

## Das Katalysatorgenom

Jens K. Nørskov\* und Thomas Bligaard\*

Die Suche nach dem Materialgenom – den Charakteristika eines Materials, die seine funktionellen Eigenschaften definieren – hat begonnen, und damit eine neue Ära der Materialforschung, in der große Mengen an Materialdaten zugänglich werden. Man hofft, dass dadurch die Entdeckung neuer Materialien deutlich beschleunigt wird – vor allem im Bereich der katalytischen Materialien, denn es besteht ein großer Bedarf an neuen Katalysatoren und Prozessen, um Treibstoffe und Chemikalien nachhaltig zu produzieren.

Katalyse wird überall zur Steuerung chemischer Prozesse genutzt: In der Natur regeln Enzyme die chemischen Fabriken in Zellen, und in der chemischen Industrie geht fast nichts mehr ohne Katalysatoren: Nahezu alle Chemikalien, einschließlich der Treibstoffe, hatten bei der Produktion Kontakt zu ihnen. Die meisten großtechnischen Katalysatoren sind Feststoffe, und die Katalyse läuft an ihrer Oberfläche ab, doch es gibt auch Bereiche, in denen wegen der höheren Selektivität von Niedertemperaturprozessen molekulare Katalysatoren vorgezogen werden.

Warum ist ein Material ein guter Katalysator und ein anderes ein schlechter? Die Suche nach dem Katalysatorgenom läuft schon seit über hundert Jahren. Heute ist der Fortschritt sehr schnell, und die Ansprüche an neue Ansätze des Katalysatordesigns sind gewaltig. Effizientere Katalysatoren würden nicht nur den Energieverbrauch

und die Abfallmengen in der chemischen Industrie reduzieren, sondern sie sind auch essenziell für den Aufbau eines vollständig neuen, nachhaltigen chemischen Sektors.

Wenn wir uns nicht weiterhin auf fossile Rohstoffe verlassen wollen, müssen wir neue, nachhaltige Wege zu Treibstoffen wie auch Chemikalien finden. In den meisten Szenarien wird auf die Energie der Sonne gesetzt (Abbildung 1), und neue katalytische Prozesse sind für deren Verwirklichung von zentraler Bedeutung.

Wie könnte das Katalysatorgenom aussehen? Es könnte eine Karte sein, in der alle möglichen Katalysatorstrukturen mit den Geschwindigkeiten aller möglichen elementaren Reaktionen bei allen möglichen Reaktionsbedingungen sowie mit der elektronischen Struktur und den spektroskopischen Daten, die die unterschiedlichen Intermediate charakterisieren, verknüpft sind. Man stelle sich vor, wir hätten alle diese Daten und effiziente Methoden, sie zu nutzen. Wir könnten dann für jede vorstellbare katalytische Reaktion Katalysatoren konstruieren, indem wir zunächst alle möglichen Reaktionspfade analysierten und danach nach dem Material suchten, dass den ausgewählten Prozess am besten katalysiert.

Doch von einer Erfüllung dieses Traumes sind wir noch weit entfernt. Die Datenmenge

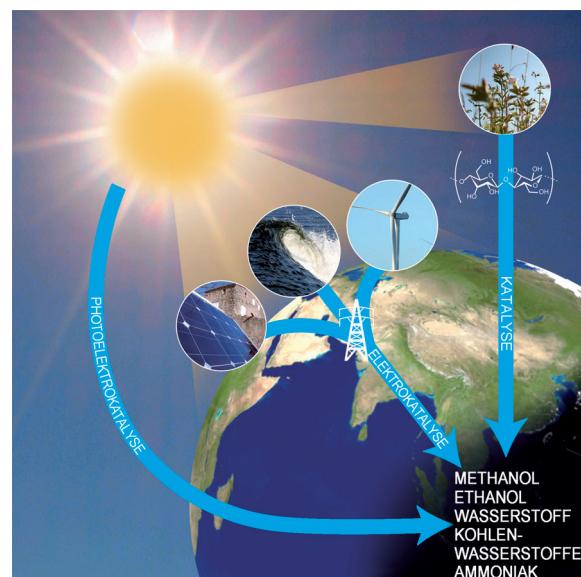


Jens K. Nørskov  
Professor, Stanford University  
und SLAC National  
Accelerator Laboratory



