

生命科学・創薬・医療領域の ユーザー視点からの 次世代計算基盤への期待

京都大学大学院医学研究科
奥野恭史

「京」から「富岳」による分子動力学計算 (MD) の進化

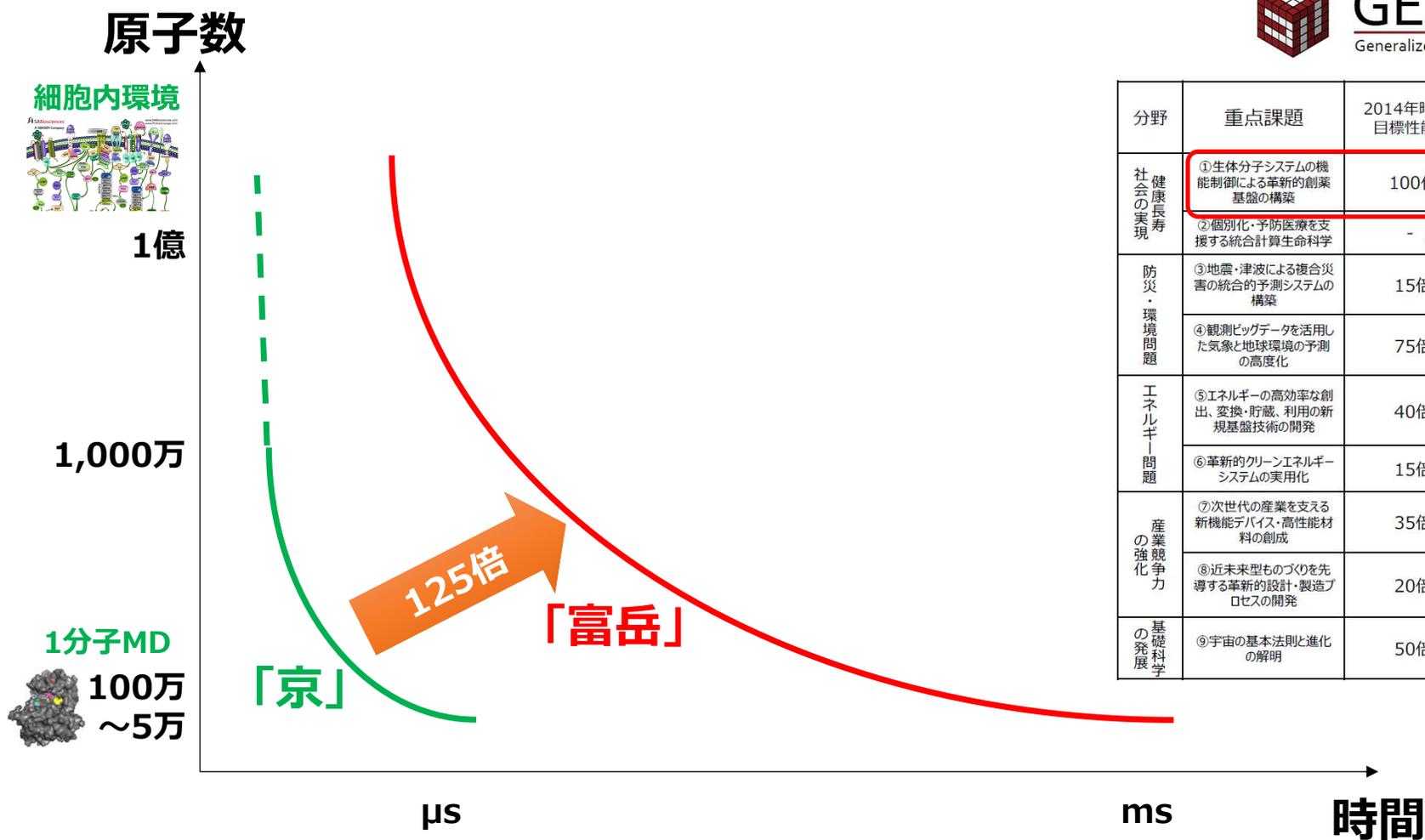
「富岳」とのコデザインにより分子動力学計算アプリGENESISで「京」の125倍（演算性能）を達成



