

PERSPECTIVE OPEN

Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods

Alfred Ludwig ^{1,2}

This perspective provides an experimentalist's view on materials discovery in multinary materials systems—from nanoparticles over thin films to bulk—based on combinatorial thin-film synthesis and high-throughput characterization in connection with high-throughput calculations and materials informatics. Complete multinary materials systems as well as composition gradients which cover all materials compositions necessary for verification/falsification of hypotheses and predictions are efficiently fabricated by combinatorial synthesis of thin-film materials libraries. Automated high-quality high-throughput characterization methods enable comprehensive determination of compositional, structural and (multi)functional properties of the materials contained in the libraries. The created multidimensional datasets enable data-driven materials discoveries and support efficient optimization of newly identified materials, using combinatorial processing. Furthermore, these datasets are the basis for multifunctional existence diagrams, comprising correlations between composition, processing, structure and properties, which can be used for the design of future materials.

npj Computational Materials (2019)5:70; <https://doi.org/10.1038/s41524-019-0205-0>

INTRODUCTION

The discovery of new (multinary) materials and their rapid optimization are necessary to enable future technological developments in areas such as sustainable energy technologies and energy-efficient processes: e.g., the envisaged hydrogen economy relies on materials discoveries like light-absorbing, catalytic and stable photoelectrodes for solar water splitting, lightweight hydrogen storage materials or noble-metal-free fuel cell catalysts. Generally, the potential for materials discoveries is high, as the periodic table offers numerous elements which can be combined in multinary material systems, offering an almost unlimited, and yet mostly unexplored search space for new materials. This is simultaneously a promise and a challenge as not all possible materials can be explored, even with the most efficient high-throughput technologies. Therefore, materials systems to be investigated in experiments need to be carefully chosen. The starting chemical elements for potential new materials can be restricted to 40 or 50 which are earth-abundant, non-toxic and sustainable.

However, engineering materials, i.e., materials which are used in applications, are usually multinary, consisting of about 10 elements, e.g., steels, metallic glasses, superalloys, so-called "high-entropy alloys" and multiple principal element alloys. Therefore, it can be expected that future engineering materials will be rather multinary than consisting of only one or two elements. However, already the number of possible combinations of chemical elements in multinary systems is immense (e.g., for quaternaries out of 50 starting elements there are more than two million combinations). Furthermore, the compositions in the

compounds (e.g., already 100 in a binary, for 1 at.% increments), phase constitution (typically multiphase materials are observed) and microstructure need to be finely tuned to specific applications to reconcile many, often contradicting, requirements. The already large unexplored composition space is thus further expanded by the processing parameter space: materials have to be synthesized and many different possibilities for processing lead to different material states (metastable, stable), microstructures, and properties, which need to be explored. Extrinsic properties of materials can be tailored by controlling microstructure using processing. Thus, the multi-dimensional search space for materials discoveries comprises both intrinsic (e.g., composition, crystal structure, functional properties (e.g., saturation magnetization in ferromagnets)) and extrinsic properties (e.g., microstructure, functional properties (e.g., coercivity in ferromagnets)).

As this search space is almost unlimited, a down-selection of the materials systems to be addressed in experiments is required. For this, there are several options: it can be based on intuition, prior knowledge, or predictions. By materials informatics, the usage of prior knowledge—stored in databases or extracted by computational means from electronically available literature—can be improved. Theory and computational materials science can provide hypotheses for where to search for new materials and make predictions of new materials, e.g., a list of predicted compositions. Input from high-throughput computational materials science is thus highly useful: with computational methods, thousands of systems can be screened in a shorter time than ever possible in experiments. From the computational results a feasible list of candidate materials for combinatorial exploration can be deduced, i.e., lists of tens or hundreds of materials.

¹Chair for Materials Discovery and Interfaces, Institute for Materials, Faculty of Mechanical Engineering, Ruhr University Bochum, Bochum, Germany and ²ZGH, Ruhr University Bochum, Bochum, Germany

Correspondence: Alfred Ludwig (alfred.ludwig@rub.de)

Received: 14 February 2019 Accepted: 4 June 2019

Published online: 10 July 2019

This perspective discusses the possibilities of combinatorial synthesis and high-throughput characterization in combination with computational methods in the endeavor to efficiently identify new materials in multi-dimensional search spaces, overcoming reliance on serendipity for discoveries and transition into a state of efficient and data-guided materials discovery. General considerations on materials discovery and experimental possibilities in thin-film combinatorial materials science (CMS) are discussed. Examples of discoveries using this approach, also in combination with computational materials science are given and the further expansion of CMS with computational methods and materials informatics which is necessary to achieve efficient materials discovery and design of materials is discussed.

MATERIALS DISCOVERY IN MULTIDIMENSIONAL SEARCH SPACES: FROM SERENDIPITOUS TO DATA-GUIDED DISCOVERY

A materials discovery could be defined as the event that a previously unknown combination of composition, crystal structure, phase constitution, microstructure and properties, i.e., a new phase or phase mixture with special properties, is identified. This is separate from materials development, where a known, i.e., already-discovered material is optimized, e.g., by compositional or processing adjustments to extend e.g., its application temperature or stability in reactive environments. In an (inverse) design of materials one would take the wanted properties of a material as the starting point and then design the material which shows these properties, e.g.¹ In contrast, a general discovery of materials explores as many as possible materials and delivers data-sets which are necessary for the design of new materials.² While exploring as many as possible materials for as many as possible properties, materials discoveries can occur. As a very high number of undiscovered materials should exist, it is expected that some of them will show unique properties. However, the multidimensional search space is immensely large, and it is unclear where to search for discoveries, making materials discovery very challenging. This might be a reason that most research is centered around known materials or base elements such as Fe, Ni, etc., limiting new discoveries.

