MATERIALS SCIENCE

Machine learning–accelerated design and synthesis of polyelemental heterostructures

Carolin B. Wahl[1,2]†, Muratahan Aykol[3]†, Jordan H. Swisher[2,4]†, Joseph H. Montoya[3], Santosh K. Suram[3], Chad A. Mirkin[1,2,4]*

In materials discovery efforts, synthetic capabilities far outpace the ability to extract meaningful data from them. To bridge this gap, machine learning methods are necessary to reduce the search space for identifying desired materials. Here, we present a machine learning–driven, closed-loop experimental process to guide the synthesis of polyelemental nanomaterials with targeted structural properties. By leveraging data from an eight-dimensional chemical space (Au-Ag-Cu-Co-Ni-Pd-Sn-Pt) as inputs, a Bayesian optimization algorithm is used to suggest previously unidentified nanoparticle compositions that target specific interfacial motifs for synthesis, results of which are iteratively shared back with the algorithm. This feedback loop resulted in successful syntheses of 18 heterojunction nanomaterials that are too complex to discover by chemical intuition alone, including extremely chemically complex biphasic nanoparticles reported to date. Platforms like the one developed here are poised to transform materials discovery across a wide swath of applications and industries.

INTRODUCTION

Nanoparticles (NPs) expand the design space of functional materials from bulk to surface and interface dominated length scales where previously unknown physical and chemical properties emerge (*1*). Efforts to explore this vast space and identify previously unidentified structures, primarily through size, shape, and composition control in an attempt to elucidate structure-function relationships, have been ongoing for decades. Using both experimental combinatorial approaches (*2–4*) and computational methods (*5–9*), new NP compositions and structures have been found for applications ranging from energy storage and conversion to therapeutics (*10–12*). Metallic NPs, for instance, show particular promise as heterogeneous catalysts for industrially critical reactions, such as oxygen reduction and evolution, hydrogen evolution, and CO_2 reduction, where complex design features at the nanoscale can have profound effects on catalyst performance (*1, 13*). NP size is a widely studied design axis, as changes in surface-to-volume ratios or surface atom coordination play a vital role in controlling properties on a per-atom basis, as in catalysis of CO_2 reduction on Au NPs (*14*). Beyond size, variation of phase morphology is emerging as another axis that can greatly expand the design space in nanostructured materials, as complex interface motifs between phases can impact properties through numerous mechanisms. For instance, catalytic activity can be enhanced by inducing lattice strain on the particle surface through a deliberately selected core (e.g., faster oxygen reduction on Pt in PtPb@Pt core-shell NPs) (*15*) or controlling the arrangement of active sites across an interface (e.g., as in tandem catalysts) (*16*). Yet, despite the potential impact interfaces can have on particle performance or properties, there are few robust and controlled methods for deliberately producing these interfaces in nanostructures.

Recent developments in the synthesis of multimetallic NPs have greatly expanded their design space (*17*), with elemental complexities reaching 10 elements in a single particle and phase motifs ranging from homogenous alloys to core-shell particles and even to heterojunction multiphasic particles with up to six unique interfaces coexisting in a single particle (*18–20*). Among the techniques, scanning probe block copolymer lithography (SPBCL), wherein a scanning probe deposits a miniscule volume of block copolymer and metal salts to form a volume-confined nanoreactor, has emerged as a particularly robust method for the mass synthesis of heterojunction NPs (*2, 6, 18, 21–24*). In SPBCL, the block copolymer serves as a solvent to facilitate the aggregation and coarsening of metal precursors into single particles, resulting in spatially defined arrays of size and composition controlled NPs. Past examples include NPs with as many as seven elements and six unique interfaces and libraries of nanostructures with more than a million positionally encoded discrete nanostructures—so-called "megalibraries" (*18*). This platform, therefore, is an excellent means for systematically investigating the properties of heterojunction NPs.

However, given the vast possible combinations of composition and particle size accessible through SPBCL and megalibraries in general (e.g., >120 million pentametallic combinations when considering particles between 10 and 20 nm in 2-nm increments), particularly when generating NP megalibraries in combinatorial synthesis (*2*), experimentation alone is insufficient to realize structures of interest under finite resources and in reasonable time scales. Accurate first-principles or atomistic-level modeling of NP properties, particularly those pertaining to interfaces, typically comes at a high computational cost that renders them viable for case studies but prohibitive for navigating the vast combinatorial space of multimetallic nanostructures. Thus, despite the emergent properties and performance gains that designer heterojunctions may impart, rational control of interfaces between distinct phases in nanomaterials remains a formidable challenge.

Machine learning, particularly in the context of Bayesian optimization (BO) (*25–27*), can help circumvent these bottlenecks by learning a relationship between the targeted property of NPs and design parameters using past experimental results and can guide the expensive and labor-intensive synthetic search through an iterative feedback loop between physical experiments and optimization algorithm.

[1]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA. [2]International Institute for Nanotechnology, Northwestern University, Evanston, IL 60208, USA. [3]Toyota Research Institute, Los Altos, CA 94022, USA. [4]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA.
*Corresponding author. Email: chadnano@northwestern.edu
†These authors contributed equally to this work.

Here, we demonstrate a highly efficient closed-loop integration of SPBCL synthesis and scanning transmission electron microscopy (STEM) imaging of multimetallic NPs with BO, focusing on accelerated discovery of isolated (single) interface, two-phase quaternary, quinary, and senary multimetallic NPs in a broader eight-dimensional (8D) element space. Our results show that the rational compositional design of heterojunctions in NPs is possible through BO when researchers "stay in the loop," paving the way for the machine learning–guided design of more complex, functional multimetallic NPs.

RESULTS
Design and simulation of an optimization approach for NP discovery

The complexity and breadth of the phase morphologies NPs can adopt increases rapidly with the number of constituent elements, n. On the basis of the phase rule, the theoretical upper bound for the number of unique interfaces in multielement NPs grows quadratically with n. Interface arrangements resulting in these heterostructured NPs can lead to unprecedented properties (1), but we initially narrow our focus to the discovery of a special subset: biphasic, single-interface NPs (SINPs) in quaternary and higher-order chemical spaces. In two- or three-element NPs, expert intuition conditioned on knowledge of phase diagrams, crystal structures, and past observations can help target SINP compositions, but this intuition does not scale well to designing higher-order, chemically diverse SINPs, and researchers can greatly benefit from a feedback loop with machine learning. More broadly, we see SINPs as a critical first target for closed-loop, machine learning–driven design of heterostructured NPs because (i) the single interface enables facile electron microscopy characterization and, in turn, robust experimental results, and (ii) developing the capability for SINP discovery can further accelerate the development of NPs for applications that are influenced by interfaces, such as catalysis or plasmonics, by controlled deconvolution of interface-property relationships.

