

The Materials Super Highway: Integrating High-Throughput Experimentation into Mapping the Catalysis Materials Genome

Jason Hattrick-Simpers · Cun Wen ·
Jochen Lauterbach

Received: 12 October 2014 / Accepted: 13 November 2014 / Published online: 27 November 2014
© Springer Science+Business Media New York 2014

Abstract The materials genome initiative (MGI) aims to accelerate the process of materials discovery and reduce the time to commercialization of advanced materials. Thus far, the MGI has resulted in significant progress in computational simulation, modeling, and predictions of catalytic materials. However, prodigious amounts of experimental data are necessary to inform and validate these computational models. High-throughput (HT) methodologies, in which hundreds of materials are rapidly synthesized, processed, and characterized for their figure of merit, represent the key experimental enabler of the theoretical aspects of the MGI. HT methodologies have been used since the early 1980s for identifying novel catalyst formulations and optimizing existing catalysts. Many sophisticated screening and data mining techniques have been developed and since the mid-1990s, this approach has become a widely accepted industrial practice. This article will provide a short history of major developments in HT and will discuss screening approaches combining rapid, qualitative primary screens via thin-film techniques with a series of quantitative screens using plug flow reactors. An illustrative example will be provided of one such study in which novel fuel-flexible sulfur tolerant cracking catalysts were developed. We will then illustrate a path forward that leverages existing HT expertise to validate, provide empirical data to and help guide future theoretical studies.

Keywords Combinatorial chemistry · Methodology and phenomena · Nanotechnology · High throughput

1 Introduction

Two major goals of the materials genome initiative (MGI) are to cut the time between discovery and commercialization of advanced materials in half, from 10–20 to 5–10 years, and to do it at lower cost [1]. This is accomplished by creating an integrated design continuum that leverages computational, experimental and open-data informatics tools, expediting the discovery of new science and breaking down obstacles to commercialization (Fig. 1). The ultimate goal of the MGI is to educate a next generation workforce capable of utilizing these tools sets to promote national security, clean energy, and human welfare. Thus far, the MGI has made significant progress in computational simulation, modeling, and predictions of materials properties [2, 3]. In the fields of catalysis, examples of the successful use of theory include the design of alloy catalysts based on DFT calculations for hydrocarbon steam reforming and acetylene hydrogenation [4–6]. However, though theory often can inform experiments, a feedback route with experimentation is necessary for validation and new hypothesis generation. As an example, a recently reported Ni–Ga based catalyst is active for CO₂ reduction to methanol, however it resides in a complicated portion of the phase diagram where competition between the δ and γ phases exists. [7]. It is therefore likely that on-stream the surface of the catalyst is quite different from that of the theoretically predicted active phase, but without experimental data characterizing these changes during operation new innovation is stymied. In addition, it is well known that theoretical studies still struggle to make accurate predictions when working in complicated gaseous environments or accounting for the effect of point defects, ion transport and interfaces in determining the reaction rates and selectivities [8–10]. Despite these well-known

J. Hattrick-Simpers · C. Wen · J. Lauterbach (✉)
Department of Chemical Engineering, Smartstate Center
for Strategic Approaches to the Generation of Electricity,
University of South Carolina, Columbia, SC 29209, USA
e-mail: lauteraj@cec.sc.edu

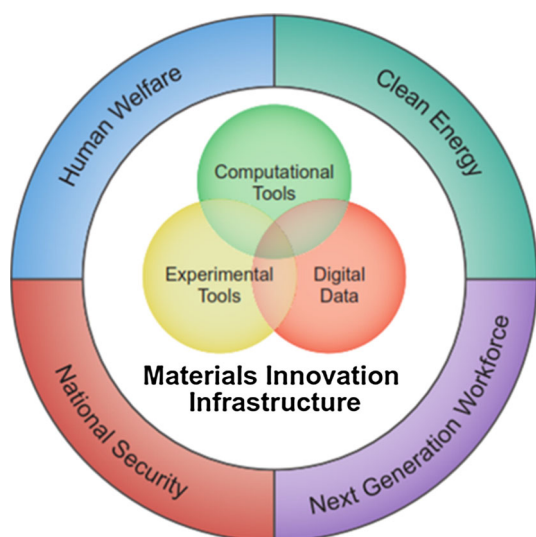


Fig. 1 MGI overview illustrating the overlap between the three core fields and the breadth of its impact on technology and the workforce [1]

caveats to theoretical studies, efforts in the MGI have largely focused on theory, disregarding the important role of experimentation.

