[74] and Higaki *et al.* successfully synthesized a novel $Au_{130-x}Ag_x(SR)_{50}$.[75] Incorporating multiple metals provides a new parameter, the metal concentration ratio, which can be tuned during NC synthesis. In 2019, Kang *et al.* used both metal exchange and co-reduction to construct a rich library of atomically precise $M_{29}(SR)_{18}(PPh_3)_4$ ranging from mono- to tetrametallic systems with M = Ag/Au/Cu/Pt/Pd. By adjusting metal ratios, the authors synthesized 21 atomically precise structures which all exhibited monodispersity.[76] Li *et al.* showed how tuning the Au/Ag ratio, during co-reduction NC synthesis, enables control over the NC structure and dopant concentration. By increasing the amount of Ag precursor, a new $Au_{23-x}Ag_x(SR)_{15}$ structure was discovered (x = 4, 5.76, and 7.44 based on the Au/Ag ratio).[77] Like many other alloy NCs reported in literature, Li *et al.* reported the amount of doped Ag as an average based on single crystal X-ray diffraction analysis. This is due to the distribution of Ag concentrations within the NCs, which reveals the variability of heterometal-doped NCs in a given chemical environment. In other words, one environment (i.e. a single set of synthesis parameters) can lead to a distribution of product NCs with different doping concentrations, thus lacking monodispersity. These results show that further exploration of alloy NCs is required to elucidate their vast materials space.

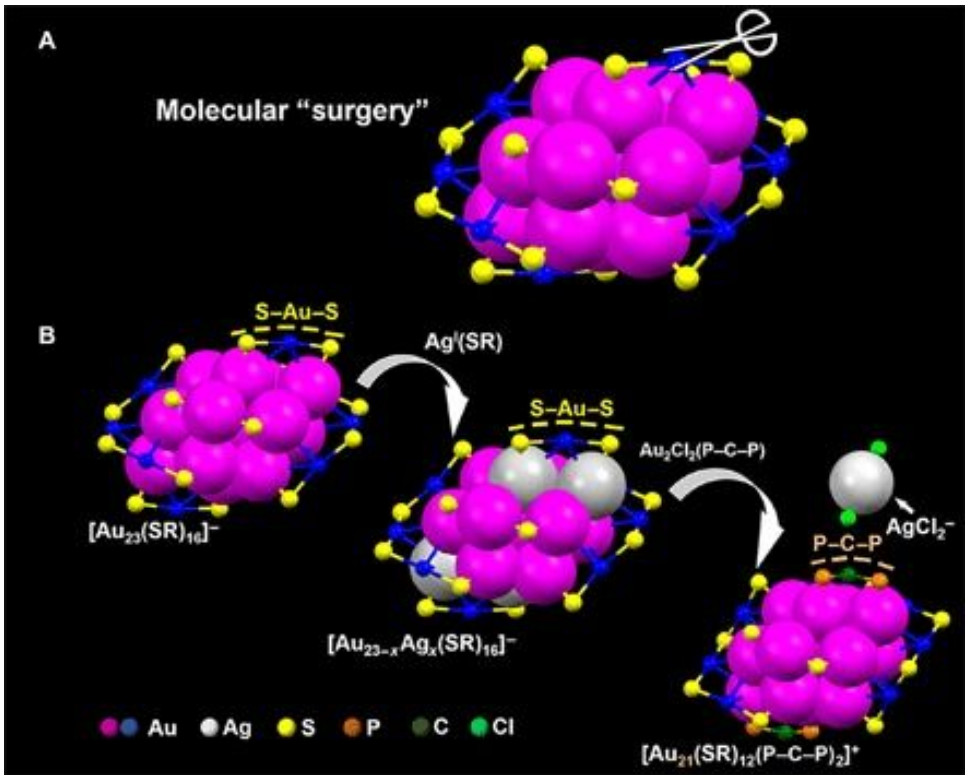


Figure 4. A) Molecular "surgery" of the $[Au_{23}(SR)_{16}]^-$ (R = C_6H_{11}) performed by replacing two monomer ligands (RS-Au-SR) with ($Ph_2PCH_2PPh_2$) B) using a two-step heterometal doping approach. Pink and blue represent core and shell (staple motif) Au atoms, respectively. Silver, yellow, orange, dark green, and light green represent Ag, S, P, C, and Cl, respectively. All other C and H atoms (i.e. R groups) are removed for clarity. Adapted with permission from ref. 79. Copyright © 2017, American Association for the Advancement of Science.

Besides forming alloy systems, heterometal doping can also expand the NC domain to new monometallic structures of different sizes. In 2017, Li *et al.* applied molecular "surgery" to the $[Au_{23}(SR)_{16}]^-$ (R = cyclo-C_6H_{11}) NC, creating a novel 21-gold-atom NC.[78] As shown in Figure 4, the transformation involved a two-step, site-specific metal exchange approach. First, the NC was doped with Ag-SR, forming $[Au_{23-x}(SR)_{16}]^-$ (x ≈ 1), with Ag found at two distinct positions on the surface of the NC core (based on determination of the structure through single crystal X-ray diffraction). Next, the doped-structure was exposed to Au_2Cl_2(P-C-P) (P-C-P = $Ph_2PCH_2PPh_2$), transforming it to $[Au_{21}(SR)_{12}(P\text{-}C\text{-}P)_2)]^+$. This new NC maintained a similar structure to the original Au_{23} by surgically replacing two monomer staple motifs with (P-C-P) units (Figure 4A). Notably, this transformation was not possible without the intermediate heterometal doping step. DFT calculations revealed that accessing the Ag-doped structure allowed for a thermodynamically downhill process that favors the formation of the novel Au_{21} NC. Moreover, a stoichiometric balance of Ag-SR was needed to maintain the mono-doped NC, as a higher concentration of Ag-SR would instead transform the system into an alloy structure, the $[Au_{25-x}Ag_x(SR)_{18}]^-$. These results further illustrate the sensitivity and overall complexity of heterometal doping of NCs to convert experimentally known NCs into new, previously undiscovered nanostructures.

