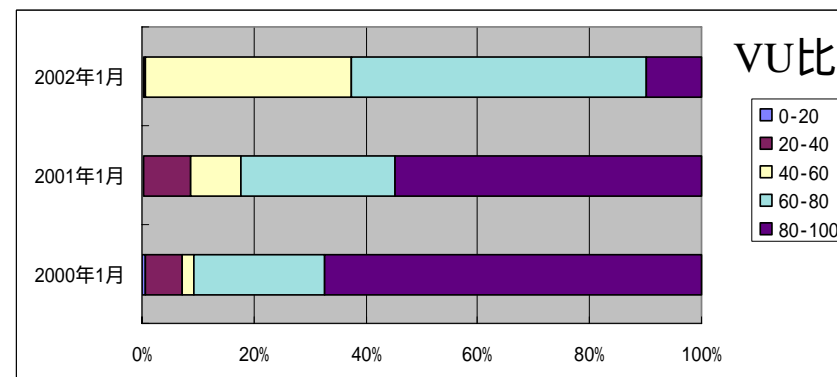
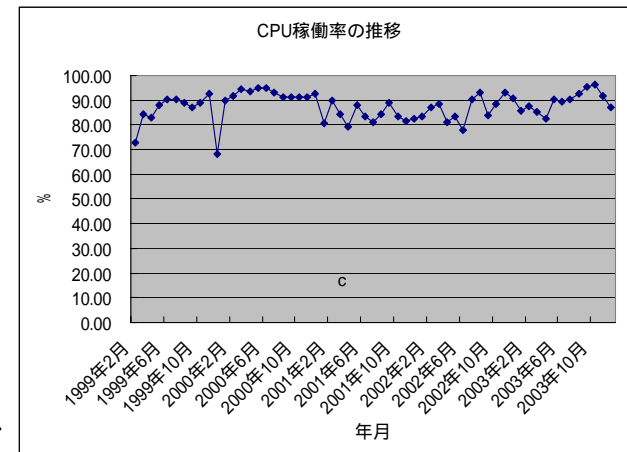


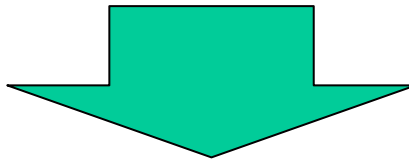
現行スーパーコンピュータシステム構築の背景

- それまでのスパコン
 - Fujitsu VPP700E/160
 - ベクトル型並列計算機(1999年導入)
 - 計算性能、メモリ容量などPCなみ
 - プログラムがベクトル化されていないと、性能はPC以下
- 稼働率は90%前後で、ジョブは常に満杯
- 主なユーザーは、計算科学分野
 - 計算化学、計算流体力学、分子動力学
- 新規のユーザーにとってベクトル化は困難
 - スカラー向きソフトの利用が増加
 - フリーソフトの利用希望の増加
- データベース検索・実験データ処理のニーズ
- 並列化できないプログラムの存在
 - 1CPUで大きなメモリが必要



新システム設計での基本的な考え方

- コスト性能比の良い計算機
- できるだけ多くのフリーソフトが利用できる環境
- 1CPUで大容量のメモリが利用できる環境
- 新規分野の開拓
 - 高エネルギー物理(実験データ処理)
 - ライフサイエンス
- 旧来のユーザーも利用できる環境
- 他の計算機センターとの連携



- 異なるサブシステムから構成される複合型システム
- Linux PCを高速ネットワークで接続したLinuxクラスタ
 - 大容量メモリ計算機(SX7)
 - グリッドコンピューティング技術を使ったジョブ制御

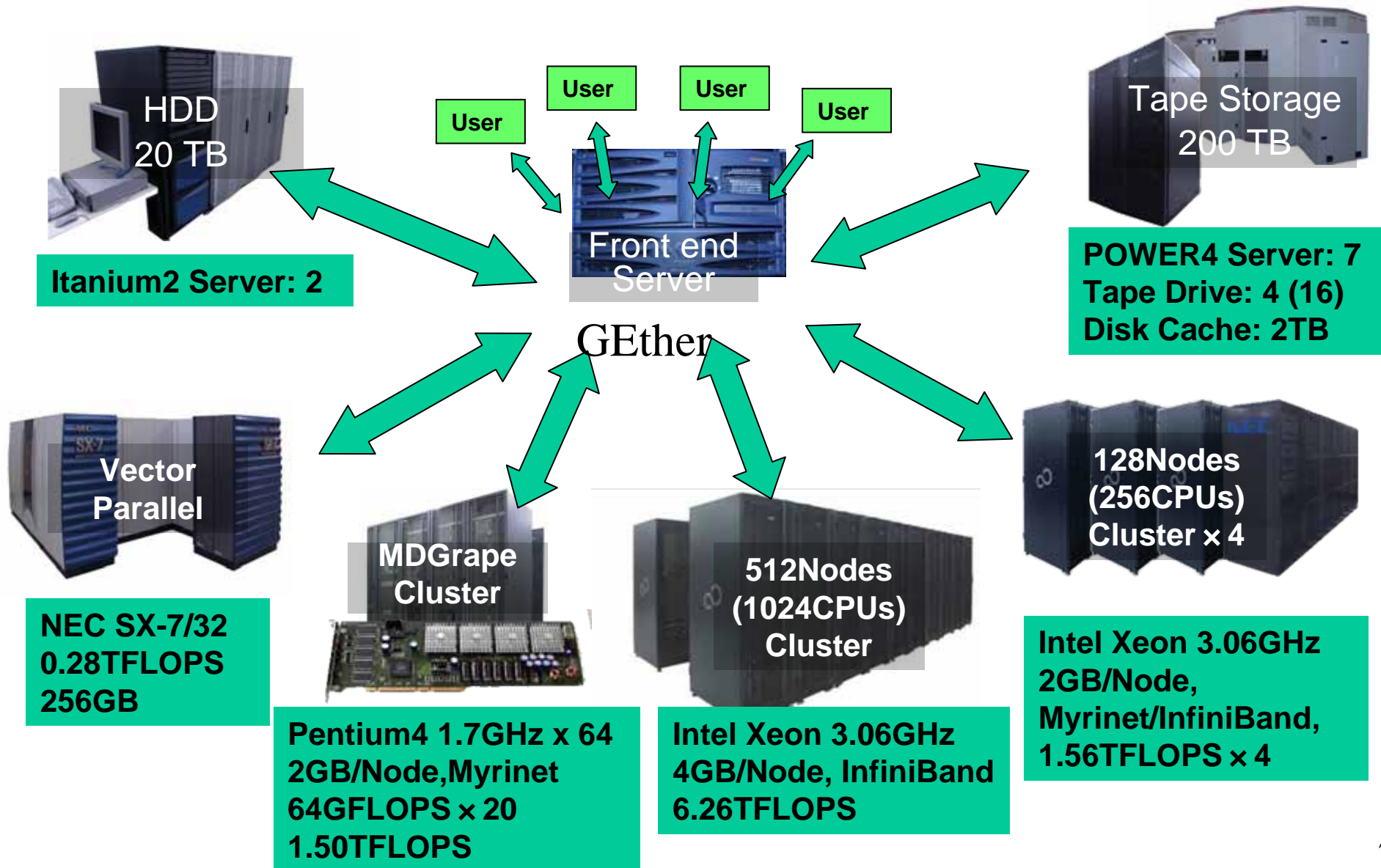
RSCCの特徴(1)

- 用途ごとに異なるサブシステムで構成
 - Linuxクラスタは6つに分割
 - 高並列のジョブ実行(通常128CPU並列, 最大1024CPU並列)
 - 長時間のジョブ実行(24時間から最長1週間)
 - 占有利用も可能(現在は、高エネ物理実験データ解析、ゲノムDB用)
 - SX-7
 - 巨大なメモリを必要とするプログラム(Linuxでは最大2GB)
 - ベクトルチューニングされたプログラム
 - VPPからLinuxクラスタへの中継的役割
- 複合システムを意識させないために
 - GRID技術の導入による計算資源の仮想化
 - ITBLプロジェクトにおける計算リソースとして利用
 - 利用者には、どのサブ・システムを使うかは明示させない
 - データはプログラム実行時にローカルHDDへコピー
 - HPCポータル: Webによるスパコンの利用

RSCCの特徴(2)

- LinuxクラスタでVPP Fortranをそのまま実行
Linux用XP Fortranの導入
- ライフサイエンス分野
常に最新のDBをLinuxクラスタのローカルHDDに配置するために、
256CPUを占有：バイオ関連DBミラーサービスと連携
Bioポータルからのジョブ実行
- 高エネルギー物理実験データ解析
高速テープライブラリ装置(HPSS)を導入
巨大な解析結果をテープライブラリ装置に直接格納
- ITBLプロジェクト
 - 計算機センターアライアンス
 - 計算資源の相互活用

RIKEN Super Combined Cluster



理論性能

	先代 (VPP700E/160PE)	現在 (Linux cluster + SX7)	比
理論ピーク性能	384 GFLOPS	Linux cluster 1 : 6.2TFLOPS	16
		Linux cluster 2 : 6.2TFLOPS	16
		MD cluster : 1.7TFLOPS	4.4
		SX-7 : 0.28GFLOPS	0.73
		合計16.1TFLOPS	41.9
記憶容量	320 GB (160PE)	PC cluster 1 : 2TB	10
		PC cluster 2 : 1TB	
		SX-7 : 256GB	
ディスク容量	4.4 TB	20TB	4.5
テープ容量	100 TB	200TB	2

LinuxクラスタでのLINPACK性能 : 8.73TFLOPS (70.4%)

世界で7位、日本で2位(2004.6)