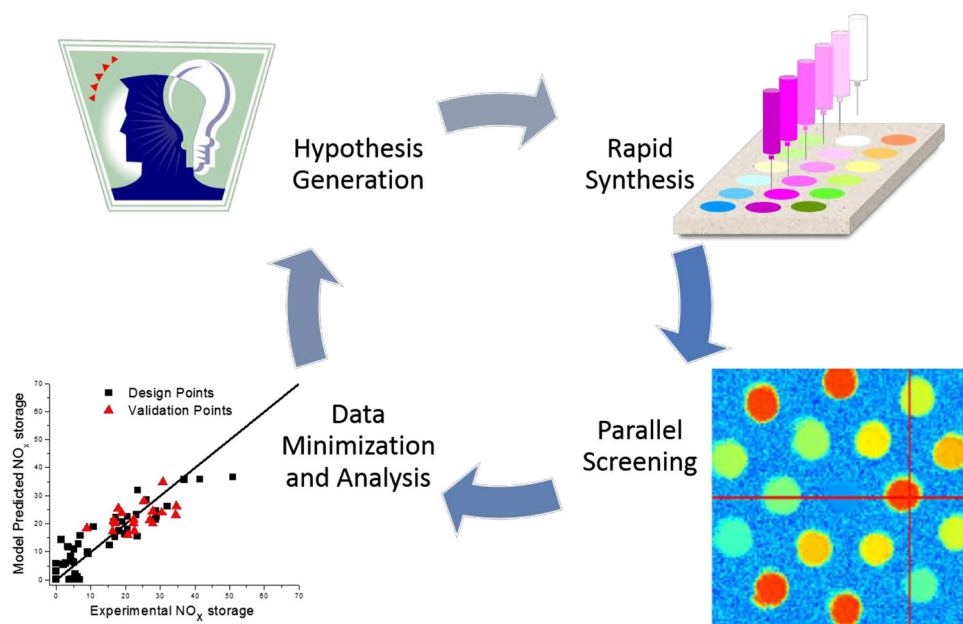
As theory becomes more routinely applied to catalysis problems, prodigious amounts of experimental data are required to inform and validate these computational models, thus powering the MGI computational “engine.” Modeling studies rely on accurate depictions of materials properties including crystal structure, composition, and surface coverage/arrangements under catalytic reaction conditions [11]. The latter property is of paramount importance, as it is often the case that the surface of a catalyst under reaction conditions (and thus the very nature of its active sites) does not resemble the equilibrium surface at ambient conditions, let alone at 0 K [12]. Moreover, the role of interfaces, in particular those between the catalyst and the support, have only recently begun to be modeled theoretically [13, 14]. When one considers, however, that these interfaces are likely as dynamic as the catalyst surface, it becomes clear that without experimental validation of the role these important parameters play in catalysis and how they change during reactions a speed bump in the “Materials Super Highway” is imminent.

High-throughput experimental (HTE) methodologies, which allow accelerated synthesis and testing of materials for optimized performance, are uniquely suited to rapidly generate high quality experimental data, and hence represent the key enabling technologies to bring the computational materials design efforts of the MGI to fruition. The field of catalysis was an early adopter of high-throughput screening technologies; the original Haber–Bosch catalyst was discovered in 1913 by systematically studying over

2,500 potential catalysts with 6,500 experiments before settling on iron as the catalyst [15]. Today, the Haber–Bosch process is responsible for synthesizing the world’s ammonia supply, mainly used for fertilizer production, which makes it possible for our globe to sustain more than 7 billion people [16]. The advent of powerful personal computers capable of automated data acquisition and analysis in the late 1980s, coupled with the marked success of pharmaceutical companies in novel drug discovery via parallelized synthesis and screening, led to a resurgence of interest in the in the 1990s. There are several outstanding recent reviews on the topic of high-throughput materials science; in particular for the field of catalysis [17]. The interested reader is particularly encouraged to consult recent reviews by Maier et al. and Green et al. [18, 19]. A schematic of the overall high-throughput approach is shown in Fig. 2, where hypothesis generation is followed by the rapid synthesis of hundreds of different catalyst formulations. The catalysts are then screened using either parallel or sequential screening techniques, and specialized data minimization/analysis software are then used to extract information from the dense data sets.

Early adopters of high-throughput technologies in catalysis included start-up companies, such as Symyx, Avantium and HTE, large chemical companies, such as DOW, BASF, and ExxonMobil, as well as a (small) number of academic labs. While these first studies were important demonstrations of the potential of high-throughput in catalysis, frequently they were plagued by irreproducible results and overly aggressive patenting of large swaths of the periodic table. Early studies often prioritized throughput in terms of synthesis and rapid reaction screening over ensuring that catalysts were created and tested in a manner consistent with the actual large-scale application. Smotkin et al. reported, for instance, that the blanket use of NaBH_4 as a reductant during production of fuel cell catalysts arrays rendered such an approach susceptible to local maxima that are largely a function of synthesis method, and are not reflective of the ultimate material figure of merit [20]. Early studies often characterized samples deposited on planar samples using scanning probe mass-spectrometry, IR thermography, or optical imaging techniques, such as fluorescence [21–23]. Many of these techniques inevitably suffer from sample cross-talk since there is no barrier preventing reactant and product spill-over between catalyst sites. There have been recent studies that sought to minimize cross-talk in planar samples by introducing physical barriers such as capillaries between individual catalysts [24]. This approach also removes a means of evaluating the catalyst-support interaction and the role of nanoparticle properties in determining activity and selectivity [13]. IR thermography has the additional caveat of not being capable of product speciation, but rather only identifying “active” catalysts by monitoring their heat signatures.

Fig. 2 Schematic representation of the high-throughput experimental cyclical workflow. Work is initiated through the generation of a set of hypotheses, rapid synthesis techniques are then used to design an array of catalytic materials. Parallel screening techniques are used to screen the catalysts for their properties, which are used to quantify figure of merit using a host of data minimization and analysis tools



Modern HT catalysis studies are often multi-faceted with multiple stages that first rapidly identify potential hits using qualitative measurements and subsequently hone in on materials of interest in increasingly greater detail under industrially important conditions [25–27]. Each study begins with an analytical search of the relevant technical literature to identify important parameters for a particular catalytic reaction. An example of such a heuristic would be the importance of the ratio of zeolite pore size to feed stock and product molecular sizes in biofuel platform chemical upgrading [28]. Once a suitable heuristic is identified, a rapid primary screen is then undertaken that uses either robotic pipetting systems or thin film depositions to create arrays of hundreds to thousands of samples on a single chip. These arrays are then screened via a qualitative primary screen, e.g., using optical techniques or thermography to identify initial leads, such as a coke inhibiting additive [29]. This is quite similar to the approaches used above with the same caveats, however, the goal is to quickly identify potential winning/losing regions and then move the most promising regions of the parameter space directly to bench-scale reactor studies. At the second stage of the screen, milligram quantities of the catalyst are produced via a host of traditional methods, which can be parallelized and automated. Such methods include wet impregnation, reverse micelles, parallel hydrothermal synthesis, etc. including impregnation onto suitable structured supports. Quantitative screening is generally undertaken via parallel reactor studies, where cross-talk between catalysts is prevented and effluent streams are analyzed separately, often via serial GC/MS or imaging FTIR [30–33]. Final screening of select catalysts can be done in

parallel pilot scale reactors, generally single-bed reactors although commercial multi-bed reactors are available, with catalysts prepared in kilogram quantities and run under true industrial conditions.