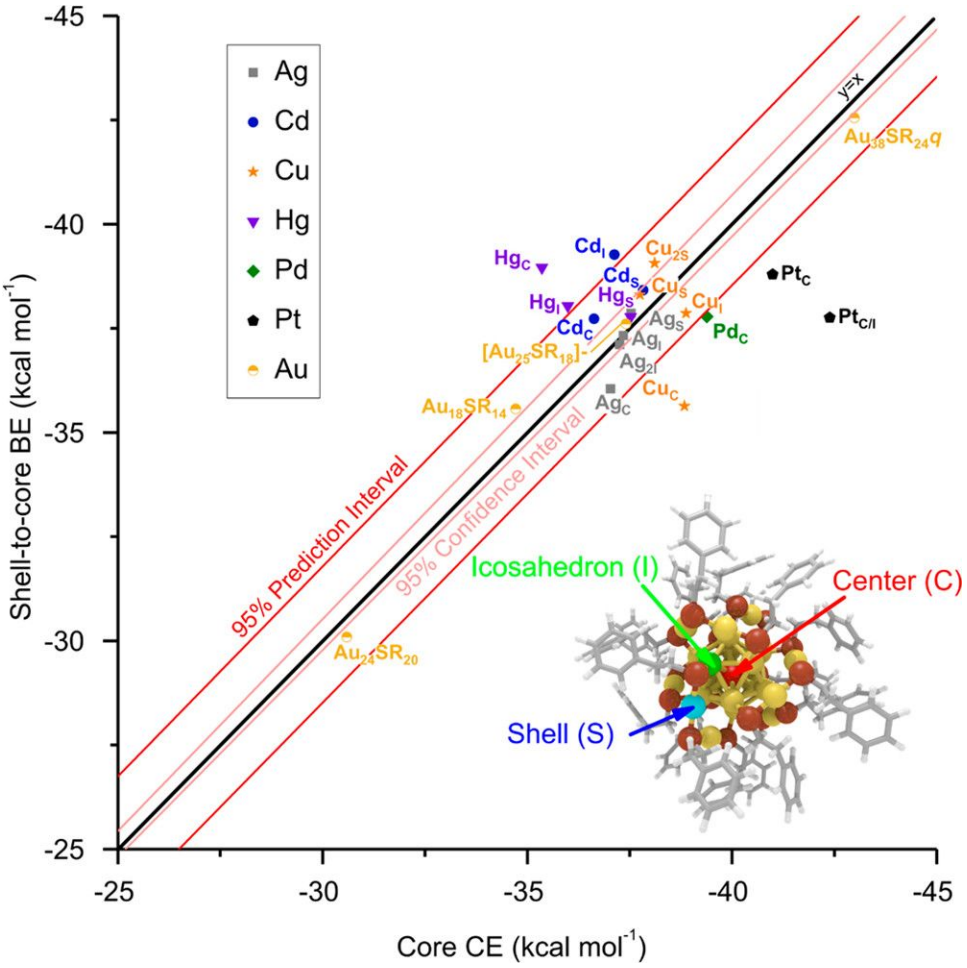


Figure 5. Thermodynamic stability model applied to the monometallic $Au_{18}(SC_6H_{11})_{14}$, $Au_{24}(SCH_2\text{-}^tBu)_{20}$, $[Au_{25}(SC_2H_4Ph)_{18}]^-$, and $Au_{38}(SC_2H_4Ph)_{24}$ as well as a series of heterometal-doped $[Au_{25-x}M_x(SC_2H_4Ph)_{18}]^q$ (x = 1, 2) (M = Ag (q = -1), Cd (q = 0), Cu (q = -1), Hg (q = 0), Pd (q = 0), and Pt (q = 0)). Inset image shows the three unique doping positions within the $[Au_{25}(SC_2H_4Ph)_{18}]^-$ NC. Core CE and shell-to-core BE were calculated using DFT. Adapted with permission from ref. 67. Copyright © 2018, American Chemical Society.

With the increased structural complexity of alloy NCs, it is essential to rationalize stability at an atomic level. Simple electron counting rules could potentially identify types of metal dopants but cannot predict the exact dopant location and concentration in the NC. To this end, the aforementioned TSM was applied to a series of heterometal-doped thiolate-protected $[Au_{25}(SC_2H_4Ph)_{18}]^q$ and $Au_{38}(SC_2H_4Ph)_{24}$ NCs.[79] First, the model was used to analyze a series of $Au_{25-x}M_x(SC_2H_4Ph)_{18}]^q$ NCs, where x = 1 or 2, q = -1 or 0 and M = Ag, Cd, Cu, Hg, Pd, or Pt. The heterometals were doped in three symmetrically distinct positions: the center and surface of the core icosahedron, labelled C and I, respectively, as well as within the dimer staple motifs (shell, S). Figure 5 reveals that the TSM accurately captures the majority of the experimentally observed doping positions, as these systems fall within (or near) the 95% prediction interval stability criterion. For example, Ag has been experimentally determined in the I, and S positions for Au_{25} NCs using x-ray diffraction,[80, 81] and their stability is captured through the TSM. Furthermore, some systems, such as Cu doped in the center position, fail to maintain the fine energy balance

criterion of the TSM – thus are predicted to be thermodynamically unstable. This is in stark contrast to simple DFT energy comparisons, which would instead indicate Cu doped in the C position as the most stable configuration. Importantly, there are cases that fall within the stability criterion but have not been experimentally confirmed (e.g. Cd in the S position). These results suggest that the NCs in question could be synthetically accessible (under appropriate experimental conditions) based on the TSM. It is important to stress that the 95% prediction interval is utilized as a metric to roughly distinguish synthetic accessibility for NCs. Due to the nature of the energetic approximations for core and shell chemical potentials, the TSM provides predictions based on synthetic accessibility rather than a quantitative comparison of stability. Furthermore, the TSM must rely on additional metrics at times to rationalize alloy NC stability. For example, geometric reconstruction during DFT relaxation can be used as a test of stability for alloy NCs.[39] Specific cases of Pd (I and S positions) and Pt (S position) are not shown in Figure 5 due to a large rearrangement in their cores during DFT relaxation,[79] suggesting their lack of stability. Similar results of agreement with experiment were observed for the case of $[Au_{38-x}Ag_x(SC_2H_4Ph)_{24}]^q$. Thus, the TSM displays its potential as a possible candidate for future alloy NC prediction.[79]