Sequential optimization strategies, such as BO, have started gaining attention in the chemical and physical sciences to solve high-dimensional and expensive optimization problems where factorial or intuition-driven experimental designs become infeasible (5, 25–30). In BO, often used with a Gaussian process (GP) prior, the next experiment in a sequence of experiments is selected by an acquisition function that combines the predictions of the GP conditioned on past results with a decision-making strategy, balancing the exploration-exploitation trade-off, to maximize the gains toward an objective (26). We start the BO-driven SINP discovery effort by curating a dataset of 148 unique NP compositions from our reports, predating this study within the 8D Au-Ag-Cu-Co-Ni-Pd-Sn-Pt space, all synthesized and characterized under a consistent SBCPL-annealing STEM workflow that is adopted in this work (see Material and Methods and the Supplementary Materials) (6, 18, 21–24). Having this dataset is critical not only for improving the performance of BO in scheduling real experiments by conditioning the model on past data but also for designing and testing the BO strategy itself, before the actual costly deployment phase.

The chemical diversity and interfacial complexity of the NPs in this initial dataset are depicted in Fig. 1 (A and B), with the number of elemental components in each particle ranging from 2 to 7 and with the number of unique interface counts ranging from 0 to 6. In Fig. 1A, we observe that the dataset is dominated by two- to four-element

NPs with zero to three interfaces, with higher interface counts appearing exclusively in higher-order systems, as expected. There are 45 SINPs in this dataset, limited to two- to four-element NPs, most abundant for ternary compositions. There are only seven quaternary SINPs and no pentanary or higher-order ones. Figure 1B shows that, except for Pt that exists in only a handful of NPs, there is no trivial relationship between the presence of a particular element in a particle and its interface count. While this dataset, to the best of our knowledge, is the largest compilation of microscopy-characterized alloy NPs, it still constitutes a highly sparse, researcher-driven sampling from independent studies, at compositions in highly disjointed subspaces of the Au-Ag-Cu-Co-Ni-Pd-Sn-Pt space. For example, approximately half of all composition pairs in the dataset share at most one common element, and of all possible two-, three-, four-, five-, six-, and seven-component spaces in the 8D design space, respectively, about 39, 64, 76, 80, 82, and 88% are empty, meaning they do not contain a single sample.

Hence, to overcome potential data inefficiencies in BO that would stem from sparsity and disjoint subspaces, we designed a domain-specific upper confidence bound (UCB) strategy and simulated its performance in BO for SINP discovery in the existing dataset before using it to help researchers in actual experiments. This UCB strategy transforms the NP compositions into a chemically informed and dynamically updated embedding space, axes of which are computable from composition and shared properties of elements (e.g., electronegativity) for all NPs rather than chemical identity of elements themselves (see Material and Methods for details). The first scenario that we simulated broadly targets the discovery of SINPs from the entire NP dataset using a seed of 25 randomly selected NPs (Fig. 1C). The second scenario targeted a more difficult search: discovery of seven SINPs among all 78 $n \geq 4$ particles and a seed composed of all $n < 4$ particles and 5 randomly selected particles from the $n \geq 4$ set (Fig. 1D). Under both scenarios, we confirmed that the new BO algorithm outperforms the "random" acquisition baseline by a substantial margin, without much depreciation in efficiency when forming small batches. Crucially, we found that the new domain-specific UCB strategy delivers up to around 10 and 40% acceleration relative to BO optimization with standard UCB methods over the compositional axis in the first and second SINP discovery scenarios, respectively (fig. S2). Overall, having built enough confidence in our BO methodology for SINPs, we next turned to actively deploying it to run a 4-week-long closed-loop experimental campaign.

Closed-loop synthesis and characterization

The workflow and results of the closed-loop iteration between experiment and BO algorithm for quaternary SINP discovery are described in Fig. 2, along with the results of a final exploratory step in Fig. 3, focusing on alternative composition-interface targets. Each round was initiated by a group of ranked target composition suggestions from the BO algorithm, which were then synthesized via a typical SPBCL NP synthesis experiment on a silicon nitride TEM membrane. The success of the synthesis experiments and the competency of the predictions were evaluated through STEM analysis. If the synthesis was successful, meaning single particles were formed and the resulting composition was near the suggested composition, three or more representative particles were quantitatively characterized by energy-dispersive x-ray spectroscopy (EDS) to determine their elemental compositions and distributions. The experimental data were then fed back into the algorithm database,
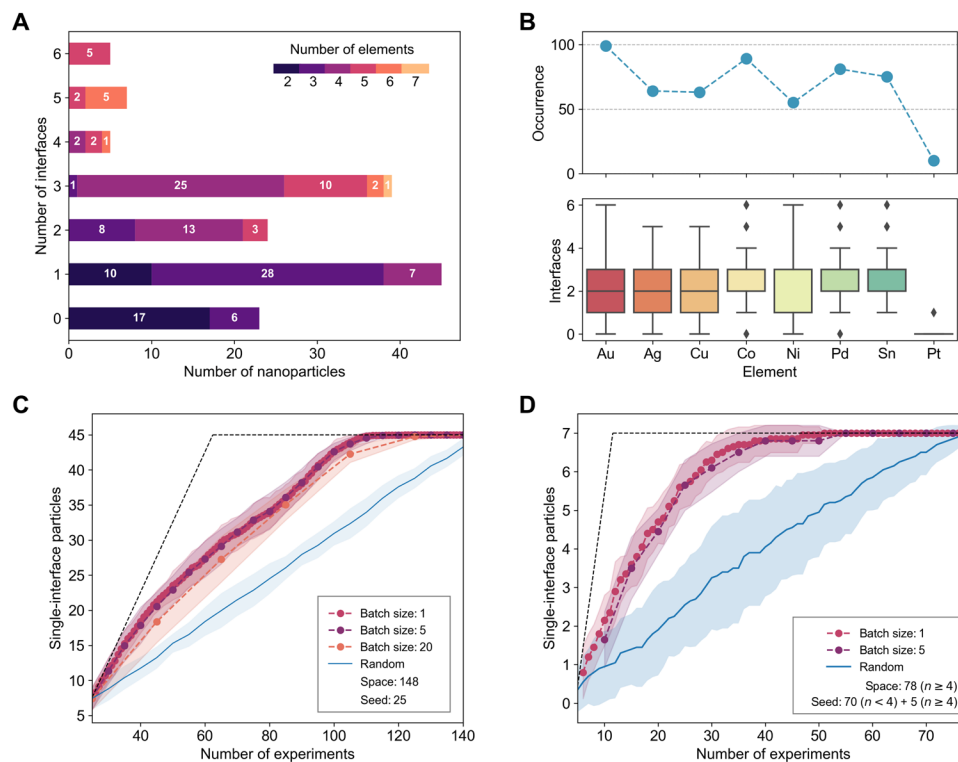
Fig. 1. Statistics of the NP dataset and simulated optimization campaigns. (**A**) Distribution of the interface counts in the initial $N = 148$ NP dataset. Each interface count is broken down into segments on the basis of the number of elements making up the particle. (**B**) Occurrence of each of the eight elements in the initial dataset and the corresponding interface count box plots for each element. Performance results for the acquisition algorithm targeting single-interface particles, simulated using the existing dataset under two different scenarios in (**C**) and (**D**), at different batch sizes and compared to the random acquisition baseline. The performance (C) with a seed of 25 randomly selected particles from the space of 148 available, with the remaining space searched by the algorithm to find single-interface particles, and (D) with a seed made of 70 particles with $n < 4$ elements and 5 randomly selected from those with $n \geq 4$, with the remaining space searched by the algorithm to find single-interface particles with $n \geq 4$. Dashed lines show ideal performance limits and shaded regions show variability, limited to $\pm\,\sigma$ for visual clarity.

