

Grid World 2005 シンポジウム

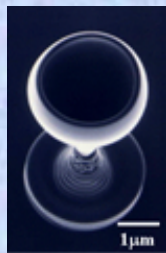
生体分子シミュレーションと GRID ASP  
- 新しいコンピュータ利用技術GRIDとの連携を目指して -

平成17年5月12日

NEC基礎・環境研究所  
高田 俊和

最新の微細加工技術

FIB-CVD法により作製した立体ナノ構造



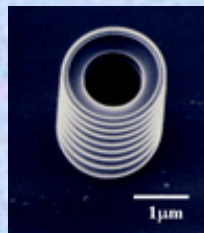
ワイングラス



コイル



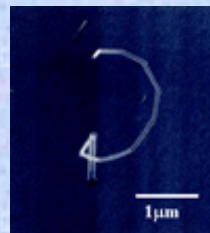
ドリル



ペロー



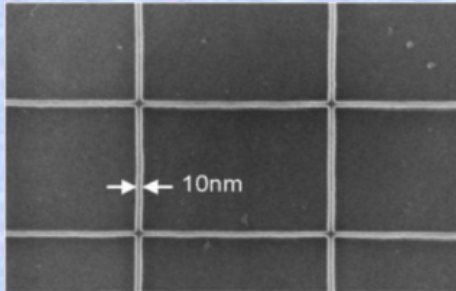
マイクロコロセウム



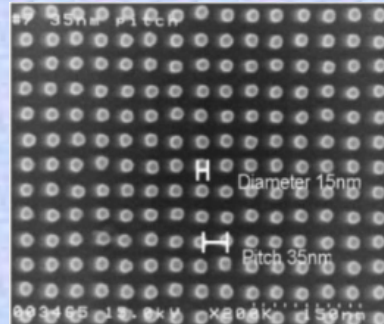
樹状突起

提供: NEC, 共同研究: 姫工大, SII

## カリックスアレーンを用いたナノ加工



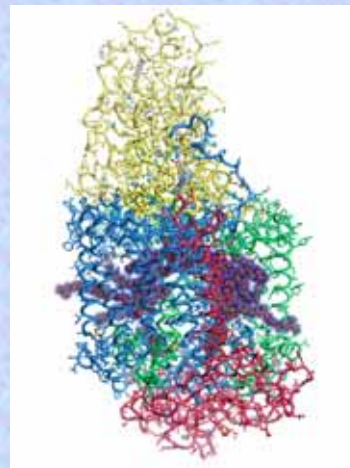
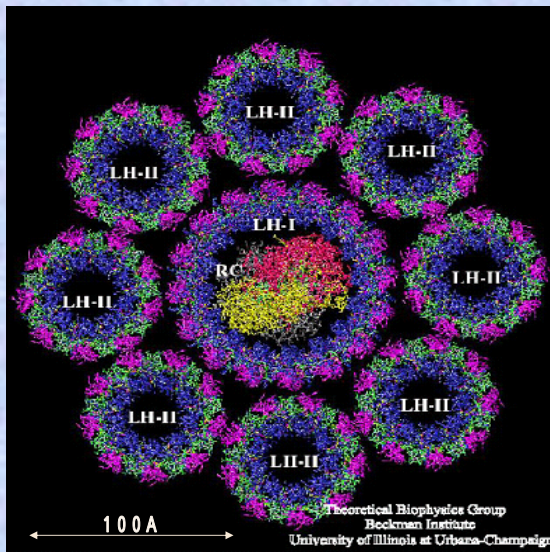
10nm ラインパターン