UNDERSTANDING PREDICTION

Due to the delay of experimental structure determination[21] compared to the discovery of magic sizes,[20] predicting NC structures has always been at the forefront of NC research. The earliest success of complete structure prediction came from Akola *et al.* in 2008, who leveraged the divide and protect theory along with inspiration from the only known structure at the time, $Au_{102}(SR)_{44}$, to accurately predict the $[Au_{25}(SR)_{18}]^-$ NC.[82] The work correctly proposed a 13-atom Au icosahedron core protected by six dimeric staple motifs (see Figure 2), which the authors validated by comparing its simulated X-ray diffraction pattern to experiment. Furthermore, the work leveraged DFT calculations to rationalize the high stability of the NC, reporting a large HOMO-LUMO gap as well as an eight-electron closed shell of delocalized valence Au electrons. This was the basis to the superatom theory of stability, which the authors (along with others) introduced later that year.[56] Since then, additional efforts have emerged in literature that follow a similar computational approach to predict NC structures. These have led to the successful prediction of $Au_{38}(SR)_{24}$,[25, 83] $Au_{130}(SR)_{50}$,[84-86] and $Au_{144}(SR)_{60}$.[87, 88] However, plenty of structure predictions solely based on a magic size composition have missed the mark when compared to the later experimental determinations,[89-91] indicating that more information is needed to elucidate unknown NC structures.

In a similar idea to the inter-NC relationship illustrated by the nano-scaling law,[52] some NCs protected by the same ligand type are related through consistent metal packing patterns, revealing a systematic size evolution.[92] By incorporating these "growth series" trends in prediction, researchers impose further constraints to identify unknown structures of known magic sizes. Arguably the most studied example[27, 92] is a set of $Au_n(SR)_m$ NCs protected by *tert*-butylbenzenethiolate (TBBT, R = Ph-^{t}Bu), known as the TBBT "magic series".[93] Besides the systematic evolution of the structures, the TBBT series also displays correlated properties.[94] Its

distinctive 1D growth pattern follows $Au_{8n+4}(SR)_{4n+8}$ for n = 2 – 6.[26, 47, 93, 95, 96] In 2013, Pei *et al.* predicted the structure and neutral charge of $Au_{44}(SR)_{28}$ based on structural patterns in the $Au_{28}(SR)_{20}$ and $Au_{36}(SR)_{24}$ NCs and DFT calculations.[57] However, the first $Au_{44}(SR)_{28}$ NC was originally synthesized with phenylthiolate (SPh) ligands and thought to have a "-2" charge.[97] Eventually, Zeng *et al.* synthesized[98] and determined the structure[93] of $Au_{44}(SPh\text{-}^tBu)_{24}$, revealing that Pei *et al.* correctly predicted the structure and neutral charge state. Additional predictions following TBBT growth patterns have been made as well. These include the $Au_{76}(SR)_{44}$ following the aforementioned 1D growth pattern[99] and the $Au_{68}(SR)_{36}$, which instead is based on a 2D growth pattern revealed by select NCs in the TBBT magic series.[100] Furthermore, Xu *et al.* recently introduced a framework that dissects the TBBT growth series to interpolate to new magic size NCs. By modulating the double-helical cores of known NCs in the TBBT series, the authors predicted eleven new structures, including four isomers of the TBBT series that exhibited highly stable electronic configurations as shown by DFT calculations.[101] The results represent important first steps towards structure prediction outside of known magic sizes. However, with the method being limited to the TBBT growth series, new approaches are needed to achieve a general NC prediction framework that is applicable to any given ligand.

