

NII 市民講座 「未来へつながる情報学」

## 化学と情報学

未来の創薬などに結びつく化学情報の体系化とは？

2008年11月6日

国立情報学研究所 情報学プリンシップ研究系

佐藤 寛子

分子と反応の広大な世界をイメージする

化学と情報学

化学情報学とは

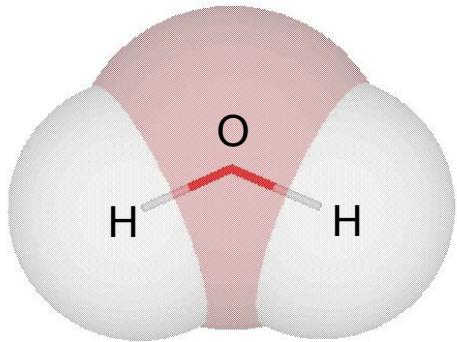
化学反応の予測への取り組み

# 化学物質の種類

この世界に存在する化学物質

この世界に存在する可能性のある化学物質

# 化学物質のスケール



水の分子  $\text{H}_2\text{O}$

1分子の大きさ

“ナノ”との比較

# 化学物質の量

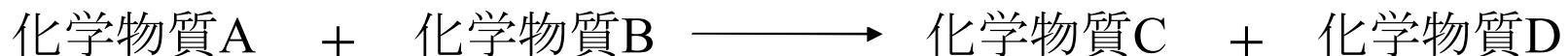
コップ 1 杯 (180 ml) の水

もし一直線に並べられたら...

1秒間に1分子ずつ数えたら...

# 化学反応

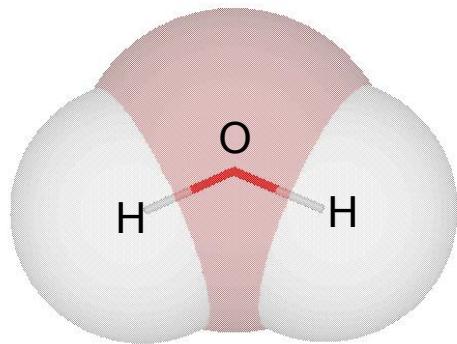
結合が切れて、新しい結合が作られる



切れる結合の可能性  
(1本,2本,...)

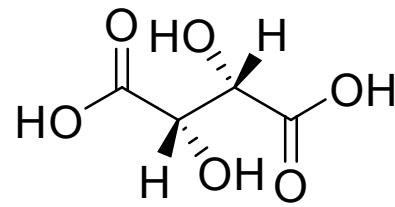
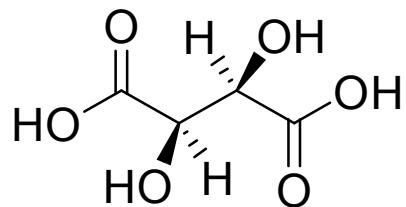
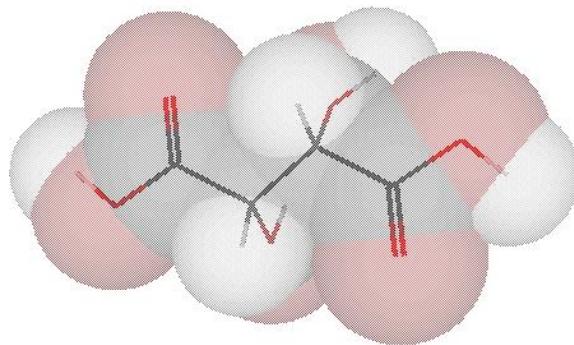
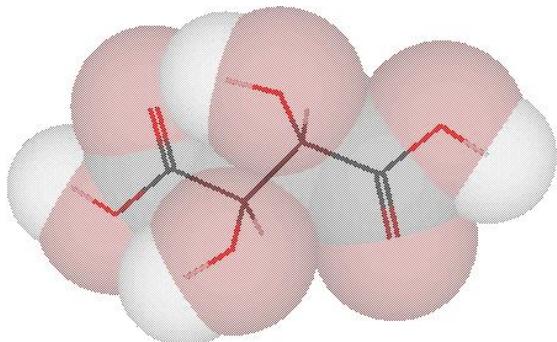
作られる結合の可能性  
(1本,2本,...)