What can we learn from materials discoveries in the past? Some of them were made serendipitously, i.e., researchers were aiming for a certain research goal, but made an unexpected interesting observation, which led them to a discovery. Here, it depends on the skills of the researchers to differentiate erroneous measurements, which only look like a discovery, from real discoveries. An example of a serendipitous discovery is the shape memory alloy NiTi.³ The discovery of quasicrystals⁴ was also related to unexpected experimental data. Other discoveries were related to surprising properties, found in materials where these properties were not at all expected: e.g., high-temperature superconductivity in oxides,⁵ superconductivity in MgB₂,⁶ metallic conductivity in special polymers,⁷ etc. The author and colleagues recently discovered a noble-metal free nanoparticulate electrocatalyst, CrMnFeCoNi, with catalytic activity for the oxygen reduction reaction.⁸ This unexpected discovery could only be made as new combinatorial synthesis and characterization methods were applied, and a multinary system was tested for a property where no results were reported before.

Thus, discoveries can be expected in (I) (compositionally) unexplored search spaces, by fabricating and characterizing parts of the unexplored composition space and (II) by performing experiments in (known) composition spaces, by testing for special, but not yet investigated functionalities. Finally (III), extreme or very special processing conditions can lead to discoveries: e.g., high-pressure synthesis⁹ or plasma processes with activated reaction partners: By reactive co-sputtering defect-tolerant semiconducting nitrides for photovoltaic and electrochemical applications were identified.¹⁰

Another discovery route is based on theory, simulations and other computational efforts. As the materials search space cannot be explored completely, predictions are necessary in order to select the composition or composition range to be investigated. High-throughput computations offer the possibility to predict new materials. From thousands of candidate materials, those which are potentially stable and may show interesting properties can be predicted. Based upon this, the combinatorial materials science approach can be used to efficiently synthesize a limited number (10–100) composition spread materials libraries (MLs) around the most promising predicted compositions. In¹¹ the authors started from 68860 materials and identified 43 new potential photocathodes for CO₂ reduction. However, calculations are frequently limited to intrinsic properties and sufficiently precise and validated experimental data for the calculations are often lacking. Such missing datasets could be delivered by high-throughput characterization of MLs. Predictions can use prior knowledge and extrapolate to data points which have not been addressed before by synthesis experiments. E.g., a discovery endeavor for new nitrides was performed by predictions from density functional theory calculations: 21 ternary nitride semiconductors were predicted and CaZn₂N₂ was realized by high-pressure synthesis.¹²

This indicates that for a future improved discovery of materials, high-quality multidimensional data needs to be generated and evaluated with high-throughput jointly in computations and experiments.

EXPLORING MULTIDIMENSIONAL SEARCH SPACES USING COMBINATORIAL SYNTHESIS AND HIGH-THROUGHPUT CHARACTERIZATION OF THIN-FILM AND NANOPARTICLE MATERIALS LIBRARIES

In order to explore this immense search space, efficient research methods and strategies are necessary: These comprise the combinatorial fabrication of MLs, their high-throughput characterization and materials informatics for analysis and visualization of the acquired multidimensional data. A ML is a well-defined set of materials—suitable for high-throughput characterization—produced in one experiment under identical conditions. Combinatorial and high-throughput methods for materials discovery and accelerated development have been developed in the last decades, e.g.^{13–15} They involve, first, identification of the chemical materials compositions and their crystallographic structure. Second, structural and functional properties can be identified and be further optimized by combinatorial processing. Finally, high-throughput methods can help to accelerate the incorporation of new materials into systems, as it is not sufficient to develop only a material by itself. It has to be developed in a system, i.e., it has to provide functionality in connection with adjacent materials and environments. Here, interface properties are important. Whereas the combinatorial approach can also be applied in bulk,¹⁶ this perspective focuses on thin-film combinatorial methods.

Synthesis of thin-film materials libraries

Up to complete ternary materials systems or large fractions of higher-component systems can be obtained by combinatorial synthesis of thin-film MLs, see Fig. 1. Most frequently, they are deposited in form of (virtual) wedge-type films to achieve well-defined composition gradients, either by co-deposition from multiple sources or multilayer deposition of nanoscale wedge-type layers. MLs fabrication by combinatorial magnetron sputter processes, e.g.,¹⁷ is useful, as sputtering is a versatile process both for science and industry: findings from screening of sputtered MLs can be transferred to industrial (thin-film) applications. The wedge-type multilayer deposition method uses computer-controlled moveable shutters to deposit nanoscale layers, oriented