ライフサイエンス分野

- バイオ関連DBミラーサービス

- バイオ関連のDBは散らばっている

- ユーザは各DBに様々なサーバを渡り歩く
- 理研内サーバに収集 ダウンロード元の一元化
- 24時間以内に最新DBを収集・公開

- DBミラーの効果

- ダウンロード作業の短縮化
- DB情報管理をリスト一覧でWeb上に公開

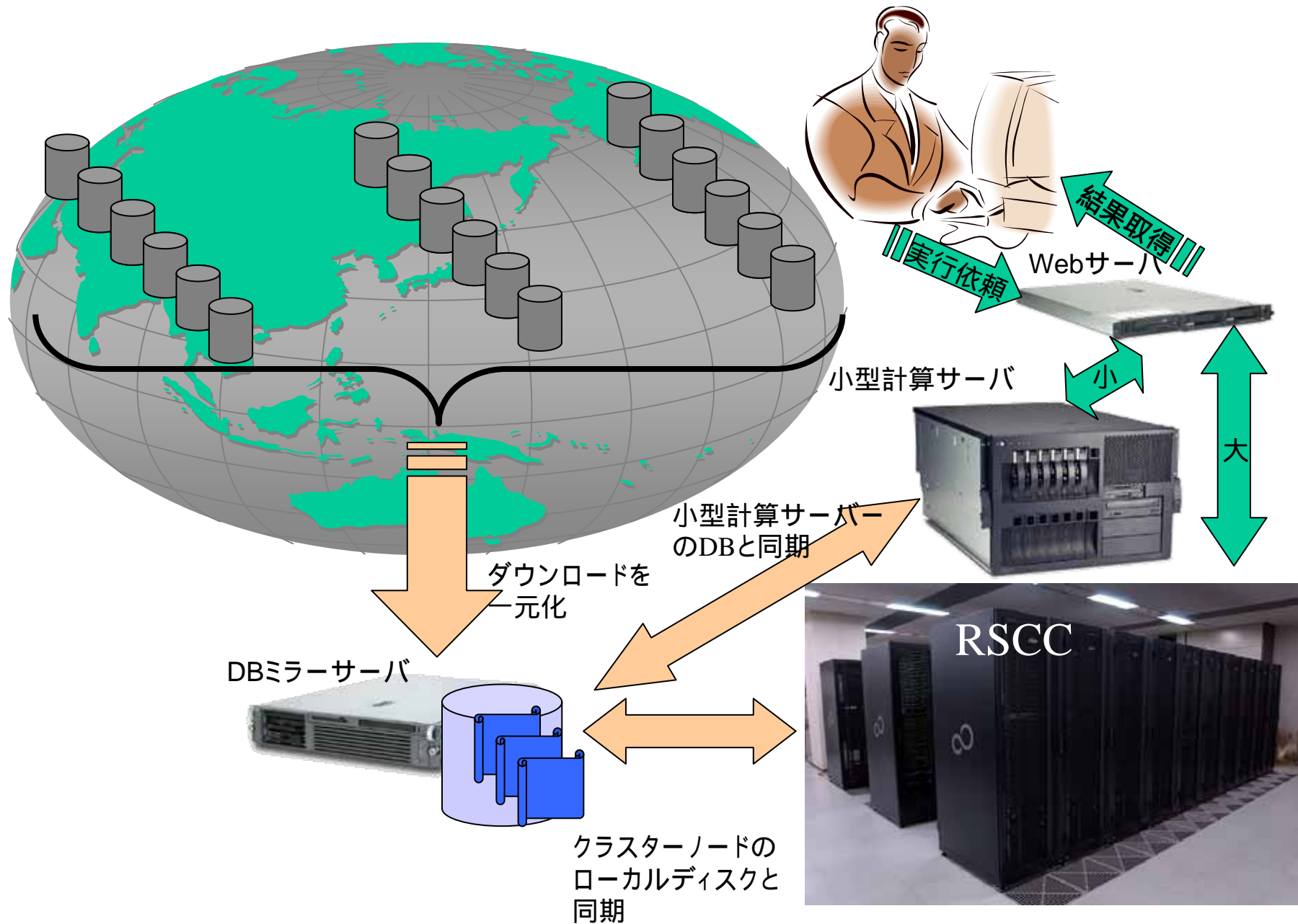
- バイオポータル

- 実験系研究者のためのDB検索サービス

- ウェブを使ったインターフェース
- 計算量小→ローカル処理
- 計算量大→RSCCで処理
- DBはバイオDBミラーと同期

ユーザーは使っている
計算機を意識しない

バイオミラーDBとバイオポータル・RSCCの関係



高エネルギー物理分野 大規模加速器実験におけるデータ処理及びシミュレーション



クラスタから高速データアクセス (HPSSシステム)



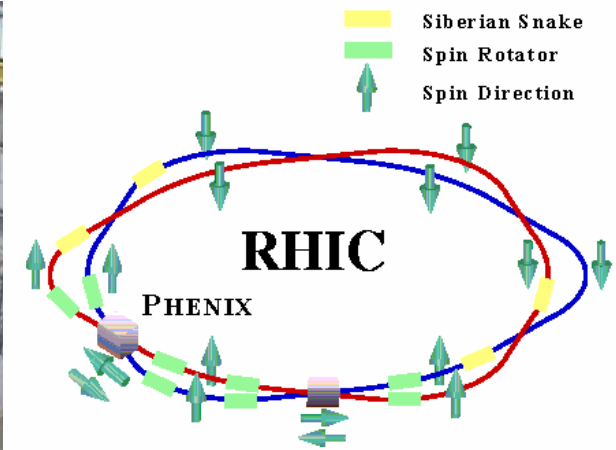
データ解析・シミュレーション
RSCC



測定器

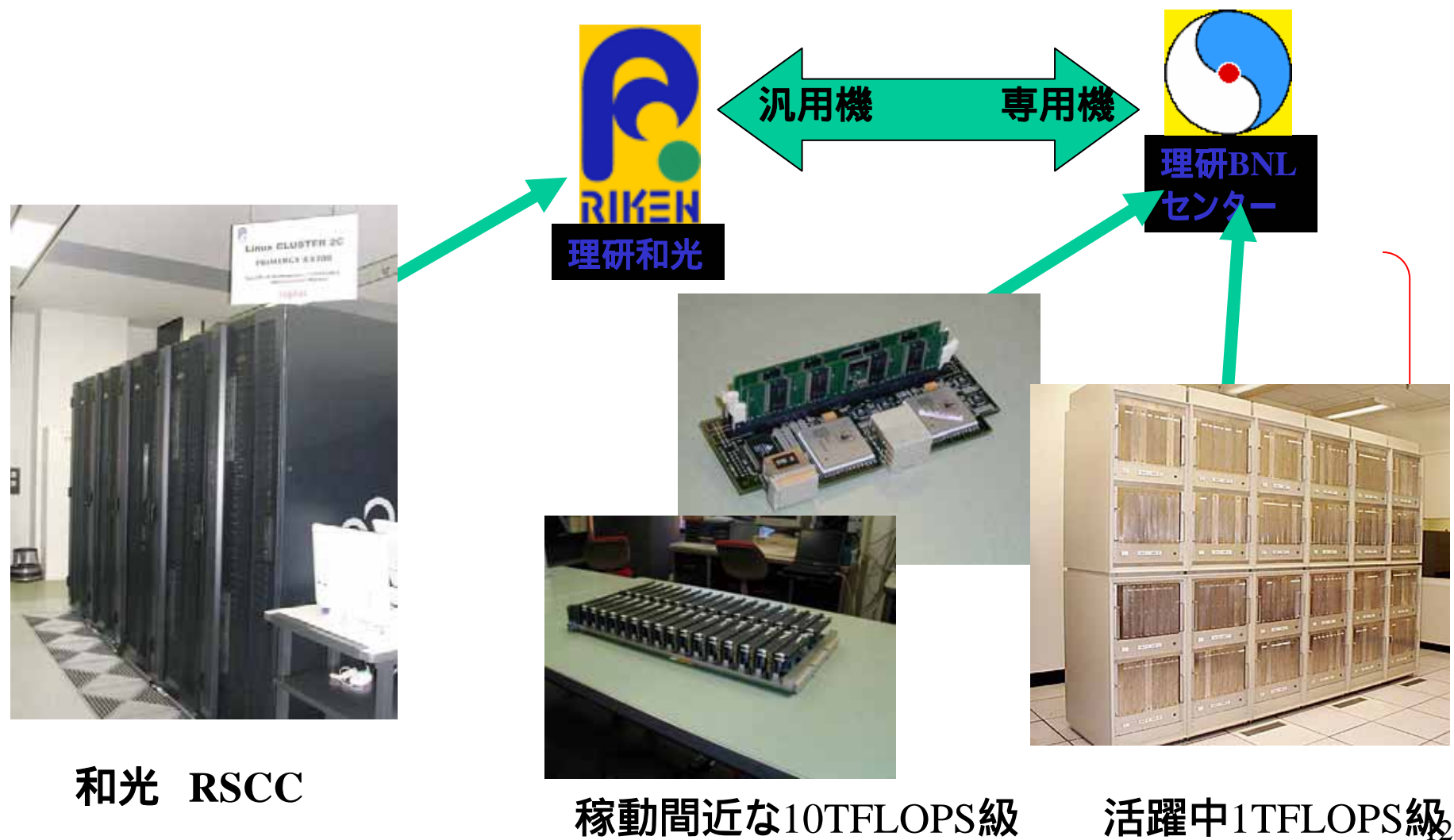


データ



加速器

高エネルギー物理分野 格子QCDシミュレーション



和光 RSCC

稼動間近な10TFLOPS級

活躍中1TFLOPS級

現行スーパーコンピュータを使った研究成果

- 正式稼働2004年5月のため、今のところ顕著な成果は出ていない
- ここ数年の代表的なコンピュータを使った研究と成果

SARSウイルス増殖抑制物質の発見 (in Silicoスクリーニング)

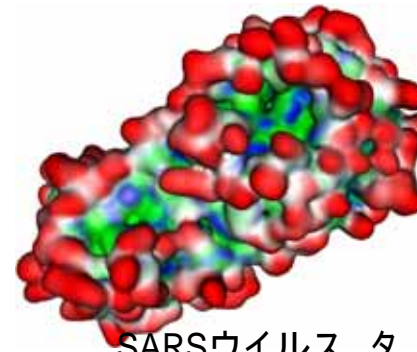
たんぱく3000:NMR、X線実験構造解析
中間質量ブラックホール形成のメカニズムを解明