# 分子はよく動く



分子全体が縦横無尽に動く  
結合が伸び縮みする  
2本の結合の角度が変わる  
結合が回転する

# 分子は立体的



結合する位置で構造が変わる

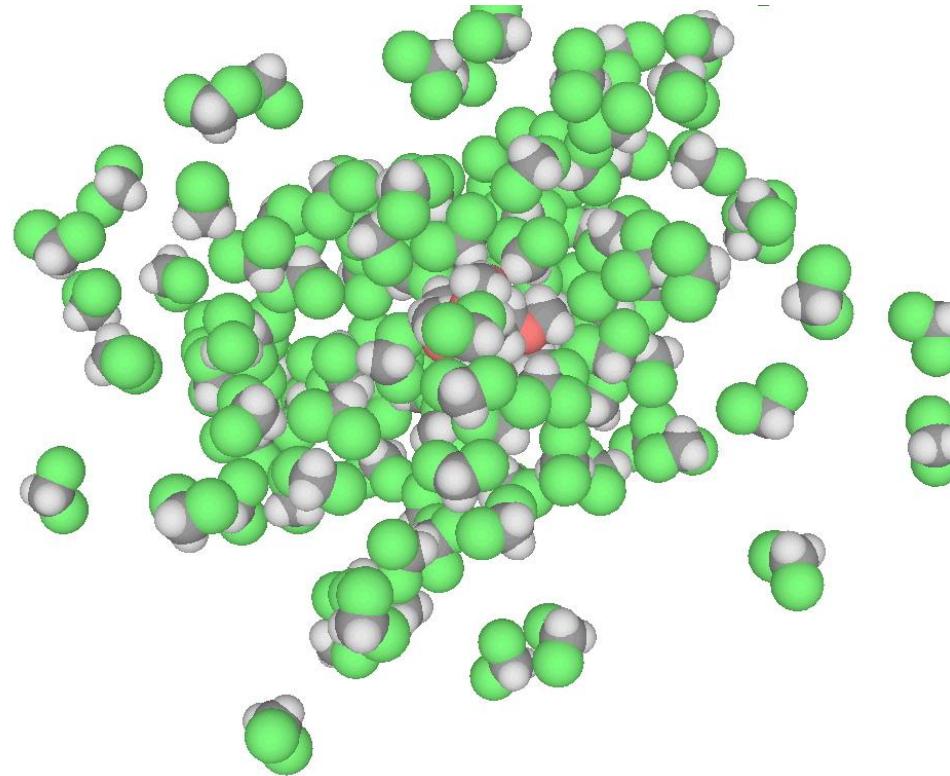
結合が回転したりすると構造が変わる

# 分子の間には力が働く

ほどよい距離だと引き合う  
近づきすぎると反発する  
離れると疎遠になる

力の起源によって性質が異なる

# 周りの環境に影響される



温度

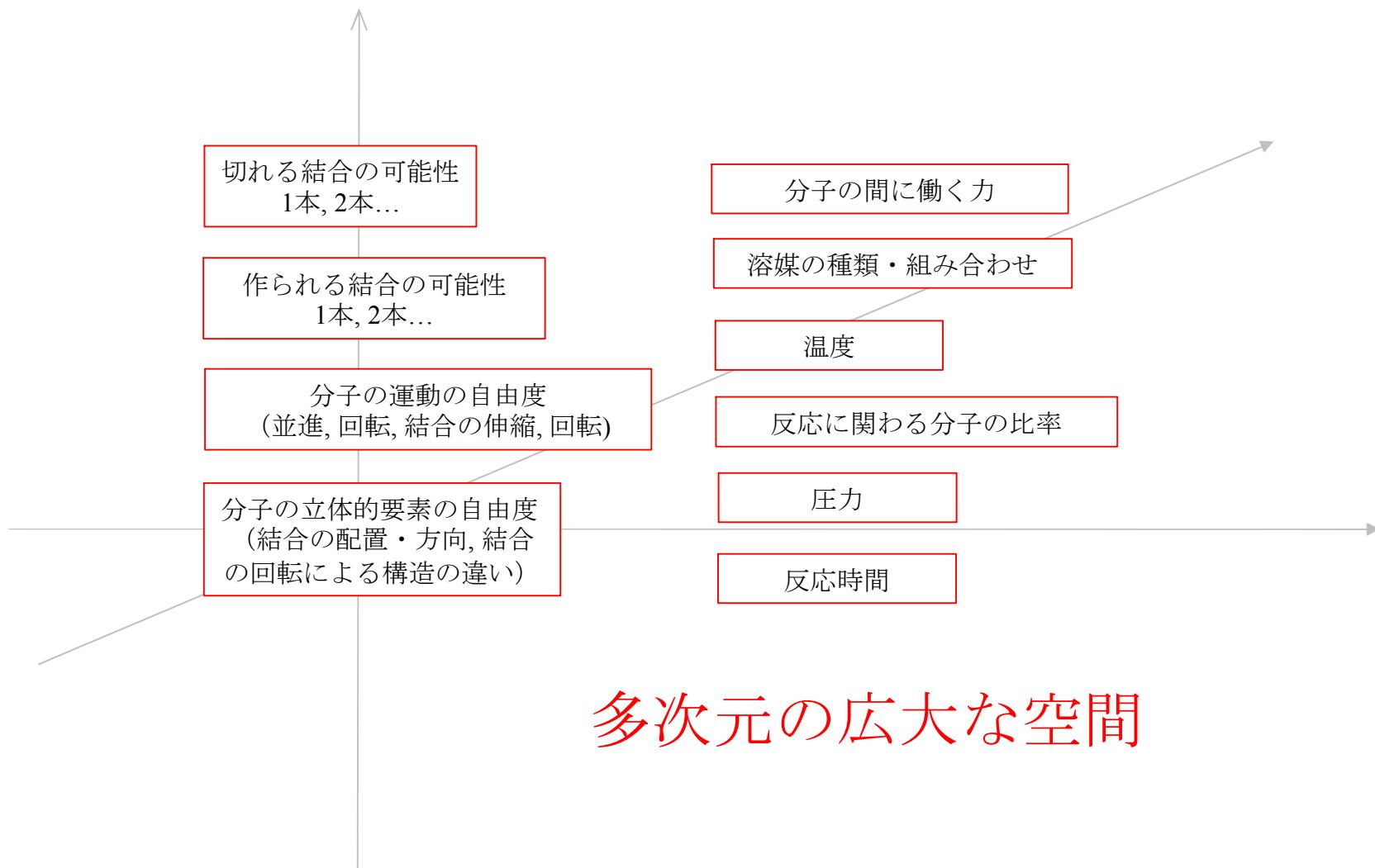
溶媒

反応に関わる分子の比率

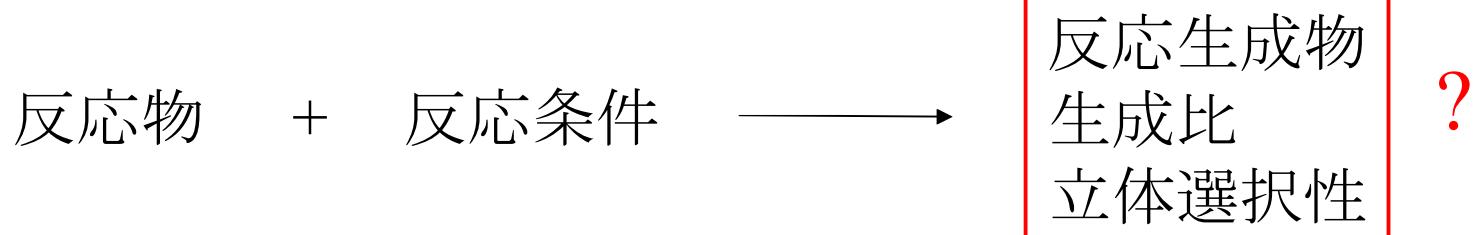
反応時間

圧力

# 化学反応の可能性の広がり



# 化学反応の予測



多次元の広大な化学反応の空間のなかで  
現実的な速さで  
現実的な正解に辿りつくこと

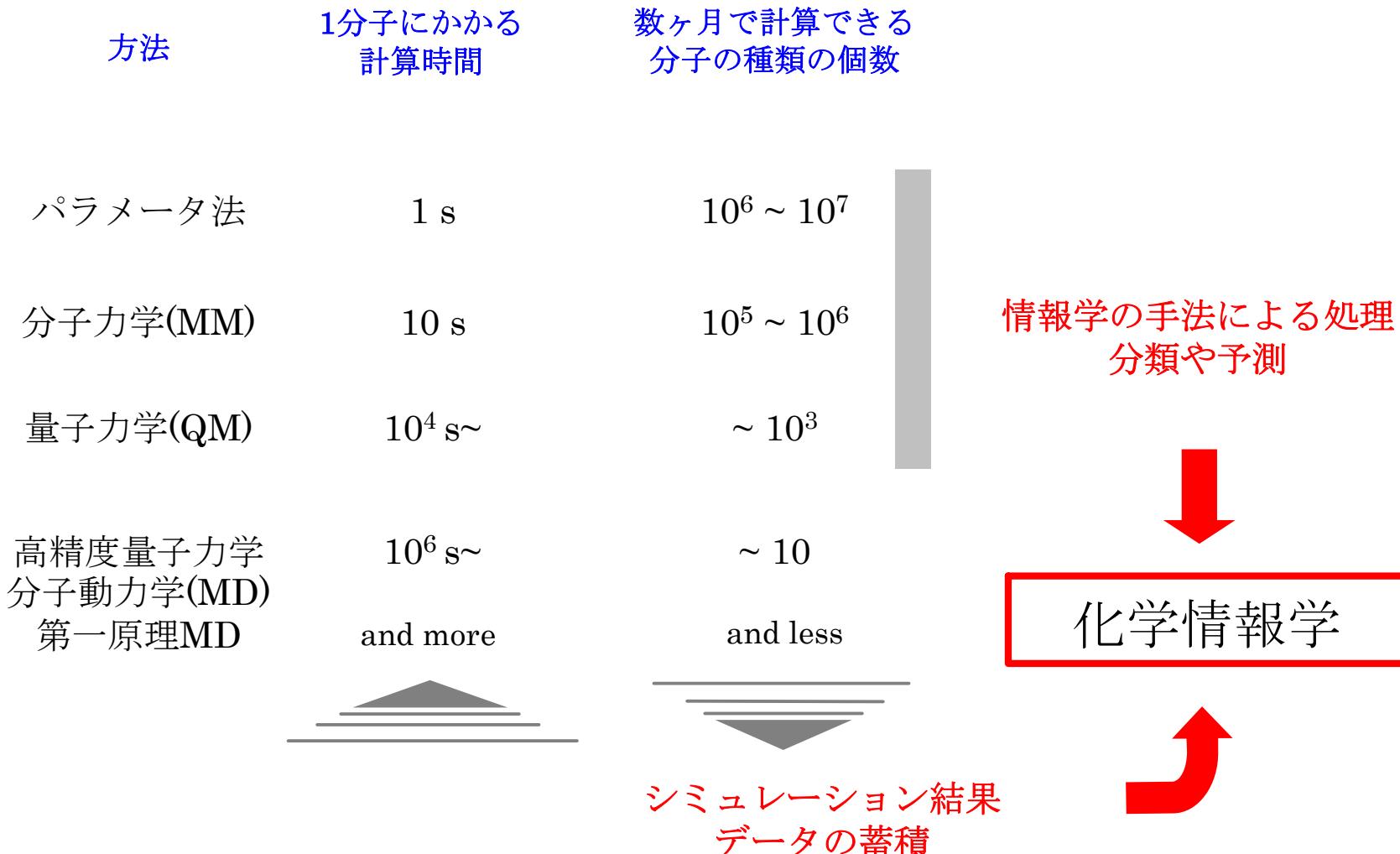
# 正解への道

夢の特効薬は “ない”

砂漠の中から一粒の宝物を探すための  
さまざまな努力

ex.) 反応経路の自動探索 GRRM (東北大、大野ら)

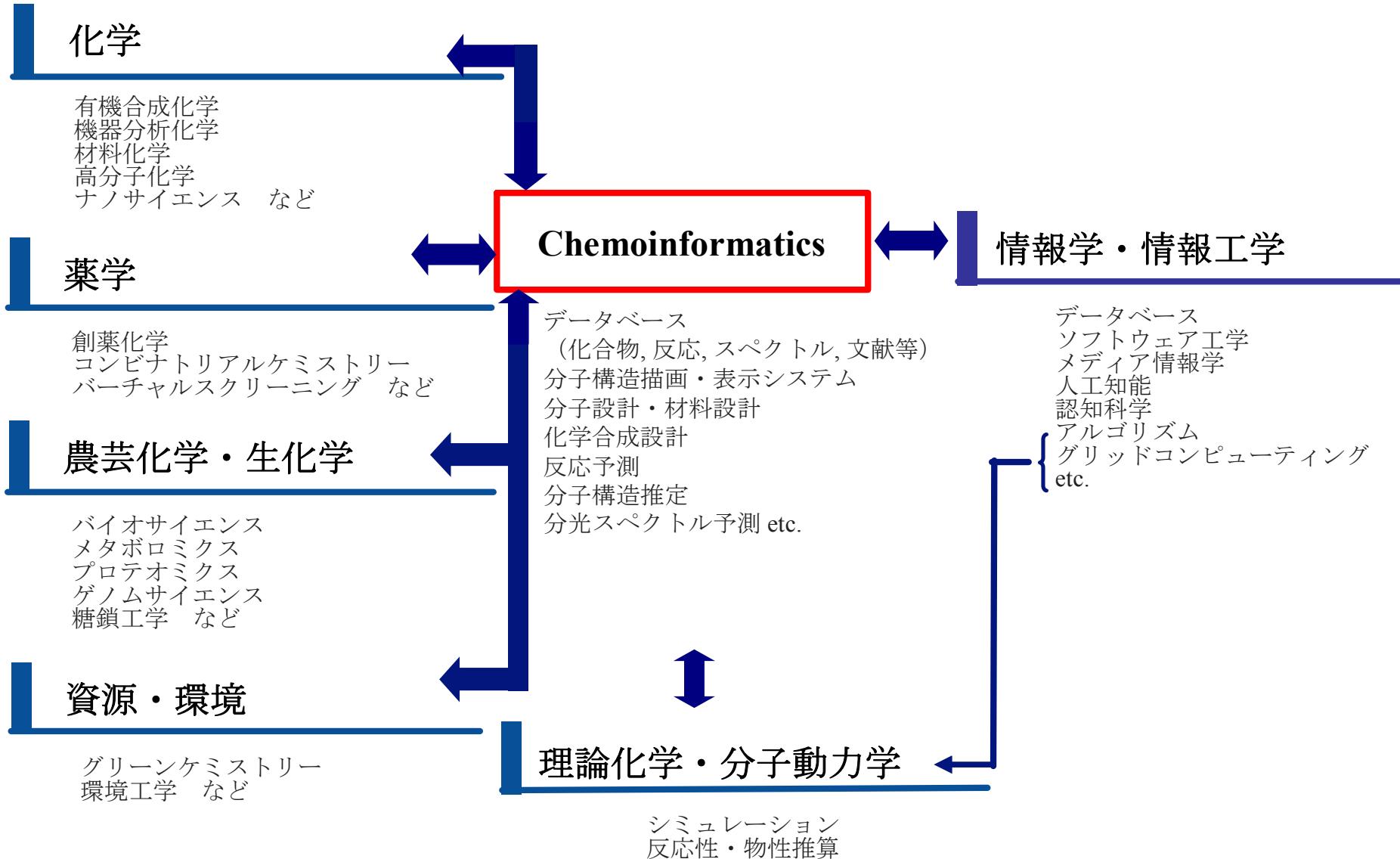
# コンピュータで分子を処理する



化学情報学とは

化学の問題を解決するための応用情報学

# 他の学問分野とのつながり



# 社会・産業とのつながり

## 化学

工業用化学製品、医薬品、化粧品、プラスチック製品、合成ゴム、化学繊維、農薬、塗料、接着剤など

## 薬学

医薬品、化粧品、農薬など

## 農芸化学・生化学

バイオ  
食品、食品添加物、香料、色素、栄養素、ビタミンなど

## 資源・環境

石油、重油、軽油、ガソリン、石炭、水素燃料、ダイオキシン、残留農薬、環境ホルモンなど

## Chemoinformatics

データベース  
(化合物、反応、スペクトル、文献等)  
分子構造描画・表示システム  
分子設計・材料設計  
化学合成設計  
反応予測  
分子構造推定  
分光スペクトル予測 etc.