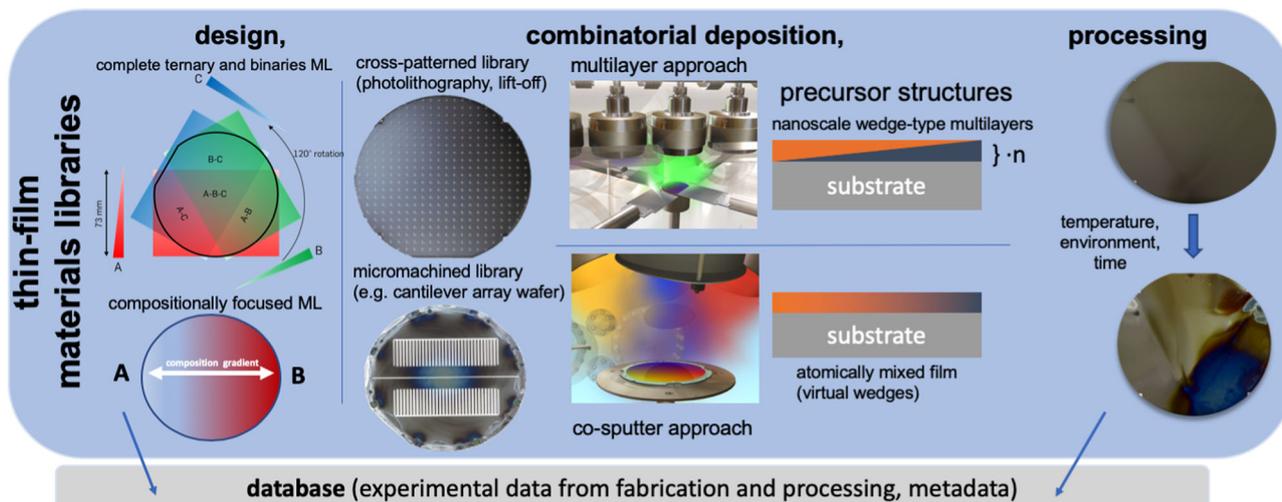


Fig. 1 Overview of thin-film ML design and fabrication. The ternary ML design uses a suitable combination of wedge-type layers with 120° symmetry to synthesize a complete ternary system, as well as its reference binary systems and the pure elements. Compositionally focused MLs can be fabricated by co-deposition. Micromachined substrates allow for special high-throughput methods (here: thickness and stress measurements). Processing in different environments is used for phase evolution. Data from design, deposition and processing of MLs should be stored in a database

to each other by 180° (binaries) or 120° (ternaries). Phase formation is achieved by post-deposition annealing at suitable temperatures where the layers can rapidly diffuse. However, the right processing (e.g., annealing) conditions need to be identified to transform the multilayered precursor structure into phases. It is also possible to synthesize “focused” compositional gradient MLs around, e.g., a predicted composition, where the composition range is tailored in a limited range. This can be achieved by co-deposition, which results in an atomic mixture in the deposited film, which makes this method suitable for the fabrication of metastable materials, if carried out at room temperature. Co-sputtered MLs contain large fractions (e.g., 20–80 at.% in a binary system) of the deposited systems.

For both approaches, due to the continuous coverage of composition space, the compositions of any predicted phases cannot be missed, as long as they are synthesizable. To achieve synthesis (if possible at all), the conditions of ML fabrication and processing need to be varied as well; i.e., the combinatorial approach needs to be extended to combinatorial processing libraries, see below.

Combinatorial deposition of MLs can also be used for studying the influence of low amounts of additional elements to base systems (“doping” studies). The thickness of the MLs should be in the range of several hundreds of nanometers to enable reliable high-throughput characterization methods and to avoid nanoscale effects (if these are not the scope of the study). Multinary nanoparticle MLs can be fabricated by co-sputtering into ionic liquids,¹⁸ which serve as liquid substrates, in which atoms are injected. If the liquids are confined in wells across a plate, a nanoparticle library is created. The nanoparticles are stabilized by suitable ionic liquids. However, next to TEM also new characterization methods need to be developed which can handle the multitude of nanoparticles of different composition made in one experiment.⁸

Creating large and consistent data-sets by high-throughput characterization of MLs

The high-throughput-high-quality characterization of multinary systems in the form of continuous composition spread thin-film MLs enables the complete and seamless determination of compositional, structural and (multi)functional properties of the

materials contained in MLs. This cannot be achieved by single-experiment approaches. The acquired large and consistent data-sets of intrinsic and extrinsic properties enable materials discoveries and efficient optimization of identified materials. Furthermore, they are the basis for multifunctional existence diagrams, comprising correlations between composition, processing, structure and properties.

Whereas in some cases completely continuous MLs are used, it is in most cases better to use a ML which is structured into measurement areas (MAs), however, without losing too much of the composition spread. In the following examples, the MLs comprise 342 MAs over a 100 mm diameter substrate (typically thermally oxidized Si wafer), see Fig. 2. The crosses, defining MAs, are fabricated by a photolithographic lift-off process to achieve a pre-patterned substrate. The numbered crosses, where no thin film is applied, are used for thickness measurements (e.g., with profilometry) and are useful for navigation on the ML when performing high-throughput characterization. However, this can only be used for low temperature deposition processes ($<200^\circ\text{C}$).

For a comprehensive materials understanding, as many properties as possible should be measured. For the selection of the applied high-throughput methods, the quality of the resulting data should be rated higher than the throughput of the measurement, i.e., high-quality-high-throughput measurements are necessary. An overview is shown in Fig. 2. Of basic importance are compositional and structural data. It is not sufficient to work with nominal compositions in MLs. High-throughput EDX (energy-dispersive x-ray analysis) determines the chemical compositions comprised in the ML sufficiently fast (e.g., 342 MAs in 8 h) with sufficient accuracy (typically ± 1 at.%). For high-throughput measurements of elements lighter than N, automated methods are available such as wavelength-dispersive x-ray analysis (WDX), x-ray photoelectron spectroscopy (XPS), Rutherford backscatter spectroscopy (RBS), and nuclear reaction analysis (NRA). These methods are more challenging with respect to data analysis but can also reveal further compositional information (e.g., oxidation state, compositional depth profiles). By combining e.g., EDX with RBS and NRA, all elements in MLs of thin film Li-ion battery materials can be determined.¹⁹ High-throughput X-ray diffraction (XRD) is used to determine the phases being present on a ML. XRD systems with microfocus source and area detector allow fast measurements with very good spatial resolution (spot size < 1 mm). For

