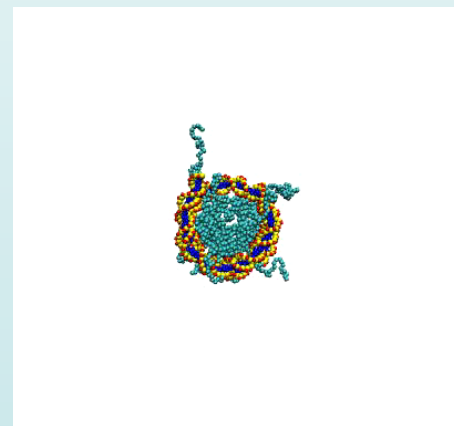
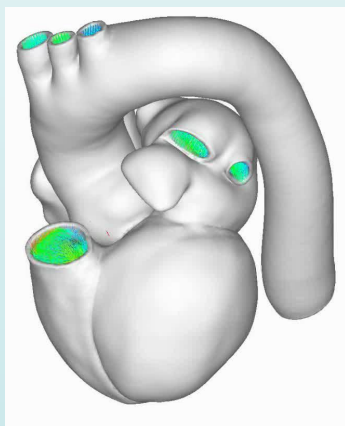


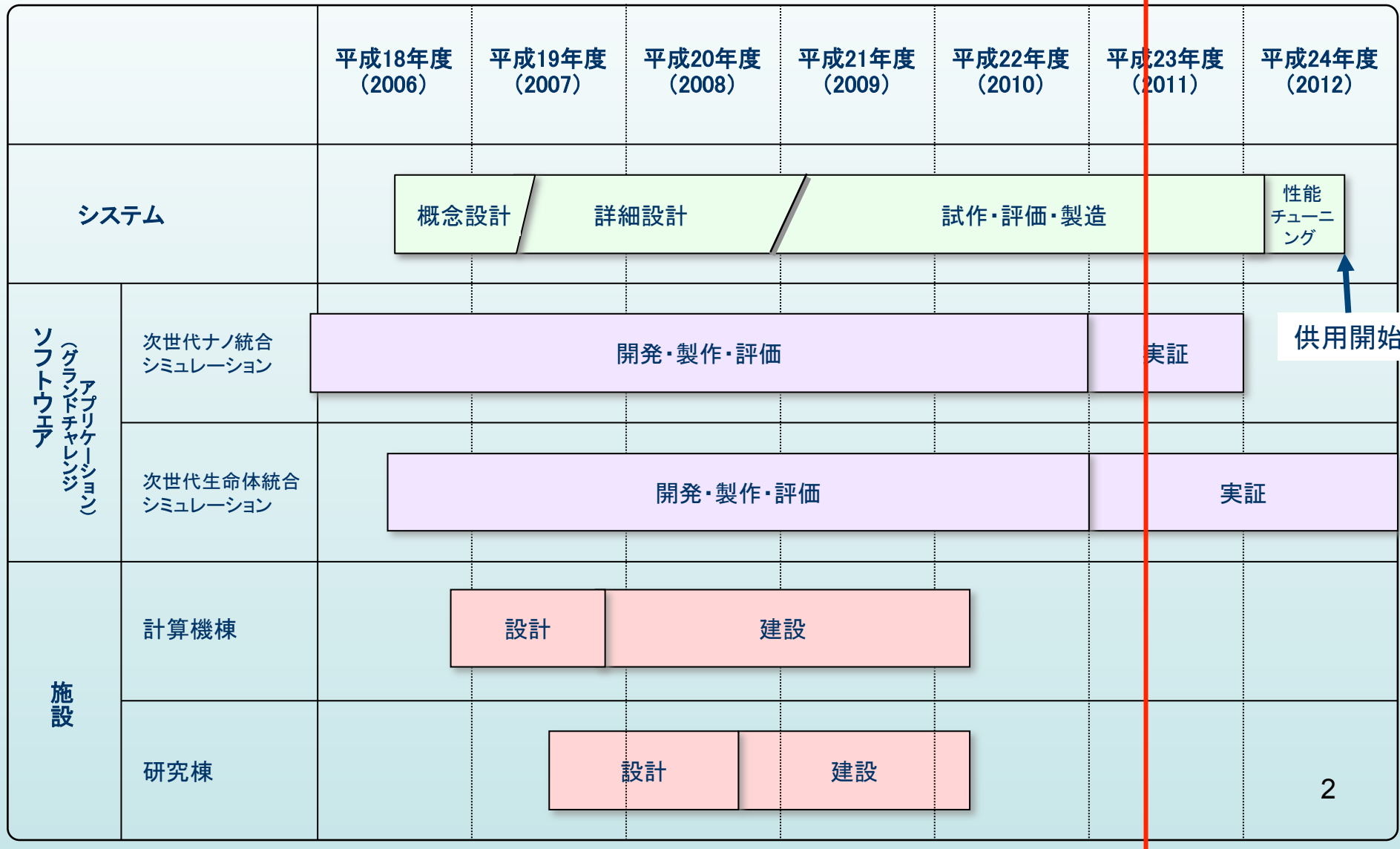
次世代スーパーコンピュータは生命と 医療にどこまで迫れるか

理化学研究所
次世代計算科学プログラム
姫野龍太郎



次世代スーパーコンピュータの研究開発プロジェクトの開発計画

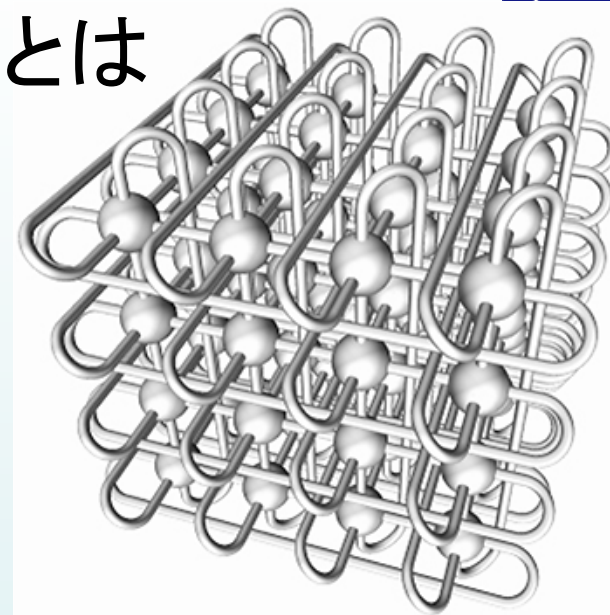
現在



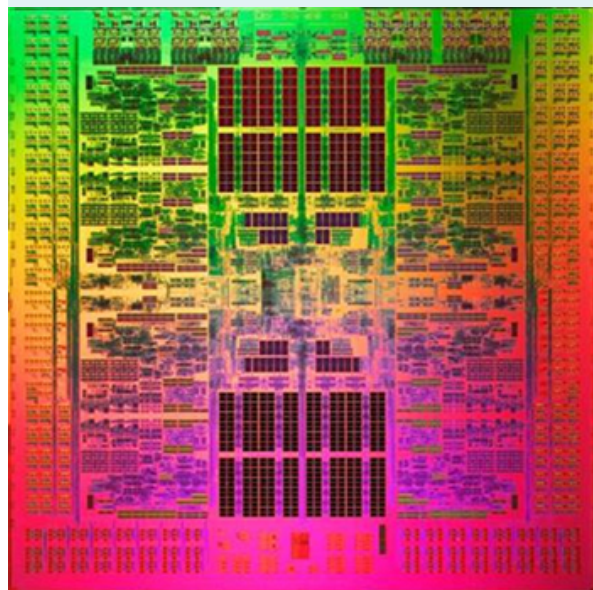
次世代スーパーコンピュータとは



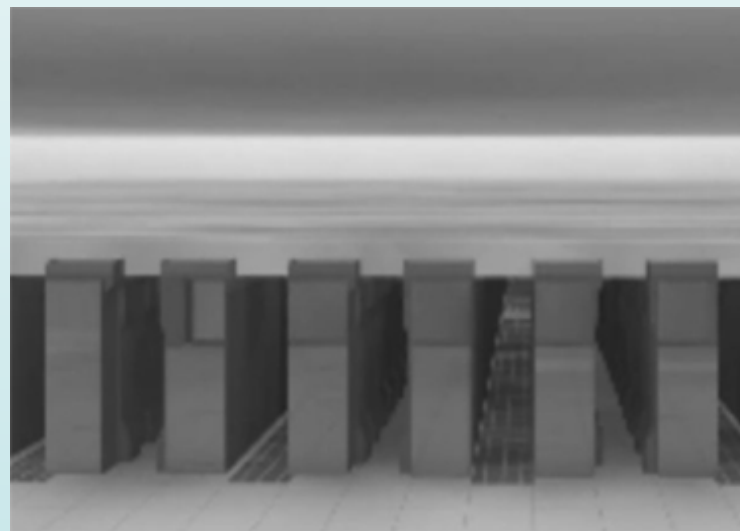
- 新しく開発したプロセッサCPU128GFLOPS
昨年試作、その時点でプロセッサ単体として世界最速
- 8万個以上集めて10PetaFLOPSを目指す
一秒で一京回(一兆の一万倍)の計算ができる
- 現在、完成に向けて次々に導入、設置中
- 現時点でも世界最速、LINPACKで8PetaFLOPS



通信用ネットワーク(イラスト)



「SPARC64™ VIIIfx」
富士通提供



3
計算機室に設置される次世代スーパーコンピュータ(イメージ図)