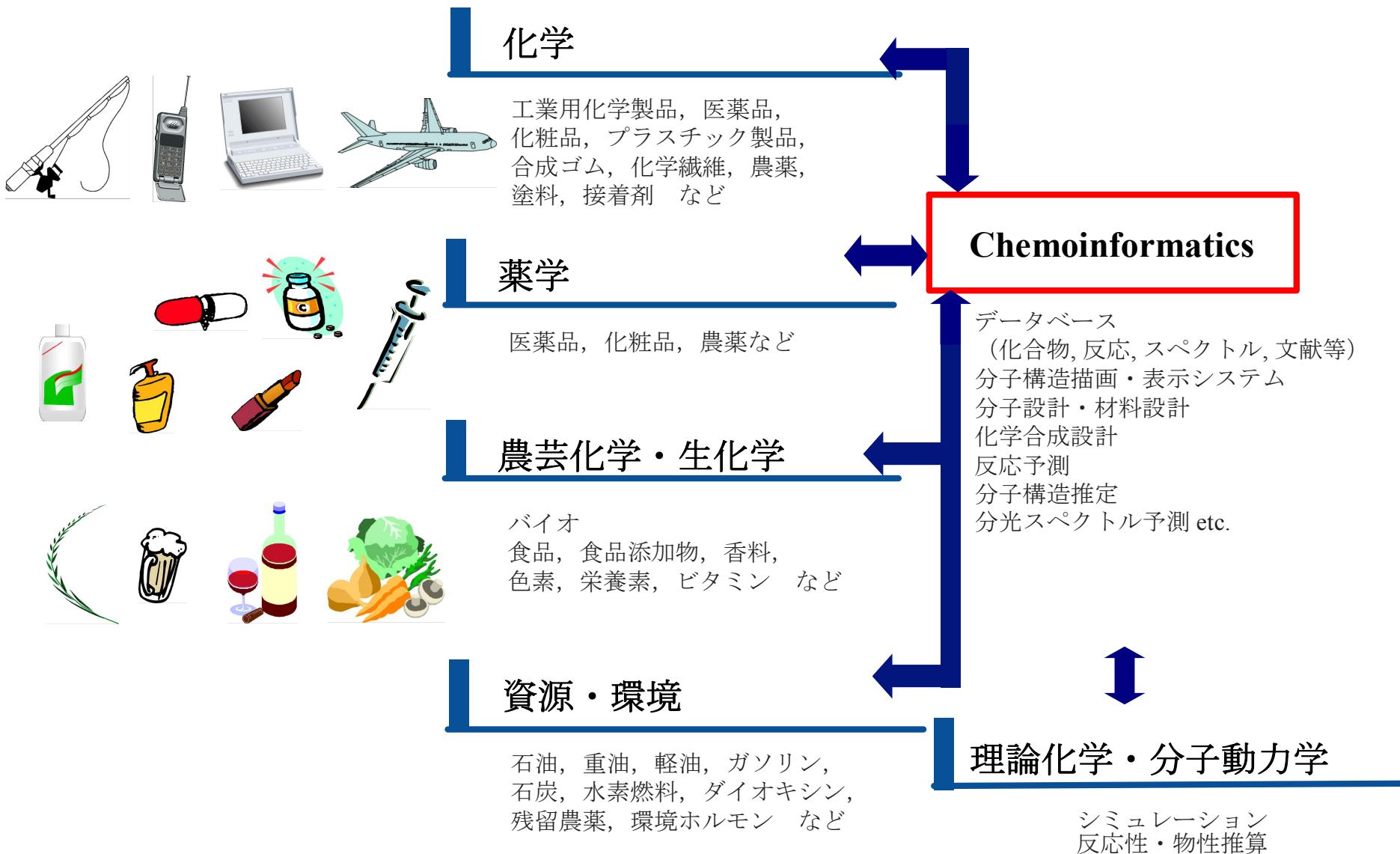
## 情報学・情報工学

データベース  
ソフトウェア工学  
メディア情報学  
人工知能  
認知科学  
アルゴリズム  
グリッドコンピューティング  
etc.

## 理論化学・分子動力学

シミュレーション  
反応性・物性推算

# 社会・産業とのつながり



# 化学の歴史と情報学

## 周期表の発見

ラボアジェ(1743-1794)の元素論

—どんな手段でも分解できない成分物質が元素である

ドルトン(1766-1844)の原子論

—原子の性質はその相対的質量によって決まる。

物質を構成する最小単位：原子(atom)

メンデレーエフによる**周期表**の発表(1869年)

—元素を原子量順に並べるとある周期で類似元素が現れる

当時未発見の元素を空欄として、その性質を予測

# メンデレーエフの周期表

|    | I  | II | III | IV | V  | VI | VII | VIII       |
|----|----|----|-----|----|----|----|-----|------------|
| 1  | H  |    |     |    |    |    |     |            |
| 2  | Li | Be | B   | C  | N  | O  | F   |            |
| 3  | Na | Mg | Al  | Si | P  | S  | Cl  |            |
| 4  | K  | Ca | ?   | Ti | V  | Cr | Mn  | Fe, Co, Ni |
| 5  | Gu | Zn | ?   | ?  | As | Se | Br  | ?          |
| 6  | Rb | Sr | Y   | Zr | Nb | Mo | ?   | Ru, Rh, Pd |
| 7  | Ag | Cd | In  | Sn | Sb | Te | I   |            |
| 8  | Cs | Ba | Dy  | Ce |    |    |     |            |
| 9  |    |    |     |    |    |    |     |            |
| 10 |    |    | Er  | La | Ta | W  |     | Os, Ir, Pt |
| 11 | Au | Hg | Tl  | Pb | Bi |    |     |            |
| 12 |    |    |     | Th | U  |    |     |            |