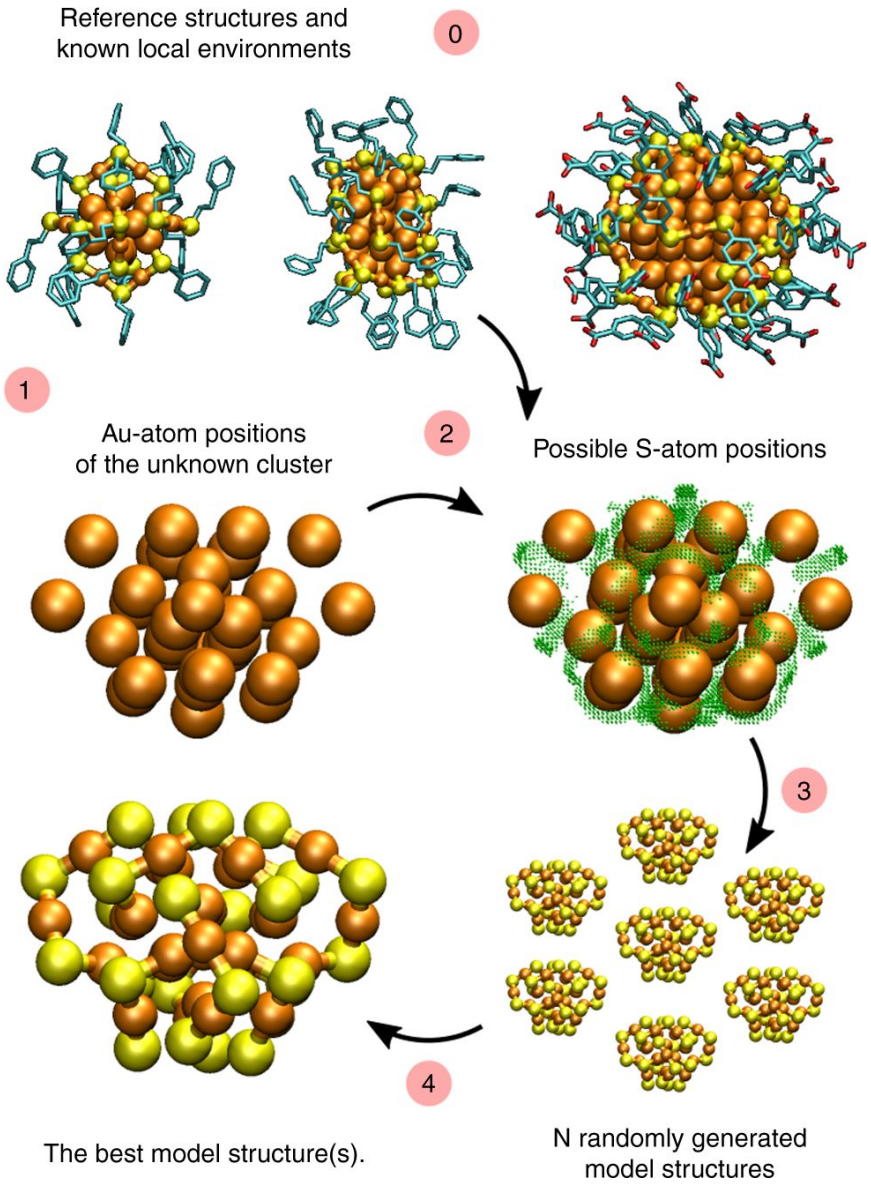


Figure 6. Schematic of the metal-ligand interface prediction algorithm for metal NCs. (0) Referencing a series of known $Au_n(SR)_m$ NCs, (1) a candidate cluster of Au atoms is created and (2) the possible S-atom positions are determined. (3) Next, the algorithm randomly generates a population of NC candidates by adding SR ligands to possible positions, and (4) the best model is determined based on structural error of the ligands relative to experimentally observed local configurations. Adapted with permission from ref. 102, Copyright © 2019, Springer Nature.

Recently, Malola *et al.* introduced a method to predict the metal-ligand interface of ligand-protected metal NCs.[102] The logic steps of the method are shown in Figure 6, with the developed algorithm leveraging experimentally observed local ligand environments by referencing previously reported Au NCs. After defining a pure Au cluster, all possible S atom positions are calculated. Next, a population of randomly generated Au-S cluster candidates are generated with no constraint to the number of S atoms added. These NC models are then ranked by structural

error of their local ligand environments, which is based on the selected experimental reference NCs. Importantly, the method accounts for the steric effects of ligands when ranking NC candidates. Since ligands have been reported to directly affect NC size[30] and properties,[103, 104] it is essential to account for full ligands when predicting new structures. The prediction framework was validated by successfully predicting nine known Au NC structures. Additionally, the framework achieved the same success when applied to Ag NCs protected by thiolate and phosphine ligands, revealing its generalizability to other metal and ligand systems. The method, although promising, relies on prior determination of the metal atom positions. Nevertheless, the work is an important step towards autonomous exploration of the NC materials space, especially NC sizes that have yet to be experimentally synthesized.

NEXT STEPS FOR NC STRUCTURE PREDICTION

Predicting structures based on known magic sizes (i.e. interpolation) has been of great value, but how do we predict new alloys or entirely new magic sizes (i.e. extrapolation)? Progress has been made leveraging systematic patterns found in the known magic sizes[101] and incorporating ligand effects,[102] but a general structure exploration methodology remains the holy grail of ligand protected NCs. A major challenge is overcoming the curse of combinatorics. The vast materials space that NCs exhibit is in large part due to the many choices of metals and ligands. Imagine transforming from $Au_{25}(SR)_{18}$ to $Ag_{25}(SR)_{18}$ by doping one Ag atom at a time, giving only 26 unique compositions (including the two monometallic cases). Within this constrained example, there are actually 33,554,432 unique structures ($\sum_{i=0}^{25}(25 \text{ - } choose \text{ - } i)$, excluding symmetry) due to the distinct positions that each metal type can take (i.e. different possible chemical orderings[105]). The problem becomes even more challenging if we expand to a trimetallic $M_{25}(SR)_{18}$ system, which has been reported in literature.[106] Moreover, removing the structure and single ligand constraints further opens the search space of candidate NCs. This ever-expanding materials space requires an automated computational framework to drive exploration and guide experimentation. Constraining the search to known magic sizes or applying simple electron counting rules would prohibit accessing NC materials space that is probably experimentally accessible and entails unexplored NCs. The framework must instead encompass automated structure generation[102] and be guided by physics-based stability metrics screening across NC size, ligand, and metal type. To this end, the TSM may act as a thermodynamic stability criterion for structure selection.[79] By additionally employing structure property relationships,[55] we can achieve an autonomous structure prediction framework that efficiently samples the materials space towards NCs with enhanced properties for specific applications.

Although challenging, we envision the following steps to achieve such a computational NC structure prediction framework. Despite the growing amount of experimental and computational studies reported, a complete NC structure, property, and synthesis parameters database is missing from literature. A centralized database of all known magic sizes and their doped derivatives would accelerate research efforts on NCs, especially the theoretical ones. This idea is

supported by the literature, as the majority of structural rules, stability models, and SPRs were developed through a collective analysis over many structures.[52, 55, 64] Providing a highly accessible and rich NC database would enable 1) development of new structure property relationships, 2) introduction of improved stability models benchmarked on the entire known NC configuration space, 3) rapid prototyping of potential structure prediction frameworks, and 4) new insights and relationships between synthesis parameters and final NC morphology.

We note that the aforementioned ideas are not new and they have been applied to other classes of nanomaterials. For example, global optimization techniques have been used to successfully predict the structure of organic clusters[107] as well as bare (unprotected) mono- and polymetallic nanoclusters.[108-110] Additionally, structure-property relationships have been previously coupled with structure exploration frameworks to search for systems with desired properties.[110, 111] Indeed, many materials databases that contain a collection of results currently exist online such as bare metal nanoclusters,[112] 2D materials,[113] and 3D crystals.[114] Due to the proven success of applying these approaches to different material classes, developing similar methodologies for ligand-protected nanoclusters is essential to advancing the field.