GENESIS

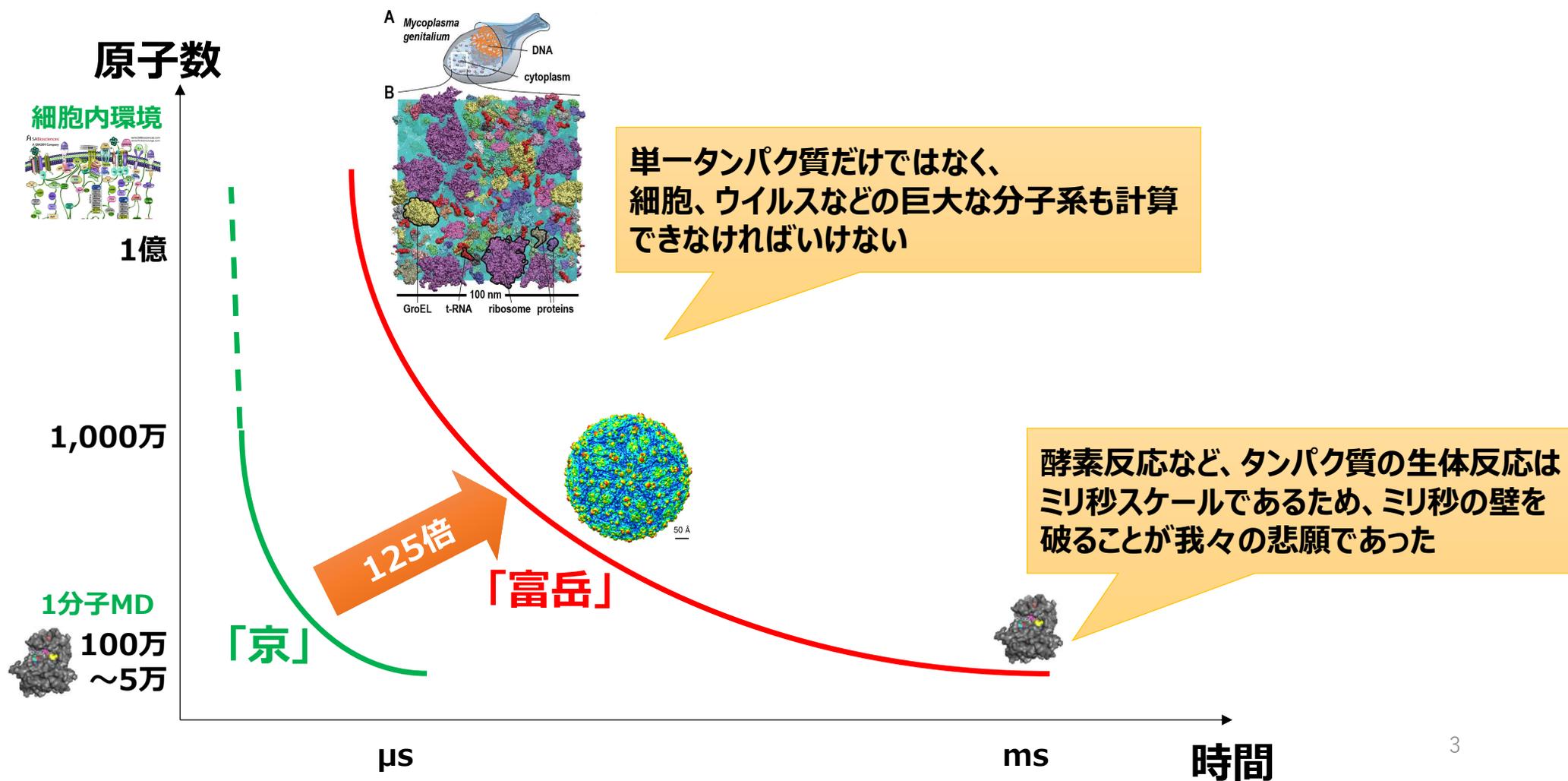
Generalized-ensemble simulation system

分野	重点課題	2014年時点の 目標性能※5	現時点の 性能見込み※6	アプリケーション
社会の実現 健康長寿	①生体分子システムの機能制御による革新的創薬基盤の構築	100倍	125倍以上	GENESIS
	②個別化・予防医療を支援する統合計算生命科学	- ※7	8倍以上	Genomon
防災・環境問題	③地震・津波による複合災害の統合的予測システムの構築	15倍	45倍以上	GAMERA
	④観測ビッグデータを活用した気象と地球環境の予測の高度化	75倍	120倍以上	NICAM+ LETKF
エネルギー問題	⑤エネルギーの高効率な創出、変換・貯蔵、利用の新規基盤技術の開発	40倍	40倍以上	NTChem
	⑥革新的クリーンエネルギーシステムの実用化	15倍	35倍以上	Adventure
産業競争力の強化	⑦次世代の産業を支える新機能デバイス・高性能材料の創成	35倍	30倍以上	RSDFT
	⑧近未来型ものづくりを先導する革新的設計・製造プロセスの開発	20倍	25倍以上	FFB
基礎科学の発展	⑨宇宙の基本法則と進化の解明	50倍	25倍以上	LQCD