In addition to the catalytic performance of the synthesized materials, a great deal of effort is devoted to obtaining fundamental understanding of the structure and surface composition of promising new catalysts. This information is just as vital as catalytic performance for guiding theoretical MGI studies as the baseline necessary for DFT studies is the bulk crystal structure and possibly the surface configuration. To this end, a number of truly HT structural characterization techniques are routinely applied including scanning synchrotron X-ray diffraction (bulk structure) [34–36], X-ray absorption spectroscopy (local coordination) [37], and X-ray photoelectron spectroscopy (surface composition and element coordination) [38]. A variety of compositional tools, such as atomic adsorption spectroscopy, inductively couple plasma analysis, and SEM/EDS are readily parallelized and are utilized regularly in HT studies [39, 40]. Conversely, measurements that provide detailed information about the microstructure of nanoparticles such as high resolution TEM are inherently slow and time consuming, and are used only for ultimate material validation [41]. This despite the fact that a truly integrated theoretical–experimental workflow may benefit from the early insights provided to each other as hits and misses are identified and related between practitioners.

Integration of high-throughput experimental and fundamental modeling efforts for catalyst design have been discussed for more than 10 years [33, 42]. Figure 3

illustrates an example workflow in catalysis where experiment and theory are used to drive both novel materials discovery and create fundamental understanding of catalyst performance. However, despite the progress of the HT catalysis field, great opportunities exist to more fully integrate its efforts into the worldwide MGI efforts. In the next section, we will provide a case study of a successful HT experimental project in academia, where a new class of catalysts were hypothesized, discovered, and optimized in less than 10 months. We will illustrate opportunities where theory and data science could have provided clear direction and further accelerated catalyst development, while extracting new knowledge that would have benefited work on dissimilar catalytic processes with similar materials. We will then generalize these opportunities and provide a view of what the next step down the “Materials Super Highway” could look like for the field of catalysis. Finally, we will show how these opportunities can be consolidated with a generalized High-throughput Experimental Center where experts in the art of HT catalysis will join with those applying HT to electronic materials, structural materials, and energy materials towards creating experimental and data mining tools benefiting all groups mutually.

2 Case Study for JP-8

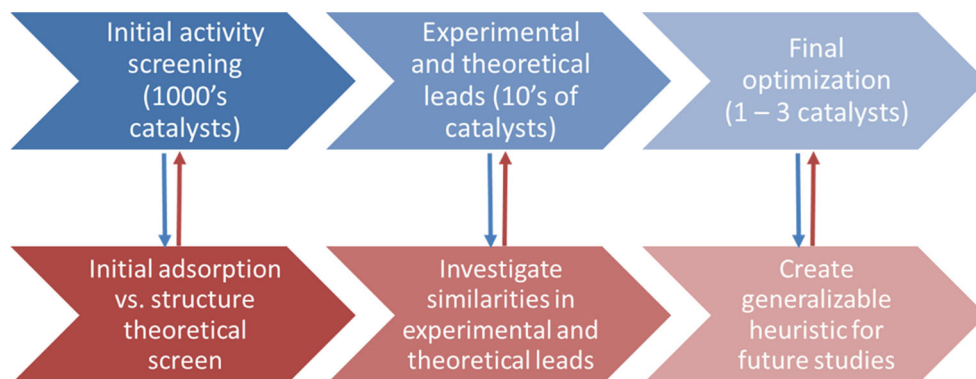
We recently applied HT techniques to develop catalysts for on-site production of LPG from diesel and JP-8 to power portable energy sources [27]. Portable energy sources are vital to a number of applications where grid power is difficult to obtain, such as cell towers in remote locations and military operations in forward operating bases. The recent advent of solid oxide fuel cells capable of operating on liquefied petroleum gas (LPG) provides an avenue towards better overall energy efficiency over the traditional diesel fuel generators currently in use. However, LPG supply lines do not exist in rural locations, and in disaster response situations it would work against humanitarian aims to

divert critical transportation resources from carrying foodstuffs and medical supplies to transporting LPG. This need motivated our group to develop a fuel flexible catalyst that can convert diesel, gasoline, or JP-8, readily available fuels in such locations with established supply lines, to LPG on-site.

Catalytic cracking is the most promising and economic process to generate LPG from diesel or JP-8. However, during direct catalytic cracking of heavy hydrocarbons at high conversion, catalyst deactivation via coke formation is inevitable [27]. Thus, one must maintain the rate of cracking while promoting air-only burn-off of the coke via additives. Additionally, although modern diesel fuels have been de-sulfurized substantially (15 ppm) to meet EPA requirements, diesel in other parts of this world may contain much higher levels of sulfur and JP-8 can contain as much as 3,000 ppm of sulfur. Sulfur deactivation of catalysts is well-known in the field [43], and often is an irreversible process, thus catalysts that can “select out” sulfur containing molecules, and thus prevent deactivation, are required.