Thomas Bligaard  
Senior Staff Scientist  
SLAC National  
Accelerator Laboratory

wäre riesig, und die experimentelle Arbeit, die für das Sammeln der Daten nötig wäre, wäre undurchführbar. Einen großen Schritt vorwärts bedeutete die Entwicklung von auf der elektronischen Struktur basierenden Methoden zur Berechnung von Katalysatoreigenschaften. Computerberechnungen eignen sich sehr gut, um systematische Daten für eine Vielzahl an Materialien und Reaktionen zu erhalten. Ein wesentliches Element dieses Prozesses ist, die (oft dichtefunktionaltheoretischen) Rechnungen anhand genauer Experimente oder aufwendiger Rechnungen für ausgewählte Systeme zu bewerten,



**Abbildung 1.** Unterschiedliche Wege zu Treibstoffen und Grundchemikalien ausgehend von Sonnenlicht. Egal ob der Energiefluss des Sonnenlichts in Form von Biomasse, durch zwischenzeitliche Stromerzeugung z. B. mit Photovoltaikzellen oder Windrädern oder direkt durch photoelektrochemische Reaktionen „geerntet“ wird, in jedem Fall wird ein effizienter Katalysator benötigt, der vorzugsweise aus auf der Erde reichlich vorhandenen Materialien hergestellt sein sollte.

[\*] Prof. J. K. Nørskov, Dr. T. Bligaard  
SUNCAT Center for Interface Science and  
Catalysis  
SLAC National Accelerator Laboratory  
und  
Department of Chemical Engineering  
Stanford University  
Stanford (USA)  
E-Mail: noriskov@stanford.edu  
bligaard@stanford.edu

um die Grenzen der Zuverlässlichkeit zu bestimmen. Die „CatApp“ ist ein Beispiel für eine Datenbank von Reaktions- und Aktivierungsenergien für elementare Prozesse auf Übergangsmetalloberflächen (siehe <http://suncat.slac.stanford.edu/catapp/> und *Angew. Chem.* **2012**, *124*, 278–280).

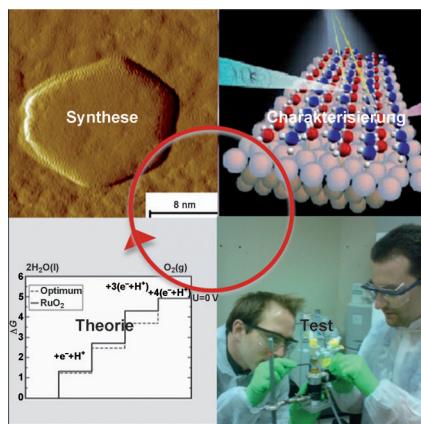
**E**s gibt große Bereiche (beispielsweise Übergangsmetallocide), in denen die konventionellen Rechenmethoden nicht genau genug sind und deshalb sehr viel mehr Arbeit in die Entwicklung von Methoden gesteckt werden muss, die ausreichend genau und effizient sind, um für das Kartieren großer Datens Mengen eingesetzt zu werden. Dazu kommt die Frage, wie man mit der Komplexität realer Katalysatoren umgehen soll – bei molekularen Katalysatoren in Lösung mit dem Einfluss des Lösungsmittels und bei Heterogenkatalysatoren mit dem jeweiligen aktiven Zentrum und dem Vorliegen mehrerer fester Phasen. Es versteht sich von selbst, dass das Ziel nicht sein kann, jedes System vollständig zu beschreiben; vielmehr muss es gelingen, die wichtigsten Kenngrößen zu extrahieren, die die katalytischen Eigenschaften beeinflussen. Sobald solche Deskriptoren bekannt sind, kann man in diesem viel kleineren Parameterraum suchen.

**W**ie identifiziert man Deskriptoren? Ein Weg ist, sich um ein Verständnis der Reaktivität anhand der elektronischen Struktur des Katalysators zu bemühen. Ein anderer ist die Suche nach Korrelationen in vorhandenen Daten. Die Katalysatoraktivität oder -selektivität, die in teuren, gewissenhaft geplanten Experimenten ermittelt wird, kann beispielsweise mit einem berechneten Deskriptor korreliert werden. Der einfach zugängliche berechnete Deskriptor eignet sich somit für die Planung neuer Experimente mit Verbindungen, für die er gute katalytische Eigenschaften vorschreibt. Ähnlich können manchmal aufwendige Rechnungen mit weit weniger aufwendigen korreliert werden, was weniger genaue, dafür aber ökonomischer erhaltene Deskriptoren für aufwendige Simulationen liefert. Beispiele sind die Paare anspruchsvollere Berechnungen der elektronischen Struktur – ein niedrigeres Theoriensi-