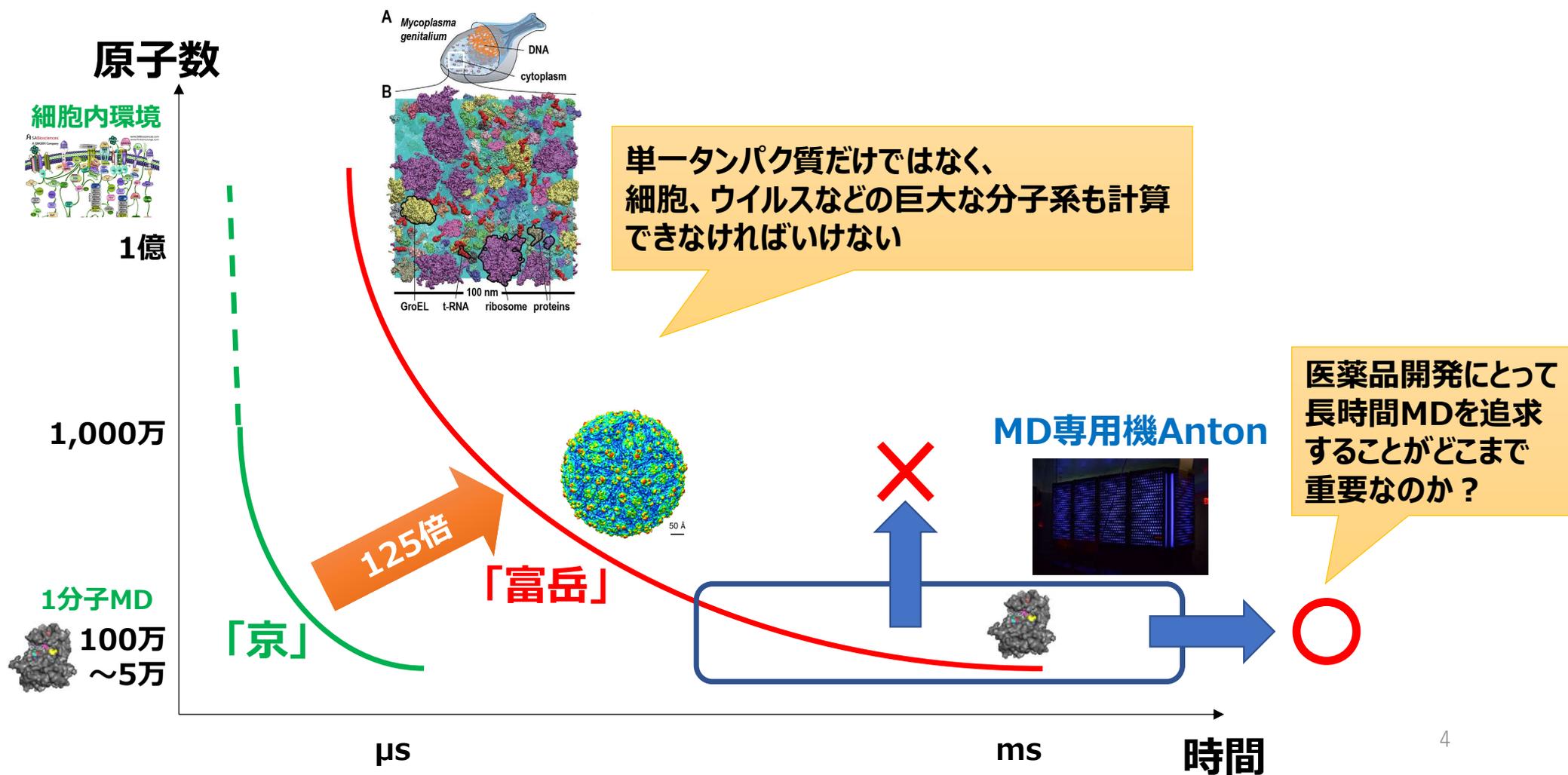
「京」から「富岳」による分子動力学計算 (MD) の進化

「富岳」とのコデザインにより分子動力学計算アプリGENESISで「京」の125倍（演算性能）を達成



「京」から「富岳」による分子動力学計算 (MD) の進化

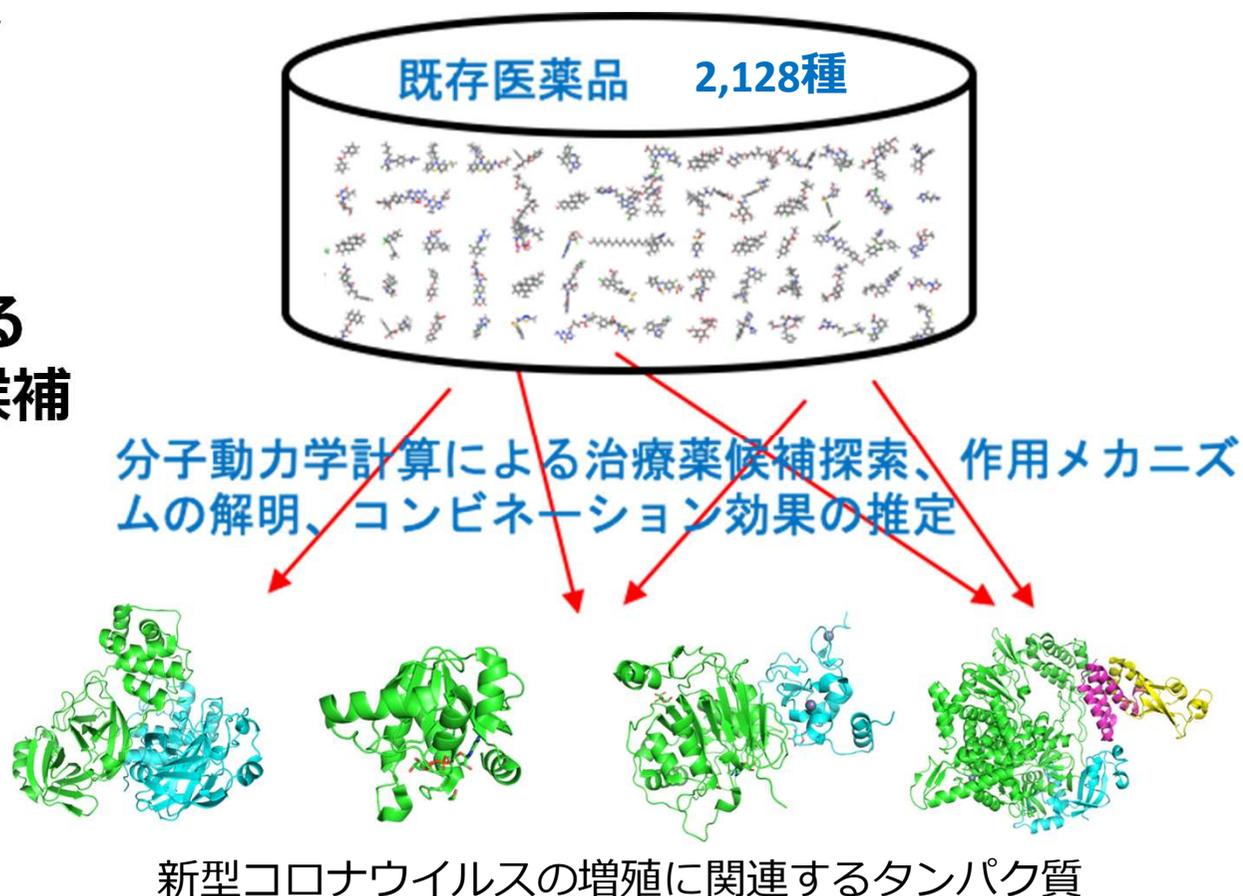
「富岳」とのコデザインにより分子動力学計算アプリGENESISで「京」の125倍（演算性能）を達成