直径15nm, ピッチ35nm

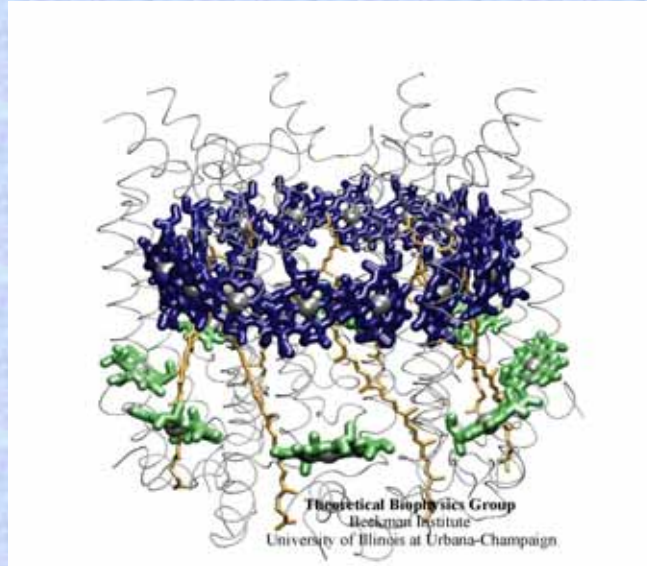
提供:NEC

## 光合成活性中心とアンテナクロロフィルの構造



活性中心とクロロフィル2量体  
- クロロフィル2量体の電子分布 -

## アンテナクロロフィルの構造と太陽光の捕獲機能

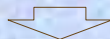


理論的考察: 住 斉 “電子と生命”, 第2章, 共立出版, 2000年

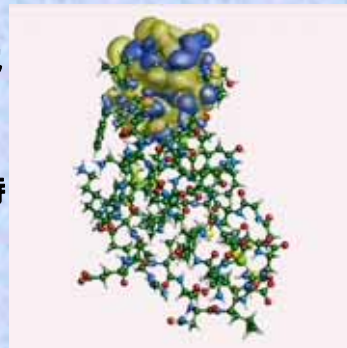
## 生体高分子シミュレーションと新産業創出

自然(生体機能)を学ぶことによる、新技術の創造

- 1) タンパク質(酵素) 触媒、創薬
- 2) 膜タンパク質 フィルター
- 3) 光合成活性中心 太陽電池



生体高分子シミュレーションへの期待



PKC + フォルボルエステル複合体  
(856原子、8672原子軌道)

