

インフォマティクスが拓く 触媒科学のパラダイムシフト

高橋 啓介 Takahashi Keisuke

北海道大学 大学院理学研究院 化学部門 准教授
2017年よりCREST研究代表者

高橋 ローレン Takahashi Lauren

北海道大学 大学院理学研究院 化学部門 特任助教



特集2

OVERVIEW ■

近年、人工知能(AI)や情報科学(インフォマティクス)の発展により、さまざまな分野でデータ科学との融合による効率的な研究手法を模索する動きが活発になっている。触媒探索の分野でいち早く実験・計算・データ科学を統合した「触媒インフォマティクス」の確立に取り組むのは、北海道大学大学院理学研究院の高橋啓介准教授だ。資源量が豊富なメタンを、工業的により有用な物質に変換する触媒開発を軸に、触媒科学のパラダイムシフトを目指す。

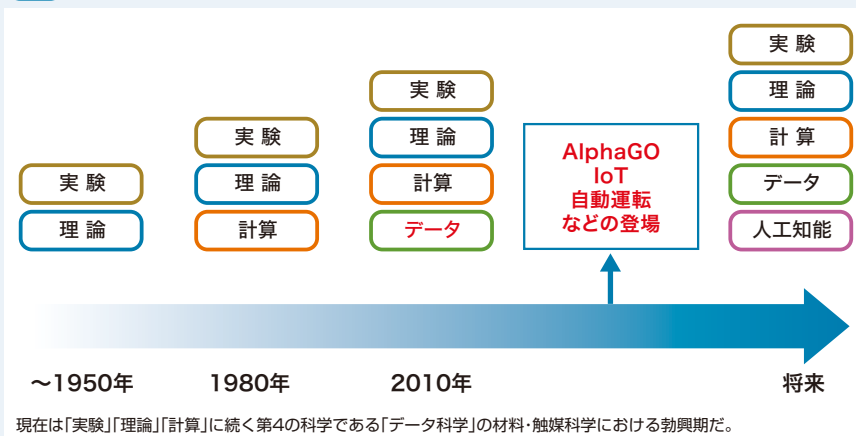
経験や直感頼みから脱却 開発が効率化、高速化

計算機やAIの発展、実験データベースの整備の進展などに後押しされ、多くの分野で融合が近年急速に進んできた。化学分野も例外ではない(図1)。一方「触媒研究にはインフォマティクスが十分に活用されていない」と指摘するのは、北海道大学大学院理学研究院の高橋啓介准教授だ。CREST「多様な天然炭素資源の活用」に資する革新的触媒と創出技術」研究領域で「実験・計算・データ科学の統合」によるメタン変換触媒の探索・発見と反応機構の解明・制御」を率いる。

化学分野におけるインフォマティクスには、これまで蓄積されてきた実験データなどを基に法則性を見いだしたり、物質情報を探索しやすい形に統合したりして、研究を加速することが期待されている。高橋さんはこれまでも、機械学習を活用した物質探索の手法「マテリアルズ・インフォマティクス(MI)」で、ある種の太陽電池の物質候補5400万件から望んだ性能を持つ11件に候補を絞り込むといった効率化を図る研究を手掛けてきた。

ではMIを触媒の探索にもすぐに応用できるかというと、そう単純ではない。一般的に、触媒は反応の前でそれ自体は変わらないと定義される。しかし実際には、反応中に状態がダイナミックに変化するので、物質の最初の情報を基に研究を進めて

図1 材料・触媒科学の歴史



も、良い結果は得られないと高橋さんは指摘する。「触媒は反応中、まるで生き物のように変化し、状態を捉えることはとても難しいのです。だからこそ、触媒インフォマティクスが確立できれば、経験や直感頼みから脱却し、触媒開発が劇的に効率化、高速化するはずだ」と語る。