グリッド上に細胞シミュレーター「E-Cell2D」を開発し、公開

ー 生体力学シミュレーション研究プロジェクト: 血流シミュレーション、インプラント・シミュレーション

ー ものづくり情報技術統合化プロジェクト: V-CADによる大規模構造解析、熱流体解析

ー 新発見鉱物 $MgSiO_3$ ポストペロブスカイト相の地震波伝播特性を初めて第一原理計算で解明し、地震波観測との比較からD層主要成分がこの新発見鉱物であるとする根拠を与えた



SARSウイルス タンパク質分解酵素



リン酸マンノース転移酵素の立体構造

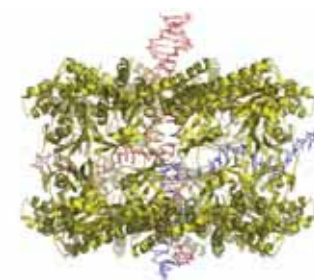
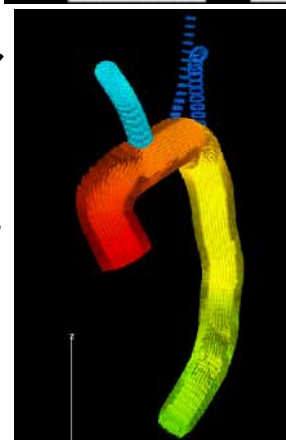
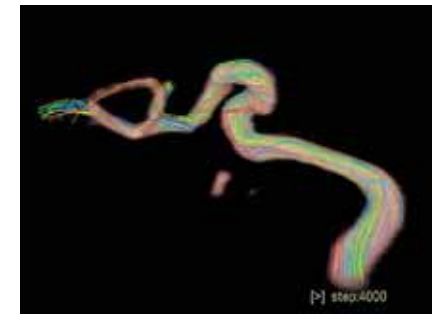
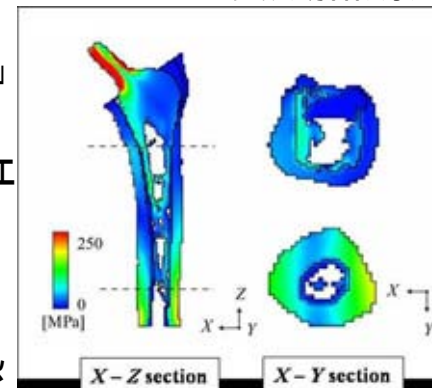
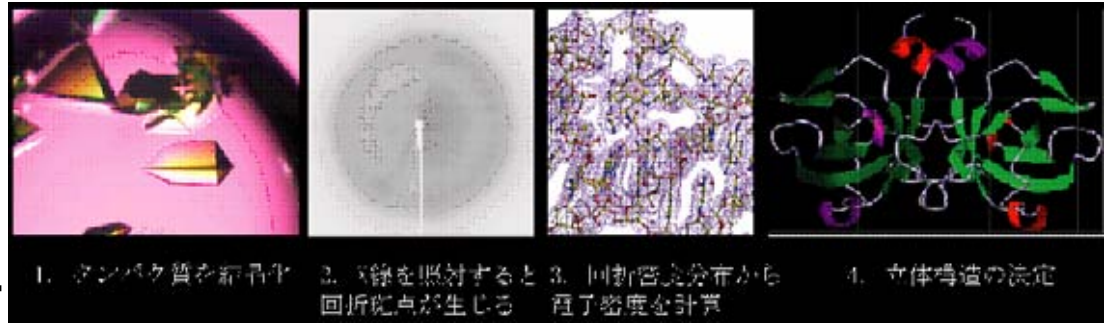


図1. 単鎖DNA-二重鎖DNA-Dmc1からなる三者複合体モデル

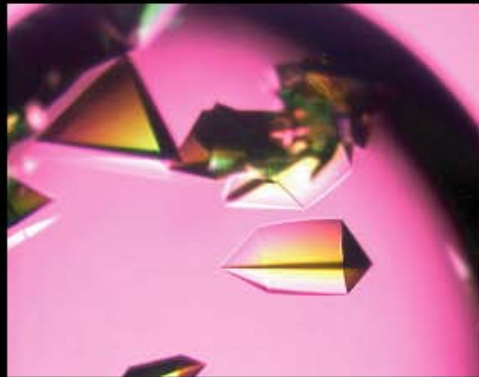
Dmc1ダブルリング(黄色)の中心の穴から向かって、上から二重鎖DNA(赤色)が通っている右のリング-リング間の隙間の穴から単鎖DNA(青色)が入ってきて、リングの中心で相同性対合反応が起こり、下の穴から組換え反応が終わった二重鎖DNAが出てくる。さらに組換わらなかった二重鎖DNAの反応側の鎖は、左のリング間の隙間から出る。

例1 タンパク質構造解析

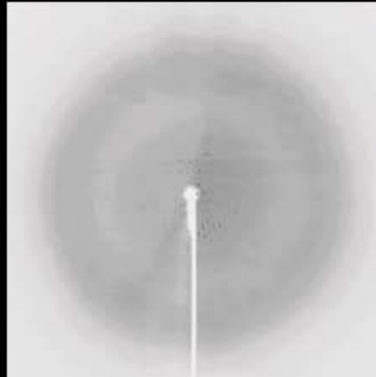
- 世界一のタンパク質構造解析ラインの確立
- タンパク3000プロジェクト
 - NMRとSPring-8を駆使して2500種を担当
 - 平成15年度末で600余を完了
- X線
 - タンパク質を結晶化し、X線回折像で解析
- NMR
 - 水溶液でOK、NMRを使って解析