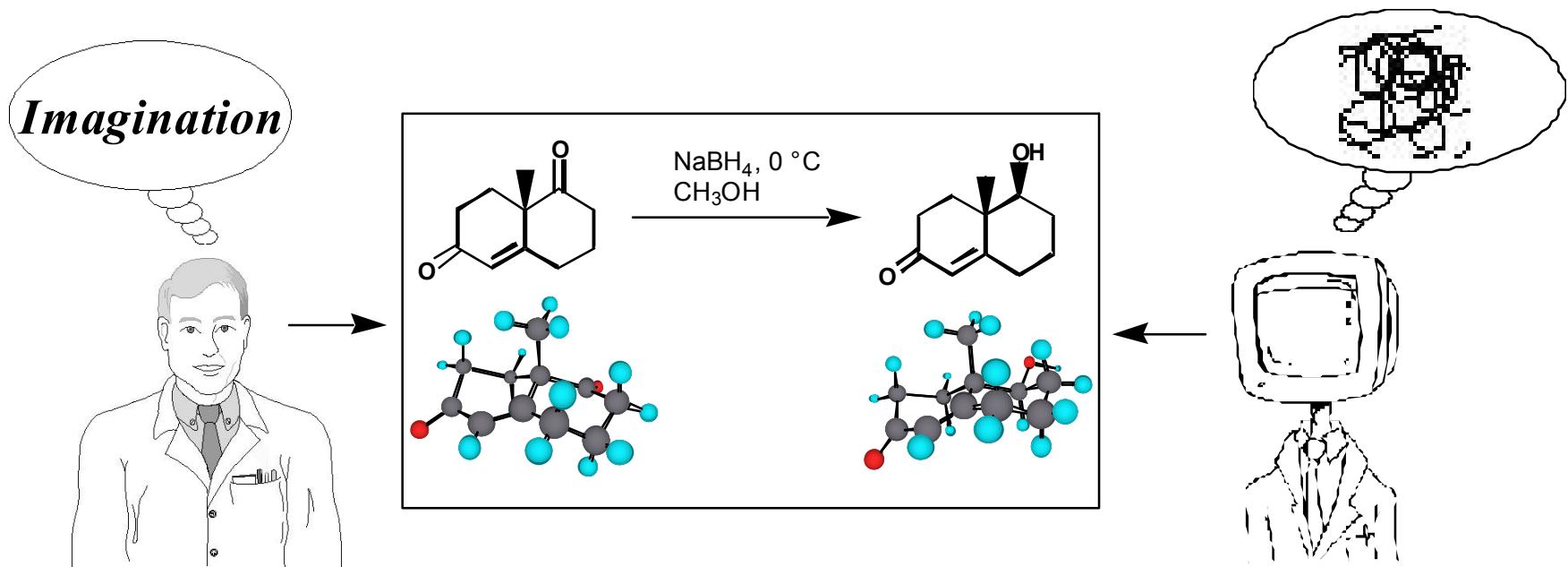
「マルチメディアで見る原子・分子の世界」(Copyright: JST)より

# 化学者と情報学

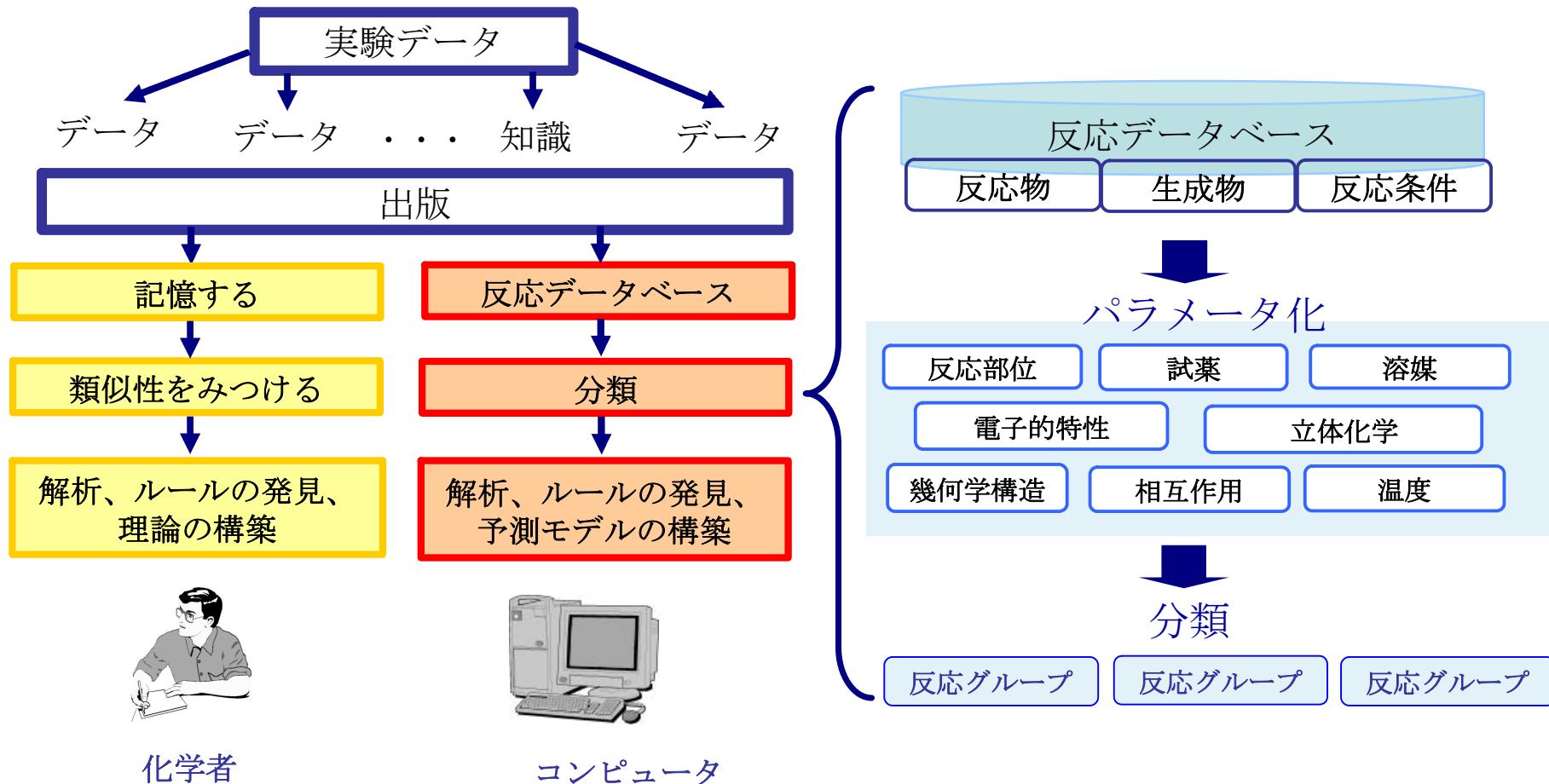


化合物データベース  
化学反応データベース  
スペクトルデータベース  
結晶構造データベース  
高分子データベース  
文献情報データベース  
など

# コンピュータのための反応表現

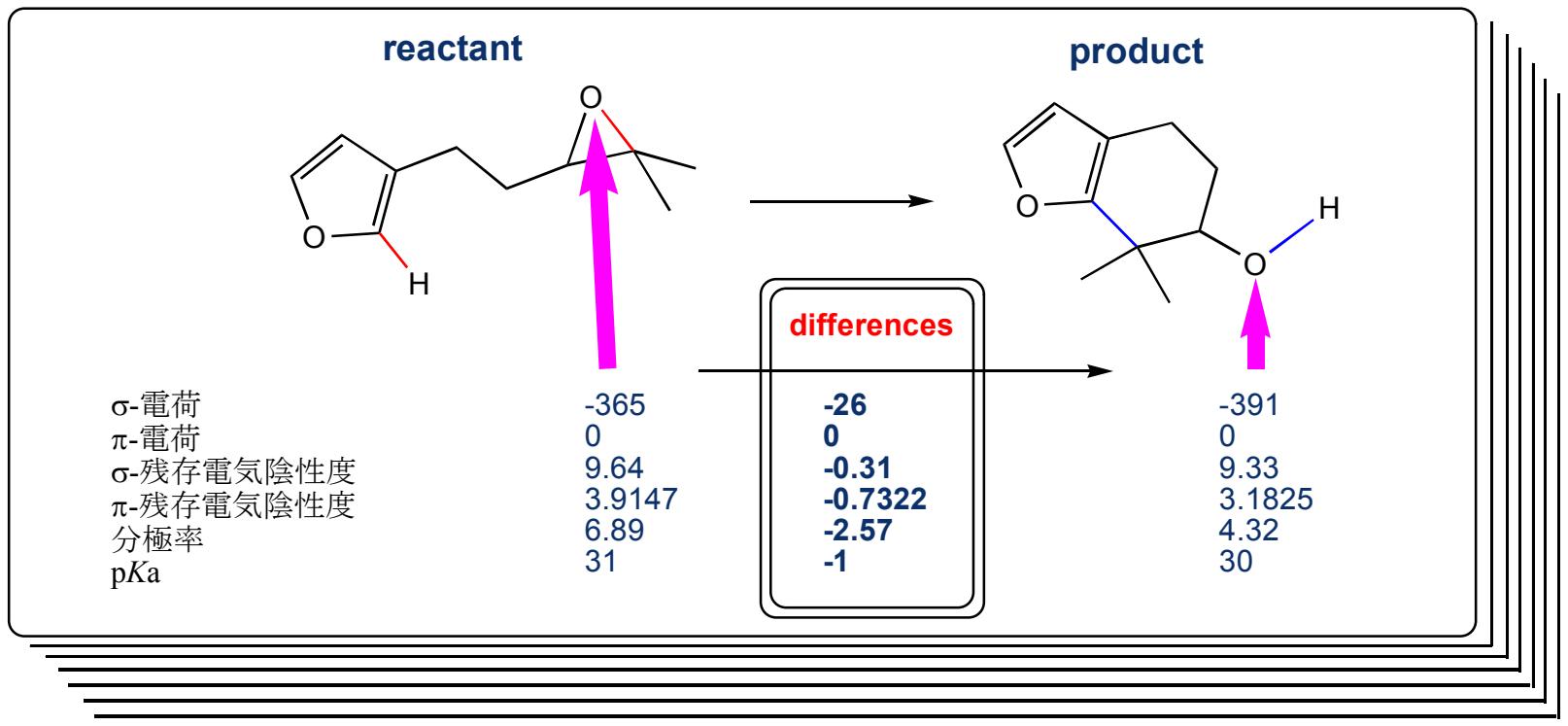


# 化学データを利用した反応の予測



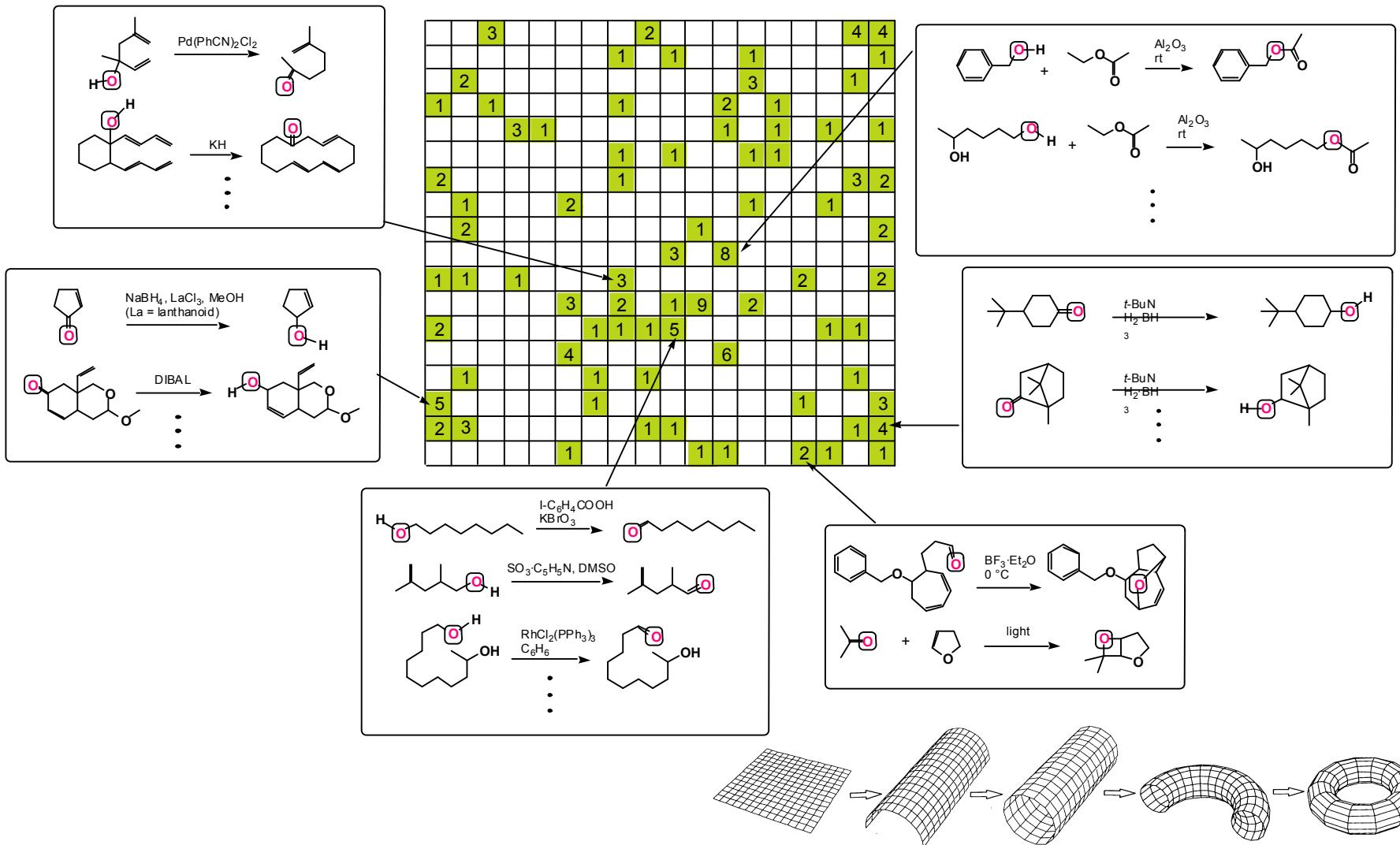
# 反応を自動的に分類する

データセット：131 反応

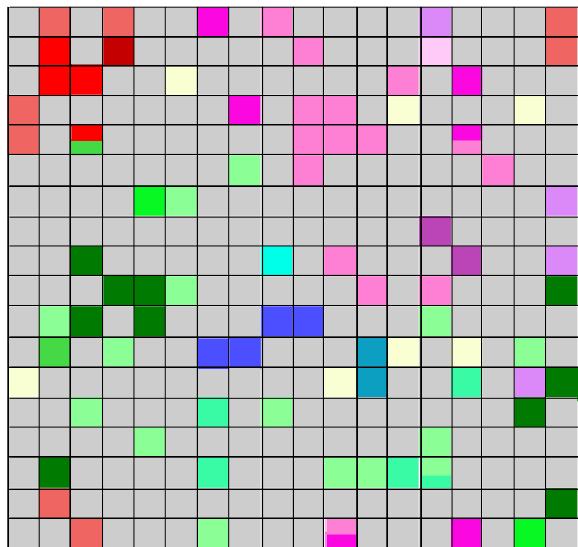


• • . ( — 切れた結合  
— 作られた結合

# 6種類のパラメータの似た反応が近くに集まる

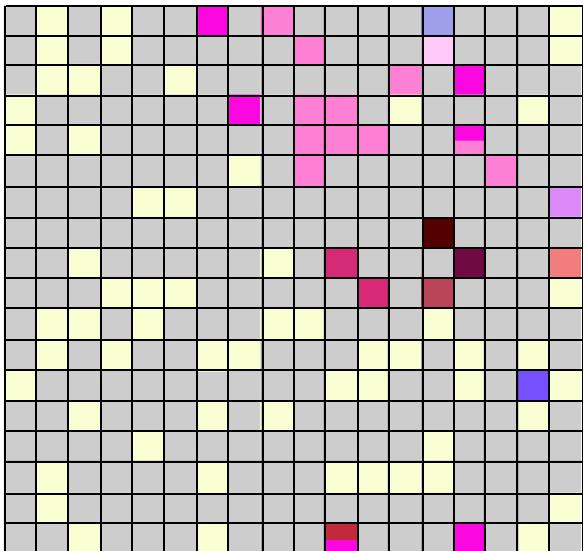


# 反応の種類で色分けする



- : reductions
- : reductive alkylations
- : aldol reactions
- : cleavage of cyclic ethers
- : cleavage of epoxides
- : cleavage of lactones
- : cleavage of ethers
- : cleavage of ester
- : oxdation of alcohols
- : oxy-Cope rearrangements
- : oxdative cleavage of grycohols
- : formation of cyclic ethers
- : formation of epoxides
- : formation of lactones
- : formation of ethers
- : formation of esters

# もっと細かく色分けする (ピンク色系の領域)



■ : cleavage of cyclic ethers



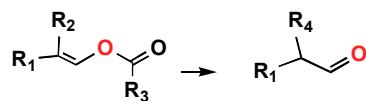
■ : cleavage of epoxides



■ : cleavage of lactones