## “生体の神秘に学ぶ”の原点



1976年出版

30年後の現在

### 構造生物学の進歩

X線、NMR、電子線、質量分析、  
遺伝子工学

### 分子シミュレーションの進歩

計算手法、分散処理、GRID、専用  
ボード、データベース、インターネット



環境は整いつつある。

## 生体分子シミュレーションの基本的考え方



インフルエンザAと薬

創薬: インヒビターによるタンパク  
質の化学反応機構のコントロール

タンパク質・リガンド相互  
作用は化学現象

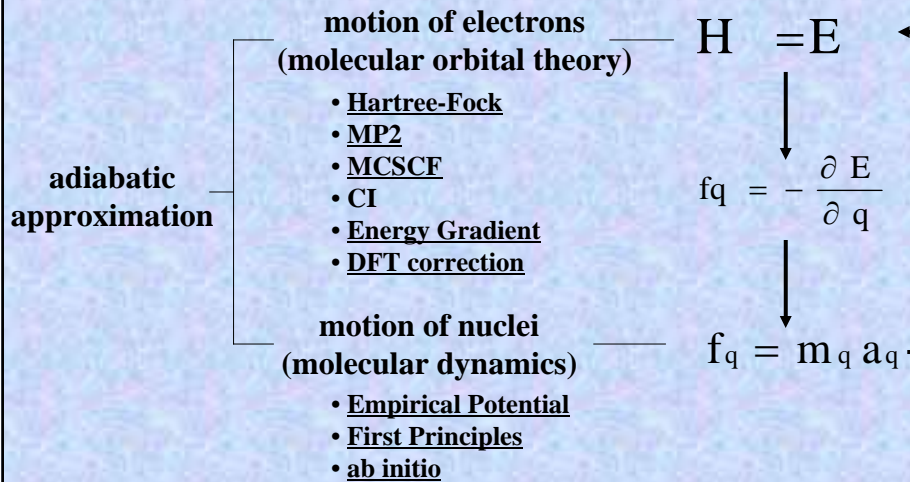
量子力学により理論的に  
予測可能

実験メカニズム解明と現象  
予測が可能



シミュレーションと実験による  
本質的な理解と応用が可能

## 生体分子シミュレーションの理論的枠組



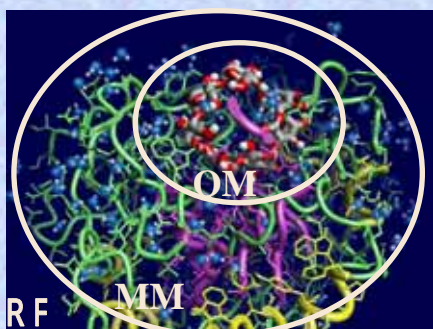
注) アンダーラインは、プロトタイププログラムで実現されている計算機能

## 生体分子シミュレーションの課題

1. 数万原子に及ぶ巨大分子      計算規模  
地球シミュレータ: ~ 10日 / 1点計算
2. 小さなエネルギー変化      高信頼性  
ポストHartree - Fock法の必要性
3. 自由エネルギーでの考察      膨大な繰返し計算  
必要な繰返し回数: ~ 1千万回

## 課題に対する対策

### QM / MM (Quantum Mechanics/Molecular Mechanics) 法の開発



QM空間: 化学反応の起こる領域  
小さくして、高速化

MM空間: QM以外の2次的領域  
近似式で、高速化



世界中でソフト開発競争

独自理論に基づく  
ソフト自主開発



新たな課題: QM・MM  
境界領域の記述が困難

## QM / MM法に現れるエネルギー項

	QM	MM	RF	