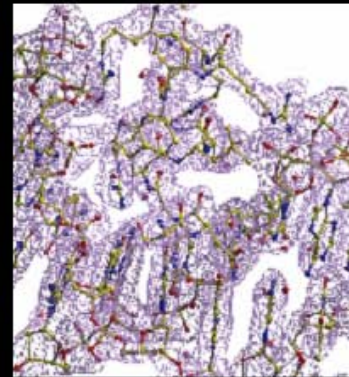
X線結晶構造回折による立体構造解析



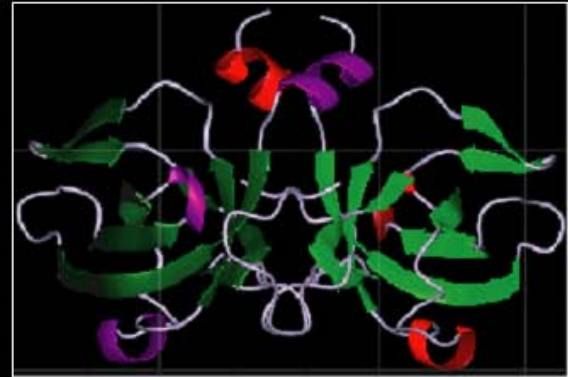
1. タンパク質を結晶化



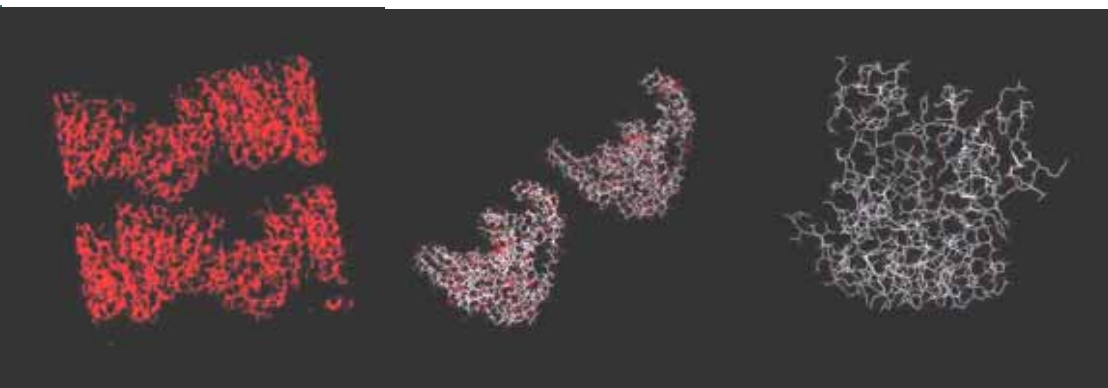
2. X線を照射すると回折斑点が生じる



3. 回折密度分布から電子密度を計算



4. 立体構造の決定



ピークサーチ

トレーシング

側鎖のアサイン

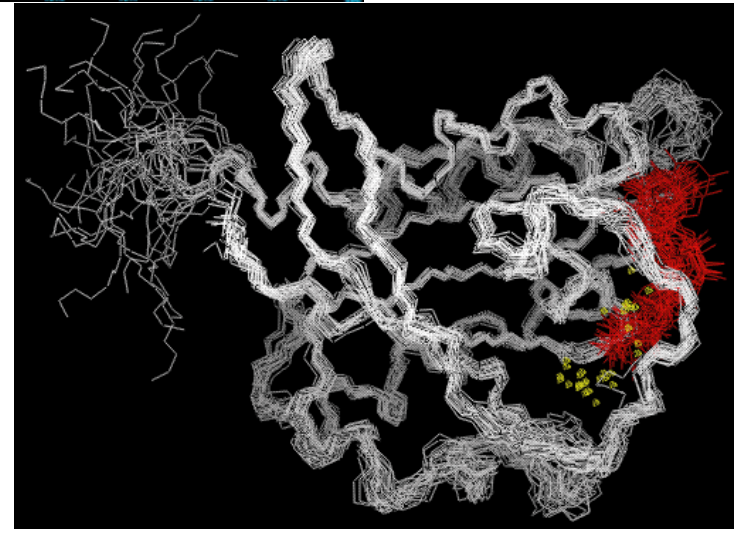
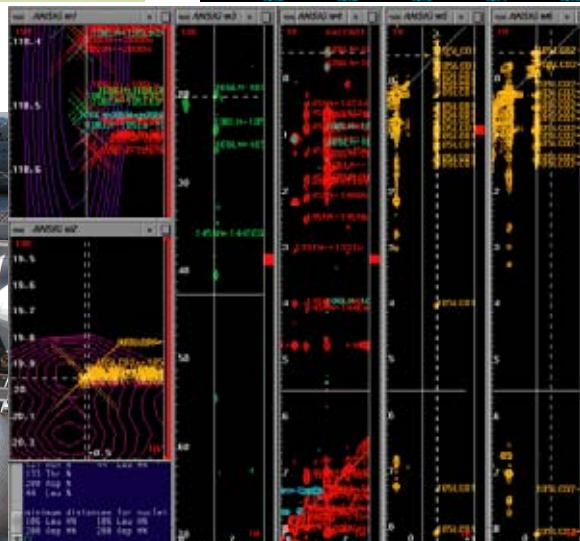
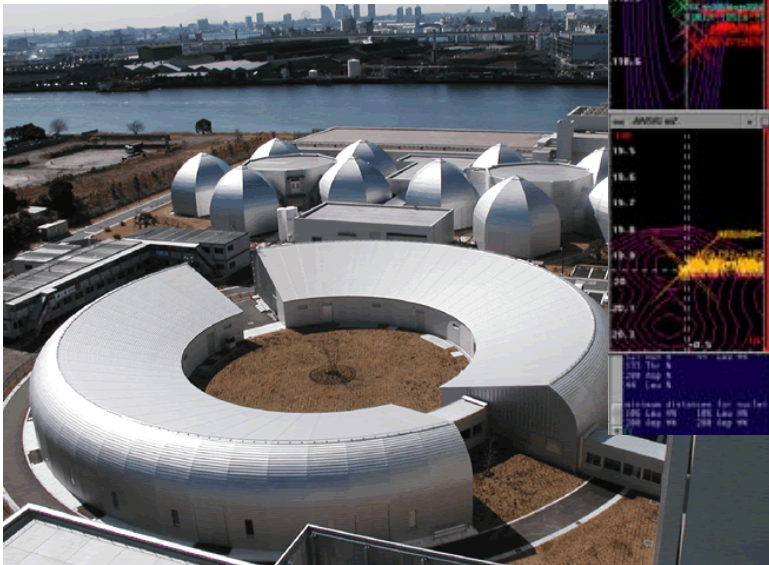
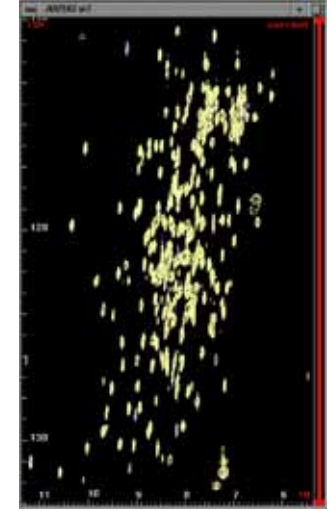
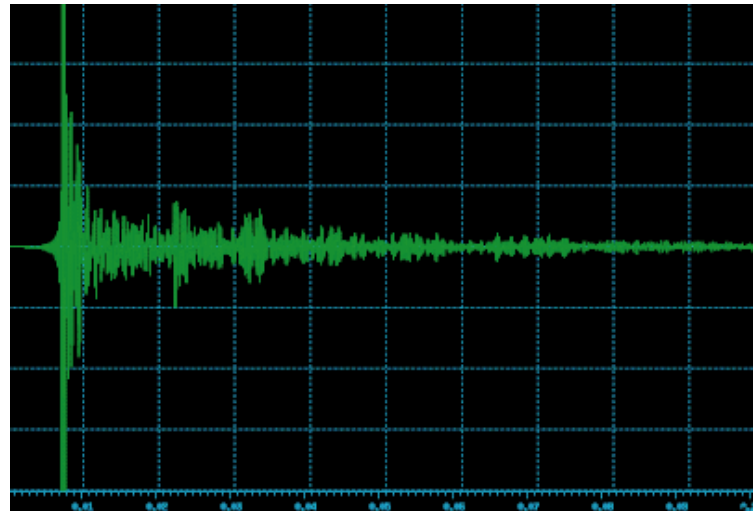
理研・播磨研究所

