AN INVITATION TO COMPUTATIONAL CHEMISTRY

コンピュータのなかの実験室

計算化学への招待



千葉工業大学 山本 典史　著

コンピュータのなかの実験室

計算化学への招待

山本 典史　著

Version 0.1.5

テオ 兄

称号「真理の織り手」
Theo：Theoretical（理論的）から

「何を計算させ、答えが正しいかを判断できるのは、化学が分かる人だけだ。」

ひとつの分子の中の
電子を、精密に計算する。
案内：第 2 章・第 4 章



足もとの波打つ床は、分子のエネルギーの地形。



水素原子の電子の雲
（第 2 章）

① **Ψ の杖**
電子の状態を表す記号 Ψ（プサイ）。量子化学の式の中心にあります。

② **片眼鏡**
ひとつの分子を、細部まで見つめます。

③ **服の飾り**
電子の雲がとる形（原子軌道）をかたどっています。

④ **フクロウの賢者**
少数精鋭の精霊。複雑な計算を 1 つずつ確実に。CPU のイメージです。

テラ 妹

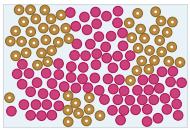
称号「流転の予言者」
Thera：Thermal（熱的）から

「コンピュータの中にも『実験室』があるんだよ！」

たくさんの分子の動きと、データを追う。
案内：第 1 章・第 3 章・第 5 章



低い谷ほど安定な形。計算化学は、この谷や峠を探ります。



水と油のつぶ（第 1 章）

① **らせんの縦ロール**
タンパク質の中のらせんをかたどった髪です。

② **砂時計**
時間とともに動いていく、分子の世界。

③ **光の鳥の大群**
数千、数万が同時にはたらく精霊。GPU のイメージです（実際の計算では、どちらの分野でも両方を使います）。

④ **透明な箱の分子**
コンピュータの中で、分子を見ます。

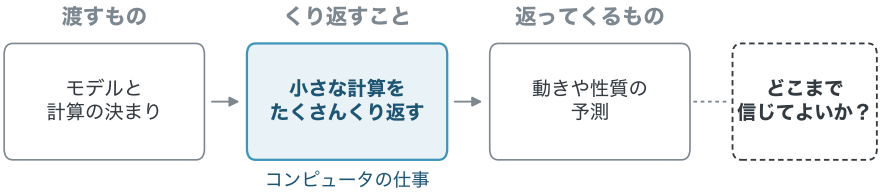
はじめに

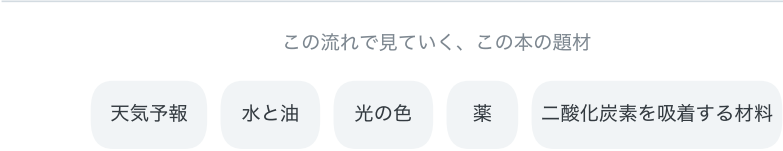


「化学の研究」と聞いて、どんな姿を思い浮かべるでしょうか。白衣を着て、フラスコに薬品を入れて振り、色が変わるのを見つめる人かもしれません。

けれども、化学の研究室には、一日の多くをパソコンの前で過ごす人もいます。画面の中で分子を動かし、分子の中の電子を計算し、まだ作っていない分子の性質を予測しています。この本は、こうした計算化学という分野への招待です。高校や大学で化学を学んでいる方や、計算化学が気になる方に向けて書きました。

この本で見ていくのは、コンピュータが何をしているかです。コンピュータに何を渡し、何をくり返させ、何が返ってくるのか。その結果は、どこまで信じてよいのか。天気予報、水と油、光の色、薬、二酸化炭素を吸着する材料を例に確かめます。





コンピュータに渡すもの（左）、くり返すこと（中）、返ってくるもの（右）。青い箱が、コンピュータの仕事です。

この本で分かること

- 計算化学とは何かを、「モデル化とシミュレーション」の言葉で説明できる

- 計算化学の 4 つのはたらき（見る・探す・予測する・学ぶ）について、コンピュータが何をくり返しているかを、例を挙げて言える
- 分子を「つぶとばね」で動かす見方と、電子から色や電荷のかたよりを求める見方のちがいを言える
- 計算で分子の性質を予測した例を 1 つ、何をモデルにしたか・何をくり返したか・どこまで信じられるか、の 3 つの問いで説明できる
- 計算の結果をうのみにせず、実験の値と比べて、外れた理由を考えられる
- 高校で学んだことが、計算化学のどこにつながっているかを言える

この本の読み方

序章で計算化学がどんな分野かを紹介し、第 1〜5 章で 4 つのはたらきを 1 つずつ取り上げます。第 6 章で計算の答え合わせをし、終章で今後の計算化学を展望します。

各章は、同じ形で進みます。章の冒頭の「この章で分かること」の箱に、その章のねらいと、つながりのある高校の単元をまとめました。本文では、まず本物の実験や観察を紹介し、次に同じことをコンピュータの中でどう行うかを説明します。章は「まとめ」で終わります。

本文のところどころにある「予想してみよう」の問いは、先を読む前に少しだけ考えてみてください。答えは、すぐ後の図と本文にあります。

「発展」の箱には、高校の範囲を超える話を入れました。読み飛ばしても、本文の流れは追えます。ほかに、まちがえやすい考え方を取り上げる「よくある誤解」と、研究での使われ方を紹介する「研究の現場から」の箱があります。案内役のテオとテラは、次の節で紹介します。

案内役のテオとテラ

この本の案内役は、「双子の計算術士」の兄テオと妹テラです。計算化学の 2 つの方向を、双子の術士に見立てたキャラクターで、著者の研究室のマスコットです。2 人の姿と得意なことは、この本の最初の見開きで紹介しました。イラストは、科学系クリエイターの揚げ鶏々（あげどりどり）さんが描きました（gensogakusho.com）。

序章・第 6 章・終章は、2 人で案内します。章の扉の絵と章の色（テオは青、テラはローズ、2 人のときは緑）が、その章の案内役の目印です。2 人は、章の初めの

短い対話と、ページのいちばん下の「ひとこと」に登場します。ひとことは余談なので、気軽に読んでください。

本のウェブページ

この本のウェブページ（以下、本のページ）で、章ごとの動く図にさわれます。分子を動かしたり条件を変えたりして、本の図で見たことを自分で確かめられます。本文は、本のページを見なくても読めるように書きました。



コンピュータのなかの実験室 — 計算化学への招待

yamlab.jp/books/icc

この本を読むために

高校の化学基礎、物理基礎、生物基礎、数学 I、情報 I の内容が分かれば、読み進められます。水素結合や対数のように、高校 2 年生ではまだ学んでいないかもしれない内容は、使うところで説明します。

プログラムを書いたことがなくても、計算のソフトにさわったことがなくても読めます。コンピュータが何をしているかは、言葉と図で追っていきます。本文に出てくる式は高校で習う形のものだけで、それより進んだ式は、名前と、何を求める式なのかだけを紹介します。

目次



はじめに i
この本で分かること i
この本の読み方 ii
案内役のテオとテラ ii
本のウェブページ iii
この本を読むために iii
序章　コンピュータのなかの実験室 1
空を観測する、空を計算する 1
天気予報を 3 つの問いで読む 3
分子でも、同じことができる 6
分子の世界は、小さくて速い 7
ノーベル化学賞でたどる計算化学 9
計算化学の 4 つのはたらきと、この本の構成 10
1 分子の動きを「見る」 13
1.1 本物の実験：水と油、氷と水と水蒸気 14
1.2 分子をつぶとばねで表す 15
1.3 1 コマずつ動かす 17
1.4 温度は熱運動の激しさ 21
1.5 水と油はなぜ分かれるか 24
2 電子を「見る」 33
2.1 電子は雲のように広がる 34
2.2 電荷のかたより 40
2.3 光の色も、作る前に分かる 43
3 薬の候補を「探す」 51
3.1 インフルエンザウイルスが増えるしくみ 52
3.2 薬の候補を実験で試す 54

3.3 活性部位に結合する 2 つの条件 55
3.4 コンピュータの中で 1 つ試す：ドッキング 59
3.5 何億個をふるいにかける：バーチャルスクリーニング 61
3.6 計算から生まれた薬 63
4 体に届くかを「予測する」 67
4.1 水になじむ原子、油になじむ原子 68
4.2 細胞膜は油の膜 71
4.3 $\log P$：水と油への分かれ方を 1 つの数で 72
4.4 コンピュータの中で水と油に分ける 75
4.5 リレンザを飲み薬に書き換える 81
5 データから「学ぶ」 87
5.1 測ったデータを並べる 88
5.2 学習：予測と実測の誤差を小さくしていく 90
5.3 伏せておいた分子で答え合わせ 95
5.4 化学で使われはじめた AI 97
6 計算を信じてよいか 103
6.1 答え合わせ：計算と実測を並べる 103
6.2 誤差の理由：モデルが省いたもの 106
6.3 計算にかかる時間 110
6.4 計算と実験は両輪 114
6.5 AI を使いこなすのは化学が分かる人 115
終章　計算化学が切り拓く未来 121
穴で CO_2 をつかまえる 122
3 つの問いで読む 123
水で答えが変わる 125
4 つのはたらきをもう一度 126
計算化学を学ぶには 127
参考文献 131
用語の手引き 139
写真の出典 143

序章

コンピュータのなかの実験室



明日の天気は、どうやって予想しているのでしょうか。テレビやスマートフォンの天気予報は、「明日の午後は雨」「3 日後に台風が近づく」と、何日も先の空のようすを伝えます。予報する人が空を見上げて、雲の形から勘で決めているわけではありません。

天気予報のもとになっているのは、コンピュータの計算です。気象庁のコンピュータは、地球をおおう大気がこれからどう動くかを、物理の法則にしたがって計算しています（文献 序-1）。コンピュータの中で、明日の空を先に作ってみているのです。この本は、同じことを分子で行う化学、つまり計算化学への招待です。

◎ この章で分かること

- 天気予報は、大気を細かいマスに区切ったモデルを作り、コンピュータにくり返し計算させて、明日の空を予測しています
- 計算化学は、同じことを分子でします。分子のモデルを作り、コンピュータにくり返し計算させて、実験の前に結果を予測します
- コンピュータの仕事は、何をモデルにしたか・何をくり返したか・どこまで信じられるか、の 3 つの問いで読めます

高校で習ったこと：モデル化とシミュレーション（情報 I）

空を観測する、空を計算する

天気を知るいちばん確かな方法は、空を観測することです。地上の観測所、気球につるして飛ばす観測の機械、航空機、船、気象衛星などで、気温・気圧・風・水蒸気の

量を測ります（文献 序-2）。気象衛星などの人工衛星は、宇宙から雲のようすを写します（図 序-1）。ただし、観測で分かるのは今の空です。明日の空は、まだ観測できません。

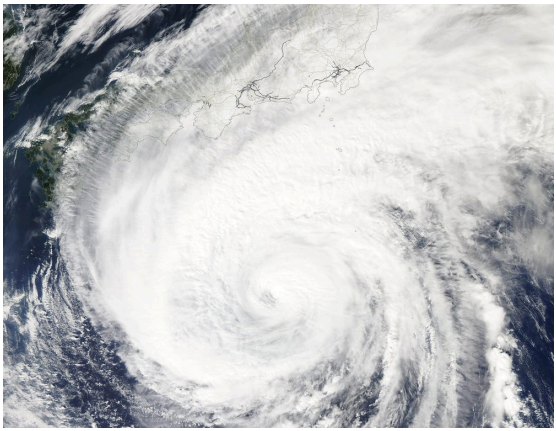


図 序-1　宇宙から見た台風（2019 年 10 月 11 日、台風第 19 号）。NASA の人工衛星 Aqua が撮影しました。左上に見えるのが日本列島です。（写真：MODIS Land Rapid Response Team, NASA GSFC, パブリックドメイン）

そこで、観測した今の空から、明日の空を計算で求めます。空気の動きは、物理の法則で決まっています。気圧の差に押された空気は、運動の法則にしたがって動きます。暖められた空気は膨らんで軽くなり、上へのぼります。水蒸気は、冷えると雲や雨になります。今の空のようすと、これらの法則が分かれば、少し先の空のようすを計算できるはずです。このように、コンピュータで大気の状態の変化を計算して天気を予測することを、数値予報と呼びます（文献 序-1）。

ただし、空気は切れ目なくつながっていて、そのままではコンピュータに渡せません。コンピュータが扱えるのは、限られた個数の数だけです。そこで、大気を縦・横・高さに細かく区切り、マスの集まりとして表します（図 序-2）。マス 1 つには、気温・風・水蒸気の量・気圧などの数を 1 組ずつ持たせます。

マスの並びを格子（grid）といいます。気象庁が地球全体の天気を予測する計算では、マスの幅は約 13 km で、高さ方向には大気を 128 層に区切っています（文献 序-2）。

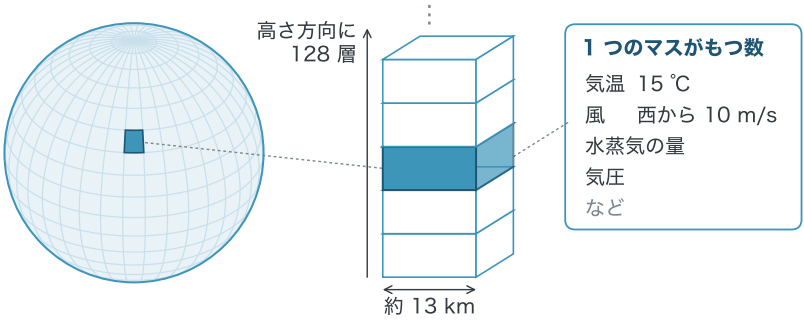


図 序-2 大気を格子のマスに区切ったようす（模式図）。地球の上の格子の線は、本物（約 13 km 間隔）よりずっと粗く描いています。マスの中の値は例です（文献 序-2）。

このように、本物の空を、コンピュータで計算できる形に置きかえたものをモデルと呼びます。

モデルを、決まった規則にしたがってコンピュータの中で動かし、どうなるかを調べることを、シミュレーションと呼びます。空を測るのが本物の観測なら、数値予報はコンピュータの中の実験です。

コンピュータの中の実験で、コンピュータが何をしているかは、次の 3 つの問いを立てるとよく分かります。何をモデルにしたか、何をくり返したか、どこまで信じられるか、です。次の節で、天気予報をこの 3 つの問いで読んでみます。

天気予報を 3 つの問いで読む

何をモデルにしたか

1 つ目の問いの答えは、前の節で見た格子のマスです。コンピュータに渡すのは、観測から作った今の大気の状態、つまり、すべてのマスの気温・風・水蒸気の量などの数です（文献 序-2）。コンピュータから返ってくるのは、何時間後、何日後のマスの数です。そこから、天気図や雨の予想を描きます。

マスに区切るときに、省いたものもあります。マスの中は 1 組の数で表すので、マスより小さなものは、1 つ 1 つを表せません（図 序-3）。マスの幅が約 13 km の計算でとらえられるのは、台風や梅雨前線のように、100 km ほどより大きな現象です。夏の夕立をもたらす積乱雲は、マスの幅を 1 km まで細かくした計算でも、1 つ 1 つを十分には表せません（文献 序-3）。

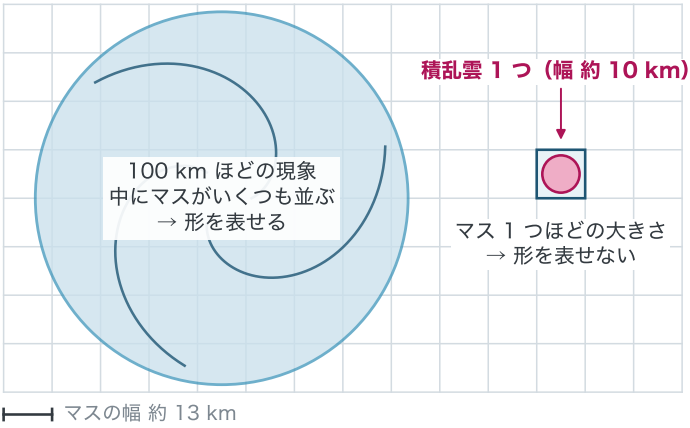


図 序-3　同じ縮尺で描いた、マスと現象の大きさ（形を簡単にした絵）。積乱雲 1 つの広がりは数 km〜十数 km で、ここでは 10 km で描きました（文献 序-4）。

何をくり返したか

2 つ目の問いに答えるために、あるマスの少し先の気温を、手で計算してみます。簡単にするために、風が気温を運ぶことだけを考えます。幅 13 km のマスが東西に 3 つ並んでいて、気温は西から順に 20 ℃、15 ℃、12 ℃ です。風は西から東へ、秒速 10 m で吹いています。

? 予想してみよう　**5 分後、真ん中のマスの気温は何 ℃ になると思いますか？**

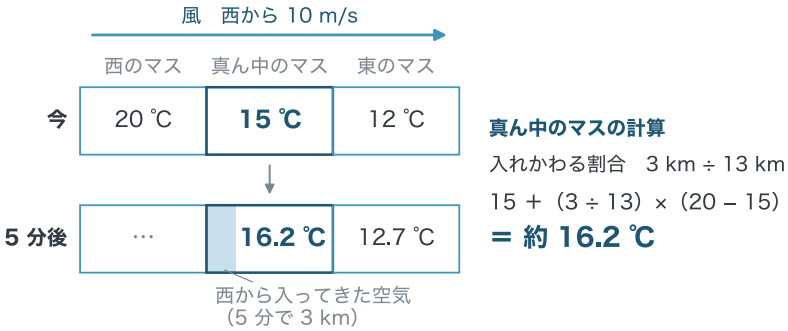


図 序-4　手で 1 回、マスの気温を 5 分進める。淡い青の部分が、5 分のあいだに西のマスから入ってきた空気です。風が気温を運ぶことだけを考えた、簡単にしたモデルです。

答えは約 16.2 ℃ です。秒速 10 m の風は、5 分（300 秒）で 3 km 進みます。真ん中のマスの幅は 13 km なので、5 分のあいだに、マスの空気のうち 13 分の 3 が、西のマスの 20 ℃ の空気と入れかわります。その分だけ、気温は 20 ℃ に近づきます。西のマスの 5 分後を求めるには、さらに西のマスの気温が必要です。

本物の数値予報では、気温だけでなく、風や水蒸気の量や気圧が互いにどう変わるかを、物理の法則の式からいっしょに計算します。それでも、計算の基本は同じです。となりのマスの数と規則から、それぞれのマスの少し先の数を求めます。

地球全体を計算するモデルは、手で 1 回やったマスの計算を、すべてのマスについて、時間を 5 分ほどずつ進めながらくり返します（図 序-5）。マスの計算は、合わせて 1 兆回を超えます。

コンピュータがなかった 1920 年ごろ、イギリスのリチャードソンは、格子に区切った大気の計算を手でやってみました。6 時間先の予報を出すのに 1 か月以上かかったうえ、計算の結果は大きく外れて、予報としては使えませんでした（文献 序-5）。コンピュータが使えるようになって、数値予報はようやく、毎日の予報に間に合う速さで計算できるようになりました。



図 序-5　手で 1 回の計算から、地球全体を計算するモデルの規模へ。マスの数は、地球の表面を 13 km 四方のマスで埋め、128 層に重ねて見積もりました（文献 序-2）。



テラ　6 時間先の空に、1 か月以上！ 答えが出るころには、その空はとっくに過ぎているね。くり返しは、コンピュータの出番だよ。

どこまで信じられるか

3 つ目の問いです。モデルは、本物の空そのものではありません。前に見たように、マスより小さな積乱雲は、1 つ 1 つを表せません。また、計算を始めるときの大気の状態は、観測しきれなかった分だけ不確かです。この始めの状態がどれだけ確かかが、予報の確かさを大きく左右します（文献 序-2）。先の予報になるほど、こうした小さな食いちがいが積み重なって、外れやすくなります。

数値予報は、初めから頼りにされていたわけではありません。気象庁がコンピュータで数値予報を始めた 1959 年から、信頼されるまでには、かなりの年月がかかりました（文献 序-5）。数値予報を信じられるものにしてきたのは、毎日の答え合わせです。

研究の現場から：予報の答え合わせ

数値予報では、出した予報を、あとで観測と比べて確かめます。次の予報を始めるときには、前の予報と、新しく届いた観測を合わせて、今の大気の状態を作り直します（文献 序-2）。

答え合わせで分かった弱点をもとに、気象庁はモデルの式を直し、マスを細かくし、コンピュータを新しくしてきました。台風の進路の予報も、年を追うごとに当たるようになっています（文献 序-2）。

計算の結果を実際と比べ、モデルを直していく。この流れは、計算化学でも同じです。第 6 章で、台風の予報の移り変わりを数で見ながら、答え合わせの話にもう一度戻ります。

分子でも、同じことができる

天気予報でコンピュータがしていることを、まとめてみます。空気の動きがしたがう物理の法則を、マスの数を 5 分先へ進める式に直し、その式をコンピュータにくり返し計算させて、明日の天気を予測します（図 序-6 の上の段）。

分子の世界でも、同じことができます。分子をつくる原子や電子も、物理の法則にしたがって動くからです。分子のモデルを作り、その規則にしたがってコンピュータにくり返し計算させれば、分子がどう動き、どんな性質をもつかを、実験の前に予測できます（図 序-6 の下の段）。これが計算化学（computational chemistry）です。「コンピュータ化学」と呼ばれることもありますが、どちらも同じ分野を指します。



図 序-6 天気予報（上）と計算化学（下）は、同じ流れで進みます。青い箱が、コンピュータの仕事です。

要点

計算化学とは、分子のモデルを作り、その規則にしたがってコンピュータにくり返し計算させ、実験の前に結果を予測することです。

計算化学は、フラスコの並ぶ実験室とは別に、コンピュータの中にもう 1 つの実験室を作ります。フラスコの中で分子を反応させる代わりに、コンピュータの中で分子のモデルを動かし、何が起こるかを調べるのです。

よくある誤解

「シミュレーションは、答えを前もってコンピュータに教えてあるから、その答えが出てくる」と考えるのは誤りです。数値予報のコンピュータは、「台風はこう進め」とは教えられていません。渡されるのは、今の大気の状態と、空気の動きの規則だけです。台風の進み方は、規則にしたがってマスの計算をくり返した結果として出てきます。分子のシミュレーションも同じです。第 1 章では、分子どうしが引き合う規則だけから、水と油が分かれていくようすを見ます。

分子の世界は、小さくて速い

では、なぜ分子を、わざわざコンピュータの中で調べるのでしょうか。理由の 1 つは、分子がとても小さく、しかもとても速く動いていることです。

まず、大きさを比べてみます。1 mm の水滴から始めて、10 倍ずつ小さくしていきます。

予想してみよう　**1 mm の水滴を地球の大きさまで拡大したとすると、その中の水分子は、どのくらいの大きさになると思いますか？**



図 序-7　大きさの物差し。目盛りを 1 つ下るごとに、大きさが 10 分の 1 になります。絵は形を簡単にしたもので、大きさは目盛りの位置で表しています（文献 序-6、序-7、序-8）。

答えは 4 m ほど、乗用車くらいです。水分子 1 個の大きさは 1000 万分の 3 mm ほどで、図 序-7 の物差しでは、いちばん下とその 1 つ上の目盛りのあいだにあります。1 mm の水滴を地球の大きさ（直径約 1 万 3000 km）にするには 130 億倍ほどに拡大するので、水分子も 4 m ほどになります。分子は、目ではもちろん、ふつうの顕微鏡でも見えません。

次に、速さです。室温（25 ℃）の水分子は、平均で秒速 590 m ほどの速さで動いています。音の速さ（秒速 350 m ほど）よりも速いのです。

ただし、まわりにぶつからずにその速さで進めたとしても、1 兆分の 1 秒（10^{-12} 秒）に進むのは、水分子 2 個ぶんほどです（図 序-8）。分子の世界の出来事は、1000 万分の 1 mm という大きさと、1 兆分の 1 秒という短い時間の中で起こっています。



テオ　水滴を地球の大きさにして、やっと車 1 台。それほど小さな分子の中の電子まで、第 2 章で計算して見せよう。

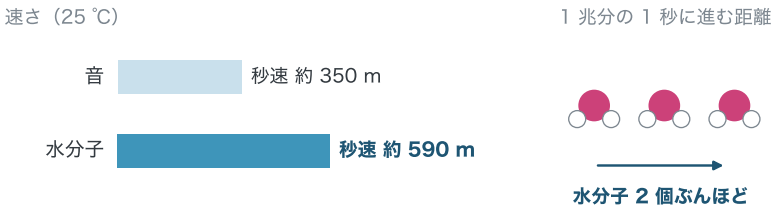


図 序-8　25 ℃ の水分子の平均の速さと音の速さ（左）。右は、その速さで 1 兆分の 1 秒のあいだに進む距離で、まわりの分子にぶつからないとしたときのものです。

これほど小さくて速いと、分子が動くようすを実験で 1 コマずつ追うのは、とてもむずかしくなります。2013 年のノーベル化学賞の発表でも、化学反応は一瞬で進むので、その一歩一歩を実験ですべて描き出すことは、ほとんど不可能だと述べられています（文献 序-9）。コンピュータの中なら、分子のモデルを 1 コマずつ止めて、原子の位置や電子の広がりを調べられます。実験では見えにくいところも、計算なら見ることができるのです。

ノーベル化学賞でたどる計算化学

分子をコンピュータで計算する研究は、1960 年代に、分子の中の電子を表す方程式を解くのにコンピュータが使われ始めたころから、大きく進みました（文献 序-10）。その歩みは、ノーベル化学賞にも表れています（図 序-9）。

賞の対象は、分子の中の電子を計算する方法（1998 年）から、タンパク質のような大きな分子を計算するモデル（2013 年）へ、さらにコンピュータによるタンパク



図 序-9　計算化学にかかわるノーベル化学賞。受賞者は姓だけを書き、授賞の理由は短くまとめました（文献 序-10、序-9、序-11）。

質の設計と、人工知能による構造の予測（2024 年）へと広がってきました。図 序-9 には、それぞれの研究を取り上げる章も書きました。

計算化学の 4 つのはたらきと、この本の構成

計算化学が化学の研究でしていることは、大きく 4 つのはたらきに分けられます（図 序-10）。小さくて速くて実験では見えにくい分子の動きや電子を、コンピュータの中で「見る」。何億もの候補の分子から、薬になりそうな分子を「探す」。分子を作る前に、その性質を「予測する」。たくさんのデータから、規則を「学ぶ」。

4 つのはたらきのうち「見る」だけは、分子の動きを調べる第 1 章と、電子を調べる第 2 章の 2 つに分けました。そのあとの第 6 章で、それまでの計算がどこまで当たるかを答え合わせします。

どの章も、同じ形で進みます。まず、その章で知りたいことを本物の実験や観察で調べ、次に「この実験を、コンピュータの中でできないか」と問いかけて、同じことをコンピュータの中で試します（表 序-1）。そして、天気予報でやったように、何をモデルにしたか、何をくり返したか、どこまで信じられるか、の 3 つの問いで、コンピュータの仕事を読みます。



図 序-10　計算化学の 4 つのはたらきと、それを受けもつ章。顔は、その章の案内役です。

テラ　電子の計算も、タンパク質の計算も、人工知能も、みんなノーベル賞につながったんだね！　この本で順に見ていくよ。

表 序-1 この本の各章で並べる、本物の実験とコンピュータの中の実験。

	本物の実験	コンピュータの中の実験
序章	空を観測する	大気をマスに区切り、空気の動きを計算する
第 1 章	水と油を振って、分かれるのを見る	分子を 1 個ずつ動かす
第 2 章	炎色反応やホタルの光の色を見る	電子を計算して、光の色を求める
第 3 章	酵素のはたらきが、薬の候補で止まるかを調べる	候補の分子を酵素に置いて、結合の強さを点数にする
第 4 章	水とオクタノール（油）を振って分ける	構造式から、水と油への分かれ方を計算する
第 5 章	測ったデータを表に集める	データから規則を学ぶ
第 6 章	実測の値で答え合わせをする	計算と実測を並べる
終章	選んだ材料を合成し、CO_2 の吸着を確かめる	たくさんの材料の穴に、CO_2 を入れてみる

天気予報も、コンピュータの中の実験だった。分子でも、やることは同じだ。





それが計算化学だね。最初は、わたしの番！ 分子を動かしてみよう。

最初に見るのは、分子の動きです。水と油は、よく振っても、しばらくすると分かれてしまいます。なぜでしょうか。第 1 章では、コンピュータの中で分子を 1 個ずつ動かして、その理由を探ります。

まとめ

- 天気予報（数値予報）は、大気を格子のマスに区切ったモデルを作り、物理の法則にしたがってマスの数を少し先へ進める計算を、コンピュータに 1 兆回を超えてくり返させて、明日の空を予測します。
- 計算化学は、同じことを分子でします。分子のモデルを作り、その規則にしたがってコンピュータにくり返し計算させ、実験の前に結果を予測します。

- 分子の世界は、1000 万分の 1 mm という大きさと、1 兆分の 1 秒という時間の世界です。実験では追いにくい動きも、コンピュータの中なら 1 コマずつ調べられます。
- コンピュータの中の実験は、何をモデルにしたか・何をくり返したか・どこまで信じられるか、の 3 つの問いで読みます。モデルは省いたもののせいで外れることがあるので、実際と比べて答え合わせをし、モデルを直していきます。



1 分子の動きを「見る」

サラダのドレッシングをよく振ると、油が細かい滴になって、全体がにごったように混ざります。ところが、しばらく置いておくと、油は上に、酢（ほとんどが水）は下に集まって、また 2 つの層に分かれてしまいます。どれだけ激しく振っても、結果は同じです。

水と油は、よく振っても、なぜやがて分かれてしまうのでしょうか。油の分子が、水の分子をはじき出しているのでしょうか。分子は小さすぎて、1 個ずつの動きを目で追うことはできません。そこで、分子を 1 個ずつコンピュータの中で動かして、そのわけを「見る」ことにします。

◎ この章で分かること

- 分子を 1 個ずつ動かす計算（分子動力学）では、原子をつぶ、結合をばねとみなし、電子を省きます
- コンピュータは、力 → 加速度 → 速さ → 位置 の計算を、時間を 1000 兆分の 1 秒ほどずつ進めながら、何百万回もくり返します
- 水と油が分かれるのは、水分子どうしが水と油の分子よりずっと強く引き合う（水素結合）からです。コンピュータの中でも確かめます

高校で習ったこと：熱運動と温度・運動の法則（物理基礎）、物質の三態（化学基礎）、分子間力と水素結合（化学）。化学の内容は、習っていなくても読めます

ドレッシングは、どれだけ振っても分かれる。理由は、びんを眺めていても分からない。





だったら、分子を 1 個ずつ動かしてみようよ。コンピュータの中でね！

1.1 本物の実験：水と油、氷と水と水蒸気

まず、本物の実験から始めます。透明なびんに水とサラダ油を入れ、ふたをしてよく振ると、ドレッシングと同じように全体が白くにごります。びんを置いておくと、油の細かい滴はしだいに集まって大きくなり、やがて上の層に戻ります（図 1-1）。

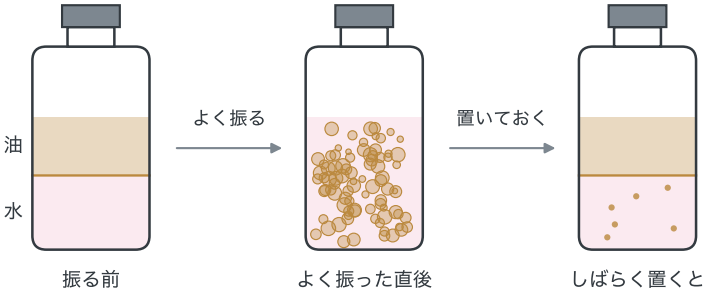


図 1-1　水と油を入れたびんを振る実験。びんと滴は、形を簡単にした絵です。

この実験で分かるのは、「水と油は混ざらない」という結果です。序章で見たように、分子は 1000 万分の 3 mm ほどの大きさしかなく、秒速数百 m で動いているので、分かれていくあいだに分子が 1 個ずつ何をしているのかは、びんを見ていても分かりません。

もう 1 つ、身近な実験があります。氷を温めると、0 ℃ で溶けて水になり、100 ℃ で沸騰して水蒸気になります。氷も水も水蒸気も、同じ水分子でできています。温度を変えるだけで、固体・液体・気体という 3 つの状態（三態）が入れかわるのです。では、氷と水と水蒸気では、分子の何がちがうのでしょうか。これも、目で見ることはできません。

本物の実験から、何が起こるかは確かに分かります。けれども、分子のレベルでなぜそうなるのかまでは見えません。この実験を、コンピュータの中でできないでしょうか。分子を 1 個ずつ動かせば、分子が何をしているのかを、コマ送りの映像のように見られるはずです。そのためには、まず、分子をコンピュータが計算できる形にする必要があります。

1.2 分子をつぶとばねで表す

分子をコンピュータの中で動かすために、まず分子のモデルを作ります。モデルは、本物を、コンピュータで計算できる形に置きかえたものでした（序章）。置きかえるときには、分かりたいことに必要なものを残し、それ以外の細部は省きます。何を残し、何を省くかを、水分子で見てみます。

本物の水分子は、1 個の O 原子と 2 個の H 原子が、共有結合でつながったものです。O–H の結合の長さは 1000 万分の 1 mm ほどで、2 本の結合は 104.5° の角度で開いています（文献 1-1）。それぞれの原子は、中心の原子核と、そのまわりに広がる電子でできています（図 1-2 の左）。

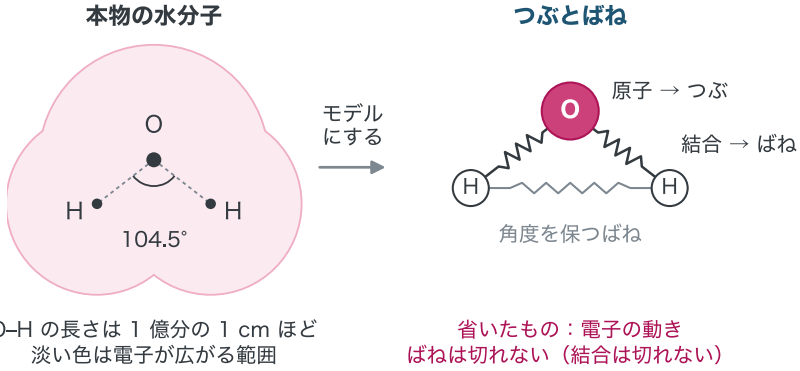


図 1-2 本物の水分子（左）と、つぶとばねのモデル（右）。左の淡い色は、電子が広がる範囲の目安です。

分子の動きを追うときに知りたいのは、原子がどこにあり、どちらへ動くかです。そこで、原子を、決まった質量をもつつぶとみなします。原子どうしを結ぶ共有結合は、のびたり縮んだりするばねとみなします。結合の角度も、開きすぎたり閉じすぎたり

しないように、ばねで保ちます。このように、原子をつぶ、結合をばねで表したモデルを、つぶとばねのモデルと呼びます（図 1-2 の右）。

このモデルは、原子の位置と、結合の長さや角度をよく表し、結合が伸び縮みする振動も表せます。一方で、電子の動きは省いています。また、本物の分子では、結合が切れたりできたりして、別の分子に変わることがあります。これが化学反応です。つぶとばねのモデルでは、ふつう、結合が切れることも新しくできることもありません（文献 1-2）。

つぶを動かすには、つぶにはたらく力を決めなければなりません。力には 2 種類あります（図 1-3）。1 つは、結合のばねの力です。ばねは、のびると縮める向きに、縮むとのばす向きに、つぶを引いたり押したりします。もう 1 つは、結合していないつぶどうしの力です。2 つのつぶは、近づきすぎると強く押し返し合い、少し離れると弱く引き合い、遠く離れるとほとんど力をおよぼしません。この力は、別々の分子の原子どうしにもはたらき、分子どうしを引き合わせます。このように分子どうしが引き合う力を、分子間力といいます。

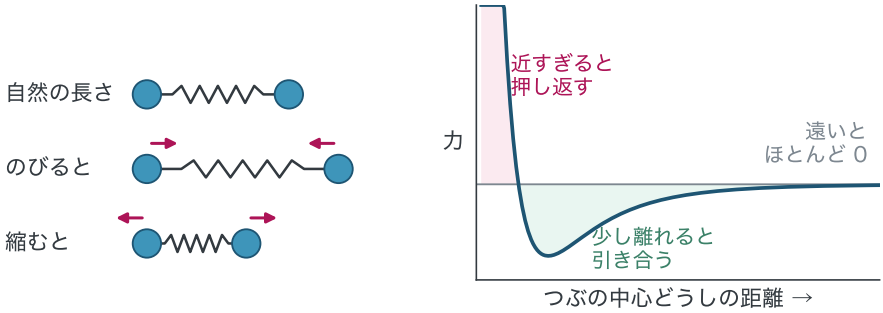


図 1-3　つぶとばねにはたらく力の規則。左は結合のばねの力（矢印が力の向き）、右は結合していないつぶどうしの力と距離の関係です。右の曲線は、この章の計算で使った式で描きました。目盛りの値は描いていません。

さらに、つぶには、わずかなプラスやマイナスの電気も割りふっておきます。プラスとマイナスのつぶは引き合い、同じ符号のつぶは反発し合います。この電気の割りふりが、この章の後半で見る水素結合のもとになります。



テオ　つぶとばねで省いた電子を計算するのが、量子化学の仕事だ。電子が雲のように広がるようすは、第 2 章で見せよう。

ばねの強さ、つぶの大きさ、引き合いの強さ、電気の量。こうした力の決め方の一式を、力場（force field）と呼びます。力場の数は、実験の結果や、電子まで考えた計算（第 2 章）に合うように決めてあります（文献 1-2）。

これで、コンピュータに渡すものがそろいました（図 1-4）。はじめのつぶの位置と速さ、力場、そして温度や、計算する時間の長さなどの条件です。たとえば水を調べたいなら、水分子をたくさん並べた箱を用意します。コンピュータは、つぶを少しずつ動かす計算をくり返し、つぶの位置を 1 コマずつ記録して返します。記録を並べるとコマ送りの映像になり、そこから温度や、分子のまざり具合などの数も読みとれます。