and new composition suggestions were generated (see Material and Methods).

To first test the efficacy of the algorithm, we selected quaternary SINPs as our first target structure for reasons discussed above. Three rounds of suggestion and synthesis were performed in this space, wherein we synthesized the top 4 ranked suggested compositions, and of the 12 suggested compositions, 11 were successfully prepared and formed the target structure, with one synthetic failure. After only one round of suggestions, the algorithm moved away from simple permutations of an existing quaternary SINP composition and suggested novel compositions, which are visualized in Fig. 2. We find that the algorithm's predictions are strongly influenced by the experimental feedback. For example, three of the four highest ranked suggestions in the second batch differed from the top 5 to 8 suggestions in the first batch, which would have become batch 2's top 4 if the rankings were not influenced by the new results. Figure 4 provides an alternative view of the 8D composition space by projecting the high-dimensional compositional embedding space used in our BO method to a more intuitive 2D visualization (see Material and Methods). First, we observe that the data in our initial dataset are distributed reasonably well throughout the large, high-dimensional design space we are searching. We also find that the algorithm targets a certain part of the chemical space in the first two rounds, exploiting regions near existing SINPs, although in the third and exploration

rounds (explained below in more detail), it ventures out of that region. On the basis of the projections of our EDS measured average compositions on the same 2D representation in Fig. 4, apart from a few visible deviations, we can track the BO suggestions in the chemical design space reasonably well.

Typical SINPs have an ellipsoidal shape with a single, straight interface whose location depends on the relative amounts of material present in the constituent phases (see, for example, Fig. 2B); however, particularly in larger particles, this desired structure is not always achieved. During particle coarsening, it is possible for smaller particles to attach to each other at different locations, leading to particles whose second phase is discontinuous or inhomogeneously distributed (Fig. 2, A1). Note that the discontinuous phase is always the one consisting of the lower-mobility metals, such as Ni and Co, which naturally take longer to form thermodynamic structures during low-temperature annealing. As a result, we hypothesized that these structures could gradually be transformed into their thermodynamic analogs through prolonged annealing. During annealing, the structure gradually approaches those that are thermodynamically favored (fig. S3). Therefore, we conclude that these particles can, for the purposes of identifying their interfaces, be treated as equivalent to ones already exhibiting the desired thermodynamic structures.

After the first three iterations of the model system resulted in competent suggestions, we explored several unprecedented phase
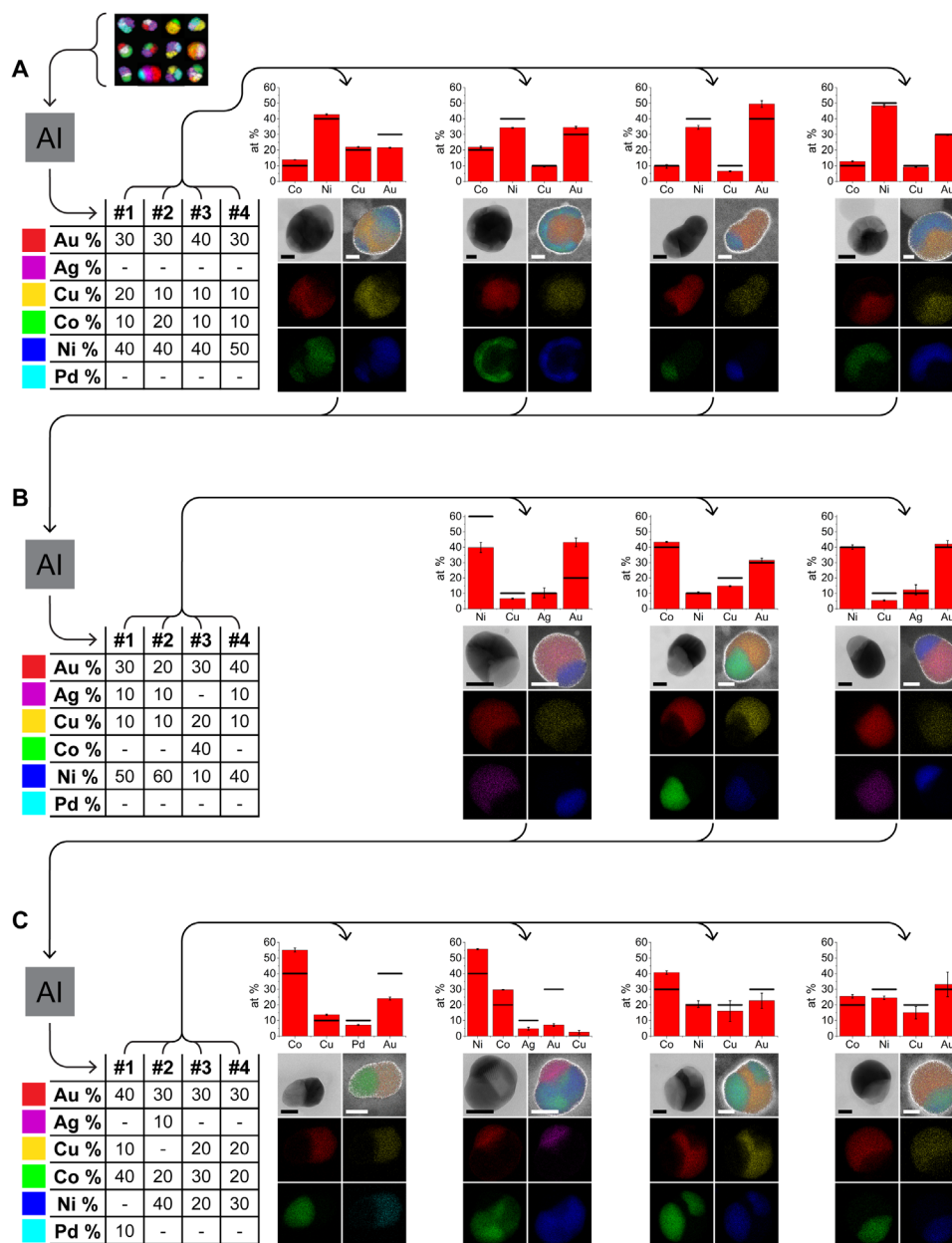
Fig. 2. Closed-loop optimization for the discovery of quaternary metallic SINPs. (**A** to **C**) Iterations 1 to 3, respectively, of the closed feedback loop process. Starting from available input data, the top compositions suggested by the machine learning algorithm are tabulated and then synthesized by SPBCL. The resulting experimental compositions, as determined by STEM-EDS, are then fed back into the algorithm for the next set of predictions. Bar graphs show elemental compositions determined by EDS averaged across multiple NPs, with error bars showing the SD and black bars representing the suggested targets. Below are annular bright-field (ABF) images and EDS maps of a representative particle for each composition. Scale bars, 20 nm. Note that particle C2 contains a small amount of Cu contamination. at %, atomic %.