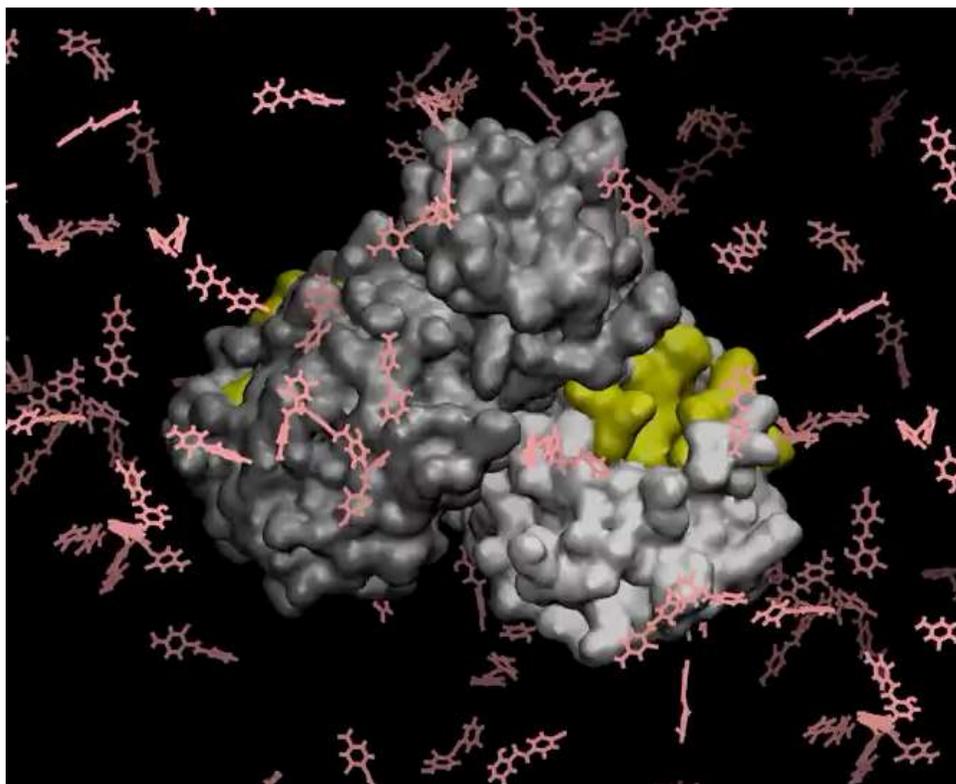
「富岳」による新型コロナウイルスの治療薬探索の例

「富岳」を用いた分子シミュレーション
(分子動力学計算)により、
現場利用されている**2,128種**の
既存医薬品の中から、
新型コロナウイルスの増殖に関連する
標的タンパク質に作用する治療薬候補
を探索する。

医薬品開発にとっては、膨大なケース数（化合物数）が計算できなければ、お話にならない。これは医薬品開発に限らず、発見を最重要とする生命科学全体にいえることである。



「富岳」による新型コロナウイルスの治療薬探索の例



「富岳」が有力候補として同定した
薬剤niclosamideが米国で治験中

Latest News
BioNTech, Fosun Pharma initiate Covid-19 vaccine candidate trial in China
Roche's Kofluzza gets FDA approval for influenza prevention
Purdue Pharma pleads guilty to three criminal charges

News

ANA Therapeutics begins trial of oral niclosamide for Covid-19 treatment

27 October 2020 (Last Updated October 27th, 2020 08:03)

Biotech firm ANA Therapeutics has begun Phase II/III trial to analyse the safety and efficacy of its oral niclosamide (ANA001) formulation for treating patients with moderate Covid-19.



1 μ s \times 2000化合物 (計2ms) を1日強でMD計算可能
数1000規模化合物のMDは世界初

しかし実際は、信頼できる結果を得るために、1ms \times 2000化合物がやりたかった

⇒ ポスト「富岳」で100倍性能になれば、10日程度で1ms \times 2000化合物が可能

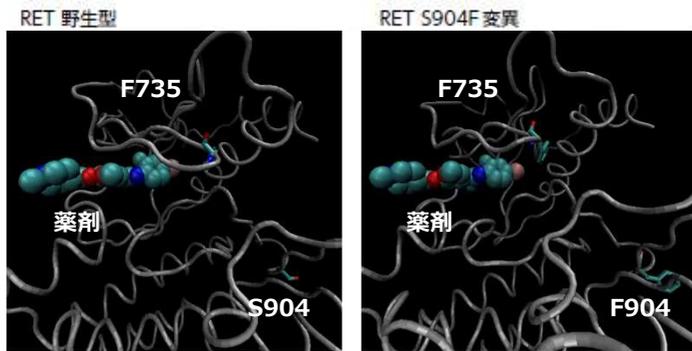
スパコン「京」によるゲノム変異に対する分子動力学計算の実績

「京」でゲノム変異による薬剤反応性（薬剤と変異タンパク質の結合親和性）の高精度予測に成功

RET 融合遺伝子上に生じるアロステリック効果を持つ二次変異
 LC-SCRUM-Japanの遺伝子スクリーニングに基づいて、分子標的治療薬に対するがんの新しい薬剤耐性メカニズムを発見

Nature Communications., 9(625), 1-9 (2018)

遺伝子変異により生じる薬剤結合部位のコンフォメーション変化



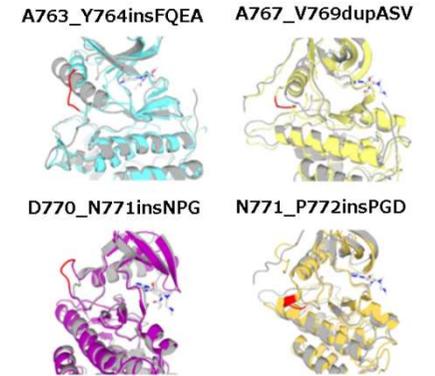
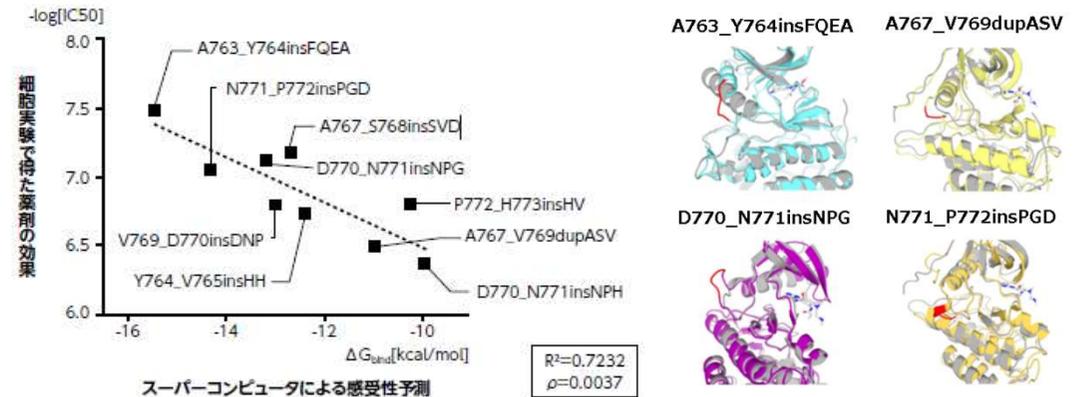
薬剤との安定な複合体構造