We developed a multi-tiered high-throughput approach combining qualitative primary screens using thin-film samples with quantitative powder screening using mg powder quantities to rapidly discover a catalyst formulation with optimized properties. The total time from project start to final catalyst selection was ~ 10 months. This project would have never been successful using traditional single sample at a time methodologies. The freedom of HT afforded our group to branch out from “safe” materials such as normal transition metal oxides to more high-risk high-reward systems such as zeolites. The screening process was broken down into three phases. First, initial identification of a base catalyst formulation with suitable conversion and selectivity was performed on powder catalyst samples using a 16-well high-throughput reactor, shown in Fig. 4 [27]. This broad initial screen vetted more than 50 solid-acid catalysts, including a series of transition metal oxides and zeolites, in less than 3 months. Small

Fig. 3 An example of a workflow encompassing HT experimentation and fundamental modeling to expedite the rate of catalyst development. *Data flows through the vertical arrows between the two thrusts via searchable databases*



pore MFI-type zeolites proved to have the highest overall conversion at the lowest temperatures although they were initially thought to be unlikely to produce sufficient yields due to sensitivity to sulfur poisoning. This is a tremendous example of HT experimentation allowing scientists to try outside of the box materials, leading to breakthrough discoveries.

Thin film composition spread studies of bimetallics were then undertaken to identify potential anti-coking additives [29]. Here, nanoparticles of the additives were deposited onto pressed zeolite pellets to simulate nanoparticles generated via wet impregnation and then images of the pellet surfaces during JP-8/air exposure were recorded. Figure 5 shows an example of a series of images of taken on a thin film sample as a function of time during JP-8 dosing. A series of four bimetallic systems were screened. The bimetallic Pt–Gd system exhibited the overall slowest coking rate, the lowest coke burn-off temperature, and was recommended for final powder verification. A Pt–Gd modified zeolite demonstrated superior time on stream performance with practically no decrease in yield over multiple cracking/regeneration cycles during final validation. Figure 6 summarizes the conversion results for catalysts studied over the course of this investigation. The best catalyst is being commercialized in a packaged plug-and-play fuel reformer targeted towards military and commercial applications [44].

During this study there were several opportunities where timely intervention by either theory or data science could

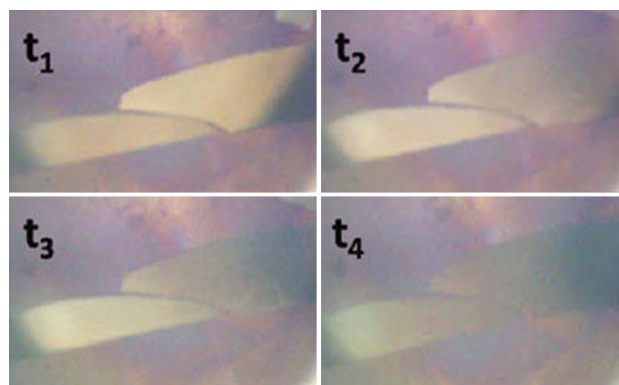


Fig. 5 Images of Pt–Gd/ZSM-5 during JP-8 cracking. Darkening on the sample corresponds to the coke deposition of coke on the catalyst surface [29]

have further reduced the time required to identify optimal catalyst compositions, and possibly further increased the overall figure of merit, as shown in Fig. 3. Theoretical understanding of the role of the transition and rare earth metals in the zeolitic framework, in particular their role in mitigating coking, could have provided guidance towards other and better alternative additives. Illumination of the role of zeolite structure and Si/Al ratio on cracking of long chain hydrocarbons might have provided additional rational design criteria for zeolite selection. Likewise, a structured database illustrating secondary catalytic properties of importance to the project such as coking resistance