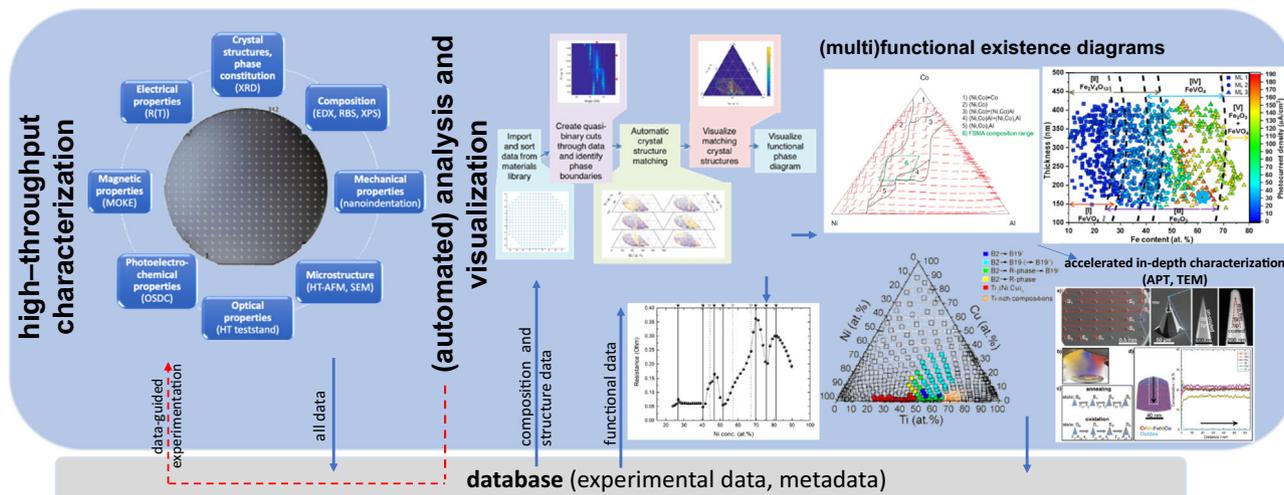


Fig. 2 High-throughput characterization of thin-film MLs using a comprehensive suite of automated characterization tools. All experimental and analyzed data, including metadata should be transferred to a database. Data is analyzed preferably by automated data-analysis and visualization tools. In order to achieve (multi)functional existence diagrams, data from the database is visualized. The most interesting MAs in MLs can be analyzed in depth with accelerated atomic-scale characterization tools. Parts of the figure were reprinted/adapted from²¹ (copyright ACS Publications, 2014),⁴⁹ (copyright ACS Publications, 2017),⁵⁸ (copyright ACS Publications, 2017),⁵⁹ (copyright Elsevier, 2012),⁶⁰ (copyright ACS Publications 2018) and³⁷ (copyright RCS, 2018)

even higher throughput and resolution synchrotron measurements can be used.²⁰ Further high-throughput characterization methods correspond to the specific material development goals, i.e., functional properties. Adequate screening parameters (e.g., electrical resistivity, Seebeck coefficient, Young's modulus, hardness, etc.) need to be defined or descriptors of the property of interest, if it cannot be directly measured. In many cases measuring one screening parameter is not enough, and a different ML design might be necessary to enable measuring all necessary parameters. High-throughput electrical resistivity measurements serve well for delivering descriptor values for properties related to it.²¹ It can be regarded as a discovery tool: with room-temperature measurements on MLs indications for different phase zones and their boundaries can be obtained. Non-linear resistance changes in temperature-dependent resistance measurements $R(T)$ yield signatures for phase transformations.²² These can be then analyzed for the few most interesting samples by XRD(T). For the determination of magnetic properties in MLs, the magneto-optical Kerr effect is suitable for high-throughput measurements of magnetic hysteresis. $R(T)$ measurements combined with high-throughput MOKE measurements allow e.g., the identification of ferromagnetic shape memory alloys.²³ Basic optical properties can be determined by photography, further properties, e.g., in automated test-stands for optical transmission (transparent substrate for MLs necessary). Photoelectrochemical properties can be measured with an automated scanning droplet cell,²⁴ however then the ML needs to be deposited on a homogeneous conductive layer such as Pt. For many materials it is also important to determine their microstructure in high-throughput, e.g., using SEM and AFM. Mechanical properties can be measured with nanoindentation,²⁵ however the microstructural variations over a ML can influence the measurement results and need to be taken into account.

EXEMPLARY RESULTS

Examples for shape memory alloys (from thin films to bulk)

The following examples illustrate discoveries made using thin-film MLs, in some cases supported by computational materials science. Findings from thin-film MLs explorations can also be transferred to

bulk (e.g., by arc-melting of hit compositions). Efficient screening in thin-film MLs reveals intrinsic properties of the materials systems, whereas extrinsic properties will differ in thin film and bulk form of the same material.