Machine learning (ML) continues to demand attention for its vast capabilities in structure and property prediction of nanomaterials.[115-118] ML models are driven by rich datasets, which they use to distill patterns and trends within a complicated parameter hyperspace. Currently, ML is missing from NC research, with the exception of two recent studies. In 2018, Panapitiya *et al.* trained a ML model (random forest) to predict CO adsorption energies on thiolate-protected Au/Ag alloy NCs. The trained model revealed the importance of Ag dopant distance from the CO adsorbate.[119] More recently, Li *et al.* trained a deep learning model to predict Au NC monodisperse synthesis protocols. To improve interpretability, the authors created synthetic data, generated from the trained deep learning model, and subsequently trained a decision tree. The new model revealed synthesis criteria (i.e. environmental factors) critical to forming monodisperse Au NCs.[120]

We believe ML has not gained traction in NC research due to the lack of an easily accessible database. However, we note that this is not an insurmountable boundary, as the two aforementioned works surpassed the problem through in-house data generation[119] and data extraction from literature.[120] Nevertheless, providing an organized NC database will accelerate future ML studies on metal NCs. We expect future studies to reveal novel structure property relationships and provide further insights towards accelerating and optimizing NC synthesis. The latter can be achieved both in terms of providing informed guesses of experimental conditions for targeted NC synthesis, as well as NC structures that can be experimentally accessed. Although some ML methods are criticized for their lack of interpretability (i.e. “black box” models with no understanding of physics), we emphasize that ML can, and should be used in concert with previously developed stability theories to form physics-based structure prediction frameworks. For example, active learning algorithms[117, 121-123] can enable a NC materials space exploration

framework that is guided by a general NC stability model and structure property relationships. Thus, the ML-aided search does not lose the important fundamental physics that dictate NC structure, stability, and resulting properties. It instead leverages this domain knowledge towards discovering new, physically and chemically relevant NC structures.

CONCLUSION

In summary, in this *Frontiers* article we highlighted the recent advances of ligand-protected metal NC research, focusing on computational efforts towards understanding the origin of magic size stability of monometallic and alloyed NCs, as well as predicting previously undiscovered NCs. We highlighted current limitations and proposed next steps to overcome challenges presented in NC structure prediction. We hope this perspective motivates the community to promote the development of a single, centralized NC database. We are currently working towards this effort and we anticipate many other research labs to follow. Compiling results from the many revolutionary works reported in ligand-protected NC literature will enable the application of ML and other data-science-based approaches. Through these methods, we can accelerate discovery and advancement of the field, making a significant step towards complete NC structural control.

CONFLICTS OF INTEREST

There are no conflicts to declare.

ACKNOWLEDGEMENTS

We gratefully acknowledge support for this work by the National Science Foundation (NSF, CBET-CAREER program) under Grant No. 1652694.

REFERENCES

1. I. Chakraborty and T. Pradeep, *Chem. Rev.*, 2017, **117**, 8208-8271.
2. M. Zhou, C. Zeng, Q. Li, T. Higaki and R. Jin, *Nanomaterials*, 2019, **9**, 933.
3. K. L. D. M. Weerawardene, E. B. Guidez and C. M. Aikens, *J. Phys. Chem. C*, 2017, **121**, 15416-15423.
4. M. J. Cowan, T. Higaki, R. Jin and G. Mpourmpakis, *J. Phys. Chem. C*, 2019, **123**, 20006-20012.
5. R. Jin, C. Zeng, M. Zhou and Y. Chen, *Chem. Rev.*, 2016, **116**, 10346-10413.
6. Y. Su, T. Xue, Y. Liu, J. Qi, R. Jin and Z. Lin, *Nano Res.*, 2019, **12**, 1251-1265.
7. S. Chattoraj, M. A. Amin, S. Mohapatra, S. Ghosh and K. Bhattacharyya, *ChemPhysChem*, 2016, **17**, 61-68.
8. X. Jiang, X. Wang, C. Yao, S. Zhu, L. Liu, R. Liu and L. Li, *J. Phys. Chem. Lett.*, 2019, **10**, 5237-5243.
9. D. Li, B. Kumari, J. M. Makabenta, A. Gupta and V. Rotello, *Nanoscale*, 2019, **11**, 22172-22181.
10. Y. Du, H. Sheng, D. Astruc and M. Zhu, *Chem. Rev.*, 2020, **120**, 526-622.
11. G. Li, C. Zeng and R. Jin, *J. Am. Chem. Soc.*, 2014, **136**, 3673-3679.
12. C. Liu, X. Ren, F. Lin, X. Fu, X. Lin, T. Li, K. Sun and J. Huang, *Angew. Chem. Int. Ed.*, 2019, **58**, 11335-11339.
13. B. Kim, H. Seong, J. T. Song, K. Kwak, H. Song, Y. C. Tan, G. Park, D. Lee and J. Oh, *ACS Energy Lett.*, 2019, **5**, 749-757.