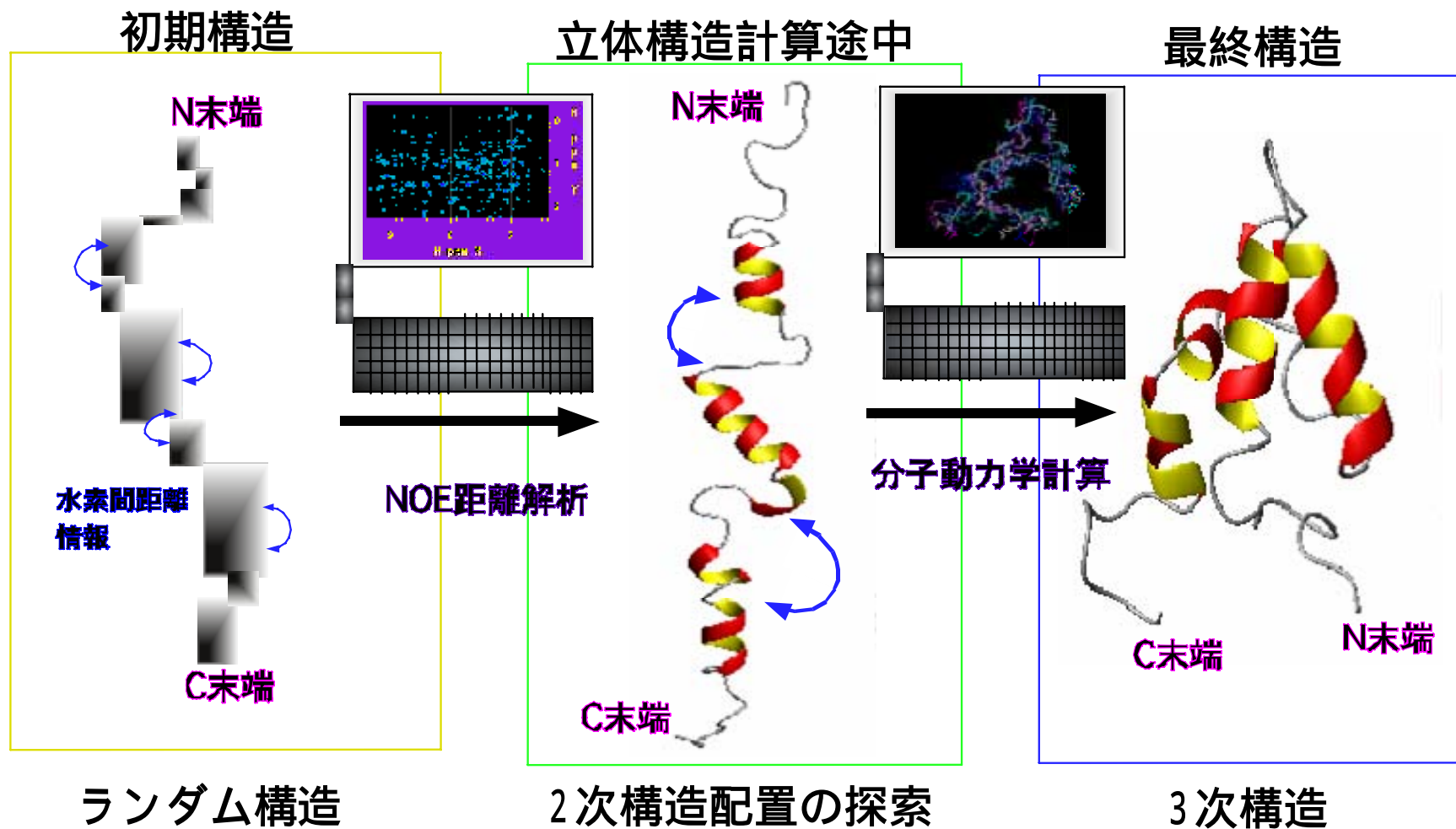
NMRによるタンパク質構造解析

測定手順



タンパク質
の試料

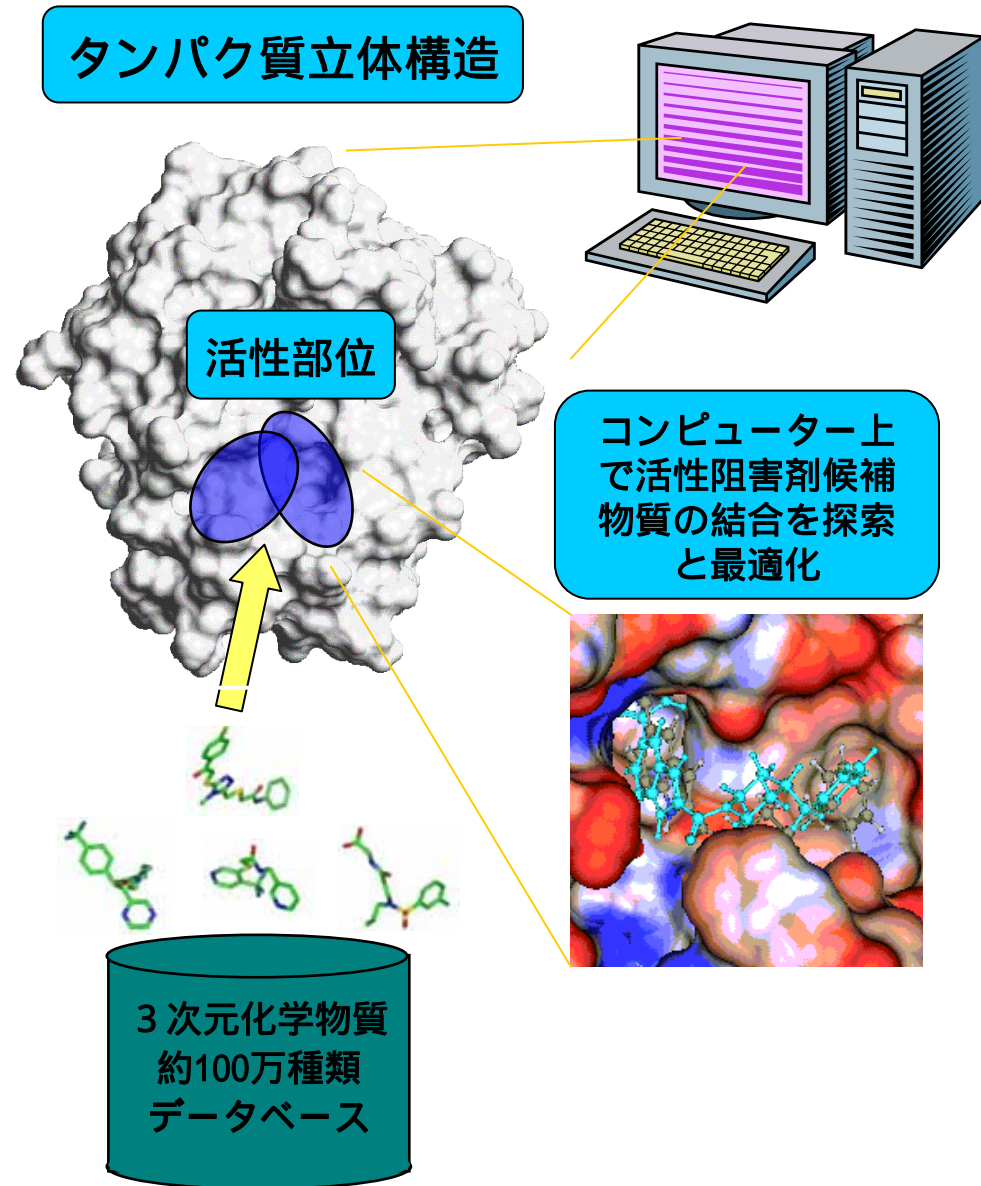




NOE 相関スペクトルから得られた水素間距離情報を満たすタンパク質の立体構造を、制限付き分子動力学計算により決定します。

註:NOE:Nuclear Overhauser Effect

例2 タンパク質立体構造に基づく合理的薬剤設計(SBDD)



疾病の多くはタンパク質の機能，
あるいはその機能のネットワークが
うまく機能しなくなることによる。

タンパク質の機能は，それと相互作用
する化学物質により制御が可能
タンパク質の活性を阻害（制御）する
化学物質（タンパク質の活性部位に
ぴったり当てはまる形）を
コンピューター上で探索。

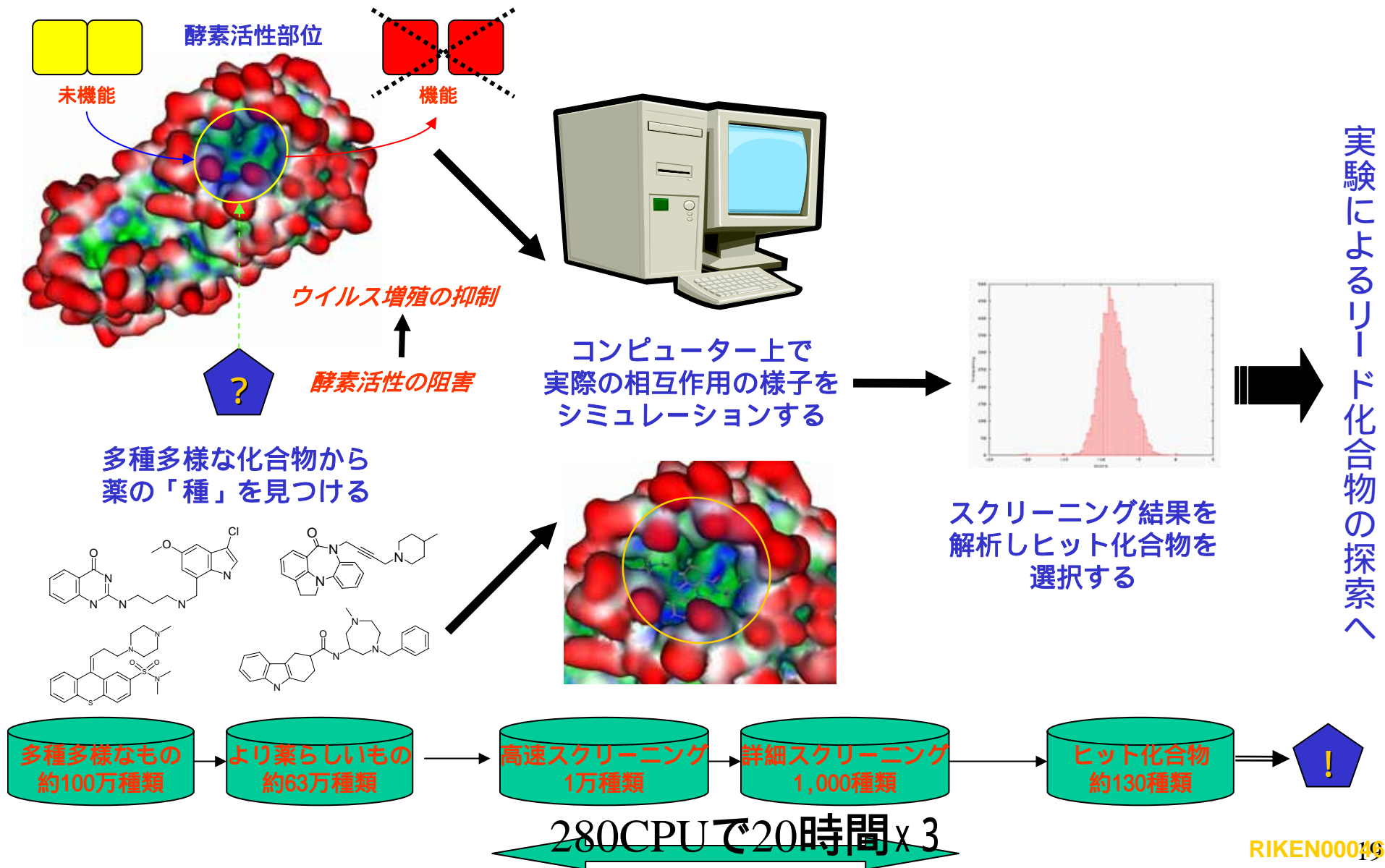
タンパク質の立体構造情報を用いる
ことで，新規薬剤候補化学物質を
効率的 / 効果的に探索，
または化学物質を最適化するよう
設計が可能になる