QM	$E_Q$			$E(Q) = \left\langle \Phi \left  -\frac{1}{2} \sum_i \nabla_i^2 + \frac{1}{2} \sum_{i \neq j} \frac{1}{r_{ij}} - \sum_{i,a} \frac{z_a}{r_{ia}} \right  \Phi \right\rangle + \frac{1}{2} \sum_{a \neq b} \frac{Z_a Z_b}{R_{ab}}$
MM	$E_{QM}$	$E_M$		$E(QM) = - \left\langle \Phi \left  \sum_{i,m} \frac{q_m}{r_{im}} \right  \Phi \right\rangle + \sum_{a,m} \frac{Z_a q_m}{r_{am}}$
RF	$E_{QR}$	$E_{MR}$	$E_R$	$E(QR) + E(MR) = - \left\langle \Phi \left  \sum_{i,o} \left( \frac{q_o}{r_{io}} \right) \right  \Phi \right\rangle + \sum_{a,o} \frac{Z_a q_o}{r_{ao}} + \sum_{m \neq o} \frac{q_m q_o}{\epsilon r_{mo}}$

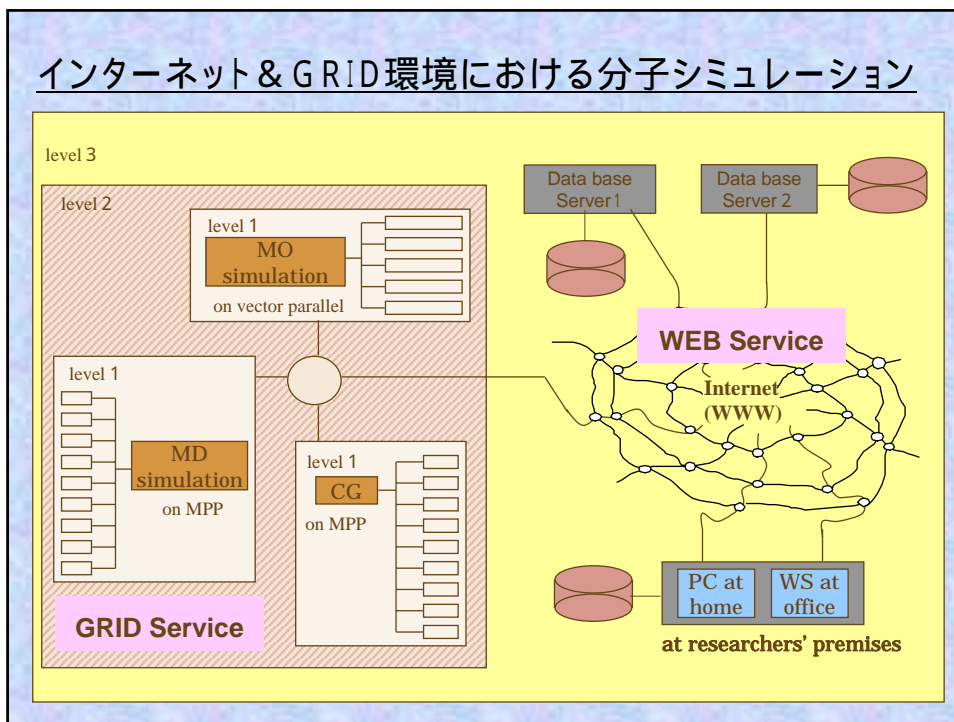
$$E(M) = \sum_{bonds} k_i^{bond} (r_i - r_{eq})^2 + \sum_{angles} k_i^{angle} (\theta_i - \theta_{eq})^2 + \sum_{dihedrals} k_i^{dihe} \left[ 1 + \cos(n_i \phi_i + \sigma_i) \right]$$