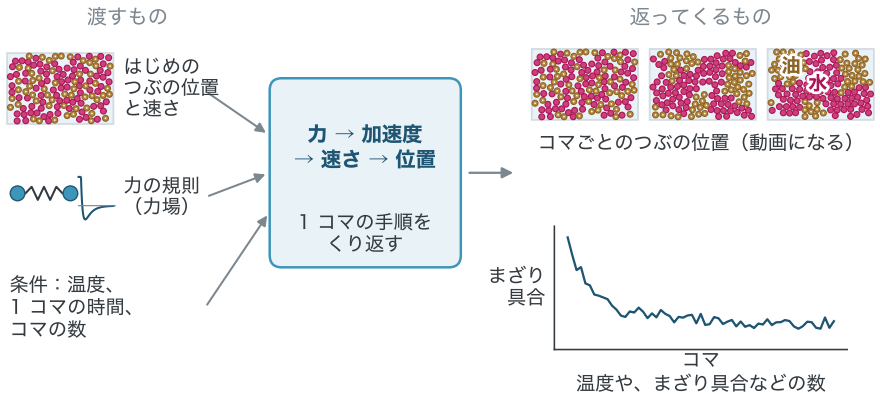


図 1-4　つぶを動かす計算で、コンピュータに渡すもの（左）と、返ってくるもの（右）。つぶの絵と線は、この章の水と油の計算（1.5 節）の出力です。

では、コンピュータは、1 コマのあいだに何を計算しているのでしょうか。

1.3 1 コマずつ動かす

つぶを動かす規則は、運動の法則です。物体に力 F がはたらくと、物体は、力を質量 m で割った加速度 a で速さを変えます。

$$F = ma$$

つぶにはたらく力が分かれば、加速度が求まります。加速度からはほんの一瞬あとの速さが、速さからはその一瞬のあとにつぶがどこにいるかが求まります。つぶが動け

ば、つぶどうしの距離が変わり、力も変わります。そこで、新しい位置で力を計算し直し、また一瞬だけ進めます。このくり返しで、つぶの動きを 1 コマずつ作っていくのです。

この 1 コマの計算を、手で 1 回やってみます。壁にばねでつながれたつぶを 1 つ考え、ばねが自然の長さになる位置を 0 とします。ばねの強さを 1、つぶの質量を 1、1 コマの時間を 0.5 とし、単位は省いて数だけで考えます。向きは符号で表し、位置が大きくなる向きをプラスとします(向きを含めた速さを速度といいますが、この章では速さと書きます)。はじめ、つぶは位置 4 に止まっています。

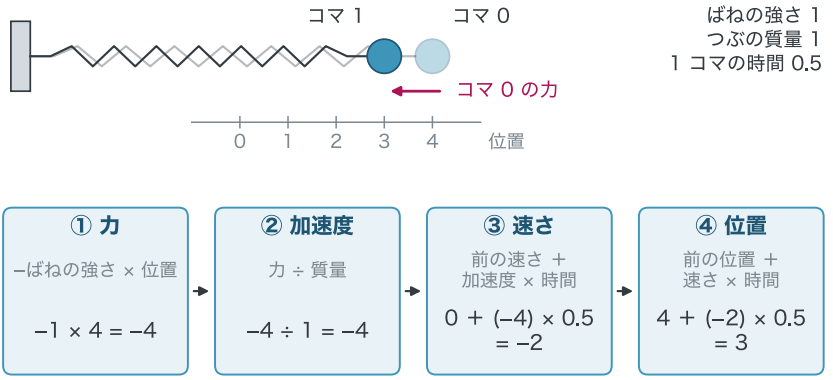


図 1-5 手で 1 回、1 コマ進める。上はつぶの位置（薄い丸がコマ 0、濃い丸がコマ 1)、下は 4 つの手順の計算です。

図 1-5 の下の 4 つの箱を、左から順に見てください。力から加速度、速さ、位置と順に求めると、つぶは位置 4 から 3 へ進みます。かけ算と足し算だけで、1 コマ分の動きが求まりました。

コマ 1 の位置 3 から、同じ 4 つの手順をもう一度やれば、コマ 2 の位置が求まります。この手順を、コンピュータに 30 回くり返させてみます（図 1-6)。

つぶは、およそ 4 と −4 のあいだを行ったり来たりしています。端（①・③・⑤)に来たときは力がいちばん大きく、速さは 0 に近くなります。真ん中（②・④）では力が 0 に近く、速さがいちばん大きくなります。手順のどこにも「振動しなさい」とは書いていません。それでも、「力 → 加速度 → 速さ → 位置」をくり返すだけで、ばねの振動があらわれました。このように、原子にはたらく力から、運動の法則

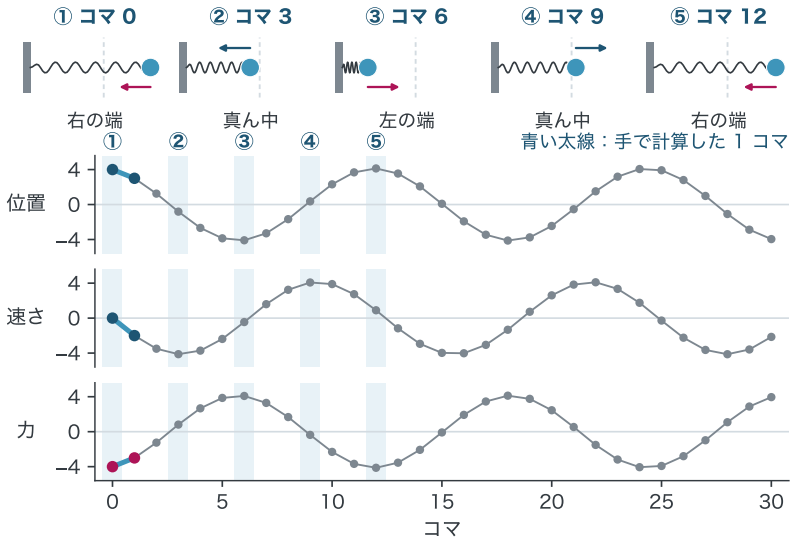


図 1-6　図 1-5 の手順を 30 回くり返したときの、つぶの位置・速さ・力（下）と、番号のコマのばねの絵（上）。絵のつぶの位置と矢印の長さは、そのコマの値です。青の矢印が速さ、ローズの矢印が力で、点線が位置 0 です。

にしたがって原子の動きを 1 コマずつ計算していく方法を、分子動力学（molecular dynamics）と呼びます。

なお、ここで使ったのは、いちばん簡単な手順です。本物の分子動力学では、速さを半コマ分ずつ 2 回に分けて新しくするなど、少していねいな手順を使い、計算の誤差がたまりにくくします。考え方は 図 1-5 と同じです。

1 コマは、どれだけ短いか

1 コマの時間は、どれだけ短くすればよいのでしょうか。パラパラ漫画と同じで、速い動きをコマ数の少ない粗い絵で描くと、途中の動きが飛んでしまいます。1 コマの計算はコマの初めの力をそのまま使うので、力が大きく変わる前に次のコマへ進まなければなりません。図 1-5 のばねのつぶで、1 コマの時間だけを変えて試してみます（図 1-7）。

1 往復を 13 コマほどに刻むと、つぶは正しい動きとほぼ重なって振動します。3 コマしか刻まないと、つぶは行き来するたびに遠くへ飛んでいき、計算が壊れました。いちばん速い動きの 1 往復を 10 コマほどに刻むのが目安です。

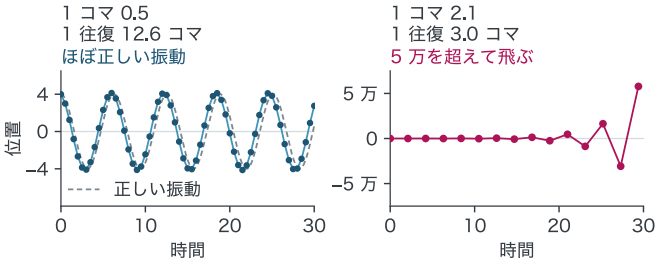


図 1-7 1 コマの時間を変えて、図 1-5 のばねのつぶを動かした結果。丸が各コマの位置、灰色の破線は、ばねの振動の正しい動きです。右は、縦の目盛りがちがいます。

分子の中でとりわけ速いのは、いちばん軽い H 原子が振動する、O–H の結合の伸び縮みです。水分子では、1000 兆分の 9 秒ほどで 1 往復します（文献 1-3）。これを 10 コマほどに刻むので、1 コマは 1000 兆分の 1 秒（10^{-15} 秒）ほどになります。この速い伸び縮みを省いて、1 コマを数倍にすることもよくあります（文献 1-2）。1 コマで進める時間の幅を、時間刻み（time step）といいます。

1 コマがこれほど短いと、少し長い時間を調べるにも、たくさんのコマが必要です（図 1-8）。

図 1-8 の目盛りは、1 つ右へ進むごとに 1000 倍になります。10 億分の 1 秒を進めるだけで、100 万コマが必要です。実際の研究では、何百万〜何十億コマの計算がふつうです。

水に囲まれた中くらいの大きさのタンパク質では、原子はおよそ 5 万個になり、コンピュータは 1 コマごとに、原子どうしの力を何百万組も計算します。2018 年の報告では、GPU（画像を描くための計算部品）を 1 枚使って、図 1-8 の 10 億コマ（100

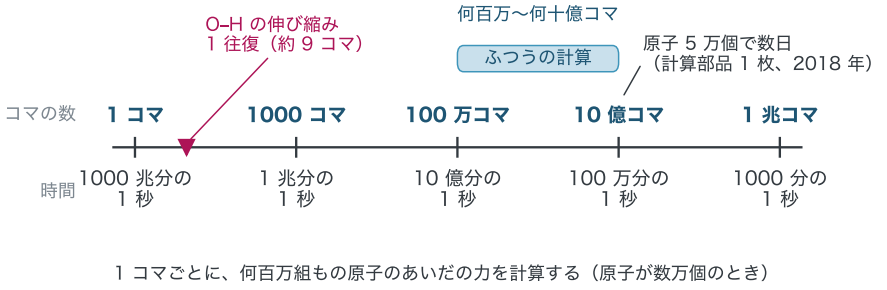


図 1-8 分子動力学の時間のものさし。1 コマを 1000 兆分の 1 秒としたときの、時間とコマの数です（文献 1-2）。

万分の 1 秒）まで進めるのに数日かかりました（文献 1-2）。手で 1 回やった 図 1-5 の計算を、5 万個の原子について、10 億回ほどくり返すのです。

要点

分子動力学で、コンピュータは「力 → 加速度 → 速さ → 位置」という 1 コマの計算を、たくさんの原子について、何百万回もくり返します。1 コマの手順は手でもできるほど単純ですが、くり返すと、振動や、分子の集まり方などの動きがあらわれます。

それでは、このくり返しで、氷・水・水蒸気のちがいが見えるかどうかを確かめます。

1.4 温度は熱運動の激しさ

分子は、止まっているように見える氷の中でも、いつも動いています。この絶え間ない動きを、熱運動と呼びます。温度は、この熱運動の激しさを表す数です。温度が高いほど、分子は平均して速く動き、運動エネルギーの平均が大きくなります。

コンピュータの中では、温度はつぶの速さで決まります。この章の計算では、つぶ 1 個あたりの運動エネルギーの平均を、そのまま温度の値とします（図 1-9）。運動エネルギーは速さの 2 乗に比例するので、全部のつぶの速さを 2 倍にすると、温度は 4 倍になります。そこでコンピュータは、何コマかごとに、全部のつぶの速さに同じ数を少しずつかけて、目標の温度に近づけます。

ここからは、本書で実際に動かした計算を見ていきます。1.2 節のモデルをさらに簡単にした、2 次元の計算です。分子 1 個を 1 つのつぶとし（ばねも電気もありません）、つぶどうしの力は 図 1-3 の右の規則だけにしました。平たい箱につぶを入れ、温度を 3 通りに変えて動かします。箱に壁はなく、右の端から出たつぶは左の端から入ります。同じ箱が上下左右に果てしなく並んでいるとみなすこのやり方を、周期境界条件（periodic boundary condition）といい、少ないつぶで広い空間をまねできます。温度の数は ℃ ではなく、つぶどうしの引き合いの強さを 1 とした目盛りです。



テラ 10 億コマも動かすなんて、ひとりじゃ無理！ GPU は、たくさんの計算をいっせいにこなすのが得意。その力を借りるんだよ。

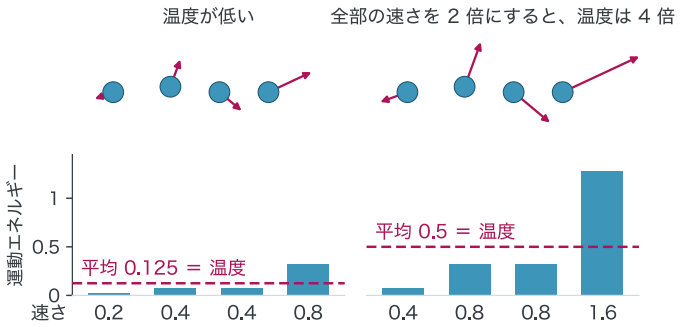


図 1-9　つぶの速さと温度（質量 1 のつぶ 4 個の例）。上の矢印の長さがつぶの速さ、下の棒がつぶごとの運動エネルギーです。

？ 予想してみよう　**温度をうんと低くすると、つぶはどのように並ぶと思いますか？**

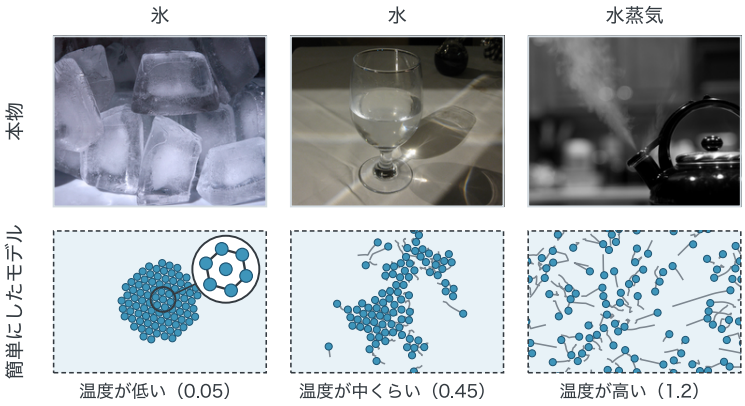


図 1-10　本物の氷・水・水蒸気（上、写真）と、コンピュータの中で 3 つの温度で動かしたつぶ（下）。水蒸気そのものは目に見えません。写真で白く見える湯気は、水蒸気が冷えてできた小さな水滴です。下の段は、つぶ 90 個を 1 万 8000 コマ動かした結果です。分子を丸 1 個にまで簡単にしたモデルで、本物の水分子の形はしていません。灰色の線はつぶが直前の 500 コマで動いた道のり、破線は箱の端です。左の枠の円は、囲んだ 7 個を拡大し、まわりの 6 個を線で結んだものです。（写真：左から liz west, CC BY 2.0、MarshyBoardwalk, CC0、Shardayyy Photography, CC BY 2.0）

答えは、六角形の並びです。温度が低いと、つぶは 1 つのかたまりになり、内側のつぶは、それぞれ 6 個のつぶに囲まれて、ほぼすき間なく並びました（図 1-10 の下の

左)。つぶは決まった場所で小さくゆれるだけで、道のりの線はほぼ見えません。これは固体の結晶です。ただし、本物の氷は、水分子が水素結合で 4 個とつながる、すき間の多い並びで、水に浮きます。どの向きにも引き合うつぶとは、ここがちがいます。温度が中くらいだと、つぶは寄り集まったまま、位置を入れかえながら動き回ります。液体です。温度が高いと、つぶはばらばらに箱いっぱいを飛び回ります。気体です。

計算のどこにも、「低い温度では結晶になりなさい」とは書いていません。つぶどうしの引き合いと、温度で決まる動きの激しさのどちらが勝つかで、3 つの状態があらわれたのです。引き合いが勝てばつぶは集まり、動きが勝てばばらばらになります。

では、同じ温度のとき、つぶはみな同じ速さで動いているのでしょうか。コンピュータは、すべてのつぶの速さを、コマごとに記録しています。その速さを集めて、速さごとの割合を数えてみます（図 1-11)。

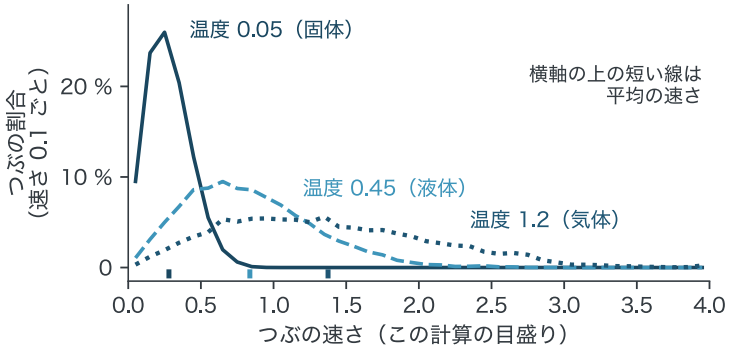


図 1-11　3 つの温度での、つぶの速さの分布（図 1-10 と同じ計算の出力)。

同じ温度でも、つぶの速さはそろっていません。遅いつぶもあれば、速いつぶもあります。温度が高いほど、分布は右へ広がり、速いつぶの割合が増えます。温度は、1 個のつぶの速さではなく、たくさんのつぶの動きの激しさを平均したものなのです。

本物の分子も同じです。室温の水の中では、水分子は平均で秒速 590 m ほどで動き、互いにぶつかり合っています。タンパク質の原子も、熱運動で少しずつゆれ動いています。第 3 章で見る、酵素が基質を受けとめる場所（活性部位）の形も、止まっているわけではありません。

ここまでで、温度によって分子の集まり方が変わることが見えました。次は、2 種類の分子を混ぜたらどうなるかを見ます。

1

1.5 水と油はなぜ分かれるか

水と油の話に戻ります。水と油のちがいを、分子どうしの引き合いで見てみます。

水分子の O 原子は、わずかにマイナスの電気を帯び、H 原子は、わずかにプラスの電気を帯びています（図 1-12 の δ− と δ+）。なぜそうなるのかは、第 2 章で、電子を見て確かめます。

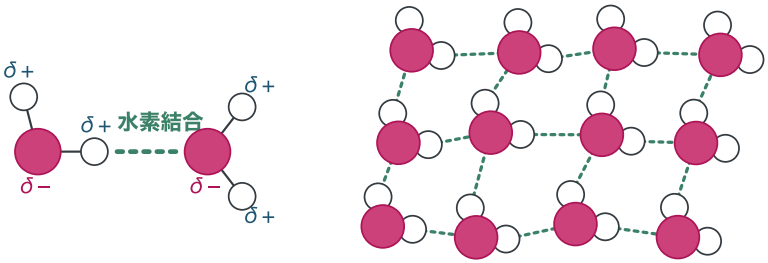


図 1-12 水分子どうしの水素結合（左）と、水の中での水素結合のつながり（右）。緑の点線が水素結合です。右は 2 次元に描いた模式図で、計算した配置ではありません。

ある水分子の H（プラス）と、となりの水分子の O（マイナス）は、互いに引き合います。このように、O や N についた H が、となりの分子の O や N と引き合う結びつきを、水素結合（hydrogen bond）と呼びます。水素結合は、分子間力の中でもとくに強い引き合いです。C と H だけでできた分子どうしも、ファンデルワールス力という分子間力で引き合いますが、原子 1 組あたりで比べると、水素結合のほうがずっと強くなります。

1 個の水分子は、2 つの H と O を使って、まわりの何個もの水分子と水素結合をつくれます。そのため水の中では、水素結合が網の目のようにつながっています（図 1-12 の右）。液体の水では、熱運動のために、このつながりが切れてはつながり直すことをくり返しています。

一方、油の分子は、ほとんど C と H だけでできています（図 1-13）。C も H も、プラスやマイナスの電気をほとんど帯びていません。そのため油の分子は、水分子と水素結合をつくれず、水分子とは弱い分子間力で引き合うだけです。

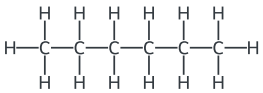
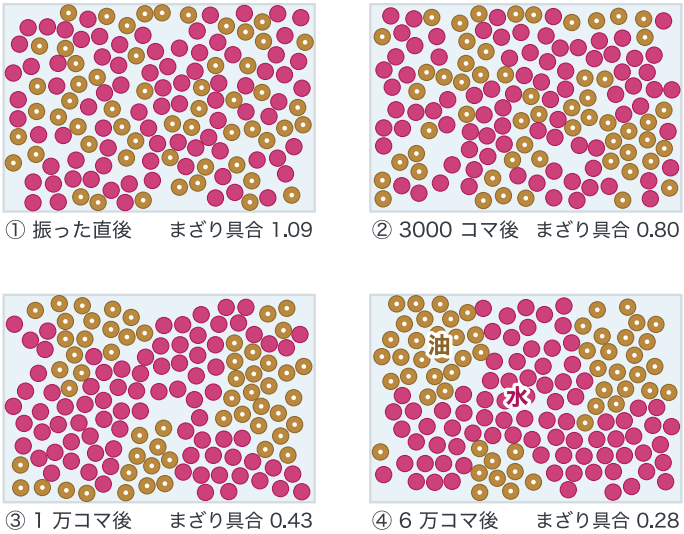


図 1-13 油の分子の例（ヘキサン）。

コンピュータの中で、水と油を混ぜる

この引き合いのちがいだけで、水と油は分かれるのでしょうか。コンピュータの中で確かめてみます。つぶは水 6 割と油 4 割の合わせて 140 個にします。温度を変えた先ほどの計算とちがって、今度の箱には壁があり、つぶは壁ではね返ります。つぶに電気は割りふらず、引き合いの強さ（図 1-3 の右）をつぶの組ごとに変えます。水と水はいちばん強くして、水素結合のかわりにします。油と油は、水と水の 10 分の 9 にしました。つぶ 1 個が表す油の分子は大きく、原子 1 組ずつの弱い引き合いも、分子 1 個あたりで合わせるとかなり強くなるからです。水と油は 10 分の 3 と、ずっと弱くしました（引き合いはあります）。つぶを箱いっぱいにでたらめに並べ、よく振った直後の状態から動かします。

予想してみよう **よく振った直後の並びから動かすと、つぶはどうなると思いますか？**



① 振った直後　まざり具合 1.09　② 3000 コマ後　まざり具合 0.80

③ 1 万コマ後　まざり具合 0.43　④ 6 万コマ後　まざり具合 0.28

図 1-14　水のつぶ（ローズ）と油のつぶ（黄、中に白い点）を、よく振った直後の並びから動かしたようす。まざり具合は、となりどうしのつぶが別の種類である割合を、でたらめに混ぜたときを 1 として表した数です。分子 1 個を 1 つのつぶにした、2 次元の簡単な計算です。

つぶは、水と油に分かれていきます。はじめは水と油のつぶが入りまじっていましたが（①）、コマが進むにつれて、水のつぶどうしが集まり、油のつぶは、水のかたまりの外へ追いやられていきます（② → ④）。まざり具合も、1 ほど（でたらめに並べても少し上下します）から 0.3 ほどまで下がりました。

ここでも、計算のどこにも「水と油は分かれなさい」とは書いていません。コンピュータは、1 コマごとに、すべてのつぶについて 図 1-5 と同じ手順をくり返しただけです。

つぶ 140 個では、つぶの組は約 1 万組あります。コンピュータは 1 コマごとに、すべての組で距離を測りますが、力を足し合わせるのは、近くにいる組だけです（図 1-15）。これを 6 万コマくり返しても、ノートパソコンで 50 秒ほどで終わりました。研究の計算でも、遠い組の力は省くか、まとめて大まかに見積もり（文献 1-4）、1 コマで力を計算する組を抑えています（文献 1-2）。

なぜ分かれるのか

分かれた理由を探るために、コンピュータが 1 コマで何を計算しているのかを、3 つのつぶを使って手でたどってみます。力の規則も、引き合いの強さも、1 コマの時間（0.005）も、140 個の計算と同じです。一直線に水・水・油のつぶを並べ、真ん中の水のつぶを動かします。左の水と右の油は、どちらも真ん中から距離 1.25 に置きます（距離は、つぶの大きさを 1 とした目盛りです）。

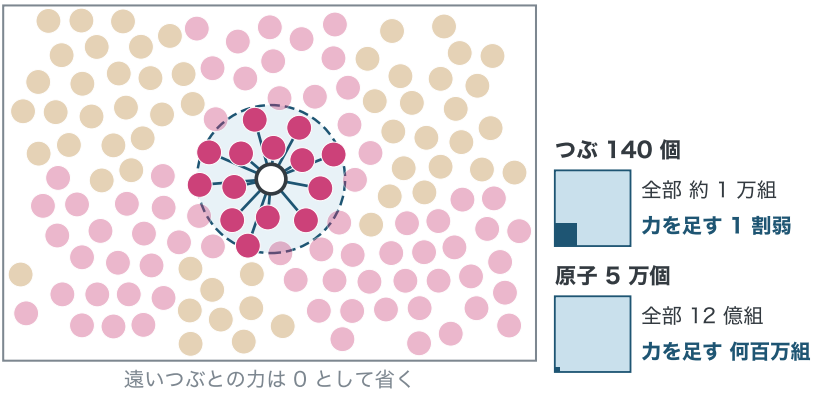


図 1-15　遠い組の力を省く。左は水と油の計算の最後のコマで、白いつぶと、破線の円（距離 2.5）の中のつぶとの組だけを線で結びました。右の四角の面積は組の数で、縮尺は上下でちがいます。原子 5 万個の濃い四角は、500 万組として描きました。

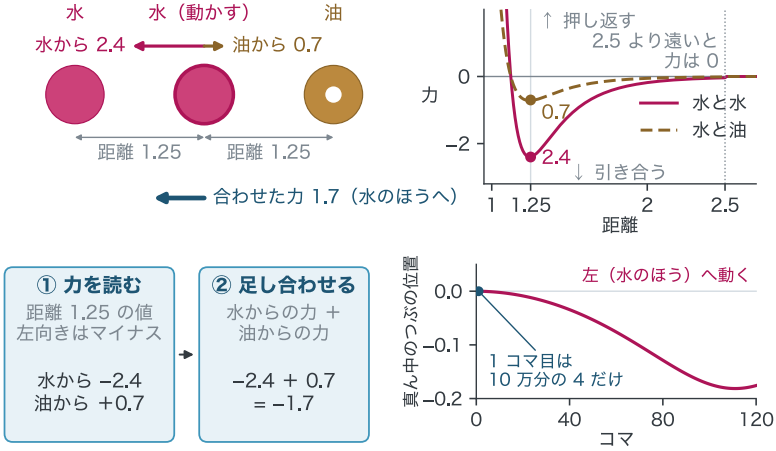


図 1-16　手で 1 回、水のつぶを 1 コマ進める。左上は 3 つのつぶと、真ん中の水のつぶにはたらく力（矢印の長さが力の大きさ）、右上はつぶどうしの力と距離の関係（図 1-3 の右と同じ式）、左下は力を求める計算、右下は左右のつぶを止めたままコマを重ねたときの真ん中のつぶの位置です。

図 1-16 の右上の曲線で距離 1.25 のところを読むと、距離は同じなのに、真ん中のつぶが水のつぶから受ける引き合う力は、油のつぶから受ける力の 3 倍あまりです（引き合いの強さの比 1 対 0.3 のとおり）。2 つの力は逆向きなので、真ん中の水のつぶは、差し引きで水のほうへ引かれます。ここから先は 図 1-5 と同じ手順です。1 コマで動くのはごくわずかですが、コマを重ねると動きが積み重なり、水のつぶに引き寄せられていきます（右下）。

このように水のつぶは、同じ距離にいる油のつぶより、水のつぶのほうへ強く引かれます。水のつぶどうしが寄り集まった並びのほうが安定なので、熱運動で動き回るうちに、水のつぶはかたまっていきます。そうなると、はみ出した油のつぶは、油どうしで集まるしかありません。油が水をはじき出すのではなく、水どうしが強く、水と油が弱くしか引き合わないために、油が外へ押し出されるのです。

本当にそうなのかを、コンピュータの中の実験で確かめてみます。本物のびんでは、水と油の引き合いの強さだけを変えることはできません。コンピュータの中な



テラ　ドレッシングを使う直前に振るのは、すぐ分かれちゃうからだね。分かれるまでのあいだに、油の滴が野菜にからむよ！

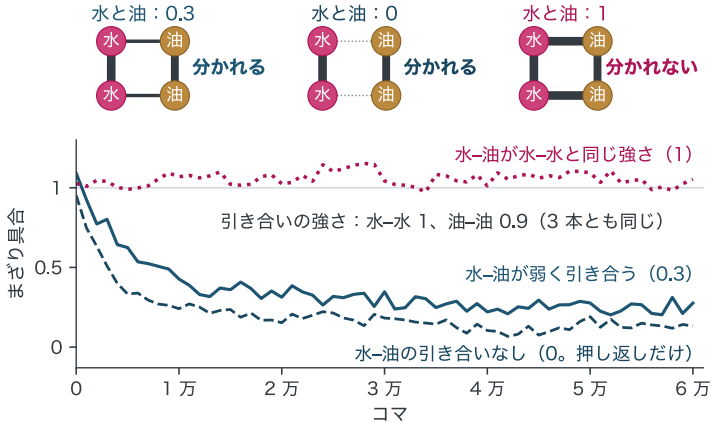


図 1-17　水と油のつぶの引き合いを 3 通りに変えたときの、まざり具合の変わり方。上の絵は引き合いの強さを線の太さで表したもので、下の 3 本の線と色が対応します。まざり具合が 1 なら、でたらめに混ざった状態です。

ら、力の規則の数を書きかえるだけです。そこで、水と油のつぶの引き合いだけを変えて、ほかの条件は同じにした計算を、さらに 2 通り行いました。1 つは、水と油のつぶの引き合いをなくし、近づきすぎたときの押し返しだけを残した計算です。もう 1 つは、水と油のつぶを、水と水のつぶと同じ強さで引き合わせた計算です（図 1-17）。

水と油の引き合いをなくすと、まざり具合は、弱く引き合うときよりも少し低くまで下がり、水と油は分かれました。反対に、水と油を水どうしと同じ強さで引き合わせると、まざり具合は 1 のままで、いつまでたっても分かれませんでした。水と油が分かれるかどうかは、水と油の引き合いがあるかないかではなく、水どうしの引き合いが水と油の引き合いよりずっと強いかどうかで決まるのです。

同じことは、水と油以外の分子にも言えます。水と水素結合をつくれる O–H や N–H の多い分子は、水分子の網の目に入りこめるので、水になじみます。C と H の多い分子は、水の外へ押し出されて、油になじみます。薬の分子がどちらになじむかは、第 4 章で見ます。

よくある誤解

「油は、水をはじく力をもっている」と考えるのは誤りです。水と油の分子のあいだにも、弱い引き合いがはたらいています。水分子どうしが水素結合で集まるために、水素結合をつくれない油の分子が、水の外へ押し出されるのです。

発展：並び方の数（エントロピー）　飛ばしても読めます

水と油が分かれるかどうかには、引き合いの強さのほかに、並び方の数もかかわります。たとえば、水 3 個と油 3 個を 1 列の 6 マスに並べる並び方は 20 通りあります。そのうち、水と油が 2 つに分かれた並びは 2 通りだけで、残りの 18 通りは混ざった並びです（図 1-18）。

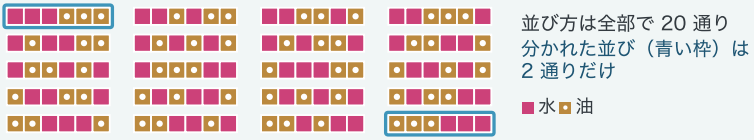


図 1-18　水 3 個と油 3 個を 6 マスに並べる、20 通りの並び方。

分子の数が増えると、この差はとても大きくなります。熱運動で分子がでたらめに動けば、数の多い、混ざった並びになりやすいのです。並び方の数の多さを表す量を、エントロピー（entropy）と呼びます。温度が高いほど、並び方の数の多さが効いてきます。この章のつぶの計算では、引き合いの強さが勝って分かれました。本物の水では、油の分子を囲む水分子が、水素結合を切らないよう限られた向きにしか並べず、並び方の数が減ります。ふだんの温度で油が水に溶けにくいのは、おもにこのためです（文献 1-5）。第 4 章の計算でも、引き合いの強さと並び方の両方を考えます。

どこまで信じられるか

図 1-17 からは、もう 1 つ大事なことが分かります。力の規則の数を 1 つ変えただけで、分かれるという結果が、分かれないという結果に変わりました。分子動力学の結果は、どんな力場を使うかしだいなのです。研究で使う力場は、実験の結果や電子まで考えた計算に合うように、ていねいに数を決めてあります。それでも、本物の力を簡単な式で置きかえたものなので、いつも正しい答えを出すとは限りません（文献 1-2）。

また、分かれるという結果が合っていても、その理由まで本物どおりとは限りません。この章のつぶの計算では、引き合いの強さのちがいで分かれました。本物の水では、発展の箱で見た並び方の数も大きくかかわります。

さらに、つぶとばねのモデルではばねが切れないので、結合が切れたりできたりする化学反応は、このままでは扱えません。また、1 コマがとても短いので、計算でき

る時間の長さにも限りがあります。どの計算をどこまで信じてよいかは、第 6 章で、ほかの計算とあわせて答え合わせをします。

研究の現場から：タンパク質が水の中で形を変えるようすを追う

第 3 章では、薬の候補の分子を、酵素の活性部位に置いて、結合の強さを点数にする計算（ドッキング）を見ます。そこでは、活性部位をほとんど動かない形として扱います。けれども本当の活性部位は、水の中で熱運動をして、少しずつ形を変えています。

そこで研究の現場では、水に囲まれたタンパク質を分子動力学で動かし、活性部位の形のちがう何コマかを取り出して、それぞれにドッキングすることがあります（図 1-19）。エイズの原因になるウイルスがもつ酵素の 1 つでは、10 億分の 2 秒の分子動力学から取り出したコマにドッキングすると、活性部位のとなりに、ときどき開く溝が見つかりました（文献 1-6）。止まった形だけからは見えなかった溝です。その後、酵素のはたらきを止める分子がこの溝に結合することが実験で確かめられ、この酵素のはたらきを止める初めての薬の開発につながったと紹介されています（文献 1-7）。

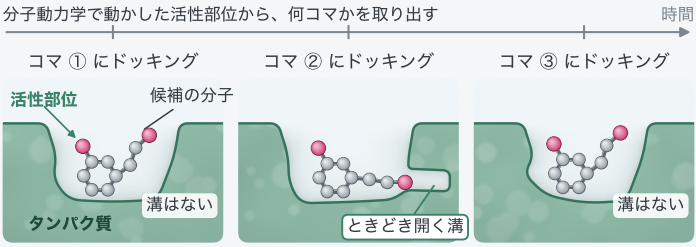


図 1-19　分子動力学で動かした活性部位の何コマかに、候補の分子をドッキングする（模式図）。緑の地がタンパク質で、その表面の淡い水色の所が活性部位です。候補の分子は、玉が原子（灰が C、ローズが O）、棒が結合です。

まとめ

- 分子動力学では、原子をつぶ、結合をばねとし、力の決め方の一式（力場）にしたがって原子を動かします。電子は省き、結合は切れません。

- コンピュータは「力 → 加速度 → 速さ → 位置」の 1 コマの計算を、時間を 1000 兆分の 1 秒ほどずつ進めて、何万個もの原子について何百万〜何十億回くり返します。
- 温度は熱運動の激しさです。温度を変えてつぶを動かすと、結晶・液体・気体の 3 つの状態があらわれます。
- 水分子どうしは、水素結合で、水と油の分子よりずっと強く引き合います。そのため、油が水をはじく力をもたなくても、水と油は分かれます。
- 分子動力学の結果は力場しだいです。結合が切れる化学反応は、つぶとばねのモデルでは扱えません。



2 電子を「見る」

第 1 章では、水分子の O はわずかにマイナス、H はわずかにプラスの電気を帯びていることを、前もって決まったこととして使いました。力場では、その電気の量をつぶに割りふります。O と H のあいだで、なぜ電気がかたよるのでしょうか。そのわけは、つぶとばねのモデルで省いた電子にあります。

電子は、光の色も決めています。花火の赤や黄色、スマホの画面の色、ホタルの黄緑色の光は、どれもその例です。花火の赤い光と黄色い光は、何がちがうのでしょうか。そもそも光の色は、何によって決まるのでしょうか。

◎ この章で分かること

- 電子は原子核のまわりに雲のように広がっていて、分子の中で電荷がどうかたよるかも、その広がり方で決まります
- コンピュータは、原子核の位置と電子の数から、電子の雲とエネルギーの段を計算し、分子の形や電荷のかたより、光の色を、分子を作る前に予測します
- 光の色は、原子や分子の中で電子がとれるエネルギーの段差で決まります。炎の原子は段差に見合う光を出し、色素の分子は段差に見合う光を受けとります

高校で習ったこと：炎色反応、電子配置、イオン化エネルギー、電気陰性度と極性（化学基礎）、光の波長（物理）。物理の内容は、習っていなくても読めます



花火の色って、火薬に色がついているの？

色を決めるのは、火薬に混ぜた金属の電子だ。電子を計算すれば、燃やす前に色が分かる。

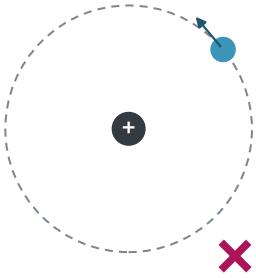


2.1 電子は雲のように広がる

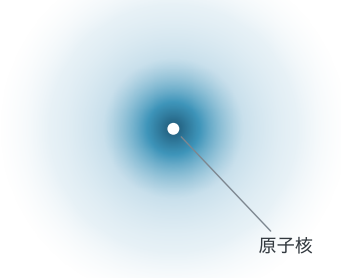
電気のかたよりも光の色も、もとをたどると電子に行き着きます。そこで、原子の中の電子が、どんな姿をしているのかを確かめておきます。原子核のまわりの決まった道を、電子が惑星のように回っている原子の絵を、見たことがあるかもしれません。