一番難しい課題に挑戦したい 重要さを増すメタン酸化反応

高橋さんが目をつけたのは、触媒研究の難問として残るメタン酸化カップリング反応(OCM)の触媒の開発だ(図2)。メタンは、天然ガスの主成分で、以前から燃料としてはもちろん、メタノールや水素ガスといった化成品の原料として工業的にも用いられてきた。構造の単純さや私たちにとっての身近さとは裏腹に、メタンは炭素原子と水素原子の結びつきが強く、化学反応の制御は難しいとされる。

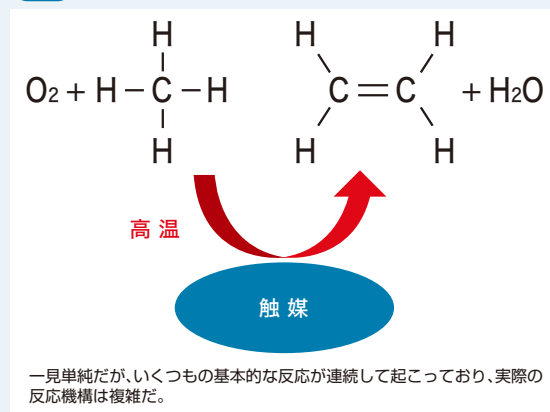
例えば、OCMIは酸素との反応を通してメタン同士をつなぎ、エタンやエチレンといった工業的により有用な物質を得るのが目的だ。化学反応を仲立ちする物質である触媒を混ぜた上、数百度の熱を加える(図2)。このとき、狙っていない複雑な反応まで

起こりやすく、結果として副生成物が多くできるという問題もある。「一番難しい課題に挑戦したいと考えました。有用な技術や方法論、知見はそういった研究に取り組むことで生まれます」。

世界中の研究者が精力的にこの課題に取り組み始めて30年以上経つものの、いまだに十分な性能を持った触媒は見つかっていない。また、2000年代後半にアメリカで起こったシェール革命の影響や、ある種の微生物の代謝からメタンガスを得る技術の進展により、OCMをはじめとしたメタンに関連する化学反応の制御の重要性はますます増している。CRESTでも革新的触媒の研究領域が発足し、まさに求めていたテーマだと迷うことなく応募した。

採択後すぐに取り組んだことの1つが、エチレンを生成するOCMに関する文献データの収集と分析だった。過去30年分、1868件のデータを機械学習で分析し、工業的に求められる30パーセント以上のエチレン収率が得られる触媒を探索した。すると、これまでに報告されたことのない56種の触媒が候補として浮上した。機械学習を活用した触媒の探索の有効性が示された形だ。また、収率の良い触媒と悪い触媒を決定づける元素の組み合わせも、傾向が大まかに把握できた。大量の実験データを深く探ることによって、触媒設計が可能となる成功イメージをつかんだときだった。