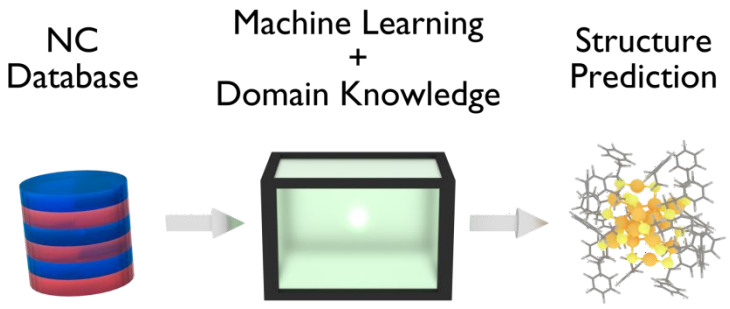
14. S. Zhao, N. Austin, M. Li, Y. Song, S. D. House, S. Bernhard, J. C. Yang, G. Mpourmpakis and R. Jin, *ACS Catal.*, 2018, **8**, 4996-5001.
15. N. Austin, S. Zhao, J. R. McKone, R. Jin and G. Mpourmpakis, *Catal. Sci. Technol.*, 2018, **8**, 3795-3805.
16. S. Zhao, R. Jin and R. Jin, *ACS Energy Lett.*, 2018, **3**, 452-462.
17. M. Brust, M. Walker, D. Bethell, D. J. Schiffrin and R. Whyman, *J. Chem. Soc., Chem. Commun.*, 1994, **0**, 801-802.
18. T. G. Schaaff, G. Knight, M. N. Shafigullin, R. F. Borkman and R. L. Whetten, *J. Phys. Chem. B*, 1998, **102**, 10643-10646.
19. T. G. Schaaff and R. L. Whetten, *J. Phys. Chem. B*, 2000, **104**, 2630-2641.
20. Y. Negishi, K. Nobusada and T. Tsukuda, *J. Am. Chem. Soc.*, 2005, **127**, 5261-5270.
21. P. D. Jadzinsky, G. Calero, C. J. Ackerson, D. A. Bushnell and R. D. Kornberg, *Science*, 2007, **318**, 430.
22. M. W. Heaven, A. Dass, P. S. White, K. M. Holt and R. W. Murray, *J. Am. Chem. Soc.*, 2008, **130**, 3754-3755.
23. M. Zhu, C. M. Aikens, F. J. Hollander, G. C. Schatz and R. Jin, *J. Am. Chem. Soc.*, 2008, **130**, 5883-5885.
24. M. Zhu, W. T. Eckenhoff, T. Pintauer and R. Jin, *J. Phys. Chem. C*, 2008, **112**, 14221-14224.
25. H. Qian, W. T. Eckenhoff, Y. Zhu, T. Pintauer and R. Jin, *J. Am. Chem. Soc.*, 2010, **132**, 8280-8281.
26. C. Zeng, H. Qian, T. Li, G. Li, L. Rosi Nathaniel, B. Yoon, N. Barnett Robert, L. Whetten Robert, U. Landman and R. Jin, *Angew. Chem. Int. Ed.*, 2012, **51**, 13114-13118.
27. Y. Pei, P. Wang, Z. Ma and L. Xiong, *Acc. Chem. Res.*, 2019, **52**, 23-33.
28. R. Jin, *Nanoscale*, 2015, **7**, 1549-1565.
29. R. Jin, H. Qian, Z. Wu, Y. Zhu, M. Zhu, A. Mohanty and N. Garg, *J. Phys. Chem. Lett.*, 2010, **1**, 2903-2910.
30. Y. Chen, C. Zeng, D. R. Kauffman and R. Jin, *Nano Lett.*, 2015, **15**, 3603-3609.
31. C. J. Zeng, *Pure Appl. Chem.*, 2018, **90**, 1409-1427.
32. T. Higaki, Q. Li, M. Zhou, S. Zhao, Y. Li, S. Li and R. Jin, *Acc. Chem. Res.*, 2018, **51**, 2764-2773.
33. Q. Yao, X. Yuan, T. Chen, D. T. Leong and J. Xie, *Adv. Mater.*, 2018, **30**, e1802751.
34. R. Jin and K. Nobusada, *Nano Res.*, 2014, **7**, 285-300.
35. W. Fei, S. Antonello, T. Dainese, A. Dolmella, M. Lahtinen, K. Rissanen, A. Venzo and F. Maran, *J. Am. Chem. Soc.*, 2019, **141**, 16033-16045.
36. N. Yan, L. Liao, J. Yuan, Y.-j. Lin, L.-h. Weng, J. Yang and Z. Wu, *Chem. Mater.*, 2016, **28**, 8240-8247.
37. S. Hossain, Y. Niihori, L. V. Nair, B. Kumar, W. Kurashige and Y. Negishi, *Acc. Chem. Res.*, 2018, **51**, 3114-3124.
38. S. Wang, Q. Li, X. Kang and M. Zhu, *Acc. Chem. Res.*, 2018, **51**, 2784-2792.
39. Q. Li, K. J. Lambright, M. G. Taylor, K. Kirschbaum, T. Y. Luo, J. Zhao, G. Mpourmpakis, S. Mokashi-Punekar, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2017, **139**, 17779-17782.
40. Q. Li, M. G. Taylor, K. Kirschbaum, K. J. Lambright, X. Zhu, G. Mpourmpakis and R. Jin, *J. Colloid Interf. Sci.*, 2017, **505**, 1202-1207.
41. J. Mato and E. B. Guidez, *J. Phys. Chem. A*, 2020, **124**, 2601-2615.
42. H. Häkkinen, M. Walter and H. Grönbeck, *J. Phys. Chem. B*, 2006, **110**, 9927-9931.
43. S. Chen, L. Xiong, S. Wang, Z. Ma, S. Jin, H. Sheng, Y. Pei and M. Zhu, *J. Am. Chem. Soc.*, 2016, **138**, 10754-10757.
44. D. Crasto, S. Malola, G. Brosofsky, A. Dass and H. Häkkinen, *J. Am. Chem. Soc.*, 2014, **136**, 5000-5005.