薬剤との安定性が低下

長時間 (1 μ sec) 分子動力学シミュレーションのスナップショット

LC-SCRUM-Japanで構築した日本最大のがん臨床ゲノムデータを
 活用しスーパーコンピュータで治療薬の効き目を予測
 がんゲノム医療における新たなツールの開発

PNAS, 116(20):10025-10030 (2019)

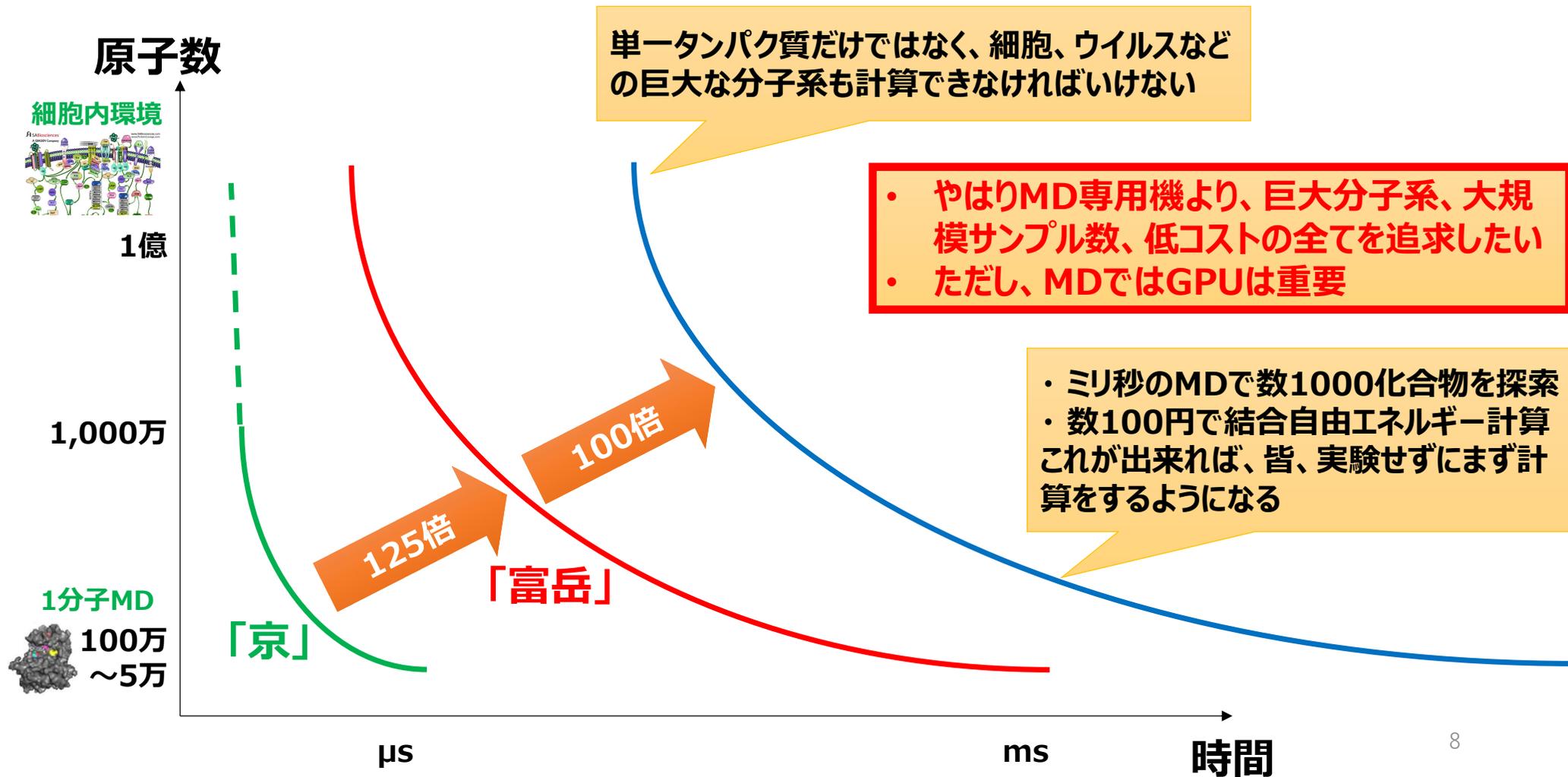


タンパク質と薬剤の結合自由エネルギーを計算を数10種類について行うのに、「京」で1か月程度必要だったが「富岳」では2日程度で計算可能 ⇒ がん、パンデミックいずれも1か月も待てない（1か月が命取り）
 ⇒ ポスト「富岳」で100倍性能になるなら、将来の変異を試行した強力な医薬品、ワクチン開発の可能かも