An example for the verification of a theoretical prediction (geometric nonlinear theory of martensite), using CMS and upscaling to bulk was the identification of quaternary shape memory alloys (SMAs) with near-zero hysteresis.²² It was predicted in²⁶ that for certain compositions in Ti-Ni-Cu-Pd the thermal hysteresis of the martensitic phase transformation should be minimal, which is important for achieving fatigue-resistant SMA and enables better actuator control. Due to the noble metal content of Ti-Ni-Cu-Pd, this study could have not been done in bulk experiments due to prohibitive cost. However, a thin-film composition spread ML around the predicted compositions delivers all necessary materials to verify the prediction in one experiment. By high-throughput characterization of the lattice parameters (XRD) and the thermal shape memory hysteresis (using $R(T)$ measurements) in dependence of composition in the MLs, the predicted SMAs with near-zero thermal hysteresis were identified.²² These results were successfully transferred to bulk scale, where near-zero thermal hysteresis was also observed. Improved functional stability was found for the bulk alloy $\text{Ti}_{50.2}\text{Ni}_{34.4}\text{Cu}_{12.3}\text{Pd}_{3.1}$. As another example,²⁷ R-phase formation in $\text{Ti}_{39}\text{Ni}_{45}\text{Cu}_{16}$ SMA thin films and bulk alloys was discovered using a thin-film composition spread ML. The discovered SMA exhibits a single B2 to R-phase transformation $> 25^\circ\text{C}$ with a thermal hysteresis width $< 1\text{K}$. The thin film results correctly predicted the bulk phase transformation behavior upon cooling from the high temperature phase. For both thin film and bulk, a two-step B2 to R-phase to B19 transformation was found. However, in this case the material is only applicable in thin-film form as the ternary bulk alloy is brittle. Multiferoic materials like ferromagnetic SMAs (FSMAs) are of interest for miniaturization and functional integration. Next to the best studied system Ni_2MnGa other FSMAs, e.g., based on $\text{Fe}_{70}\text{Pd}_{30}$, exist. In this system it was found by high-throughput experimentation and ab initio calculations that the addition of Cu improves the alloy properties.²⁸

Examples for identification of new phases and assessment of phase/existence diagrams

The experimental identification of all (meta)stable phases in at least ternary systems and their mapping in existence diagrams would be beneficial, not only for these systems themselves, but also as basis for then going to increased chemical complexity. The discovery of a ternary μ -phase in Co-Ti-W²⁹ based on high-throughput XRD and resistivity data and verified by TEM on bulk samples as well as DFT calculations, is an example for improving CALPHAD calculations by incorporation of these new findings. Then, improved CALPHAD calculations can be used for e.g., superalloy investigations. New non-cubic ferromagnetic phases for permanent magnet applications can be searched for in Fe-Co-X thin-film MLs by high-throughput MOKE: in a Fe-Co-Nb composition spread, a new magnetic compound (Fe,Co)₃Nb with C36 structure, embedded in a FeCo-based matrix was identified.³⁰

Increased chemical complexity

Increased chemical complexity can lead to improved functional properties e.g.³¹ Compositional effects on the light-induced oxygen evolution in co-deposited Bi(V-Mo-X)O₄ MLs (X = Ta, W, Nb) were unraveled using an optical scanning droplet cell. Bi(V-Mo-W)O₄ and Bi(V-Mo-Nb)O₄ MLs show ten times higher photocurrents compared to a BiVO₄ reference ML. For Cr-Mn-Fe-Co-Ni nanoparticles it was discovered⁸ that five elements are necessary to achieve the best electrocatalytic properties, whereas for all quaternary or lower subsets of this system, the electrocatalytic properties were worse. It needs to be investigated by high-throughput methods if equiatomic quinary nanoparticles show the best properties or if other compositions (within the Cr-Mn-Fe-Co-Ni or other systems) are better. This can only be achieved by combinatorial synthesis of nanoparticle libraries and suitable high-throughput characterization tools. A paradigm change is envisaged³² and these experiments should be supported by theory, a first example being.³³

Microstructure libraries

The microstructure of thin films is determining their extrinsic properties. Therefore, not only synthesizability and stability of new materials should be addressed by combinatorial materials science, but also the microstructural diversity, which is closely linked to processing. This means for each composition of interest, the search space is extended by the different microstructures a material can be created in. To address this challenge “combinatorial processing” can be applied to create “processing libraries”. These contain e.g., composition spreads around predicted compositions, synthesized simultaneously at different temperatures. This can be realized by gradient and step heater methods e.g.^{34,35} These MLs comprise films of the same composition in different microstructures: amorphous, nano- or microcrystalline, single- or multi-phase constitution, etc.

FROM DISCOVERY TO UNDERSTANDING

Discovered materials need to be studied in depth to unravel their crystal structure and the origin of their properties. Also here, accelerated experimentation is desired. Next to synchrotron measurements, advanced (in situ) TEM methods are promising, e.g., automated diffraction tomography combined with precession electron diffraction.³⁶ Furthermore, the temperature and environment-dependent phase evolution of complex materials can be studied in an accelerated way using combinatorial processing platforms,³⁷ see Fig. 2. Combinatorial processing platforms are created by co-deposition of multinary thin films on nanoscale Si-tip arrays. Thus, many identical nanoscale “reactor volumes” are formed, allowing for fast diffusion and reaction and

immediate observation of the product phases on the atomic scale by atom-probe tomography (APT) and TEM. This allows for an accelerated mapping of the phase space of multinary systems with regards to stability against decomposition into phases (e.g., from a solid solution to intermetallics) and as well against reactions with the environment, e.g., oxidation of alloys.³⁸

REQUIREMENTS FROM AND INTERACTIONS WITH COMPUTATIONAL METHODS AND MATERIALS INFORMATICS

Based on the above examples of interactions between CMS and computational materials science and further examples from literature e.g.,³⁹ it is discussed how computational methods can further accelerate experimental discovery of new phases as well as what can be provided by high-throughput experimentation in the laboratory to improve computational materials science. Cooperation between experimental and computational groups should be intensive, with a very good knowledge of each other's methods.