45. C. Liu, T. Li, G. Li, K. Nobusada, C. Zeng, G. Pang, N. L. Rosi and R. Jin, *Angew. Chem. Int. Ed.*, 2015, **54**, 9826-9829.
46. Z. Gan, J. Chen, J. Wang, C. Wang, M.-B. Li, C. Yao, S. Zhuang, A. Xu, L. Li and Z. Wu, *Nat. Commun.*, 2017, **8**, 14739.
47. C. Zeng, C. Liu, Y. Chen, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2014, **136**, 11922-11925.
48. Z. Luo, V. Nachammai, B. Zhang, N. Yan, D. T. Leong, D. E. Jiang and J. Xie, *J. Am. Chem. Soc.*, 2014, **136**, 10577-10580.
49. L. E. Marbella, D. M. Chevrier, P. D. Tancini, O. Shobayo, A. M. Smith, K. A. Johnston, C. M. Andolina, P. Zhang, G. Mpourmpakis and J. E. Millstone, *J. Am. Chem. Soc.*, 2015, **137**, 15852-15858.
50. G. Mpourmpakis, S. Caratzoulas and D. G. Vlachos, *Nano Lett.*, 2010, **10**, 3408-3413.
51. D. E. Jiang, M. L. Tiago, W. Luo and S. Dai, *J. Am. Chem. Soc.*, 2008, **130**, 2777-2779.
52. A. Dass, *Nanoscale*, 2012, **4**, 2260-2263.
53. N. A. Sakthivel and A. Dass, *Acc. Chem. Res.*, 2018, **51**, 1774-1783.
54. M. Rambukwella, N. A. Sakthivel, J. H. Delcamp, L. Sementa, A. Fortunelli and A. Dass, *Front. Chem.*, 2018, **6**, 330.
55. M. J. Cowan and G. Mpourmpakis, *Nanoscale Adv.*, 2019, **1**, 184-188.
56. M. Walter, J. Akola, O. Lopez-Acevedo, P. D. Jadzinsky, G. Calero, C. J. Ackerson, R. L. Whetten, H. Gronbeck and H. Hakkinen, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 9157-9162.
57. Y. Pei, S. Lin, J. Su and C. Liu, *J. Am. Chem. Soc.*, 2013, **135**, 19060-19063.
58. A. Das, T. Li, K. Nobusada, C. Zeng, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2013, **135**, 18264-18267.
59. D. M. P. Mingos, *Dalton Trans.*, 2015, **44**, 6680-6695.
60. L. Cheng, Y. Yuan, X. Zhang and J. Yang, *Angew. Chem. Int. Ed.*, 2013, **52**, 9035-9039.
61. Z. Y. Ma, P. Wang, L. Xiong and Y. Pei, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2017, **7**.
62. W. W. Xu, B. Zhu, X. C. Zeng and Y. Gao, *Nat. Commun.*, 2016, **7**, 13574.
63. W. W. Xu, X. C. Zeng and Y. Gao, *Acc. Chem. Res.*, 2018, **51**, 2739-2747.
64. M. G. Taylor and G. Mpourmpakis, *Nat. Commun.*, 2017, **8**, 15988.
65. T.-A. D. Nguyen, Z. R. Jones, B. R. Goldsmith, W. R. Buratto, G. Wu, S. L. Scott and T. W. Hayton, *J. Am. Chem. Soc.*, 2015, **137**, 13319-13324.
66. A. V. Nagarajan, R. Juarez-Mosqueda, M. J. Cowan, R. Jin, D. R. Kauffman and G. Mpourmpakis, *SN Appl. Sci.*, 2020, **2**, 680.
67. D. E. Jiang and S. Dai, *Inorg. Chem.*, 2009, **48**, 2720-2722.
68. X. Kang, H. Chong and M. Zhu, *Nanoscale*, 2018, **10**, 10758-10834.
69. X. Wei, X. Kang, S. Wang and M. Zhu, *Dalton Trans.*, 2018, **47**, 13766-13770.
70. S. Wang, L. Xiong, G. Sun, L. Tang, J. Zhang, Y. Pei and M. Zhu, *Nanoscale Adv.*, 2020, **2**, 664-668.
71. A. Ghosh, O. F. Mohammed and O. M. Bakr, *Acc. Chem. Res.*, 2018, **51**, 3094-3103.
72. S. Wang, H. Abroshan, C. Liu, T. Y. Luo, M. Zhu, H. J. Kim, N. L. Rosi and R. Jin, *Nat. Commun.*, 2017, **8**, 848.
73. H. Yang, Y. Wang, H. Huang, L. Gell, L. Lehtovaara, S. Malola, H. Häkkinen and N. Zheng, *Nat. Commun.*, 2013, **4**, 2422.
74. Y. Song, Y. Li, H. Li, F. Ke, J. Xiang, C. Zhou, P. Li, M. Zhu and R. Jin, *Nat. Commun.*, 2020, **11**, 478.
75. T. Higaki, C. Liu, D. J. Morris, G. He, T.-Y. Luo, M. Y. Sfeir, P. Zhang, N. L. Rosi and R. Jin, *Angew. Chem. Int. Ed.*, 2019, **58**, 18798-18802.
76. X. Kang, X. Wei, S. Jin, Q. Yuan, X. Luan, Y. Pei, S. Wang, M. Zhu and R. Jin, *Proc. Natl. Acad. Sci. U.S.A.*, 2019, **116**, 18834-18840.
77. Y. Li, T.-Y. Luo, M. Zhou, Y. Song, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2018, **140**, 14235-14243.