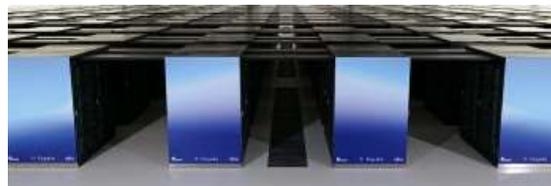
産業利用、臨床利用にとって、もっと重要なことは、利用料金！
 「京」の有償利用では、タンパク質－薬剤の1ペアあたり30万円であったが、「富岳」では同じ換算で1.5万円
 ⇒ ポスト「富岳」でさらに20分の1の800円になれば、もう誰も実験をしなくなるくらいすごいことになる

分子動力学計算 (MD) に関するポスト「富岳」への期待

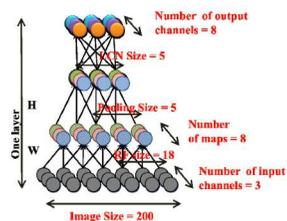
トータル演算性能で、分子動力学計算が「富岳」の100倍以上は欲しい



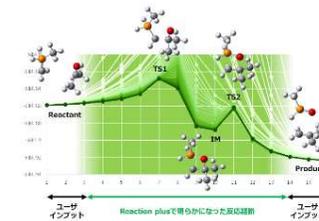
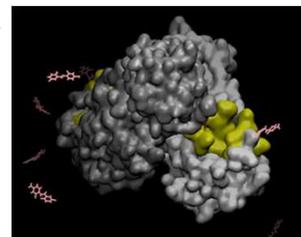
生命科学・創薬・医療においてシミュレーション×AI×実験の統合フレームを構築することが世界的優位性を生む一つの戦略



データサイエンス・AI



シミュレーション



必要最小限の実データによる誤差補正

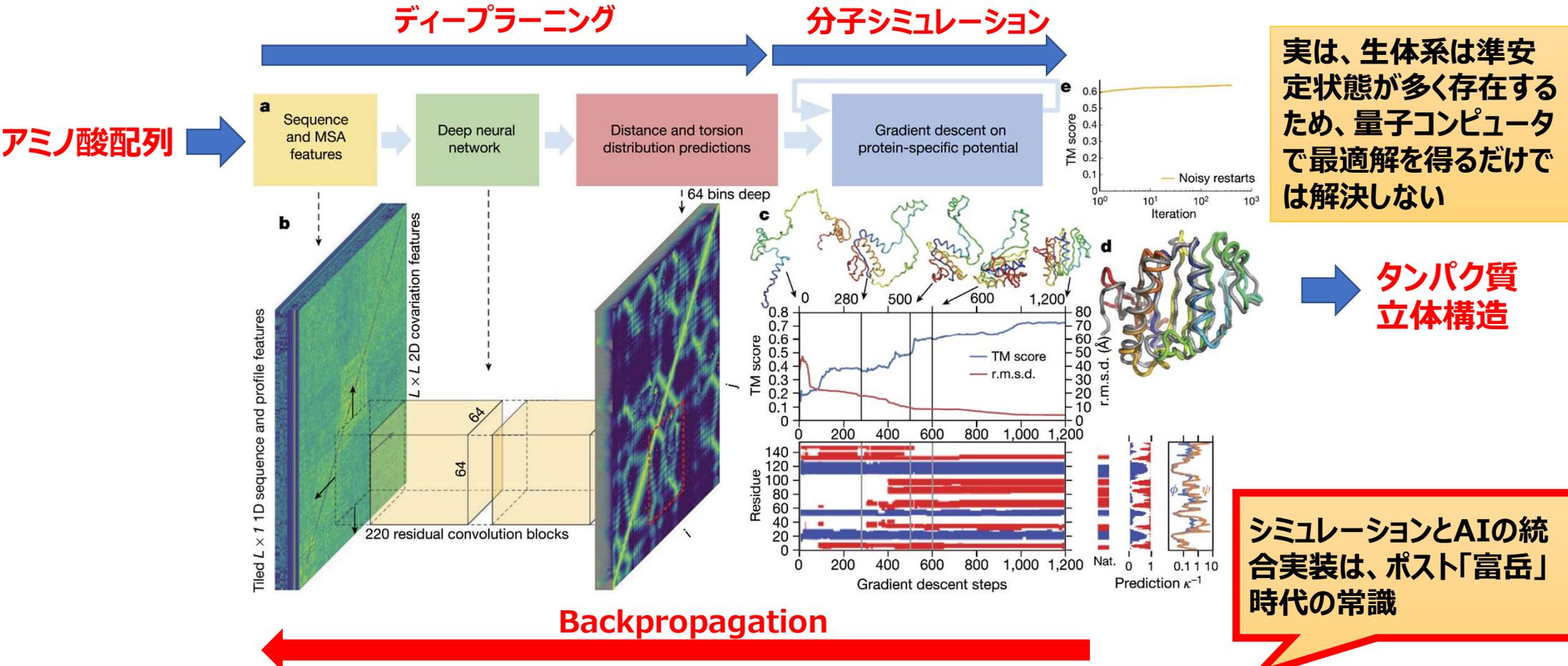
ビッグデータ (実験、臨床、文献)