- 8.16PetaFLOPS, 2位のスパコンの3倍を超える性能
 - 実効効率の良さ
- 消費電力の低さでも評価



Fujitsu Lunch Party at ISC2011





コンピュータの速度の進歩の歴史



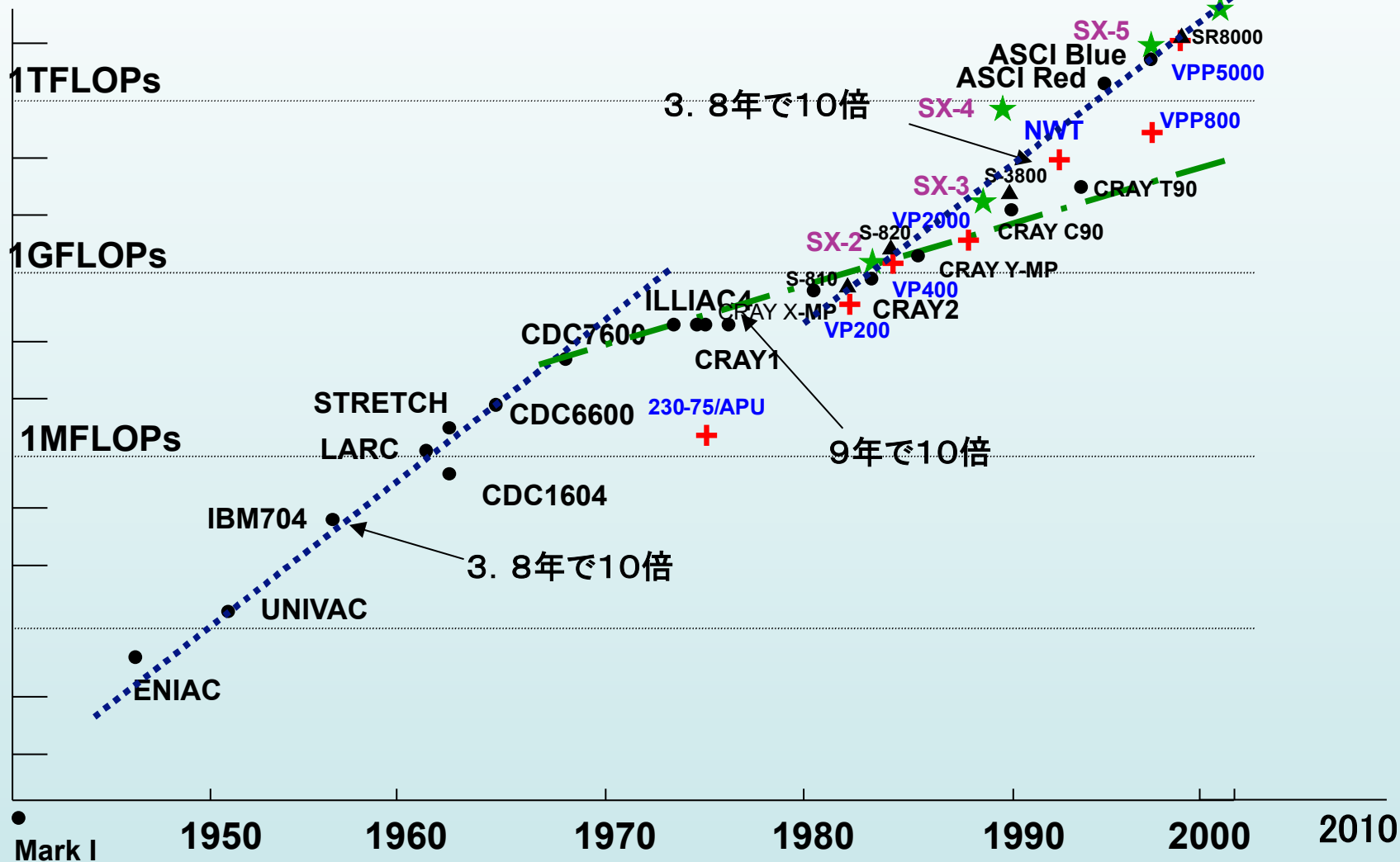
1000兆回/秒
10¹⁵M
1PFLOPs

1兆回/秒
10¹²
1TFLOPs

10億回/秒
10⁹
1GFLOPs

百万回/秒
10⁶
1MFLOPs

1千回/秒
10³

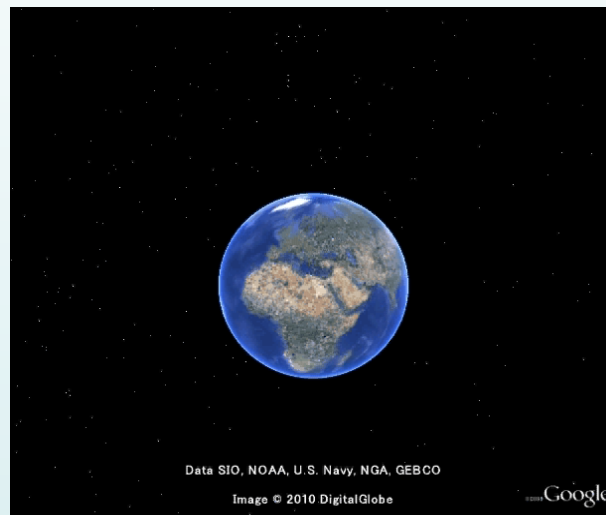


Source:L. Smarr Top500 site

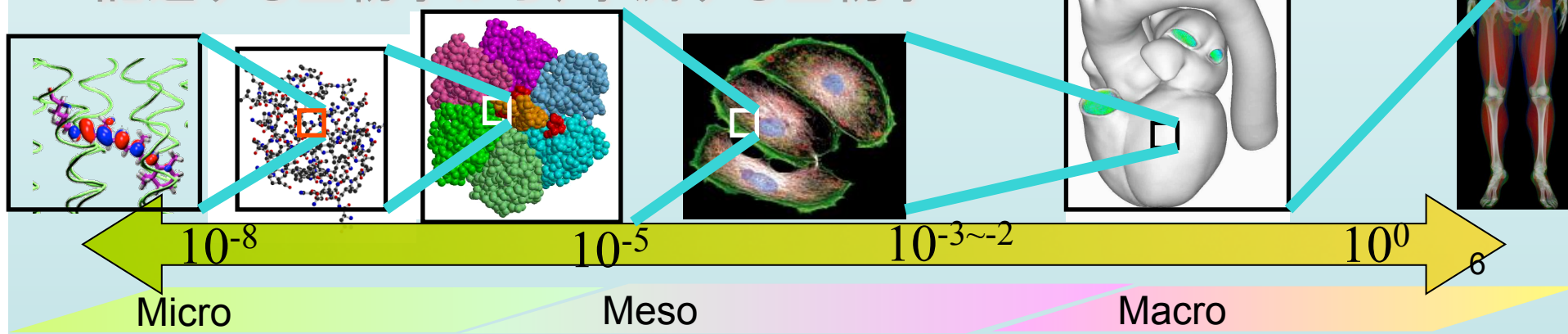
動機

生命現象は最も複雑で難しい問題

複雑で美しい振舞いを示す**超**多体系多階層問題

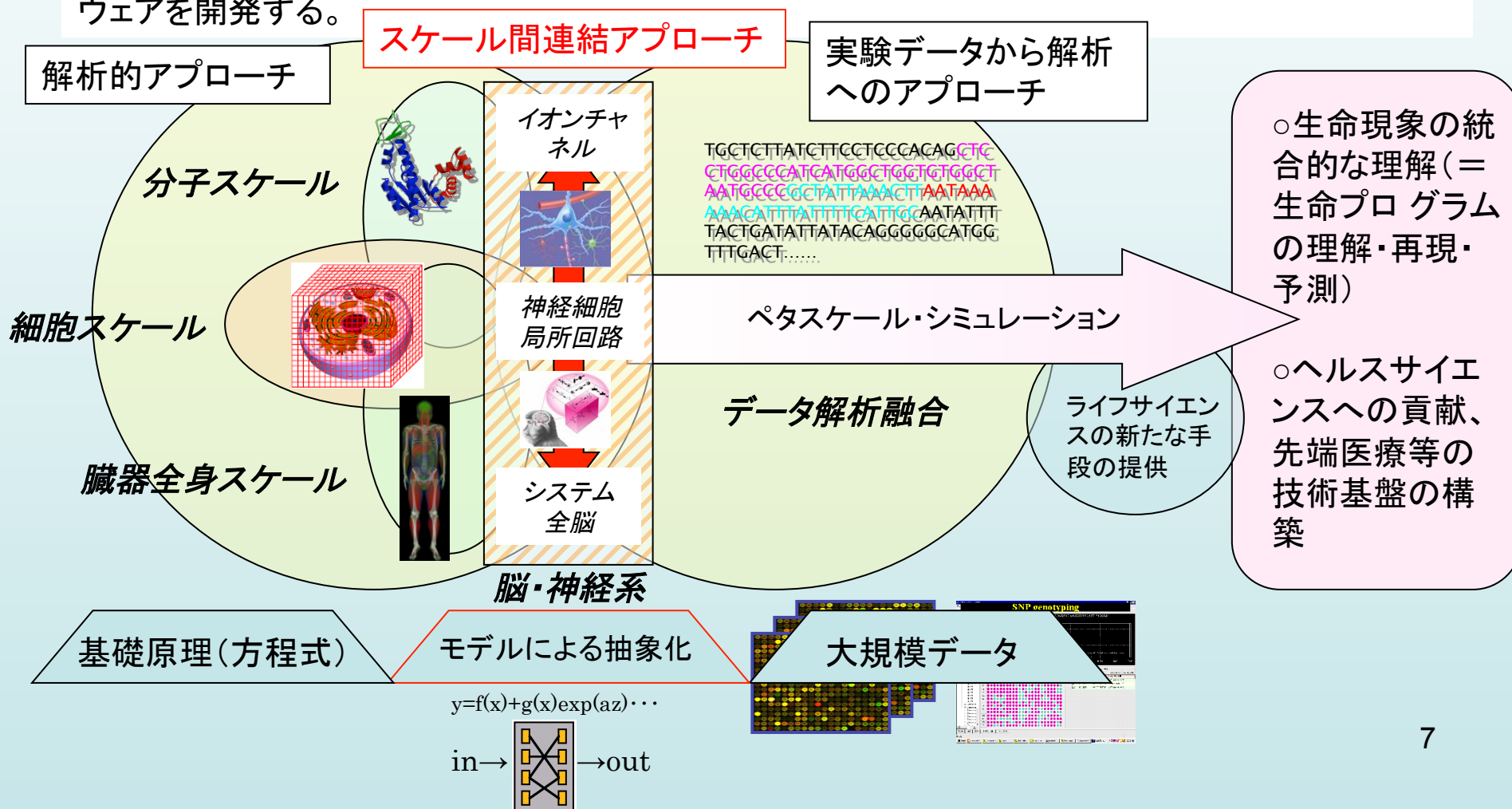


スーパーコンピュータを使って、この複雑な生命現象を解析
記述する生物学から、予測する生物学へ



研究開発の概要と達成目標

基礎方程式に基づく解析的アプローチと、大量の実験データから未知の法則に迫る実験データから解析へのアプローチ、さらには多階層を連結するアプローチにより、異なるスケールの研究と実験データを統合的かつ有機的に結びつけ、ペタスケールという桁違いの性能を持つスーパーコンピュータの性能をフルに発揮し、生体で起こる種々の現象を理解し医療に貢献するためのソフトウェアを開発する。

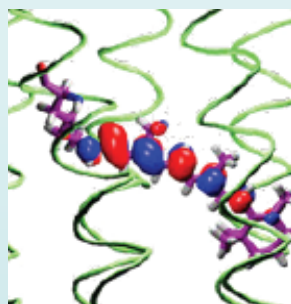


開発目標:量子化学計算(QM)・分子動力学計算(MM)・粗視化モデル計算(CG)を総合化する技術を開発、タンパク質や細胞の機能発現過程のシミュレーションにつなげ、細胞スケールとの有機的な連携を図る。

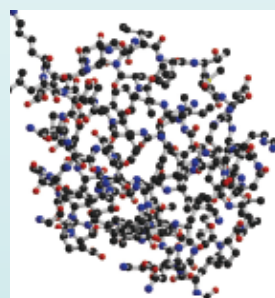
ターゲット:多剤排出トランスポータ、脂肪酸代謝酵素反応

開発計画:QM、MM、CGの各スケールでプログラムの開発を進めるとともに、それらの手法を結合したQM/MM、MM/CG法によってマルチスケールシミュレーションを実現するためのプログラムを開発する。

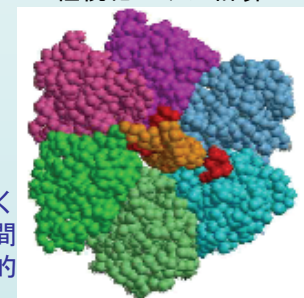
量子化学計算



分子動力学計算



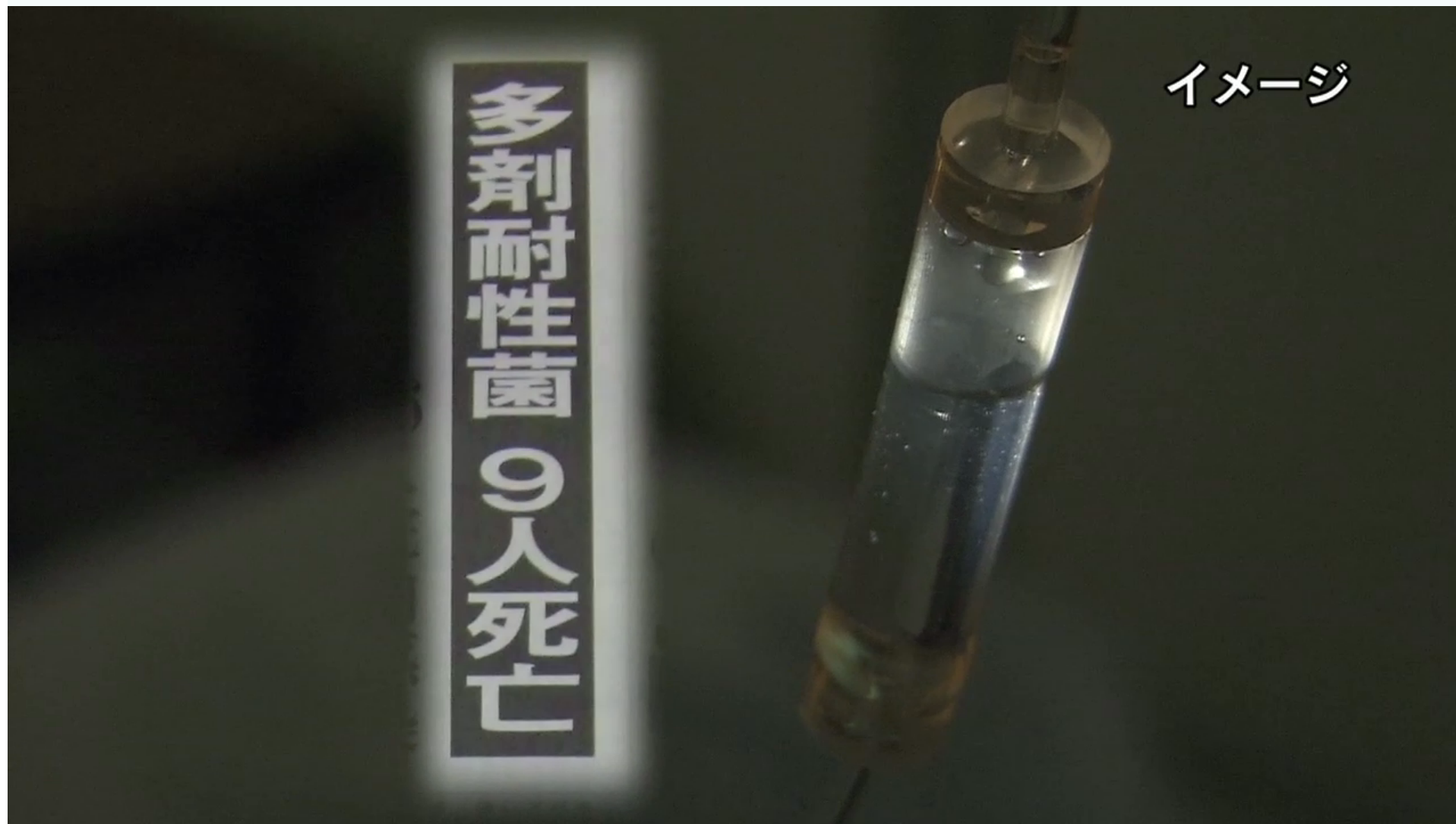
粗視化モデル計算



←
より小さく
より短時間
より厳密

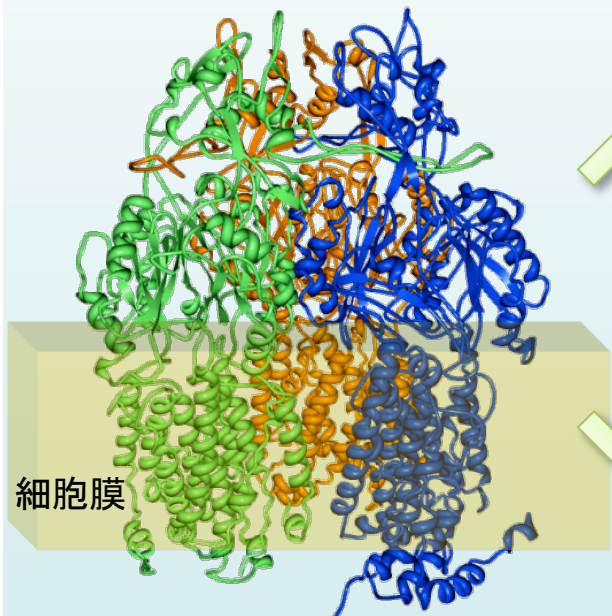
→
より大きく
より長時間
より近似的

進捗状況:QM/MM/CGの3階層それぞれで計算方法をプログラム化し、テストを続けながら開発、従来のスパコンでは8000並列程度まで性能が出るまでになっている。4月から京を使ってソフトのテスト中



分子スケールの
ISLAM スーパーコンピューティング

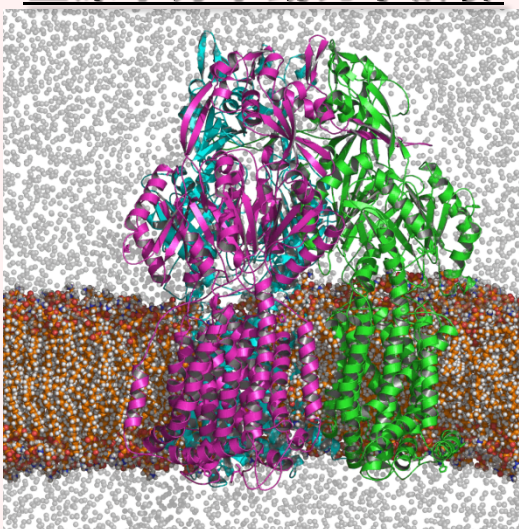
多剤排出トランスポーター



細菌の細胞膜にある多剤排出トランスポーターが、その形を変えながら薬剤を細胞外に排出することで薬剤耐性化を引き起こしている。そのメカニズムを明らかにして耐性化を防ぐ

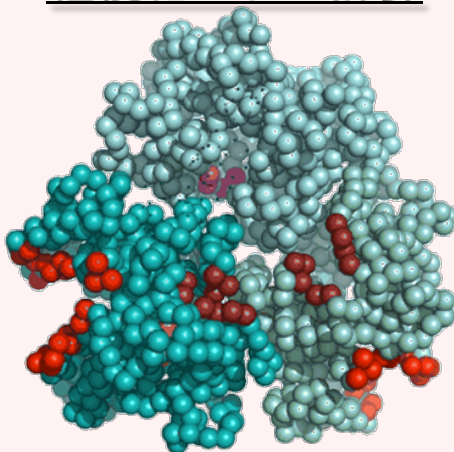
10 TFLOPS (H22-23)

全原子分子動力学計算



数百ナノ秒の計算で、形を変える初期過程の仕組みを明らかにした

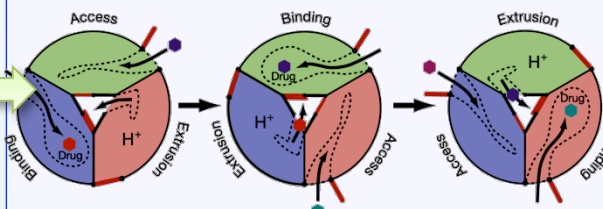
粗視化モデル計算



構造変化をさせることで薬剤排出をシミュレートすることに成功

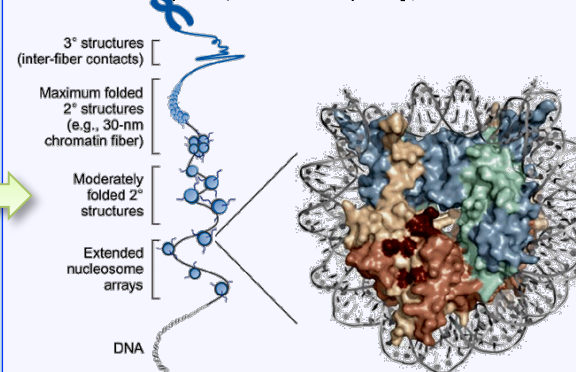
10 PFLOPS (H24)

多剤排出トランスポーターの全過程経路シミュレーション



数十マイクロ秒の計算によって、粗視化モデルが実現した構造変化の全過程のシミュレーションを全原子モデルで実現する

ヌクレオソームの粗視化モデルシミュレーション



エピジェネティクスが染色体に与える影響をシミュレートすることで、染色体構造の転写制御での役割を明らかにする

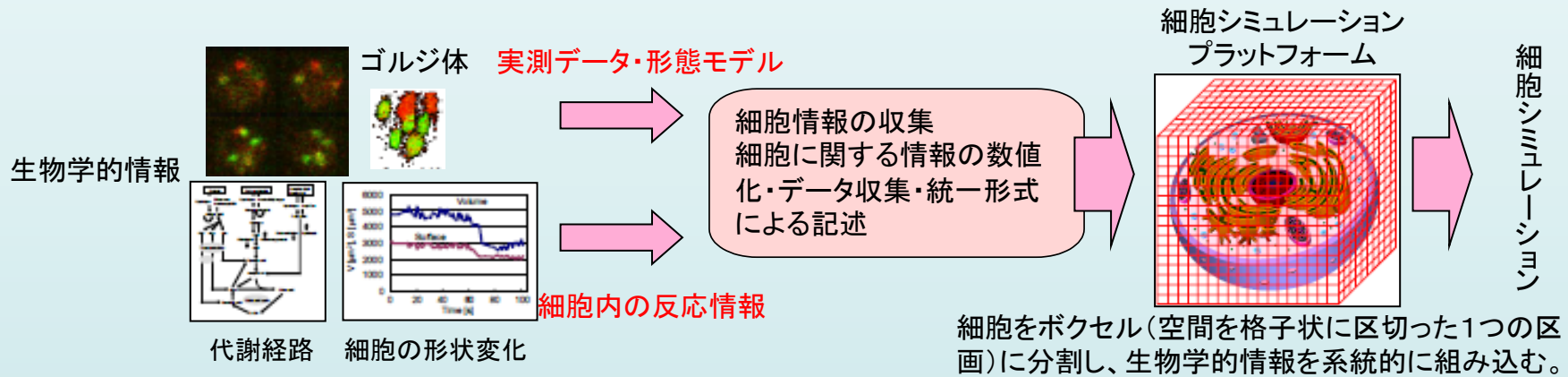


細胞スケール研究開発



開発目標: ペタフロップス級の計算力を活かして細胞を100万のボクセルに区画、そこに包括的な実証データを組み込んだ肝細胞・肝小葉のシミュレーションの実現を目指す。薬剤等の作用や副作用の予測を行うシミュレーションの開発につなげる。

開発計画: 生物の空間的情報を入れ込むことが出来るプラットフォームの開発と平行して、プラットフォームに実装するためのデータ収集とモデルの構築を進める。また、プラットフォームと各種データ・モデルとの有機的連携を図る。



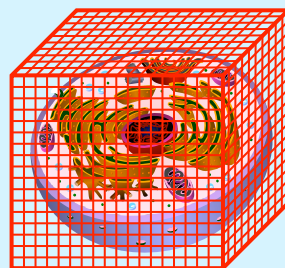
進捗状況: 細胞を約3万個のボクセルに分割、複数のオルガネラを表現し、移流拡散方程式による物質移動をシミュレーションできる細胞シミュレーションプラットフォームを開発、E-CELL3をベースにした代謝反応がオルガネラ間も含めてシミュレーションできるようになった。8192並列までの性能テストを終了。京でテスト中¹¹

細胞シミュレーションプラットフォーム RIKEN Cell Simulator (RICS)

～細胞内の場を考慮した生化学反応シミュレーター～

膜透過 + 拡散 + 生化学反応 (代謝)

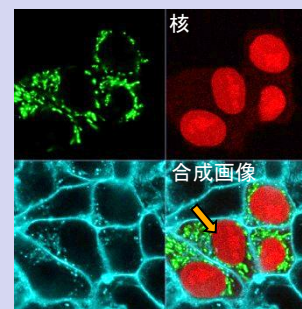
- ・連続体として細胞内を表現
- ・空間をボクセル空間で表現
- ・細胞内小器官などの複雑な空間構造にも対応
- ・複数細胞を同時に計算可能
- ・細胞膜の機能を表現可能