Whereas databases of (high-throughput) calculated materials data, e.g.,^{40,41} exist and are accessible, databases of high-throughput experimental data are rare.⁴² For a close cooperation between computational and experimental materials science groups it would be useful to have a joint database, or at least databases which can be used to fulfill both areas' needs: e.g., download large and consistent experimental datasets as input data for high-throughput computations or use calculated data as input for down-selection of possible materials system for a combinatorial exploration. Materials informatics⁴³ has a huge potential in closing the gap between experimental and computational groups. Digitalized research data management is a challenge for high-throughput experimentation groups, e.g., to store all data (and meta-data) which is produced during fabrication and, most of all, high-throughput characterization of MLs. A database of experimental datasets, acquired in a consistent manner by high-throughput experimentation, would be advantageous: as MLs are fabricated in single experiments, all materials in the ML are very well comparable and due to the continuous composition gradients, the quality and consistency of the data is assured. ML can also include reference materials (e.g., the three binaries of a ternary system, Fig. 1) and thus the quality of the datasets can be very high. However, it is important that *all* results are included: i.e., also “negative” results, e.g., a synthesis experiment which was not successful, or results of high-throughput characterization which did not show the desired properties. This is necessary, e.g., if the data is intended to be used for machine learning. Data management should include the FAIR guiding principles: data needs to be findable, accessible, interoperable, and reusable, both for humans and computers.⁴⁴ However, there are still issues with data ontology, data quality and thus also with usability and transferability. For a future efficient exchange of data between experimental and computational groups to enable data-driven materials discovery, meta-data and data formats need to be agreed upon, e.g.⁴⁵ Furthermore, input and output from databases (Figs 1, 2) for improving the combinatorial workflow should be at least semi-automated, as scientists usually don't like to put data in databases “by hand”. The quality of data to be put in joint database could also be improved by fabricating and measuring nominally identical MLs of newly identified systems in different labs and by different methods to assure their properties: In⁴⁶ the authors report on the reproducibility of results from high-throughput measurements for the MLs of the same system at different laboratories, and on the associated sample and data exchange standards.

Finally, for exploring multinary systems, data produced by CMS is so overwhelming, that it cannot be handled by scientists alone, and the development and use of new computational tools is a necessity. Visualization of compositions and properties in complex multinary materials systems is difficult but necessary for data

analysis, e.g.^{47,48} New software has to be developed which will lead to the establishment of functional phase or existence diagrams (including metastable phases) for multinary alloys for a future design of materials.

For high-throughput XRD several software developments already accelerated phase analysis in MLs.^{49–54} Instead of having to analyze hundreds of diffractogram, e.g., by non-negative matrix factorization a set of e.g., tens of similar patterns can be obtained which are then the start for an in-depth analysis. Concepts for using machine learning for data-guided experimentation were introduced, e.g.⁵⁴ These need to be adapted and further developed for data-driven high-throughput experimentation: Usually all MAs in a ML, e.g., 342 MAs on a 100 mm diameter substrate, are measured sequentially. For a set of several parameters to be measured, this sums quickly up to a huge amount of MAs. One idea is to use first the fastest high-throughput measurements, which identify rapidly areas of interest in a ML, where then slower methods are applied.²¹ Another idea is to reduce the number of MAs to be measured, by using design of experiment concepts combined with machine learning. However, it might be argued if the number of necessary MAs could be reduced while still achieving qualitatively the same, or nearly the same result. Instead of measuring MAs inside a region of interest, it would be more efficient to measure only boundaries between regions. A possibility is to start measuring arbitrary MAs and then measure MAs which are most distant from the first data. If this works, more of the search space can be explored in the same time. Instead of arbitrary starting MAs also MAs based on prior knowledge or intuition could be selected. Finally, predictions of synthesizability^{55,56} would be highly useful for a further reduction of the number of necessary experiments.

CONCLUSIONS AND OUTLOOK

Whereas combinatorial deposition methods for thin-film MLs are well-established, the further automatization of high-quality high-throughput characterization methods need to be continued to accelerate materials discovery. Furthermore, schemes where MLs are not only characterized for one property, but rather comprehensively for as many as possible functional properties need to be developed. Setting up a library of MLs would be useful, e.g., for the use of well-characterized MLs for (high-throughput) measurements and further analysis by external groups for their specific purposes. The implementation of research data management is of highest importance to enable cooperation between experimental and computational groups.

As an actionable plan for the future: (I) research groups from computational and experimental high-throughput research should get in much closer cooperation. High-throughput experiments should be planned and be performed from the beginning as joint efforts with close links across all stages of materials development. (II) For enabling discovery of complex materials, the challenges of materials data science, e.g., data-guided experimentation in immense search spaces, visualization and (automated) analysis of high-dimensional data, need to be addressed. For this a new generation of materials data scientists have to be educated. Currently researchers with the required skills are rare. (III) Results of materials discovery and design should be published such that data of the new materials and how they were developed are accessible. As a new form of visualization interactive multifunctional existence diagrams are proposed. These comprise all data of a system with respect to composition, crystal structure and microstructures of metastable and stable phases, all processing data, and the functional properties. (IV) Further aspects relate to the automation of experiments towards autonomous experimentation by integrating artificial intelligence approaches for hypothesis generation and their high-throughput testing.⁵⁷

Further advances are expected by using augmented and virtual reality tools for data analysis.

ACKNOWLEDGEMENTS

The author acknowledges all current and former members of the research group, all cooperation partners and funding sources, especially DFG.