よくある誤解

「電子は、原子核のまわりの決まった道を、惑星のように回っている」と考えるのは誤りです。電子の居場所は、1 本の道としては決まりません。分かるのは、どのあたりで見つかりやすいかという広がりです（図 2-1）。惑星の絵は、K 殻、L 殻のような電子の層を数えるのには便利ですが、電子の広がり方を表してはいません。



惑星のように回る絵（誤り）



電子の雲（水素原子）

図 2-1　惑星のように回る絵（左）と、電子の雲（右）。右は、水素原子の電子が見つかりやすい所ほど濃く塗った図で、原子核を通る平面での切り口です。左右の大きさの比は合わせていません。

図 2-1 の右のように、電子は原子核のまわりに、雲のようにぼんやりと広がっています。原子核の近くほど、電子は見つかりやすく、遠くへ行くほど見つかりにくくなります。電子が見つかりやすい所ほど濃く塗ったこの広がりを、電子の雲と呼びます。

電子がとれるエネルギーは、とびとびの値に限られています。電子は、決まった高さの段にしかいられません。この段を、エネルギー準位（energy level）と呼びます。電子配置の K 殻、L 殻、M 殻も、この段を大まかに分けたものです。電子の雲の形は、段ごとにちがいます（図 2-2）。

電子の雲の形と、エネルギー準位の高さは、シュレディンガー方程式（Schrödinger equation）という式を解くと求まります。電子と原子核のあいだ、電子どうしのあいだには、静電気力（プラスとマイナスの電荷が引き合い、同じ符号の電荷どうしが反発し合う力）がはたらいています。シュレディンガー方程式は、これらの力のもとで、電子の雲がどんな形に広がり、どれだけのエネルギーをもつかを決める式です。この式を解いて、原子や分子の電子の雲とエネルギーを求める分野を量子化学（quantum chemistry）、コンピュータで解くことを量子化学計算といいます。

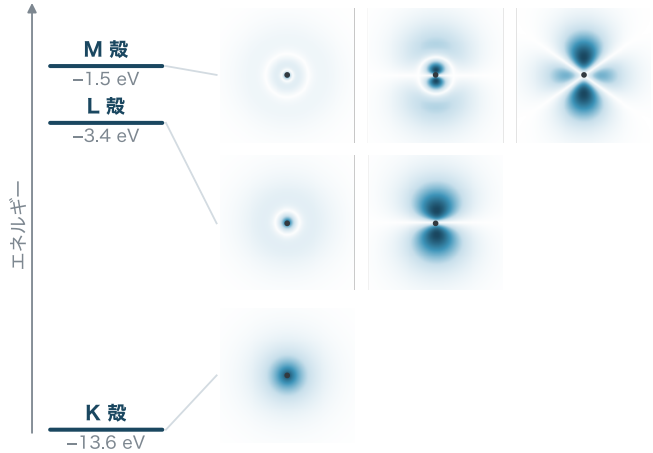


図 2-2 水素原子のエネルギー準位（左。高さは本当の値）と、段ごとの電子の雲（右）。雲は、シュレディンガー方程式の答えを原子核を通る平面で切ったもので、色の濃い所ほど電子が見つかりやすい所です。黒い点が原子核で、縮尺は段ごとにちがいます。

コンピュータに渡すのは、原子核の種類と位置、電子の数と、使う近似だけです（図 2-3）。返ってくるのは、電子の雲とエネルギー準位です。そこから、分子の形、電荷のかたより、光の色が分かります。この章では、この 3 つを順に見ていきます。

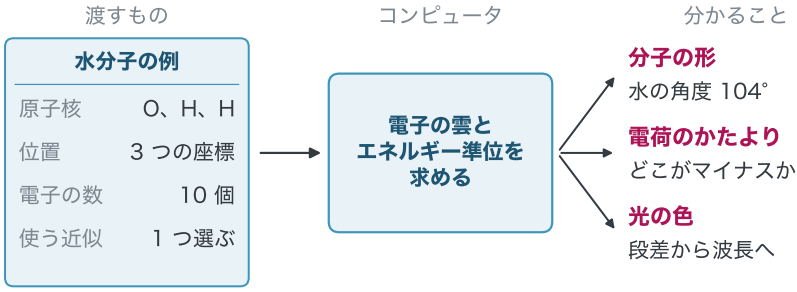


図 2-3　量子化学計算で、コンピュータに渡すもの（左）と、そこから分かること（右）。水の角度は、本書で計算した値です。

何をモデルにしたか

量子化学計算では、原子核と電子を、次のように扱います。

原子核は、電子よりずっと重く、陽子 1 個でも電子の約 1840 倍の質量があります（文献 2-1）。そのため、電子が動き回るあいだ、原子核はほとんど止まって見えます。そこで、原子核は決まった位置に止めておき、電子だけを計算します。第 1 章のつぶとばねとは逆に、ここでは原子核の動きのほうを省きます。

電子どうしのかかわり方は、そのまま計算するには複雑すぎるので、近似します。近似（approximation）とは、本当の答えに近く、計算しやすい答えを使うことです。よく使う近似では、それぞれの電子が、ほかの電子 1 個 1 個からではなく、ほかの電子の雲を平均した力を受けると考えます。

何をくり返したか

コンピュータは、何を手がかりにして、電子の雲を求めるのでしょうか。手がかりは、エネルギーです。電子の雲は、エネルギーがいちばん低くなる形に落ち着きます。そこでコンピュータは、雲の形を変えてはエネルギーを計算し、いちばん低くなる雲を探します。

電子のエネルギーは、2 つの分を足したものです。1 つは、原子核との引き合いの分（位置エネルギー）です。雲を原子核に近づけるほど、電子と原子核は静電気力で

強く引き合うので、この分は下がります。もう 1 つは、電子の動きの分（運動エネルギー）です。電子には、狭い所に閉じこめるほど激しく動くという性質があります。電子のようなごく小さな粒に特有の性質で、ここでは決まりとして使います。そのため、雲を縮めるほど、この分は上がります。雲の大きさは、下がる分と上がる分のかね合いで決まります。

手で 1 回、電子を 1 個だけもつ水素原子で試してみます。雲の形は 図 2-1 の右と同じく、原子核から離れるほど薄くなる形に決め、大きさだけを変えます。雲の大きさの目安を、雲の半径と呼ぶことにします。大きさを 3 通りに変えて、2 つの分と合計を計算しました。

予想してみよう　**電子と原子核は引き合っています。電子の雲をどんどん小さくしていくと、電子は原子核に落ちてしまうと思いますか？**

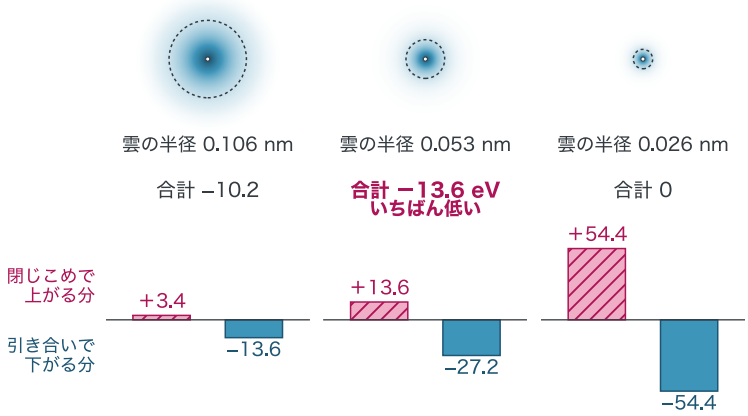


図 2-4　水素原子の電子の雲の半径と、エネルギー（eV、本書で計算した値）。上は雲の切り口で、同じ縮尺です。点線の円が雲の半径で、原子核からの距離のうち、電子がいちばん見つかりやすい距離を使いました。

答えは、落ちない、です。エネルギーの合計は下がり続けるのではなく、ある大きさでいちばん低くなり、それより小さくすると上がります（図 2-4）。図の真ん中と右を比べてみます。雲の半径を半分にすると、引き合いで下がる分は 2 倍になりますが、閉じこめで上がる分は 4 倍になります。雲を小さくしすぎると、上がる分のほうが勝ってしまうのです。

合計がいちばん低くなるのは、雲の半径が 0.053 nm のときで、そのエネルギーは −13.6 eV でした。このエネルギーは、電子が原子核から遠く離れて止まっているときを 0 として測っています。つまり、水素原子から電子を取り去るには 13.6 eV が必要です。これは、水素のイオン化エネルギー（13.6 eV、1 mol あたり 1312 kJ）と一致します（文献 2-2）。雲の大きさを、エネルギーがいちばん低くなるように選ぶだけで、実測の値を言い当てられました。

電子がたくさんあるとき

水素原子の電子は 1 個でしたが、水分子には電子が 10 個あり、マイナスの電荷どうしで押し合います。そのため、1 個の電子の雲がどこに広がるかは、ほかの電子の雲がどこにあるかで変わります。寒い日のたき火に集まる人たちに似ています。火に近づきたいけれど、混んだ所は避けたい。どこに立つのがよいかは、ほかの人の立つ場所で変わります。

そこでコンピュータは、図 2-5 のように、くり返し計算して解きます。まず、電子の雲を仮に置きます（①）。次に、その雲から電子が受ける力を計算し（②）、その力のもとで、水素原子のときと同じように、エネルギーがいちばん低くなる雲を求め直します（③）。雲が変わると力も変わるので、② と ③ を、雲が変わらなくなるまでくり返します。ただし、たき火の人たちとちがって、電子が実際にこう動くのではありません。これは、答えを探す計算の手順です。

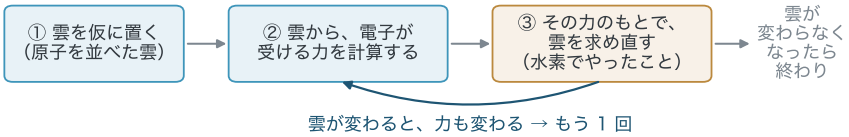


図 2-5　電子の雲を求めるくり返し。

水分子で、くり返しのようすを追ってみます（図 2-6）。はじめは、O 原子と H 原子の雲をそのまま並べた雲です。たき火でいえば、O は大きな火、2 つの H は小さな火です。1 回目の計算では、雲は O の原子核に強く引かれて、O のまわりに集まりました（O のまわりがローズ）。ところが、集まった電子どうしは押し合います。この押し合いは、集まる前の雲から計算した力には、まだ入っていませんでした。2 回目の計算ではそれが入るので、雲は O から押し戻されました（O のまわりが青）。こ

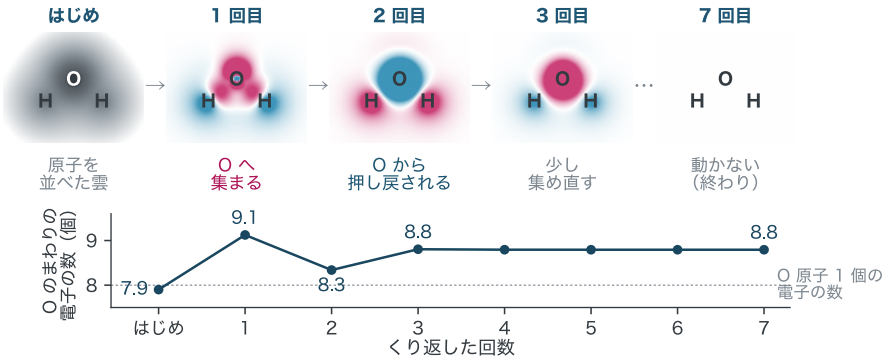


図 2-6 水分子で計算したときの、回ごとの雲の変わり方（本書で計算した値）。上の段の「はじめ」は、原子を並べた雲の切り口で、電子が見つかりやすい所ほど濃く塗りました。1 回目からは、その回に雲が変わった所を、分子の平面での切り口で塗りました。ローズは前の回より電子が増えた所、青は減った所です。下の段の数は、電子 10 個を原子ごとに分けたときの、O のまわりの電子の数で、分け方によって少し変わります。

うして行き来しながら、変わり方は小さくなっていきます。7 回目には、変わった所がなくなり、図が白くなりました。雲と力のつじつまが合ったのです。

最後に O のまわりにいる電子は 8.8 個で、O 原子 1 個の電子の 8 個より多くなりました。このかたよりは、2.2 節で見ます。



テオ 仮に置いては、求め直す。地道だが、水分子なら 7 回でつじつまが合う。$\sqrt{2}$ を 1.4、1.41 と詰めていくのと同じ進め方だ。

こうして得られた水分子の電子の雲が、図 2-7 です。原子核 3 個と電子 10 個を渡すと、コンピュータは、分子全体に広がる電子の雲を返します。O のまわりが特に濃くなっています。

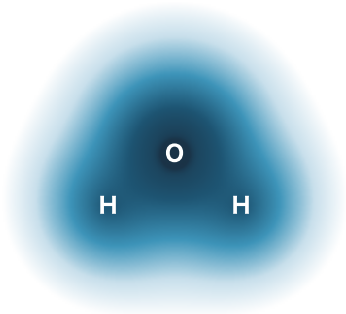


図 2-7　水分子の電子の雲（分子の平面での切り口）。濃い所ほど電子が見つかりやすく、原子核の近くは塗り方を調整しています。

原子核の位置を少しずつ動かして、エネルギーがいちばん低くなる位置を探せば、分子の形も求まります。水分子で計算すると、H–O–H の角度は 103.9° になりました。第 1 章で見た本物の水分子の 104.5° に近い値です。

要点

コンピュータは、原子核の種類と位置、電子の数を受けとり、電子の雲を仮に置いては、エネルギーがいちばん低くなるように求め直すことを、変わらなくなるまでくり返します。返ってくる電子の雲とエネルギー準位から、光の色や、分子の中の電荷のかたよりが分かります。

2.2 電荷のかたより

では、章の初めの問い「O と H のあいだで、なぜ電気がかたよるのか」に答えます。雲が濃いからといって、そこがマイナスに帯電しているとは限りません。O の原子核には、電子 8 個分のプラスの電荷があるからです。手がかりは電気陰性度です。電気陰性度は、結合をつくる電子を、原子がどれだけ強く引きつけるかの目安です（図 2-8）（文献 2-3）。

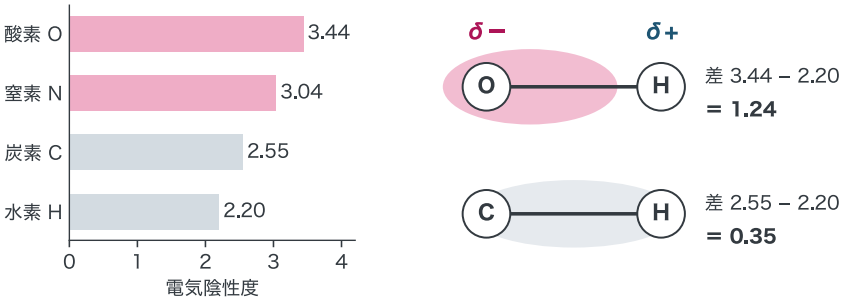


図 2-8 H・C・N・O の電気陰性度（左）と、O–H・C–H の結合の電荷のかたより（右）。右の楕円は結合の電子を表し、O–H では O の側に寄っています（文献 2-3）。

O の電気陰性度は 3.44 で、H の 2.20 よりずっと大きく、差は 1.24 です。そのため、O と H が共有している結合の電子は O の側に引き寄せられ、O はわずかにマイナス（δ−）に、H はわずかにプラス（δ+）に帯電します。これが、極性のある結合です。結合の電子がこうして一方に引き寄せられて生じる電気のかたよりを、電荷のかたよりと呼びます。一方、C と H の差は 0.35 しかないので、C–H の結合には、電荷のかたよりがほとんどありません。

電気陰性度は、原子ごとの大まかな目安です。電子の雲を計算すれば、分子のどこがどれだけかたよっているかを確かめられます。雲がごく薄くなる所を分子の表面と決め、表面のそれぞれの場所で、電子の雲のマイナスと原子核のプラスのどちらの影響が強いかを計算して、色で塗り分けるのです。

予想してみよう　**水（O–H）、アンモニア（N–H）、メタン（C–H）の表面を、マイナスに寄る所はローズ、プラスに寄る所は青で塗ります。メタンはどう見えると思いますか？**

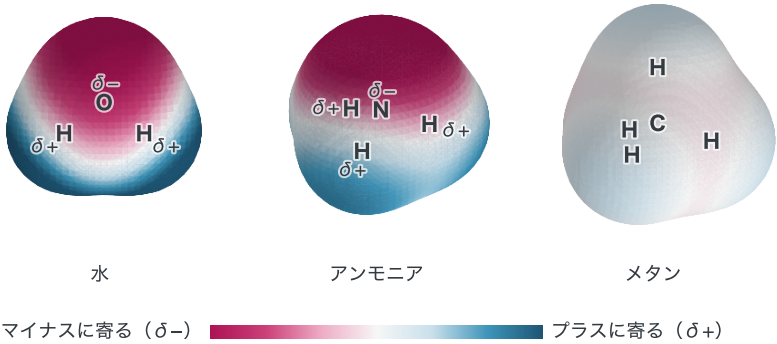


図 2-9　量子化学計算で、電子の雲の表面を電荷のかたよりで塗り分けた図。原子の記号は、その奥にある原子核の位置です。3 つの分子で同じ色の目盛りを使っています。

答えは、ほとんど白です。メタンの表面には、目立ったかたよりがありません（図 2-9）。水では O の側がローズ（δ−）、2 つの H の側が青（δ+）になりました。O の側では電子の雲のマイナスが原子核のプラスに勝ち、H の側ではプラスが勝っています。O と H が共有する電子が O の側へ引き寄せられたようすが、計算で目に見える形になったのです。アンモニアでも、N の側がローズ、3 つの H の側が青です。第 1 章の力場でつぶに割りふった δ− と δ+ を、電子の雲から確かめたことになります。力場でつぶに割りふる電気の量も、実験の結果や、こうした電子の計算に合うように決めてあります。

この電荷のかたよりが、第 1 章で見た水素結合のもとです。水分子のプラスに寄った H は、となりの水分子のマイナスに寄った O と、静電気力で引き合います。水素結合をつくれない油の分子が水の外へ押し出されるのは、このためです。O–H や N–H をもつ分子は水になじみやすく、C–H ばかりの分子は油になじみやすくなります。このちがいは、薬が体の中で水と油のどちらになじむか、という第 4 章の話につながります。電荷のかたよりは、薬と酵素が引き合う力のもとにもなります（図 2-10）。

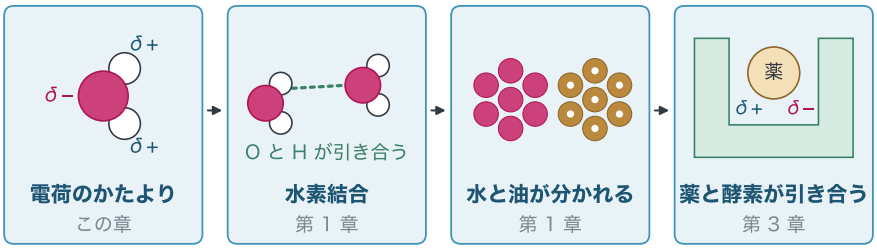
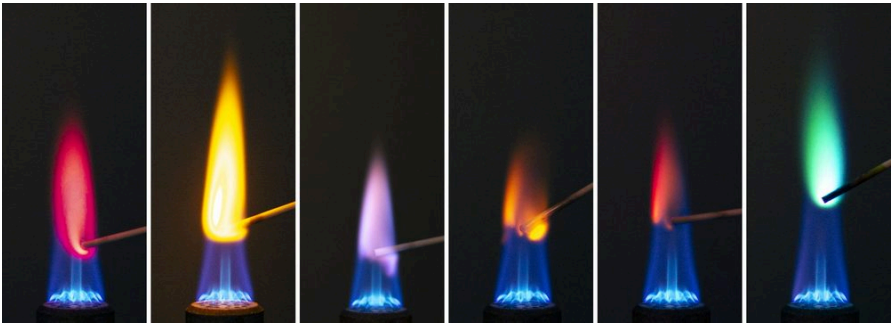


図 2-10　電荷のかたよりから、薬と酵素の引き合いまでのつながり（模式図）。

2.3 光の色も、作る前に分かる

電子の雲とエネルギー準位が分かると、光の色も分かります。ナトリウムをふくむ水溶液を白金線につけて、ガスバーナーの炎に入れると、炎は黄色になります。リチウムなら赤、銅なら青緑です。炎の色は、入れた元素によって決まっています（図 2-11）。これが炎色反応で、花火の色も、この反応を使って出しています。



リチウム	ナトリウム	カリウム	カルシウム	ストロンチウム	銅
赤	黄	赤紫	橙赤	紅	青緑

図 2-11　炎色反応の写真。試料をつけた棒を、ガスバーナーの炎に入れました。炎の下の青い部分は、ガスそのものが燃える炎です。（写真：S k y r, CC BY 4.0）

光は波の一種で、波の山から山までの長さを波長といいます。光の色は、この波長で決まります。人の目に見える光の波長は、紫のおよそ 380 nm から、赤のおよそ 780 nm までです。1 nm（ナノメートル）は、100 万分の 1 mm です。ナトリウムの炎からは 589 nm の黄色の光が、リチウムの炎からは 671 nm の赤の光が出ます（文献

2-2)。元素ごとに決まった波長の光が出るのは、電子のエネルギーの段が、元素ごとに決まっているからです。

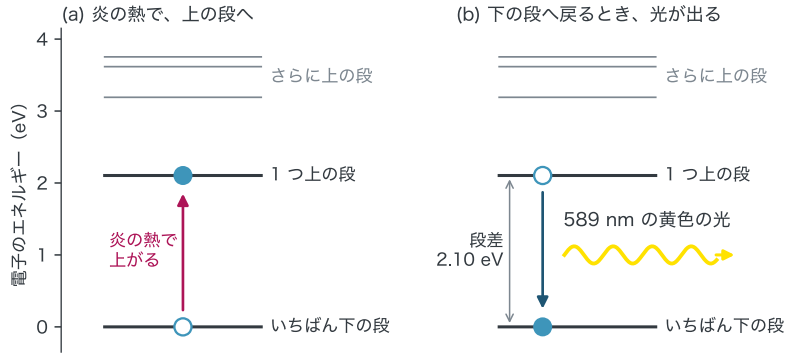


図 2-12 ナトリウム原子で、いちばん外側にある電子のエネルギー準位（横線）。丸が電子で、白抜きの丸は電子が移る前の位置です（文献 2-2）。

炎の中で、ナトリウム原子のいちばん外側の電子は、熱からエネルギーを受けとって上の段へ上がります（図 2-12 の (a))。上の段は不安定なので、電子はすぐに下の段へ戻ります。このとき、上下の段のエネルギーの差（段差）に、ちょうど等しいエネルギーをもつ光が出ます（(b))。電子 1 個のエネルギーはとても小さいので、電子ボルト（eV）という単位で表します。1 eV は、電子 1 個を 1 V の電圧で加速したときに得るエネルギーです。ナトリウムの段差は 2.10 eV で、出てくる光が 589 nm の黄色です。

光 1 つ分のエネルギーと、その光の波長は反比例します。段差を eV で、波長を nm で表すと、2 つをかけた値は、いつも約 1240 になります。つまり、段差が分かれば、割り算 1 回で光の波長が求まります。

$$\text{波長（nm）} = 1240 \div \text{段差（eV）}$$



テオ みそ汁がふきこぼれると、コンロの炎が黄色くなる。あれも食塩のナトリウムが出す光だ。台所でも炎色反応は起きている。

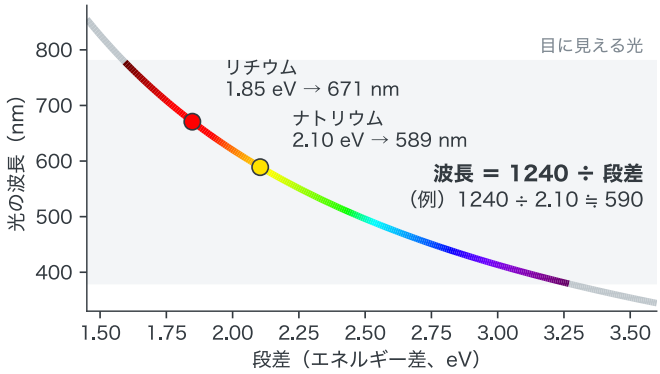


図 2-13 段差と光の波長の関係。曲線の色は、その波長の光の色です。灰色の部分は、目に見えない光です。丸は、ナトリウムとリチウムの実測の値です（文献 2-2）。

リチウムの段差は 1.85 eV です。段差が大きいほど、光の波長は短くなり、色は赤から黄、緑、青、紫の側へ移っていきます（図 2-13）。

手で 1 回、計算してみます。1240 をナトリウムの段差 2.10 eV で割ると、1240 ÷ 2.10 = 590 です。589 nm の黄色の光を、ほぼそのまま言い当てられました。

ただし、ここで使った段差は、炎の光を測った結果から逆に求めた値です。そこで、段差をコンピュータで求めます。ナトリウム原子の段差を計算すると 2.09 eV で、割り算をすると 593 nm の黄色になりました。実測の 589 nm に近い値が、炎に入れる前に得られます。

分子でも同じです。電子の雲とエネルギー準位を計算すれば、光とやりとりするエネルギーの段差が分かり、割り算で光の波長が分かります。原子核の位置も、計算で分子の形を求めて決められるので、まだ作っていない分子でも計算できます。

例として、炭素の鎖に二重結合と単結合が交互に並んだ分子を考えます。二重結合が 1 個のエチレン（C_2H_4）から、鎖を 1 段ずつ伸ばして（図 2-14）、二重結合が 8 個の分子まで、段差に見合う光の波長をコンピュータで計算しました。

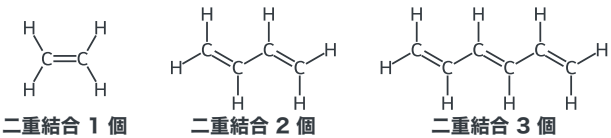


図 2-14 炭素の鎖のはじめの 3 つ（左がエチレン）。C と H をすべて書きました。

予想してみよう **炭素の鎖を長くしていくと、段差に見合う光の波長は、長くなると思いますか、短くなると思いますか？**

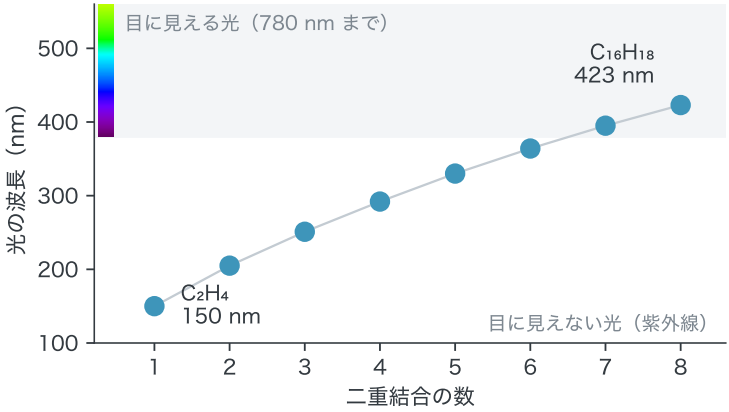


図 2-15 二重結合と単結合が交互に並んだ炭素の鎖の、段差に見合う光の波長（本書で計算した値）。左の色の帯は目に見える光の範囲で、帯の色は、その波長の光の色です（分子そのものが見える色ではありません）。

答えは、長くなる、です。鎖が長くなるほど、電子の雲が鎖にそって遠くまで広がれるようになり、段差が小さくなります。段差が小さくなれば、割り算の答えの波長は長くなります。エチレンでは 150 nm で、目に見えない紫外線です。鎖を伸ばしていくと、波長は少しずつ長くなり、二重結合が 7 個の分子（395 nm）で、目に見える光の紫の端に入りました（図 2-15）。

段差に見合う光が目に見える光の範囲に入ると、分子は色づいて見えます。ただし、光のやりとりの向きは、炎とは逆です。炎の中の原子は、段差に見合う光を出すので、その光の色に光ります。色素の分子は、日光のような白い光の中から段差に見合う光を受けとり、私たちは残りの光の色を見ます。植物に多い色素の β-カロテンは、二重結合が 11 個つながった長い炭素の鎖をもち（文献 2-4）、青い光を受けとるので、赤橙色に見えます。



テラ にんじんの橙色も、β-カロテンの色だよ！ 長い鎖が青い光を受けとるから、残りの橙色が目に届くんだ。

作る前に、光る色を選ぶ

スマホやテレビの画面には、分子に電気を流して光らせる、有機 EL というしくみを使ったものがあります。2016 年に報告された研究では、光らせる分子の候補を 160 万個考え、そのうちの 40 万個を超える分子の光り方を、量子化学計算で求めました（図 2-16）（文献 2-5）。見つかった有望な分子の中から選んだ分子だけを実際に作り、光らせて確かめています。



図 2-16　有機 EL で光らせる分子を、計算でふるいにかけた例（2016 年に報告）。帯の幅は形だけで、数の比ではありません（文献 2-5）。

40 万個の分子を計算するのは、簡単ではありません。量子化学計算は、電子の数が増えると、急に重くなるからです。炭素の鎖の分子で、電子の雲を求める計算にかかった時間を、著者のパソコンで測りました。電子が 16 個のエチレンは 0.5 秒ほどで終わりますが、電子が 170 個の分子では 4 分近くかかりました。電子の数が 2 倍になると、時間はおよそ 5 倍になります。

電子の雲を求めるだけで 1 分子 4 分かかるとすると、40 万個の分子を 1 個ずつ順に計算すると、休まず続けて 3 年ほどになります。第 4 章で見る、分子の性質をもっと簡単に見積もる計算は 1 分子 0.0002 秒ほどなので、それに比べると桁ちがいに重い計算です。タンパク質をまるごと計算するのは、今のコンピュータでも大変です。分子が大きくなると時間がどう増えるかは、第 6 章で測って確かめます。

どこまで信じられるか

量子化学計算の答えは、選んだ近似によって変わります。図 2-17 は、ナトリウム原子の段差を、電子どうしのかかわりの見積もり方がちがう 3 つの近似で計算し、波長にしたものです。

実測の 589 nm に対して、計算の答えは 557 nm から 628 nm まで分かれました。近似の選び方によって、計算した色は、黄緑から橙まで変わります。原子 1 個でも

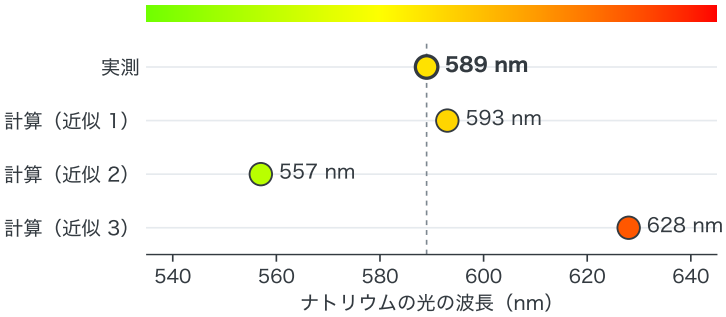


図 2-17 ナトリウム原子の段差を 3 つの近似で計算して求めた光の波長と、実測の波長（文献 2-2）。近似 1〜3 は、電子どうしのかかわりの見積もり方がちがう方法です。

これだけちがうので、分子の計算では、実測と比べて、その種類の分子によく合う近似を選んで使います。

ここでの計算は、分子のまわりの水や、熱による分子の動き（第 1 章）も省いています。第 6 章では、ほかの計算とあわせて、計算がどれだけ当たるかを答え合わせします。

電荷のかたよりは、薬の分子とウイルスのタンパク質が引き合う強さにもかかわります。では、何億個もの分子の中から、ウイルスのタンパク質と強く引き合う分子を、作って試す前にコンピュータで選べないでしょうか。

研究の現場から：1 億年前のホタルの光

ホタルの光は、ルシフェリンという分子が、ルシフェラーゼという酵素のはたらきで反応して出る光です。光の色は、ルシフェラーゼの形によって決まります（文献 2-6）。

2020 年、中部大学の大場裕一さんたちは、今のホタルの遺伝子をもとに、祖先のホタルのルシフェラーゼを復元しました。今のホタルに共通する祖先は、約 1 億年前に現れたと考えられています。復元した祖先の酵素を光らせて測ると、光がいちばん強い波長は 548 nm の緑でした。今のゲンジボタルの酵素が出す光（562 nm）より、少し短い波長です（図 2-18）。祖先の酵素では、光る分子が入るすき間が狭く、光る分子が動きにくくなっていることも、酵素の原子の並びを調べて分かりました。

なぜ色が変わるのかは、量子化学計算で調べられています。光る分子と、まわりのアミノ酸をいっしょに計算した研究では、計算した光の色が実測に近くな

り、まわりのアミノ酸などの電荷が、光の色を決めるうえで大事なはたらきをしていることが示されました（文献 2-7）。祖先と今の酵素を同じように計算して比べれば、どのアミノ酸の入れかわりが、緑から黄緑へ色を変えたのかを調べられます。電荷のかたよりは、タンパク質の中の分子のふるまいも左右しているのです。

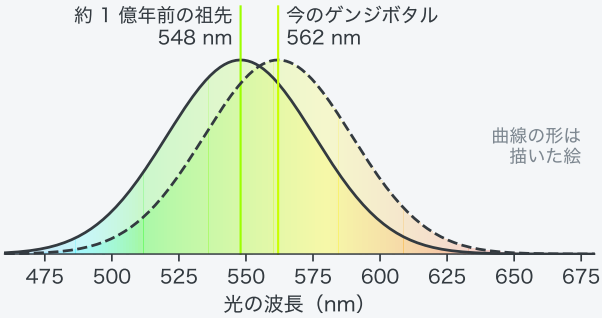


図 2-18　約 1 億年前のホタルの祖先と、今のゲンジボタルの酵素が出す光（文献 2-6）。曲線がいちばん高い所の波長は論文の値で、曲線の形は描いた絵です。

まとめ

- 電子は惑星のように回っているのではなく、原子核のまわりに雲のように広がっています。これを計算で求めるのが量子化学です。
- コンピュータは、電子の雲をエネルギーがいちばん低くなるように求め直すことを、変わらなくなるまでくり返します。電子の雲から、分子の形と電荷のかたより、光の色が分かります。
- 計算で塗り分けると、水の O はマイナス、H はプラスに寄り、メタンにはほとんどかたよりがありません。これが水素結合のもとです。
- 光の色は、エネルギー準位の段差で決まります。段差（eV）と波長（nm）をかけると約 1240 なので、段差が分かれば割り算で波長が分かります。
- 作る前に分子の光の色を予測できますが、電子が増えると計算は急に重くなり、近似の選び方で計算した色も変わります。



3 薬の候補を「探す」

インフルエンザの薬は、どうやって生まれるのでしょうか。インフルエンザの薬にかぎらず、薬は、何十万個もの化合物を作って試した中から、ようやく見つかります。新しい薬が 1 つできるまでには、9〜17 年かかります（文献 3-1）。

候補の分子を作って試す前に、当たりそうな分子を選べないでしょうか。第 1 章と第 2 章では、分子の動きと電子をコンピュータで調べました。この章では、その考え方を薬の候補探しに生かし、コンピュータが何億個もの分子を試して候補を絞りこむしくみを紹介します。

◎ この章で分かること

- ウイルスの酵素で反応が起こる場所（活性部位）に結合する分子は、酵素のはたらきを止め、薬の候補になります
- 分子が活性部位に結合するには、形が合うことと、引き合うことが必要です
- コンピュータは、分子を活性部位に置いて結合の強さを点数にする計算をくり返し、何億個もの分子から、作って試す候補を絞りこみます

高校で習ったこと：酵素と基質（生物基礎）、アミノ酸とタンパク質（化学）。化学の内容は、習っていなくても読めます

何十万個を作って、1 つずつ試す。それでは時間がかかりすぎる。





だから、作る前にコンピュータの中で試すんだよ！

3.1 インフルエンザウイルスが増えるしくみ

薬の候補を探す前に、何を止めればウイルスが広がらなくなるのかを知る必要があります。インフルエンザウイルスが細胞に入り、増えて、外へ出ていくまでを追ってみます（図 3-1）。

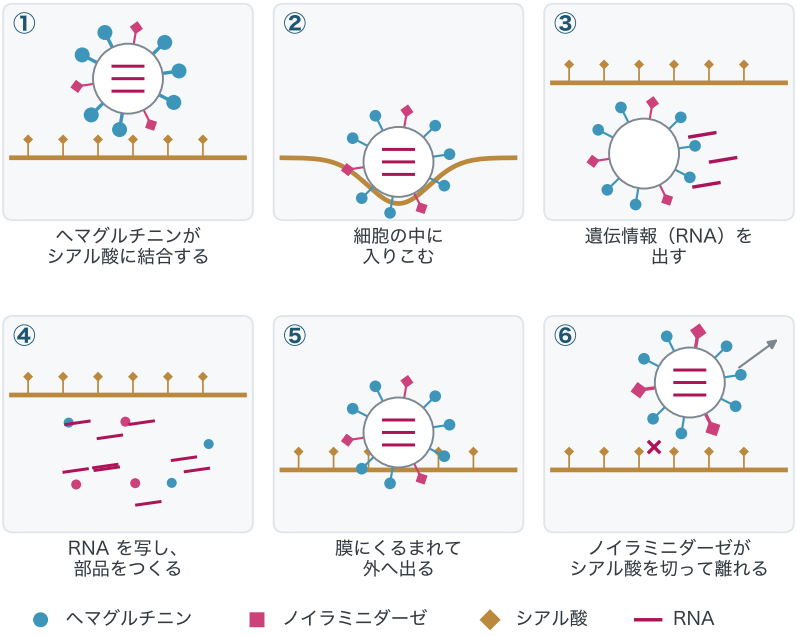


図 3-1　インフルエンザウイルスが増えるサイクル。① でヘマグルチニン、⑥ でノイラミニダーゼがはたらきます。ウイルスと細胞は、簡単にした絵です。

ウイルスの表面には、2 種類のタンパク質が突き出ています（図 3-2）。細胞の表面には糖の鎖がたくさん生えていて、その端にはシアル酸という分子がついています。1 つ目のタンパク質は、このシアル酸に結合して、ウイルスを細胞につなぎとめます。これをヘマグルチニンと呼びます（図 3-1 の ①）。