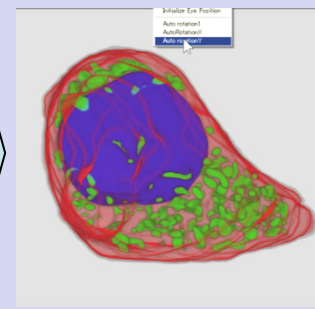
RICSを使用すると、細胞の局所(一部分)で起こっている現象と、細胞の集合体の解析が可能となる！！

→ 臓器と細胞をつなぐシミュレーションを目指しています。

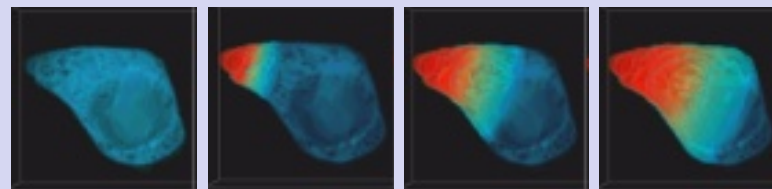
顕微鏡画像から作成した3次元形状内での計算



共焦点レーザー顕微鏡画像
(HepG2細胞: 肝臓細胞)

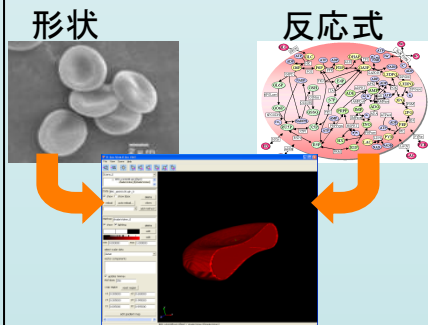


3次元再構築した細胞形状



時間

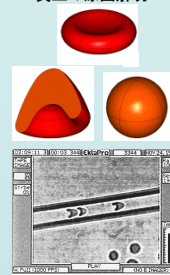
RICSの計算イメージ



細胞チームのターゲット (RICSを用いた計算)

赤血球

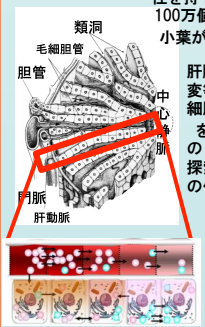
形状の変化による酸素運搬の仕組みを探る
貧血の原因解明



細い管を通過するときの赤血球形状

肝臓細胞 (肝小葉)

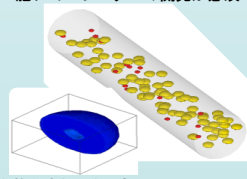
部位特異的な代謝特性を持つ肝細胞が約100万個集合して肝小葉ができている



肝臓ガン、肝硬変等で、一部の細胞が機能障害を起こした時の・治療方法の探索や薬物作用の仕組みを探る

血小板

血小板は血管の傷を修復する。血小板が血管からの作用によって活性化すると、血栓を形成し心筋梗塞、脳梗塞などの疾病の発症につながる。メカニズムの理解と新薬開発には細胞スケールと臓器全身スケールを連携する血小板細胞シミュレーターの開発が必須

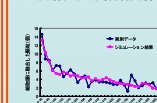


血管を流れる血小板のシミュレーション (左: 細胞レベル、右: 臓器レベル)

膵臓β細胞

日本人を含めたアジア人の糖尿病はインスリン分泌の不足が主な原因である。インスリンは膵臓β細胞から分泌される。インスリン分泌のメカニズムを顕微鏡画像のデータをもとにモデル化し、糖尿病の原因を探る

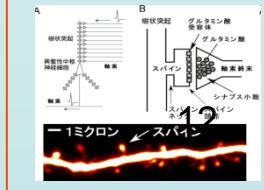
マウス膵臓β細胞からのインスリン分泌 (全反射蛍光顕微鏡画像)



グローバルに異なる膵臓β細胞のインスリン分泌をリアルタイムに可視化する装置が構築されています。

神経細胞

神経細胞では細胞膜が興奮することによって情報が伝達される。その機能障害は聴覚や記憶障害などに影響を与える。特に神経細胞でのシナプス可塑性は記憶に重要であり、そのメカニズムを探る。



1マイクロン スパイン



ISIRI

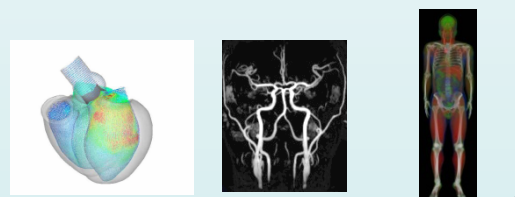
臓器全身スケール研究開発



RIKEN

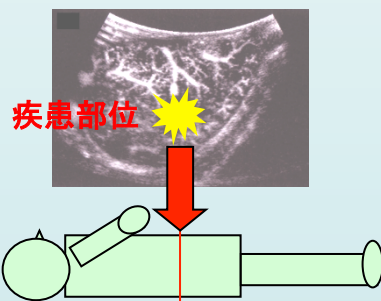
開発目標: ペタフロップス級の計算量を用いて、数時間で実行可能な臓器全身シミュレーションの構築を目指す。これにより、病態予測や治療予測等の医療支援ツールの開発につなげる。

開発計画: 血管網、各種臓器、全身を3次元的に再現した臓器全身モデルを構築し、病態予測や治療予測を行う臓器全身モデルを次世代スパコンに実装、循環器系疾患を始めとする種々の病態のシミュレーションと、ガンなどの治療支援シミュレーション(放射線・集束超音波など)につなげる。

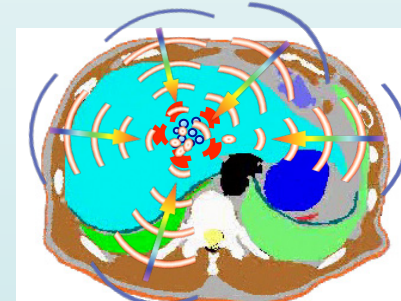


臓器全身モデルの構築

治療・診断の例(超音波を用いた悪性腫瘍の治療予測)



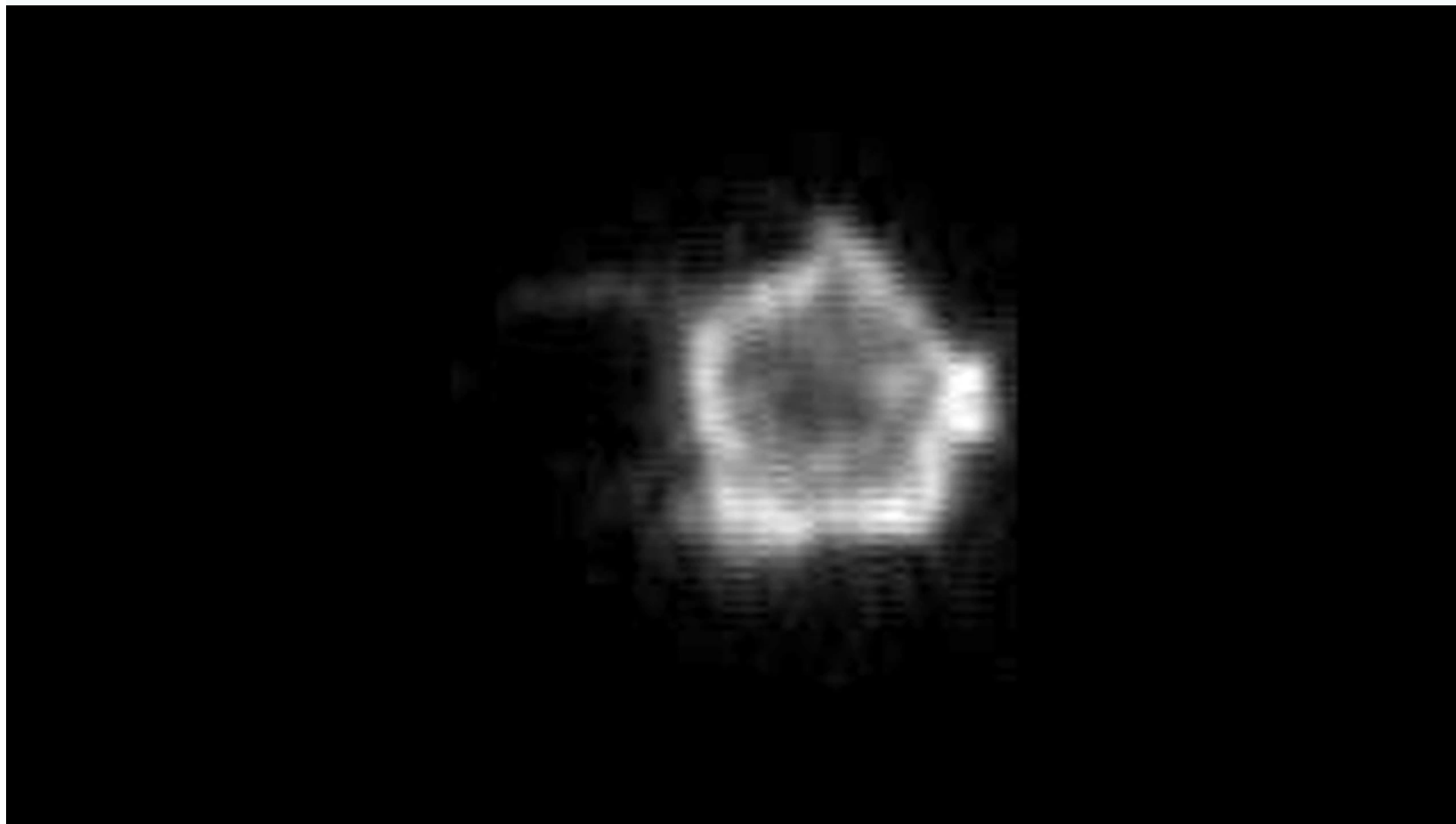
個別患者のモデルを作成



超音波伝搬シミュレーション

進捗状況: 1mm分解能の全身ボクセルモデルを開発すると同時に新しい構造流体連成手法と多媒質内の超音波伝搬解析手法を開発した。これらにより、外科手術シミュレーションや収束超音波・重イオンビームによるガンの治療等へつなげる基盤を構築。従来のスパコンで8000並列までの性能を確認。京でのテスト中

血小板の凝集と血栓の生成



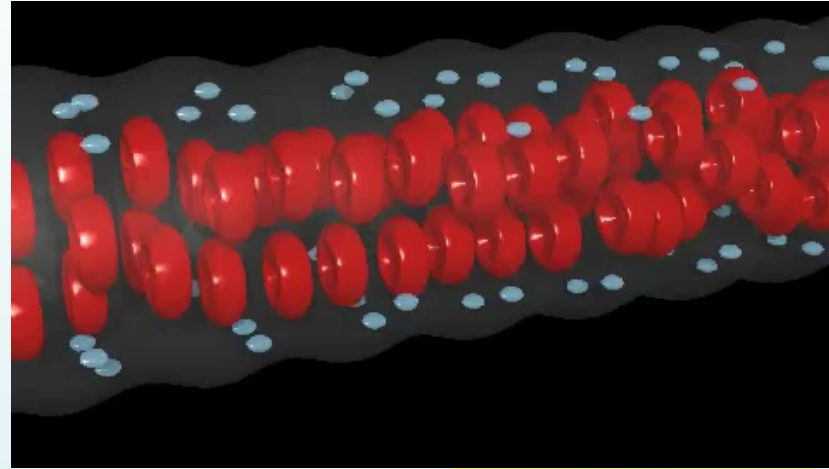


ISLAM

血栓症のマルチスケール・マルチフィジックスシミュレーション



多数の赤血球や血小板を含んだ血流計算

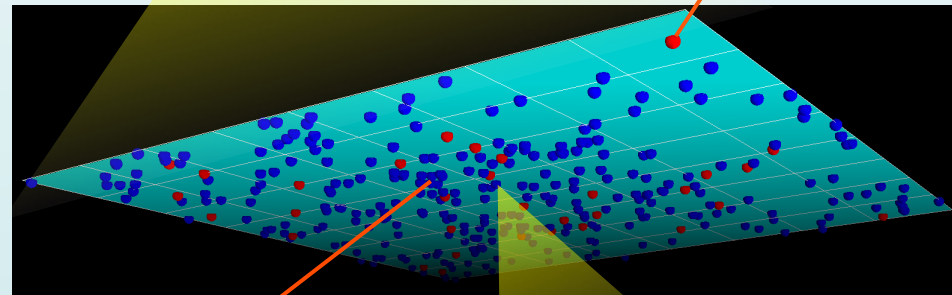


vessel wall

red blood cell

platelet

血小板, 血管壁間の分子間結合に関する動的モンテカルロ計算



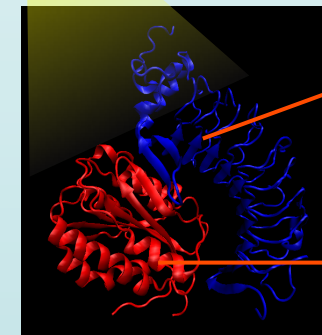
GPIIb/IIIa-vWF bond

GPIIb/IIIa

現状: 毛細血管が対象

「京」では: 心筋梗塞の対象となる直径数mm程度の冠動脈の計算が可能

リガンド-受容体分子の相互作用に関する分子動力学計算



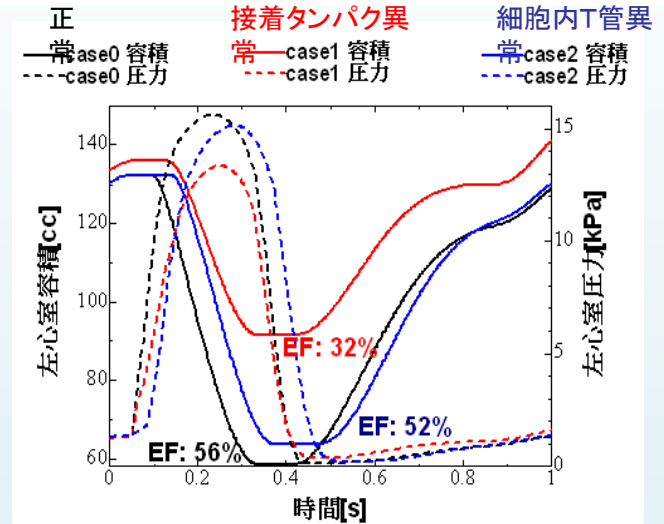
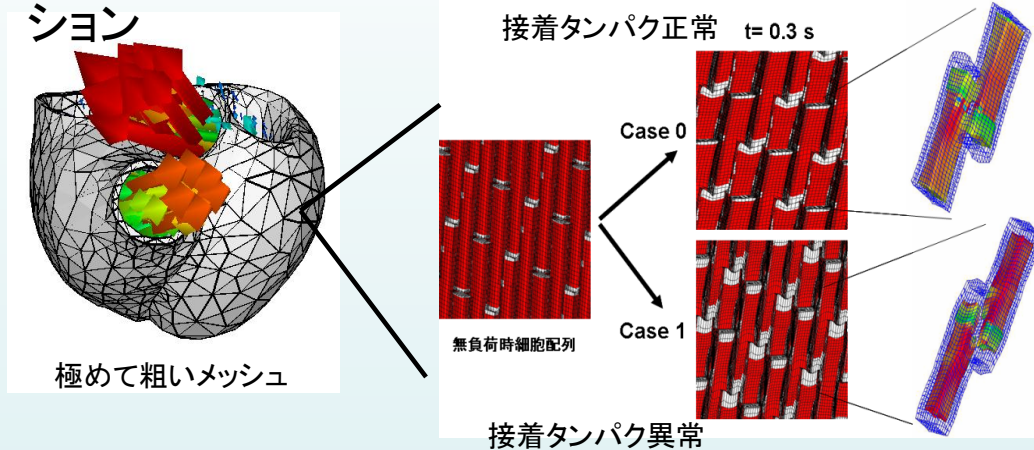
GPIIb/IIIa N-terminal

vWF 15 A1 domain

マルチスケール・マルチフィジックス心臓シミュレータ **UT-Heart**

マイクロレベルの異常と心疾患の関係を合理的に解明し、医学・医療に貢献する日本の独自技術

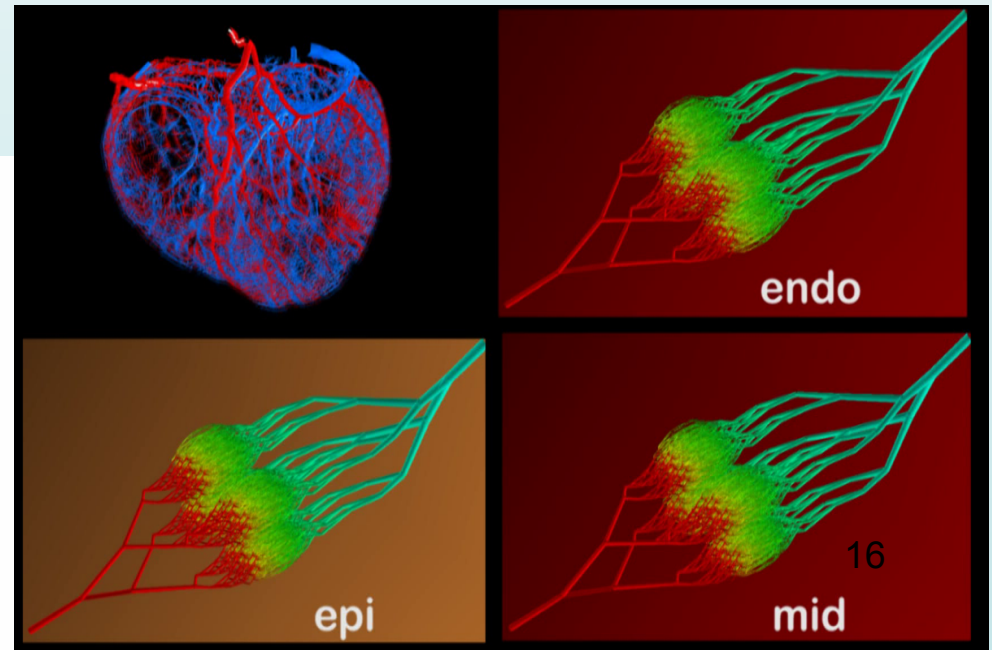
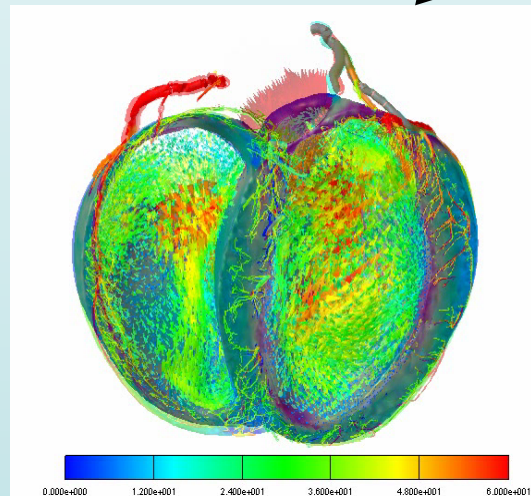
RICC 8000コアによる予備的シミュレーション