78. Q. Li, T. Y. Luo, M. G. Taylor, S. Wang, X. Zhu, Y. Song, G. Mpourmpakis, N. L. Rosi and R. Jin, *Sci. Adv.*, 2017, **3**, e1603193.
79. M. G. Taylor and G. Mpourmpakis, *J. Phys. Chem. Lett.*, 2018, **9**, 6773-6778.
80. Q. Li, S. Wang, K. Kirschbaum, K. J. Lambright, A. Das and R. Jin, *Chem. Commun.*, 2016, **52**, 5194-5197.
81. C. Kumara, C. M. Aikens and A. Dass, *J. Phys. Chem. Lett.*, 2014, **5**, 461-466.
82. J. Akola, M. Walter, R. L. Whetten, H. Hakkinen and H. Gronbeck, *J. Am. Chem. Soc.*, 2008, **130**, 3756-3757.
83. O. Lopez-Acevedo, H. Tsunoyama, T. Tsukuda, H. Hakkinen and C. M. Aikens, *J. Am. Chem. Soc.*, 2010, **132**, 8210-8218.
84. Y. Negishi, C. Sakamoto, T. Ohyama and T. Tsukuda, *J. Phys. Chem. Lett.*, 2012, **3**, 1624-1628.
85. A. Tlahuice-Flores, U. Santiago, D. Bahena, E. Vinogradova, C. V. Conroy, T. Ahuja, S. B. H. Bach, A. Ponce, G. Wang, M. José-Yacamán and R. L. Whetten, *J. Phys. Chem. A*, 2013, **117**, 10470-10476.
86. Y. Chen, C. Zeng, C. Liu, K. Kirschbaum, C. Gayathri, R. R. Gil, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2015, **137**, 10076-10079.
87. O. Lopez-Acevedo, J. Akola, R. L. Whetten, H. Grönbeck and H. Häkkinen, *J. Phys. Chem. C*, 2009, **113**, 5035-5038.
88. N. Yan, N. Xia, L. Liao, M. Zhu, F. Jin, R. Jin and Z. Wu, *Sci. Adv.*, 2018, **4**, eaat7259.
89. N. K. Chaki, Y. Negishi, H. Tsunoyama, Y. Shichibu and T. Tsukuda, *J. Am. Chem. Soc.*, 2008, **130**, 8608-8610.
90. D.-e. Jiang, M. Walter and J. Akola, *J. Phys. Chem. C*, 2010, **114**, 15883-15889.
91. Y. Pei, R. Pal, C. Liu, Y. Gao, Z. Zhang and X. C. Zeng, *J. Am. Chem. Soc.*, 2012, **134**, 3015-3024.
92. X. Du, J. Chai, S. Yang, Y. Li, T. Higaki, S. Li and R. Jin, *Nanoscale*, 2019, **11**, 19158-19165.
93. C. Zeng, Y. Chen, K. Iida, K. Nobusada, K. Kirschbaum, K. J. Lambright and R. Jin, *J. Am. Chem. Soc.*, 2016, **138**, 3950-3953.
94. M. Zhou, C. Zeng, M. Y. Sfeir, M. Cotlet, K. Iida, K. Nobusada and R. Jin, *J. Phys. Chem. Lett.*, 2017, **8**, 4023-4030.
95. C. Zeng, T. Li, A. Das, N. L. Rosi and R. Jin, *J. Am. Chem. Soc.*, 2013, **135**, 10011-10013.
96. C. Zeng, Y. Chen, C. Liu, K. Nobusada, N. L. Rosi and R. Jin, *Sci. Adv.*, 2015, **1**, e1500425.
97. R. C. Price and R. L. Whetten, *J. Am. Chem. Soc.*, 2005, **127**, 13750-13751.
98. C. Zeng, Y. Chen, G. Li and R. Jin, *Chem. Commun.*, 2014, **50**, 55-57.
99. Z. Ma, P. Wang, G. Zhou, J. Tang, H. Li and Y. Pei, *J. Phys. Chem. C*, 2016, **120**, 13739-13748.
100. P. Wang, X. Sun, X. Liu, L. Xiong, Z. Ma and Y. Pei, *J. Phys. Chem. Lett.*, 2017, **8**, 1248-1252.
101. W. W. Xu, X. Duan and X. C. Zeng, *J. Phys. Chem. Lett.*, 2020, **11**, 536-540.
102. S. Malola, P. Nieminen, A. Pihlajamäki, J. Hämäläinen, T. Kärkkäinen and H. Häkkinen, *Nat. Commun.*, 2019, **10**, 3973.
103. R. Juarez-Mosqueda and G. Mpourmpakis, *Phys. Chem. Chem. Phys.*, 2019, **21**, 22272-22282.
104. R. Senanayake and C. M. Aikens, *Phys. Chem. Chem. Phys.*, 2020, **22**, 5272-5285.
105. R. Ferrando, J. Jellinek and R. L. Johnston, *Chem. Rev.*, 2008, **108**, 845-910.
106. S. Hossain, T. Ono, M. Yoshioka, G. Hu, M. Hosoi, Z. Chen, L. V. Nair, Y. Niihori, W. Kurashige, D. E. Jiang and Y. Negishi, *J. Phys. Chem. Lett.*, 2018, **9**, 2590-2594.
107. J. Hernández-Rojas and F. Calvo, *Front. Chem.*, 2019, **7**, 573.
108. N. M. Isenberg, M. G. Taylor, Z. Yan, C. L. Hanselman, G. Mpourmpakis and C. E. Gounaris, *Mol. Syst. Des. Eng.*, 2020, **5**, 232-244.
109. Z. Yan, M. G. Taylor, A. Mascareno and G. Mpourmpakis, *Nano Lett.*, 2018, **18**, 2696-2704.
110. R. Ferrando, A. Fortunelli and R. L. Johnston, *Phys. Chem. Chem. Phys.*, 2008, **10**, 640-649.
111. M. Wang, X. Hu, D. N. Beratan and W. Yang, *J. Am. Chem. Soc.*, 2006, **128**, 3228-3232.

112. D. J. Wales, J. P. K. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin, F. Calvo, J. Hernandez-Rojas and T. F. Middleton, *http://www-wales.ch.cam.ac.uk/CCD.html*.
113. S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, J. Gath, K. W. Jacobsen, J. Jørgen Mortensen, T. Olsen and K. S. Thygesen, *2D Mater.*, 2018, **5**, 042002.
114. S. Gražulis, D. Chateigner, R. T. Downs, A. F. T. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck and A. Le Bail, *J. Appl. Crystallogr.*, 2009, **42**, 726-729.
115. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev and A. Walsh, *Nature*, 2018, **559**, 547-555.
116. P. Schlexer Lamoureux, K. T. Winther, J. A. Garrido Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen and T. Bligaard, *ChemCatChem*, 2019, **11**, 3581-3601.
117. K. A. Brown, S. Brittman, N. Maccaferri, D. Jariwala and U. Celano, *Nano Lett.*, 2020, **20**, 2-10.
118. J. Dean, M. G. Taylor and G. Mpourmpakis, *Sci. Adv.*, 2019, **5**, eaax5101.
119. G. Panapitiya, G. Avendano-Franco, P. Ren, X. Wen, Y. Li and J. P. Lewis, *J. Am. Chem. Soc.*, 2018, **140**, 17508-17514.
120. J. Li, T. Chen, K. Lim, L. Chen, S. A. Khan, J. Xie and X. Wang, *Adv. Intell. Syst.*, 2019, **1**, 1900029.
121. Y. H. Tang and W. A. de Jong, *J. Chem. Phys.*, 2019, **150**, 044107.
122. T. D. Loeffler, T. K. Patra, H. Chan, M. Cherukara and S. K. R. S. Sankaranarayanan, *J. Phys. Chem. C*, 2020, **124**, 4907-4916.
123. C. Dai and S. C. Glotzer, *J. Phys. Chem. B*, 2020, **124**, 1275-1284.

TOC



Developing a centralized database for ligand-protected nanoclusters can fuel machine learning and data-science-based approaches towards theoretical structure prediction.