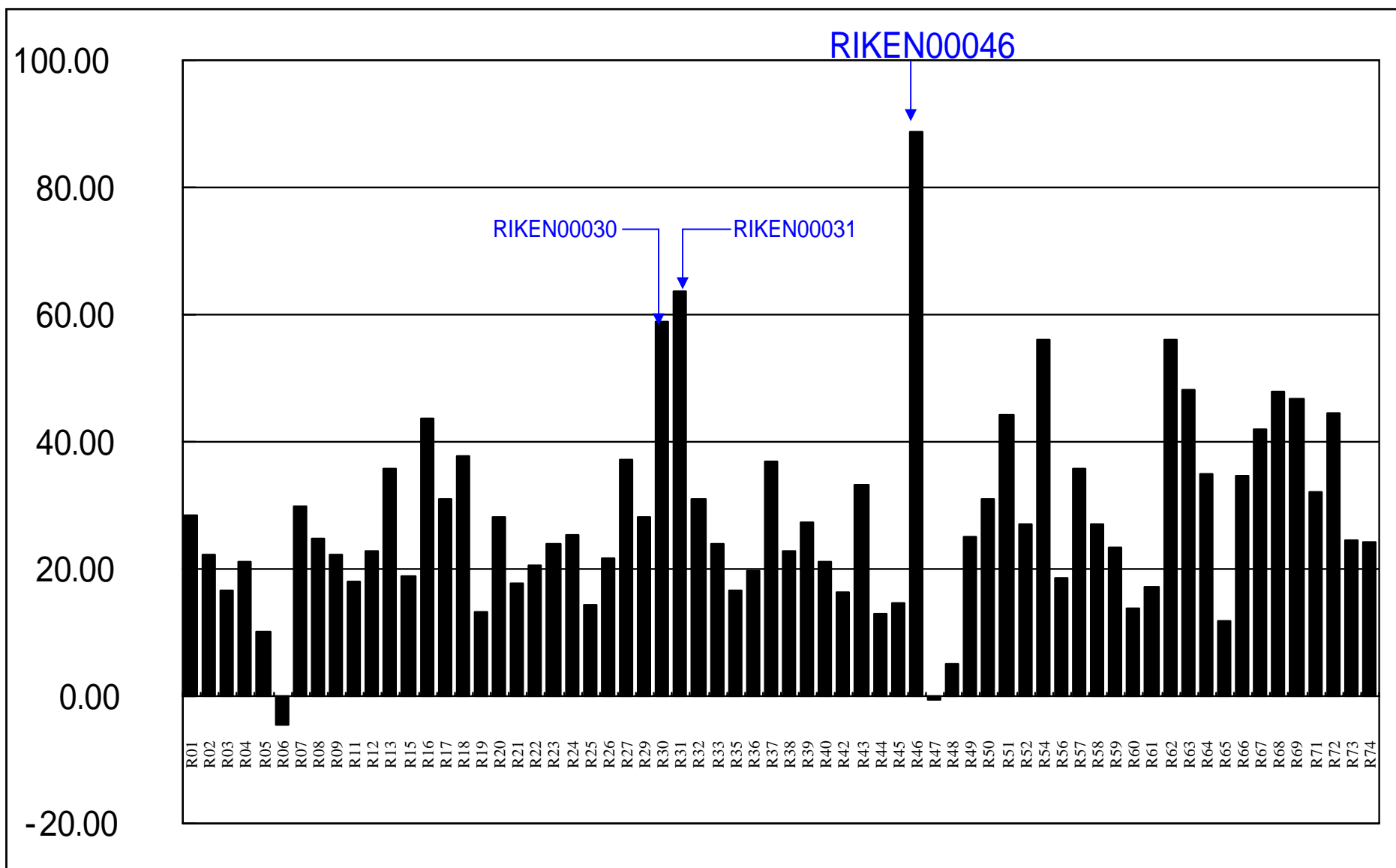
新規薬剤開発のスピード化，ローコスト化。

In silico スクリーニングによるヒット化合物の探索

標的タンパク質： SARSウイルス タンパク質分解酵素（ウイルスの増殖に関与）

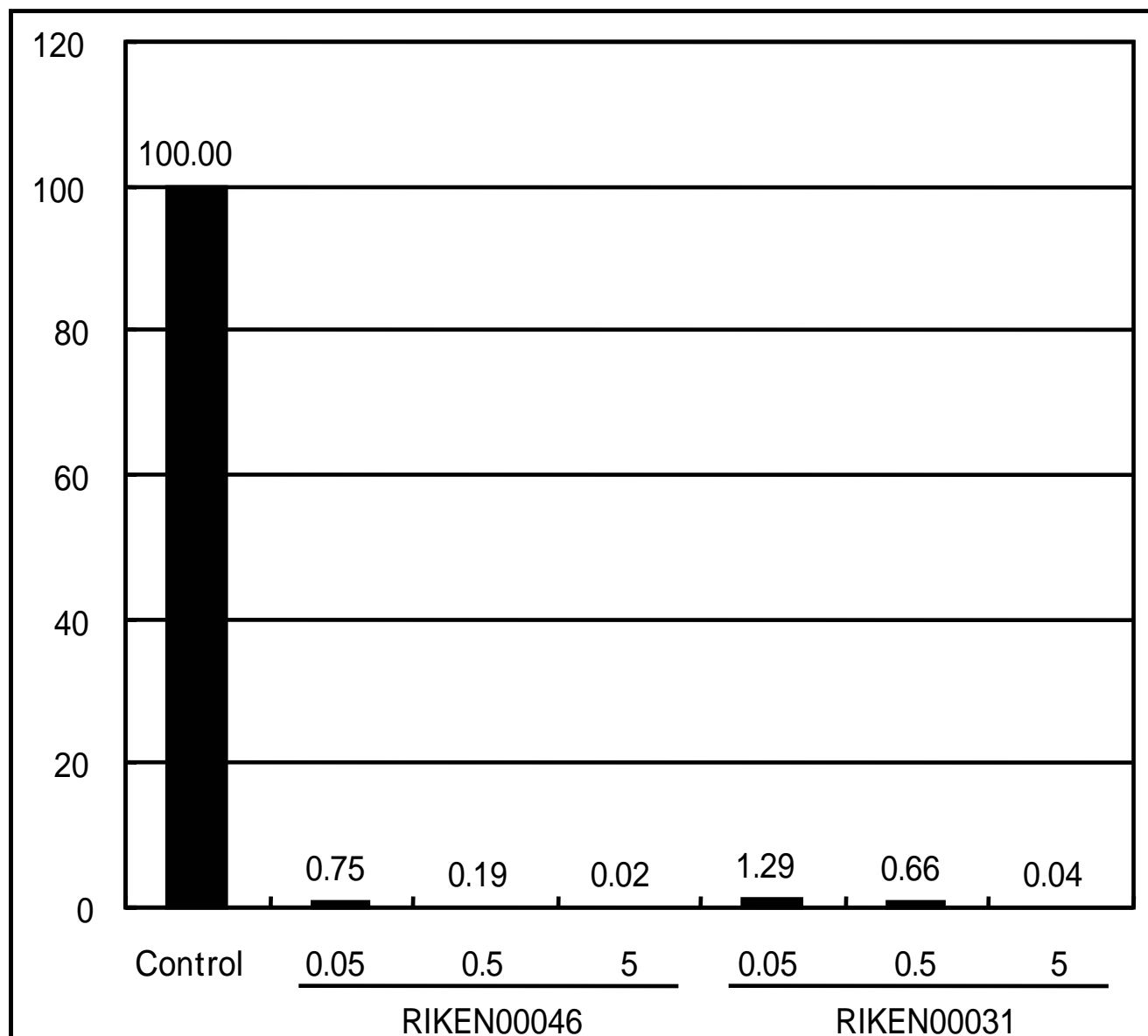


化合物によるウイルス感染細胞変性 (CPE) 抑制効果の比較



* ウイルスが感染していない細胞で、化合物を投与しなかった時の細胞の活性を100%とする

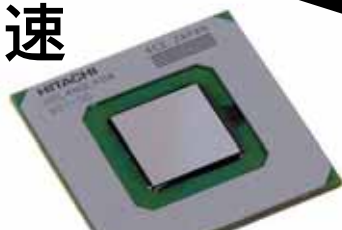
上位2化合物によるSARSウイルス増殖阻害効果の比較



例3 MDGRAPE-3の開発

タンパク質構造

X線結晶構造解析
の加速



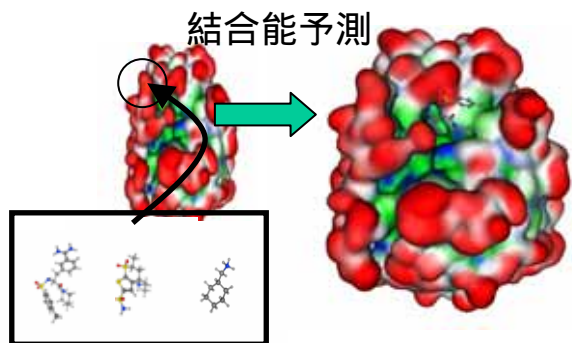
世界最速
230GFLOPS/chip

フォールディング

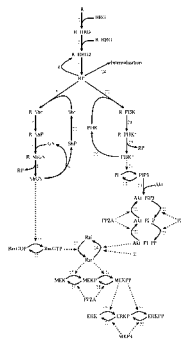


ペタフロップス計算機による分子シミュレーション

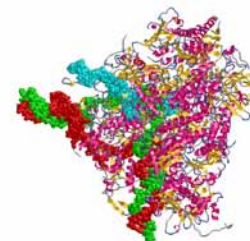
創薬



生体分子
ネットワーク

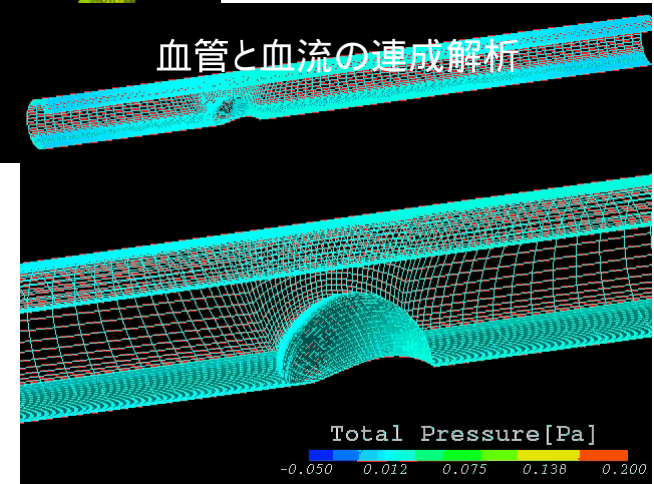
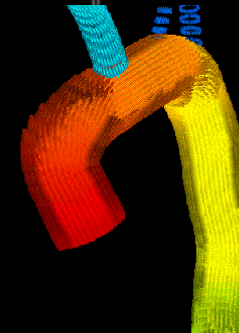
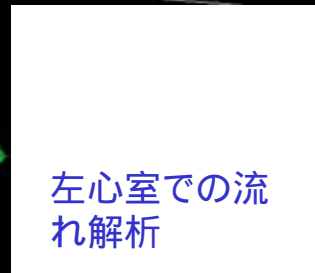
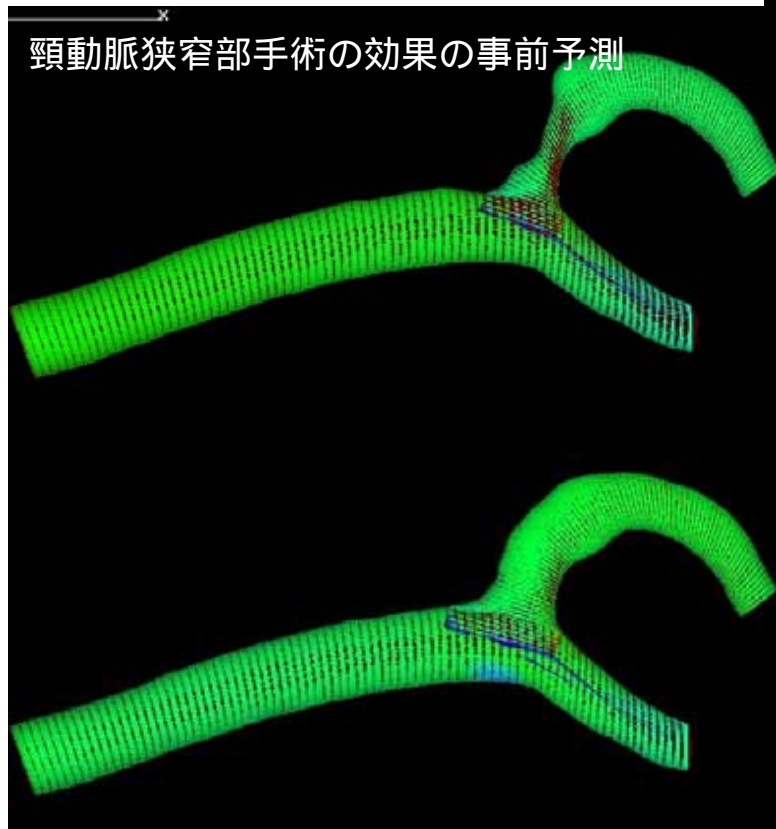
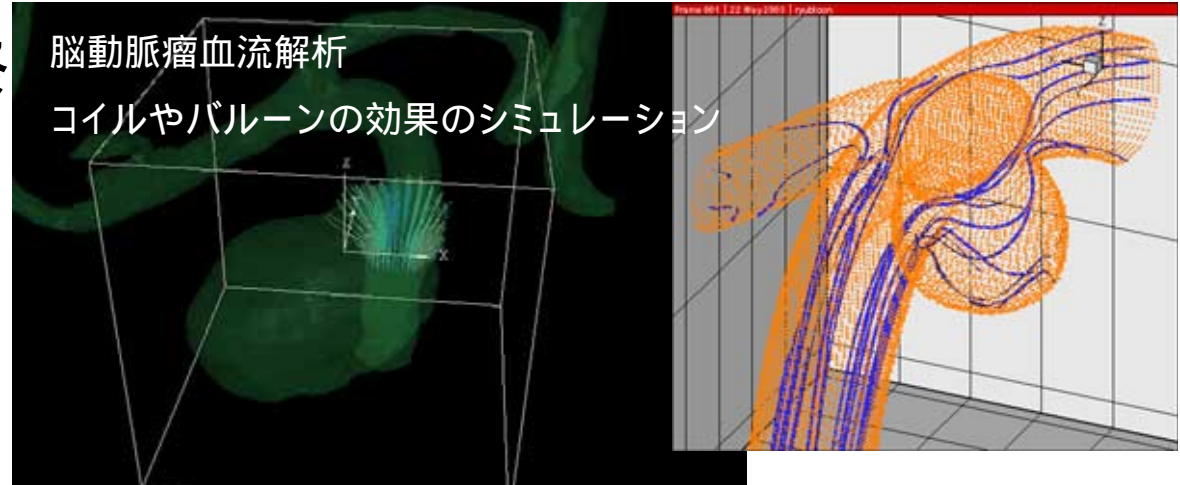