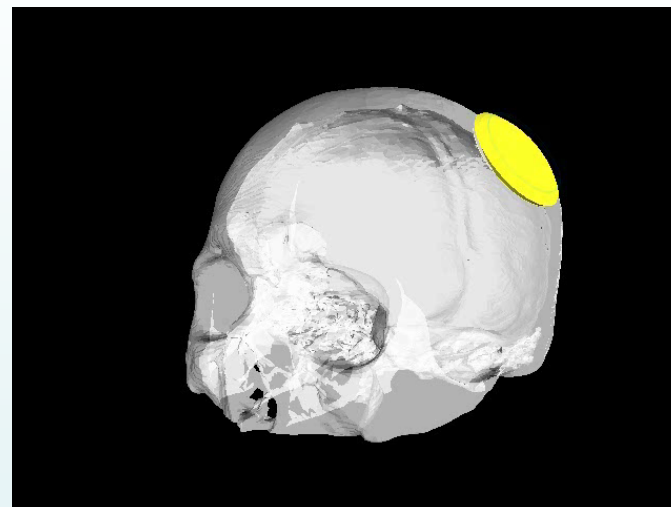
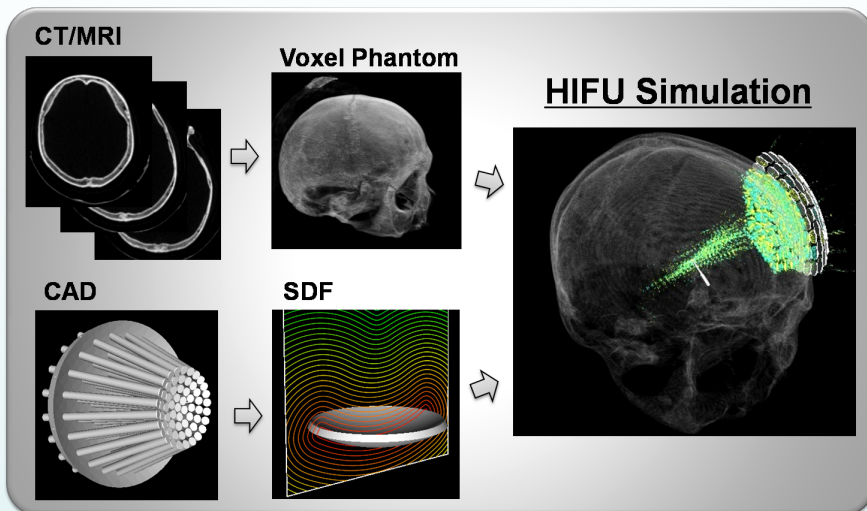


京では更に「毛細管まで再現した冠循環」と「細胞内の代謝機構」を導入し、虚血性心疾患など多様な病気を再現

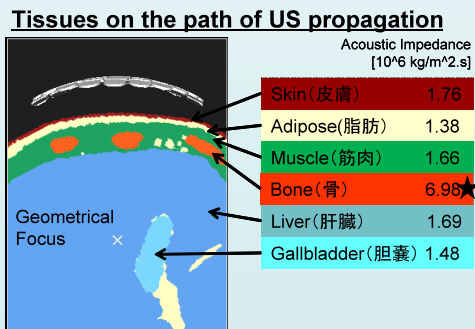
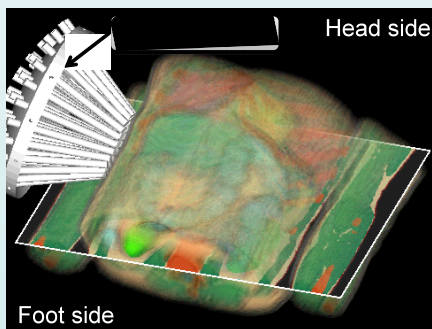
正常な冠循環：大動脈から入った血液は4拍程で毛細血管床に至る

右の動画は冠状動脈（左回旋枝）閉塞により、始めは酸素、やがてATPが欠乏する急性心筋梗塞のシミュレーション。
色は収縮力を表す
可視化：富士通（株）

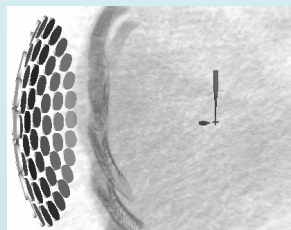




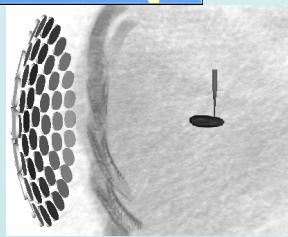
医用画像データとCADデータを直接利用するHIFUシミュレータ



実機設計のための
詳細シミュレーション



焦点制御なし



焦点制御あり

時間反転法による肝腫瘍焼灼シミュレーション

現状: 低解像度・ミリ秒スケール
超音波伝播シミュレーション

「京」では: 高解像度・実機設計
用詳細計算 & 秒スケール腫瘍
焼灼シミュレーション



データ解析融合研究開発



ISL:IM

RIKEN

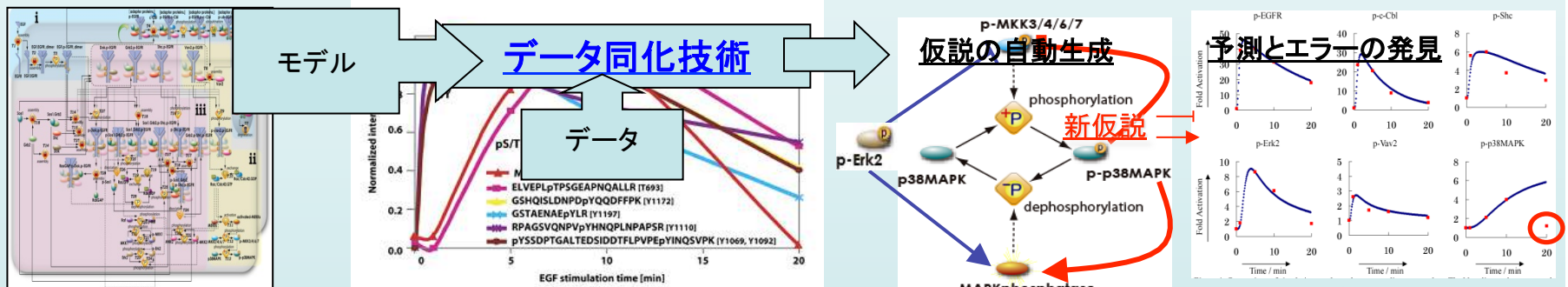
達成目標:ペタフロップス級の計算によって、創薬ターゲット探索や個人差を考慮した医療のための基盤情報技術の構築を目指す。これにより、ヒト全遺伝子を対象とした創薬ターゲット遺伝子探索の実現につなげる。

開発計画:遺伝子ネットワーク推定・タンパク質ネットワーク推定・ゲノム多型解析技術等の研究、ならびにデータ同化技術の開発を進め、開発した解析法やデータ同化技術を次世代スパコンに実装し、データ解析とシミュレーションを融合した技術を開発する。

シミュレーションモデルの構築

関連するデータをモデルに同化

病気のシステム的理解による医療開発イノベーション



(癌などの病気に関わっているモデル)

(関連たんぱく質の時系列データ)

データエラーの発見

進捗状況:分子のネットワークを地図として抽出するための方式として、大規模遺伝子ネットワーク探索及びタンパク質構造に基づく相互作用予測の研究を行い、新たな技術開発の成功と新たな並列プログラムの開発により、これらのプログラムが8000を超えるcoreで稼動するようになった。



ISL&M

大規模ゲノム多型と表現型データの関連解析

(角田達彦, 三澤計治, 藤本明洋)



RIKEN

ゲノムワイド関連解析による
疾患関連・薬剤応答遺伝子の
全ゲノム上の網羅的探索

現時点では個々のSNPを探

索

爆発的展開を示す GWAS
(Genome-wide
Association Study):
*Nature Genetics*誌
の55%を占める(2010年)

複数のSNP, その他の多くの多様性を用い探索する時代

- ・ **ハプロタイプ**解析(近くにあるSNPどうしをまとめて解析)
- ・ **複数の遺伝子**による**相乗効果**の網羅的な解析
- ・ **次世代シーケンサーデータ**によるさらなる多様性を探索

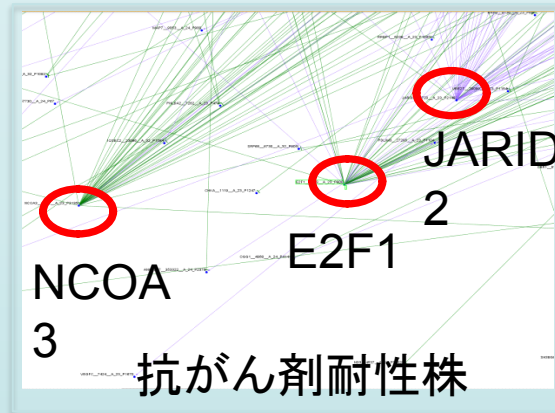
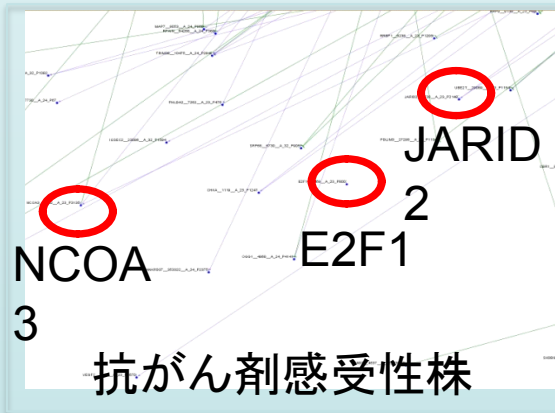
数ヶ月～年かかる計算が, 「京」上では**数十分**に!

疾患関連・薬剤応答遺伝子の**包括的な**発見へ



スパコンで焙り出すがんの黒幕たちのネットワーク

- TCGA, GEO, Sanger Center, NCIなどのデータベースには、肺がんや乳がんなど、**ISLIM** 主ながん種について、抗がん剤応答をはじめとして、様々な実験条件で計測された膨大な数の遺伝子発現プロファイルデータが公開されている。今後、次世代シーケンサーによるRNA解析で、データは激増する。
- 遺伝子発現プロファイルデータセットから全遺伝子を漏れなく対象として、高精度の遺伝子ネットワークを推定することは膨大な計算を必要とするが、精度を担保した高並列で動作するアルゴリズムはなく、これまでその計算は限定的であった。
- 2011年、大規模遺伝子ネットワーク推定ソフトウェアパッケージSIGNは、超並列化に適したネットワーク推定法と並列アルゴリズムを同時に開発することにより、この困難を解決し、これまでの小規模・モザイク精度レベルから飛躍して、全遺伝子規模・高精度ネットワーク推定を可能にした。
- **期待される成果**: がんに関する数百の主要なデータセットを京で一挙に解析し、薬剤応答、薬剤耐性、転移、浸潤、EMT、個々人のがんのシステムなど、がんに関する網羅的な大規模遺伝子ネットワーク地図を世界で初めて作成することができる。これはがん研究のGoogle Mapに例えられる。

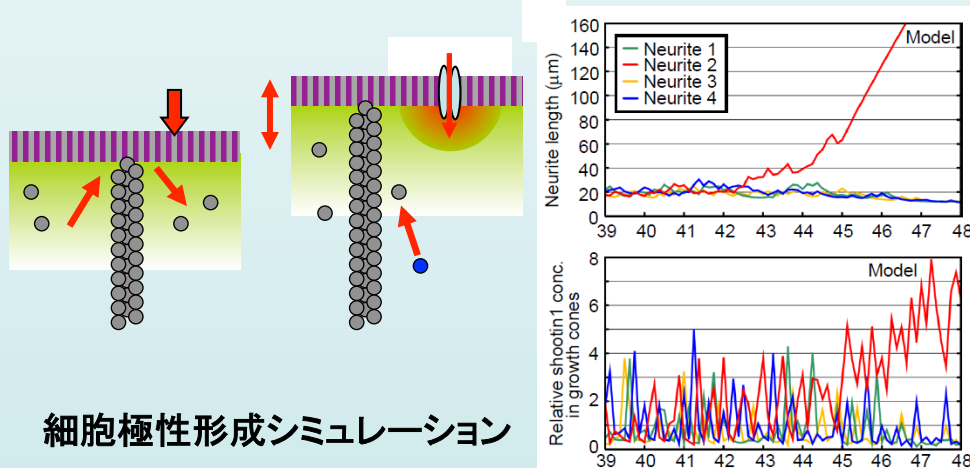


抗がん剤Gefitinib感受性肺がん細胞株と耐性を獲得した肺がん細胞株をEGF(上皮細胞成長因子)で刺激した遺伝子発現プロファイルデータセットから推定したがんの遺伝子ネットワークを比較したもの一部。左(感受性)と右(耐性)で、たとえばE2F1とよばれるアポトーシスや細胞周期を制御している転写因子のハブ構造が大きく異なっていることがわかる(ここで点は遺伝子を表し、線(A→B)は遺伝子Aと遺伝子Bの転写物間に活性や抑制の関係があることを表している)。

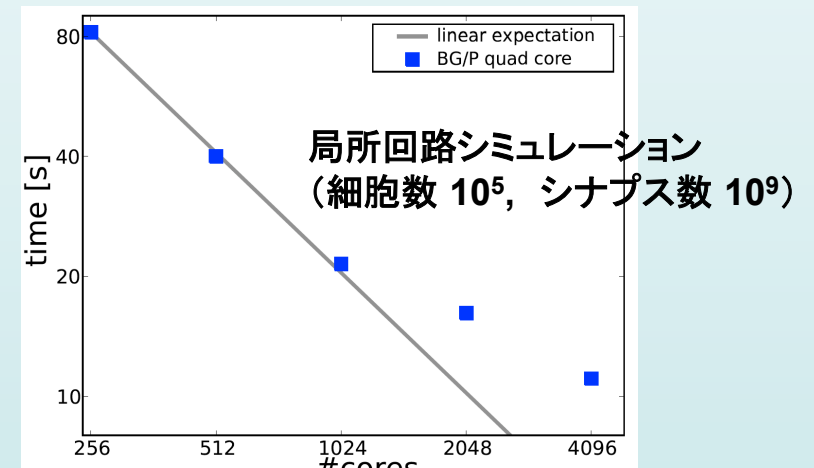
開発目標: 脳全体レベルでの入出力関係を再現できるモデルとシミュレーションソフトウェアを開発、次世代スーパーコンピュータの性能を引き出して、脳の柔軟な環境適応能力である脳の発達と学習の再現を目指す。

開発計画: 神経細胞と局所回路のシミュレータ、および、脳のシミュレータとして昆虫の匂い情報処理回路モデル、ほ乳類の網膜モデルとを開発、次世代スーパーコンピュータに適したプログラムとすることで、昆虫の嗅覚系情報処理では実時間で、ほ乳類の視覚系情報処理は60倍の時間で計算可能とする。