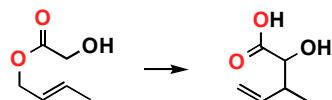
■ : cleavage of esters



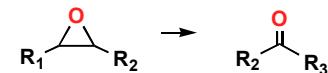
■ : cleavage of esters with halogens



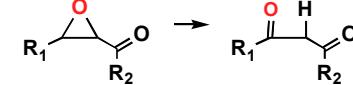
■ : Claisen rearrangement



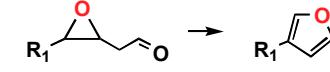
■ : rearrangement of epoxides



■ : formation of diketones by Pd catalyzed reaction of *a, b*-epoxyketones



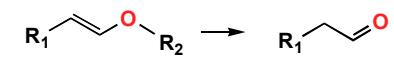
■ : formation of furans by cleavage of epoxides



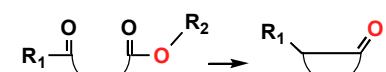
■ : oxidation of aromatics



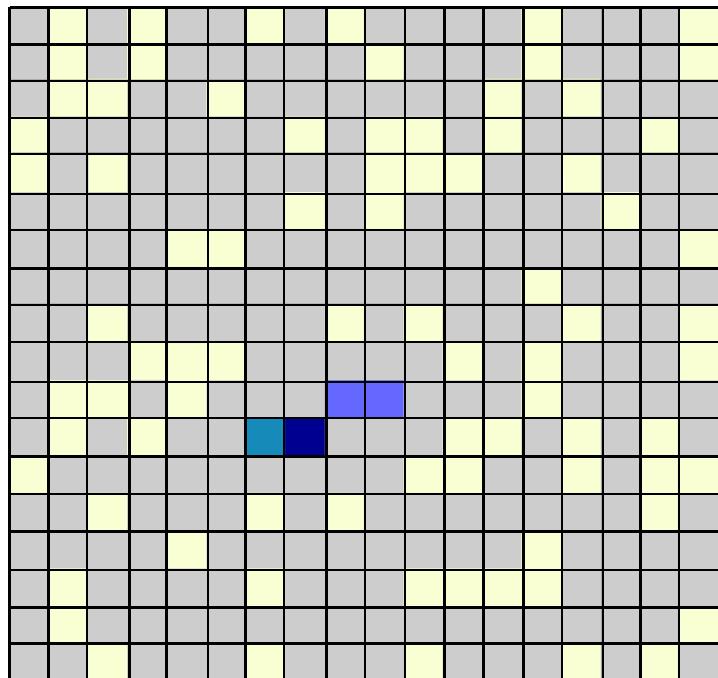
■ : hydrolysis of enol ethers



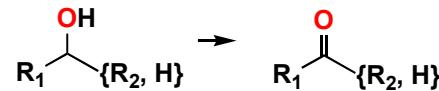
■ : McMurry coupling



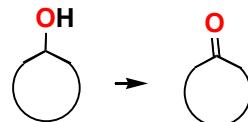
もっと細かく色分けする（青色系の領域）



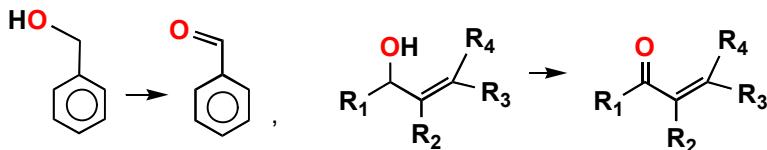
 : oxydation of primary alcohols and acyclic secondary alcohols



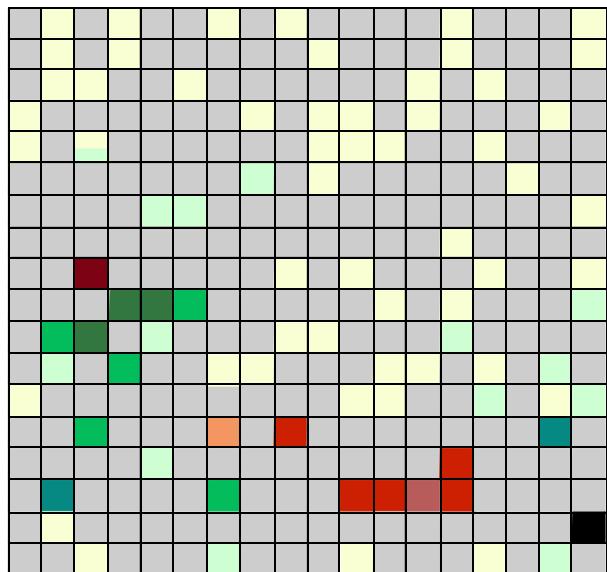
: oxydation of cyclic secondary alcohols



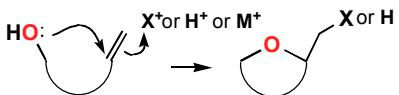
 : oxydation of benzyl alcohols and aryl alcohols



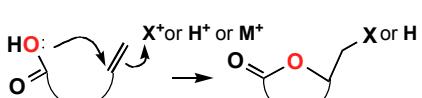
# もっと細かく色分けする (緑色系の領域)