veau, kompliziertere Berechnungen der freien Energie – Berechnungen von Grundzustandsenergien oder schwieriger zu erhaltende Reaktionsbarrieren – Adsorptionsenergien. Wenn die Korrelationen genau genug sind, können solche Vereinfachungen die Zahl der Systeme, die rechnerisch zugänglich sind, um mehrere Größenordnungen erhöhen und so die Datenfülle drastisch steigern und das Durchführen einer Suche ermöglichen. Kontinuierlich bessere theoretische Verfahren haben mit der Zeit ebenfalls zu einer höheren Genauigkeit der berechneten Katalysatoreigenschaften geführt. Je komplexer die analysierten katalytischen Prozesse werden und je größer die Menge an berechneten Daten ist, desto wahrscheinlicher wird, dass maschinelle Lernverfahren eine wichtigere Rolle beim Formen des Katalysatorgenoms spielen werden. Maschinelles Lernen lässt sich grob definieren als ein Satz von Methoden, die das Ziel haben, (weniger offensichtliche) Muster in großen Datens Mengen zu identifizieren oder (vernünftige) Vorschläge auf der Basis unvollständiger Datensätze zu machen.

**D**a alle Ansätze, die Entdeckung neuer Katalysatoren zu beschleunigen, auf die Verfügbarkeit großer Datens Mengen setzen, dürfte das Katalysatorgenom zunächst primär eine Datenbank berechneter Eigenschaften sein, die mit experimentellen Schlüsseldaten für das Bewerten und das Festlegen von Korrelationen angereichert ist. Das Katalysatorgenom sollte jedoch als mehr als nur die zugrundeliegenden Daten gesehen werden. Da das für die Katalyse wesentliche Wissen auf der Kenntnis der Korrelationen und Beziehungen zwischen katalytisch wesentlichen Datensätzen beruht und diese häufig noch entdeckt werden müssen, ist das Katalysatorgenom auch eine Sammlung der notwendigen Konzepte, Analysenwerkzeuge, Suchmethoden und Lernalgorithmen, um Daten zu generieren, wo es noch keine gibt.

**A**lle Bemühungen, die Entdeckung zu beschleunigen, haben das zentrale Ziel, sowohl viele Forschungsgruppen als auch verschiedene Verfahrensansätze zu integrieren. Damit ist eine zentrale Aufgabe das Schaffen einer Plattform,



**Abbildung 2.** Prinzip eines integrierten Katalysatorgenomikansatzes zur Entdeckung neuer katalytischer Materialien. Synthese, Charakterisierung und Testen werden rasch mit Rechnungen verflochten. Das Wechselspiel treibt nicht nur die Theorie in Richtung höherer Genauigkeit und erlaubt die Validierung und Evaluierung der Unsicherheit von Berechnungen, es funktioniert auch umgekehrt, d.h., die Theorie treibt die Experimente an die Grenzen einer Steuerung auf atomarer Ebene.

auf der die Forschungsgruppen und die Verfahrensansätze interagieren können (Abbildung 2). Fortschritt wird nur möglich sein, wenn die ganze Forschergemeinschaft zusammenarbeitet. Wahrscheinlich wird es mehrere Datenbanken oder Plattformtechnologien geben. Für den Erfolg entscheidend ist, dass sie alle transparent formatiert sind, damit eine Software auf alle verfügbaren Daten zugreifen und sie nutzen kann. Da wohl keine Gruppe den kompletten Überblick darüber oder eine vollständige Vorstellung davon hat, wohin sich das Katalysatorgenom entwickeln wird, sollte unserer Meinung nach ein internationales offenes Wissenschaftlerforum geschaffen werden, dessen Aufgabe es ist, einige „Standards“ für die Integration von Katalysatordaten und -methoden zu definieren. Ein wichtiges Thema wird dabei die Behandlung proprietärer Daten sein. Damit das Katalysatorgenom als Mittel für die Entdeckung neuartiger technischer Katalysatoren wichtig wird, muss beispielsweise die Industrie Zugang zum gesamten „öffentlichen Genom“ haben und die Möglichkeit, es nahtlos mit ihrem eigenen „geschützten Genom“ für spezifische Prozesse oder Katalysatoren zu kombinieren, ohne fürchten zu müssen, dass ihre Daten öffentlich zugänglich werden.