神経マルチフィジクスシミュレータ

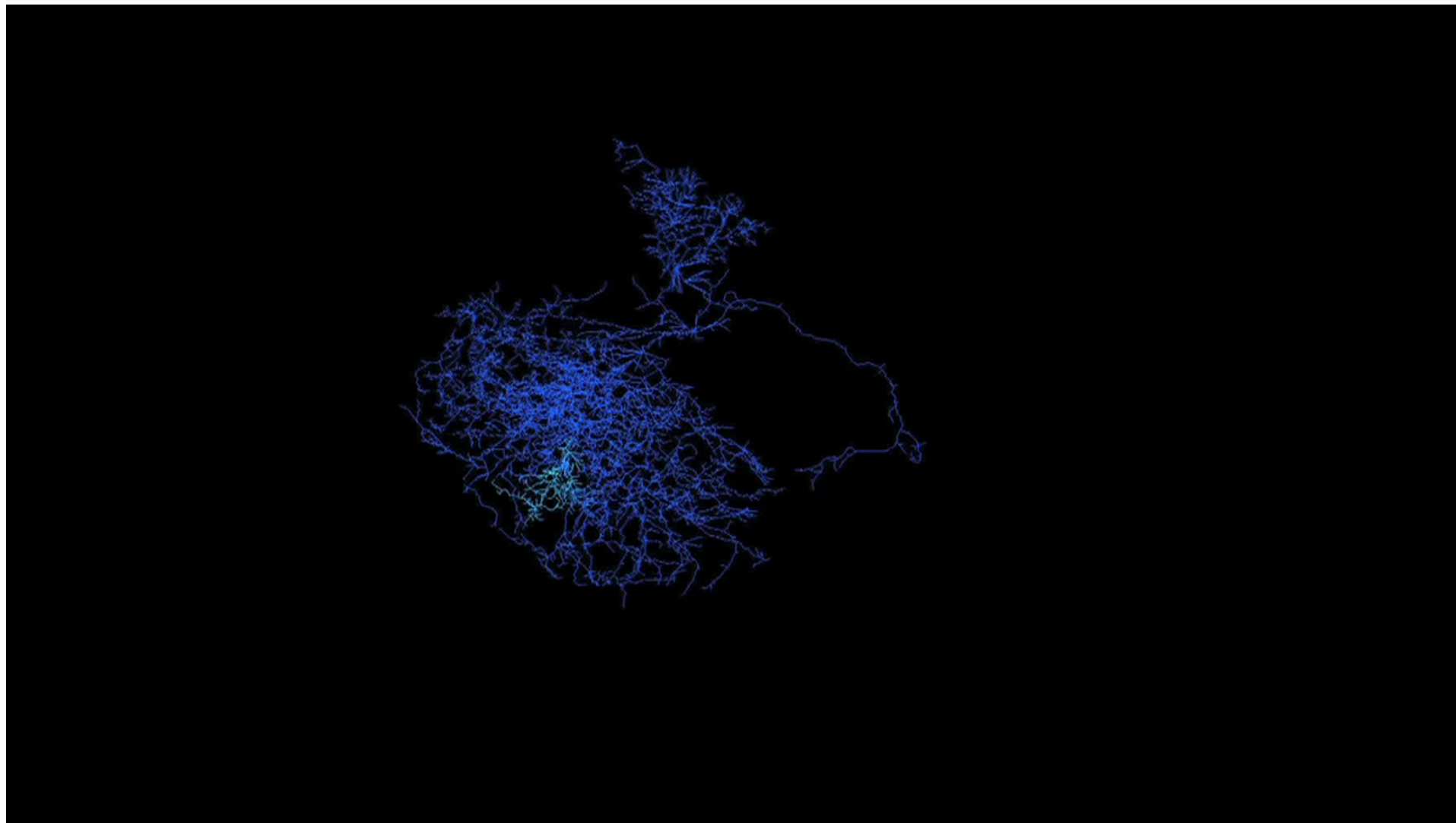


局所回路シミュレータNEST



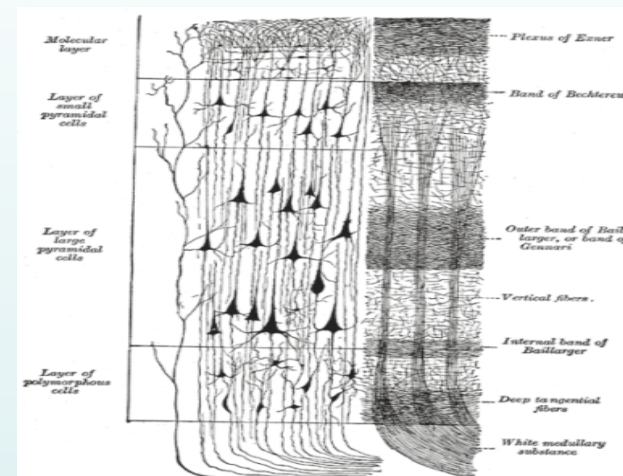
進捗状況: 本チームは2008年10月に発足、昨年度は神経マルチフィジックスシミュレータでは神経細胞の極性形成について、分子レベルからのシミュレーションができた。また、局所回路シミュレータNESTでは 10^5 個の神経細胞、 10^9 個のシナプスからなる大規模局所回路のシミュレーションができた。並列性能に関しては32000並列までのテストを終了、京でのテスト中

神経系シミュレーション(蚕蛾)

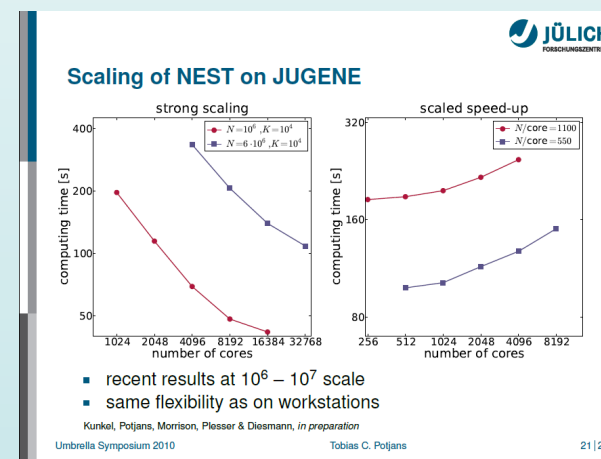


局所回路シミュレータNESTによる 大脳皮質シミュレーション

- シミュレーションの達成目標：大脳皮質の100コラムに相当する1千万神経細胞、および神経細胞間を接続するシナプス1千億による信号処理の再現・予測を行う。
- アプリケーションの開発状況：すでに、Eulich research center のEugeneを用いて、16384並列までの性能の向上を確認済み。部分稼働している京で、12000並列の実行を確認済み。
- 京の完成後には、数十万並列によるシミュレーションの実行が実現可能となり、世界最速の大脳皮質シミュレーションを実行し、その機能的役割の解明を目指す。

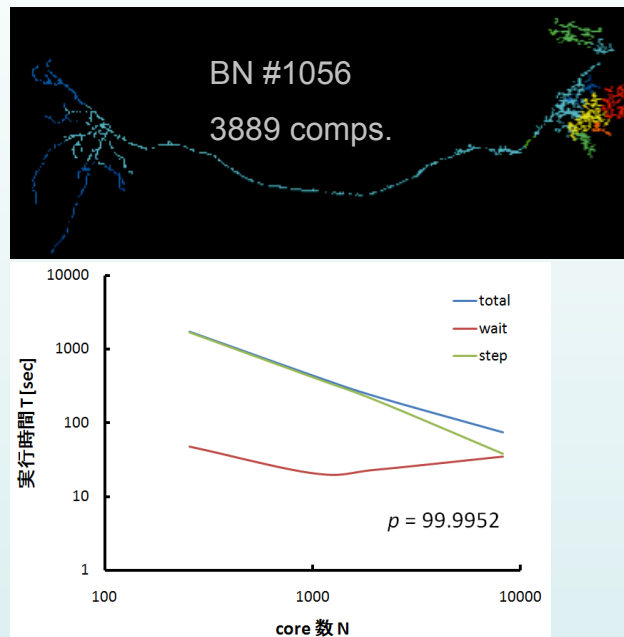


大脳皮質の構造模式図



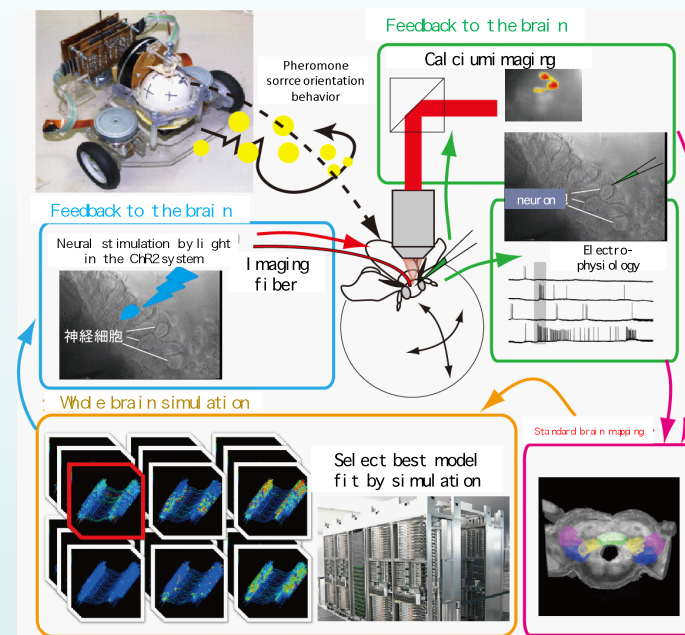
並列数と計算性能
(Provided by Tobias
Potjans)

無脊椎動物神経回路シミュレータIOSSIMによる 嗅覚情報処理シミュレーション



マルチコンパートメントモデルと並列性能

- 京のコンパイルは成功・ハイブリッド並列化も成功
- テストネットワーク(各ニューロンは1ニューロンに接続)はRICC8192並列までは良好な並列性能(図: P=99.9952)

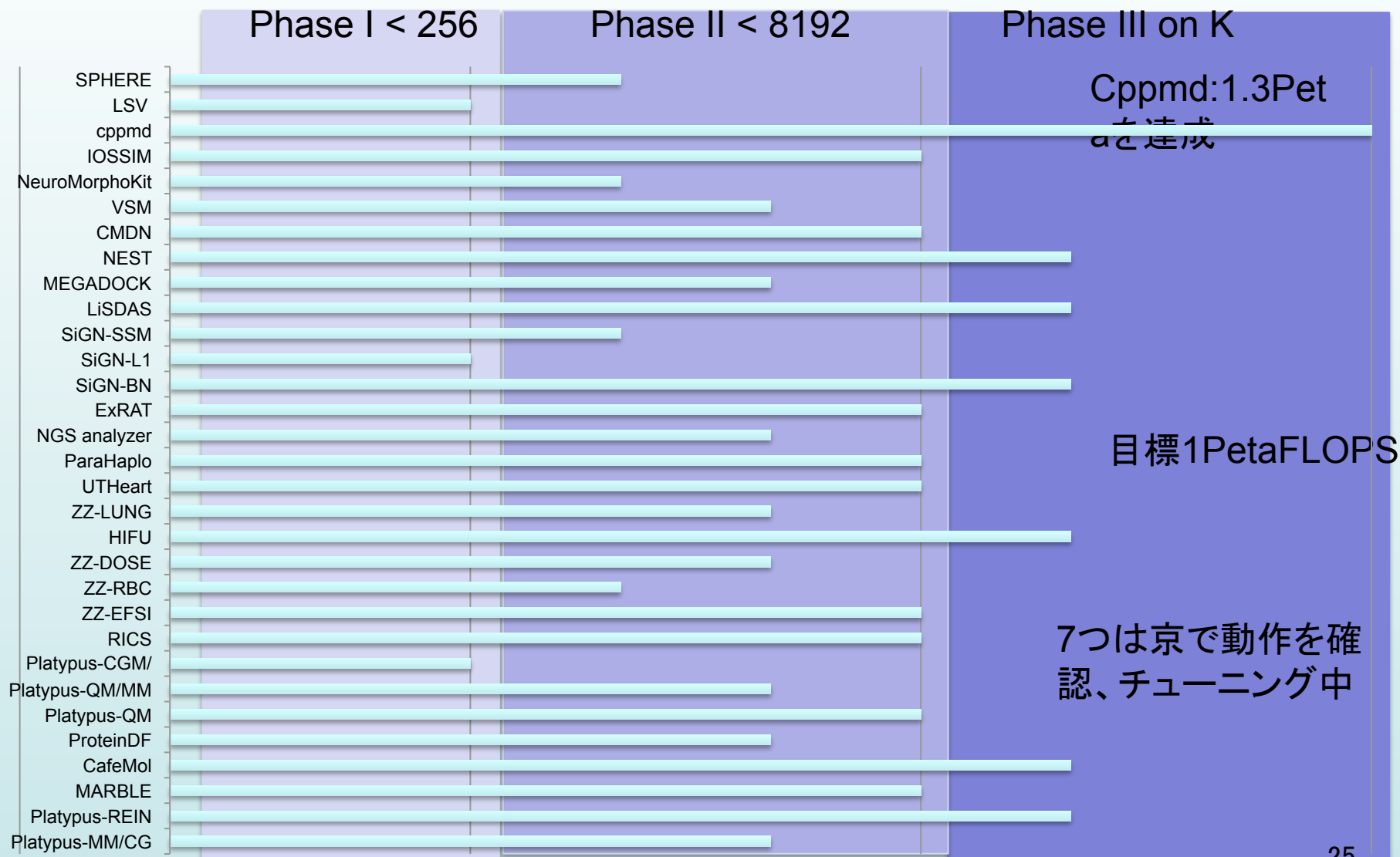


京を用いたシミュレーションの目標

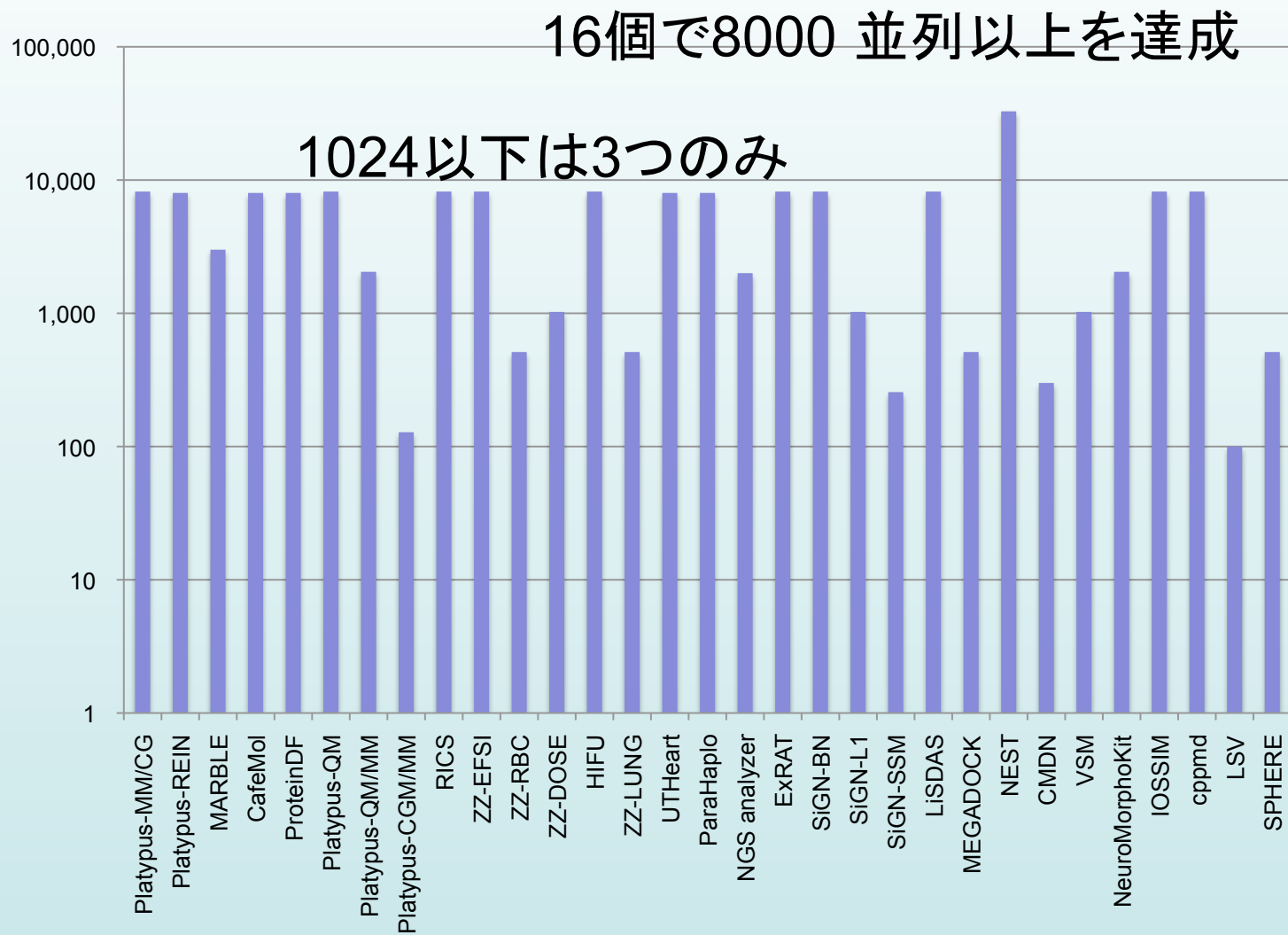
目標: 匂い受容から行動指令までの3000-4000ニューロンを形態を含めて詳細に再現したモデルによるリアルタイムシミュレーション

将来展望: 生理実験と組み合わせることで生体の神経活動のリアルタイムな推定・制御の可能性を切り開く(図)