spaces to test the capabilities of the algorithm (Fig. 3). In particular, we changed our targets to ones that had no analogs in the existing dataset (i.e., quinary and senary SINPs), forcing the algorithm to extrapolate from the existing data. Starting with quinary SINPs, the algorithm's top 5 suggested candidates were all variations of AuCuCoNiPd; however, due to a scientific curiosity and a desire to maximize the search space, we selected the two highest ranked variations of AuCuCoNiPd, the two highest ranked variations of AuAgCuCoNi, and one variation of AuCuCoNiPd that was markedly different

than the other suggested compositions. Again, in each case, the algorithm provided competent suggestions (Fig. 3A and fig. S4), although there was no existing data to draw from. A related previously reported AuAgCuCoNi composition was a triphasic, two-interface particle (*22*), yet the suggested compositions reported here all yielded the target biphasic structure, highlighting the capabilities of the algorithm to explore different regions of phase space.

Next, we went multiple steps further by targeting both a senary SINP and two-interface particles. We selected the highest ranked
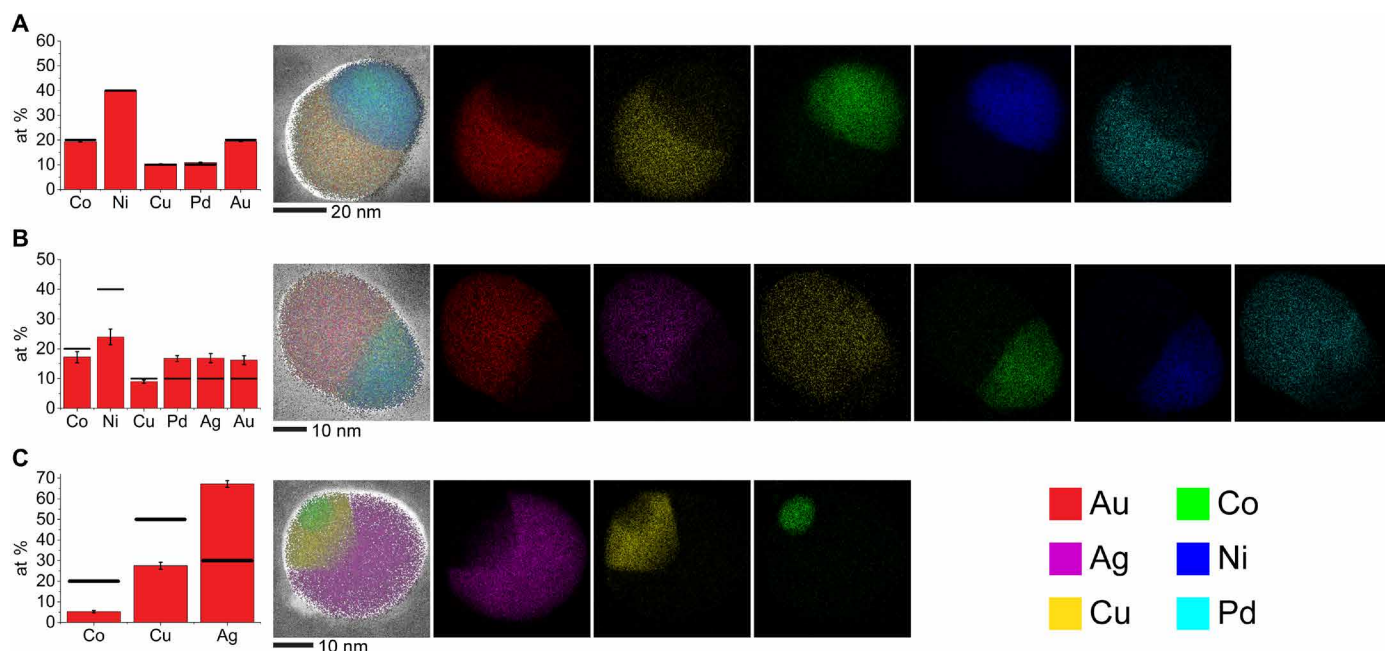
Fig. 3. STEM-EDS analysis of exploratory NP compositions suggested by the optimization agent. Elemental compositions and EDS maps of a representative (**A**) quinary SINP, (**B**) senary SINP, and (**C**) ternary two-interface NP. Bar graphs show elemental compositions determined by EDS averaged across multiple NPs, with error bars showing the SD and black bars representing the suggested targets.
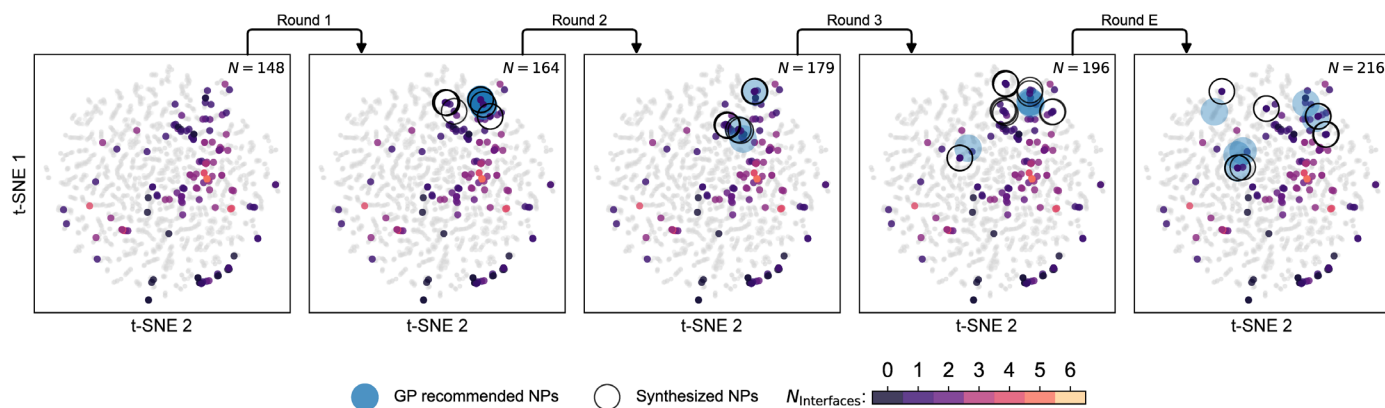