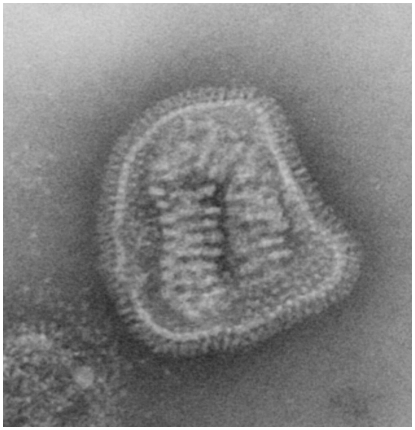


図 3-2 電子顕微鏡で見たインフルエンザウイルス。表面にとげのように並んでいるのが、ヘマグルチニンとノイラミニダーゼです。（写真：Cynthia Goldsmith（CDC）, パブリックドメイン）

細胞の中で増えたウイルスは、細胞の膜をかぶって外へ出てきます。ところが、新しいウイルスのヘマグルチニンも細胞の表面のシアル酸に結合してしまうので、このままではウイルスは細胞から離れられません。そこで、もう 1 つのタンパク質であるノイラミニダーゼが、細胞の表面にある糖の鎖からシアル酸を切りはなします。こうしてウイルスは細胞から離れ、まわりの細胞へ広がっていきます（図 3-1 の ⑥）（文献 3-2）。

ノイラミニダーゼのように、ほかの分子の反応を進めるタンパク質を酵素と呼びます。酵素が反応させる相手の分子は、基質といいます。ノイラミニダーゼの基質は、端にシアル酸がついた糖の鎖で、ノイラミニダーゼは、シアル酸と糖の鎖をつなぐ結合を切ります。

酵素の表面には、基質を受けとめて反応を進めるくぼみがあります。この場所を活性部位（active site）と呼びます。ノイラミニダーゼの活性部位に別の分子が結合して、シアル酸が入れなくなれば、ノイラミニダーゼはシアル酸を切れません（図 3-3）。新しいウイルスは細胞の表面に結合したままになり、まわりへ広がれなくなります。

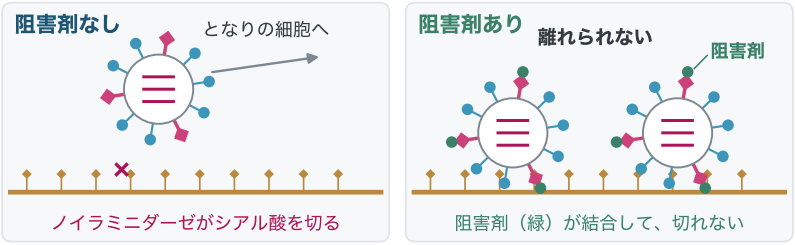


図 3-3　ノイラミニダーゼのはたらき（左）と、阻害剤のはたらき（右）。× は、ノイラミニダーゼが切ったシアル酸です。

このように、酵素のはたらきを止めることを阻害といい、止める分子を阻害剤と呼びます。つまり、インフルエンザの薬を探すには、ノイラミニダーゼの阻害剤を探せばよいのです。

3.2 薬の候補を実験で試す

ある分子がノイラミニダーゼの阻害剤になるかどうかは、実験で確かめられます。ノイラミニダーゼに切られると光るように、基質に目印の分子をつないでおき、2 つを試験管で混ぜます。光の強さを測れば、基質がどれだけ速く切られているかが分かります（図 3-4）。

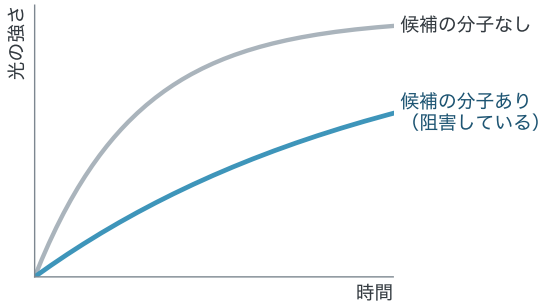


図 3-4　阻害の実験で、時間がたつにつれて光が強くなるようす。考え方の図なので、目盛りの値は描いていません。

候補の分子を入れたとき、光が強くなる速さが落ちれば、その分子はノイラミニダーゼを阻害しています。阻害の強さは、どれだけの量を入れれば速さが半分になるかで比べます。

ただし、この実験をするには、候補の分子を 1 つずつ作らなければなりません。冒頭で見たように、薬は何十万個もの化合物を試した中から生まれます。日本の製薬会社が 2015〜2019 年の 5 年間に、あらゆる病気の薬の候補として合成した化合物は、545,967 個です（文献 3-1）。

予想してみよう　**545,967 個の化合物のうち、薬として承認されたのは何個だと思いますか？**

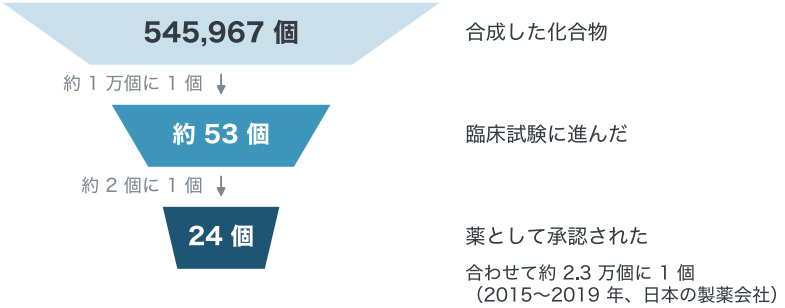


図 3-5　薬ができるまでのじょうご（2015〜2019 年、日本の製薬会社。あらゆる病気の薬の合計）。帯の幅は形だけで、数の比ではありません（文献 3-1）。

答えは 24 個です（図 3-5）。およそ 2.3 万個を作って、ようやく 1 個が薬になる計算です。作る前に、当たりそうな分子をコンピュータの中で選べれば、合成する分子の数を減らせます。そのためには、分子が活性部位に結合するとはどういうことかを、はっきりさせる必要があります。

3.3 活性部位に結合する 2 つの条件

酵素と基質の関係は、よく鍵と鍵穴にたとえられます。鍵穴に入って錠を回せるのは、決まった形の鍵だけです。同じように、活性部位の形に合う分子だけが、活性部位に入れます。

ただし、形が合うだけでは足りません。活性部位に入った分子は、まわりの原子と引き合わないと、すぐに外へ出ていってしまいます。分子が活性部位に結合するには、形が合うことと、引き合うことの 2 つが必要です（図 3-6）。

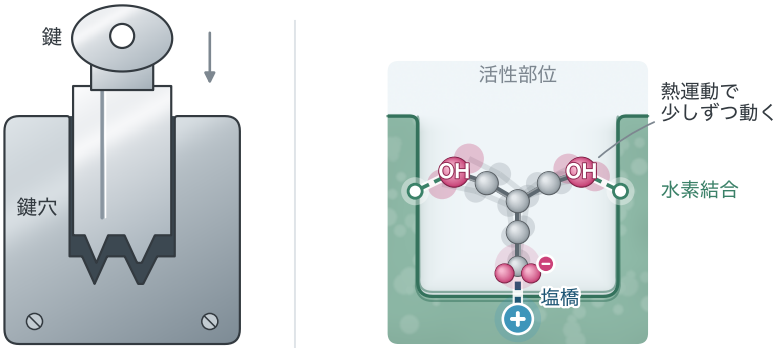
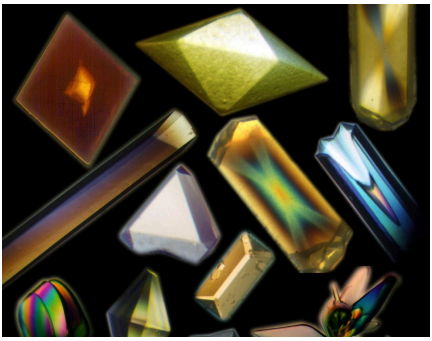


図 3-6　活性部位に結合する 2 つの条件。鍵と鍵穴のたとえ（左。断面の絵）は、形が合うことだけを表します。右は、タンパク質（緑の地）の表面の活性部位に分子が結合したようすの模式図です。玉の色は元素（灰が C、ローズが O）、緑の点線は水素結合、青の太い点線は塩橋（プラスとマイナスのイオンの引き合い。このあと説明）です。うすく重ねた形は、熱運動で少し動いたときの位置です。

鍵と鍵穴のたとえには、合わない所もあります。金属の鍵穴とちがって、本当の活性部位も分子も、熱運動で少しずつ形を変えながら動いています（図 3-6 の右のうすい形。第 1 章）。

どの原子がどこにあるかが分からなければ、形が合うかどうかも、どこで引き合うかも分かりません。タンパク質の中で原子が 3 次元に並んだようすを、立体構造と呼びます。立体構造は、タンパク質の結晶に X 線を当てて調べる X 線結晶構造解析という方法で決めます。ノイラミニダーゼの立体構造は 1983 年に決まり、活性部位の形が分かりました（文献 3-3）。

図 3-7 の (a) は、X 線結晶構造解析に使うタンパク質の結晶、(b) は、こうして決めたノイラミニダーゼの立体構造です。



(a) タンパク質の結晶



(b) ノイラミニダーゼの立体構造

図 3-7 (a) X 線を当てて調べるタンパク質などの結晶。偏光で撮ったので色がついて見えます。(b) H5N1 型のウイルスのノイラミニダーゼ（立体構造のデータベース PDB の番号 2HU4（文献 3-4)）。同じ鎖が 4 本集まってできています。原子の表面を描き、4 本の鎖を 4 通りの青で塗り分けました。ローズの粒は、4 つの活性部位に結合した薬です（薬の名前は 3.6 節で紹介します)。（(a) の写真：NASA Marshall Space Flight Center（結晶：Alex McPherson）, パブリックドメイン）

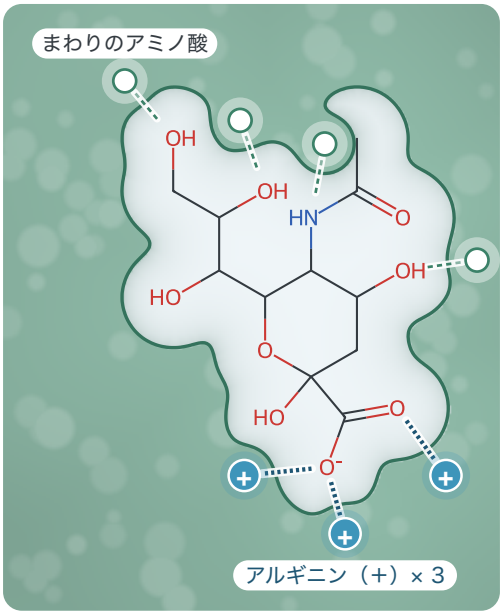
活性部位で分子が引き合うしくみのうち、この章では主な 2 つを取り上げます。1 つは第 1 章で見た水素結合、もう 1 つは第 2 章で出てきた静電気力（プラスとマイナスの電荷が引き合う力）です。実際には、炭素と水素の多い部分どうしが集まることも、結合を助けます（3.6 節)。

ノイラミニダーゼの活性部位には、プラスの電荷をもつアルギニンというアミノ酸が 3 つ並んでいます。シアル酸のカルボキシ基（COOH）は、体の中の水では H^+ を放して、マイナスの電荷をもつ COO^- になっています。この COO^- が 3 つのアルギニンと引き合うことで、シアル酸は活性部位に結合します（文献 3-5)。このような、プラスとマイナスのイオンどうしの結びつきを、塩橋（salt bridge）と呼びます（図 3-8)。塩橋では、静電気力と水素結合がいっしょにはたらきます。

図 3-8 のシアル酸は、構造式の書き方を簡単にしてあります。分子が大きくなると、C と H をすべて書いた構造式は読みにくくなります。そこで、C と、C についた H を省きます。線の折れ目と端には C があり、C の手のうち線で書いていない



テオ 第 2 章では、電子の広がり方で電荷が決まるのを見た。その電荷どうしの引き合いが、ここでは結合の決め手になる。



シアル酸

図 3-8　ノイラミニダーゼの活性部位に結合したシアル酸（模式図）。緑の地がタンパク質、淡い水色の所が活性部位です。シアル酸は、線の折れ目と端に C がある書き方（骨格構造式。次の段落で説明）で描いています。青の太い点線が塩橋、緑の点線が水素結合、白い丸がまわりのアミノ酸です。

分には H がついていると考えます。O や N と、O や N についた H は、そのまま書きます。このような書き方を、骨格構造式と呼びます（図 3-9）。この本では、ここから先の構造式を骨格構造式で描きます。

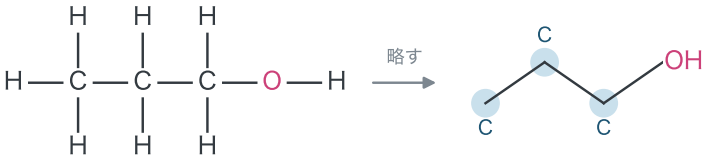


図 3-9　1-プロパノールの構造式の 2 つの書き方。右の骨格構造式では、青い丸の場所（線の折れ目と端）に C があります。

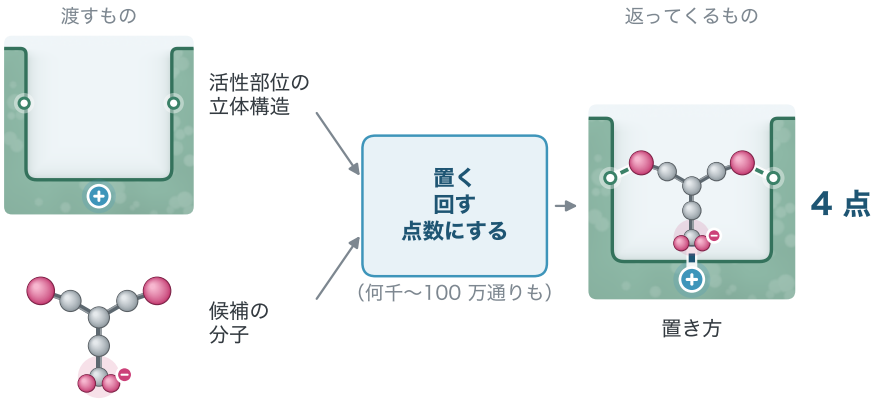


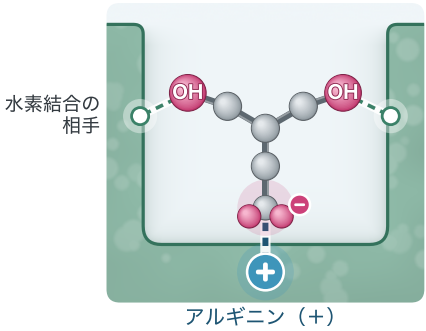
図 3-10　ドッキングで、コンピュータに渡すもの（左）と、返ってくるもの（右）。活性部位と分子は、考え方を見せるための簡単な絵で、線と玉の意味は 図 3-6 と同じです。

3.4 コンピュータの中で 1 つ試す：ドッキング

形が合うかどうかと、どこで引き合うかは、原子の位置が分かれば計算できます。そこで、候補の分子を活性部位に置いてみて、結合の強さを点数にする計算を考えます。これをドッキング（docking）と呼びます。

ドッキングでは、コンピュータに 2 つのものを渡します（図 3-10）。1 つは、X 線結晶構造解析で決めた、活性部位の立体構造です。もう 1 つは、試したい候補の分子です。分子は、単結合のまわりで原子のまとまりが回るので、形を少しずつ変えられます。コンピュータは、候補の分子を活性部位の中に置き、向きや形を変えながら結合の強さを点数にして、いちばん点数の高い置き方と、その点数を返します。

点数の付け方を、手で 1 回やってみます。ここでは、考え方が分かるように、簡単な点数の決め方を使います。水素結合が 1 本できるごとに 1 点、塩橋が 1 本できるごとに 2 点を加えます。原子どうしがぶつかって重なる所が 1 か所あるごとに、3 点を引きます。



点数		
水素結合	2 本 × +1	+2
塩橋	1 本 × +2	+2
ぶつかり	0 か所 × −3	0
合計		4 点

図 3-11　手で 1 回、点数を付ける。緑の点線が水素結合、青の太い点線が塩橋です。点数は本文の簡単な決め方で数えました。

図 3-11 の置き方では、水素結合が 2 本、塩橋が 1 本でき、ぶつかりはありません。点数は 1＋1＋2＝4 点です。同じ分子でも、向きを変えると、引き合う相手が入れかわったり、壁にぶつかったりして、点数が上下します（図 3-12）。

実際のドッキングでも、考え方は同じです。ただし、点数はもっとていねいな式で求めます。形の重なり、水素結合、静電気力などの項を、原子どうしの距離から計算して足し合わせます。この点数の式を、スコア関数と呼びます。実際のソフトの多くは、点数をエネルギーとして出し、小さい（マイナスに大きい）ほど強く結合すると見積もります。この本では、分かりやすくするために、高いほどよい点数にしました。

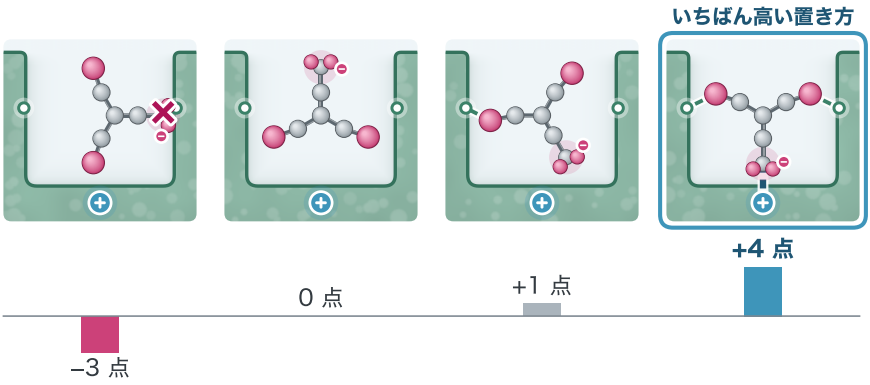


図 3-12　同じ分子の、置き方のちがう 4 通りと、その点数（点数の低い順）。× は壁にぶつかった原子です。

候補を 4 つ用意して、それぞれの点数を比べてみます（図 3-13）。

予想してみよう **大きすぎる分子、小さすぎる分子、形は合うが引き合わない分子、形も引き合いも合う分子。点数がいちばん高いのはどれだと思いますか？**

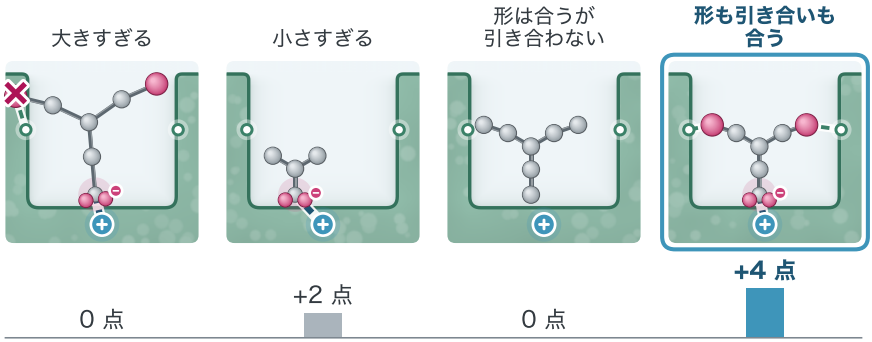


図 3-13 4 つの候補を、それぞれいちばん点数の高い置き方で置いたときの点数。分子の全体が活性部位に入る置き方だけを試しました。

答えは、形も引き合いも合う分子です。大きすぎる分子は活性部位の壁にぶつかり、小さすぎる分子は、水素結合の相手のいる壁まで、原子が届きません（OH を付ける場所もありません）。この簡単な決め方では、形がぴったりでも、引き合う原子をもたない分子の点数は上がりません。

3.5 何億個をふるいにかける：バーチャルスクリーニング

ドッキングでは、向きと形を変えながら、たくさんの置き方を試します。2019 年に報告された研究では、細菌の酵素へのドッキングで、分子 1 個あたり平均で約 4,000 の向きと、それぞれの向きで約 280 の形を試しました（文献 3-6）。合わせると、コン



テラ 手で 1 回やってみたから、しくみはもう分かったね！ あとはこれを、コンピュータが 100 万回以上くり返すだけ。

ピュータは分子 1 個につき 100 万通りを超える置き方を点数にしています。図 3-11 で手で数えた 1 回を、100 万回以上くり返すのです。

この計算を、たくさんの分子でくり返します。候補になる分子を集めておき、1 つずつドッキングして、点数の高いものだけを残します。このコンピュータの中でのふるい分けを、バーチャルスクリーニング（virtual screening）と呼びます（図 3-14）。

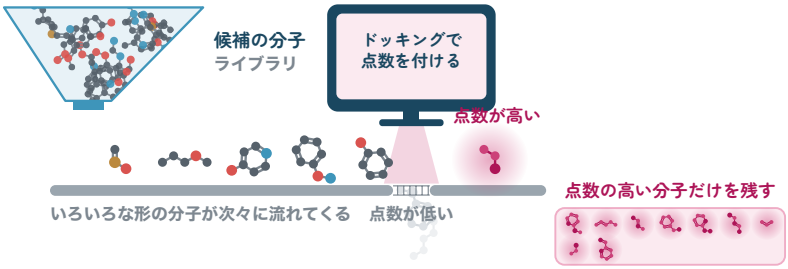


図 3-14　バーチャルスクリーニングのイメージ（簡単にした絵）。分子の形と数は、本物のものではありません。

2019 年の研究では、注文すれば作れる 1 億個を超える分子を集め、細菌の酵素と、もう 1 つのタンパク質の 2 つを相手にドッキングしました（文献 3-6）。細菌の酵素では、9,900 万個を試しています（図 3-15）。

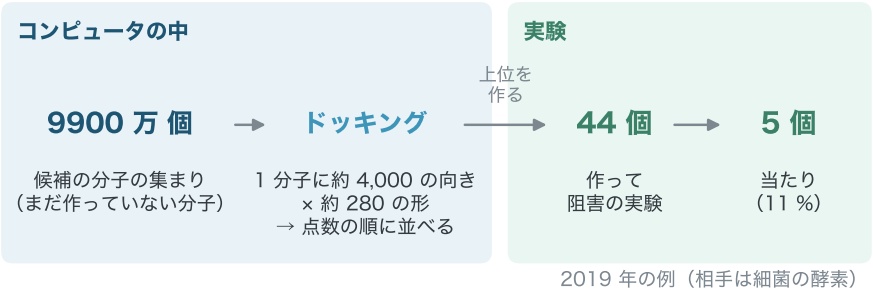


図 3-15　バーチャルスクリーニングのふるい（2019 年の例）（文献 3-6）。

2020 年には、10 億個を超える分子をふるいにかけた報告もあります（文献 3-7）。その研究では、分子 1 個のドッキングに平均 15 秒かかるとして、全体の時間を見積もっています。コア（計算をする部品）1 個では 475 年かかりますが、1 万個のコア

を同時に使えば、約 2 週間で終わる計算です。

ここでふるいにかけた分子の多くは、まだ作られたことのない分子です。2019 年の計算では、注文すれば作れることが分かっている分子を集めて、ライブラリにしました。コンピュータの中で選んでから作るので、何億個もの分子をすべて合成する必要はありません。合成するのは、点数が上位の分子だけです。

では、点数が上位の分子は、本当に阻害剤になるのでしょうか。細菌の酵素では、点数の上位から選んで作った分子を実験で試し、その中から阻害剤が見つかりました（図 3-15）。ただし、上位の分子でも、8〜9 割は外れです（文献 3-6）。計算では当たりに見えたのに、実験では結合しない分子を、偽の当たり（偽陽性）と呼びます。ドッキングでは、分子の置き方は比較的よく当たりますが、結合がどれだけ強いかや、分子の順位はあまり当たりません（文献 3-8）。そのため、バーチャルスクリーニングで候補を絞り、最後は実験で確かめます。

偽の当たりが出る理由は、モデルが省いたものにあります（図 3-16）。どのくらい当たるのかは、第 6 章で、ほかの計算もあわせて答え合わせをします。

3.6 計算から生まれた薬

実は、今も使われているインフルエンザの薬のうち 2 つは、活性部位の立体構造をもとに、コンピュータを使って設計されました。

1 つ目は、ザナミビル（商品名リレンザ）です。研究者たちは、ノイラミニダーゼの立体構造をコンピュータに渡し、活性部位のどこに、どんな原子のまとまりを置くと強く引き合うかを計算しました（文献 3-9、3-10）。

その結果、環にある O–H のうち 1 つ（図 3-17 の点線の丸）を、プラスの電荷をもつまとまりに置きかえると、結合が強くなると予測されました。この予測にしたがって、N と H がたくさんついたグアニジノ基をつけたのが、ザナミビルです（文献 3-2）。



テラ 大掃除も、ひとりより家族全員で分けたほうが早く終わるよね。10 億個の分子も、1 万個のコアで分ければ 2 週間！

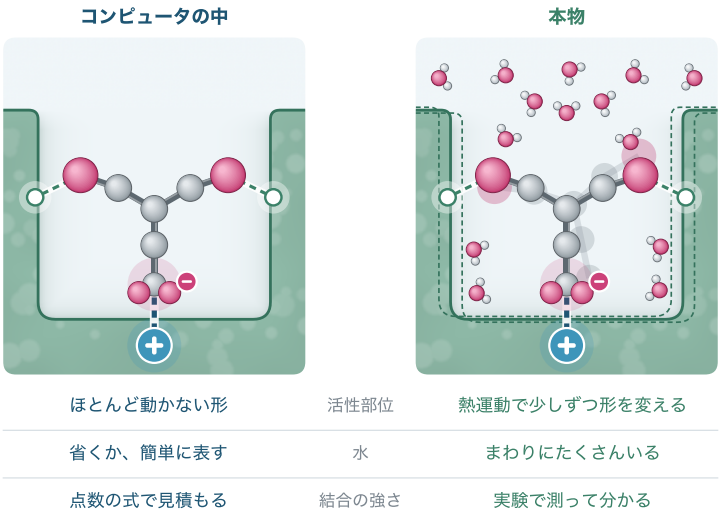


図 3-16　ドッキングの模型（左）と本物（右）のちがい。右の点線とうすい分子は動いたときの位置で、動きの幅と水分子の数は説明のための絵です。タンパク質の動きを追う計算は、第 1 章で見ました。

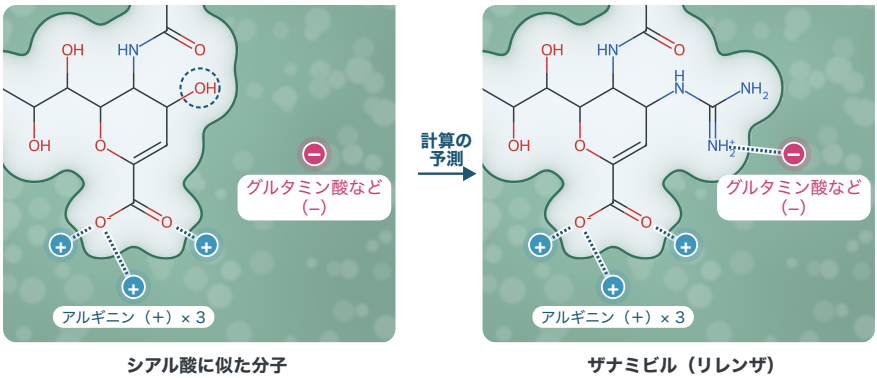
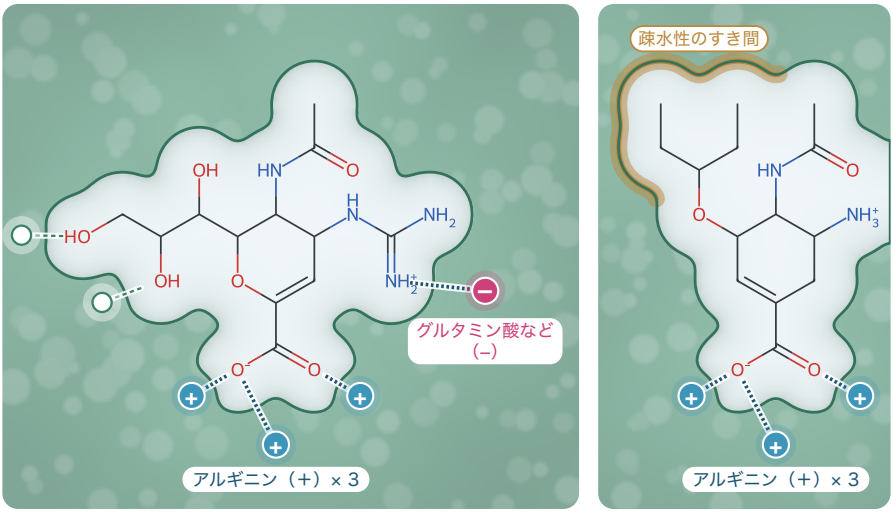


図 3-17　環の O–H を、グアニジノ基に置きかえる（環のまわりを切り抜いた模式図）。線と丸の意味は 図 3-8 と同じです。

2 つ目のオセルタミビル（商品名タミフル）も、活性部位の立体構造をもとに設計されました。分子の環から伸びた枝を、側鎖と呼びます。ザナミビルでは O–H が 3 つ並んだ側鎖がある場所に、オセルタミビルは、炭素と水素の多い側鎖をもっています。この側鎖が、活性部位の中にある疎水性の小さなすき間に入りこむので、オセル



ザナミビル（リレンザ） オセルタミビル（タミフル）

図 3-18 ザナミビル（左）とオセルタミビル（右）が、ノイラミニダーゼの活性部位に結合したようす（模式図）。線と丸の意味は 図 3-8 と同じです。黄色の帯が疎水性のすき間です。オセルタミビルは、飲んだ後に体の中で一部が外れた形で描いています（第 4 章）。

タミビルは活性部位に強く結合します（文献 3-11）。疎水性は、水になじみにくく、油になじみやすい性質で、第 4 章でくわしく見ます。

どちらの薬も、同じ活性部位に結合して、ノイラミニダーゼを阻害します。ところが、リレンザは粉末を吸いこむ吸入薬で、タミフルはカプセルを飲みこむ経口薬です。作用する相手が同じなのに、なぜ薬を体に入れる道筋がちがうのでしょうか。その理由は、次の章で考えます。

★ 要点

コンピュータは、活性部位の立体構造をもとに、候補の分子が結合する強さを点数にします。何億個もの分子をふるいにかけ、作って試す価値のある候補を絞りこみます。ただし点数には外れがあるので、最後は実験で確かめます。

! よくある誤解

「ドッキングの点数が高ければ、薬になる」と考えるのは誤りです。点数は結合の強さの見積もりで、偽の当たりも混じります。それに、活性部位に強く結合しても、薬として使えるとは限りません。飲んだ薬が、体の中で作用する場所まで

届かなければ効かないからです。体に届くかどうかを計算で見積もる方法は、第 4 章で見ます。

研究の現場から：ウイルスの変異と薬剤耐性

ウイルスが増えるとき、遺伝情報が少しずつ変わり、タンパク質のアミノ酸が入れかわることがあります。これを変異といいます。変異によって薬が効きにくくなることを、薬剤耐性と呼びます。

タミフルが効きにくくなったウイルスでは、ノイラミニダーゼのアミノ酸が 1 つ入れかわっていました。変異したノイラミニダーゼの立体構造を調べると、タミフルの側鎖が入る疎水性のすき間が、うまくできなくなっていました（文献 3-12）。そのためタミフルは結合しにくくなりましたが、リレンザは変わらず強く阻害しました。リレンザはシアル酸によく似た形をしているので、このような変異の影響を受けにくいと考えられています（文献 3-2）。

研究の現場では、分子動力学（第 1 章）を使って、アミノ酸の入れかわりや薬の結合に、タンパク質が原子のレベルでどう変化するかを調べています（文献 3-13）。ドッキングと組み合わせれば、変異したウイルスにも効く薬を選ぶ手がかりが得られます。

まとめ

- インフルエンザウイルスは、酵素であるノイラミニダーゼがシアル酸を切りはなすことで、細胞から離れて広がります。活性部位に結合してノイラミニダーゼを阻害する分子が、薬の候補になります。
- 分子が活性部位に結合するには、形が合うことと、水素結合や静電気力（塩橋）などで引き合うことの 2 つが必要です。
- コンピュータは、活性部位の立体構造と候補の分子を受けとり、向きと形を変えて置き方を試し、結合の強さを点数にします（ドッキング）。1 分子で 100 万通りを超える置き方を試すこともあります。
- これを何億個もの分子でくり返し、上位の候補だけを作って実験で確かめます（バーチャルスクリーニング）。点数には偽の当たりが混じり、上位でも当たりは 1〜2 割ほどです。



4

体に届くかを「予測する」

冬にインフルエンザにかかって病院に行くと、第 3 章で見たリレンザやタミフルが出されることがあります。2 つの薬は、どちらもウイルスがもつ酵素ノイラミニダーゼの活性部位（基質を受けとめて反応を進める場所）に結合して、そのはたらきを止めるのでした。

ところが、2 つの薬は、体への入れ方がちがいます。リレンザは、粉末を専用の器具で口から吸いこむ吸入薬で、薬はのどや気管の表面に直接届きます。タミフルは、カプセルを飲みこむ経口薬（飲み薬）で、薬は腸から吸収されて血液に入り、全身をめぐります（図 4-1）。作用する相手は同じなのに、薬を体に入れる道筋がちがうのです。なぜリレンザは飲み薬にならず、吸入薬になっているのでしょうか。

◎ この章で分かること

- 水へのなじみやすさ（親水性）と油へのなじみやすさ（疎水性）のつり合いが、薬を飲み薬にできるかどうかを左右します
- 水と油へのなじみやすさを、$\log P$ という 1 つの数で表せます
- コンピュータは、原子の寄与を足す方法と分子動力学を使い分けて、$\log P$ を予測します

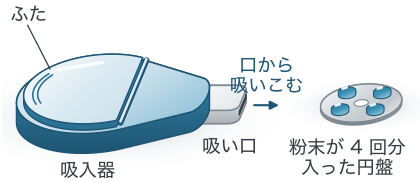
高校で習ったこと：電気陰性度（化学基礎）、細胞（生物基礎）、水素結合・エステル（化学）、対数（数学 II）。化学と数学 II の内容は、習っていなくても読めます



同じ相手に結合するのに、どうして吸入薬と飲み薬に分かれるの？

飲み薬は、腸の細胞の膜を通らなければならない。通れるかどうかを、分子を作る前にコンピュータで調べよう。





(a) リレンザ（吸入薬）



(b) タミフル（経口薬）

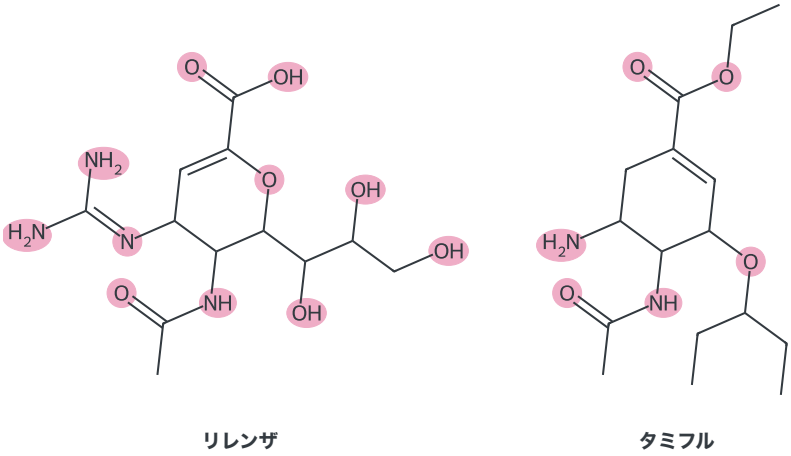
図 4-1　リレンザとタミフル。(a) 粉末の薬が入った円盤を吸入器にセットし、吸い口から吸いこみます（形を簡単にした絵）。(b) シートに入ったカプセルと、取り出したカプセル。（写真：Vantey, パブリックドメイン）

4.1 水になじむ原子、油になじむ原子

砂糖をコーヒーに入れると、すぐに溶けます。サラダ油は、水といくらかき混ぜても、しばらくすると水の上に浮いてきます。分子には、水になじみやすいものと、油になじみやすいものがあります。

リレンザとタミフルは、どちらなのでしょうか。図 4-2 は 2 つの分子の構造式で、O と N の原子に色を付けてあります。第 3 章で見たように、この構造式では C と、C についた H を省いて書いています。線の折れ目と端には C があります。

予想してみよう **リレンザとタミフル、水になじみやすいのはどちらだと思いますか？**



リレンザ　　タミフル

図 4-2　リレンザ（左）とタミフル（右）の構造式。O と N の原子をローズで塗りました。活性部位に結合する骨格が重なるように、2 つの向きをそろえてあります。

答えはリレンザです。O と N を数えると、リレンザには合わせて 11 個、タミフルには 6 個あります。このように原子の種類を見分けて数えることは、コンピュータにもできます。4.4 節では、コンピュータが原子を 1 つずつ見分け、水へのなじみやすさを数にして足していくようすを見ます。

O と N が多いと、なぜ水になじむのでしょうか。O と N は電気陰性度が大きく、結合の電子を強く引きつけます。そのため O–H や N–H の結合では、O や N がわずかにマイナスに、H がわずかにプラスに帯電します。第 2 章では、この電荷のかたよりを計算で確かめました。

水分子も O–H の結合をもっています。分子の O–H にある H（プラス）は、水分子の O（マイナス）と引き合います。反対に、水分子の H が、分子の O や N と引き合うこともあります。これが第 1 章で見た水素結合です。水素結合をつくれる場



テオ　第 2 章で計算した O–H の電荷のかたよりが、ここでは水へのなじみやすさを決めている。電子の計算は、溶け方を考える手がかりにもなる。

所が多い分子ほど、水分子にしっかり囲まれて安定になり、水に溶けこみます。このように、分子やイオンが水分子に囲まれて安定になることを、水和（hydration）と呼びます。