図2 エチレンを生成するメタン酸化カップリング反応



実験装置は自分たちで設計 3日で1万件以上のデータ

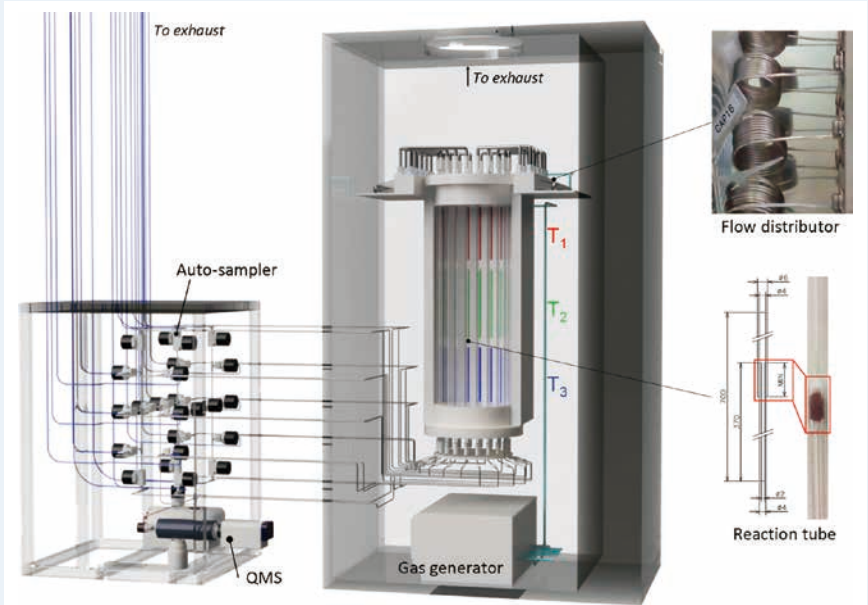
一貫性のある手順や条件で、大量の実験を行うのに欠かせないのが「ハイスループット実験装置」だ。実験に必要な種々の操作や測定を高速かつ自動的に行う装置で、すでに創薬の研究などでは広く活用されている。しかし、実験条件が広範囲に及び触媒研究において、オーダーメイドのハイスループット装置を開発するのは、技術面でも金銭面でも容易ではない。

例えば、数百度といった過酷な条件では、触媒が燃えたり溶けたりして装置にダメージを与えることもある。研究チームの1人でもある北陸先端科学技術大学院大学先端科学技術研究科の谷池俊明教授を中心に、研究の大黒柱となるこの装置を全員で意見を出し合いながら、自作することにした。できあがった装置は、20種類の触媒を異なる条件で一度に反応させることができるという(図3)。数千件以上のデータを1日で、自動的に集めることも可能だ。

実際に、装置の完成後に投稿した20年の論文では、59種類の触媒を用いて1万2708件のデータを得たことを報告した。過去30年分のデータより1桁多く、そして手法が厳密に制御され、一貫した質の高いデータをたった3日で得られることを証明したのだ。ただし、これほどのデータ量になると新たな苦労も生じる。

機械学習というと、データがあれば簡単に結果を手に行けるというイメージを持たれがちだ。しかし、機械が学習できる形にデータを前処理する必要があり、それを担うのは人間だ。高橋さんはやや困ったような笑顔を浮かべ、こう話す。「私は出てきたデータを全て読み、不適切なデータがないか、適切な分析手法になっているかを見極めていきます。ですが、これほどの量のデータとなると簡単ではありません。ハイスループット実験装置

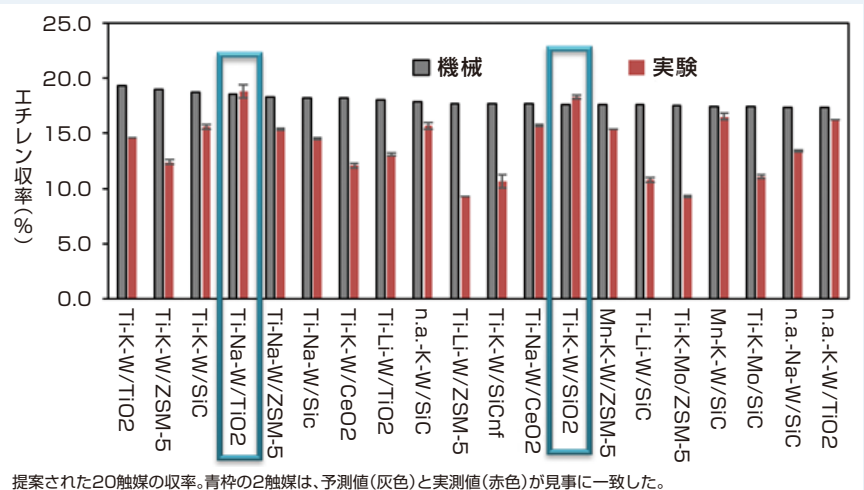
図3 ハイスループット実験装置



右側の装置の円筒部分には20本の管が入っていて、それぞれに触媒をセットし、自動的に実験を行うことができる。

Reprinted (adapted) with permission from ACS Catalysis. Copyright 2022 American Chemical Society.