Fig. 4. Characteristics of the 8D NP space and its BO-guided exploration for discovery of SINPs. This 2D projection of our composition-embedded BO design space (see Material and Methods) is generated using the *t*-distributed stochastic neighbor embedding (t-SNE) method. Gray circles represent the entire compositional search space of our 8D compositional grid (see Material and Methods). Experimental data (both initial and acquired in the loop) are color-coded with the measured interface counts. Acquisition suggestions by the algorithm (large open circles) and synthesized suggestions (large blue circles) are overlaid on the same projection to provide visual insights on how the search space is traversed with the aid of the BO algorithm. Round E corresponds to the "exploratory" search toward modified targets carried out upon completion of the closed-loop iteration for quaternary SINPs.

suggestion for the senary SINP from the BO model conditioned on all data that we collected up to this point, and we successfully synthesized, to the best of our knowledge, the most chemically complex single-interface metallic NP ever prepared (Fig. 3B). In our attempt to synthesize two-interface particles, we selected the highest ranked permutation of an AgCuCo composition (again with BO conditioned on all past data), which successfully formed the target structure (Fig. 3C).

During this final effort to explore two-interface NPs, a particular high-ranking suggestion in the Au-Ag-Cu space caught our attention as it conflicted with our intuition and bulk phase diagrams and, hence, would serve as a case study for understanding the underlying algorithmic and experimental causes of these conflicts. As shown in Fig. 5A, we confirmed that NPs synthesized at this nominal composition were off the two-interface target. Further inspection revealed that the GP algorithm relied on two nearby points in this ternary
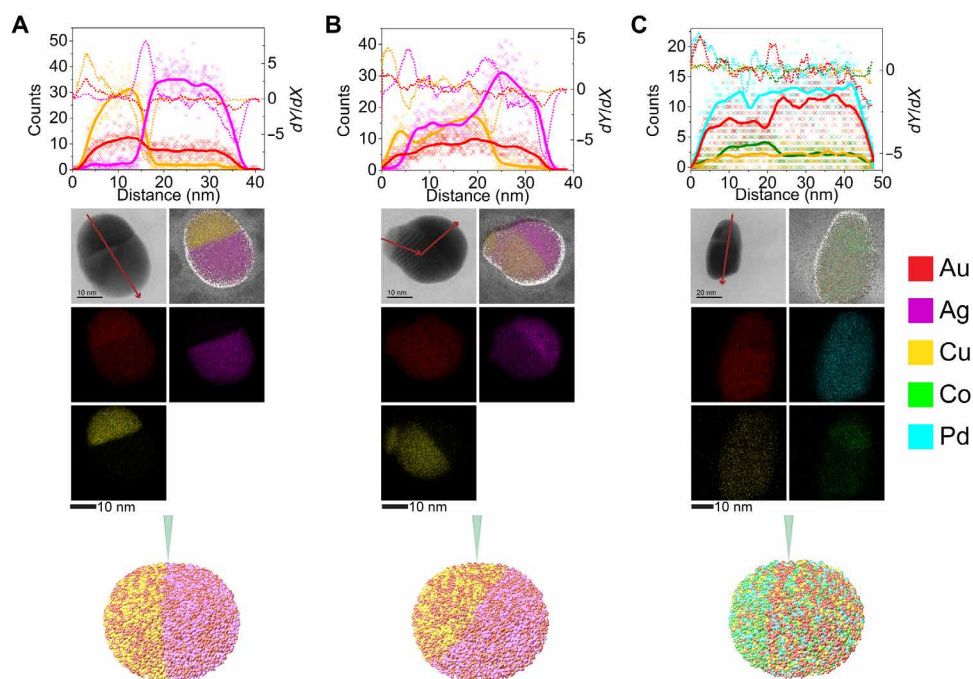
Fig. 5. STEM-EDS characterization of interfaces in different biphasic NPs. EDS line profiles taken along the red arrows in the ABF images, corresponding EDS maps, and schematic models of particle orientation with respect to the electron beam of (**A** and **B**) $Au_{10}Ag_{30}Cu_{60}$ and (**C**) $Au_{30}Cu_{10}Co_{20}Pd_{40}$ NPs. Crosses and solid curves represent raw and smoothed data, respectively, and dotted curves are the corresponding gradients.

space in our dataset with a larger than average uncertainty in composition and phase structure. Characterization of numerous AuAgCu particles reveals that the inherently random orientation of the synthesized NPs on a surface can easily lead to ambiguities in characterization (Fig. 5, A and B) when interfaces are not oriented along the optical axis of the microscope, making a region of two overlapping phases indistinguishable from a third phase by imaging or EDS. Moiré fringes can indicate phase overlap (e.g., Fig. 5B and fig. S3C), but their absence does not prove the existence of another phase. Peaks in gradients of compositional line profiles can help locate interfaces [e.g., one versus two peaks in Fig. 5 (A and B), respectively], but by aligning its interface with the optical axis, the particle in Fig. 5B is identified as having a single interface (fig. S3E). Overall, without reliable statistical averages over the particle orientations or additional characterization techniques, particles such as those in Fig. 5 (A and B) can easily be interpreted as having more than one type of interface, leading to ambiguity in input data. We confirmed that, upon removal of the two such suspect AuAgCu points from our NP dataset, the BO algorithm no longer suggests any AuAgCu compositions for a two-interface target (not one in the top 100 suggestions in ternary space). The addition of the new AuAgCu data does not immediately eliminate this off-target suggestion, indicating that it would require more BO iterations, and hence, keeping the researcher in the loop for the identification and elimination of "bad" data is a more cost-effective remedy.