一方、C と H は電気陰性度の差が小さく、C–H の結合にはほとんど電荷のかたよりがありません。水分子は C–H と水素結合をつくれないので、C–H の多い部分は水になじまず、油になじみます。図 4-3 で、そのちがいを見てください。

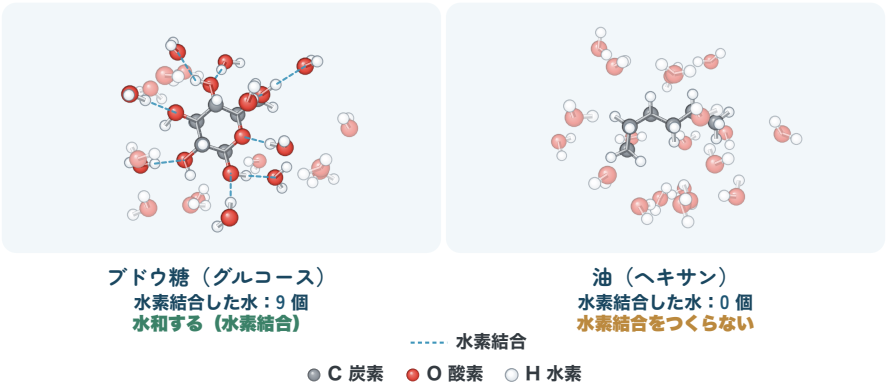


図 4-3　ブドウ糖（グルコース、左）とヘキサン（右）のまわりの水分子を描いた図。O–H の多いブドウ糖は、水分子と水素結合（点線）でつながって水和します。炭素と水素だけでできたヘキサンとは、水分子は水素結合をつくりません。

水になじむ性質を親水性（hydrophilicity）、水になじまず油になじむ性質を疎水性（hydrophobicity）と呼びます。セッケンの「親水基」「疎水基」も、同じ考え方です。疎水性は「水をはじく力をもつ」という意味ではありません。水分子どうしが水素結合で強く引き合うので、水素結合をつくれない部分が、水の外へ押し出されるのです（第 1 章）。

リレンザの側鎖（第 3 章）には、ブドウ糖と同じように O–H が 3 つ並んでいます。リレンザの親水性が高いのは、O–H の並んだこの側鎖と、体の中の水で電荷をもつグアニジノ基・カルボキシ基のためです。では、親水性の高い分子は、なぜ経口薬（飲み薬）に向かないのでしょうか。

4.2 細胞膜は油の膜

経口薬が、作用する場所にたどり着くまでの道筋を追ってみます（図 4-4）。口から飲みこんだ薬は、胃を通って小腸へ送られ、小腸の壁をおおう細胞を通り抜けて血液に入ります。これを吸収（absorption）と呼びます。吸収された薬は、血液にのって全身をめぐり、ウイルスが増えているのどや気管の組織に届きます。

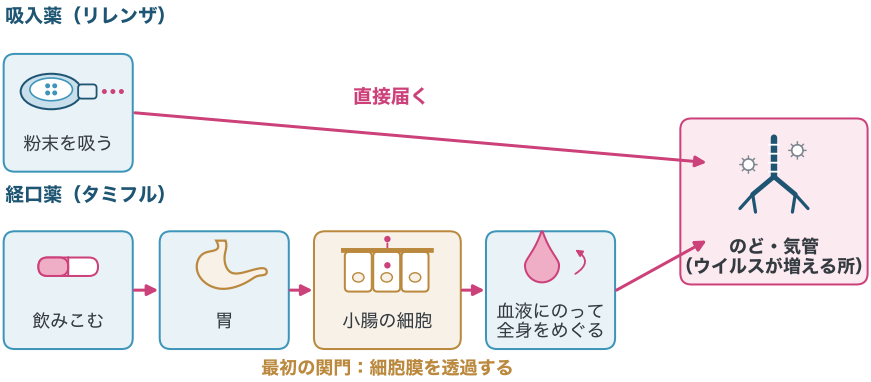


図 4-4　吸入薬（上）と経口薬（下）の、作用する場所までの道筋。経口薬は、小腸の細胞を透過して血液に入り、全身をめぐってから、のどや気管に届きます。絵は形を簡単にしたものです。

この道筋で最初の関門になるのが、小腸の細胞です。細胞は細胞膜で包まれているので、薬は細胞膜を通り抜けないと血液に入れません。膜を通り抜けることを、透過（permeation）といいます。細胞膜の中ほどは、油に似た部分がすき間なく並んだ層になっています。

図 4-5 のように、親水性の高い分子は、膜の油の層に溶けこみにくいので、透過しにくいのです。細胞膜には、決まった物質だけを運び入れるタンパク質もありますが、ここでは膜そのものを透過する道だけを考えます。

親水性の高いリレンザは、小腸の細胞の膜をほとんど透過できません。リレンザは、飲みこんだ量のうち、吸収されて全身の血液に届く割合が、平均約 2 % しかありません（文献 4-1、4-2）。そのためリレンザは、飲み薬ではなく吸入薬として、ウイルスが増えているのどや気管に直接届けるように作られています（図 4-4）。

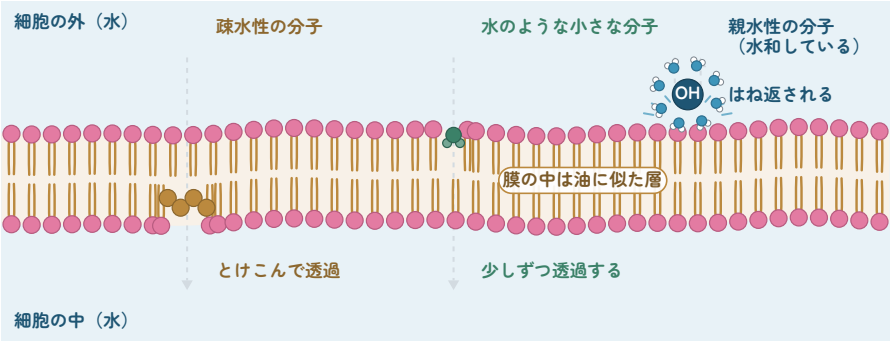


図 4-5 細胞膜を透過しようとする 3 つの分子。疎水性の分子は、膜の油の層に溶けこんで透過します。水のようにとても小さな分子は、少しずつ透過します。親水性の分子は、水和したまま、膜の油の層にはね返されます。

分子が経口薬になれるかどうかを見分けるには、親水性と疎水性の度合いを、数で比べられると便利です。そのための数が $\log P$ です。

4.3 $\log P$：水と油への分かれ方を 1 つの数で

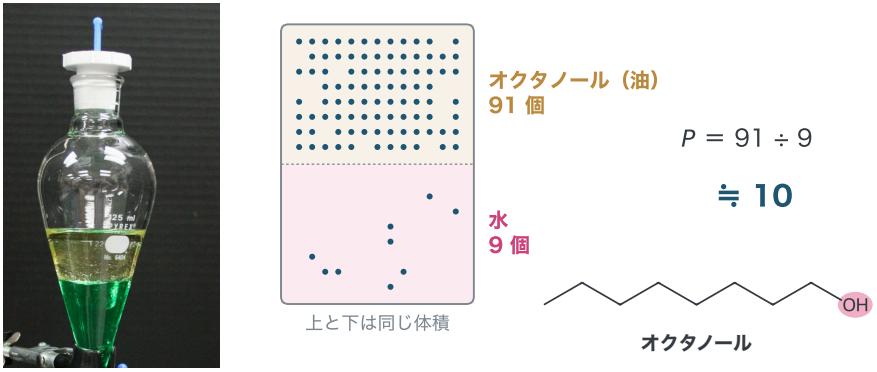
水と、オクタノールというアルコール（図 4-6 の (b) の右下）を、同じ量ずつ容器に入れます。2 つは混ざらず、上下の層に分かれます（図 4-6 の (a)）。オクタノールは細胞膜の油の層に見立てた液体で、どの分子も同じ条件で比べられるように、これを使う決まりになっています。ここに調べたい分子を少し溶かして、よく振ってからしばらく置きます。すると分子は油の層（上）と水の層（下）に分かれ、やがてそれぞれの層に溶けている量が変わらなくなります。この状態を平衡（equilibrium）と呼びます。平衡では、分子の動きが止まっているわけではありません。油から水へ移る分子と、水から油へ移る分子の数がつり合って、見かけの量が変わらなくなっているのです。

平衡に達したとき、油の層と水の層で分子の濃度を比べた値を P と書きます。

$$P = \frac{\text{油の層の濃度}}{\text{水の層の濃度}}$$



テラ ふたをしたびんの水も、表面では分子が飛び出したり戻ったりしているよ。量が変わらないのは、出入りがつり合っているから！



(a) 分液漏斗の 2 つの層　　(b) 層に分かれた分子

図 4-6　(a) 分液漏斗に入れた油（上の黄色い層）と、色をつけた水（下の緑の層）。油はオクタノールではありませんが、同じように 2 層に分かれます。(b) 水とオクタノールに分けた分子（紺の丸）。上の層がオクタノール（油）、下の層が水です。右下は、オクタノールの構造式です。
((a) の写真：PRHaney, CC BY-SA 3.0)

上下の層の体積が同じなら、それぞれの層にいる分子を数えれば、濃度の比が分かります（図 4-6 の (b)）。

P の値は、分子によって大きくちがいます。リレンザでは油と水の濃度の比がおよそ 1 対 6,000、ビタミン A ではおよそ 30 万対 1 です。これほどちがう数を比べやすくするために、「10 を何回かけた数か」で数えます。1 より小さい数は、10 で何回割った数かを、マイナスで数えます。この回数が $\log P$ です（図 4-7）。

図 4-7 からは、次の 3 つが読みとれます。$\log P$ が 0 なら、油と水に半分ずつ分かれます。プラスなら油に、マイナスなら水に多く分かれます。そして、1 増えるごとに 油：水 の比が 10 倍になります。ただし、油の層にいる分子の割合は 10 倍にはならず、91 %、99 % と 100 % に近づいていきます。$\log P$ は、P の常用対数です。

回数は、整数とはかぎりません。リレンザの 1 対 6,000 は、1 対 1,000（10 で 3 回割る）と 1 対 10,000（4 回割る）のあいだなので、$\log P$ も -3 と -4 のあいだの -3.79 になります。

それでは、身近な分子の $\log P$ を並べてみます。その前に、予想の手がかりとして、ビタミン C とビタミン A の構造式を見ておきます（図 4-8）。

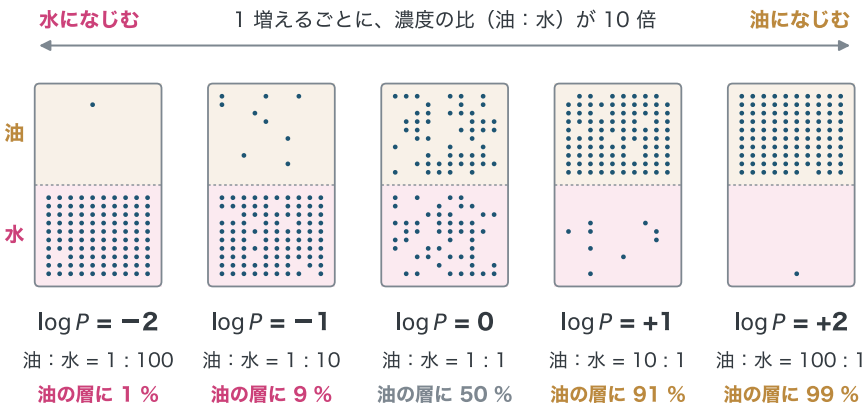
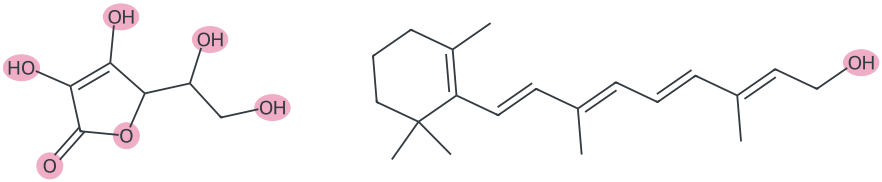


図 4-7　log P と、水と油への分かれ方。どの容器も、同じ体積の油の層（上）と水の層（下）に、分子 100 個を分けて描きました。log P が 1 増えるごとに、油と水の濃度の比（油：水）が 10 倍になります。



ビタミン C　　ビタミン A

図 4-8　ビタミン C（左）とビタミン A（右）の構造式。O の原子をローズで塗りました。

予想してみよう　**ビタミン C とビタミン A、油になじむのはどちらだと思いますか？**

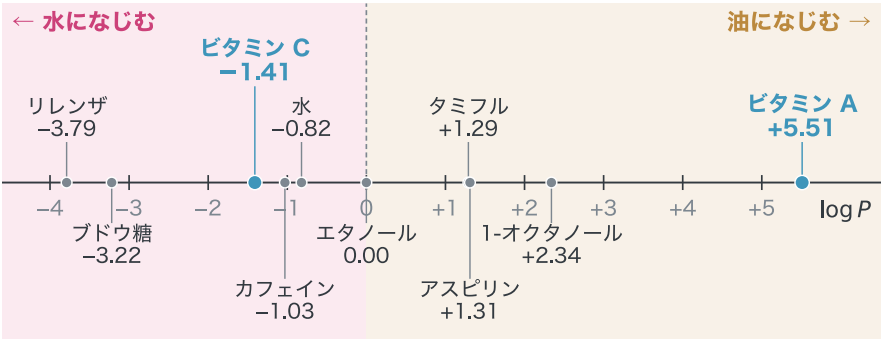


図 4-9　身近な分子の log *P*。0 より左（ローズ）ほど水に、右（黄）ほど油になじみます。青い丸が、予想で取り上げたビタミン C と A です。値は計算で見積もったものです（4.4 節）。

ビタミン C は −1.41 で水の側、ビタミン A は +5.51 で油の側にいます。栄養の話でよく聞く「水溶性ビタミン」「脂溶性ビタミン」の区別と同じ向きです。リレンザの −3.79 は、ブドウ糖（−3.22）よりもさらに水の側にあります。

log *P* は、水とオクタノールを振って分ける実験で測れます。ただし、この実験には、調べたい分子そのものが手元に必要です。ところが、第 3 章でコンピュータの中でふるいにかけた分子は、ほとんどがまだ作られていない分子です。作る前に、この実験をコンピュータの中でできないでしょうか。

実は、図 4-9 の値も、実験で測ったものではありません。コンピュータが、分子の構造式だけから見積もった値です。どうやって見積もるのかを、次の節で説明します。

4.4　コンピュータの中で水と油に分ける

コンピュータの中で分子を水と油に分ける方法は、大きく 2 つあります。実験のデータから決めた値を足す方法と、分子のあいだにはたらく力から計算する方法です。順に見ていきます。

コンピュータに分子を渡す

まず、コンピュータに分子を渡します。コンピュータは、構造式の絵を見て分子を読みとるわけではないので、どの原子がどの原子とつながっているかを表にして渡します（図 4-10）。この表を結合表（connection table）と呼びます。結合表は数と記号の並びなので、コンピュータでそのまま読みこめます。

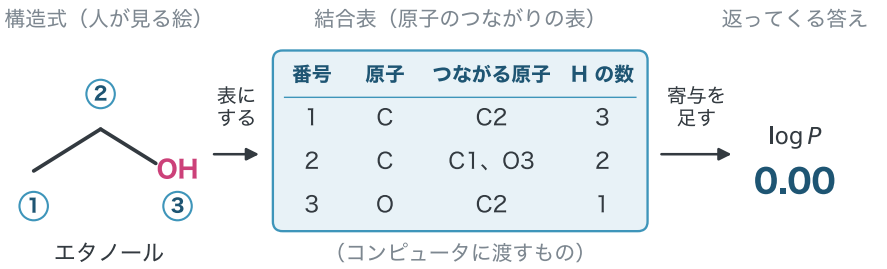


図 4-10 コンピュータに渡すもの（中）と、返ってくるもの（右）。エタノールの構造式（左）の丸数字が、結合表の番号です。

原子の寄与を足す：原子寄与法

分子量は、原子量を足せば求まります。log P にも、これに近い性質があります。分子の log P を、原子ごとの分担分に分けて考えられるのです。

分子の log P のうち、各原子が受けもつ分を、その原子の寄与（contribution）といいます。寄与を足して log P を見積もる方法が、原子寄与法です。

寄与は、4.1 節で見た「O–H は親水性、C–H は疎水性」を数にしたものです。分子を疎水性の側へ寄せる原子はプラス、親水性の側へ寄せる原子はマイナスの寄与をもちます。日常の「寄与」は「役に立つこと」の意味でも使いますが、ここではプラスにもマイナスにも働く、全体への「受けもち分」のことです。

寄与が原子量とちがうのは、同じ C でも、どの原子とつながっているかで寄与が変わることです。そこで、原子を種類とつながる相手で細かく分けます。この区分を原子タイプ（atom type）と呼びます。図 4-10 の結合表に「つながる原子」の列があるのは、原子タイプを見分けるためです。

図 4-11 の上の段は、エタノール（CH_3–CH_2–OH）です。H は小さく、いつも 1 つの原子にだけつながっているので、C や O とそこについた H をひとまとめにして（CH_3、CH_2、OH）、まとまりごとに寄与を示しました。図の 3 つの寄与を足すと −0.01 で

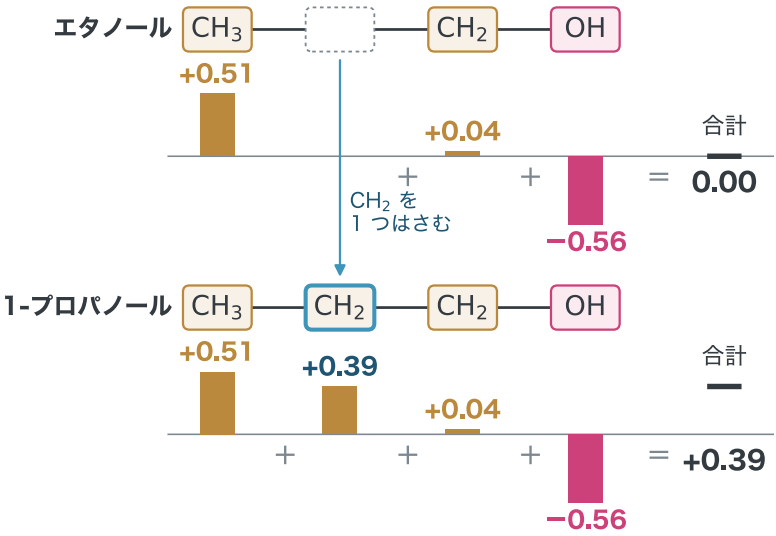


図 4-11 エタノール（上）と 1-プロパノール（下）の、原子のまとまりごとの寄与の足し算。黄の棒はプラス（疎水性の側）、ローズの棒はマイナス（親水性の側）の寄与です。青い枠が、はさんだ CH_2 です。

すが、これは寄与を小数第 2 位までに丸めたためのずれで、丸める前に足すと 0.00 です。エタノールは、水と油にほぼ同じだけなじむと見積もれます。

同じ CH_2 でも、となりの原子で寄与が変わります。となりに O がある CH_2 は +0.04 ですが、C と C にはさまれた CH_2 は +0.39 です。下の段の 1-プロパノール（CH_3–CH_2–CH_2–OH）は、エタノールにこの CH_2 を 1 つはさんだ分子なので、$\log P$ はその分だけ上がって +0.39 になります。

原子タイプごとの寄与の値は、$\log P$ を実験で測ったおよそ 1 万個の分子に合うように決めてあります。このように、実測に合うように決めて計算に使う数を、パラメータ（parameter）といいます。コンピュータは、原子タイプごとのパラメータの表を前もって持っていて、渡された結合表の 1 行ずつについて、その表から寄与を探して足します。パラメータをデータから決める考え方は、第 5 章でくわしく取り上げます。

寄与を足すだけなので、手でもできるほど簡単な計算です。そのため、コンピュータならとても速く終わります（図 4-12）。

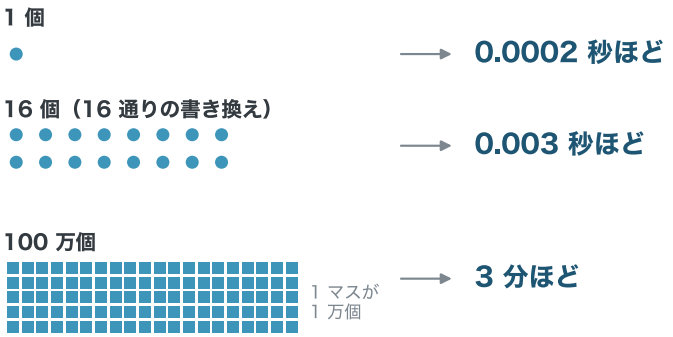


図 4-12　原子寄与法で $\log P$ を出すのにかかる時間。著者のノートパソコンで 1 分子あたりの時間を測り、分子の数をかけました。

100 万個の分子でも 3 分ほどなので、第 3 章のように何億個もの分子をふるいにかけるときにも、$\log P$ をまとめて出せます。実際、2005 年に公開された分子の集まり（約 73 万個）では、すべての分子に、計算で見積もった $\log P$ と分子量が付けてあります（文献 4-3）。

このように、分子をデータとして扱い、構造と性質の関係から性質を予測する分野を、ケモインフォマティクス（chemoinformatics）と呼びます。原子寄与法は、その基本的な方法の 1 つです。第 5 章の機械学習も、この分野でよく使われています。

＋ 発展：クリッペン法　飛ばしても読めます

この本の $\log P$ は、すべてクリッペン法という原子寄与法で計算しています。1999 年にワイルドマンとクリッペンが発表した方法で、原子を 68 の原子タイプに分け、それぞれの寄与を、9,920 個の分子の実測値に合うように決めました（文献 4-4）。計算には、無料で使えるソフト RDKit を使いました。原子ではなく、もう少し大きなまとまりごとに寄与を決める方法もあります。方法によって、見積もりの値は少しずつちがいます。

分子を動かす：分子動力学による計算

2 つ目の方法は、本物の実験を、そのままコンピュータの中で再現するものです（図 4-13）。第 1 章で見た分子動力学を使って、水分子をたくさん並べた箱と、オクタノールの分子を並べた箱を作ります。それぞれの箱に調べたい分子を 1 つ入れ、ま

わりの分子といっしょに動かします。そして 1 コマごとに、調べたい分子がまわりの分子とどれだけ強く引き合っているかを計算し、平均します。

4.1 節では、分子が水分子に囲まれて安定になることを水和と呼びました。調べたい分子は、水の箱では水和し、油の箱でもオクタノールの分子に囲まれて安定になります。どちらの箱のほうが安定かを比べ、その差から $\log P$ を求めます。油の箱のほうが安定なほど、分子は油の層に多く分かれ、$\log P$ は大きくなります。

ただし、安定さには、引き合う強さだけでなく、まわりの分子の並び方（第 1 章の発展）もかかわります。この 2 つを合わせた安定さの目安を自由エネルギー（free energy）といい、実際の計算ではこれを求めます。

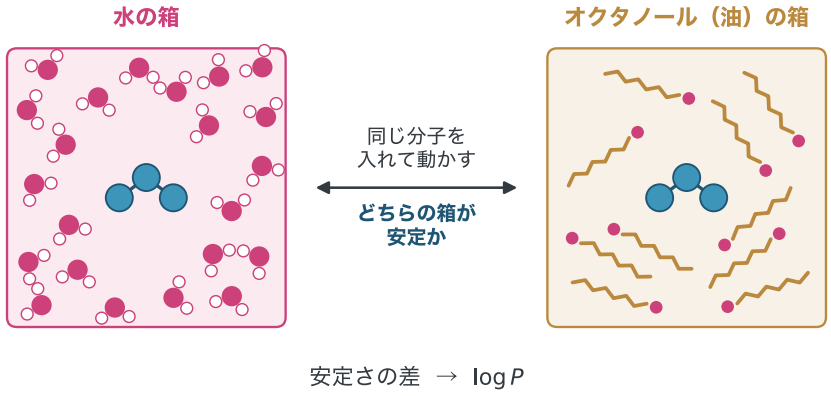


図 4-13 水の箱とオクタノールの箱に、同じ分子（青）を入れたところ。分子の並びは説明のための絵で、計算で求めた配置ではありません。

この方法は、寄与の表の代わりに、第 1 章の力場の数を使います。力場は原子どうしにはたらく力を決める数なので、寄与の表にない、初めて見る形の分子にも使いやすいのです。ただし、力場の数のでき具合によっては、予測が外れます。また、分子を何百万コマも動かして平均をとるので、原子寄与法よりずっと時間がかかります（図 1-8）。第 2 章の量子化学を使って、水の中と油の中での安定さを求める方法もあります。

2 つの方法を比べる

表 4-1 に、2 つの方法を並べました。

表 4-1　コンピュータの中で水と油に分ける 2 つの方法。

	原子寄与法	分子動力学による計算
もとにするもの	実験のデータ	分子のあいだの力
速さ	1 分子 0.0002 秒ほど	何百万コマも動かすので遅い
得意なこと	たくさんの分子をまとめて調べる	初めて見る形の分子も調べる
苦手なこと	パラメータを決めた分子に似ていない分子	たくさんの分子を調べる

2 つの方法は、目的に合わせて使い分けます。たくさんの分子をまとめて調べるなら原子寄与法、寄与の表にない、初めて見る形の分子を調べるなら分子動力学です。

ただし、力から計算する方法のほうが、いつもよく当たるとは限りません。2020 年に報告された予測の腕比べでは、11 個の分子の $\log P$ を、実験の値を伏せたまま、世界の研究グループが予測しました（図 4-14）（文献 4-5）。予測と実験の値の差（予測の誤差）を比べると、分子動力学などのつぶとばねの方法では、誤差の大きさが、データから決める方法の 2 倍ほどありました。水とオクタノールの $\log P$ は、実験で測った値がとてもたくさんあるので、データから決める方法がその分よく当たった、と報告されています。分子動力学は、そうしたデータの少ない、初めて見る形の分子を調べるために使います。

原子寄与法でも分子動力学でも、見積もった値には誤差があります。原子寄与法の誤差の大きさは、第 6 章で、身近な分子を使った答え合わせで確かめます。

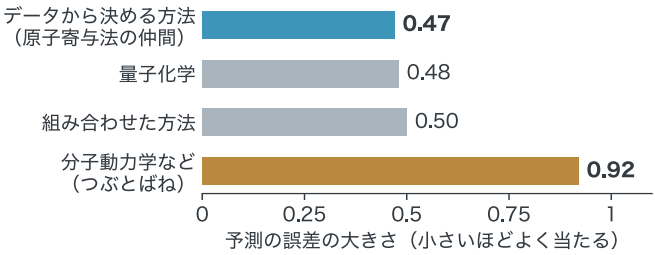


図 4-14　$\log P$ の予測の腕比べ（2020 年に報告。11 個の分子）。棒の長さは、方法の種類ごとに、よく当たった上位 5 つの予測について、誤差の大きさを表す目安です（誤差を 2 乗して平均し、平方根をとった数）。この数が 1 なら、油と水の濃度の比を 10 倍見誤ります（文献 4-5）。

4.5 リレンザを飲み薬に書き換える

ここまでで、リレンザを経口薬にするための方針が見えてきました。活性部位に結合する骨格（分子の本体になる環と、活性部位に結合する部分）は残したまま、親水性の部分を疎水性の部分に置きかえて、分子の構造を書き換えます。そうすれば $\log P$ が上がり、小腸の細胞の膜を透過しやすくなるはずです。

書き換える場所は、次の 4 か所です（図 4-15）。

① 側鎖：O–H が 3 つ並んだ側鎖を、O を 1 つはさんだ炭素と水素の側鎖に
② グアニジノ基：N と H がたくさんついた部分を、小さなアミノ基（NH_2）に
③ 環：環の中の O を C にし、環の二重結合を O のあった側へ移す
④ カルボキシ基：COOH の H を、炭素の鎖（CH_2CH_3）に置きかえて、エステルに

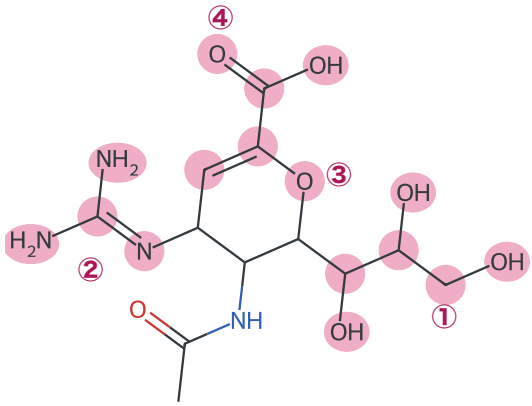


図 4-15 リレンザで書き換える 4 か所（ローズ）。番号は上と同じです。

4 か所それぞれについて、書き換えるか書き換えないかを選ぶと、組み合わせは $2 \times 2 \times 2 \times 2 = 16$ 通りあります。16 個の分子について寄与を手で足すのは大変ですが、コンピュータなら、前の節の速さで 0.003 秒ほどです。16 通りをすべて計算して、並べてみます。



テオ 16 通りなら、ひとつずつ全部確かめても、すぐに終わる。全部を並べれば、勘に頼らずに比べられる。

予想してみよう　**4 か所のうち、log P をいちばん大きく上げるのはどこだと思いますか？**

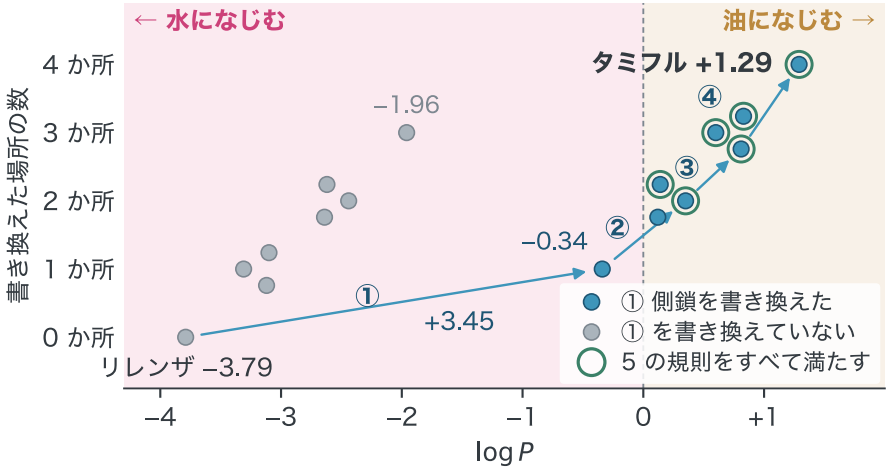


図 4-16　リレンザの 4 か所を書き換えた 16 通りの分子の log P。縦は、書き換えた場所の数です。緑の輪は、経口薬の目安「リピンスキーの 5 の規則」（分子量や、N と O の数などの 4 つの目安。この章の最後の 表 4-2）をすべて満たす分子です。青い矢印は、① → ② → ③ → ④ の順に 1 か所ずつ書き換えた道です。

答えは ① の側鎖です。図 4-16 では、① を書き換えた分子（青い丸）と、書き換えていない分子（灰色の丸）が、はっきり 2 つに分かれています。① を書き換えずにほかの 3 か所をすべて書き換えても（灰色の右端）、① だけを書き換えた分子（青の左端）に及びません。① だけで log P は 3.45 上がり、油と水の濃度の比はおよそ 3,000 倍、油の側に寄ります。

4 か所すべてを書き換えると log P は +1.29 になり、分子はタミフル（オセルタミビル）と同じになります。どの順に書き換えても、最後は +1.29 です。log P は合わせて 5 ほど上がり、油と水の濃度の比はおよそ 10 万倍に変わりました。

コンピュータは、こうして計算した 16 通りの候補を、ほかの目安でもふるいにかけられます。経口薬の目安としてよく使われるのが、この章の最後で紹介する「5 の規則」です。16 通りのうち、その 4 つの目安をすべて満たすのは 6 個（緑の輪）で、どれも ① を書き換えた分子でした。

要点

活性部位への結合にかかわる骨格は残したまま、膜の透過しやすさにかかわる $\log P$ を変えます。コンピュータで書き換えの組み合わせをすべて計算すれば、分子を作る前に、どこを書き換えるとよいかを比べられます。

ここでは、話を分かりやすくするために、2 つのことを簡単にしています。

1 つ目は、書き換えの順と理由です。図 4-16 の青い矢印の道は、考え方を示すためのものです。タミフルが実際にこの順で作られたわけではありません。また、第 3 章で見たように、実際の側鎖は、$\log P$ を上げるためというより、活性部位の中にある疎水性の小さなすき間に入りこみ、活性部位に強く結合するように選ばれたものです（文献 4-6）。

2 つ目は、④ のエステルです（図 4-17）。COOH は、体の中の水では H^+ を放してイオン（COO^-）になり、細胞膜をほとんど透過できません。エステルにすると、カルボキシ基の部分はイオンにならないので、小腸の細胞の膜を透過できます（文献 4-7）。

ところが、第 3 章で見たように、活性部位の 3 つのアルギニン（プラス）と塩橋で結びつくのは、マイナスの COO^- です（文献 4-8、4-9）。エステルのままでは、活性部位にうまく結合できません。そこでタミフルは、飲んで膜を透過した後に、体の中でエステルが外れて COOH の形に戻るように作られています（文献 4-10、4-11）。このように、体の中で効く形に変わるように作った薬を、プロドラッグと呼びます。

ノイラミニダーゼはウイルスの表面にあるので（第 3 章）、血液に届いた後の薬は、細胞の中に入らなくても活性部位に結合できます。

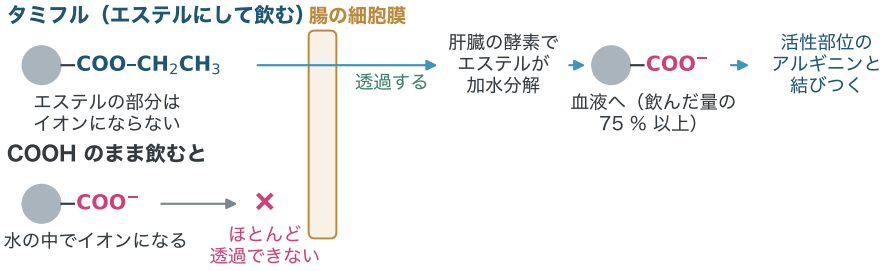


図 4-17　タミフル（プロドラッグ）が体の中でたどる道（上）と、COOH のまま飲んだときの道（下）。灰色の丸は、活性部位に結合する骨格です。75 % 以上は、飲んだ量のうち、効く形（エステルが外れた形）で全身の血液に届く割合です（文献 4-10）。

なお、log P はイオンになる前の形どうしで比べる数なので、エステルにしてイオンにならないようにした効果は、log P には表れません。④ の書き換えで上がるのは 0.48 だけです。

よくある誤解

「log P は大きいほど、よい経口薬になる」と考えるのは誤りです。log P が大きすぎると、こんどは水に溶けにくくなります。経口薬は、まず胃や腸の中の水に溶けないと、細胞膜までたどり着けません。反対に、4.3 節のリレンザ（−3.79）のように小さすぎると、細胞膜を透過できません。log P には、ちょうどよい範囲があるのです。小さすぎる側に決まった数はありませんが、大きすぎる側には、次の「5 の規則」という目安があります。

研究の現場から：経口薬の目安「リピンスキーの 5 の規則」

1997 年にリピンスキーらは、臨床試験まで進んだ多くの分子を調べ、腸から吸収されやすいかどうかを 4 つの目安にまとめました（文献 4-12）。4 つのうち 2 つ以上を超える分子は、腸から吸収されにくいことが多いのです。表 4-2 に、4 つの目安と、リレンザ・タミフルの値を並べました。

表 4-2　「5 の規則」の 4 つの目安と、リレンザ・タミフル。○ は目安の内側、× は目安を超えたことを示します。

目安	リレンザ	タミフル
計算で見積もった log P が 5 まで	−3.79 ○	+1.29 ○
分子の重さ（分子量）が 500 まで	332 ○	312 ○
N–H と O–H の数（N や O についた H の数）が 5 まで	9 ×	3 ○
N と O の数が 10 まで	11 ×	6 ○

どの目安も 5 の倍数なので、「5 の規則」という名前が付いています。これは決まりではなく目安で、これにあてはまらない薬もあります。それでも、候補の分子を作る前に、計算でこの目安を確かめてふるいにかける方法は、薬の研究で広く使われています。

まとめ

- O–H や N–H は水と水素結合をつくるので親水性、C–H の多い部分は疎水性です。
- 経口薬は、小腸の細胞の膜を透過して吸収されます。親水性が高すぎる分子は透過しにくく、リレンザは飲みこんでも全身の血液に届く割合が約 2 % なので、吸入薬になっています。
- 水と油へのなじみやすさは、$\log P$ で表します。$\log P$ は、平衡に達したときの濃度の比 P の常用対数で、1 増えると、濃度の比が 10 倍になります。
- コンピュータは、分子を結合表にして受けとり、原子タイプごとの寄与を足して（原子寄与法）、1 分子 0.0002 秒ほどで $\log P$ を見積もります。遅いけれど力から計算する分子動力学と、目的に合わせて使い分けます。
- 16 通りの書き換えをすべて計算すると、① 側鎖の書き換えが $\log P$ をいちばん大きく上げます。4 か所すべてを書き換えると、$\log P$ が +1.29 のタミフルになります。



5

データから「学ぶ」

第 4 章では、コンピュータが原子のまとまりごとの寄与を足して、エタノールの $\log P$ をおよそ 0.00 と見積もりました。

では、こうした寄与の値（パラメータ）は、だれが、どうやって決めたのでしょうか。この方法（クリッペン法）の寄与は、原子タイプの数だけ、68 個あります。研究者が 1 つずつ考えて決めたのでしょうか。

実は、68 個の寄与は、$\log P$ を実験で測った 9,920 個の分子のデータに合うように、計算で決められました（文献 5-1）。コンピュータは、どうやってデータから規則を見つけるのでしょうか。

◎ この章で分かること

- 測ったデータの表から、コンピュータが規則を見つけることを、機械学習と呼びます
- コンピュータは、規則で予測した値と実測の値の誤差を計算し、誤差が小さくなるようにパラメータを少しずつ直すことを、何千回もくり返します
- 学んだ規則で、まだ測っていない分子を予測できます。ただし、学んだデータの外では外れます