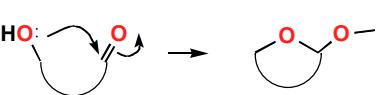
■ : haloetherification type reactions



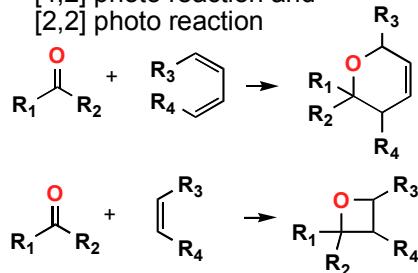
■ : halolactonization type reactions



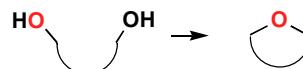
■ : intramolecular acetalizations



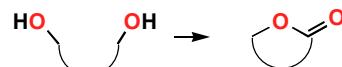
■ : [4,2] photo reaction and [2,2] photo reaction



■ : cyclic etherification of diols



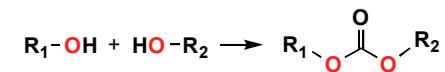
■ : oxidative lactonization of diols



■ : acetylation of alcohols



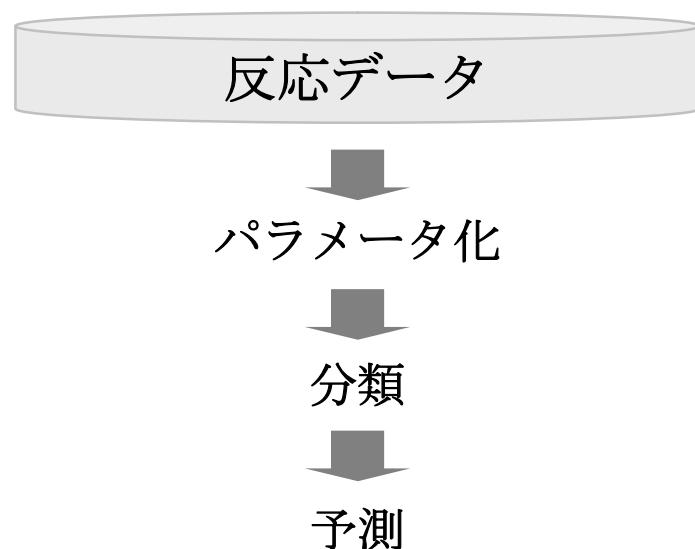
■ : formation of dialkyl carbonates from alcohols



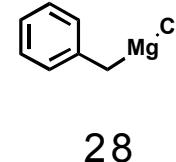
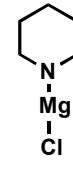
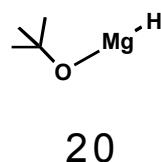
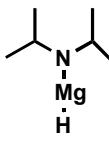
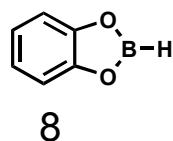
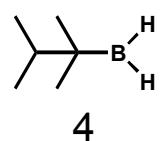
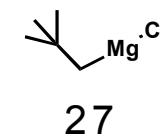
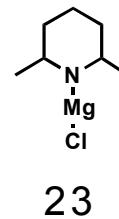
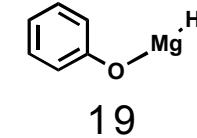
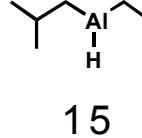
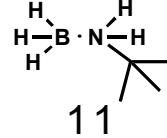
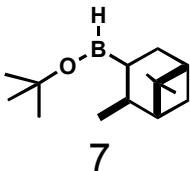
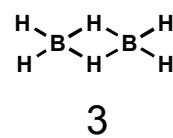
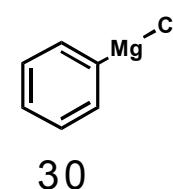
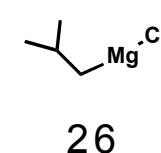
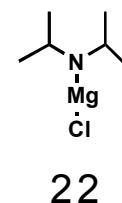
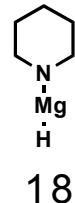
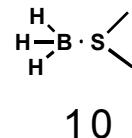
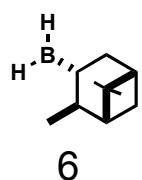
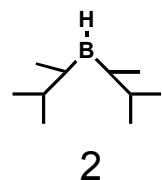
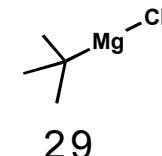
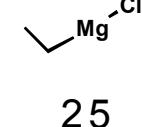
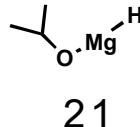
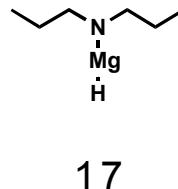
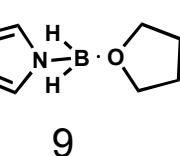
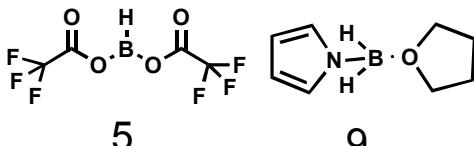
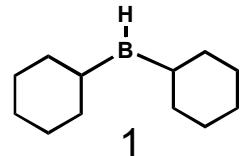
■ : others