We emphasize, however, that homogeneously distributed elements in high-order NPs can bring about more and nearly intangible interfaces (Fig. 5C). In certain cases, large percentages of miscible elements lead to these "soft" interfaces, wherein the transition from one phase to the other is less obvious. The examples of AuAgCu

(Fig. 5A) and AuCuCoPd (Fig. 5C) illustrate the "hard" and soft interface extremes, respectively. This observation highlights shortcomings in defining what constitutes an interface. Clearly, a discontinuous change in crystal structure is an interface, but for coherent interfaces with continuous changes in structure, there is no universally accepted answer for exactly how steep a gradient in composition and/or lattice parameter must be to be classified as such. For the purposes of this work, all interfaces are classified equivalently, but the nature of the interface poses an interesting structural motif that, with the requisite characterization capabilities, could be investigated in more detail.

Considering the success of our closed-loop implementation and our in-depth inspection of the AuAgCu case, we identify several opportunities for improvement for BO-driven NP-interface discovery: (i) acquiring reliable data for building predictive models for BO, (ii) keeping the "researcher in the loop" when the experiment throughput allows for inspection of suggestions and interventions, and (iii) improving the algorithmic capabilities in all stages of BO. Methods such 4D STEM would enable complete mapping of crystal structure throughout individual particles and, therefore, remove ambiguities otherwise present in EDS maps and allow for simultaneous identification of composition, morphology, and crystal structure. While researchers can always intervene, algorithmic improvements that would bring ambiguous input data to their attention would make interventions and, in turn, overall BO more efficient. Although our domain-specific, composition-embedded BO strategy improves upon more standard UCB methods, the optimization methodology certainly has room for improvement, particularly in active research areas (*31*, *32*) such as accounting for the uncertainty in input compositions and targets in BO.

DISCUSSION

In this work, we have demonstrated that machine learning–assisted optimization is a viable approach to identifying nanostructures with target morphological qualities, even with limited datasets. This platform allowed us to synthesize new and complex materials, including, to the best of our knowledge, the most complex biphasic NP ever made, with success rates far greater than random selection (18 of 19) and, in more complex circumstances, greater than a keen chemical intuition. The limitations for expansion of such a design feedback loop to higher-throughput synthesis and experimentation, and into more complex properties, such as catalytic activity or stability, are almost exclusively experimental. In this work, we synthesize one material at a time using dip-pen nanolithography (DPN) to accommodate the electron microscopy characterization; parallelization of SPBCL to synthesize thousands or millions of unique materials in a single experiment is possible and has been demonstrated (2). As we develop new characterization techniques to extract more information from megalibraries of nanomaterials at higher throughputs, and as targeted properties become increasingly more complex in future work, the benefits of using such a machine learning aid to guide the search will increase exponentially, as the capabilities of an algorithm can scale unlike those of human intuition.

MATERIALS AND METHODS

Materials and chemicals

Chloroauric acid hydrate (HAuCl$_4$ xH$_2$O; ≥99.9% trace metals basis), cobalt(II) nitrate hexahydrate [Co(NO$_3$)$_2$ 6H$_2$O; 99.999%], copper(II) nitrate hydrate [Cu(NO$_3$)$_2$ xH$_2$O; 99.999%], nickel(II) nitrate hexahydrate [Ni(NO$_3$)$_2$ 6H$_2$O; 99.999%], palladium(II) nitrate dihydrate [Pd(NO$_3$)$_2$ 2H$_2$O; ~40% Pd basis], silver nitrate (AgNO$_3$; anhydrous, ≥99.999% trace metals basis), nitric acid (HNO$_3$; 70%, redistilled, ≥99.999% trace metals basis), hexamethyldisilazane (HMDS; 99.9%), and hexane (anhydrous, 95%) were purchased from Sigma-Aldrich. Poly(ethylene oxide)-*block*-poly(2-vinylpyridine) (PEO-*b*-P2VP; M_n = 2.8-*b*-1.5 kg/mol, polydispersity index = 1.11) was purchased from Polymer Source Inc. DPN 1D pen arrays (type M, without a gold coating) were purchased from Advanced Creative Solutions Technology LLC. Silicon nitride membranes for TEM (amorphous, thickness of 15 nm) were purchased from Ted Pella Inc.

NP synthesis

Silicon nitride membranes were functionalized with HMDS by vapor deposition in a closed chamber containing an HMDS/hexanes mixture (1:1 v/v) overnight at room temperature; rinsed with toluene, isopropanol, and acetone; and dried with N$_2$. Precursor inks were prepared by dissolving PEO-*b*-P2VP and metal precursors at desired ratios in deionized water, such that the final polymer concentration was 5.0 mg/ml and the total metal to pyridyl group molar ratio was 1:32. The inks were then acidified with 2.0 μl of 1.0 M HNO$_3$ per milliliter of ink and left to shake overnight in the dark. DPN pen arrays were plasma-treated at 30 W for 60 s and mounted onto an atomic force microscopy (Park Systems XE-150) in a chamber kept at 20°C and a relative humidity of 80%. Pen arrays were inked by bringing the pen array into contact with a dropcast ink reservoir, and nanoreactors were deposited onto the functionalized silicon nitride membranes by repeatedly bringing the cantilever tips in contact with the surface. Samples were thermally annealed in a tube furnace with an H$_2$ flow of 100 standard cubic centimeters per minute

using the following protocol: ramped to 150°C at 2.2°C/min, held at 150°C for 6 hours, ramped to 300°C at 2.5°C/min, held at 300°C for 6 hours, ramped to 500°C at 3.3°C/min, held at 500°C for 12 hours, and then cooled to room temperature.

NP characterization

All STEM and EDS characterization was carried out on a JEOL JEM-ARM200CF equipped with a cold field-emission gun operated at 200 kV, a CEOS hexapole probe aberration corrector, and two EDS detectors using a JEOL beryllium holder. High-angle annular dark-field (HAADF) and annular bright-field (ABF) images were acquired with a convergence angle of 20.6 mrad and a collection angle of 68 to 280 mrad (HAADF) or 7.5 to 17 mrad (ABF). The Lα peaks of Ag, Au, and Pd and the Kα peaks of Co, Cu, and Ni were used for EDS quantification and mapping. Elemental maps were generated from at least 100 frames with a pixel dwell time of 20 μs and processed to show net x-ray counts using a 3 × 3 kernel with one full pixel overlap. EDS quantification was carried out using the standardless Cliff-Lorimer correction method with the open source Python library HyperSpy (33). For each composition, multiple NPs were quantified to obtain a representative average composition.