高校で習ったこと：データの分析（散布図・相関）（数学 I）、データの活用（情報 I）

68 個の寄与を、データに合わせて決める。手で合わせるのは無理だ。

テオ

テラ

そこで、コンピュータにデータから学ばせるんだよ！

5.1 測ったデータを並べる

$\log P$ を 1 回測るには、分子そのものを用意し、水とオクタノールを振って平衡にし、それぞれの層の濃度を分析する手間がかかります（4.3 節）。それでも、世界中の研究室が長い年月をかけて、いろいろな分子の $\log P$ を測ってきました。1989 年に、サングスターは、簡単な有機化合物 611 個について、それまでに報告された $\log P$ を集めて表にしました（文献 5-2）。同じ分子でも、測る人や方法によって、値は少しずつちがいます（図 5-1）。サングスターは、報告された値を比べて、分子ごとにいちばん確からしい値を決め、推奨値としました。推奨値には、本当の値がこのくらいの幅に入るだろう、という不確かさが付いています。この章では、この推奨値を実測の値として使います。

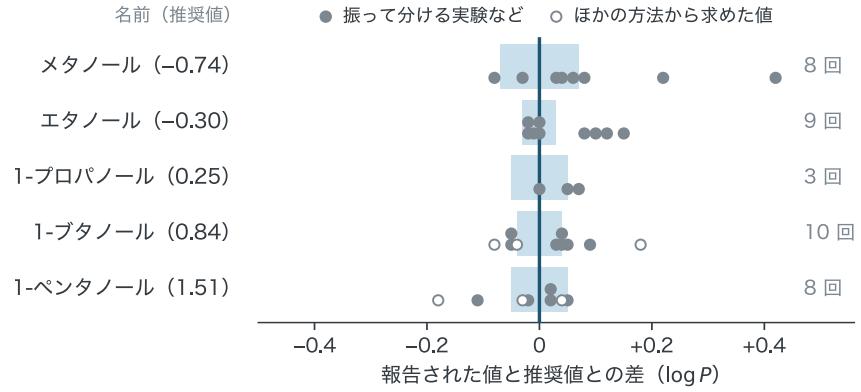


図 5-1　直鎖のアルコール 5 つについて報告された $\log P$ の値。横の位置は、それぞれの値と推奨値との差です。同じ値の丸は縦に積み、振って分ける実験とは別の測定から求めた値は白抜きの丸にしました。青い縦線が推奨値、淡い青の帯が不確かさの幅です。原典で疑わしいとされた値は除きました（文献 5-2）。

図 5-1 の 5 つの分子は、どれも炭素が 1 列につながった鎖の端に、OH が 1 つついたアルコールです。ちがうのは、鎖の炭素の数だけです。炭素の数と $\log P$ の関係を見るために、散布図に並べてみます。

予想してみよう　**炭素が 1 個増えるごとに、$\log P$ はどう変わると思いますか？**

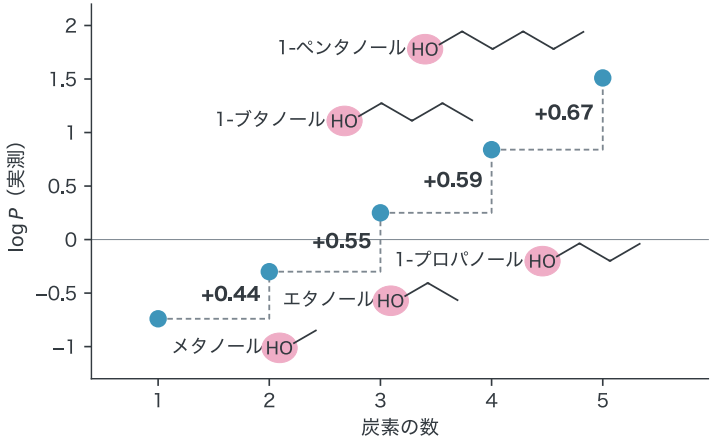


図 5-2　直鎖のアルコールの、炭素の数と $\log P$（実測）の散布図。名前の横は骨格構造式で、O–H の O をローズで塗りました。点線の階段の数は、炭素が 1 個増えたときの $\log P$ の増加分です（文献 5-2）。

答えは、炭素が 1 個増えるごとに、$\log P$ はおよそ 0.5 ずつ上がる、です。CH_2 が 1 つ増えるたびに（図 5-2 の構造式）、分子は油になじみやすくなります。散布図の丸は、右上がりのほぼ一直線に並んでいて、炭素の数と $\log P$ には強い正の相関があります。

ただし、増加分は +0.44 から +0.67 までばらついています。測った値には不確かさがあるので、丸はぴったり一直線には並びません。では、5 つの丸にいちばんよく合う直線は、どれでしょうか。定規を当てて線を引くと、人によって少しずつちがう線になります。「いちばんよく合う」を数で決めて、コンピュータに線を探させることにします。

このように、測ったデータの表から、規則を表す数（パラメータ）を、コンピュータに計算で見つけさせることを、機械学習（machine learning）と呼びます。計算化

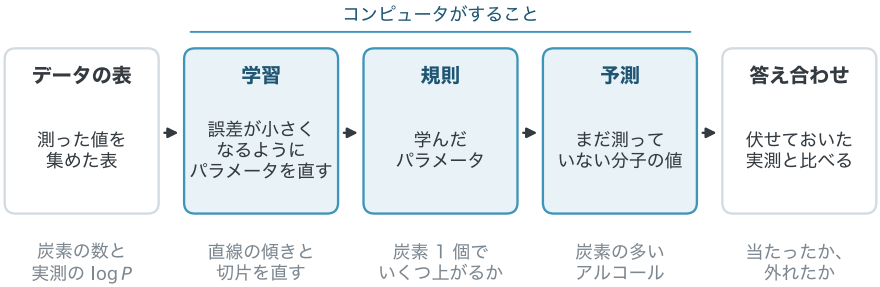


図 5-3　機械学習の流れ。青の枠がコンピュータのすることです。箱の下の灰色の字は、この章のアルコールの例です。

学の 4 つのはたらきでは「学ぶ」に当たります。見つけた規則を使えば、まだ測っていない分子の値も予測できます（図 5-3）。

次の節では、図 5-3 の「学習」のところで、コンピュータが何をくり返しているのかを見ます。

5.2 学習：予測と実測の誤差を小さくしていく

コンピュータに渡すもの

まず、コンピュータに渡すものを決めます（図 5-4）。1 つは、データの表です。分子の名前はそのままでは計算に使えないので、分子を「炭素の数」という数で表します。もう 1 つは、規則の形です。ここでは、規則の形を直線にします。

$$\log P = a \times \text{炭素の数} + b$$

a は直線の傾き、b は切片です。a と b の値を変えると、直線は傾いたり、上下に動いたりします。コンピュータの仕事は、5 つの丸にいちばんよく合う a と b を探して、返すことです。a と b は、第 4 章の寄与と同じく、データに合うように決める数、つまりパラメータです。

分子を炭素の数だけで表したことで、省いたものもあります。鎖の枝分かれや、OH の位置、ほかの原子のまとまりです。この規則が使えるのは、OH を 1 つもつ直鎖のアルコールだけです。

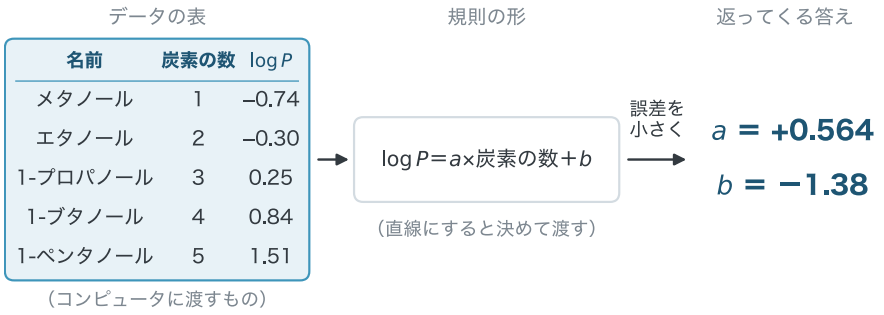


図 5-4　コンピュータに渡すもの（左と中）と、返ってくるもの（右）。

手で 1 回、直線を動かす

直線がどれだけ丸に合っているかを、数で表します。それぞれの分子で、直線で予測した値から実測の値を引くと、予測の誤差が出ます。誤差はプラスにもマイナスにもなるので、そのまま足すと打ち消し合います。そこで、分散を求めるときと同じように、誤差を 2 乗してから平均します。この数を、平均二乗誤差（mean squared error）といいます。直線が丸に合うほど、平均二乗誤差は小さくなります。

何も学んでいない直線から始めます。傾きも切片も 0 の直線、つまり、どの分子も $\log P = 0$ と予測する直線です。この直線では、平均二乗誤差は 0.74 です（図 5-5 の左）。

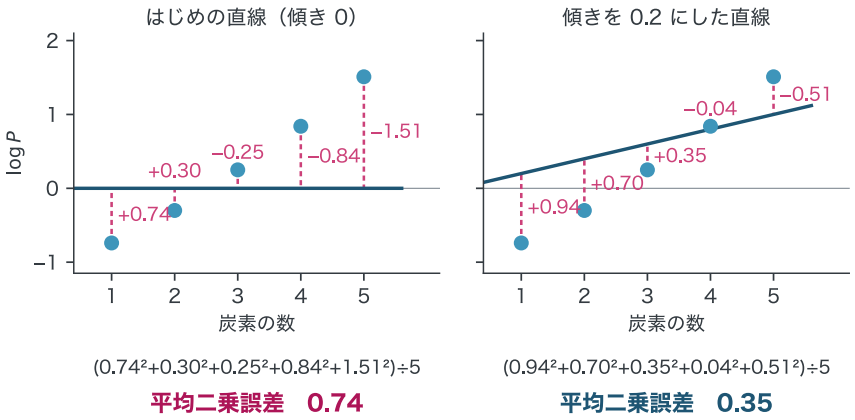


図 5-5　手で 1 回、直線を動かす。丸が実測、紺の線が直線、ローズの点線と数が予測の誤差（予測 − 実測）です。下は平均二乗誤差の計算です。

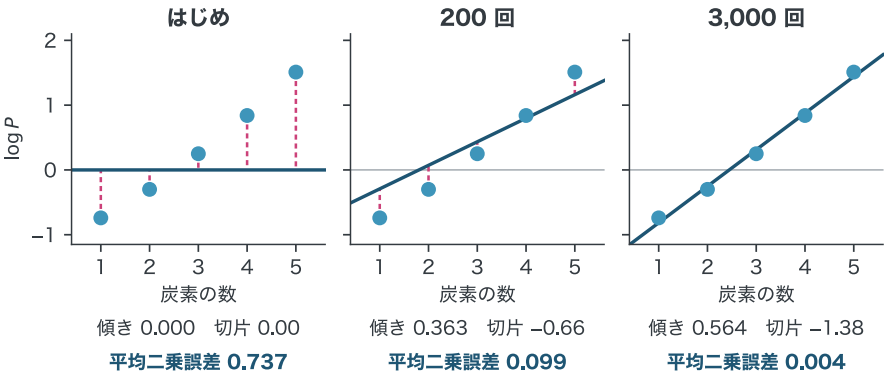


図 5-6　コンピュータの学習で、直線が丸に寄っていくようす。はじめ、200 回、3,000 回くり返したときの直線です。線と丸の意味は 図 5-5 と同じです。

ここで、直線を 1 回だけ動かしてみます。切片はそのままで、傾きを 0 から 0.2 にします。すると、平均二乗誤差は 0.35 に下がりました（図 5-5 の右）。誤差が小さくなったので、この向きに動かしたのはよかったことになります。

これをくり返せば、直線は丸に近づいていきます。ただし、手でやると、どちらの向きにどれだけ動かすかを、毎回考えなければなりません。

コンピュータは何千回もくり返す

コンピュータは、この 1 回を、決まった手順でくり返します。いまの a と b を少し変えたら平均二乗誤差が増えるか減るかを計算し、誤差が下がる向きへ、a と b を少しだけ動かします。これを 3,000 回くり返します（図 5-6）。

くり返すたびに、直線は丸に寄り、ローズの点線は縮んでいきます。平均二乗誤差は、はじめの 0.74 から、3,000 回で 0.004 まで下がりました（図 5-7）。

このように、予測と実測の誤差を計算しては、誤差が小さくなるようにパラメータを直すことをくり返して、データに合う規則を見つけていくことを、学習（learning）と呼びます。日常の「学習」とちがって、コンピュータは、なぜ炭素が増えると $\log P$ が上がるのかを理解しているわけではありません。誤差が小さくなる数を、計算で探しているだけです。



テラ　九九を覚えていても、かけ算の意味が分かっているとは限らないよね。「なぜ」を考えるのは、化学が分かる人の仕事だよ！

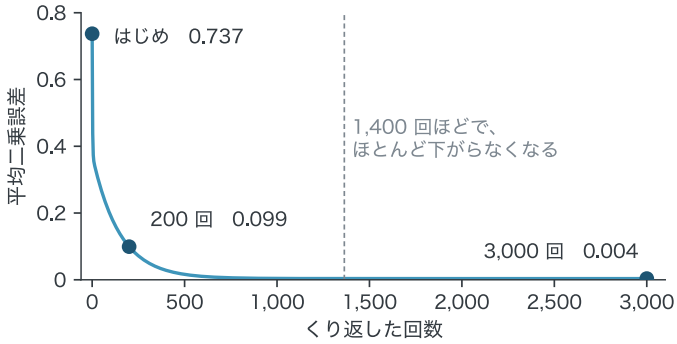


図 5-7　くり返した回数と、平均二乗誤差。

誤差の地図を下る

学習のようすは、地図で見ることもできます。a と b の組を 1 つ決めると、直線が 1 本決まり、平均二乗誤差が 1 つ決まります。そこで、横に a、縦に b をとり、平均二乗誤差の等しいところを線で結ぶと、地図の等高線のような図になります（図 5-8）。

コンピュータは、いまいる場所で坂がいちばん急に下る向きを計算し、その向きへ少しずつ進みます。はじめは手で動かしたのと同じように傾き a を大きくし、その後は谷に沿って下り、谷の底にたどり着きます。坂の勾配（急さ）をたよりに下っていくので、この方法を勾配降下法（gradient descent）と呼びます。たどり着いた谷の底は、$a = +0.564$、$b = -1.38$ でした。

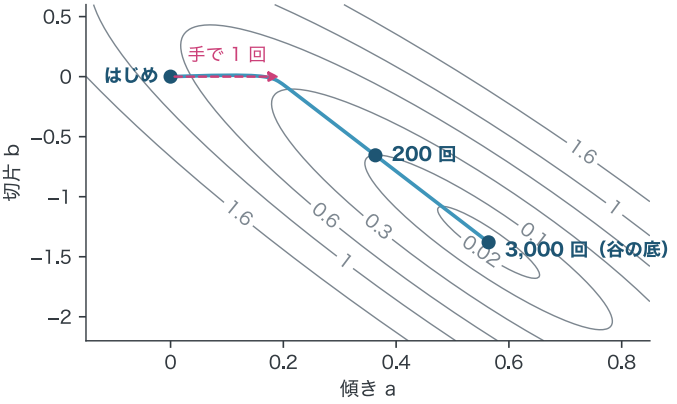


図 5-8　誤差の地図と、学習がたどった道。横が傾き a、縦が切片 b で、灰色の線は平均二乗誤差の等高線（数は平均二乗誤差）です。青の線が、コンピュータの学習でたどった道、ローズの矢印が、手で 1 回動かした分です。

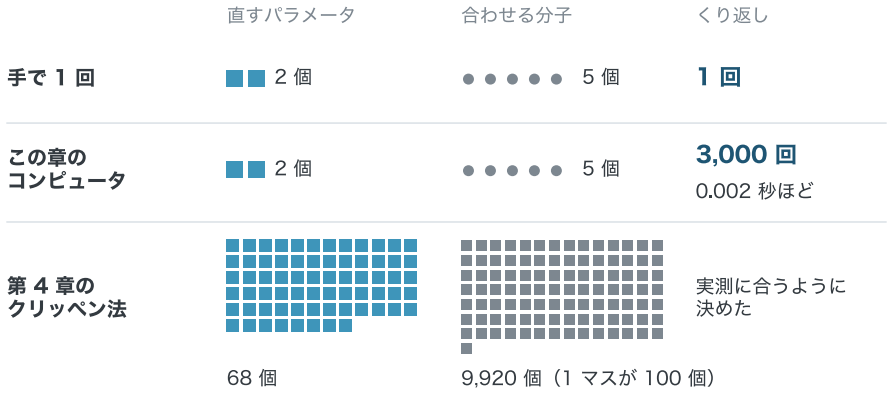


図 5-9　学習の規模。クリッペン法の 68 個と 9,920 個は、文献の値です（文献 5-1）。この章の 3,000 回の時間は、著者のノートパソコンで測りました。

規則が直線のときは、谷の底を 1 回の計算で求める式もあり、それで求めても同じ a と b になります。坂を下る方法は、規則の形がもっと複雑で、そうした式がないときにも使えます。

要点

コンピュータは、規則の形（ここでは直線）を受けとり、予測と実測の誤差が小さくなるように、パラメータを少しずつ直すことを何千回もくり返します。これが学習です。第 4 章の原子寄与法の寄与も、同じ考え方で、たくさんの分子の実測から決めた数です。

手で 1 回から、何千回・何千個へ

手で直線を 1 回動かすのと同じことを、コンピュータは 3,000 回くり返しても、0.002 秒ほどで終えます（図 5-9）。クリッペン法は、$\log P$ を次の式で見積もります。

$$\log P = a_1 \times \text{タイプ 1 の数} + a_2 \times \text{タイプ 2 の数} + \cdots + a_{68} \times \text{タイプ 68 の数}$$

a_1 は原子タイプ 1 の寄与、「タイプ 1 の数」は分子の中にあるタイプ 1 の原子の数です。この章の直線では、炭素の数にかける a は 1 つだけでした。クリッペン法の式は、その a が 68 個に増えた形です。この 68 個の寄与は、$\log P$ を実測した 9,920 個の分子で、予測と実測の誤差が小さくなるように決められました（文献 5-1）。直すパラメータも合わせる分子もずっと多くなりますが、考え方は同じです。

5.3 伏せておいた分子で答え合わせ

学習で見つかった規則は、$\log P = 0.564 \times$ 炭素の数 $- 1.38$ です。炭素が 1 個増えるごとに、$\log P$ は 0.564（およそ 0.56）ずつ上がる、という規則です。この規則で、学習に使わなかった分子の $\log P$ を予測してみます。

サングスターの表には、炭素が 6〜10 個と 12 個の直鎖のアルコールも載っています。答え合わせのために、これらの実測の値は伏せておき、学習には使いませんでした。第 4 章の予測の腕比べで、実験の値を伏せたまま予測させたのと同じやり方です。

予想してみよう　**炭素が 10 個の 1-デカノールの $\log P$ を、学んだ規則で予測すると、いくつになるでしょうか。実測と合うと思いますか？**

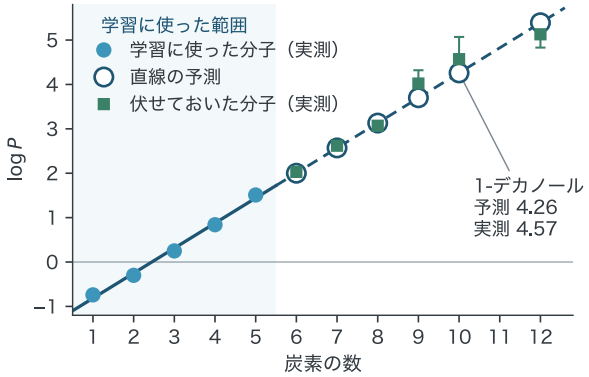


図 5-10　伏せておいた分子の予測と実測。学習に使った範囲の外では、直線を破線でのばしました。緑の四角の縦の線は、実測の不確かさの幅です（文献 5-2）。

答えは $0.564 \times 10 - 1.38$ で、およそ 4.26 です。伏せておいた実測は 4.57 でした。差の 0.31 は、実測の不確かさ 0.5 の範囲に収まるので、合っていると言えます。ほかの分子でも、予測と実測の差は、実測の不確かさと同じくらいか、それより小さく

なりました（図 5-10）。

5 つの分子だけから学んだ規則で、学習に使った範囲の外にある分子まで、かなりよく予測できました。直鎖のアルコールでは、CH_2 が 1 つ増えるたびに、$\log P$ がほぼ同じだけ上がるからです。

ここで、第 4 章の寄与を思い出してください。原子寄与法では、C と C にはさまれた CH_2 の寄与は +0.39 でした。この章で学んだ規則では、炭素 1 個（CH_2 1 つ）で +0.56 です。同じ「CH_2 1 つぶん」なのに、値がちがいます。

規則を複雑にしすぎると

規則の形も、予測の当たり外れを左右します。直線の代わりに、5 つの丸をすべてぴったり通る曲線を引くこともできます。4 次式（x^4 までふくむ式）の曲線です。学習に使った 5 つの分子では平均二乗誤差が 0 になり、直線よりもよく合うように見えます。ところが、この 4 次式で伏せておいた 1-デカノールを予測すると、15.16 になります（実測 4.57、図 5-11）。

測った値には、不確かさの分だけ小さなばらつきがあります。4 次式は、そのばらつきまで写しとるように曲がってしまい、学習に使った範囲の外で大きく外れたのです。この例では範囲の外で目立ちましたが、範囲の中の新しいデータでも外れることがあります。このように、規則を複雑にしすぎて、学習に使ったデータには合うの

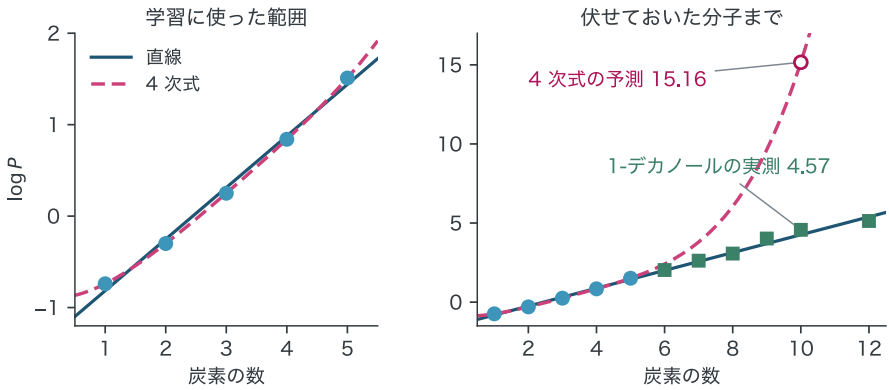


図 5-11　直線と 4 次式。左は学習に使った範囲、右は伏せておいた分子（緑の四角）までの範囲です。白抜きの丸は、4 次式による 1-デカノールの予測です。



テオ　答えを先に見てから解いた問題では、本当に解ける力があるかは分からない。伏せておくのは、そのためだ。

に、新しいデータを外すことを、過学習（overfitting）と呼びます。過学習に気づくには、この章のように、一部のデータを伏せておいて答え合わせをします。

よくある誤解

「学んだ規則は、どんな分子にも使える」と考えるのは誤りです。この章の規則は、炭素の数だけで $\log P$ を決めます。ブドウ糖は炭素が 6 個なので、規則で予測すると +2.00 になります。ところが、ブドウ糖は O–H を 5 つもつので、実測はおよそ −3 です（文献 5-3）。規則が外れたのは、OH を 1 つもつアルコールしか学んでいないからです。学んだデータの範囲の外では、規則は当たるとは限りません（図 5-12）。

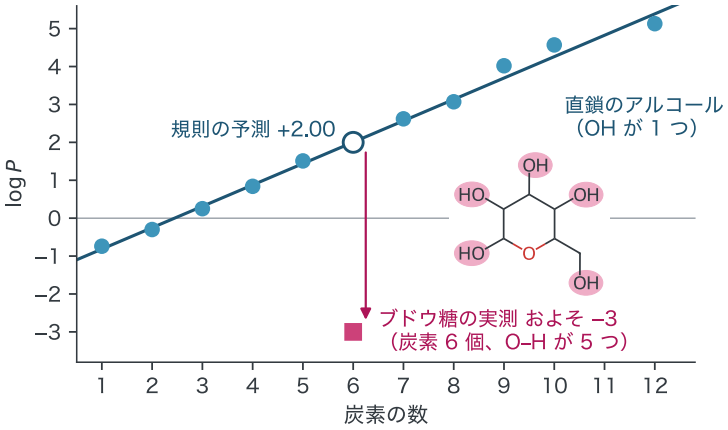


図 5-12 データの外の分子。青い丸が直鎖のアルコールの実測、紺の線が学んだ規則、白抜きの丸が規則によるブドウ糖の予測、ローズの四角がブドウ糖の実測です。右下は、ブドウ糖の構造式で、O–H の O をローズで塗りました（文献 5-2、5-3）。

5.4 化学で使われはじめた AI

人のように判断したり、文章を書いたりするコンピュータのしくみを、人工知能（AI）といいます。質問に答えて文章を書く AI も、たくさんの文章から、次に来る言葉を予測する規則を学んでいます。

AI の規則の形は、直線よりずっと複雑です。よく使われるのは、簡単な計算を何層にも重ねた、パラメータをとても多くもつ規則の形です。これをニューラルネット

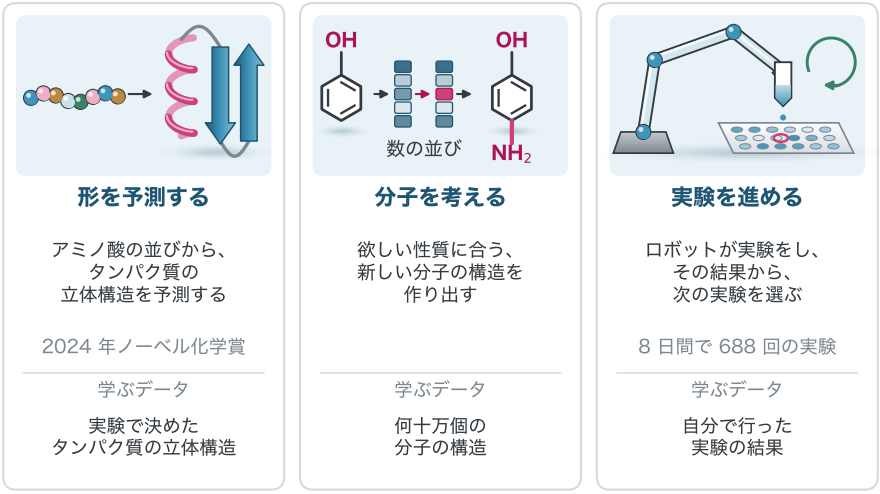


図 5-13 化学で使われる AI の 3 つの例。上の絵は、それぞれの AI が何を受けとり、何を返すかを簡単に表したものです。左の絵のローズのらせんと青の矢印は、図 5-14 と同じ描き分けです。真ん中の絵のローズは、変えた数と、新しくできた枝です。

ワーク（neural network）と呼びます。それでも、データを渡し、予測と正解の誤差を小さくするようにパラメータを直していく、という学習の考え方は、この章の直線と同じです。

化学でも、機械学習による AI が、いろいろな場面で使われはじめています（図 5-13）。

1 つ目は、タンパク質の形を予測する AI です。立体構造を実験で 1 つ得るのに何年もかかることも多く、これまでに実験で決められた立体構造は約 20 万です（文献 5-4）。2020 年に発表された AlphaFold2（アルファフォールド 2）は、アミノ酸の並びだけから、図 5-14 のような、鎖が折りたたまれた形を予測します。答えを伏せた予測の腕比べ（第 6 章）では、多くのタンパク質で、実験で決めた構造に近い精度で予測しました（文献 5-5）。そして、知られているタンパク質のほぼすべての立体構造を予測しました。実験で決めた約 20 万に対して、予測はおよそ 2 億です。序章で紹介した 2024 年のノーベル化学賞の半分は、AlphaFold2 を作ったハサビスとジャ

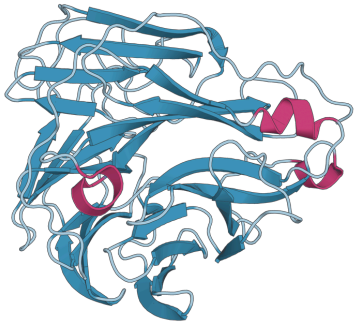


図 5-14 タンパク質の立体構造の例。第 3 章のノイラミニダーゼの鎖 1 本を、リボンのように描きました（PDB 2HU4（文献 5-6））。ローズはらせん、青はシートと呼ばれる形です。

ンパーに贈られました（文献 5-4）。

2 つ目は、新しい分子を考える AI です。分子を数の並びに置きかえ、その数を少しずつ変えて、欲しい性質に近い新しい分子の構造を作り出します（文献 5-7）。

3 つ目は、実験を進める AI です。2020 年に報告された例では、ロボットが、光を当てると水から水素を作る材料で、はじめの 6 倍の活性をもつ組み合わせを見つけました（文献 5-8）。

どの AI も、この章の直線と同じく、データから学んだ規則で予測します。そのため、学んだデータの範囲の外では外れることがあり、最後は実験で確かめます。それでは、学んだ規則や、第 4 章の原子寄与法は、実際にどこまで当たるのでしょうか。同じ CH_2 1 つぶんの +0.39 と +0.56 は、どちらを信じればよいのでしょうか。第 6 章で、計算と実測を並べて答え合わせをします。

研究の現場から：学習した規則で、新しい抗菌薬の候補を探す

抗菌薬が効かない細菌が増えていて、新しい抗菌薬が求められています。2020 年に報告された研究では、機械学習で抗菌薬の候補を探しました（文献 5-9）。



テラ タンパク質の鎖には、らせんのような決まった形がよく出てくるよ。どう折りたたまれるかは、おもにアミノ酸の並びで決まるんだ！

研究者たちは、分子を実験で試して、大腸菌の増殖を止めるかどうかのデータを集めました。このデータから、分子の構造を受けとって増殖を止めるかどうかを予測する規則を学習させ、別の分子の集まりから候補を選びました（図 5-15）。その 1 つのハリシンは、これまでの抗菌薬とは構造が大きくちがう分子でした。第 3 章のバーチャルスクリーニングでは、ドッキングの点数で分子をふるいにかけました。この研究では、実験のデータから学んだ規則で、分子をふるいにかけたのです。それでも、最後に効き目を確かめたのは実験でした。

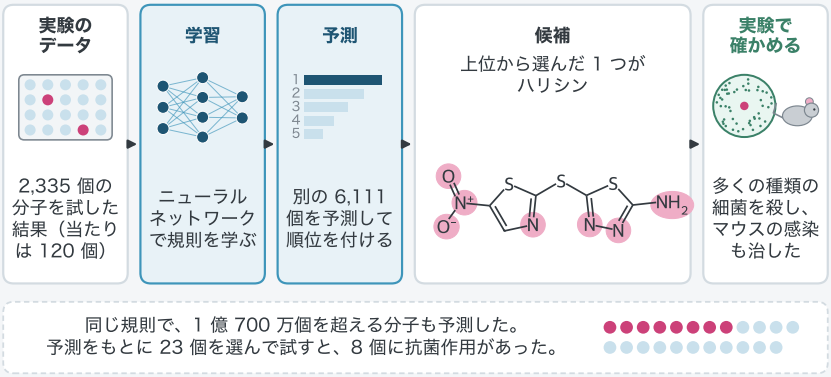


図 5-15　学習した規則で抗菌薬の候補を探した流れ。青の枠がコンピュータのすることです。「候補」の箱はハリシンの構造式で、O と N をローズで塗りました（文献 5-9）。ほかの箱の絵は簡単にしたものです。下の段の丸は試した 23 個で、ローズが抗菌作用のあった 8 個です。

まとめ

- $\log P$ の実測の値は、実験の積み重ねです。同じ分子でも測るたびに少しずつちがうので、集めた値から推奨値を決めて使います。
- 測ったデータの表から、規則を表すパラメータをコンピュータに見つけさせることを、機械学習と呼びます。直鎖のアルコールでは、炭素が 1 個増えるごとに $\log P$ が 0.564 上がる、という規則が見つかりました。
- コンピュータは、予測と実測の平均二乗誤差が小さくなる向きにパラメータを少しずつ動かすことを、3,000 回くり返して学習します。3,000 回でも 0.002 秒ほどです。
- 学んだ規則は、伏せておいた分子をよく予測しました。ただし、学んだデータの外の分子では外れ、規則を複雑にしすぎると新しい分子を外します（過学習）。だから、最後は実験で確かめます。

6

計算を信じてよいか



第 4 章では、コンピュータがエタノールの寄与を足して、$\log P$ を 0.00 と見積もりました。ところが、水とオクタノールに分けて実際に測った値（実測）は −0.30 です（文献 6-1）。第 5 章でも、炭素 1 個ぶんの $\log P$ が、第 4 章の寄与では +0.39、実測から学んだ規則では +0.56 と食いちがっていました。

計算と実測が食いちがうとき、どちらを信じればよいのでしょうか。そもそも、コンピュータの計算はどのくらい当たるのでしょうか。この章では、これまでの章でコンピュータが出した答えを、実験の値と突き合わせます。

◎ この章で分かること

- コンピュータの計算を実測と並べると、誤差の大きさが分かります。外れ方の理由は、モデルが省いたものにあります
- 省いたものの少ないモデルほど、分子が大きくなると計算時間が急に増えます。だから目的に合わせてモデルを選びます
- コンピュータの計算で候補を絞り、実験で確かめ、測った値でモデルを直します。計算と実験は両輪です

高校で習ったこと：データの分析（数学 I）

6.1 答え合わせ：計算と実測を並べる

まず、実測の −0.30 という値を見直してみます。第 5 章で見たように、1989 年にサングスターは、いくつもの研究室が測った $\log P$ を集めて表にしました。同じ分子で

も、測った値には幅があります。サングスターの表には、エタノールについて 9 つの報告が並んでいます。そこに計算の値を重ねたのが 図 6-1 です。

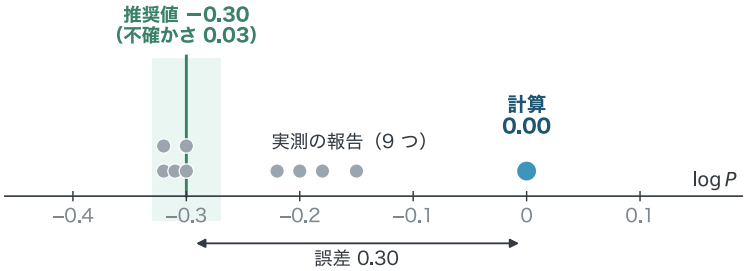


図 6-1　エタノールの $\log P$ の、計算（青）と実測。灰色の丸は、サングスターの表に並んだ 9 つの実測の報告で、同じ値は縦に積みました。緑の線と帯は、推奨値と、その不確かさです（文献 6-1）。

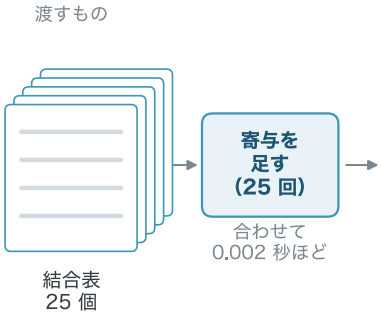
報告ごとに測った温度や層の濃度の測り方などがちがうので、報告された値は −0.32 から −0.15 まで散らばっています。サングスターの推奨値は −0.30（不確かさ 0.03）で、第 5 章で学習に使ったのもこの値でした。

実測にも、このくらいの幅はあります。それでも、計算の 0.00 は、9 つのどの報告よりも油の側にあります。計算は、実測より 0.3 ほど大きく見積もったのです。$\log P$ が 0.3 ちがうと、油と水の濃度の比は 2 倍ちがいます。2 を 3 回かけると 8 で、ほぼ 10 倍です。10 倍は $\log P$ で 1 なので、2 倍はその 3 分の 1 の 0.3 ほどに当たります。計算では、エタノールは油の層と水の層に半分ずつ分かれますが、実測では、油の層にいるのはおよそ 3 分の 1 です。

では、ほかの分子ではどうでしょうか。水、お酒のエタノール、お酢の酢酸、防虫剤のナフタレン、コーヒーのカフェインなど、身近な分子を 25 個選び、第 4 章と同じクリッペン法で $\log P$ を計算しました（図 6-2）。コンピュータは 25 個の結合表を受けとり、1 つずつ寄与を足して、25 個の $\log P$ を返します。かかった時間は、25 個を合わせても 0.01 秒未満でした。

実測の値には、公開されているデータ集の値を使いました。計算から実測を引いた差が、その分子の誤差です。誤差がプラスなら、計算は実測より油の側に見積もり、マイナスなら水の側に見積もったことになります。

25 個の答え合わせを、横の軸に実測、縦の軸に計算をとった散布図にしてみます。計算と実測がぴったり合えば、丸は斜めの線「計算 ＝ 実測」の上に乗ります。



分子	コンピュータが返す値 計算	実験で測った値 実測	誤差
水	−0.82	−1.38	**+0.56**
エタノール	0.00	−0.30	**+0.30**
ヘキサン	+2.59	+3.90	**−1.31**
アスピリン	+1.31	+1.19	**+0.12**
グリシン	−0.97	−3.21	**+2.24**
（ほか 20 個）			誤差 ＝ 計算 − 実測