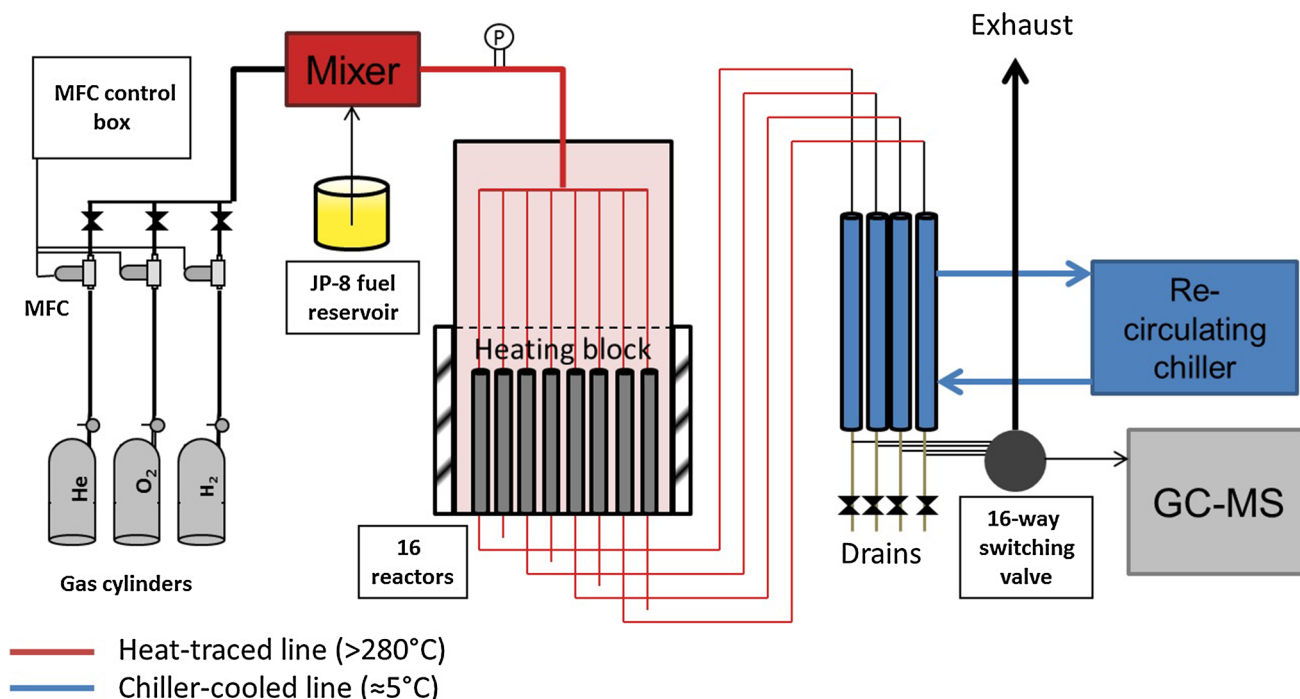


Fig. 4 High throughput reaction system for JP-8 cracking catalyst screening with 16 parallel reactors. End products from each reactor are analyzed separately by on-line GC–MS

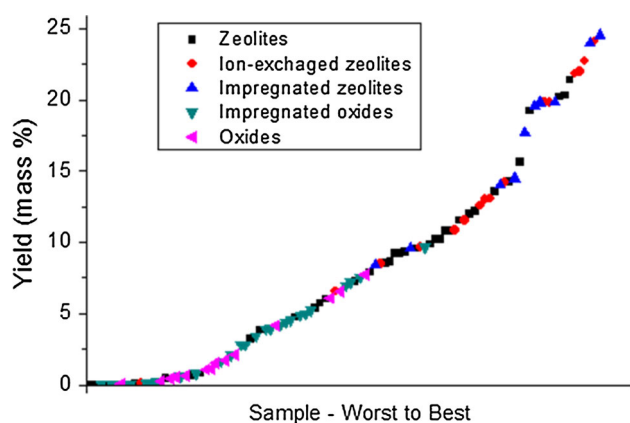


Fig. 6 An overview of the catalytic performance of more than 100 catalysts tested in a 9 month period illustrating the progression from lowest to highest yield and its correlation with catalyst structure

of transition metals, coking rates of zeolites, and their use in direct cracking of even simple hydrocarbon mixtures would have been of tremendous use. In fact many of these secondary characteristics of interest for the catalyst eventually selected have been investigated previously but the results are scattered throughout the technical literature in myriad papers, extended abstracts, and buried in patents or worse, unreported due to perceived failure. A coalesced and searchable version of this technical literature would have greatly accelerated hypothesis creation and materials development.

3 Catalyzing the Next Stage in the Materials Super Highway

To this moment, in the field of HT catalysis approaches that synthesize and screen catalysts using scalable techniques have been used only by a limited portion of the community. Many synthesis techniques are capable of producing sufficient quantities of catalyst for a bench-scale reactor, but would be commercially prohibitive, owing to excessive solvent or reductant usage or the use of expensive precursors. Also to meet the commercialization aspirations of the MGI, catalysts should not be screened for an abstract figure of merit under semi-realistic conditions. Instead it would be beneficial to embrace the full complexity of catalytic systems with synthesis techniques, inexpensive and commercially viable supports and reaction conditions (pressures, temperature, gas-hour space velocities) comparable to those used at industrial-scale. Likewise, if the purpose of catalyst development is to create catalysts that can be used to manufacture useful chemicals at industrial scales the use of surrogate feedstocks could lead to unexpected roadblocks at the development stage arising due to exposure to more complicated feedstocks. Even

fundamental HT studies could benefit from running in more realistic conditions, although this would necessitate more complex data analytics to extract mechanistic understanding.

Future HT-MGI studies could be substantially empowered if an emphasis were to be placed on understanding the linkages between theoretical identification of a new catalyst and its actual experimental synthesis. This will require the creation of multi-scale modeling tools that can be coupled with in situ experimental tools for monitoring catalyst composition, size, morphology, and elemental coordination to create new catalyst design rules. It is an almost trivial matter to build a statistical design based on design of experiments and use that to identify a condition that will synthesize the desired materials, however, turning that into transferable knowledge that can be generalized to new systems is a great challenge yet to be addressed.

Even once a material can be produced after theoretical identification, if its catalytic properties and structure are not well understood then the theory-experimental cycle is not closed. The current state-of-the-art in HT experimental catalysis techniques are largely results driven, this is particularly the case for parallel reactor studies, and thus provide limited insight into why one material or material synthesis procedure produces markedly different activity and selectivity. The current standard in the HT community is to perform detailed structural investigations *ex situ* or at substantially reduced pressure/temperature/concentration of the catalyst to try to extract information regarding how the catalyst “reacts” to its environment. It is well understood in the field, however, that catalyst active sites can be generated or changed under reaction conditions [45, 46].