機能解析
発病メカニズム
バイオ・ナノテク

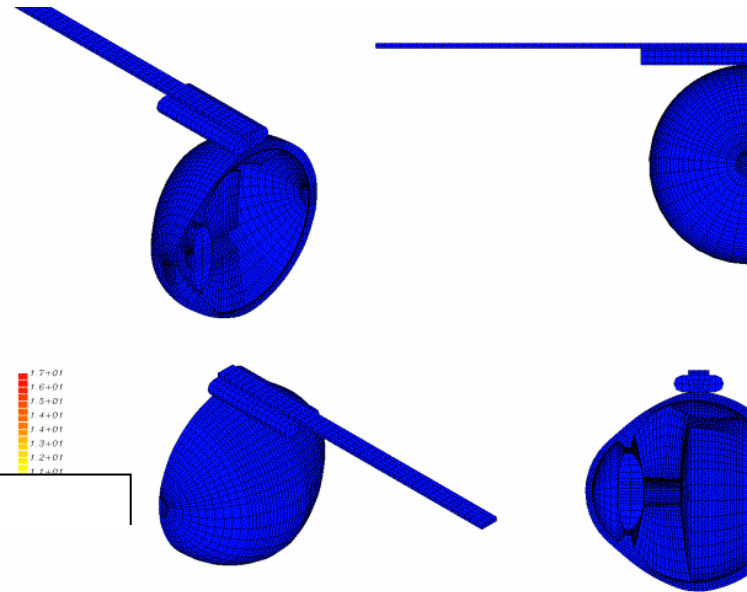
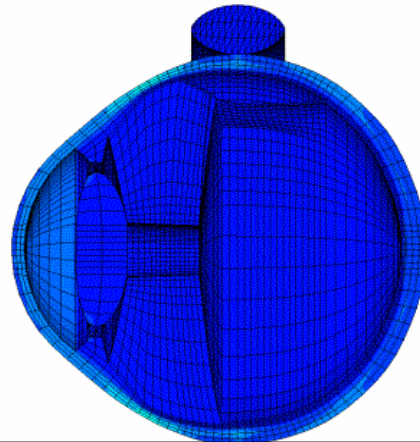
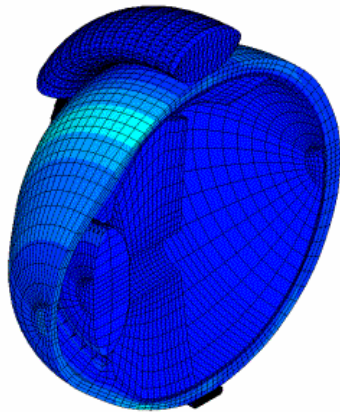