図4 ハイスループット実験データとディープラーニングによって発見された触媒



提案された20触媒の収率。青枠の2触媒は、予測値(灰色)と実測値(赤色)が見事に一致した。

を導入したゆえの大変さですね」。

自前の実験装置で大量のデータを取得できるメリットは、偏りを除くことができる点にもある。従来は研究者が試行錯誤の末に、独自に編み出したノウハウや知見を駆使した実験の結果データを利用するより方法がなかった。しかし、機械学習に使うデータとしては、どうしても偏りが生じてしまう。研究者が気付くことができない、あるいは見落としているような意外な触媒を見いだすためには、偏りのあるデータはなるべく少ない方がよい。

データベースを無料公開 世界の研究者の右腕に

そこで高橋さんらは、3種類の金属が原子の数では1:1:1の比で混ざっているような固体の触媒について、考えられ得る36万通り以上の組み合わせからランダムに抽出した291種類を135条件で評価した。得られた4万件近いデータを機械学習で分析し、触媒として正常に機能すると予想された20種類の触媒を見いだした。さらに詳細な検討を行ったところ、8割にあたる16種類が良好な

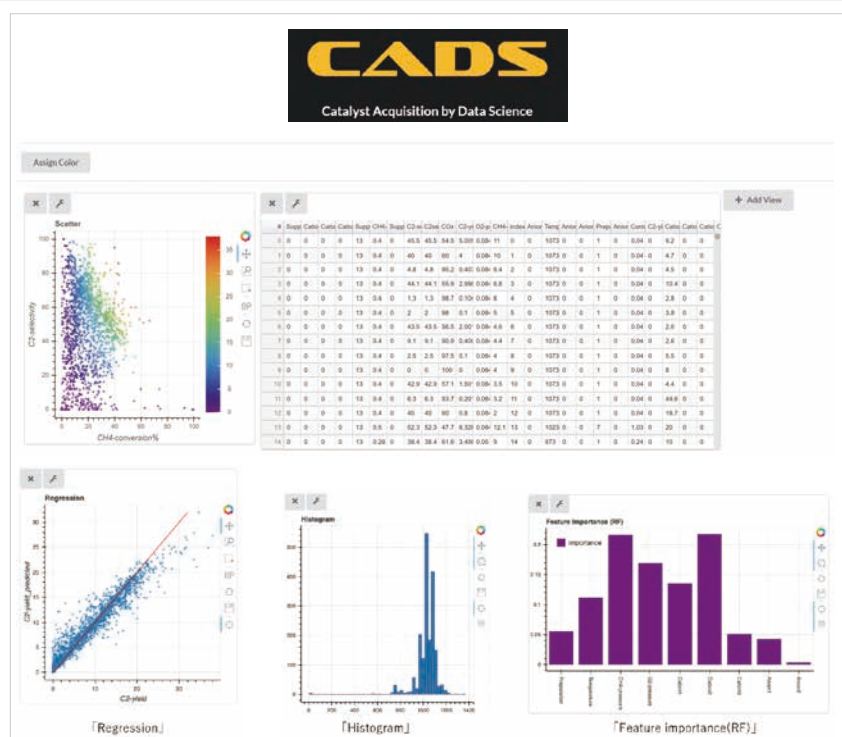
結果を示した。中には未報告の触媒も含まれており、改めて手法の有用性が示された(図4)。

OCMの触媒探索に有用なビッグデータを生み出し、有効性を示すことに成功した高橋さんは、そのデータベースを誰もが使いやすい形で無料公開することにした。開発したのは「CADS(<https://cads.eng.hokudai.ac.jp/>)」と名づけたプラットフォームだ(図5)。ユーザー自身がAIや機械学習に詳しくなくとも、クリック1つで、必要なデータを図表で表示させたり、それらに関連づけたり、機械学習を行うこともできる。現在、日本だけでなく米中独といった触媒研究が盛んな国々でも活用されており、総利用人数は4000人にも及ぶ。