図 6-2　答え合わせで、コンピュータに渡すもの（左）と、返ってくる値（右の表の青い列）。表は 25 個のうち 5 個を示しました。

予想してみよう　**25 個の分子のうち、誤差の大きさが 0.5 より小さい分子は、いくつあると思いますか？**

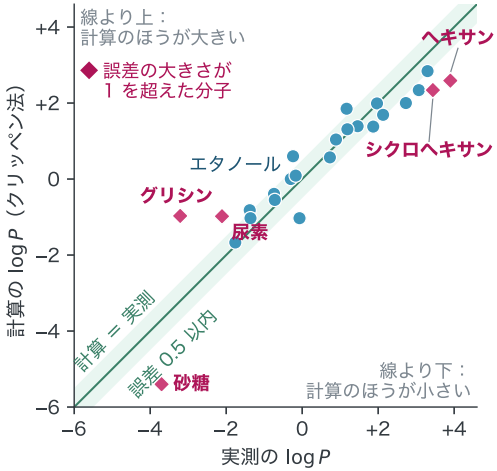


図 6-3　身近な分子 25 個の log P の、計算と実測の散布図。淡い緑の帯は、誤差の大きさが 0.5 より小さい範囲です。ローズの ◆ が、誤差の大きさが 1 を超えた分子です。実測は、直鎖のアルコール 3 個（メタノール、エタノール、1-オクタノール）がサングスターの推奨値、ほかの 22 個がデータベース PubChem に載っている 1995 年のデータ集の値です（文献 6-2、6-3）。

答えは 14 個です。25 個のうち半分ほどが、斜めの線をはさむ帯に入りました。丸は全体として斜めの線にそって並び、相関係数は 0.92 です。水になじむ分子か、油

になじむ分子かは、計算でほぼ見分けられています。一方で、誤差の大きさが 1 を超えた分子も 5 個ありました。

誤差を 1 つの数にまとめるには、25 個の誤差の大きさ（マイナスの符号を外した値）を平均します。この数を、誤差の大きさの平均と呼びます。クリッペン法では、誤差の大きさの平均は 0.62 でした。0.62 は 0.3 のおよそ 2 つ分なので、$\log P$ が 0.62 ちがうと、油と水の濃度の比は、2 倍をさらに 2 倍して、およそ 4 倍ちがいます。これが、「原子寄与法はどこまで当たるか」という問いへの答えの 1 つです。

外れた分子には、何か共通するところがあるのでしょうか。次の節で、誤差の理由を探ります。

6.2 誤差の理由：モデルが省いたもの

図 6-4 は、25 個の分子を、誤差の順に並べたものです。

計算が水の側に外れた分子を上から見ていくと、ヘキサンやシクロヘキサンのように、炭素と水素だけでできた分子が並んでいます。トルエンや 1-オクタノールも、C–H の多い分子です。反対の端で、計算が油の側に大きく外れたのは、グリシンと尿素です。外れ方にくせがあるのは、モデルが何かを省いているからです。

炭素 1 個ぶんの寄与：+0.39 と +0.56

C–H の多い分子で外れる理由を、第 5 章の直鎖のアルコールで確かめます。第 5 章では、メタノールから 1-ペンタノールまでの実測から、炭素が 1 個増えるごとに $\log P$ が +0.56 増えるという規則を学びました。第 4 章の寄与の足し算では、C と C にはさまれた CH_2 1 つで +0.39 です。2 つを実測と並べたのが 図 6-5 です。



テラ　pH も、1 ちがうと水素イオンの濃度が 10 倍ちがうよね。$\log P$ も、同じ読み方をする目盛りなんだよ！

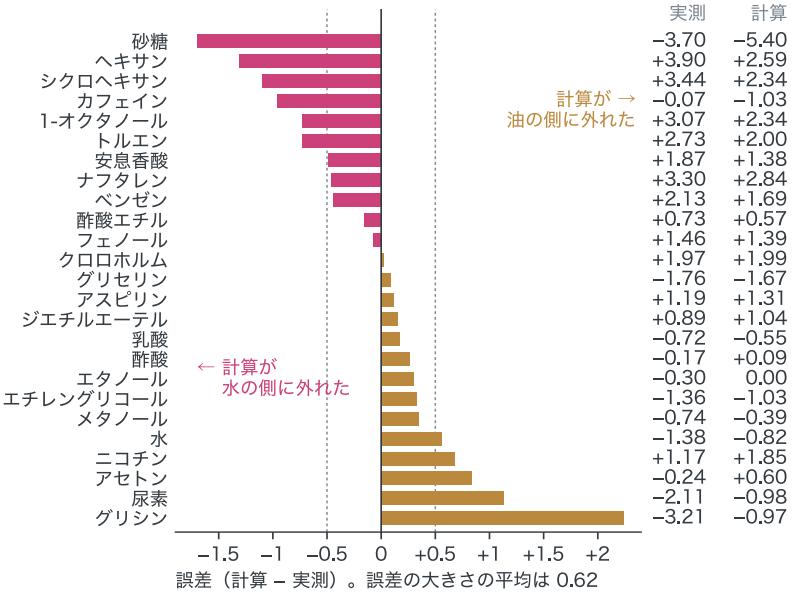


図 6-4　25 個の分子の誤差（計算 − 実測）。ローズの棒は計算が水の側に、黄の棒は計算が油の側に外れた分子です。点線は誤差 ±0.5 です。

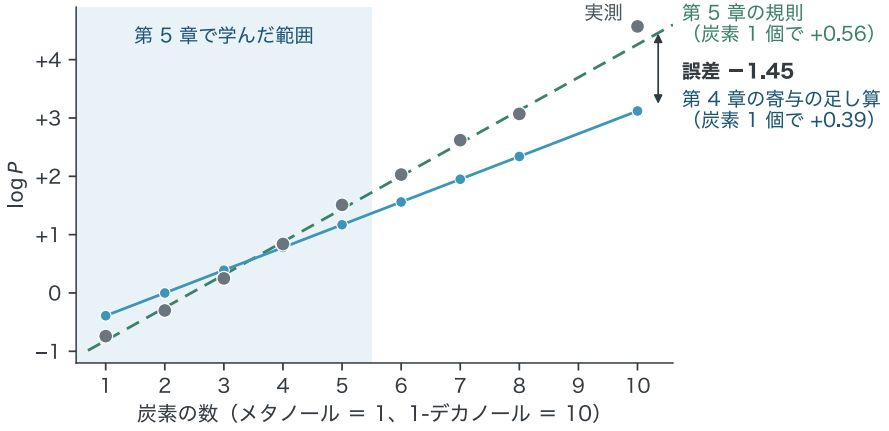


図 6-5　直鎖のアルコールの炭素の数と $\log P$。灰色の丸が実測（サングスターの推奨値）、青の線と丸が第 4 章の寄与の足し算、緑の破線が第 5 章で学んだ規則です（文献 6-1）。

炭素が増えるほど、寄与の足し算は実測より小さくなっていきます。CH_2 が 1 つ増えるたびに、$0.56 - 0.39 = 0.17$ ずつ差が開くからです。1-デカノール（炭素 10 個）

では、誤差は −1.45 になりました。第 5 章の規則なら、学んだ範囲の外の 1-デカノールでも、誤差は −0.31 です。1-デカノールは実測の不確かさも 0.5 と大きいので、規則はその幅の中で当たっています。

なぜ、寄与の足し算は CH_2 を小さく見積もるのでしょうか。クリッペン法では、ヘキサンの CH_2 も、シクロヘキサンの環の CH_2 も、薬の分子で C と C にはさまれた CH_2 も、同じ寄与で数えます。そして 68 の原子タイプの寄与は、およそ 1 万個の分子の実測に、まとめて合うように決めてあります（第 4 章）。いろいろな分子に少しずつ合わせた値なので、直鎖のアルコールや、炭素と水素だけの分子にぴったり合うとは限りません。

一方、第 5 章の規則は、直鎖のアルコールのデータだけから学んだので、直鎖のアルコールにはよく合います。そのかわり、ほかの種類の分子には使えません。どちらかが正しいのではなく、2 つのモデルは省いたものがちがうのです。第 4 章の寄与は、どんな分子にも使えるかわりに、分子の種類ごとのちがいを省いています。第 5 章の規則は、直鎖のアルコールしか学んでいません。だから、直鎖のアルコールの $\log P$ を知りたいなら +0.56 の規則を、いろいろな分子をまとめて見積もるなら +0.39 の寄与を使い、誤差の大きさを頭に入れておきます。

砂糖（スクロース）では、計算が実測より 1.70 も水の側に外れました。砂糖の分子には、O–H が 8 つも近くに並んでいます（図 6-6）。寄与の足し算では、原子をとなりの原子とのつながりだけで見分けて、寄与を 1 つずつ足します。となりより遠くの原子どうしが分子の中で影響し合っても、その分は数に入りません。

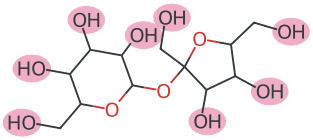


図 6-6　砂糖（スクロース）の構造式。O–H の O をローズで塗りました。

尿素（誤差 +1.13）とカフェイン（誤差 −0.96）も大きく外れましたが、その理由を説明するには、この本より進んだ化学が必要です。いちばん外れたグリシンは、6.5 節で見ます。

ドッキングの答え合わせ

第 3 章のドッキングも、答え合わせをしておきます。2019 年の大きなバーチャルスクリーニングでは、ある酵素に 9,900 万個の分子をドッキングし、点数の上位から選んだ 44 個を作って試したところ、阻害剤だったのは 5 個（11 %）でした（図 6-7）（文献 6-4）。同じ研究のもう 1 つのタンパク質では、点数がいちばん高い範囲で当た



図 6-7　ドッキングの点数の上位から選び、作って実験で試した 44 個の分子（2019 年の例）（文献 6-4）。

りの割合が 24 % ほどあり、点数が下がるほど当たりは減りました（文献 6-4）。上位の分子でも、当たりは 1〜2 割ほどです。点数で、分子 1 つ 1 つの結合の強さや順位まで当てるのはむずかしいのです（文献 6-5）。それでも、点数の式は、活性部位が動くことやまわりの水を省くかわりに、何億個もの分子を計算できるほど速くなっています。外れは多くても、点数は当たりの多い範囲を選び出すふるいとしてはたらき、9,900 万個の中から作って試す 44 個を選んで、5 個の阻害剤を見つけられたのです。

各章のモデルが省いたもの

ここまでの章で見たコンピュータの中の実験を、3 つの問い（何をモデルにしたか・何をくり返したか・どこまで信じられるか）でまとめ直したのが 図 6-8 です。

どのモデルも、何かを省いています。物理の法則から計算する方法も、省いたもののぶんだけ外れます。第 2 章のナトリウムでは、電子どうしのかかわりの近似を変えると、計算した光の波長は 557 nm から 628 nm まで分かれました（実測は 589 nm、図 2-17）。第 4 章の腕比べでは、分子動力学などのつぶとばねの方法で求めた $\log P$ の誤差が、データから決める方法の 2 倍ほどありました（図 4-14）。

省くことは、手抜きではありません。省くことで計算が速くなり、何億個もの分子を調べたり、何百万コマも分子を動かしたりできるようになります。大切なのは、何を省いたかを知っておくことです。そうすれば、計算が外れたときに、どこを疑えばよいかが分かります。

要点

コンピュータの答えには、モデルが省いたもののぶんだけ誤差があります。実測と答え合わせをすれば、誤差の大きさと、外れやすい分子の種類が分かります。

それなら、何も省かずに、ていねいに計算すればよいのではないでしょうか。そこで問題になるのが、計算にかかる時間です。

コンピュータの中の実験	何をモデルにしたか（省いたもの）	何をくり返したか	どこまで信じられるか
 第 1 章・見る **分子動力学**	原子をつぶ、結合をばねにする（電子は省く）	力 → 加速度 → 速さ → 位置を 1 コマずつ。何百万コマも	力の決め方しだい。結合が切れる反応は扱えない
 第 2 章・見る **量子化学**	原子核を止めて、電子だけを考える。電子どうしのかかわりは近似する	電子の雲を計算し、エネルギーの段差から光の色を求める	近似の選び方で、色の予測に誤差が出る。大きな分子は計算が重い
 第 3 章・探す **ドッキング**	活性部位はほぼ動かない形に。水は省く。結合の強さは点数の式で見積もる	置く・回す・点数にするを、1 分子で 100 万通り以上	上位でも当たりは 1〜2 割。偽の当たりが混じる
 第 4 章・予測する **ケモインフォマティクス**	原子寄与法：原子をとなりの原子だけで見分けて、寄与を足す	寄与の足し算を、1 分子 0.0002 秒ほどで	誤差の大きさの平均 0.62。炭素と水素ばかりの分子や、双性イオンは苦手
 第 5 章・学ぶ **機械学習**	規則を直線に、分子を数（炭素の数）にする	直線を少し動かして誤差を測るのを、何千回もくり返す	学んだデータの範囲の外では、外れることがある

図 6-8 第 1〜5 章の、コンピュータの中の実験と 3 つの問いの答え。左の小さな絵は、それぞれのモデルを簡単にしたものです。

6.3 計算にかかる時間

省いたものの少ない、ていねいなモデルほど、計算に時間がかかります。どのくらいかかるのかを、著者のノートパソコンで確かめてみました。メタン、エタン、プロパン、……と、炭素の数を 1 つずつ増やした、炭素と水素だけの鎖の分子を用意し、分子 1 個の計算にかかる時間を 3 つのやり方で測りました。第 4 章の原子寄与法と、第 2 章の量子化学の計算を、精度のちがう 2 つの近似（ふつうの精度・高い精度）で行ったものです。

予想してみよう **分子の大きさが 2 倍になると、量子化学の計算時間は何倍になると思いますか？**

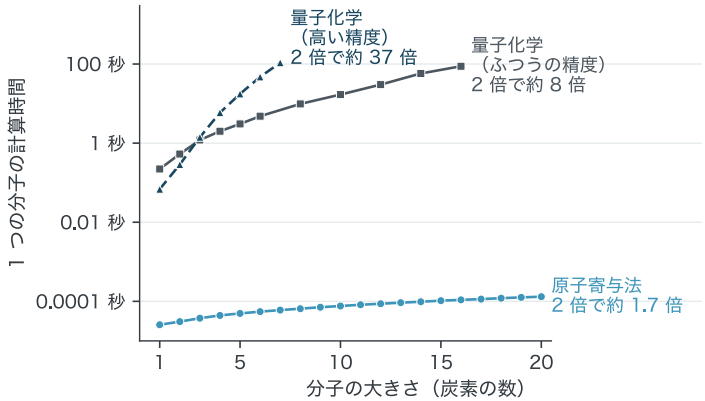


図 6-9 分子の大きさ（炭素の数）と、分子 1 個の計算にかかる時間。縦の目盛りは、1 つ上がるごとに 100 倍になります。炭素 1〜2 個の小さな分子では、計算の準備にかかる時間が目立つので、ふつうの精度の線が高い精度の線より上にあります。著者のノートパソコンのコア 1 つで測りました。ふつうの精度は第 2 章の近似 2 と同じ近似です。時間も増え方も、分子や計算機によって変わります。

答えは、ふつうの精度で 8 倍ほど、高い精度で 37 倍ほどです。線の傾きが急なものほど、時間が大きく増えます。量子化学の線は、原子寄与法の線よりずっと急です。原子寄与法では、結合表の行が増えたぶん、足す寄与の数が増えるだけだからです。

量子化学の時間が急に増えるのは、電子どうしのかかわりを計算するからです。電子の数が 2 倍になると、電子 2 個の組み合わせの数はおよそ 4 倍になります。実際の計算には、組み合わせの数のほかにも分子が大きくなるほど増える手間があるので、ふつうの精度で測った時間は 4 倍より大きく、8 倍ほどになりました。高い精度の近似は、電子どうしのかかわりをもっとていねいに数えるので、さらに急に増えます。炭素 6 個のヘキサンでは、ふつうの精度で 5 秒ほど、高い精度で 47 秒ほどかかりました。原子寄与法なら 0.0001 秒もかかりません。



テオ クラスの人数が 2 倍になると、全員どうしの握手の回数はおよそ 4 倍になる。電子の組み合わせも、同じ数え方で増える。

スーパーコンピュータでも

もっと速いコンピュータを使えば、どんな分子でも計算できるのでしょうか。日本のスーパーコンピュータ「富岳」（図 6-10）には、計算をする部品（コア）が約 763 万個あります（文献 6-6）。コア 1 つが 1 秒間にできる計算の回数は、著者のノートパソコンと同じくらいです。富岳の速さは、コアの数から来ているのです。



図 6-10 スーパーコンピュータ「富岳」（理化学研究所 計算科学研究センター、神戸市）。並んだ箱の 1 つ 1 つに、計算機がたくさん入っています。（写真：Barsaka2, CC0）

仮に、富岳のすべてのコアを、1 つの計算にむだなく使えたとします。実際には、コアどうしが計算の結果をやりとりする時間などがあるので、これはいちばんうまくいったときの見積もりです。それでも、計算は 763 万倍速くなるだけです。そのとき、計算できる分子の大きさが何倍になるかを、図 6-9 で測った増え方から見積もりました（図 6-11）。

原子寄与法なら、時間は分子の大きさにほぼ比例するので、計算できる分子の大きさも、ほぼ同じだけ大きくなります。ところが量子化学では、763 万倍速くなっても、ふつうの精度で 200 倍ほど、高い精度では 10〜20 倍ほどの大きさの分子までしか計算できません。ふつうの精度では、大きさが 2 倍で時間が 8 倍（$2 \times 2 \times 2$）なので、大きさが 200 倍なら時間は 800 万倍（$200 \times 200 \times 200$）になり、763 万倍の速さを使い切ってしまうからです。分子が大きくなると計算時間が急に増えるので、コンピュータが速くなっても、計算できる分子はわずかしか大きくならないのです。富岳のようなスーパーコンピュータでも、何でもていねいに計算できるわけではありません。

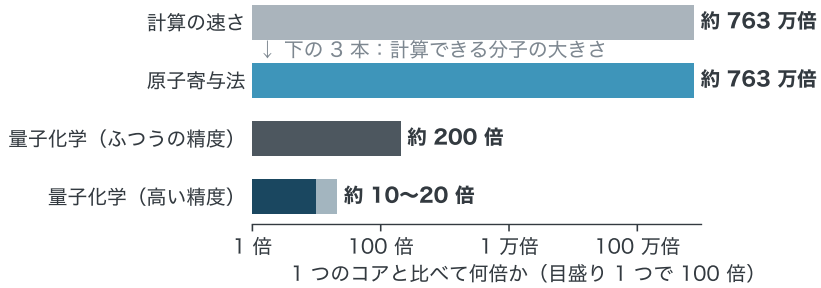


図 6-11 富岳のすべてのコアをむだなく使えたとき、コア 1 つと同じ時間で計算できる分子の大きさが何倍になるか。図 6-9 の増え方から見積もりました。高い精度の幅は、炭素 7 個までの増え方をのばすと 20 倍ほど、もっと大きな分子での増え方では 10 倍ほどになることを表します。

目的に合わせてモデルを選ぶ

そこで研究の現場では、目的に合わせてモデルを選びます。何億個もの分子をふるいにかけるときは、原子寄与法やドッキングのような速いモデルを使います。誤差はありますが、何億個をまとめて調べられることのほうが大事だからです。ふるいに残った分子は、分子動力学や量子化学のような、省いたものの少ない遅いモデルで調べ直すこともあります。そして最後に残った候補を、実際に作って実験で確かめます（図 6-12）。第 4 章で、原子寄与法と分子動力学を使い分けたのも、同じ考え方です。

実験で確かめた結果は、そこで役目を終えるわけではありません。次の節では、実験の結果が計算に戻ってくるところを見ます。

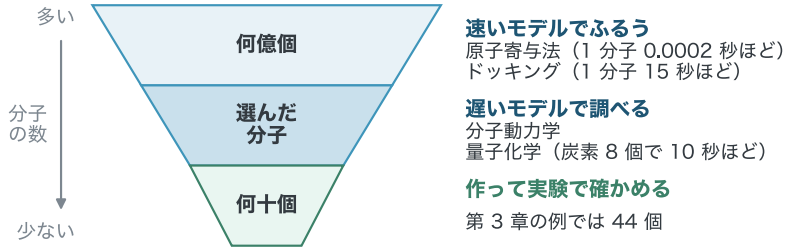


図 6-12 速いモデルでふるい、遅いモデルで調べ、実験で確かめる。じょうごの段の幅は形だけで、分子の数の比ではありません。ドッキングの時間は第 3 章の例です。

6.4 計算と実験は両輪

実験で測った値は、計算のモデルを改良するデータになります。第 4 章の寄与も、第 5 章の規則も、実験で測った値から決めたものでした。測った値が増えれば、学習し直して、パラメータや規則を新しくできます。直したモデルで、また候補を絞ります。こうして、計算と実験は輪になって回ります（図 6-13）。

要点

計算は実験の前に候補を絞り、実験は計算の答えを確かめて、モデルを直すデータを生みます。計算と実験は、どちらかが要らなくなる関係ではなく、輪になって回る両輪です。

天気予報も、同じようにして当たるようになってきました。序章で見たように、数値予報は、出した予報をあとで観測と比べる答え合わせをくり返し、モデルを直してきました。その積み重ねを、台風の進路予報の誤差で見てみます。気象庁は、予報した台風の中心と実際の中心との距離を 1 年分平均して、予報の誤差を毎年公表しています（図 6-14）（文献 6-7）。

いまでは、3 日先を予報したときの誤差が、1980 年代に 1 日先を予報したときの誤差より小さくなっています。その年に予報しやすい台風が多いかどうかで、誤差は年ごとに上下しますが、長い目で見ると、予報の精度は上がってきました。気象庁

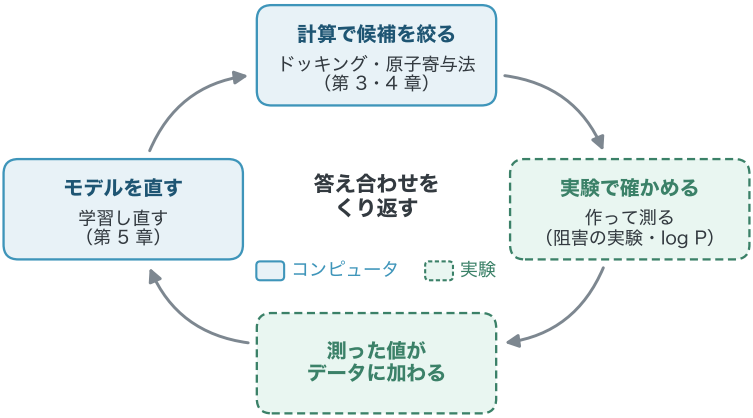


図 6-13　計算と実験が回るサイクル。青い実線の箱がコンピュータの仕事、緑の破線の箱が実験の仕事です。

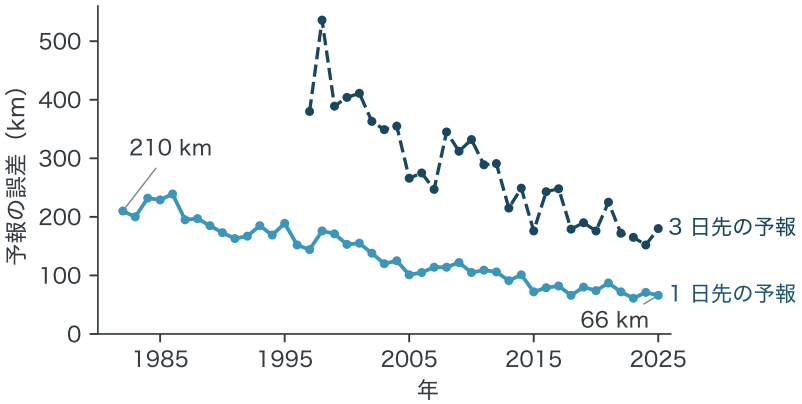


図 6-14 台風の進路予報の誤差（年平均）の移り変わり。気象庁の公表値です（文献 6-7）。

は、その理由として、数値予報のモデルの改良や、その使い方の改善などを挙げています。

答え合わせをくり返し、外れた理由を探してモデルを直す。天気予報も計算化学も、こうして進歩してきました。

! よくある誤解

「計算があれば、実験はいらない」と考えるのは誤りです。この章で見たように、計算には誤差があり、ドッキングの上位の候補でも、実験で試すと多くは外れでした。それに、計算に使う寄与や規則は、実験で測った値から決めています（第 4 章・第 5 章）。実験がなければ、計算の答え合わせも、モデルの改良もできません。

6.5 AI を使いこなすのは化学が分かる人

最後に、答え合わせでいちばん外れたグリシンを見ます。計算は −0.97、実測は −3.21 で、誤差は +2.24 です。なぜ、ここまで外れたのでしょうか。

グリシンは、いちばん小さなアミノ酸です。アミノ酸は、水の中では、カルボキシ基の COOH が H^+ を放して COO^- になり、アミノ基の NH_2 が H^+ を受けとって NH_3^+ になります。このように、1 つの分子の中にプラスとマイナスの両方の電荷をもつイオンを、双性イオンと呼びます（図 6-15）。



コンピュータに渡した形
計算 −0.97

水の中の形（双性イオン）
実測 −3.21

図 6-15　グリシンの、コンピュータに渡した形（左）と、水の中の形（右）。O と N をローズで塗りました。

ところが、コンピュータに渡した結合表は、イオンになる前の形でした。第 4 章では、$\log P$ はイオンになる前の形どうしで比べる数だと述べました。けれどもグリシンは、水の中ではほぼいつも双性イオン（分子全体の電荷は 0）で、イオンになる前の形をほとんどとりません。だから、グリシンの $\log P$ を測るときに油と水に分かれるのは、双性イオンです。コンピュータには、グリシンが実際にはほとんどとらない形を渡したことになります。

双性イオンの COO^- や NH_3^+ は、水分子にしっかり囲まれて水和するので、グリシンは油の層にはほとんど移りません。そのため、実測の $\log P$ はとても小さくなります。計算にはこの形のちがいが入っていないので、油の側に大きく外れたと考えられます。

この外れに気づき、理由を説明できたのは、「アミノ酸は水の中で双性イオンになる」という化学の知識があったからです。コンピュータは、渡された結合表のとおりに寄与を足すだけです。渡した形が水の中の形とちがっていても、コンピュータはそのまま数を返します。答えがおかしいと気づき、モデルが何を省いたせいかを考えるのは、使う人の仕事です。

この知識があると、第 4 章のリレンザの数にも目が向きます。リレンザも水の中では双性イオンで、−3.79 はグリシンと同じ向きにずれた見積もりだと考えられます（図 6-16）。それでも、「水になじみすぎて、飲み薬に向かない」という結論は変わりません。

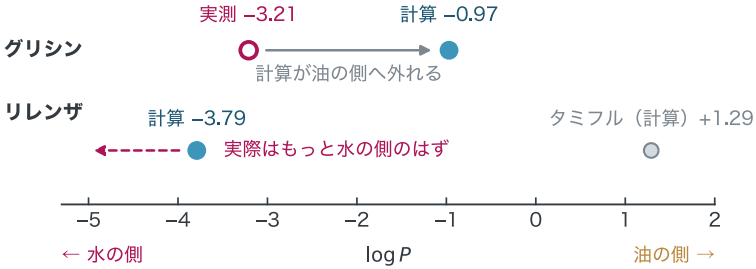


図 6-16　イオンになる前の形で計算した数と、実際のずれの向き。リレンザの矢印は向きだけを示しています。

第 5 章の機械学習や、人工知能（AI）でも同じです。学んだデータの範囲の外でも、AI はもっともらしい数を返します。その数を信じてよいかどうかを判断するには、化学の知識が必要です。計算や AI がどれだけ速くなっても、何を計算させるかを決め、答えを確かめる仕事は、化学が分かる人に残ります。



AI がいれば、化学の勉強はいらなくなる？

逆だ。何を計算させ、答えが正しいかを判断できるのは、化学が分かる人だけだ。





化学とコンピュータと AI。3 つそろうと強いね。

研究の現場から：答えを伏せた予測の腕比べ

計算の答え合わせを、研究者全体で公平に行うしくみがあります。タンパク質の立体構造の予測では、1994 年から 2 年ごとに、CASP と呼ばれる腕比べが開かれています（文献 6-8、6-9）。手順は 図 6-17 のとおりです。

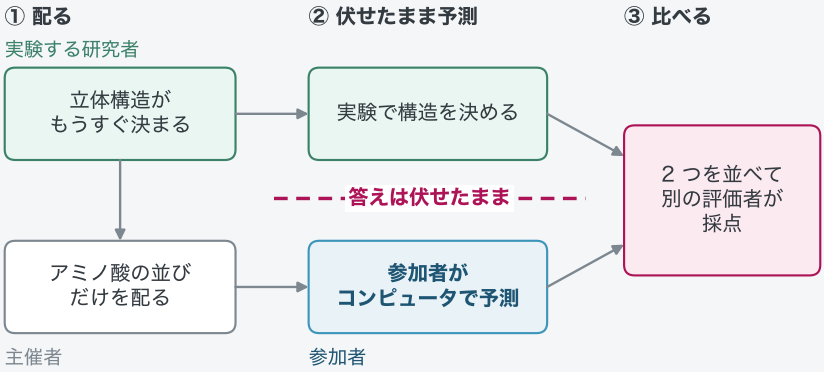


図 6-17 腕比べの流れ。緑の箱が実験、青の箱がコンピュータのすることです。

答えを伏せておけば、予測する人が、知っている答えにモデルを合わせることはできません。2020 年の第 14 回では、AlphaFold2（第 5 章）の予測が、多くのタンパク質で、実験で決めた構造に匹敵する正確さに達しました（文献 6-10、6-11）。第 4 章で見た $\log P$ の予測の腕比べも、同じ考え方で開かれたものです。予測が本当に当たるかどうかは、答えを知らない分子で確かめてはじめて分かります。

この本では、薬の候補を探し、体に届くかを予測する場面で、計算と実験が組むようすを見てきました。同じ組み方は、新しい材料を探すときにも使われています。終章では、小さな穴に二酸化炭素をとどめる材料を、作る前にコンピュータで絞りこむ研究を見ます。

まとめ

- コンピュータの計算を実測と答え合わせすると、誤差の大きさが分かります。クリッペン法の $\log P$ は、身近な分子 25 個で、誤差の大きさの平均が 0.62 でした。実測にも、報告ごとの幅があります。
- 外れ方の理由は、モデルが省いたものにあります。炭素と水素ばかりの分子や、グリシンのように水の中でイオンになる前の形をほとんどとらない分子（双性イオン）は、原子寄与法の苦手な分子の例です。
- 省いたものの少ないモデルほど、分子が大きくなると計算時間が急に増えます。スーパーコンピュータでも何でも計算できるわけではないので、目的に合わせて速いモデルと遅いモデルを使い分けます。
- コンピュータで候補を絞り、実験で確かめ、測った値でモデルを直します。計算と実験は両輪で、答えがおかしいと気づけるのは、化学が分かる人です。

終章

計算化学が切り拓く未来



発電所や工場の煙突（図 終-1）から出る排ガスには、CO_2 のほかに、窒素や水蒸気がたくさん混じっています。この混ざったガスから、CO_2 だけを取り出せないでしょうか。CO_2 と窒素は、どちらも目に見えない小さな分子です。この 2 つを分けるには、分子と同じくらい小さな穴をもつ材料が役に立ちます。



図 終-1　発電所の煙突（アメリカ）。（写真：Dhudson4god, パブリックドメイン）

ここまでの章では、主に薬を例に、コンピュータが何をしているかを見てきました。最後の章では、材料に目を向けます。新しい読み方は要りません。序章で天気予報を読んだ 3 つの問いで、CO_2 をつかまえる材料を探す計算を読んでいきます。そのうえで、この本全体をふり返ります。

◎ この章で分かること

- 金属イオンと有機分子でできた結晶（MOF）は、小さな穴の壁に CO_2 をとどめます
- コンピュータは、合成する前に、何十万個もの候補から材料を絞りこみます
- 初めて出会う計算も、この本の 3 つの問いで読めます

高校で習ったこと：結晶の構造（化学）。習っていなくても読めます

穴で CO_2 をつかまえる

金属イオンと有機分子を溶液の中で混ぜると、金属イオンを角にして、長い有機分子が橋をかけるようにつながっていきます。こうしてできた結晶の中には、分子が入れるほどの小さな穴が、規則正しく並びます（図 終-2）。このような結晶を、金属有機構造体（metal–organic framework、略して MOF）と呼びます（文献 終-1）。結晶の中の小さな穴を、細孔（pore）といいます。

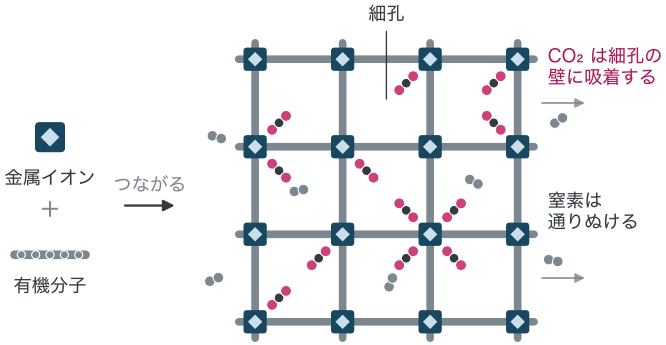


図 終-2　MOF のつくりと、細孔に CO_2 がとどまるようす（模式図）。紺の四角が金属イオン、灰色の棒が有機分子です。CO_2 は 3 つの丸（O・C・O）、窒素は灰色の 2 つの丸で描いています。

ガスの分子が細孔に入ると、壁の原子と引き合って、壁ぎわにとどまることがあります。このように、気体の分子が固体の表面に集まってとどまることを、吸着（adsorption）と呼びます。細孔の形と壁の原子を選べば、CO_2 はよく吸着し、窒素はほとんど吸着せずに通りぬける MOF を作れます。細孔の壁の面積はとても広く、1999 年に作られた MOF の 1 つは、わずか数グラムの中に、サッカー場ほどの広さの表面をもっていました（文献 終-2）。

2025 年のノーベル化学賞は、「金属有機構造体の開発」で、北川進、リチャード・ロブソン、オマー・ヤギーの 3 人に贈られました（文献 終-1）。MOF を作り、その可



テラ　冷蔵庫の脱臭剤の炭も、小さな穴の壁ににおいの分子を吸着しているよ。MOF なら、細孔の形まで設計できる！

能性を示した研究への賞で、計算の賞ではありません。けれども、作る MOF を選ぶところで、計算化学が使われています。実物の結晶は、図 終-3 のように見えます。

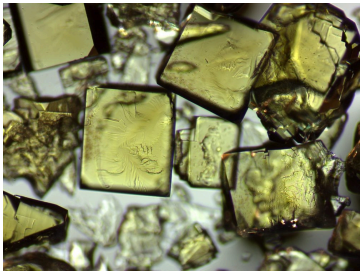


図 終-3 MOF の結晶（MOF-5 の仲間）。（写真：Marcin Wiszniewski（CNBCh）, CC BY-SA 4.0）

金属イオンと有機分子の組み合わせは数えきれないほどあり、化学者がこれまでに作った MOF だけでも何万種類にのぼります（文献 終-1）。1 つずつ合成して、CO_2 をどれだけ吸着するかを測っていたのでは、とても追いつきません。どの組み合わせがよいのかを、作る前に知る方法はないでしょうか。

3 つの問いで読む

作る前に選ぶための計算を、序章と同じ 3 つの問いで読んでみます。

何をモデルにしたか

金属イオンと有機分子がつながった MOF の本体を、骨格といいます。骨格の原子の位置は、部品をつないで組み立てれば、まだ作られていない MOF でも決められます。骨格はほとんど動かないものとして原子をその位置に止めておき、CO_2 は、第 1 章と同じように原子をつぶとして扱います。CO_2 と骨格の原子が引き合う強さやぶつかりは、原子どうしの距離から計算し、電子の動きは省きます。



テオ この計算は電子を省いた。CO_2 が細孔の壁と引き合う強さを、電子から求めるなら、第 2 章の量子化学の出番だ。

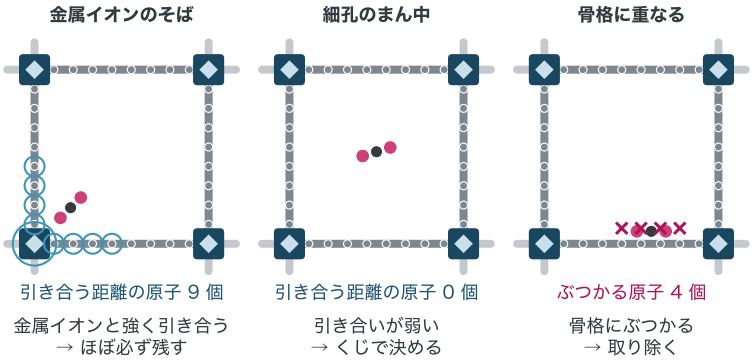


図 終-4　細孔の 3 か所に CO_2 を置いてみる（考え方を見せるための、平面の細孔の模型）。青い輪は、置いた CO_2 と引き合う距離にある骨格の原子、× は CO_2 とぶつかる原子です。

何をくり返したか

細孔のでたらめな場所に CO_2 を 1 つ入れてみるか、入っている CO_2 を 1 つ取り出してみます。そのとおりにするかどうかは、引き合いの強さと圧力から、くじで決めます（図 終-4）。コンピュータは、この 1 回を何百万回もくり返します。

やがて、細孔に入っている CO_2 の数は、ある数のまわりで増えたり減ったりするようになります。この数の平均が、その圧力での吸着量です。CO_2 がどこにとどまりやすいかも見えてきます。

コンピュータは、この計算を何十万個もの MOF で行えます。2012 年に報告された研究では、102 種類の部品から 13 万 7953 個の MOF を組み立て、都市ガスの主成分であるメタンをどれだけためられるかを計算しました（文献 終-3）。それまでに知られていたどの材料よりも多くためると予測された MOF が 300 個以上見つかり、そのうちの 1 つを合成すると、予測どおりの量をためました。第 3 章のバーチャルスクリーニングと同じように、コンピュータの中でふるいにかけ、上位だけを作って確かめたのです。

どこまで信じられるか

コンピュータの中で組み立てた MOF が、実際に作れるとは限りません。作れたとしても、こわれずに使えるかどうかは、また別の問題です。1989 年に作られた最初の MOF も、すぐにこわれてしまう不安定なものでした（文献 終-2）。骨格を止めたので、ガスが入ると形を変えるやわらかい MOF も、この計算では扱いにくくなります。