In terms of the HT structural characterization of catalysts, great advances in the state-of-the-art are on the horizon, which will permit ever greater amounts of detailed information to be generated rapidly. For instance, techniques for HT screening of nanoparticle morphology and size are being developed by parallel efforts such as Scalable Nano-manufacturing, which will greatly accelerate the information feedback cycle in DoE experiments [47]. Meanwhile, techniques to speed up data acquisition and analysis via HRTEM are currently being developed that will permit high quality data to be taken and interpreted at a rate commensurate with that of sample generation [48]. With such techniques in hand, one could envision the creation of catalyst array samples where tens to hundreds of different synthesis conditions are probed to yield information regarding their effect on particle shape, morphology, and possibly activity in a next, next generation gas-dosing cell. Great strides have been made in performing in situ EXAFS and XANES experiments to monitor local elemental coordination in parallel, the best example of this to date being the Miller group’s 6-shooter

(Fig. 7) [37, 49, 50]. However, such systems often involve substantial beam path-lengths and can only operate at ambient pressure conditions, next generation XAFS/XANES high-speed instruments operating under industrially relevant conditions would significantly broaden existing understanding of how the chemical environment alters catalyst coordination. A similar opportunity exists for in situ x-ray diffraction studies, particularly those at national synchrotron sources, where bulk catalyst crystal structure under elevated pressures and temperatures would provide information on bulk structural changes applicable to a broad range of fields. In all cases, if the ability were present to simultaneously integrate online chemical analytics with structural and/or chemical characterization techniques a new understanding of the link between the observed catalyst performance and changes to its microstructure could be obtained.

Such a drastic increase in the quantity of in situ chemical, structural, and catalytic information generated per experiment will render traditional means of analyzing data onerous, if not impossible. Moreover, the issues of data quality, reproducibility of results, and effective dissemination of terabyte quantities of data must be addressed to maximize return of investment for public and private entities. Modern HT in situ diffraction studies now produce in excess of 100 GBs of diffraction patterns in a single day that then need to be correlated to composition, synthesis condition, activity, and an assortment of other measurement techniques. To accomplish this in a reasonable amount of time, one can look to the fields of bioinformatics, DFT, and functional materials for examples of methods of organizing, archiving and querying large-scale databases. Databases like AflowLib and Materials Project use JavaScript Object Notation (JSON) tags to organize data and permit simple keyword or complex interdependencies to be searched via intuitively designed websites

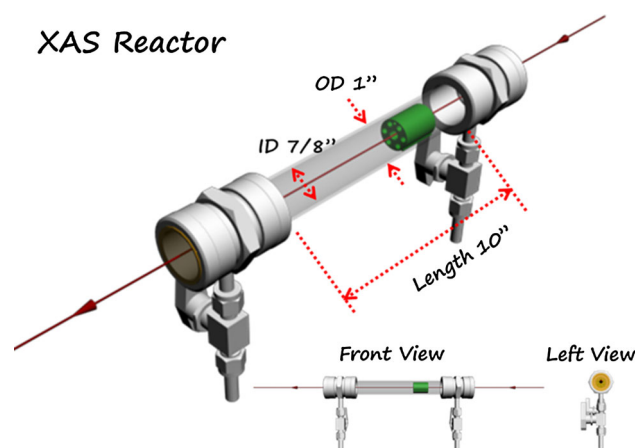


Fig. 7 Schematic of the 6-shooter sample holder (green part) in situ XAFS/XANES reactor with dimensions of the reactor marked

[51, 52]. Examples of such databases for experimentally determined materials properties exist such as MatNav [53], however, the scope of such databases is often limited. It is necessary to begin developing such a searchable database for the HT catalysis community. A problem arises, however, in that while calculations performed within a single entity such as AFlowLib or Materials Project are largely self-consistent, experimental values of a particular property (say catalytic activity) are sensitive not only to synthesis and processing of the catalyst, but also to the activation and catalytic reaction condition. It is therefore a necessity to establish standard experimental procedures, and methods for systematically reporting all metadata that will permit valid cross-comparison of catalytic properties. One such method for this would be the development of catalyst standards for important reactions and associated reaction conditions that can be used to calibrate individual reactors. The mandate for accumulating, curating and storing the obtained catalytic data may be most effective if it is assigned to an individual entity. In the US, a government agency such as NIST with an established record of accomplishment in collaborating with industry, an existing mandate to promote commerce, and a track record of hosting materials property databases, would be a natural fit for such a database.

Merely producing a database is a meaningless enterprise, however, if there do not exist sufficient resources for extracting new knowledge from it [54]. Moreover, performing hundreds of in situ studies on catalysts which are not active for a particular reaction could potentially clog the database with meaningless data. A two-pronged attack would help to maximize the amount of information provided by each data point. Firstly, new methods of analyzing structural, spectral, and catalytic data that utilize recent

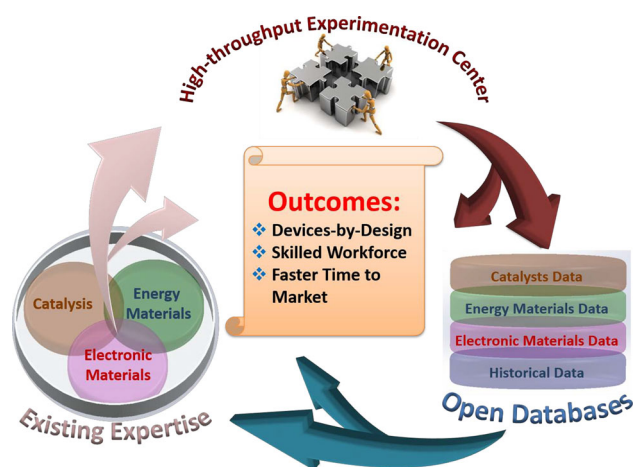


Fig. 8 Envisioned High-throughput Experimentation Center demonstrating that common materials characterization needs being served by a “brick and mortar” center. Established catalysis specific tools would form a foundation by providing open access

advances in machine learning would be of great use for establishing correlations between composition, processing and catalytic activity in large open databases. The functional materials community currently leads the HT field in this respect with freely available software such as CombiView that can use a number of cluster analysis tools as well as supervised machine learning algorithms [55–57]. Such approaches can be adopted and modified by the HT catalysis community to provide a means of linking structure, processing, and functionality. Secondly, these techniques can not only benefit end of the experiment analysis but can also be applied in real-time to facilitate on-the-fly design of experiments. As XANES, XRD, and catalytic activity data are being taken, data can be analyzed and clustered into regions of interest, permitting an optimal distribution of data density that prioritizes not “winning catalysts” but those materials and conditions that reside on the edge of being effective.