# 分類から予測へ

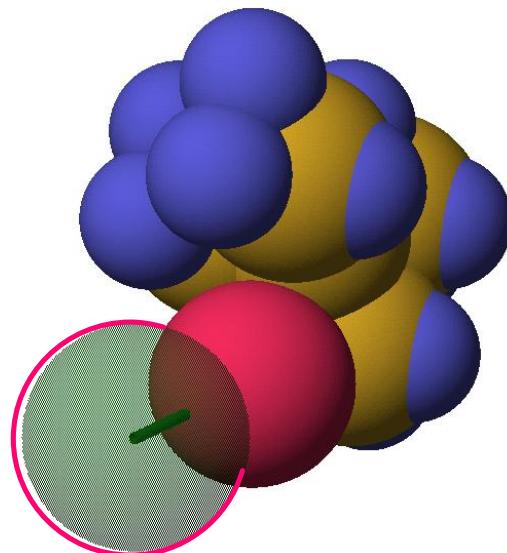
化学反応の試薬の機能を予測する



# 分類した試薬分子

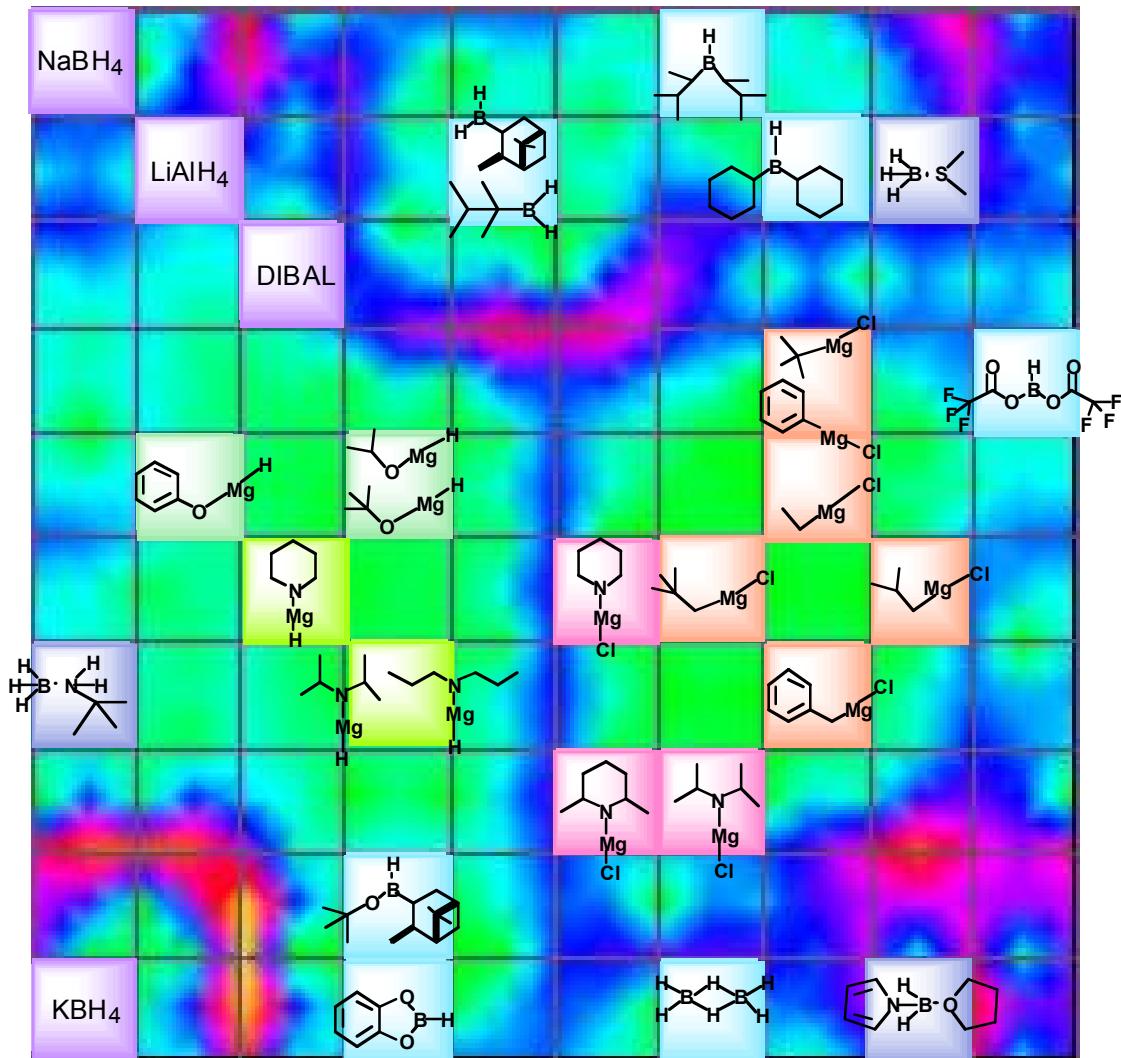


# 分子のパラメータ化

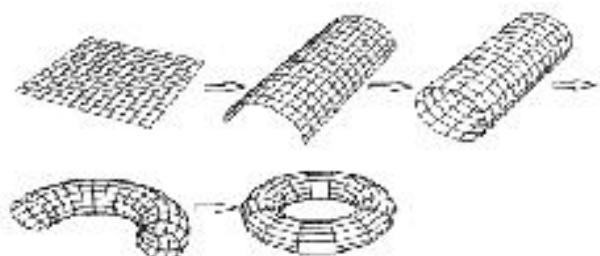


- 静電的な相互作用に基づくパラメータ
- 立体的な相互作用に基づくパラメータ
- 表面積に基づくパラメータ

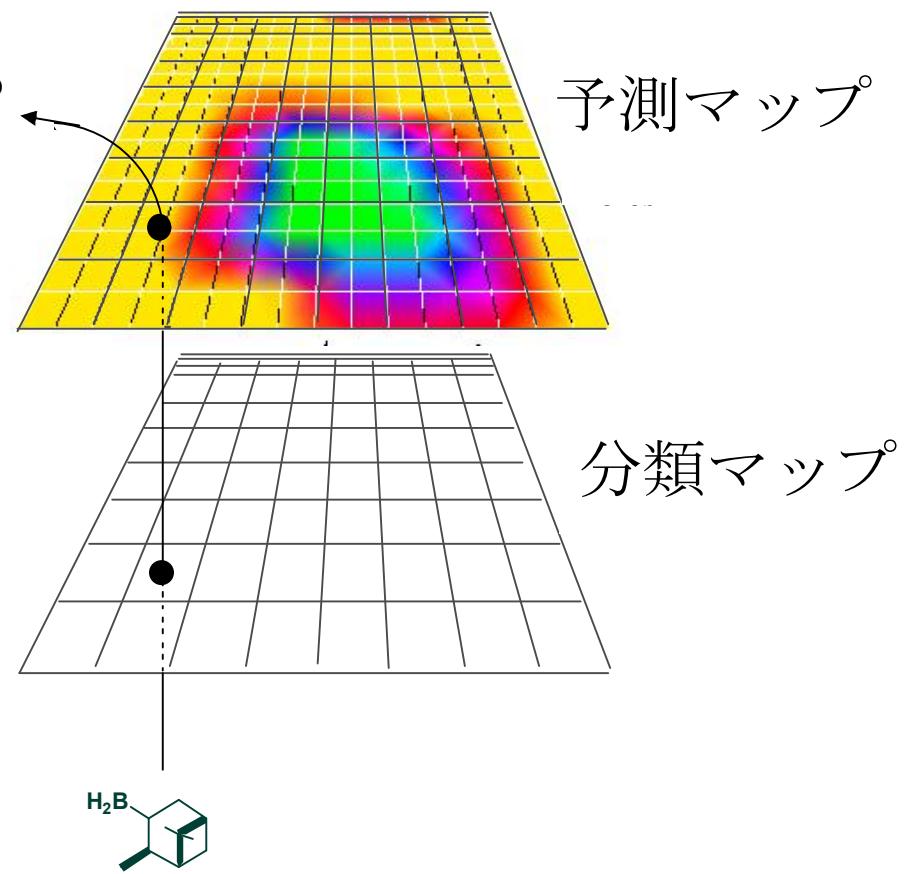
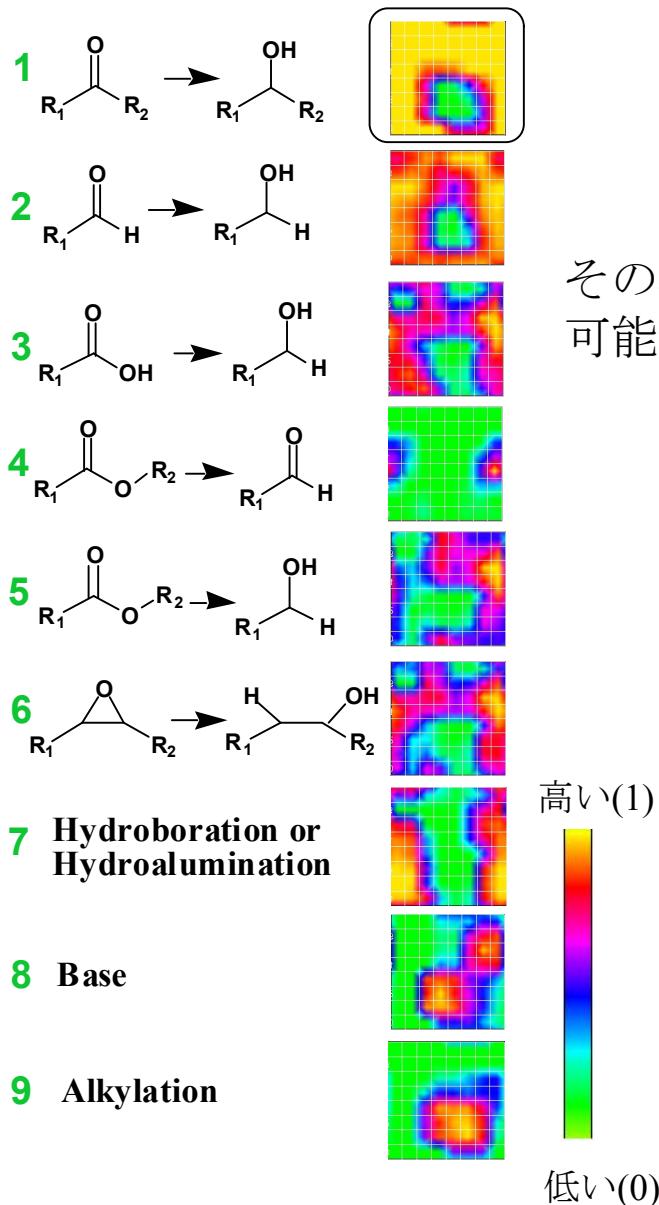
# 分類の結果



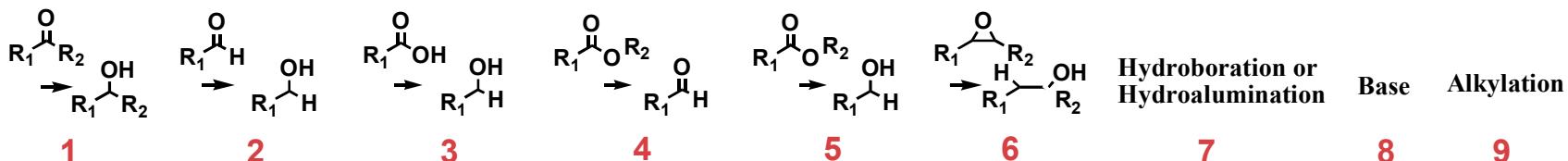
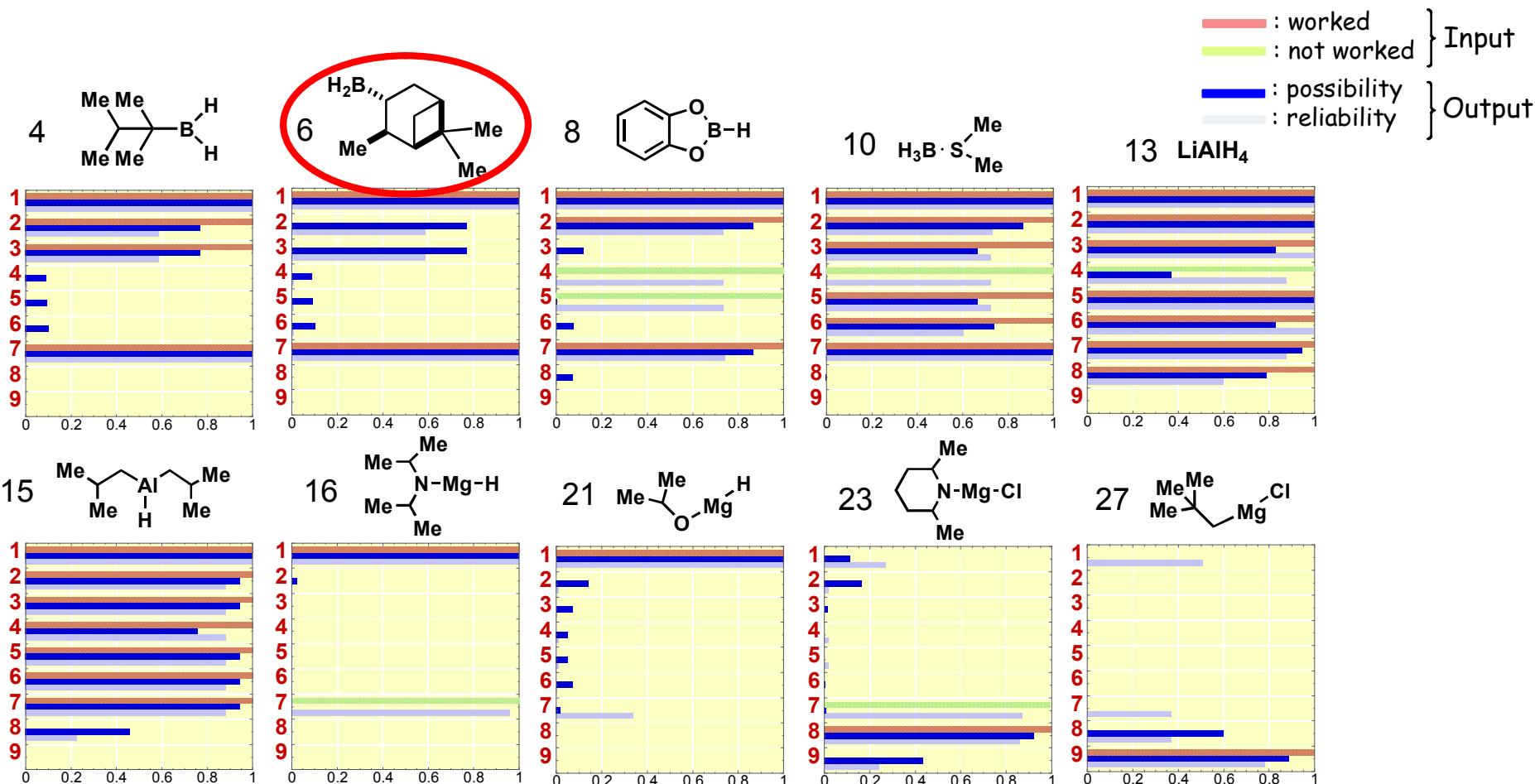
- █ boranes
- █ borane complexes
- █ hydrides
- █ reducing agents containing N atom
- █ reducing agents containing O atom
- █ bases
- █ Grignard reagents