生命科学・創薬・医療にとっては、高性能MD単体よりもこの統合フレームが容易に実装・実行できる次世代機を開発することが最も重要と考える

DeepMindによるタンパク質の構造予測AI “Alpha-Fold”の例

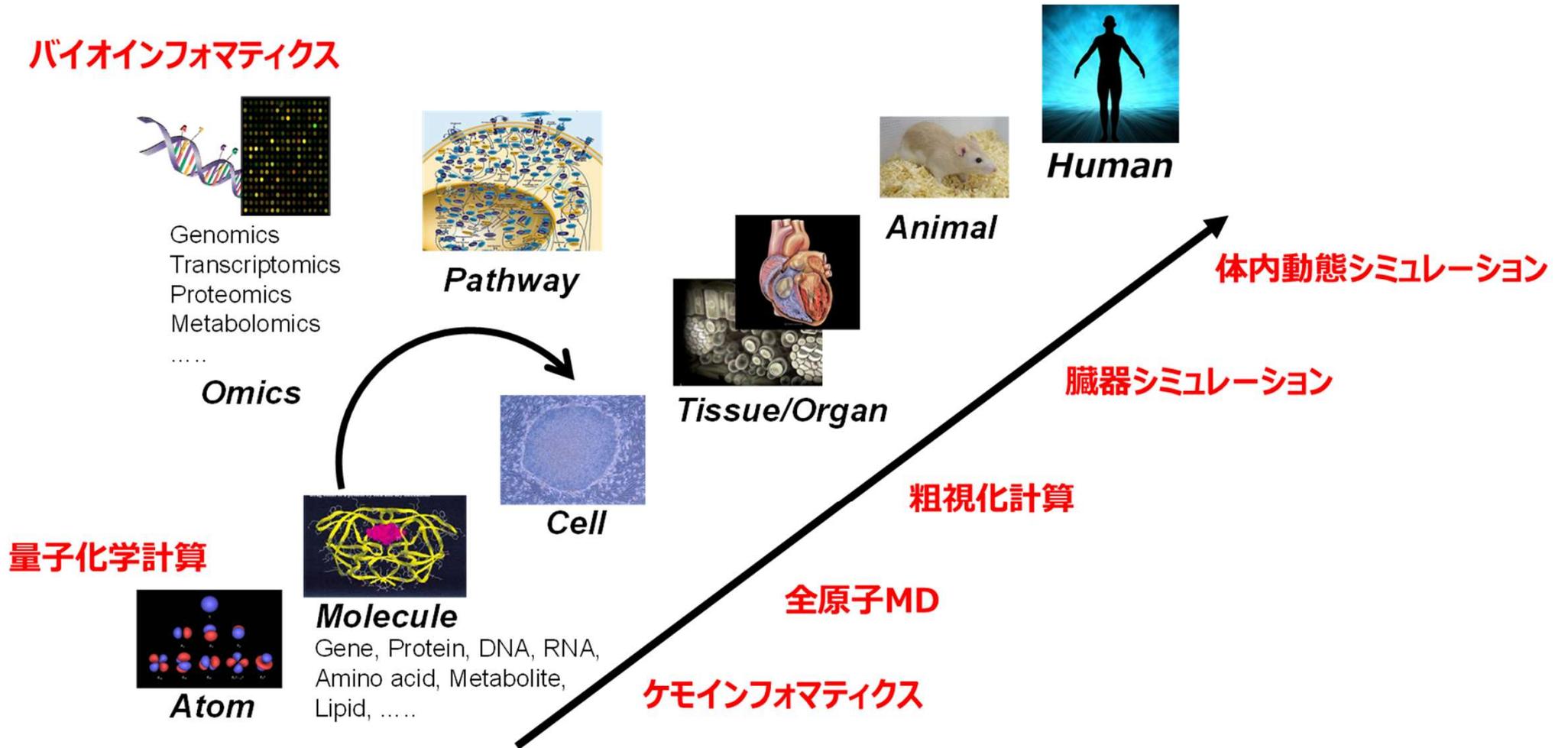
超高速MDがあればタンパク質構造予測が可能になるとわれてきたが、AIが強力な威力を示した



Alpha-Foldは、ディープラーニングと分子シミュレーションは別々に計算しているが、AIとシミュレーションを統合実装して、分子シミュレーション後にend to endで学習することまで考えれる

生命科学、創薬、医療にとっては高性能MDだけでは不十分

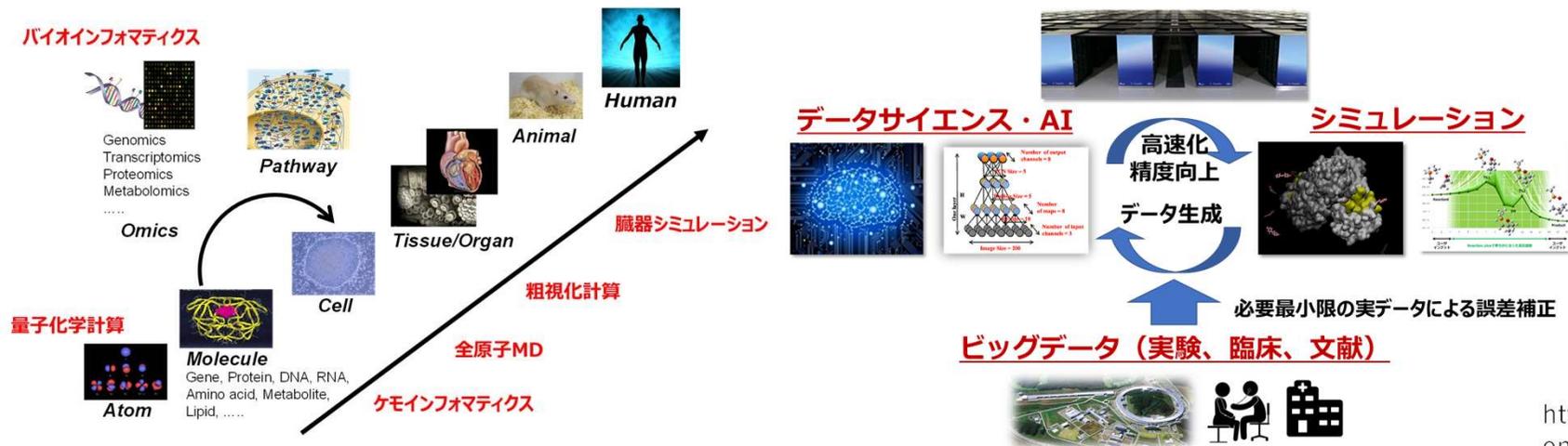
生命科学、創薬、医療にとって高性能MDは重要であるが、必要条件に過ぎない。
次世代は複数のアプリを使いこなし、組み合わせることにより高度な問題を解く必要がある。