BO algorithm design

Our BO strategy derives from the GP-UCB method (34). In GP-UCB, the next point of acquisition x_t at step t from a decision space D ($x \in D$) in a sequence of experiments is decided as x_t = argmax $[\mu_{t-1}(x) + \sqrt{\beta_t} \, \sigma_{t-1}(x)]$, which provides an exploration-exploitation trade-off by combining the predictions, $\mu(x)$, with the predicted uncertainties $\sigma(x)$ weighted by $\sqrt{\beta_t}$. Decisions made using this approach after the first selection in a batch are made under information poor conditions and may lead to over- or underexploration of parts of D, which can, in part, be mitigated by reconditioning the GP using the past selections in the batch (35, 36) to update the uncertainty estimates.

In our context, NP optimization is to be carried out in an 8D space where $x \coloneqq [c_{Ag}, c_{Au}, \ldots, c_{Sn}]$ and c_i is the molar fraction of element i. In practice, many compositions x are confined to a lower-dimensional subspace of D, and hence, information from fully or partially disjoint subspaces (e.g., Au-Cu-Pd versus Ag-Ni-Sn) cannot be used efficiently in BO. We therefore map x to an embedding space r using a transformation g that converts the design axes from elemental compositions to descriptors derived from stoichiometry and properties of elements, hence shared across all NPs. Because the relevance of descriptors is not known a priori, for each batch B, g is parameterized by data available up to that point (t' being the latest measurement; $t' < t$) and used to dynamically generate the embedding space $r_B(x) = g(x; x_1, x_2, \ldots, x_{t'})$ from x. Thus, to compose batch B, acquisitions are decided as

$$r_{B,t} = \text{argmax} \, [\mu_{t-1}(r_B) + \sqrt{\beta_t} \, \sigma_{t-1}(r_B)]$$

and repeated by reconditioning GP with past selections until the desired batch size is reached. To recast our target-interface-count problem as maximization, we define $y_t \coloneqq -|y'_t - n|$ as our BO target, where the STEM interface count of a particle is y'_t and n is the targeted interface count (e.g., n = 1 for SINPs).

We parameterize g for a batch B as follows: (i) create a large set of m descriptors for each composition $x_1, x_2, \ldots, x_{t'}$ measured until B; a $t' \times m$ matrix of scalars; (ii) standardize each column; and (iii) find a transformation to a more compact N_{PCA}-dimensional embedding

space using principal components analysis (PCA). This parameterized g is then used to transform the search space $\boldsymbol{x}$ to $\boldsymbol{r}_B$ for batch B. N_{PCA} is tuned a priori with the aim of finding a small enough space ($N_{\text{PCA}} \ll m$), where GP provides sufficient accuracy and BO can be performed efficiently.

Hyperparameters of the optimization framework

We use the compositional descriptors ($m = 132$) described by Ward *et al.* (*37*, *38*), which are derived from statistical aggregations of stoichiometries and various elemental properties, in our implementation of g. We select $N_{\text{PCA}} = 20$ for the active closed-loop experiment-algorithm iterations on the basis of fig. S1 (see Supplementary Text). We use the radial basis function kernel $k(\boldsymbol{r},\boldsymbol{r}') = \sigma^2 \exp\left(-\frac{1}{2\ell^2}\|\boldsymbol{r} - \boldsymbol{r}'\|^2\right)$ with all GP models, where σ, ℓ, and the noise variance are optimized using the *bfgs* algorithm. Parameter β_t is dynamically set using the theorem 1 of (*34*) with a scaling factor of 0.05. We use the GPy library for implementation of GPs in BO, with the rest of the machine learning and data transformation steps handled through the scikit-learn library. The BO workflow is carried out using the camd python library.

Compositional search space design

For the closed-loop optimization, we generate a discretized search space D by creating a grid of compositions with 10% increments on each composition axis, i.e., Ag, Au, Cu, Co, Ni, Pd, and Sn. We use data from Pt-bearing particles in the BO seed, and hence, Pt is also included as a composition axis in D, but these particles are not targeted in the search itself. We remove compositions from the search space that are closer than 5% in at least one of the composition axes to those that already exist in our initial NP dataset and those that are sent to and received from experimental workflow in each iteration.