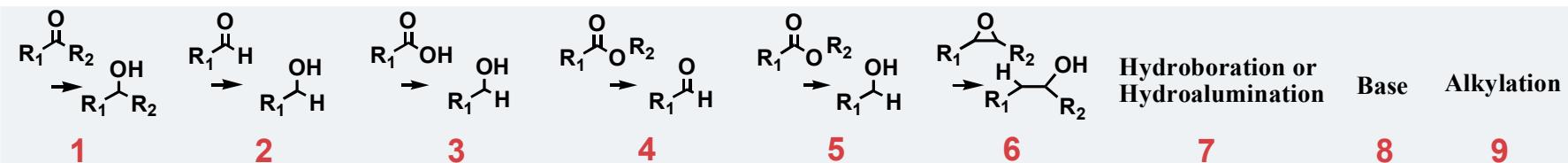
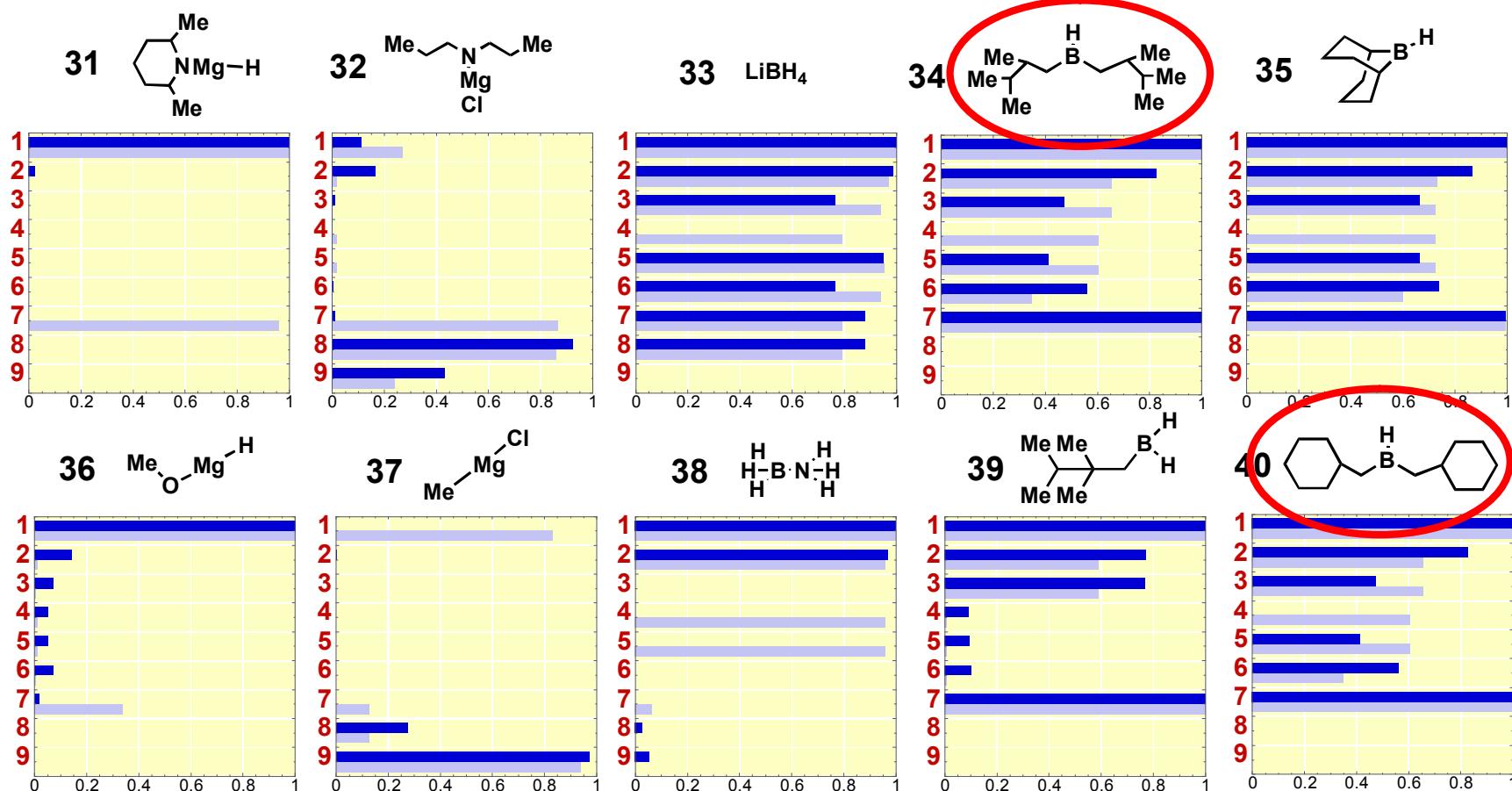
# 予測モデルをつくる



# 予測結果 — 分類に用いた分子

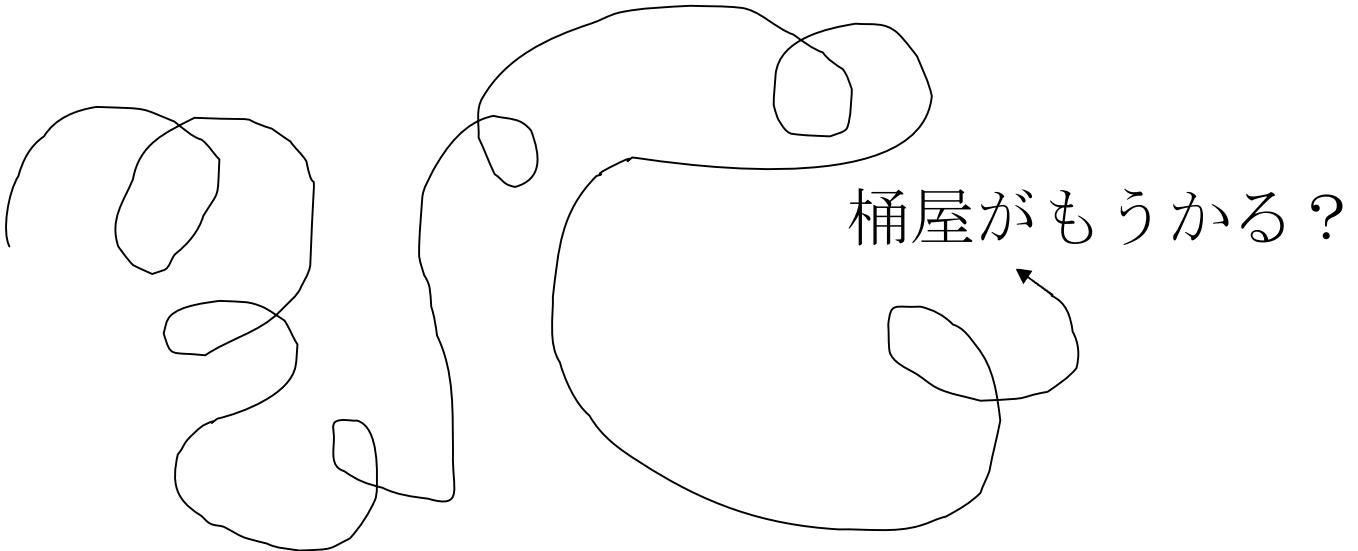


# 予測結果



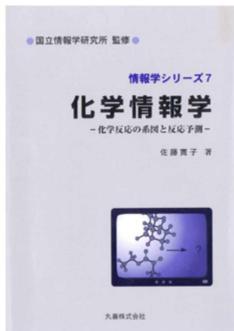
# 最後に：基礎研究と風と桶屋

風が吹けば...



- 必ず桶屋が儲かるわけではない
- 桶屋が儲かる方法は1つではない
- 風が吹くと様々なことが起こりうる
- 桶屋をもうけさせるために風を吹かせるわけではない
- 詳しい過程がわからないと因果関係は理解できない

# もっと勉強するなら



化学情報学—化学反応の系図と反応予測—  
佐藤 寛子 著  
丸善株式会社 293 ページ, 2003  
ISBN 4621071971



ケモインフォマティクス 予測と設計のための化学情報学  
J. Gasteiger, T. Engel 編  
船津公人, 佐藤寛子, 増井秀行 訳  
丸善株式会社 638 ページ, 2005  
ISBN 4621075527

“Chemoinformatics A Textbook” J. Gasteiger, T.Engel, Wiley-VCH, 2003, ISBN 3 527 30681 1

“An Introduction to Chemoinformatics” A. R. Leach, V. J. Gillet, Kluwer Academic Publishers, 2003,  
ISBN 1 4020 1347 7

“Chemoinformatics Concepts, Methods, and Tools for Drug Discovery” J. Bajorath Humana Press, 2004,  
ISBN 1 58829 261 4