ソフトの開発状況

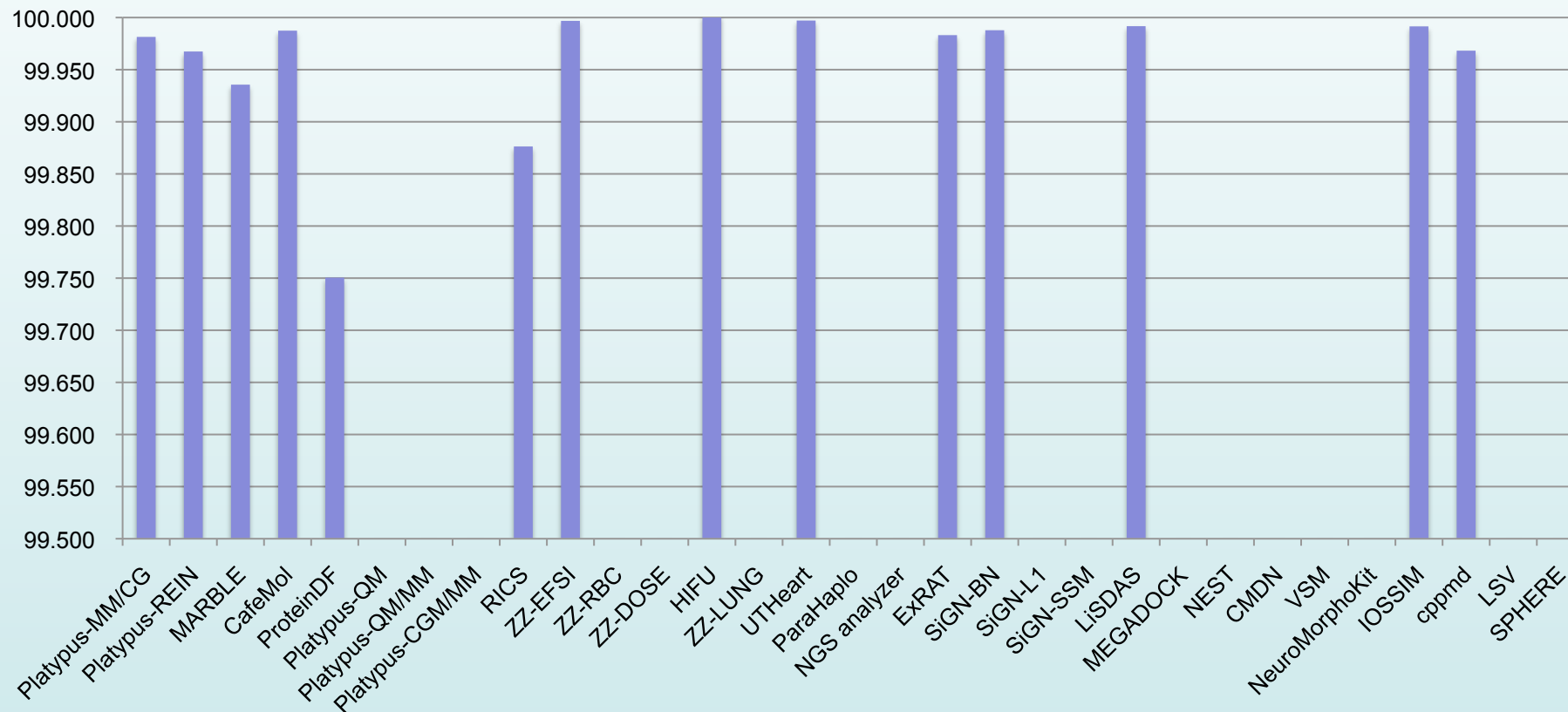


並列の達成度



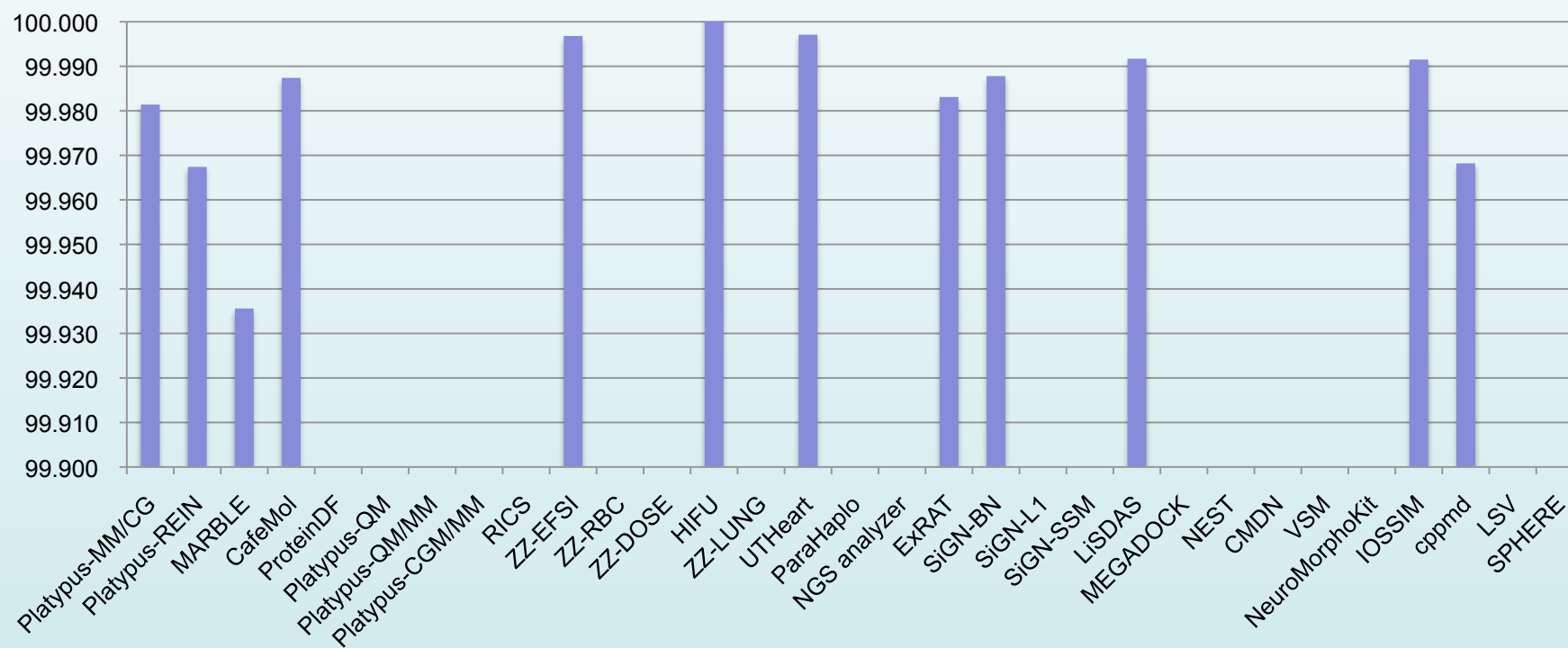
並列化率(1)

14個のアプリで99.75%



並列化率(2)

5個で99.99%以上

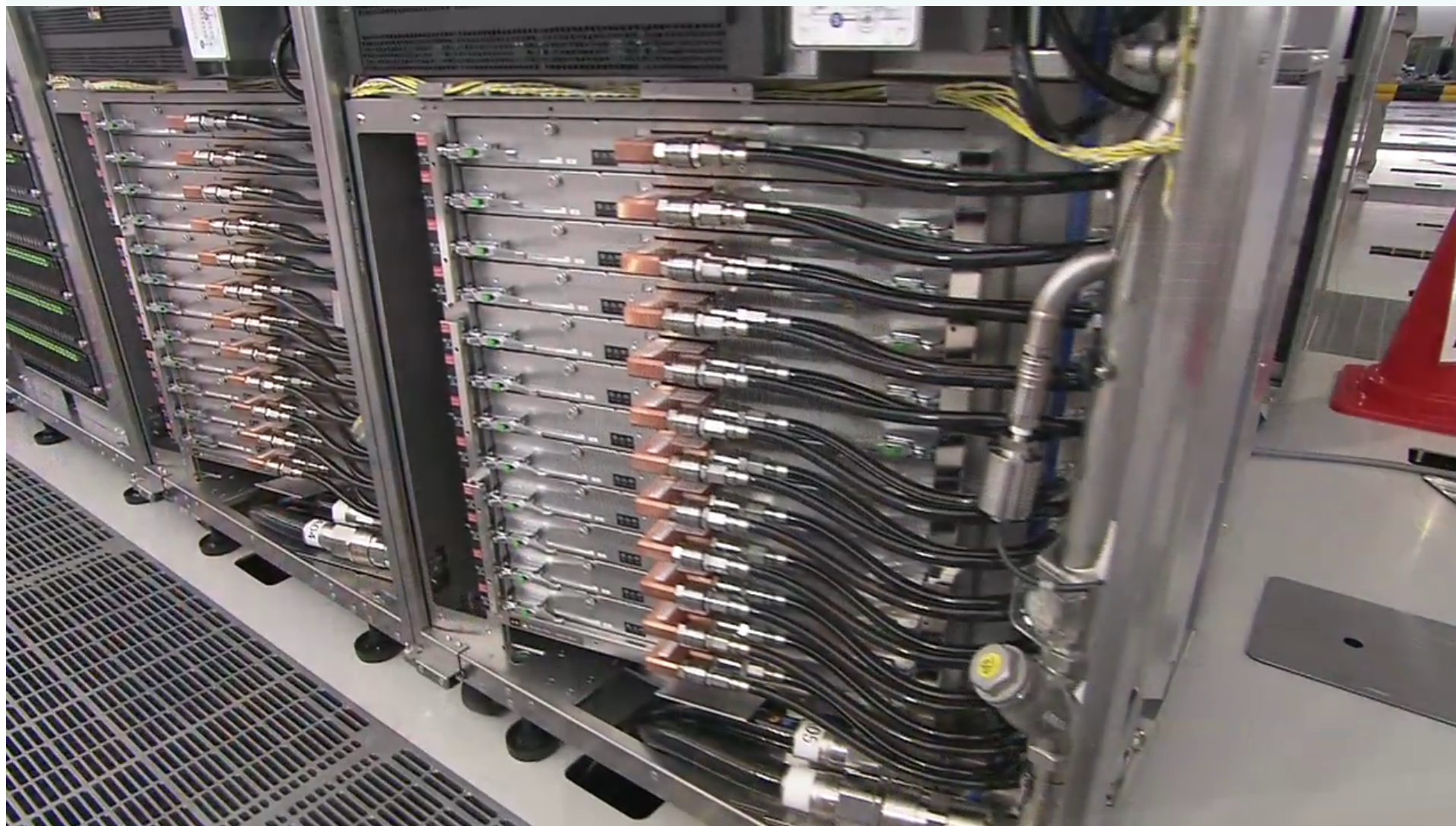




京に最適化した分子動力学ソフト: cppmd



- 1.3PetaFLOPS, 実効効率37.2%(3月末)

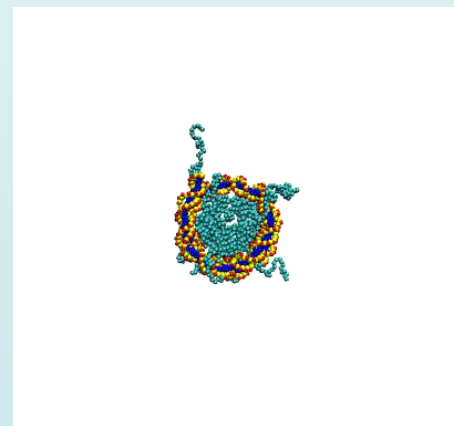
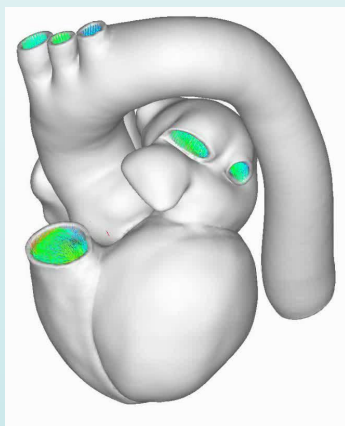


	アプリケーション名	略称	開発責任者	言語	並列化の方法
分子	密度汎関数法に基づくタンパク質全電子波動関数計算	ProteinDF	佐藤(東大)	C++	MPI, OpenMP
	粗視化モデル計算	CafeMol	高田(京大)	Fortran90	MPI, OpenMP
	全原子分子動力学計算	MARBLE	池口(横浜市大)	C	MPI
	マルチコピー・マルチスケール分子シミュレーション法開発の基盤となるクラスライブラリ	Platypus-MM/CG	木寺(横浜市大・理研)	C++	MPI
	ハイブリッドQM/MM反応自由エネルギー計算	Platypus-QM/MM-FE	林(京大)	FORTTRAN77	GAMESSのsocksライブラリ
	レプリカ交換分子動力学計算インターフェイス	Platypus-REIN	杉田(理研)	Fortran90	MPI
	粗視化モデル計算/分子動力学計算	Platypus-CGM/MM	中村(阪大)	C++, Fortran90	MPI
	量子化学計算/分子動力学計算	Platypus-QM/MM	中村(阪大)	FORTTRAN77, Fortran90	MPI
	量子化学計算	Platypus-QM	中村(阪大)	FORTTRAN77, Fortran90	MPI
細胞	細胞シミュレーションプラットフォーム	RICS	横田(理研)	Fortran, C, C++	Sphere(MPI)
臓器全身	全身ボクセルシミュレーション(ボクセル構造流体連成解析プログラム)	SPH3D	高木(東大・理研)	Fortran, C, C++	Sphere(MPI)
	マルチスケール・マルチフィジックス心臓シミュレーション	UTHeart	久田(東大)	Fortran	MPI&OpenMP
	低侵襲治療シミュレーション (ボクセル超音波伝播プログラム)	HIFU	高木・松本(東大)	Fortran90	MPI&OpenMP
	微小循環シミュレータ (埋め込み境界法による微小循環プログラム)	ZZ-RBC	高木(東大・理研)	Fortran90	Sphere(MPI)
	重粒子線治療シミュレーション	ZZ-DOSE	高木(東大・理研)	Fortran90	MPI
	肺呼吸・肺循環シミュレーション	ZZ-LUNG	和田(阪大)	Fortran, C, C++	Sphere(MPI)

アプリのリスト(2)

データ解析融合	ハプロタイプ関連解析に於ける統計検定を行うためのソフトウェア	ParaHaplo	鎌谷(理研)	C	MPI, OpenMP
	大規模遺伝子制御ネットワーク推定プログラム	SiGN	宮野(東大)	C99	MPI, EP (SGE)
	再帰的正則化法による生体内分子の大規模ネットワーク推定プログラム	L1GN	宮野(東大)	R	EP (SGE)
	状態空間モデルによる時系列データからの遺伝子ネットワーク推定プログラム	SSM	宮野(東大)	C, C++, JAVA	MPI, OpenMP
	データ解析融合プラットフォームの開発	SBiP	宮野(東大)		プラットフォーム
	網羅的タンパク質ドッキング解析プログラム	MEGADOCK	秋山(東工大)	C++	MPI
	生命体データ同化プログラム	LiSDAS	樋口(統数研)	Fortran90, C, C++	MPI, OpenMP
	次世代シーケンス解析プログラム	NGS analyzer	鎌谷(理研)	perl, shell	EP
	拡張RAT法による2SNP組合せの全ゲノム関連解析ソフトウェア	ExRAT	鎌谷(理研)	C++	MPI
脳神経	Neural Simulation Tool	NEST	Diesman(理研)		
	Cortical Microcircuit Developed on NEST - Interneuron	CMDN	深井(理研)	SLI, C++	MPI
	神経細胞形態シミュレーションキット	NeuroMorphoKit	石井(京大)	MATLAB, C/C++	逐次
	視覚系シミュレーションのための共有プラットフォーム	VSM	臼井(理研)		プラットフォーム
	昆虫嗅覚系全脳シミュレータ	IOSSIM	神崎(東大)		実験系アプリ
基盤	大規模並列用MDコアプログラム	cppmd	泰地(理研)	C++	MPI
	大規模仮想化合物ライブラリー	VLSVL	船津(東大)		EP
	分散並列大規模データ可視化システム	LSV	小野(理研)	C++, C	MPI
	アプリケーションミドルウェア	SPHERE	小野(理研)	C++,C	MPI

次世代スーパーコンピュータで なにを狙っているか



それぞれの成果アピール時期

これらのアプリは準備状況などで適宜見直して行く



第一走者

cppmd (分子動力学)
 ProteinDF
 CafeMol (ヒストンDNA)
 UTHart
 SPH3D
 ParaHaplo
 NEST (神経回路網)

完成後すぐに成果をアピール



第二走者

多剤排出トランスポータ
 MARBLE
 Platypus
 CafeMol
 RICS:肝細胞
 HIFU ZZ-DOSE重イオン粒子
 SIGN SSM L1GNE
 LiSDAS
 MEGADOCK
 CMDN:大脳皮質局所回路