データの公開に際しては、慎重な声もあったという。その中でもチーム内で対話を重ねて公開に至った背景には、データや技術のオープンソース化に対する高橋さんの熱意がある。Linuxはオープンソースのプログラムを理念とするオペレーティングシステム(OS)として有名だが、このユーザーコミュニティに高橋さんは若い頃から身を置いてきた。ユーザー間での活発な議論が、科学技術の発展に重要だと感じてきたという。

また、化学の分野でも特許などを取らなかったことが後続の研究や産業での応用につながり、革新を起こした研究もある。10年に触媒研究でノーベル化学賞を受賞した北海道大学の鈴木章名誉教授がその代表例だ。「触媒研究の新たな方法として定着するためには、ユーザーが議論でき、よりよいアイデアがあれば改良もできるオープンソース化が重要です。有用なデータを提供する人もいでしょうし、CADSが研究者の右腕に育ってくれたらうれしいですね」と高橋さんは語る。

図5 触媒プラットフォーム「CADS」



CADSにはOCMをはじめとしたさまざまなデータが登録されている。誰もがデータを公開・管理・利用できる、触媒研究のデータセンターを目指している。

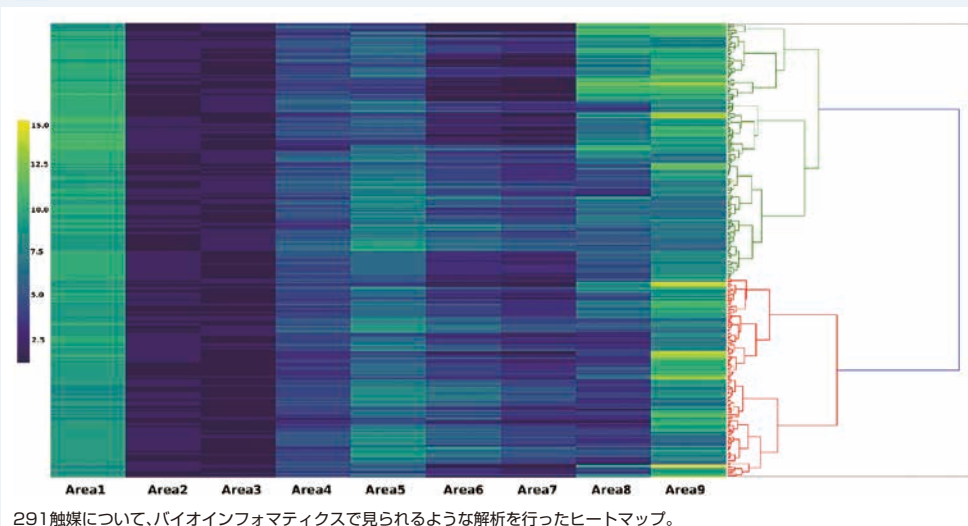
「遺伝子」も設計に有用 自然言語処理などを駆使

続いて高橋さんらは「触媒遺伝子」というユニークな成果を生み出した。生物の遺伝子のように、触媒の機能を規定する因子を指す概念だ。「従来の元素記号による表記では、組成が異なる

触媒同士は、同じ機能を持っていますが、全くの別物として扱われてきました。触媒遺伝子は、同じ機能を持つ触媒は同じ遺伝子を持つとします。この特有の遺伝子を塩基配列のような形で表現することができれば、触媒開発に応用できると考えました」と話す。

例えば、OCM触媒のFe-Nd-Tb-

図6 高橋さんが提案した触媒遺伝子



291触媒について、バイオインフォマティクスで見られるような解析を行ったヒートマップ。

Reprinted (adapted) with permission from Journal of Physical Chemistry Letters. Copyright 2022 American Chemical Society.