ADDITIONAL INFORMATION

Competing interests: The author declares no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Zunger, A. Inverse design in search of materials with target functionalities. *Nat. Rev. Chem.* **2**, 0121 (2018).
2. Alberi, K. et al. The 2019 materials by design roadmap. *J. Phys. D Appl. Phys.* **52**, 013001 (2019).
3. Kauffman, G. B. & Mayo, I. The story of Nitinol: The serendipitous discovery of the memory metal and its applications. *Chem. Educ.* **2**, 1–21 (1997).
4. Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. Metallic phase with long-range orientational order and no translational symmetry. *Phys. Rev. Lett.* **53**, 1951–1953 (1984).
5. Bednorz, J. G. & Müller, K. A. Possible high T_c superconductivity in the Ba-La-Cu-O system. *Z. Phys. B Condens. Matter* **64**, 189–193 (1986).
6. Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. Superconductivity at 39 K in magnesium diboride. *Nature* **410**, 63–64 (2001).
7. Chiang, C. K. et al. Electrical conductivity in doped polyacetylene. *Phys. Rev. Lett.* **39**, 1098–1101 (1977).
8. Löffler, T. et al. Discovery of a multinary noble metal free oxygen reduction catalyst. *Adv. Energy Mater.* **8**, 1802269 (2018).
9. Walsh, J. P. S., Clarke, S. M., Meng, Y., Jacobsen, S. D. & Freedman, D. E. Discovery of FeBi₂. *ACS Cent. Sci.* **2**, 867–871 (2016).
10. Zakutayev, A. Design of nitride semiconductors for solar energy conversion. *J. Mater. Chem. A* **4**, 6742–6754 (2016).
11. Singh, A. K., Montoya, J. H., Gregoire, J. M. & Persson, K. A. Robust and synthesizable photocatalysts for CO₂ reduction: a data-driven materials discovery. *Nat. Commun.* **10**, 443 (2019).
12. Hinuma, Y. et al. Discovery of earth-abundant nitride semiconductors by computational screening and high-pressure synthesis. *Nat. Commun.* **7**, 11962 (2016).
13. Xiang, X.-D. et al. A combinatorial approach to materials discovery. *Science* **268**, 1738–1740 (1995).
14. Koinuma, H. & Takeuchi, I. Combinatorial solid-state chemistry of inorganic materials. *Nat. Mater.* **3**, 429–438 (2004).
15. Green, M. L., Takeuchi, I. & Hatrick-Simpers, J. R. Applications of high throughput (combinatorial) methodologies to electronic, magnetic, optical, and energy-related materials. *J. Appl. Phys.* **113**, 231101 (2013).
16. Zhao, J.-C. Combinatorial approaches as effective tools in the study of phase diagrams and composition–structure–property relationships. *Prog. Mater. Sci.* **51**, 557–631 (2006).
17. Ludwig, A., Zarnetta, R., Hamann, S., Savan, A. & Thienhaus, S. Development of multifunctional thin films using high-throughput experimentation methods. *Inter. J. Mater. Res.* **99**, 1144–1149 (2008).
18. König, D., Richter, K., Siegel, A., Mudring, A.-V. & Ludwig, A. High-throughput fabrication of Au-Cu nanoparticle libraries by combinatorial sputtering in ionic liquids. *Adv. Funct. Mat.* **24**, 2049–2056 (2014).
19. Borhani-Haghighi, S. et al. High-throughput compositional and structural evaluation of a Li_x(Ni_xMn_{1-x}Co₂)O₂ thin film battery materials library. *ACS Comb. Sci.* **15**, 401–409 (2013).
20. Gregoire, J. M. et al. High-throughput synchrotron X-ray diffraction for combinatorial phase mapping. *J. Synchrotron Radiat.* **21**, 1262–1268 (2014).
21. Thienhaus, S., Naujoks, D., Pfetzinger-Micklich, J., König, D. & Ludwig, A. Rapid identification of areas of interest in thin film materials libraries by combining electrical, optical, X-ray diffraction, and mechanical high-throughput measurements: A case study for the system Ni–Al. *ACS Comb. Sci.* **16**, 686–694 (2014).
22. Zarnetta, R. et al. Identification of quaternary shape memory alloys with near-zero thermal hysteresis and unprecedented functional stability. *Adv. Funct. Mater.* **20**, 1917–1923 (2010).
23. Takeuchi, I. et al. Identification of novel compositions of ferromagnetic shape-memory alloys using composition spreads. *Nat. Mater.* **2**, 180–184 (2003).