ユーザーファーストならAll in Oneスパコン

これからの生命科学・創薬・医療を対象にした計算は分子動力学計算だけでは不十分であり、AI、量子化学計算、粗視化、臓器、体内動態シミュレーション、バイオインフォマティクス、ケモインフォマティクスなど多岐の計算アプリを組合せる必要がある。

- そのため、複数の既存アプリを個別に使う場合でも、全く使い勝手の違う専用スパコンを単にネットワークで接続した異機種環境は、ユーザー側からすると非常に使い勝手が悪い。ユーザーにとっては同一の使用感で使えるユーザーファーストなマシンであってほしい。
- さらに、それぞれのアプリを独立開発・独立利用するだけでなく、それぞれのアプリが無駄な苦勞なく統合化できる統合化環境を、マシン側に具備できるかが非常に重要に思う。
- また統合化環境を使いこなせる次世代のIT人材育成も非常に重要。
- したがって、ポスト「富岳」開発では、個別のコーデザインだけではなく、アプリとマシンの統合を平行で行っていく必要があると考える。次世代のキーワードは「統合化」



<https://www.clicknf5.org/gadgets/smartphones-technology-evolution/10338>

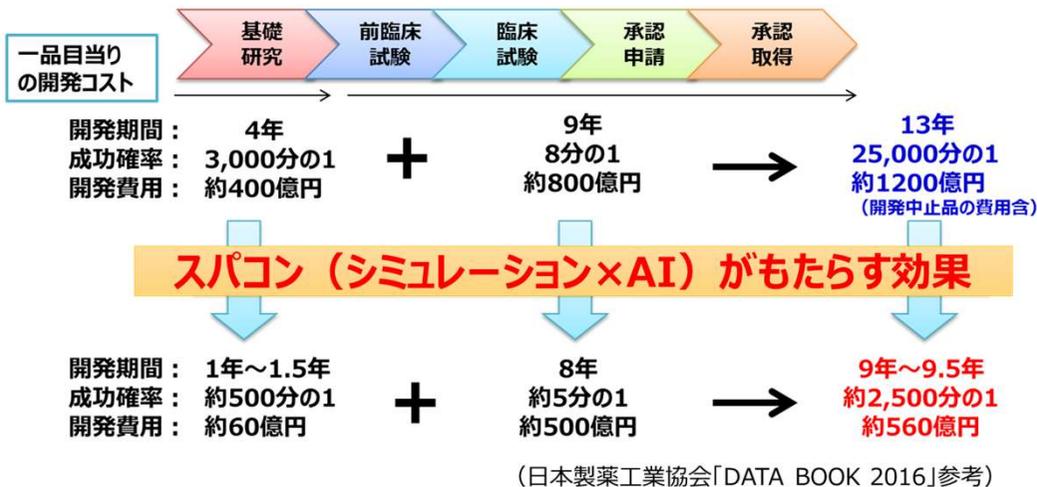
生命科学・創薬・医療の 新時代を切り拓くポスト「富岳」を切望する

スパコンが命を守る時代に

国民の健康・安全保持と医療費削減

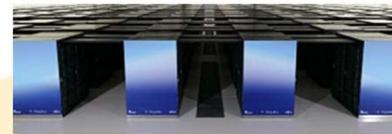
開発期間：4年短縮

開発費：業界全体で1.2兆円削減（1品目あたり600億円削減）

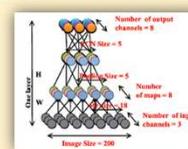


スパコンがサイエンスそのものを作る時代に

第5の科学（統合科学）

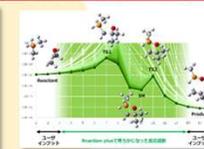
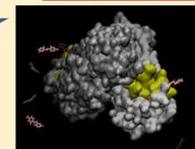


第4の科学（データ科学）



第2の科学（理論科学）

第3の科学（計算科学）



必要最小限の実データによる誤差補正

第1の科学（実験科学）



まとめ：生命科学・創薬・医療のユーザー視点からのポスト「富岳」への期待

- 分子動力学計算においては、「富岳」時代よりもトータル演算性能100倍以上の向上が欲しい。注意すべきは、高速性能だけでなく、大規模対応も必要
- 多くの化合物や実験条件の検討を行いたいため、超高速になっても1回の計算が高額であることは好ましくない。高速であるに越したことはないが、計算費用が安価であれば実験の代わりにまず計算を誰もがするようになる。例えば、「富岳」より10倍しか高速化できなくても、1000倍安く計算できるなら、ユーザーは一気に増える。
- 現在は、創薬＝分子動力学と思われがちだが、生命科学・創薬・医療を対象にした計算は分子動力学計算だけでは不十分であり、AI、量子化学計算、粗視化、臓器、体内動態シミュレーション、バイオインフォマティクス、ケモインフォマティクスなど多岐の計算アプリを組合せる必要がある。
- 生命科学・創薬・医療においてシミュレーション×AI×実験の統合フレームを構築することが世界的優位性を生む一つの戦略と考える。
- 特に、ポスト「富岳」時代では、これらの多種多様なアプリを独立に扱うのではなく、統合的に実装・利用できる統合化環境を備えたマシンが必要と考える。
- さらに、統合化環境を使いこなす世界をリードできる次世代のIT人材育成が最重要。その一環として、気象、材料、流体など、異分野の計算・情報専門家との情報交流、人材交流が、ポスト「富岳」に向けてより一層重要になる。
- ハードとソフトの研究開発が注目されがちだが、ユーザーにとっては、実際の運用やサポート体制が充実していることも非常に重要である。ユーザーファーストの運用とサポートが実践可能かどうかも含め、スパコン仕様は策定されるべきである。
- 最後に情報系人材がいつまで経っても足りないのと同様に、スパコンを使いこなす計算人材も圧倒的に少ない。そもそも「富岳」からポスト「富岳」に向けて、計算・情報人材の育成とポストの増員をもっと強化しなければならない。