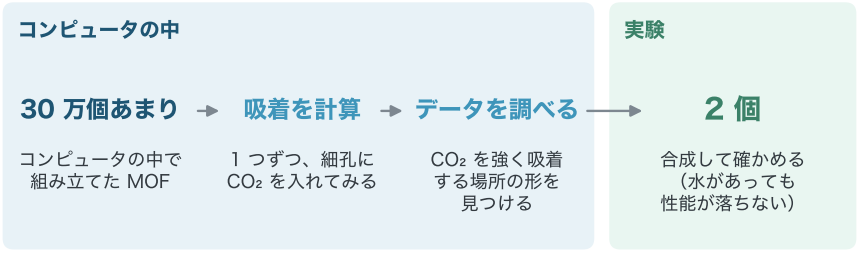


図 終-5　湿った排ガスから CO_2 を回収する MOF を探した例（2019 年）（文献 終-4）。

省いたものは、ほかにもあります。本物の排ガスには水蒸気が混じっていて、水が CO_2 と同じ場所を取り合うことがあります（文献 終-4）。

予想してみよう　**金属イオンのそばは、CO_2 と強く引き合う場所の 1 つです。水蒸気の混じった排ガスから CO_2 を取り出すときも、金属イオンのそばの場所をもつ MOF を選べばよいでしょうか？**

水で答えが変わる

実際の研究者たちは、金属イオンのそばを選びませんでした。2019 年に報告された研究を見てみます（図 終-5）。

この研究では、コンピュータの中で組み立てた 30 万個を超える MOF について、CO_2 の吸着を計算しました（文献 終-4）。多くの MOF は、乾いたガスでは CO_2 と窒素をよく分けますが、水蒸気が混じると、CO_2 が吸着する場所を水が先に占めてしまいます。

そこで、計算で得たたくさんのデータを調べ、CO_2 を強く吸着する場所の形を、いくつかの種類に分けました。すると、金属イオンのそばの場所は、水も引きつけやすい傾向がありました。選んだのは、いちばん疎水性の高い形（第 4 章）で、ベンゼン環のような平らな環が 2 枚、すき間をあけて重なった、金属イオンをふくまない形です。

この形をもつ MOF を 2 つ合成すると、2 つとも、水が混じっても CO_2 を吸着する性能が落ちず、市販の材料のいくつかを上回りました。実際の工場で使えるかどうかの試験は、これからの課題とされています。

強く引き合う場所を選ぶという素直な見通しは、水が混じると通用しませんでした。その見通しを直したのは、たくさんの計算のデータと、合成して確かめる実験です。第 6 章で見たように、計算には外れがあるので、実験と組み合わせて使います。

4 つのはたらきをもう一度

序章では、計算化学のはたらきを、見る・探す・予測する・学ぶの 4 つに分けました。MOF を探す研究には、この 4 つがすべて入っています（図 終-6）。細孔の中で CO_2 がどこにとどまるかを見て、吸着する場所で CO_2 が引き合う強さを、電子まで計算して求めます（文献 終-4）。何十万個もの MOF の中から候補を探し、合成する前に吸着量を予測します。さらに、計算で得たデータから、よく吸着する場所の形を学びます。

図 終-6 のまん中の列を、上から読んでみてください。どの章でも、コンピュータは、決めた規則で同じ計算をくり返していました。力から次の位置を求める、電子の雲を求め直す、分子を置いて点数にする、寄与を足す、誤差が小さくなるように規則を直す、といったことです。どれも 1 回分なら手でもできる計算ですが、コンピュー



図 終-6　この本で見た 4 つのはたらきと、各章でコンピュータがくり返したこと（中）、MOF を探す研究での使われ方（右）。絵は簡単にしたものです。

タは、それを何百万回、何億回とくり返します。

要点

計算化学とは、分子のモデルを作り、その規則にしたがってコンピュータにくり返し計算させ、実験の前に結果を予測することです。何を省いたモデルなのかを知り、最後は実験で確かめれば、計算は、作る前に候補を絞りこむ力になります。

計算化学が切り拓くのは、作る前に選べる化学です。薬の候補も、CO_2 を回収する材料も、何十万個も作って試すことはできません。コンピュータの中で試して絞りこみ、実験で確かめる。この組み合わせで、化学者が調べられる分子と材料の数は、ずっと増えていきます。

計算化学を学ぶには

計算化学を学ぶのに、特別な入口はありません。高校の化学・物理・数学・情報の続きを大学で学んでいけば、その先に計算化学があります（図 終-7）。

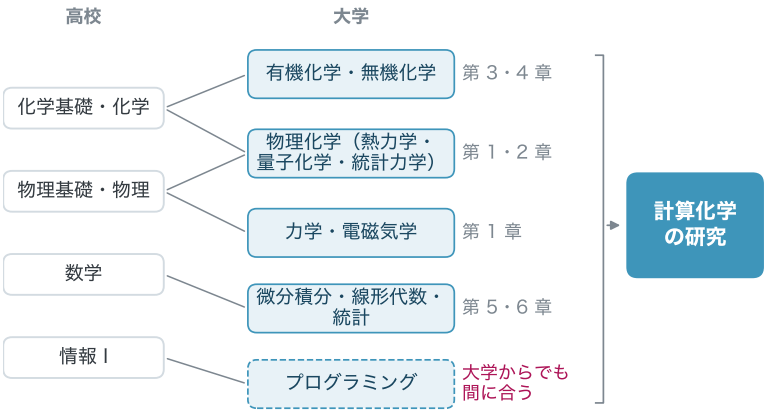


図 終-7　高校の科目から、大学で学ぶ科目を通って、計算化学へ。大学の科目の名前は、化学系の学科でよく見られるものです。右の小さな字は、つながりの深いこの本の章です。



テラ　1回1回は、この本でみんなも手で解いてきた計算だよ。コンピュータの仕事の中身は、もう知っているよね！

プログラミングは、今できなくても心配はいりません。大学に入ってから、あるいは研究室に入ってから身につける人がたくさんいます。それより先に、分子を見る目を育ててください。計算の答えがおかしいことに気づけるのは、化学が分かっている人です。

著者の研究室（やまラボ）では、MOF の細孔の形によって CO_2 の吸着がどう変わるかや、やわらかい MOF の細孔が開くしくみを、分子動力学で調べています。研究室のことは、「はじめに」で紹介した本のページからたどれます。



薬を探すのも、CO_2 をつかまえる材料を探すのも、コンピュータがしていることは同じなんだね。

モデルを決めて、同じ計算をくり返す。最後は実験で確かめる。どの章でも、そうだった。





じゃあ、プログラミングはあとからでいいから、まずは化学をしっかりやろう！

まとめ

- MOF は、金属イオンと有機分子がつながった結晶で、細孔の壁に CO_2 を吸着します。2025 年のノーベル化学賞は、MOF を作った研究に贈られました。
- コンピュータは、骨格をほとんど動かないものとし、CO_2 をつぶとして、細孔に CO_2 を出し入れする計算を何百万回もくり返して、吸着量を見積もります。
- 何十万個もの MOF をコンピュータの中で組み立てて計算し、合成する候補を絞りこめます。ただし、合成できるか、こわれないか、水が混じっても使えるかは、実験で確かめます。
- 見る・探す・予測する・学ぶの 4 つのはたらきは、どれも、モデルを決めて同じ計算をくり返す仕事です。

最初の絵を、もう一度

絵のなかの部品は、この本で見てきたものです。今なら、読めるはずです。



① **Ψ の杖**
電子の雲を、計算で求める（第 2 章）

② **砂時計**
1 コマずつ、時間を進める（第 1 章）

③ **鳥の大群**
同じ計算を、同時にくり返す（第 1 章）

④ **らせんの髪**
タンパク質の形を予測する（第 5 章）

⑤ **波打つ床**
エネルギーのいちばん低い形を探す（第 2 章）

何をモデルにしたか、何をくり返したか、どこまで信じられるか。次は、この 3 つを自分で問う番だ。

テオ



精霊は、もう 1 種類いるんだよ。量子コンピュータの精霊は、まだ誰も使いこなせていない。次に挑むのは、あなたかも！

参考文献



本文で「(文献 4-1)」のように引いた文献を、章ごとに並べました。DOI は、論文を探すときの番号です（https://doi.org/ の後に続けると、論文のページが開きます)。

序章

序-1　気象庁「数値予報とは」. https://www.jma.go.jp/jma/kishou/know/whitep/1-3-1.html

序-2　気象庁「令和 7 年度 数値予報解説資料集」（2025 年）. 1.3 データ同化、1.7.1 全球解析、1.7.2 全球モデル、4.1 数値予報システムの概要一覧表（2025 年 9 月 30 日現在）. https://www.jma.go.jp/jma/kishou/books/nwpkaisetu/R7/1_7_2.pdf

序-3　気象庁「気象に関する数値予報モデルの種類」. https://www.jma.go.jp/jma/kishou/know/whitep/1-3-4.html

序-4　気象庁「積乱雲に伴う激しい現象の住民周知に関するガイドライン」（2015 年）. https://www.jma.go.jp/jma/kishou/know/toppuu/cb_guideline.pdf

序-5　気象庁「数値予報の歴史」. https://www.jma.go.jp/jma/kishou/know/whitep/1-3-2.html

序-6　T. Noda, “Native morphology of influenza virions”, *Front. Microbiol.* **2**, 269 (2012). doi:10.3389/fmicb.2011.00269

序-7　J. N. Varghese, W. G. Laver, P. M. Colman, “Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution”, *Nature* **303**, 35–40 (1983). doi:10.1038/303035a0

序-8　A. R. Hoy, P. R. Bunker, “A precise solution of the rotation bending Schrödinger equation for a triatomic molecule with application to the water molecule”, *J. Mol. Spectrosc.* **74**, 1–8 (1979). doi:10.1016/0022-2852(79)90019-5（値は NIST Computational Chemistry Comparison and Benchmark Database の水の構造の実験値から）

序-9　The Royal Swedish Academy of Sciences, “The Nobel Prize in Chemistry 2013”, press release (2013 年 10 月 9 日). NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2013/press-release/

序-10　The Royal Swedish Academy of Sciences, “The Nobel Prize in Chemistry 1998”, press release (1998 年 10 月 13 日). NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/1998/press-release/

序-11 The Royal Swedish Academy of Sciences, “The Nobel Prize in Chemistry 2024”, press release (2024 年 10 月 9 日). NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2024/press-release/

第 1 章

1-1 A. R. Hoy, P. R. Bunker, “A precise solution of the rotation bending Schrödinger equation for a triatomic molecule with application to the water molecule”, *J. Mol. Spectrosc.* **74**, 1–8 (1979). doi:10.1016/0022-2852(79)90019-5（値は National Institute of Standards and Technology, Computational Chemistry Comparison and Benchmark DataBase の、水の実験の構造の表から. https://cccbdb.nist.gov/）

1-2 S. A. Hollingsworth, R. O. Dror, “Molecular dynamics simulation for all”, *Neuron* **99**, 1129–1143 (2018). doi:10.1016/j.neuron.2018.08.011

1-3 National Institute of Standards and Technology, NIST Chemistry WebBook, SRD 69「Water」の振動のデータ. 元は T. Shimanouchi, *Tables of Molecular Vibrational Frequencies, Consolidated Volume I*, NSRDS-NBS 39 (1972). https://webbook.nist.gov/cgi/cbook.cgi?ID=C7732185&Mask=800

1-4 M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, 2nd ed. (Oxford University Press, 2017). doi:10.1093/oso/9780198803195.001.0001（5 章「Some tricks of the trade」、6 章「Long-range forces」）

1-5 R. L. Baldwin, “Temperature dependence of the hydrophobic interaction in protein folding”, *Proc. Natl. Acad. Sci. USA* **83**, 8069–8072 (1986). doi:10.1073/pnas.83.21.8069

1-6 J. R. Schames, R. H. Henchman, J. S. Siegel ほか, “Discovery of a novel binding trench in HIV integrase”, *J. Med. Chem.* **47**, 1879–1881 (2004). doi:10.1021/jm0341913

1-7 J. D. Durrant, J. A. McCammon, “Molecular dynamics simulations and drug discovery”, *BMC Biol.* **9**, 71 (2011). doi:10.1186/1741-7007-9-71

第 2 章

2-1 P. J. Mohr, D. B. Newell, B. N. Taylor ほか, “CODATA recommended values of the fundamental physical constants: 2022”, *Rev. Mod. Phys.* **97**, 025002 (2025). doi:10.1103/RevModPhys.97.025002

2-2 A. Kramida, Yu. Ralchenko, J. Reader, NIST ASD Team, NIST Atomic Spectra Database (version 5.12), National Institute of Standards and Technology (2024). Na I・Li I の準位と線、H I・He I・He II のイオン化エネルギー（2026 年 10 月 22・23 日に参照）. doi:10.18434/T4W30F

2-3 National Center for Biotechnology Information, PubChem “Periodic Table of Elements”（電気陰性度、ポーリングの値。2026 年 10 月 22 日に参照）. https://pubchem.ncbi.nlm.nih.gov/periodic-table/

2-4 National Center for Biotechnology Information, PubChem Compound Summary for CID 5280489, Beta-Carotene（2026 年 10 月 22 日に参照）. https://pubchem.ncbi.nlm.nih.gov/compound/5280489

2-5 R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel ほか, “Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach”, *Nat. Mater.* **15**, 1120–1127 (2016). doi:10.1038/nmat4717

2-6 Y. Oba, K. Konishi, D. Yano ほか, “Resurrecting the ancient glow of the fireflies”, *Sci. Adv.* **6**, eabc5705 (2020). doi:10.1126/sciadv.abc5705

2-7 N. Nakatani, J. Hasegawa, H. Nakatsuji, “Red light in chemiluminescence and yellow-green light in bioluminescence: color-tuning mechanism of firefly, *Photinus pyralis*, studied by the symmetry-adapted cluster-configuration interaction method”, *J. Am. Chem. Soc.* **129**, 8756–8765 (2007). doi:10.1021/ja0611691

第 3 章

3-1 厚生労働省「第 7 回 医薬品の迅速・安定供給実現に向けた総合対策に関する有識者検討会 資料 6 事務局参考資料」（2023 年 1 月 13 日）. 11〜13 枚目（2015〜2019 年の数を使う）. 元は日本製薬工業協会「DATA BOOK 2021」. https://www.mhlw.go.jp/content/10807000/001036959.pdf

3-2 J.-J. Shie, J.-M. Fang, “Development of effective anti-influenza drugs: congeners and conjugates – a review”, *J. Biomed. Sci.* **26**, 84 (2019). doi:10.1186/s12929-019-0567-0

3-3 J. N. Varghese, W. G. Laver, P. M. Colman, “Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution”, *Nature* **303**, 35–40 (1983). doi:10.1038/303035a0

3-4 R. J. Russell, L. F. Haire, D. J. Stevens ほか, “The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design”, *Nature* **443**, 45–49 (2006). doi:10.1038/nature05114（立体構造のデータ：PDB 2HU4）

3-5 S. Evteev, D. Nilov, A. Polenova ほか, “Bifunctional inhibitors of influenza virus neuraminidase: molecular design of a sulfonamide linker”, *Int. J. Mol. Sci.* **22**, 13112 (2021). doi:10.3390/ijms222313112

3-6 J. Lyu, S. Wang, T. E. Balius ほか, “Ultra-large library docking for discovering new chemotypes”, *Nature* **566**, 224–229 (2019). doi:10.1038/s41586-019-0917-9

3-7 C. Gorgulla, A. Boeszoermenyi, Z.-F. Wang ほか, “An open-source drug discovery platform enables ultra-large virtual screens”, *Nature* **580**, 663–668 (2020). doi:10.1038/s41586-020-2117-z

3-8 M. Su, Q. Yang, Y. Du ほか, “Comparative assessment of scoring functions: the CASF-2016 update”, *J. Chem. Inf. Model.* **59**, 895–913 (2019). doi:10.1021/acs.jcim.8b00545

3-9 M. von Itzstein, W.-Y. Wu, G. B. Kok ほか, “Rational design of potent sialidase-based inhibitors of influenza virus replication”, *Nature* **363**, 418–423 (1993). doi:10.1038/363418a0

3-10 M. von Itzstein, J. C. Dyason, S. W. Oliver ほか, “A study of the active site of influenza virus sialidase: an approach to the rational design of novel anti-influenza drugs”, *J. Med. Chem.* **39**, 388–391 (1996). doi:10.1021/jm950294c

3-11 C. U. Kim, W. Lew, M. A. Williams ほか, “Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity”, *J. Am. Chem. Soc.* **119**, 681–690 (1997). doi:10.1021/ja963036t

3-12 P. J. Collins, L. F. Haire, Y. P. Lin ほか, “Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants”, *Nature* **453**, 1258–1261 (2008). doi:10.1038/nature06956

3-13 S. A. Hollingsworth, R. O. Dror, “Molecular dynamics simulation for all”, *Neuron* **99**, 1129–1143 (2018). doi:10.1016/j.neuron.2018.08.011

第 4 章

4-1 L. M. R. Cass, C. Efthymiopoulos, A. Bye, “Pharmacokinetics of zanamivir after intravenous, oral, inhaled or intranasal administration to healthy volunteers”, *Clin. Pharmacokinet.* **36** (Suppl. 1), 1–11 (1999). doi:10.2165/00003088-199936001-00001

4-2 World Health Organization, Prequalification Unit, “Notes on the design of bioequivalence study: Zanamivir” (2021 年 9 月). https://extranet.who.int/prequal/sites/default/files/document_files/BE_Zanamivir_September2021.pdf

4-3 J. J. Irwin, B. K. Shoichet, “ZINC — a free database of commercially available compounds for virtual screening”, *J. Chem. Inf. Model.* **45**, 177–182 (2005). doi:10.1021/ci049714+

4-4 S. A. Wildman, G. M. Crippen, “Prediction of physicochemical parameters by atomic contributions”, *J. Chem. Inf. Comput. Sci.* **39**, 868–873 (1999). doi:10.1021/ci9903071

4-5 M. Işık, T. D. Bergazin, T. Fox ほか, “Assessing the accuracy of octanol-water partition coefficient predictions in the SAMPL6 Part II log P Challenge”, *J. Comput. Aided Mol. Des.* **34**, 335–370 (2020). doi:10.1007/s10822-020-00295-0

4-6 C. U. Kim, W. Lew, M. A. Williams ほか, “Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity”, *J. Am. Chem. Soc.* **119**, 681–690 (1997). doi:10.1021/ja963036t

4-7 W. Li, P. A. Escarpe, E. J. Eisenberg ほか, “Identification of GS 4104 as an orally bioavailable prodrug of the influenza virus neuraminidase inhibitor GS 4071”, *Antimicrob. Agents Chemother.* **42**, 647–653 (1998). doi:10.1128/AAC.42.3.647

4-8 T. Scior, K. Cuanalo-Contreras, A. A. Islas ほか, “Targeting the human influenza A virus: the methods, limitations, and pitfalls of virtual screening for drug-like candidates including scaffold hopping and compound profiling”, *Viruses* **15**, 1056 (2023). doi:10.3390/v15051056

4-9 S. Evteev, D. Nilov, A. Polenova ほか, “Bifunctional inhibitors of influenza virus neuraminidase: molecular design of a sulfonamide linker”, *Int. J. Mol. Sci.* **22**, 13112 (2021). doi:10.3390/ijms222313112

4-10 Genentech, Inc., TAMIFLU (oseltamivir phosphate) prescribing information, U.S. National Library of Medicine, DailyMed (2025). https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=ee3c9555-60f2-4f82-a760-11983c86e97b

4-11 中外製薬株式会社「タミフルカプセル 75 電子添文」(2024 年 10 月改訂 第 2 版)、医薬品医療機器総合機構（PMDA）. https://www.pmda.go.jp/PmdaSearch/iyakuDetail/450045_6250021M1027_1_40

4-12 C. A. Lipinski, F. Lombardo, B. W. Dominy ほか, “Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings”, *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997). doi:10.1016/S0169-409X(96)00423-1

第 5 章

5-1 S. A. Wildman, G. M. Crippen, “Prediction of physicochemical parameters by atomic contributions”, *J. Chem. Inf. Comput. Sci.* **39**, 868–873 (1999). doi:10.1021/ci9903071

5-2 J. Sangster, “Octanol-water partition coefficients of simple organic compounds”, *J. Phys. Chem. Ref. Data* **18**, 1111–1229 (1989). doi:10.1063/1.555833

5-3 National Center for Biotechnology Information, PubChem Compound Summary for CID 5793, D-Glucose, “LogP”（2026 年 10 月 22 日に閲覧）. https://pubchem.ncbi.nlm.nih.gov/compound/5793. 値の元は M. F. Mazzobre, M. V. Román, A. Feo Mourelle ほか, “Octanol–water partition coefficient of glucose, sucrose, and trehalose”, *Carbohydr. Res.* **340**, 1207–1211 (2005). doi:10.1016/j.carres.2004.12.038

5-4 The Royal Swedish Academy of Sciences, “The Nobel Prize in Chemistry 2024”, press release（2024 年 10 月 9 日）と popular information “They have revealed proteins’ secrets

through computing and artificial intelligence”（2024 年 10 月）. https://www.nobelprize.org/prizes/chemistry/2024/

5-5 J. Jumper, R. Evans, A. Pritzel ほか, “Highly accurate protein structure prediction with AlphaFold”, *Nature* **596**, 583–589 (2021). doi:10.1038/s41586-021-03819-2

5-6 R. J. Russell, L. F. Haire, D. J. Stevens ほか, “The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design”, *Nature* **443**, 45–49 (2006). doi:10.1038/nature05114（立体構造のデータ：PDB 2HU4）

5-7 R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud ほか, “Automatic chemical design using a data-driven continuous representation of molecules”, *ACS Cent. Sci.* **4**, 268–276 (2018). doi:10.1021/acscentsci.7b00572

5-8 B. Burger, P. M. Maffettone, V. V. Gusev ほか, “A mobile robotic chemist”, *Nature* **583**, 237–241 (2020). doi:10.1038/s41586-020-2442-2

5-9 J. M. Stokes, K. Yang, K. Swanson ほか, “A deep learning approach to antibiotic discovery”, *Cell* **180**, 688–702 (2020). doi:10.1016/j.cell.2020.01.021

第 6 章

6-1 J. Sangster, “Octanol-water partition coefficients of simple organic compounds”, *J. Phys. Chem. Ref. Data* **18**, 1111–1229 (1989). doi:10.1063/1.555833

6-2 S. Kim, J. Chen, T. Cheng ほか, “PubChem 2025 update”, *Nucleic Acids Res.* **53**, D1516–D1525 (2025). doi:10.1093/nar/gkae1059（実測の log P は、各化合物のページの LogP の欄。2026 年 10 月に取得）

6-3 C. Hansch, A. Leo, D. Hoekman, *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*, American Chemical Society (1995).

6-4 J. Lyu, S. Wang, T. E. Balius ほか, “Ultra-large library docking for discovering new chemotypes”, *Nature* **566**, 224–229 (2019). doi:10.1038/s41586-019-0917-9

6-5 M. Su, Q. Yang, Y. Du ほか, “Comparative assessment of scoring functions: the CASF-2016 update”, *J. Chem. Inf. Model.* **59**, 895–913 (2019). doi:10.1021/acs.jcim.8b00545

6-6 理化学研究所 計算科学研究センター「富岳のシステム構成」. https://www.r-ccs.riken.jp/fugaku/system/

6-7 気象庁「台風予報の精度検証結果」「台風進路予報の年平均誤差の一覧表」（2025 年までの値）. https://www.data.jma.go.jp/typhoon/verification/index.html

6-8 J. Moult, J. T. Pedersen, R. Judson ほか, “A large-scale experiment to assess protein structure prediction methods”, *Proteins* **23**, ii–v (1995). doi:10.1002/prot.340230303

6-9 Protein Structure Prediction Center, “CASP experiments”（CASP1 1994 年〜CASP17 2026 年の一覧）. https://predictioncenter.org/

6-10 A. Kryshtafovych, T. Schwede, M. Topf ほか, “Critical assessment of methods of protein structure prediction (CASP)—Round XIV”, *Proteins* **89**, 1607–1617 (2021). doi:10.1002/prot.26237

6-11 J. Jumper, R. Evans, A. Pritzel ほか, “Highly accurate protein structure prediction with AlphaFold”, *Nature* **596**, 583–589 (2021). doi:10.1038/s41586-021-03819-2

終章

終-1 The Royal Swedish Academy of Sciences, “The Nobel Prize in Chemistry 2025”, press release (2025 年 10 月 8 日). https://www.nobelprize.org/prizes/chemistry/2025/

終-2 The Royal Swedish Academy of Sciences, “They have created new rooms for chemistry”, popular information (2025 年 10 月). https://www.nobelprize.org/prizes/chemistry/2025/

終-3 C. E. Wilmer, M. Leaf, C. Y. Lee ほか, “Large-scale screening of hypothetical metal–organic frameworks”, *Nat. Chem.* **4**, 83–89 (2012). doi:10.1038/nchem.1192

終-4 P. G. Boyd, A. Chidambaram, E. García-Díez ほか, “Data-driven design of metal–organic frameworks for wet flue gas CO_2 capture”, *Nature* **576**, 253–256 (2019). doi:10.1038/s41586-019-1798-7

用語の手引き



本文に出た専門の言葉を、50 音順に並べました。括弧の中は、初めて出てくる章です。

エネルギー準位、段差　（第 2 章）
電子がとれるエネルギーの段がエネルギー準位。2 つの準位の差を段差と呼び、電子が段を降りるとき、段差に見合った波長の光が出る。段を上るときは、その波長の光を受けとる。

塩橋　（第 3 章）
プラスのイオンとマイナスのイオン（たとえばアルギニンと COO^-）が近くで結びつくこと。静電気力と水素結合がいっしょにはたらく。

エントロピー　（第 1 章の発展）
並び方の数の多さを表す量。分子がまざるか分かれるかにもかかわる。

過学習　（第 5 章）
規則を複雑にしすぎて、学習に使ったデータには合うのに、新しいデータを外すこと。

活性部位　（第 1 章）
酵素が基質を受けとめて反応を進める、表面のくぼみ。多くの薬は、ここに結合して酵素のはたらきを止める。

機械学習、学習　（第 5 章）
データの表から、規則を表す数（パラメータ）を計算で見つけること。第 5 章では、予測と実測の誤差が小さくなるように、パラメータを少しずつ動かすことをくり返した。

偽陽性　（第 3 章）
計算では当たりに見えたのに、実験では結合しない分子（偽の当たり）。

吸収、透過　（第 4 章）
薬が腸などから血液に入ることが吸収。膜を通り抜けることが透過。

吸着、吸着量　（終章）
気体の分子が固体の表面に集まってとどまること。体に取りこまれる吸収とはちがう。

寄与、原子寄与法　（第 4 章）
$\log P$ のうち、原子のまとまりが受けもつ分が寄与。寄与を足して $\log P$ を見積もる方法が原子寄与法。

近似　（第 2 章）
本当の答えに近く、計算しやすい答えを使うこと。近似の選び方で、計算した色が少し変わる。

金属有機構造体（MOF）　（終章）
金属イオンを有機分子でつないだ、細孔の多い結晶。2025 年のノーベル化学賞の対象。

計算化学　（序章）
分子のモデルを作り、その規則にしたがってコンピュータにくり返し計算させ、実験の前に結果を予測する化学。コンピュータ化学とも呼ぶ。

結合する　（第 3 章）
分子が活性部位などに結びつくこと。

結合表　（第 4 章）
どの原子がどの原子とつながっているかの表。コンピュータに分子を渡すときの形。

ケモインフォマティクス （第 4 章）
分子をデータとして扱い、構造と性質の関係から性質を予測する分野。原子寄与法や機械学習が使われる。

原子タイプ （第 4 章）
原子を、種類とつながる相手で分けた区分。

格子 （序章）
大気などを区切ったマスの並び。

酵素、基質 （序章）
酵素は、ほかの分子の反応を進めるタンパク質。酵素が反応させる相手の分子が基質。

勾配降下法 （第 5 章）
誤差の地図の坂を下る向きに、パラメータを少しずつ動かすことをくり返す方法。

誤差 （第 4 章）
予測（計算）と実験の値の差。計算から実測を引くので、プラスにもマイナスにもなる。

誤差の大きさの平均 （第 6 章）
誤差の大きさ（マイナスの符号を外した値）の平均。計算の答え合わせの物差し。

骨格 （第 4 章）
分子のつながった本体。MOF の本体も骨格と呼ぶ（終章）。

骨格構造式 （第 3 章）
C と、C についた H を書かずに略した構造式。線の折れ目と端に C がある。

細孔 （終章）
MOF の結晶の中の小さな穴。

細胞膜 （第 4 章）
細胞を包む膜。中が油の層なので、水になじむ分子は通り抜けにくい。

散布図 （第 5 章）
2 つの量の組を、横と縦の位置で丸として並べたグラフ。

シアル酸 （第 3 章）
細胞の表面の糖の鎖の端にある分子。ヘマグルチニンが結合し、ノイラミニダーゼが切りはなす。

時間刻み、コマ （第 1 章）
分子動力学で、1 回に進める時間の幅が時間刻み、1 回分の計算が 1 コマ。1 コマで進める時間は 1000 兆分の 1 秒ほど。

自由エネルギー （第 4 章）
引き合う強さと、まわりの分子の並び方を合わせた、安定さの目安。

周期境界条件 （第 1 章）
箱の端から出たつぶが反対の端から入るようにして、同じ箱が果てしなく並んでいるとみなすやり方。

シュレディンガー方程式 （第 2 章）
電子の雲の形とエネルギーを決める式。これを解くのが量子化学。

人工知能（AI） （序章）
人のように判断したり、文章を書いたりするコンピュータのしくみ。機械学習は、その中心の考え方。

親水性、疎水性 （第 3・4 章）
水になじむ性質と、水になじみにくく油になじむ性質。疎水性は、水をはじく力をもつという意味ではない。

推奨値 （第 5 章）
何回も報告された実測値から、評価した人が決めた、いちばん確からしい値。

水素結合 （第 1 章）
O–H や N–H の H（わずかにプラス）が、となりの分子の O や N（わずかにマイナス）と引き合う結びつき。分子間力の中でとくに強い。

水和 （第 4 章）
分子やイオンが水分子に囲まれて安定になること。

数値予報 （序章）
コンピュータで大気の状態の変化を計算して、天気を予測すること。

スコア関数 （第 3 章）
結合の強さを見積もる点数の式。

スーパーコンピュータ （第 6 章）
たくさんの計算のコアをつないだ、とても速い計算機（富岳など）。

静電気力 （第 2 章）
プラスとマイナスの電荷が引き合い、同じ符号の電荷どうしが反発する力。

双性イオン （第 6 章）
1 つの分子の中に、プラスとマイナスの両方の電荷をもつイオン（グリシンなどのアミノ酸）。

阻害、阻害剤 （第 3 章）
酵素のはたらきを止めることが阻害、止める分子が阻害剤。

つぶとばね （第 1 章）
原子をつぶ、共有結合をばねとみなす分子のモデル。電子を省くので、結合は切れない。

電気陰性度、極性 （第 2 章）
結合の電子を引きつける強さが電気陰性度。差の大きい結合では電荷がかたより、極性が生まれる。

電子の雲 （第 2 章）
電子がどこにいそうかの広がり。電子は、原子核のまわりを惑星のように回っているのではない。

電子ボルト（eV） （第 2 章）
電子 1 個を 1 V で加速したときに得るエネルギー。段差（eV）× 波長（nm）はおよそ 1240。

吸入薬、経口薬 （第 4 章）
口から吸いこむ薬が吸入薬、飲みこむ薬が経口薬（飲み薬）。

ドッキング （第 1 章（研究の現場から））
候補の分子を活性部位に置いて、結合の強さを点数にする計算。

ニューラルネットワーク （第 5 章）
簡単な計算を何層にも重ねた、パラメータをとても多くもつ規則の形。文章を書く AI などで使われる。

熱運動 （第 1 章）
分子のたえまない動き。温度が高いほど激しい。

ノイラミニダーゼ （第 3 章）
インフルエンザウイルスの表面の酵素。細胞の表面の糖の鎖からシアル酸を切りはなし、ウイルスを細胞から離す。

バーチャルスクリーニング （第 3 章）
候補の分子の集まりをコンピュータの中でドッキングにかけ、作って試す候補を絞りこむこと。

波長 （第 2 章）
光の波の山から山までの長さ。光の色を決める。

パラメータ （第 4 章）
実測に合うように決めた、計算に使う数。第 5 章では、データから学習して決める。

プロドラッグ （第 4 章）
体の中で、効く形に変わるように作った薬。タミフルは、エステルが外れてから効く。

分子間力 （第 1 章）
分子どうしが引き合う力。水素結合は、その中でとくに強い。

分子動力学 （第 1 章）
原子をつぶとばねで表し、力から次の一瞬の位置を求めることをくり返して、分子の動きを計算する方法。

平均二乗誤差 （第 5 章）
予測と実測の差を 2 乗して平均した数。小さいほど、規則がデータに合っている。

平衡 （第 4 章）
反対向きの変化がつり合って、見かけの変化がなくなった状態。第 4 章では、分子が水の層と油の層のあいだを行き来する数がつり合う。止まっているのではない。

ヘマグルチニン （第 3 章）
インフルエンザウイルスの表面のタンパク質。細胞の表面のシアル酸に結合して、ウイルスを細胞につなぎとめる。

変異、薬剤耐性 （第 3 章）
遺伝情報が変わり、タンパク質のアミノ酸が入れかわることが変異。変異で薬が効きにくくなることが薬剤耐性。

モデル、シミュレーション （序章）
分かりたいことのために細部を省いた姿がモデル。モデルをコンピュータで動かして、ふるまいを調べることがシミュレーション。

力場 （第 1 章）
つぶとばねにはたらく力の決め方の一式。

立体構造、X 線結晶構造解析 （第 3 章）
タンパク質の中の原子の 3 次元の並びが立体構造。結晶に X 線を当てて立体構造を決める方法が X 線結晶構造解析。

量子化学 （第 2 章）
電子の雲とエネルギーを、計算で求める分野。

リレンザ、タミフル （第 3 章）
インフルエンザの薬（一般名ザナミビル、オセルタミビル）。どちらもノイラミニダーゼの阻害剤。

ログ P（$\log P$） （第 4 章）
分子を水とオクタノール（油）に分けたときの、油の層の濃度 ÷ 水の層の濃度を P と書く。$\log P$ はその常用対数で、0 より大きいほど油になじみやすく、小さいほど水になじみやすい。

写真の出典



本書の写真は、すべてウィキメディア・コモンズで公開されているものです。写真のページは、commons.wikimedia.org/wiki/File: の後にファイル名を続けたアドレスにあります（空白は _ に置きかえます）。
どの写真も、切り抜きと縮小をしました。炎色反応の写真は、6 枚を切り抜いて横に並べました。
奥付の権利表示は、これらの写真にはおよびません。写真の利用条件は、下の表のライセンスに従います。CC BY-SA の写真を切り抜いたものも、もとと同じライセンスです。

図 序-1	MODIS Land Rapid Response Team, NASA GSFC、Typhoon Hagibis off Japan (MODIS 2019-10-12).jpg、パブリックドメイン
図 1-10 上左	liz west、Cloudy ice cubes.jpg、CC BY 2.0
図 1-10 上中	MarshyBoardwalk、Glass of water refracting light.jpg、CC0
図 1-10 上右	Shardayyy Photography、93 365 - 04 03 11 - Kettle Steaming (5587696584).jpg、CC BY 2.0
図 2-11	S k y r、Flammenfärbung Lithium (Li).jpg、Flammenfärbung Natrium (Na).jpg、Flammenfärbung Kalium (K).jpg、Flammenfärbung Calcium (Ca).jpg、Flammenfärbung Strontium (Sr).jpg、Flammenfärbung Kupfer (Cu) ohne Halogene.jpg、CC BY 4.0
図 3-2	Cynthia Goldsmith（CDC）、Influenza virus particle 8430 lores.jpg、パブリックドメイン
図 3-7 (a)	NASA Marshall Space Flight Center（結晶：Alex McPherson）、Protein crystals grown in space.jpg、パブリックドメイン
図 4-1 (b)	Vantey、Tamiflu 75mg JP 2.jpg、パブリックドメイン
図 4-6 (a)	PRHaney、Separatory funnel with oil and colored water.jpg、CC BY-SA 3.0
図 6-10	Barsaka2、RIKEN R-CCS Fugaku.jpg、CC0
図 終-1	Dhudson4god、HarrisonPowerStationViewFromEast.jpg、パブリックドメイン
図 終-3	Marcin Wiszniewski（CNBCh）、Zdjęcie IRMOF-3.jpg、CC BY-SA 4.0

ライセンスの全文は、次のアドレスにあります。

CC BY 2.0	creativecommons.org/licenses/by/2.0/
CC BY 4.0	creativecommons.org/licenses/by/4.0/
CC BY-SA 3.0	creativecommons.org/licenses/by-sa/3.0/
CC BY-SA 4.0	creativecommons.org/licenses/by-sa/4.0/
CC0	creativecommons.org/publicdomain/zero/1.0/

コンピュータのなかの実験室

発行日　2026 年 11 月 6 日　第 0.1.5 版発行
著　者　山本 典史

© 2026　山本 典史

本書の著作権は著者が保有します。本書は、学習と教育のために、無償で閲覧・ダウンロード・印刷できます。著者の許可なく、本書の全部または一部を再配布・転載・改変・販売することはできません（著作権法が認める引用を除きます）。写真の利用条件は、巻末の「写真の出典」のとおりです。授業や勉強会で紹介するときは、ファイルを配らずに、本のページを案内してください。本書は改訂を続けています。引用するときは版番号を添えてください。最新版・改訂履歴・誤りの報告の窓口は、本のページにあります。

著者紹介

山本 典史（やまもと のりふみ）　千葉工業大学 工学部 応用化学科 教授。博士（理学）。1999 年 九州大学 理学部 化学科 卒業、2004 年 九州大学 大学院 総合理工学研究科 修了。日本学術振興会 特別研究員（東京大学）、岐阜大学 助教、名古屋大学 特任助教、千葉工業大学 助教・准教授を経て、2022 年より現職。専門は計算化学。量子化学・分子動力学・機械学習などの手法で、コンピュータを使って分子を解析している。化学の学びを身近にすることにも取り組んでいる。

テオとテラのイラスト

揚げ鶏々（あげどりどり）　科学系クリエイター。元素を擬人化した創作『元素楽章』を手がける。

gensogakusho.com

本のページ

章ごとの動く図にさわれます。

yamlab.jp/books/icc

Version 0.1.5