例4 生体力学シミュレーション: 誰でも名医

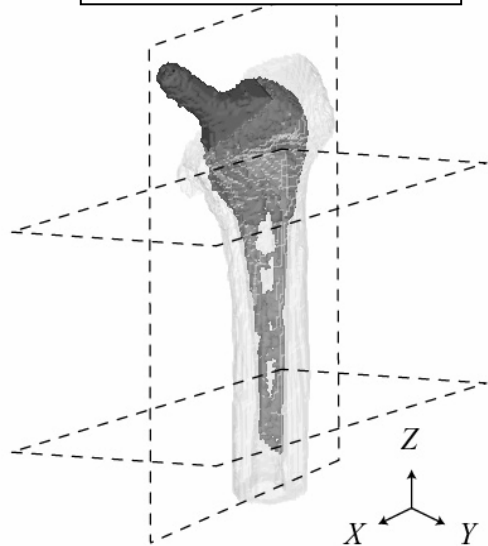
- 第一期2004年3月で終了、現在第二期
- マクロな立場からのアプローチ



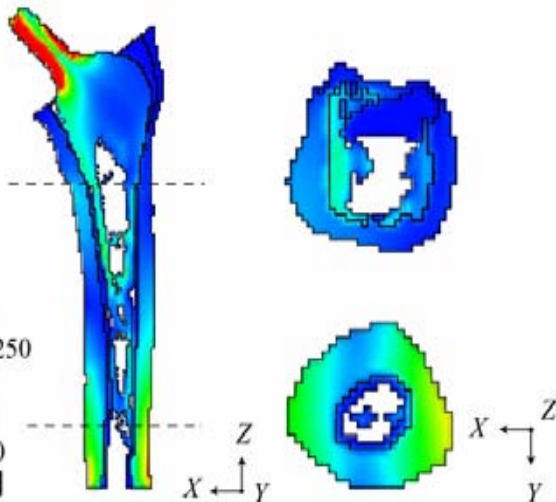
眼球の網膜剥離手術と人工関節のシミュレーション



最適人工関節の選択



Stem Shape



X-Z section

X-Y section

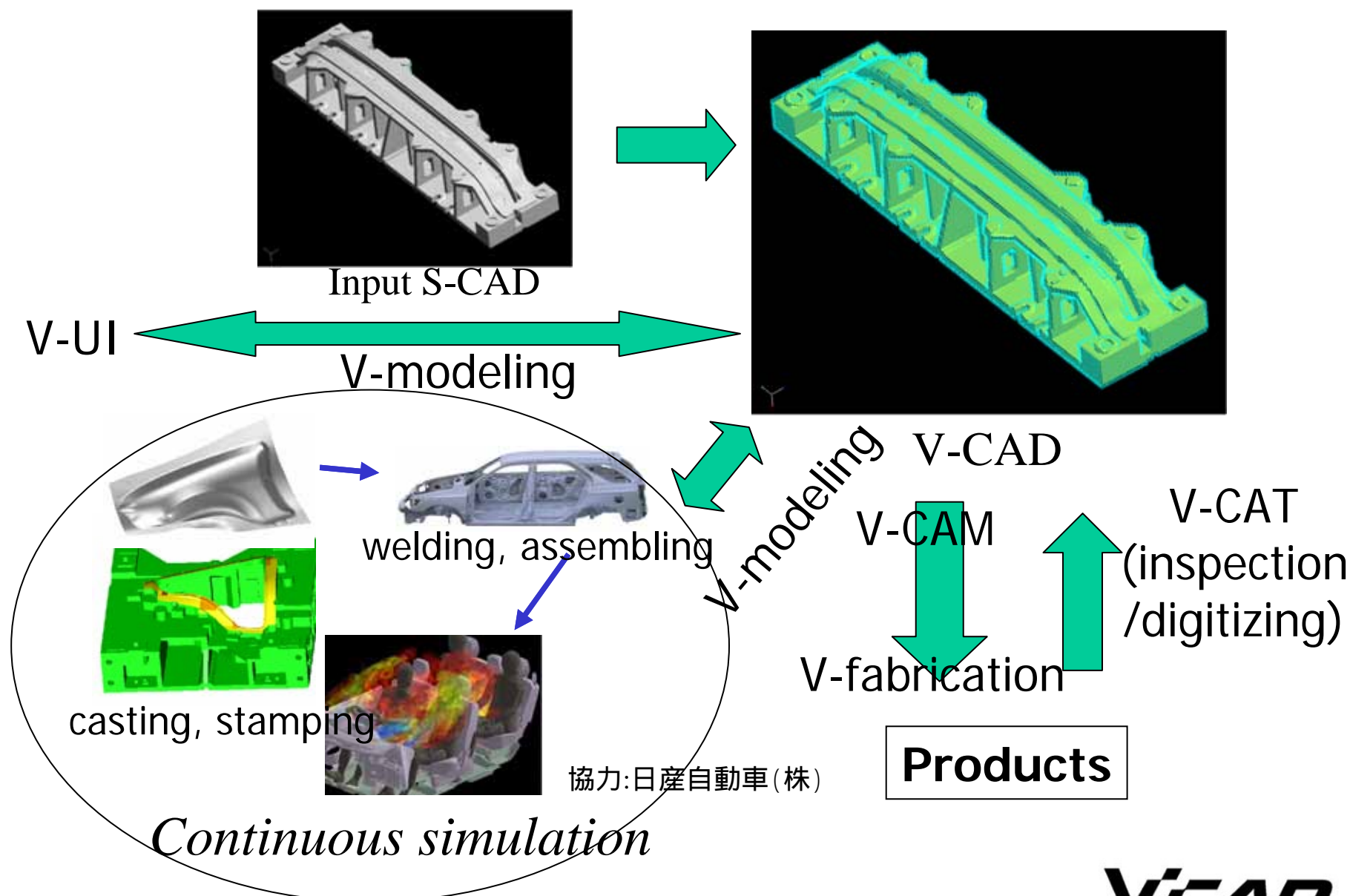
Equivalent Stress

最適手術方法の検討



カテーテル手術シミュレータ

例5 ものづくり情報技術統合化プロジェクト

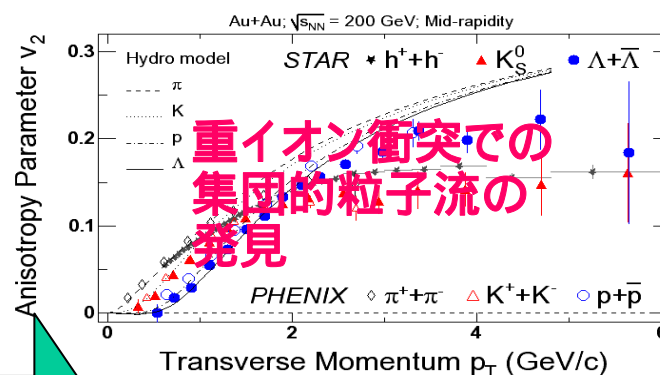


例6 クラスタ計算機を利用した成果
高エネルギー原子核実験

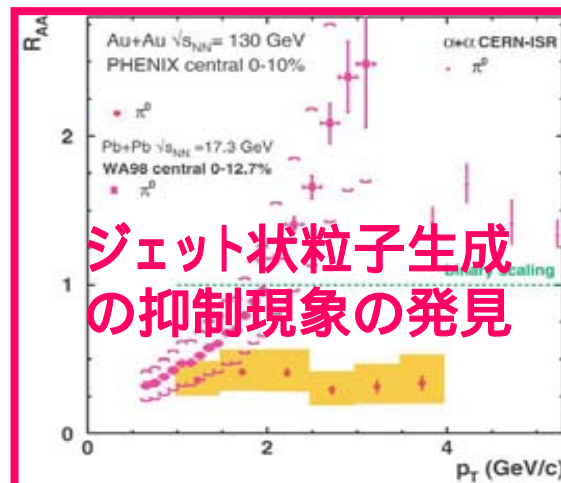
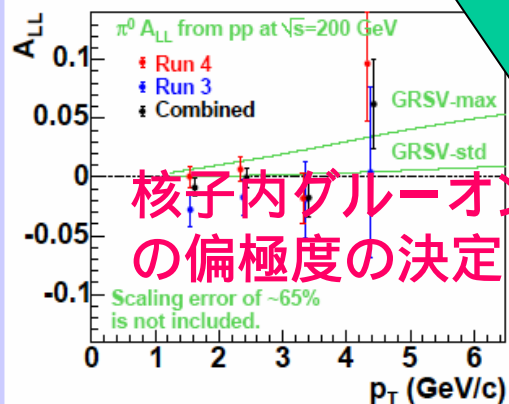


大量の実験
データ

核物質中のメソンの
質量変化の発見



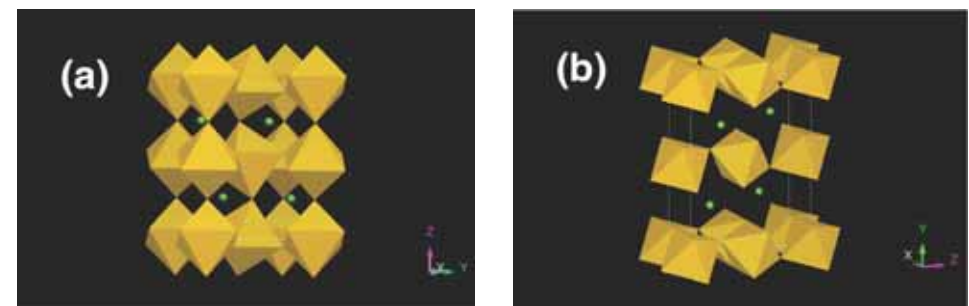
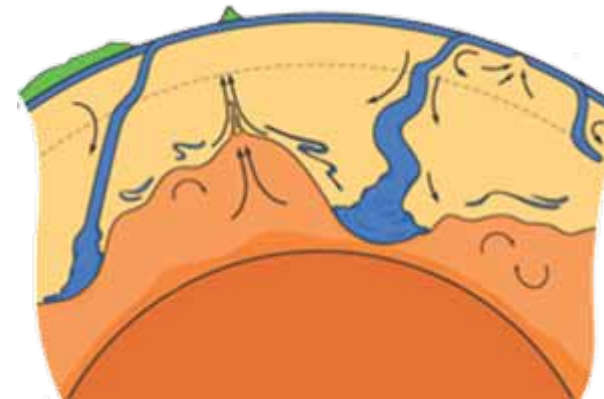
RSCCでの解析



例7 地球マントル最深部D“層の 第一原理分子動力学計算

- 地球最深部は見るできない。大規模第一原理計算が重要
- 新発見鉱物 MgSiO_3 ポストペロブスカイト相の地震波伝播特性を初めて第一原理計算で解明。
- 地震波観測との比較は、D“層主要成分がこの新発見鉱物であるとする根拠を与える。
- ポストペロブスカイトによるD“層の熱流体的不安定性は巨大マントル上昇流の種になる。
- 巨大マントル上昇流は海底埋蔵の大量のメタンハイドレートを融解させ地球温暖化の原因となる。
2.5億年前の生物大量絶滅

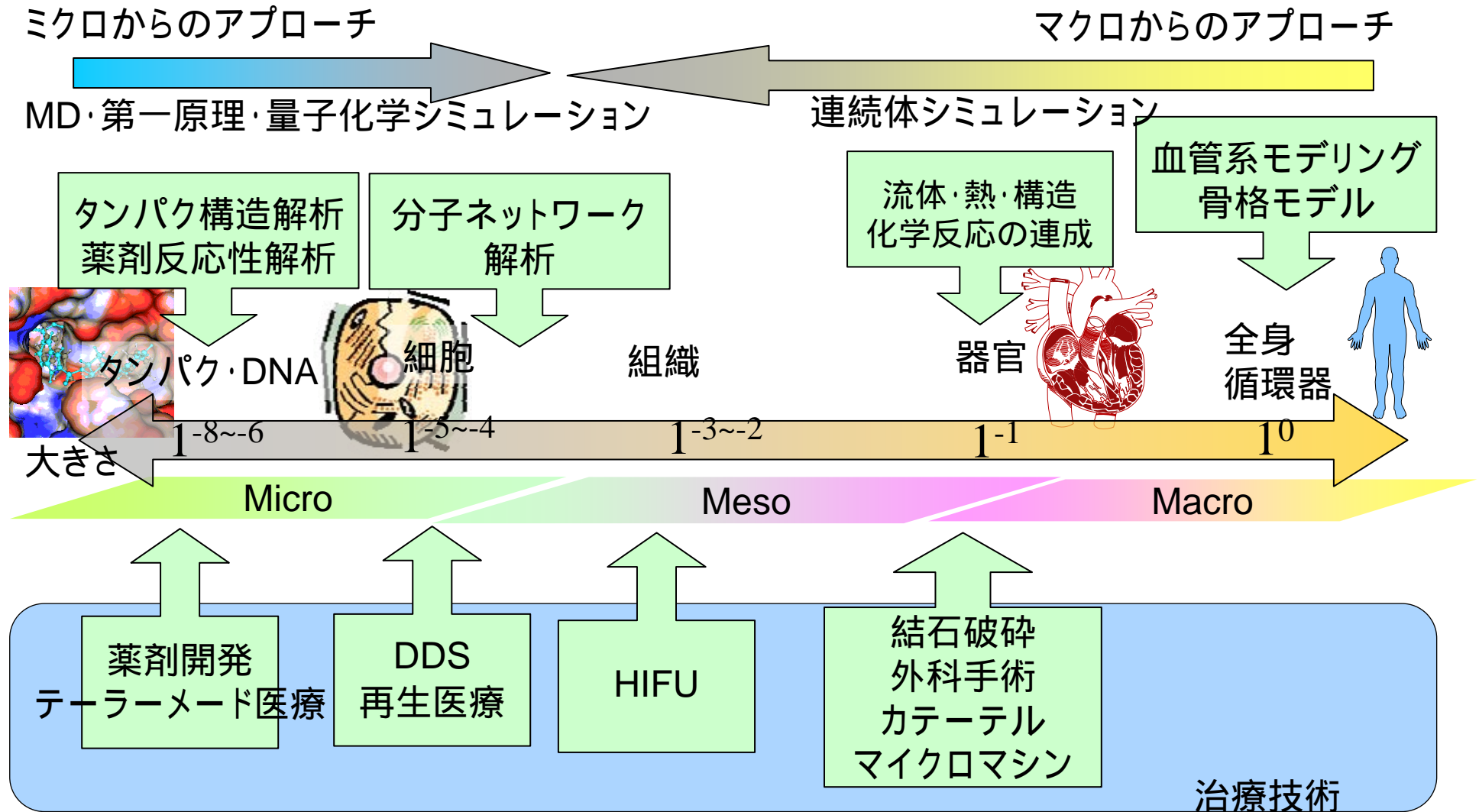
Figure by Krieger (2004)



(a) MgSiO_3 ペロブスカイトと(b)新発見鉱物
 MgSiO_3 ポストペロブスカイトの結晶構造

T. Iitaka et al., Nature 430, 442-445 (2004)₂₇

ライフサイエンス分野のシミュレーションの広がり



今後のシミュレーション

- 仮定を置いて演算量を少なくした計算 → 原理に忠実な計算
 - 分子軌道法 → 分子動力学 → 第一原理計算
- 個々のミクロなシミュレーションを統合したマクロなシミュレーションへ → マルチレベル
 - レベルを繋ぐモデル化が必要
 - 分子の反応 → 分子ネットワーク → 個体
 - 部品 → 完成品、部品 → 加工機全体
- 単一の現象のシミュレーションから複合したシミュレーションへ → マルチフィジックス
 - 単機能の解析 → 互いに関連する複合解析
 - 構造/流体 → 構造・熱・流体・化学反応
- 実験データの処理 → 逆問題としてシミュレーション
 - タンパク質構造解析
 - 高エネルギー物理