It would also be worthwhile to establish an accepted metric for estimating the economics of experimental studies. Very few academic experimentalists consider cost-benefit analysis in the process of designing a new set of experiments. While we do not suggest this criteria be used as the basis of funding for academic research, this is a skill set that could guide future industrial studies by providing a ball park return on investment for new research projects. On the other hand, systems level engineering perspectives could benefit funding agencies, by providing the argument against technologies that, though technically feasible, would not provide sufficient impact to merit large programs. Such tools could be used to provide guidance for the exploration of new catalytic systems by helping to determine if an exploitative study of existing catalytic materials might be more cost-effective than an explorative study into exotic catalysts and synthesis techniques.

To effect a large scale change in the manner by which new materials are conceived will require the creation of high throughput centers spanning large number of fields, including catalysis (Fig. 8). Leveraging the existing network of practitioners of high-throughput experiments can create an “HTE materials superhighway,” and as a result maximize scientific impact and return on public investment. The network can be mobilized to address the urgent, high impact, and materials-constrained technologies mentioned earlier. Moreover, although each field has its own particular figure of merit, many of the characterization techniques required in the different fields are common e.g. diffraction and high resolution TEM for bulk structural information. There are also similar gaps in the ability to create standards to acquire, process, archive, and effectively probe truly large databases containing data and metadata generated by disparate groups. These mutual goals make it possible to create centralized facilities with

expertise in designing, standardizing, and formatting combinatorial datasets, which can then be used to inform experimentalists and theorists across multiple fields of research. Brick and mortar physical centers located at synchrotron light sources are natural fits for generalized structural investigations; while virtual centers specialized catalysts synthesis and parallel bed catalytic reactors would be established in academic and industrial labs with pre-existing expertise. Centralizing data collection will help de facto in the creation of data acquisition, processing, and formatting standards, with multiplicative increases in the generalizability of each new data set. With these tools in hand, a new era of accelerated catalyst discovery, optimization, and commercialization would be enabled, solving a large number of pressing industrial and environmental issues, while providing the data that will drive a revolution in knowledge based catalyst design.

Acknowledgements Hattrick-Simpers would like to acknowledge the support by the US National Science Foundation under Grant DMR 1439054.

References

1. Holdren JP (2011) Materials genome initiative for global competitiveness. Office of Science and Technology Policy and National Science and Technology Council, Washington, D.C
2. Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson KA (2013) *APL Mater* 1:011002
3. Wang L, Liu J (2013) *Front Energy* 7:317–332
4. Besenbacher F, Chorkendorff I, Clausen BS, Hammer B, Mollenbroek AM, Nørskov JK, Stensgaard I (1998) *Science* 279:1913–1915
5. Honkala K, Hellman A, Remediakis IN, Logadottir A, Carlsson A, Dahl S, Christensen CH, Nørskov JK (2005) *Science* 307:555–558
6. Studt F, Abild-Pedersen F, Bligaard T, Sørensen RZ, Christensen CH, Nørskov JK (2008) *Science* 320:1320–1322
7. Studt F, Sharafutdinov I, Abild-Pedersen F, Elkjær CF, Hummelshøj JS, Dahl S, Chorkendorff I, Nørskov JK (2014) *Nat Chem* 6:320–324
8. Rodriguez JA, Ma S, Liu P, Hrbek J, Evans J, Perez M (2007) *Science* 318:1757–1760
9. Cargnello M, Doan-Nguyen VVT, Gordon TR, Diaz RE, Stach EA, Gorte RJ, Fornasiero P, Murray CB (2013) *Science* 341:771–773
10. Kwak JH, Hu J, Mei D, Yi CW, Kim DH, Peden CHF, Allard LF, Szanyi J (2009) *Science* 325:1670–1673
11. Nørskov JK, Abild-Pedersen F, Studt F, Bligaard T (2011) *Proc Natl Acad Sci* 108:937–943
12. Tao F, Dag S, Wang LW, Liu Z, Butcher DR, Bluhm H, Salmeron M, Somorjai GA (2010) *Science* 327:850–853
13. Saavedra J, Doan HA, Pursell CJ, Grabow LC, Chandler BD (2014) *Science* 345:1599–1602
14. Giusepponi S, Celino M (2013) *Int J Hydrog Energ* 38:15254–15263
15. Chen P (2003) *Angew Chem Int Ed* 42:2832–2847
16. Erisman JW, Sutton MA, Galloway J, Klimont Z, Winiwarter W (2008) *Nat Geosci* 1:636–639