24. Sliozberg, K. et al. High-throughput screening of thin-film semiconductor materials libraries I: System development and case study for Ti-W-O. *ChemSusChem* **8**, 1270–1278 (2015).
25. Zarnetta, R., Kneip, S., Somsen, Ch & Ludwig, A. High-throughput characterization of mechanical properties of Ti-Ni-Cu shape memory thin films at elevated temperatures. *Mat. Sci. Eng. A* **528**, 6552–6557 (2011).
26. James, R. D., & Zhang, Z. A way to search for multiferoic materials with “unlikely” combinations of physical properties. (eds A. Planes., L. Mañosa. & A. Saxena) *Magnetism and structure in functional materials*. 159–175 (Springer, Berlin, Heidelberg, 2005).
27. Zarnetta, R. et al. R-phase formation in $\text{Ti}_{39}\text{Ni}_{45}\text{Cu}_{16}$ shape memory thin films and bulk alloys—discovered by combinatorial methods. *Acta Mater* **57**, 4169–4177 (2009).
28. Hamann, S. et al. The ferromagnetic shape memory system Fe-Pd-Cu. *Acta Mater* **58**, 5949–5961 (2010).
29. Naujoks, D. et al. Identification of a ternary μ -phase in the Co-Ti-W system—An advanced correlative thin film-bulk combinatorial materials investigation. *Acta Mater* **138**, 100–110 (2017).
30. Alexandrakis, V. et al. Combinatorial development of Fe-Co-Nb thin film magnetic nanocomposites. *ACS Comb. Sci.* **17**, 698–703 (2015).
31. Gutkowski, R. et al. Unraveling compositional effects on the light-induced oxygen evolution in $\text{Bi}(\text{V-Mo-X})\text{O}_4$ material libraries. *Energ. Environ. Sci.* **10**, 1213–1221 (2017).
32. Löffler, T. et al. Towards a paradigm shift in electrocatalysis using complex solid solution nanoparticles. *ACS Energy Lett.* **4**, 1206–1214 (2019).
33. Batchelor, T. A. A. et al. High-entropy alloys as a discovery platform for electrocatalysis. *Joule* **3**, 834–845 (2019).
34. Caskey, C. M., Richards, R. M., Ginley, D. S. & Zakutayev, A. Thin film synthesis and properties of copper nitride, a metastable semiconductor. *Mater Horiz.* **1**, 424–430 (2014).
35. Stein, H. et al. A structure zone diagram obtained by simultaneous deposition on a novel step heater: a case study for Cu_2O thin films. *Phys. Status Solidi A* **212**, 2798–2804 (2015).
36. Mugnaioli, E., Gorelik, T. & Kolb, U. “Ab initio” structure solution from electron diffraction data obtained by a combination of automated diffraction tomography and precession technique. *Ultramicroscopy* **109**, 758–765 (2009).
37. Li, Y. J., Savan, A., Kostka, A., Stein, H. S. & Ludwig, A. Accelerated atomic-scale exploration of phase evolution in compositionally complex materials. *Mater Horiz.* **5**, 86–92 (2018).
38. Li, Y., Kostka, A., Savan, A. & Ludwig, A. Atomic-scale investigation of fast oxidation kinetics of nanocrystalline CrMnFeNiCo thin films. *J. Alloy. Compd.* **766**, 1080–1085 (2018).
39. Zakutayev, A. et al. Theoretical prediction and experimental realization of new stable inorganic materials using the inverse design approach. *J. Am. Chem. Soc.* **135**, 10048–10054 (2013).
40. Jain, A. et al. The materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
41. Curtarolo, S. et al. AFLOWLIB.ORG: a distributed materials properties repository from high-throughput ab initio calculations. *Comp. Mater Sci.* **58**, 227–235 (2012).
42. Zakutayev, A. et al. An open experimental database for exploring inorganic materials. *Sci. Data* **5**, 180053 (2018).
43. Rajan, K. Materials informatics. *Mater Today* **8**, 38–45 (2005).
44. Wilkinson, M. D. et al. The FAIR Guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
45. Ghiringhelli, L. M. et al. Towards efficient data exchange and sharing for big-data driven materials science: metadata and data formats. *npj Comput. Mater* **3**, 46 (2017).
46. Hatrick-Simpers, J. R. et al. An inter-laboratory study of Zn-Sn-Ti-O thin films using high-throughput experimental methods. *ACS Comb. Sci.* **21**, 350–361 (2019).
47. Suh, C. et al. Visualization of high-dimensional combinatorial catalysis data. *J. Comb. Chem.* **11**, 385–392 (2009).
48. Rickman, J. M. Data analytics and parallel-coordinate materials property charts. *npj Comput Mater* **4**, 5 (2018).
49. Stein, H. S., Jiao, S. & Ludwig, A. Expediting combinatorial dataset analysis by combining human and algorithmic analysis. *ACS Comb. Sci.* **19**, 1–8 (2017).
50. Long, C. J., Hatrick-Simpers, J., Murakami, M., Srivastava, R. C. & Takeuchi, I. Rapid structural mapping of ternary metallic alloy systems using the combinatorial approach and cluster analysis. *Rev. Sci. Instrum.* **78**, 072217 (2007).
51. Suram, S. K. et al. Automated phase mapping with AgileFD and its application to light absorber discovery in the V-Mn-Nb oxide system. *ACS Comb. Sci.* **19**, 37–46 (2017).
52. Iwasaki, Y., Kusne, G. & Takeuchi, I. Comparison of dissimilarity measures for cluster analysis of X-ray diffraction data from combinatorial libraries. *npj Comput. Mater* **3**, 4 (2017).
53. Stanev, V. et al. Unsupervised phase mapping of X-ray diffraction data by non-negative matrix factorization integrated with custom clustering. *npj Comput. Mater* **4**, 43 (2018).
54. Kusne, A. G. et al. On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Sci. Rep.* **4**, 6367 (2014).
55. Aykol, M., Dwaraknath, S. S., Sun, W. & Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Sci. Adv.* **4**, eaaq0148 (2018).
56. Kim, E. et al. Materials synthesis insights from scientific literature via text extraction and machine learning. *Chem. Mater* **29**, 9436–9444 (2017).
57. Tabor, D. P. et al. Accelerating the discovery of materials for clean energy in the era of smart automation. *Nat. Rev. Mater* **3**, 5–20 (2018).
58. Decker, P., Naujoks, D., Langenkaemper, D., Somsen, C. & Ludwig, A. High-throughput structural and functional characterization of the thin film materials system Ni-Co-Al. *ACS Comb. Sci.* **19**, 618–624 (2017).
59. Zarnetta, R., Buenconsejo, P. J. S., Savan, A., Thienhaus, S. & Ludwig, A. High-throughput study of martensitic transformations in the complete Ti-Ni-Cu system. *Intermetallics*. **26**, 98–109 (2012).
60. Kumari, S. et al. Combinatorial synthesis and high-throughput characterization of Fe-V-O thin film materials libraries for solar water splitting. *ACS Comb. Sci.* **20**, 544–553 (2018).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019