完成後、1年以内に成果をアピール



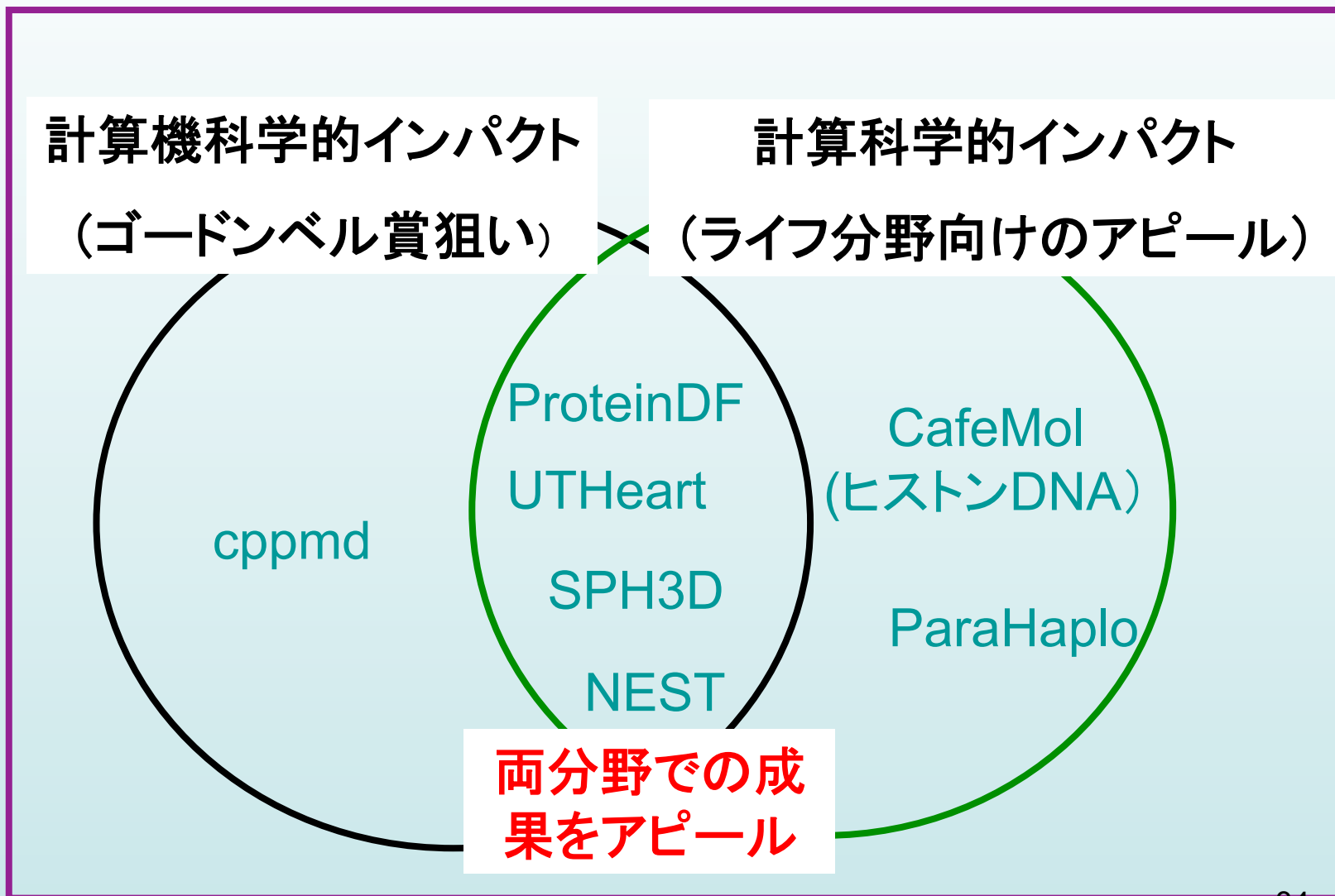
第三走者

神経系統合シミュレータ
 細胞から臓器へ
 分子から細胞へ

中長期目標

スケールを越えた統合的理解に向けた取り組み

第一走者の分類

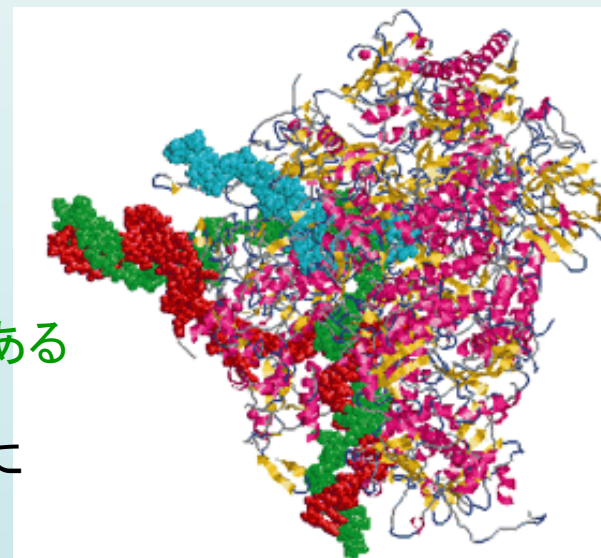


世界最速の分子動力学計算でGordonBell賞を目指す

— タンパク質の機能を高速で長時間シミュレーションするための計算技術 —

ソフトウェアの特徴

- 分子動力学計算ソフトウェア: **10万コア以上の超並列化** (100万-1000万原子系で) 10Petaflopsで1億原子級のシミュレーションを実現し、実効性能で最速を目指す
- 生体分子の運動の長時間シミュレーション
100万原子系のサブミリ秒シミュレーションを目指す
(対象例) 細胞膜構造、核膜輸送、ウイルスなど
【従来の100倍以上の時間スケール】
- これまで不可能だった時間スケールで、生理的に意味のあるタンパク質の大きな構造変化を捉えることにより
生命システムの原子レベルでの構成原理を理解すると共に
創薬に必要なレベルの計算精度を実現する

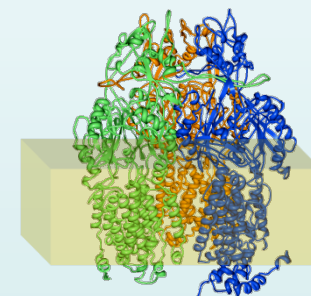


研究代表者: 木寺 詔紀 (理化学研究所)

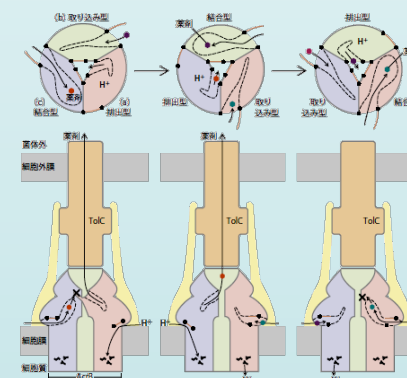
抗生物質が効かなくなる メカニズムの理論的解明を目指す

ソフトウェアの特徴

- 電子スケール～原子・分子スケール～タンパク質全体スケールの多階層(量子化学・分子動力学・粗視化モデル)の連成計算によって巨大タンパク質の動作機序を明らかにする。
- 計算対象: 細胞膜にある多剤排出トランスポーター
3つのタンパク質が結合した巨大タンパク質。
各々、取り込み型 → 結合型 → 排出型
と順繰りに構造を変えながら、薬剤化合物を排出すると考えられている。そのメカニズムを探る
- 細菌やがん細胞の薬剤耐性化の主要な原因の1つであるメカニズムの解明につながり、薬剤耐性化しにくい抗生物質や、排出を妨げる新薬の設計等への波及が期待できる。



多剤排出トランスポーター-AcrB

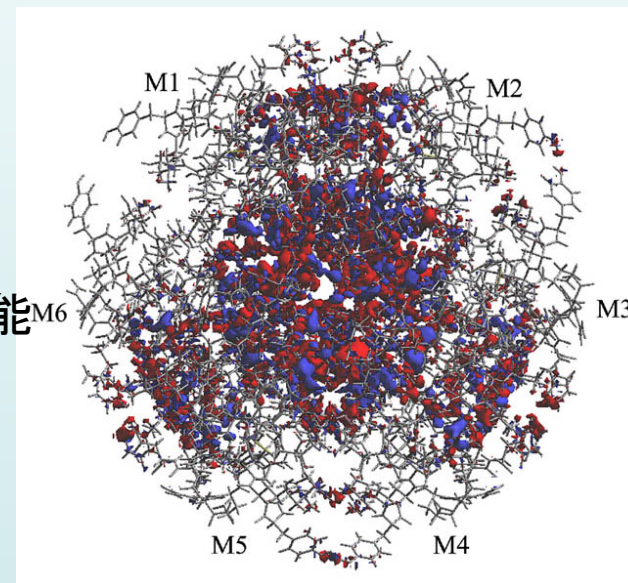


世界最大のタンパク質全電子計算を目指す

— タンパク質の電子レベルでの反応を正確に解明するための計算科学技術 —

ソフトウェアの特徴

- 密度汎関数法による全電子波動関数計算ソフト
タンパク質の励起状態を1万並列超で計算
- 巨大なタンパク質の励起状態や電子移動の計算が可能
(1千残基、10万軌道、励起状態計算)
→ Protein Data Bankの99%以上が計算可能に
- ミネラルを含むタンパク質(酵素やヘムタンパク質)の
働きを解明するのに有効
- 電子が主役の高速反応が詳細に解明可能になり、
生体模倣光電子素子、人工光合成などへの応用が期待
される



(成果の例)

インスリンの凝集に伴う電子の再分布
6量体と単量体の全電子分布の差を描いたもの
10Petaではアンテナタンパクなどが
ターゲット

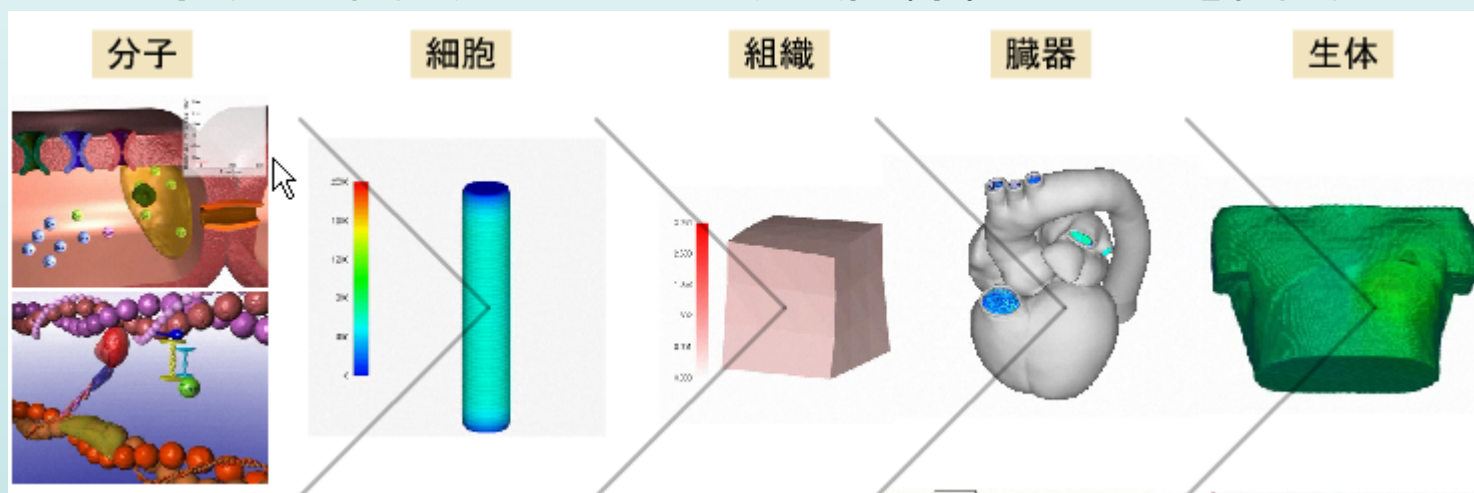
世界最大規模の 心筋細胞レベルからの心臓全体シミュレーションを目指す

— 世界でも前例のないバーチャル心臓を世界最大規模でコンピュータ上に再現 —

ソフトウェアの特徴

○分子レベルの電気化学、力学現象を再現した、心筋細胞のモデルを使って心臓を組み上げ、心臓の動きと血液の拍動を再現するシミュレーション・ソフト

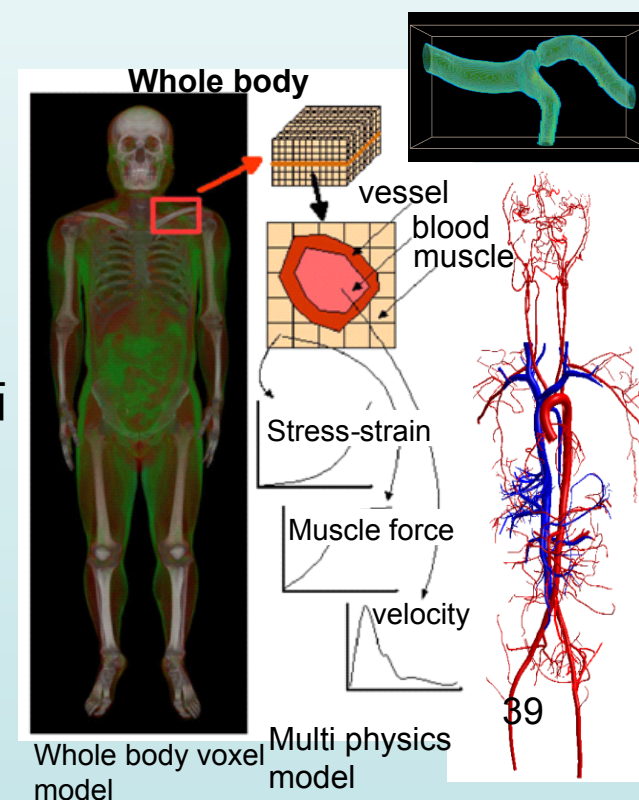
○各種心臓病(不整脈や拡張型心筋症など)の基になっている因子を細胞内で起きる分子レベルの挙動から検討することが可能。医療・創薬への応用を検討。



世界最大規模の 血栓の形成、輸送、梗塞の同時シミュレーションを目指す

— 医療応用に向け、やわらかな人体の解析・予測に適した計算科学技術 —
ソフトウェアの特徴

- 大きな血管から毛細血管までの赤血球、血小板を入れた計算ができる構造流体連成ソフトウェア
- 医療データ(CTスキャン, MRI等)から直接シミュレーションを行うことができ、個人毎の診断に即応してシミュレーションが利用可能に
- 血栓形成・閉塞による心筋梗塞, 脳梗塞の予測と薬効評価
- 動脈瘤・動脈硬化症等の循環器系疾患の予測と治療法の検討
- 全身血管網シミュレータとの連成による血圧波形(血圧測定)からの循環器系疾患の予測



世界最大規模の大脳皮質神経回路の活動シミュレーションを目指す

— 単純な神経細胞モデルから実際の脳の活動が説明できるかを解明 —

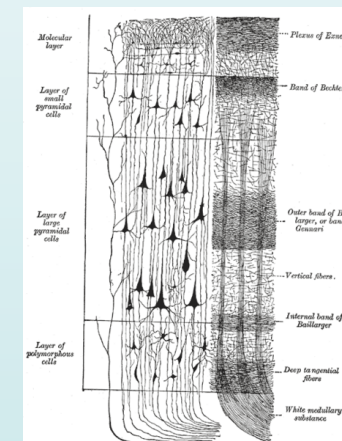
ソフトウェアの特徴

○大脳皮質の100コラムに相当する1千万神経細胞、神経細胞間を接続するシナプス1千億の信号処理を再現・予測

(従来の世界トップの100倍の規模)

○超高並列プログラム(現時点で32,000並列)

次世代スパコン上では数十万並列を達成、**世界最速を目指す**



大脳皮質の情報処理単位
1コラム (1mm³) に
10万個の神経細胞

膨大な長さのDNAが核の中に収納され、 必要なときに読み出せる謎に挑む

— 細胞核内でのDNAのヒストンへの巻き付きと、ほどこメカニズムの解明 — ソフトウェアの特徴

○規模・時間的に巨大な生体分子の現象をシミュレーションするための粗視化モデル計算

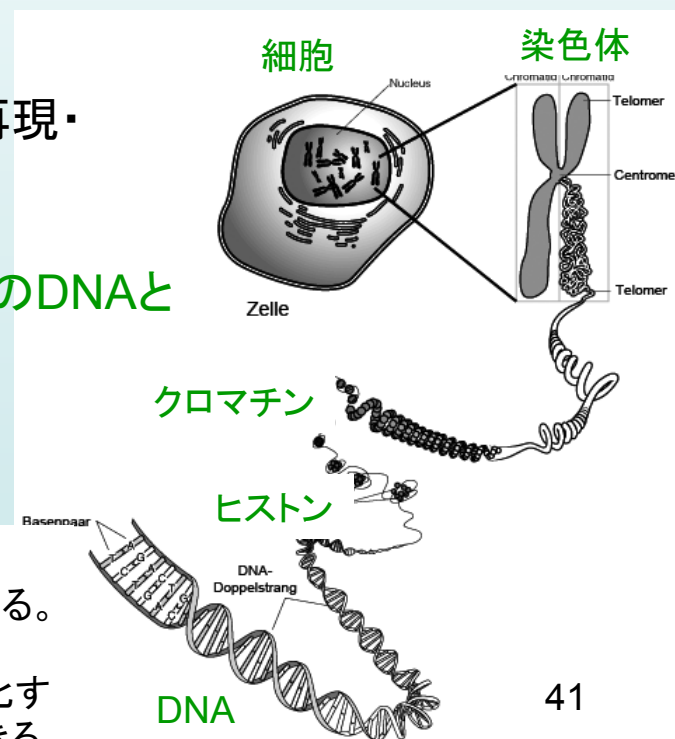
○世界最大1千万原子規模・世界最長のミリ秒の現象を再現・予測を目指す

○学術的なインパクトを与える、遺伝情報が転写される際のDNAとヒストンのダイナミクスを再現

細胞中で膨大な長さのDNAはヒストンに巻き付き、コンパクトに核に収納されている(クロマチン構造)