17. Hagemeyer A, Volpe AF (2014) Modern applications of high throughput r&d in heterogeneous catalysis. Bentham Science
18. Maier WF, Stöwe K, Sieg S (2007) *Angew Chem Int Ed* 46:6016–6067
19. Green ML, Takeuchi I, Hatrick-Simpers JR (2013) *J Appl Phys* 113:231101
20. Smotkin ES, Díaz-Morales RR (2003) *Ann Rev Mater Res* 33:557–579
21. Reichenbach HM, An H, McGinn PJ (2003) *Appl Catal B* 44:347–354
22. Scheidtmann J, Weiß PA, Maier WF (2001) *Appl Catal A* 222:79–89
23. Reddington E, Sapienza A, Gurau B, Viswanathan R, Sarangan S, Smotkin ES, Mallouk TE (1998) *Science* 280:1735–1737
24. Kondratyuk P, Gumuslu G, Shukla S, Miller JB, Morreale BD, Gellman AJ (2013) *J Catal* 300:55–62
25. Hustad PD, Kuhlman RL, Arriola DJ, Carnahan EM, Wenzel TT (2007) *Macromolecules* 40:7061–7064
26. Arriola DJ, Carnahan EM, Hustad PD, Kuhlman RL, Wenzel TT (2006) *Science* 312:714–719
27. Bedenbaugh JE, Kim S, Sasmaz E, Lauterbach J (2013) *ACS Combi Sci* 15:491–497
28. Wen C, Barrow E, Hatrick-Simpers J, Lauterbach J (2014) *Phys Chem Chem Phys* 16:3047–3054
29. Yang K, Bedenbaugh J, Li H, Peralta M, Bunn JK, Lauterbach J, Hatrick-Simpers J (2012) *ACS Combi Sci* 14:372–377
30. Snively CM, Oskarsdottir G, Lauterbach J (2001) *Angew Chem Int Ed* 40:3028–3030
31. Snively CM, Oskarsdottir G, Lauterbach J (2001) *Catal Today* 67:357–368
32. Snively CM, Lauterbach J (2002) *Spectroscopy* 17:26
33. Caruthers JM, Lauterbach JA, Thomson KT, Venkatasubramanian V, Snively CM, Bhan A, Katare S, Oskarsdottir G (2003) *J Catal* 216:98–109
34. Hunter D, Osborn W, Wang K, Kazantseva N, Hatrick-Simpers J, Suchoski R, Takahashi R, Young ML, Mehta A, Bendersky LA, Lofland SE, Wuttig M, Takeuchi I (2011) *Nat Commun* 2:518
35. Cui J, Chu YS, Famodu OO, Furuya Y, Hatrick-Simpers J, James RD, Ludwig A, Thienhaus S, Wuttig M, Zhang Z, Takeuchi I (2006) *Nat Mater* 5:286–290
36. Gregoire JM, Van Campen DG, Miller CE, Jones RJR, Suram SK, Mehta A (2014) *J Synchrotron Radiat* 21:1262–1268
37. D'Souza L, Regalbuto JR, Miller JT (2008) *J Catal* 254:157–169
38. Diaz-Quijada GA, Peytavi R, Nantel A, Roy E, Bergeron MG, Dumoulin MM, Veres T (2007) *Lab Chip* 7:856–862
39. Low JJ, Benin AI, Jakubczak P, Abrahamian JF, Faheem SA, Willis RR (2009) *J Am Chem Soc* 131:15834–15842
40. Zhang Y, McGinn PJ (2012) *J Power Sources* 206:29–36
41. Wen C, Dunbar D, Zhang X, Lauterbach J, Hatrick-Simpers J (2014) *Chem Commun* 50:4575–4578
42. Dellamorte JC, Barteau MA, Lauterbach J (2009) *Surf Sci* 603:1770–1775
43. Bartholomew CH (2001) *Appl Catal A* 212:17–60
44. Lauterbach J, Glascock M, Bedenbaugh J, Chien CY, Jangam A, Salim S, Kim S, Tilburg R (2013) US patent appl 20130041198
45. Nishihata Y, Mizuki J, Akao T, Tanaka H, Uenishi M, Kimura M, Okamoto T, Hamada N (2002) *Nature* 418:164–167
46. Tao F, Grass ME, Zhang YW, Butcher DR, Renzas JR, Liu Z, Chung JY, Mun BS, Salmeron M, Somorjai GA (2008) *Science* 322:932–934
47. Tracy JB, Crawford TM (2013) *MRS Bull* 38:915–920
48. Vogt T, Dahmen W, Binev P (2012) *Modeling nanoscale imaging in electron microscopy*. Springer, New York
49. Pazmiño JH, Miller JT, Mulla SS (2011) Nicholas Delgass W, Ribeiro FH. *J Catal* 282:13–24
50. Wang H, Miller JT, Shakouri M, Xi C, Wu T, Zhao H, Akatay MC (2013) *Catal Today* 207:3–12
51. Curtarolo S, Setyawan W, Wang S, Xue J, Yang K, Taylor RH, Nelson LJ, Hart GLW, Sanvito S, Buongiorno-Nardelli M, Mingo N, Levy O (2012) *Comp Mater Sci* 58:227–235
52. Persson K <http://www.materialsproject.org>. Accessed 18 Nov 2014
53. Ashino T, Oka N (2007) *Data Sci J* 6:S847–S852
54. Hatrick-Simpers JR, Green ML, Takeuchi I, Barron SC, Joshi AM, Chiang T, Davydov A, Mehta A, Fulfilling the promise of the materials genome initiative via high-throughput experimentation (MRS 2014)
55. Takeuchi I, Long CJ, Famodu OO, Murakami M, Hatrick-Simpers J, Rubloff GW, Stukowski M, Rajan K (2005) *Rev Sci Instrum* 76:062223
56. Long CJ, Hatrick-Simpers J, Murakami M, Srivastava RC, Takeuchi I, Karen VL, Li X (2007) *Rev Sci Instrum* 78:072217
57. Kusne AG, Gao T, Mehta A, Ke L, Nguyen MC, Ho KM, Antropov V, Wang CZ, Kramer MJ, Long C, Takeuchi I (2014) *Sci Rep* 4:6367