REFERENCES AND NOTES

1. C. Xie, Z. Niu, D. Kim, M. Li, P. Yang, Surface and interface control in nanoparticle catalysis. *Chem. Rev.* **120**, 1184–1249 (2020).
2. E. J. Kluender, J. L. Hedrick, K. A. Brown, R. Rao, B. Meckes, J. S. Du, L. M. Moreau, B. Maruyama, C. A. Mirkin, Catalyst discovery through megalibraries of nanomaterials. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 40–45 (2019).
3. A. Ludwig, Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods. *NPJ Comput. Mater.* **5**, 70 (2019).
4. X.-D. Xiang, X. Sun, G. Briceño, Y. Lou, K.-A. Wang, H. Chang, W. G. Wallace-Freedman, S.-W. Chen, P. G. Schultz, A combinatorial approach to materials discovery. *Science* **268**, 1738–1740 (1995).
5. H. C. Herbol, M. Poloczek, P. Clancy, Cost-effective materials discovery: Bayesian optimization across multiple information sources. *Mater. Horiz.* **7**, 2113–2123 (2020).
6. L. Huang, P.-C. Chen, M. Liu, X. Fu, P. Gordiichuk, Y. Yu, C. Wolverton, Y. Kang, C. A. Mirkin, Catalyst design by scanning probe block copolymer lithography. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3764–3769 (2018).
7. A. Mannodi-Kanakkithodi, M. K. Y. Chan, Computational data-driven materials discovery. *Trends Chem.* **3**, 79–82 (2021).
8. A. R. Oganov, G. Saleh, A. G. Kvashnin, *Computational Materials Discovery* (Royal Society of Chemistry, 2018), pp. 456.
9. M. Witman, G. Ek, S. Ling, J. Chames, S. Agarwal, J. C.-C. Wong, M. Allendorf, M. Sahlberg, V. Stavila, Data-driven discovery and synthesis of high entropy alloy hydrides with targeted thermodynamic stability. *Chem. Mater.* **33**, 4067–4076 (2021).
10. M. E. Davis, Z. Chen, D. M. Shin, Nanoparticle therapeutics: An emerging treatment modality for cancer, in *Nanoscience and Technology* (Co-Published with Macmillan Publishers Ltd, 2009), pp. 239–250.
11. J. Liu, H. He, D. Xiao, S. Yin, W. Ji, S. Jiang, D. Luo, B. Wang, Y. Liu, Recent advances of plasmonic nanoparticles and their applications. *Materials* **11**, 1833 (2018).
12. L. Liu, A. Corma, Metal catalysts for heterogeneous catalysis: From single atoms to nanoclusters and nanoparticles. *Chem. Rev.* **118**, 4981–5079 (2018).
13. J. A. Trindell, Z. Duan, G. Henkelman, R. M. Crooks, Well-defined nanoparticle electrocatalysts for the refinement of theory. *Chem. Rev.* **120**, 814–850 (2020).
14. W. Zhu, R. Michalsky, Ö. Metin, H. Lv, S. Guo, C. J. Wright, X. Sun, A. A. Peterson, S. Sun, Monodisperse Au nanoparticles for selective electrocatalytic reduction of CO_2 to CO. *J. Am. Chem. Soc.* **135**, 16833–16836 (2013).
15. L. Bu, N. Zhang, S. Guo, X. Zhang, J. Li, J. Yao, T. Wu, G. Lu, J.-Y. Ma, D. Su, X. Huang, Biaxially strained PtPb/Pt core/shell nanoplate boosts oxygen reduction catalysis. *Science* **354**, 1410–1414 (2016).
16. Y. Yamada, C.-K. Tsung, W. Huang, Z. Huo, S. E. Habas, T. Soejima, C. E. Aliaga, G. A. Somorjai, P. Yang, Nanocrystal bilayer for tandem catalysis. *Nat. Chem.* **3**, 372–376 (2011).
17. W.-T. Koo, J. E. Millstone, P. S. Weiss, I.-D. Kim, The design and science of polyelemental nanoparticles. *ACS Nano* **14**, 6407–6413 (2020).
18. P.-C. Chen, M. Liu, J. S. Du, B. Meckes, S. Wang, H. Lin, V. P. Dravid, C. Wolverton, C. A. Mirkin, Interface and heterostructure design in polyelemental nanoparticles. *Science* **363**, 959–964 (2019).
19. R. Ghosh Chaudhuri, S. Paria, Core/shell nanoparticles: Classes, properties, synthesis mechanisms, characterization, and applications. *Chem. Rev.* **112**, 2373–2433 (2012).
20. T. Li, Y. Yao, Z. Huang, P. Xie, Z. Liu, M. Yang, J. Gao, K. Zeng, A. H. Brozena, G. Pastel, M. Jiao, Q. Dong, J. Dai, S. Li, H. Zong, M. Chi, J. Luo, Y. Mo, G. Wang, C. Wang, R. Shahbazian-Yassar, L. Hu, Denary oxide nanoparticles as highly stable catalysts for methane combustion. *Nat. Catal.* **4**, 62–70 (2021).
21. P.-C. Chen, G. Liu, Y. Zhou, K. A. Brown, N. Chernyak, J. L. Hedrick, S. He, Z. Xie, Q.-Y. Lin, V. P. Dravid, S. A. O'Neill-Slawecki, C. A. Mirkin, Tip-directed synthesis of multimetallic nanoparticles. *J. Am. Chem. Soc.* **137**, 9167–9173 (2015).
22. P.-C. Chen, X. Liu, J. L. Hedrick, Z. Xie, S. Wang, Q.-Y. Lin, M. C. Hersam, V. P. Dravid, C. A. Mirkin, Polyelemental nanoparticle libraries. *Science* **352**, 1565–1569 (2016).
23. P.-C. Chen, J. S. Du, B. Meckes, L. Huang, Z. Xie, J. L. Hedrick, V. P. Dravid, C. A. Mirkin, Structural evolution of three-component nanoparticles in polymer nanoreactors. *J. Am. Chem. Soc.* **139**, 9876–9884 (2017).
24. P.-C. Chen, Y. Liu, J. S. Du, B. Meckes, V. P. Dravid, C. A. Mirkin, Chain-end functionalized polymers for the controlled synthesis of sub-2 nm particles. *J. Am. Chem. Soc.* **142**, 7350–7355 (2020).
25. P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon, W. C. Chueh, Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020).
26. B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, Taking the human out of the loop: A review of bayesian optimization. *Proc. IEEE* **104**, 148–175 (2016).
27. B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, A. G. Doyle, Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590**, 89–96 (2021).
28. F. Häse, L. M. Roch, C. Kreisbeck, A. Aspuru-Guzik, Phoenics: A Bayesian optimizer for chemistry. *ACS Cent. Sci.* **4**, 1134–1145 (2018).
29. J. H. Montoya, K. T. Winther, R. A. Flores, T. Bligaard, J. S. Hummelshøj, M. Aykol, Autonomous intelligent agents for accelerated materials discovery. *Chem. Sci.* **11**, 8517–8532 (2020).
30. T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, K. Tsuda, COMBO: An efficient Bayesian optimization library for materials science. *Mater. Discov.* **4**, 18–21 (2016).
31. M. Aldeghi, F. Häse, R. J. Hickman, I. Tamblyn, A. Aspuru-Guzik, Golem: An algorithm for robust experiment and process optimization. arXiv:2103.03716 (2021).
32. R. Oliveira, L. Ott, F. Ramos, paper presented at the Proceedings of Machine Learning Research, 2019.
33. F. de la Peña, E. Prestat, V. T. Fauske, P. Burdet, T. Furnival, P. Jokubauskas, M. Nord, T. Ostasevicius, K. E. MacArthur, D. N. Johnstone, M. Sarahan, J. Lähnemann, J. Taillon, T. Aarholt, V. Migunov, A. Eljarrat, J. Caron, S. Mazzucco, B. Martineau, S. Somnath, T. Poon, M. Walls, T. Slater, N. Tappy, N. Cautaerts, F. Winkler, G. Donval, J. C. Myers. (Zenodo, 2020).
34. N. Srinivas, A. Krause, S. M. Kakade, M. Seeger, Gaussian process optimization in the bandit setting: No regret and experimental design. arXiv:0912.3995 (2009).
35. E. Contal, D. Buffoni, A. Robicquet, N. Vayatis, in *Machine Learning and Knowledge Discovery in Databases* (Springer Berlin Heidelberg, 2013), pp. 225–240.
36. T. Desautels, A. Krause, J. W. Burdick, Parallelizing exploration-exploitation tradeoffs in Gaussian process bandit optimization. *J. Mach. Learn. Res.* **15**, 3873–3923 (2014).
37. L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials. *NPJ Comput. Mater.* **2**, 16028 (2016).

38. L. Ward, A. Dunn, A. Faghaninia, N. E. R. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, A. Jain, Matminer: An open source toolkit for materials data mining. *Comput. Mater. Sci.* **152**, 60–69 (2018).