遺伝子が発現する際には、適切な場所のDNAをヒストンからほどこ、遺伝情報を転写している。

iPS細胞の機能は、クロマチン構造による転写制御機構を初期化することによって達成されるといわれている。そのなぞにも挑戦できる



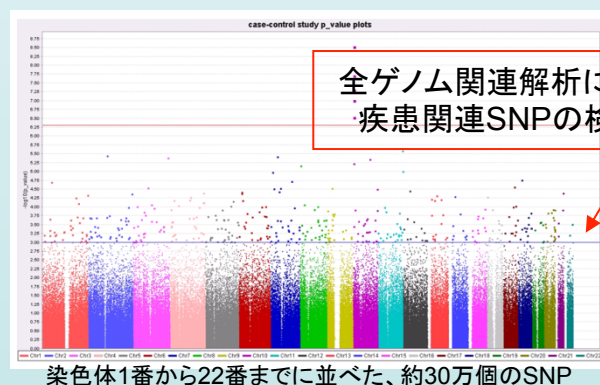
疾患の原因遺伝子を一挙に発見することを目指す

— 個人の遺伝情報と疾病、薬物反応性の関連を明らかに —

ソフトウェアの特徴

- 複数の遺伝子や環境要因(身長、体重、喫煙、飲酒など)の組み合わせによって決まる疾病や薬物の効き方、副作用などの関連を網羅的に解析
- 全ゲノム(30億塩基)の個人ごとの遺伝情報の違いの中から、疾病に関連する遺伝情報を、網羅的に精度よく探し出すことが可能に

SNP遺伝子型と疾患との関連を示すアーミテージ検定のP値の $-\log P$ 値

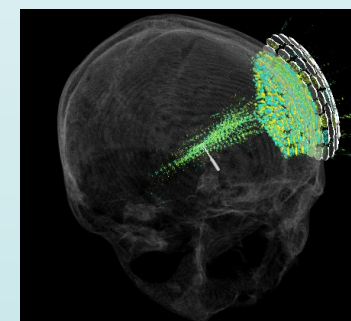
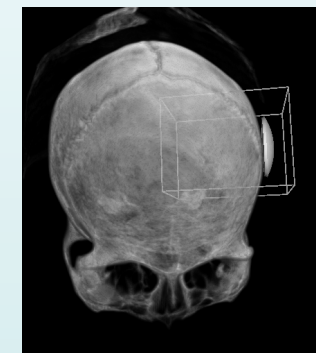


超音波治療機器の設計開発・安全性の確認に貢献

— 患部だけを焼く集束超音波治療の安全な適用に向けた計算科学技術 —

ソフトウェアの特徴

- 人体内(骨や臓器など音響特性の異なる多媒質体内)の超音波伝搬をシミュレーションするソフトウェア
- 医療画像データから開発されたボクセル人体モデルを使ってシミュレーション。HIFU(強力集束超音波)を使った治療機器の超音波の集束状況の高精度な制御をシミュレーション可能
- 現在の超音波治療では対応困難な肝臓などの体内深部にある臓器や脳など骨で囲まれた部位に対して、集束超音波で腫瘍を焼灼する医療機器の開発に利用。シミュレーション結果をもとに超音波の照射方法を検討し、手術計画を立案可能

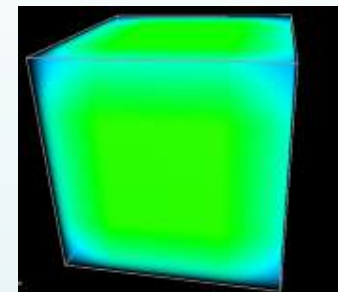


頭蓋骨越しのHIFU照射シミュレーション

反応—拡散—膜透過の連成計算を100万Voxel空間で実現!
(100³)

小規模モデルの計算時間(赤血球)

2×2×2(=8)ボクセルを逐次(1CPU)で計算
1000steps/197秒 Δt=1ms (実時間1秒)



2×2×2ボクセル
oxyHbを表示

赤血球形状モデル

64×64×25(=102400)ボクセルを逐次(1CPU)で計算
512並列:1.36h Δt=0.05ms
実時間1分:600h 10万並列:3hr

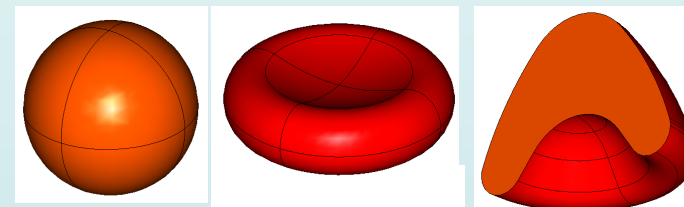


64×64×25ボクセル

肝細胞詳細モデル

100×100×100(百万)ボクセルを10万並列)で計算

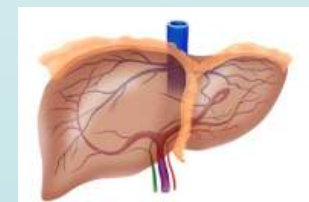
フル代謝モデルを想定 Δt=0.05ms
実時間1分 :10万並列:30hr



肝小葉モデル(代謝骨格モデル)

50×50×500(百万)ボクセルを10万並列)で計算

フル代謝モデルを想定 Δt=0.5ms
実時間10分 :10万並列:40hr





ISLIM



RIKEN



JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY
1990



KYOTO UNIVERSITY
KYOTO JAPAN
FOUNDED 1897



大阪大学
OSAKA UNIVERSITY



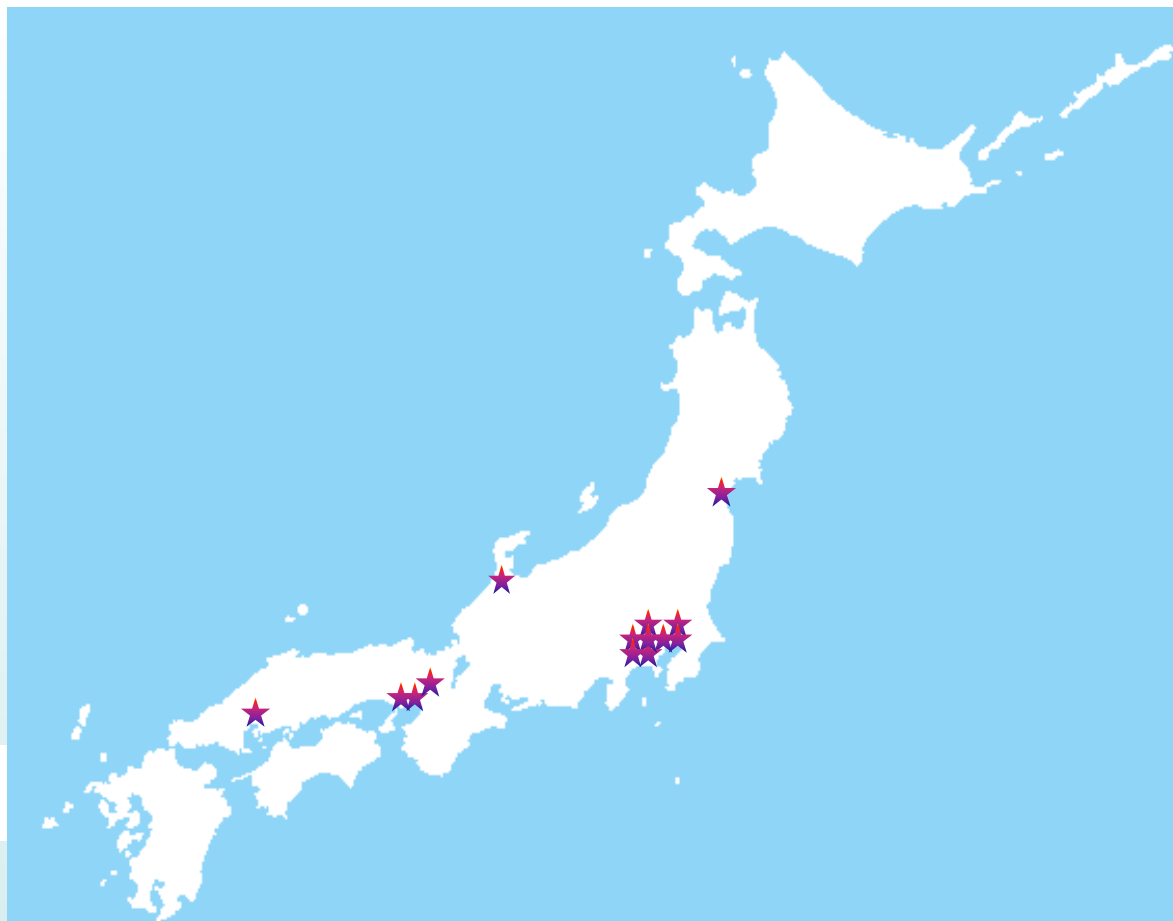
KOBE
UNIVERSITY



広島大学



大学共同利用機関法人
情報・システム研究機構
Research Organization of Information and Systems



13機関：東京大学、大阪大学、京都大学、東北大学、北陸先端大学、東海大学、千葉大学、慶応大学、横浜市大、東工大、広島大学、統計数理研究所、理化学研究所
約200名の研究者（ポスドク約80名）



RIKEN



東北大学



国立大学法人 千葉大学
National University Corporation
Chiba University



東京大学
THE UNIVERSITY OF TOKYO

Keio University



1858
CELESTIS GLADIO FORTIORE



東京工業大学
Tokyo Institute of Technology



TOKAI
UNIVERSITY
EDUCATIONAL SYSTEM



公立大学法人 横浜市立大学
YOKOHAMA CITY UNIVERSITY



Contributors



Workshop in 2009

1st Joint Workshop on Cor
7-8 Jul 2008

Workshop in 2008



ライフサイエンス分野における 計算科学研究開発の世界の現状

1) 世界的な研究の中心は

- 三極：アメリカと欧州（+豪州、ニュージーランド）、日本

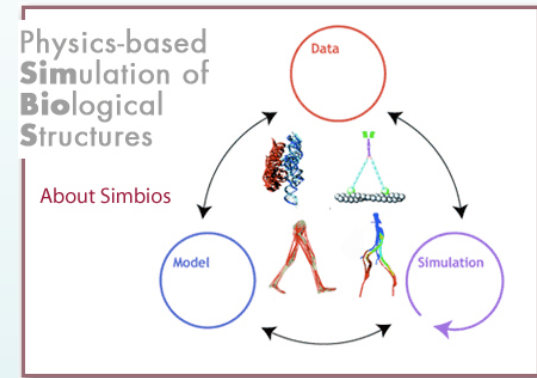
アメリカ: DoE、NIH、DoD、民間企業（ベンチャー、製薬企業）が主、DoEが基礎科学（生物学+計算科学）、NIHが医療を担当するため、互いの協力は希薄。計算科学と医療をつなぐ動きがない。

欧州: EUがスポンサーとなり、IT技術を医療に生かす取り組みを推進。一例: VPH(Virtual Physiological Human)、個別患者の治療が目標。大学・医療機関・製薬企業・医療機器メーカーが参加、マッチングファンド、5年で数百億円規模。脳科学ではスイスがBlueBrainプロジェクトを推進。世界をリード。



2) 海外での研究開発の今後の方向

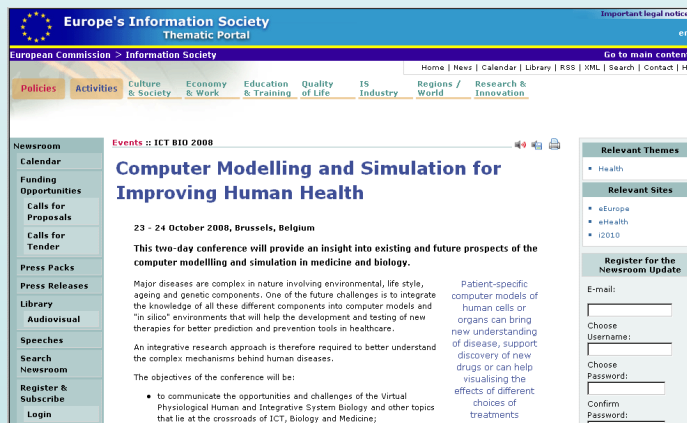
- 世界的に取り組まれているのは
 - タンパクの構造や変化などの解析
 - 大量の実験データに基づく多次元因子解析など
- 現在の研究開発でホットな話題はマルチスケール
 - NIHもEUでも二つ以上のスケールをつなぐ研究開発にファンディング
 - 日本はグランドチャレンジとして既にマルチスケールに挑戦
 - 方程式に忠実にモデル化し、スケールを超える
 - 膨大な計算量は次世代スパコンで解消



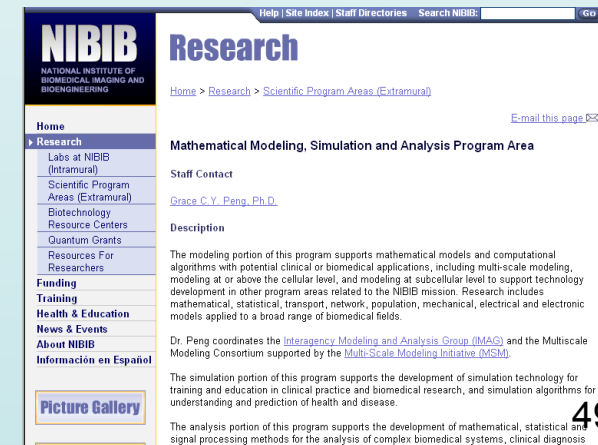
<http://simbios.stanford.edu/>

欧米では医療機関で使うことを前提に簡略化したモデル化。
(脳科学では例外的にスパコンを指向)

戦略分野化で、日本が世界に貢献



http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=3956



<http://www.nibib.nih.gov/Research/ProgramAreas/MathModeling>

3) 国際比較

- 分子スケール

- アメリカは超並列計算で先行
- 粗視化モデルによる計算、全電子を入れたタンパク質の量子化学計算では日本がリード

- 細胞スケール

- システムバイオロジー(E-cell等)は日本発
- 細胞群から臓器へ、分子から細胞へのアプローチはどこもこれから

- 臓器・全身スケール

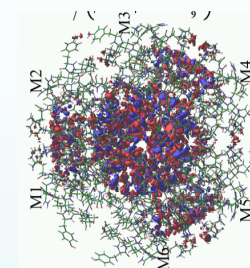
- 国際的な研究開発体制が構築されつつある
- 日本は心臓モデル、健常者からの詳細人体モデルでリード

- バイオインフォマティクス

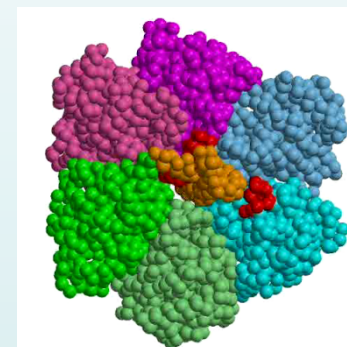
- 個人の多様性と医療をつないでゆくところでは日本は優位
- 次世代超高速シーケンサー等による巨大データに直面、世界的な課題

- 脳科学

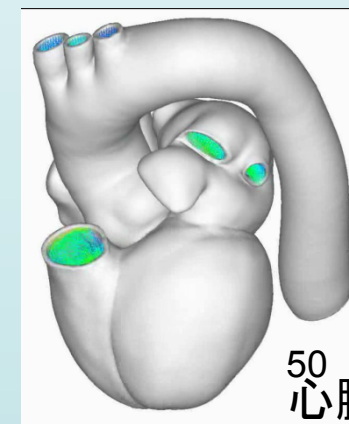
- スーパーコンピュータによるシミュレーションではBlueBrainなどが先行
- しかし、計算論的神経科学では日本が世界をリード、BlueBrainではこの観点が今は抜けている



全電子計算では世界をリード(ProteinDF)



粗視化モデルによる分子モータ



50
心臓

3) 国際比較(続き)

数万並列規模の計算能力をもったソフトウェア開発を推進

今後、次世代スパコンに向けた並列性能向上を加速

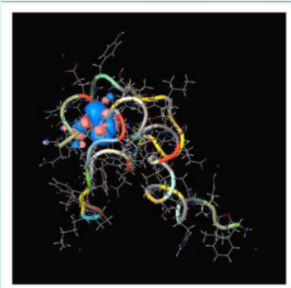
⇒分子・細胞・臓器全身の各スケールと実験データ解析、脳神経系を統合して研究開発

★世界的にも全くない試み

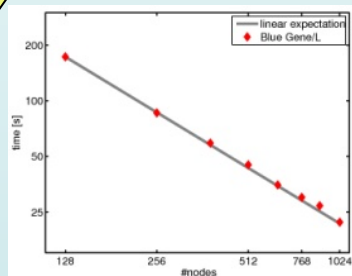
同時に行うことにより、手法やソフトを互いに利用し、開発を加速

⇒次世代スパコンで世界をリードする

↓ チーム間の連携



全電子計算での並列性能 (ProteinDF) 2500コアまでの並列を確認



脳の局所回路シミュレーションでの並列性能 (BlueGeneLで1000並列までの速度向上を確認)

受提供	データ解析融合	分子スケール	細胞スケール	臓器全身スケール	脳神経系
データ解析融合	肺ガンと薬	ターゲットタンパクの特定	生体分子ネットワークモデル	画像処理パラメータ推定	パラメータ推定(モデル化)
分子スケール	タンパク-リガンド	代謝・膜	膜での輸送モデル	血栓形成	神経可塑性・変性・イオンチャンネル
細胞スケール	生体分子ネットワークデータ	代謝反応・膜での物質輸送データ	肝細胞	複数細胞から臓器へ・血栓形成	細胞プラットフォーム供給
臓器全身スケール	医療画像データ	データ提供・ドラッグデリバリー	血流・分泌物輸送	循環器系	全身モデル
脳神経系	神経細胞データ・局所回路データ	可塑性や神経変性に関するデータ	成長・変形モデル	運動制御臓器制御	51 細胞から高次機能までの一貫