$$+ \sum_{a,m} 4 \epsilon_{am} \left[ \left( \frac{\sigma_{am}}{r_{am}} \right)^{12} - \left( \frac{\sigma_{am}}{r_{am}} \right)^6 \right] + \sum_{m,n} 4 \epsilon_{mn} \left[ \left( \frac{\sigma_{mn}}{r_{mn}} \right)^{12} - \left( \frac{\sigma_{mn}}{r_{mn}} \right)^6 \right] + \sum_{i \neq j} \frac{q_i q_j}{\epsilon r_{ij}}$$

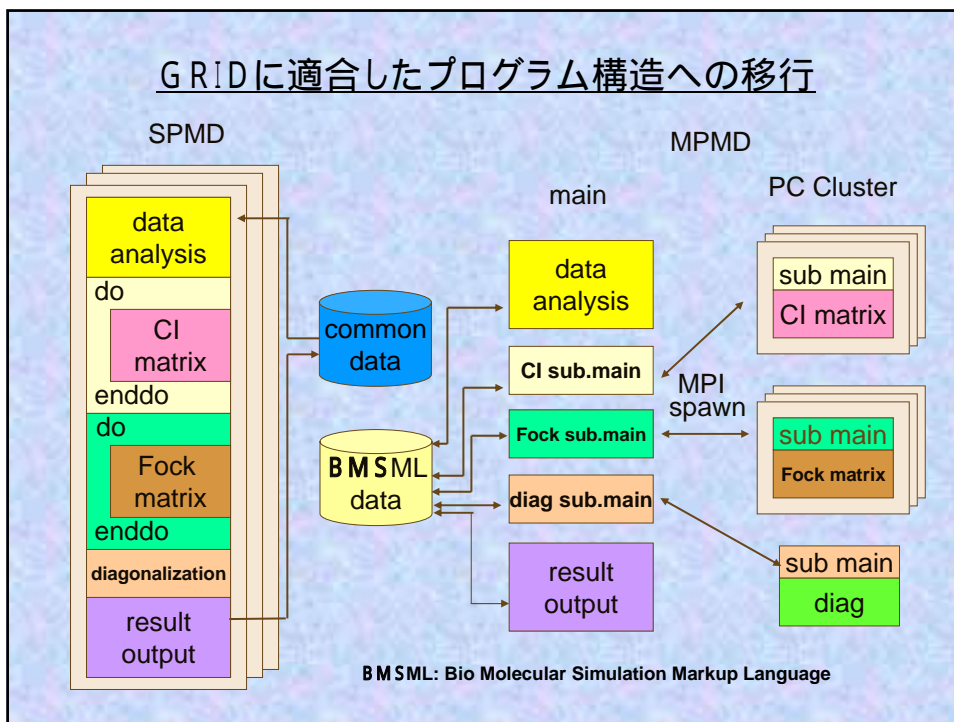
複数の計算要素を連成

グリッドに適合

## インターネット & GRID環境における分子シミュレーション



## GRIDに適合したプログラム構造への移行



## XMLによるコンポーネントプログラムインターフェース

CMLComp in CML

<http://cml.sourceforge.net/schema/cmlComp/>

QCML (Quantum Chemistry ML) in AbiGrid

<http://www.cineca.it/abigrid/workArea/QCMLdoc.html>

BMSML (Bio-Molecular Simulation ML) in BioGrid

This will be opened soon at <http://www.biogrid.jp/>

For a large matrix which can not be stored at a time in memory

1.1 1.2 | 1.3 1.4  
2.1 2.2 | 2.3 2.3

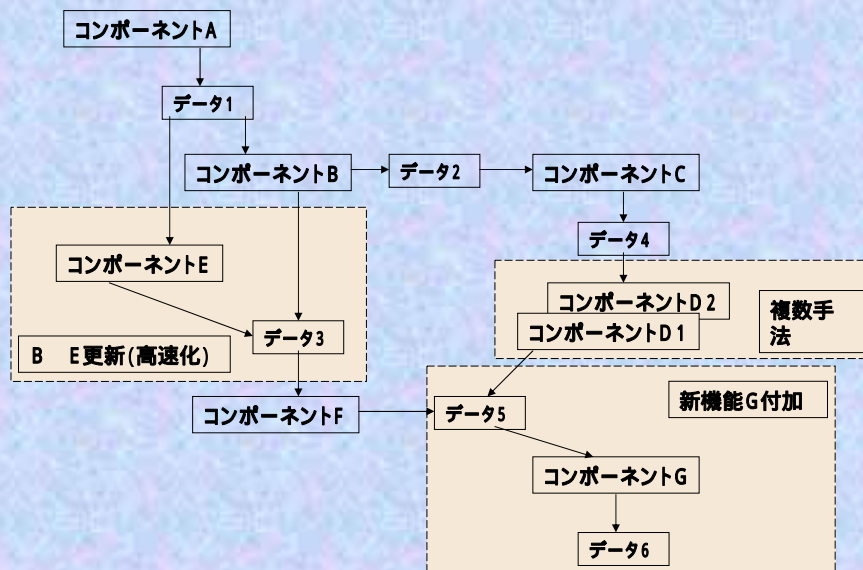
-----  
3.1 3.2 | 3.3 3.4

4.1 4.2 | 4.3 4.4

```
<blockMatrix matrixID="001" content="blockContainer" numElement="4">  
<bmsDouble2D content="block1" blockID="1" dataNum1D="2" dataNum2D="2"  
startBlock1D="1" startBlock2D="1">1.1 1.2 2.1 2.2</blockMatrix>  
<bmsDouble2D content="block2" blockID="2" dataNum1D="2" dataNum2D="2"  
startBlock1D="1" startBlock2D="3">1.3 1.4 2.3 2.4</blockMatrix>  
<bmsDouble2D content="block3" blockID="3" dataNum1D="2" dataNum2D="2"  
startBlock1D="3" startBlock2D="1">3.1 3.2 4.1 4.2</blockMatrix>  
<bmsDouble2D content="block4" blockID="4" dataNum1D="2" dataNum2D="2"  
startBlock1D="3" startBlock2D="3">3.3 3.4 4.3 4.4</blockMatrix>  
</blockMatrix>
```

XML to describe relations between components is under development

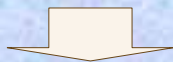
## コンポーネントプログラミングとワークフロー



## シェルスクリプトによる実行イメージ

```
#!/bin/sh

1. LAM/MPIの起動
lamboot v ~/lamhost-4
2. HF用初期MO係数の生成
mpirun lamd O huds_mpi_initmo_gbe > zzlog_uds_mpi_initmo_gbe 2>&1
3. CASSCF用MO係数の生成
mpirun lamd O huds_mpi_rhf > zzlog_uds_mpi_rhf 2>&1
4. CASSCFの実行
mpirun lamd O huds_mpi_casscf > zzlog_uds_mpi_casscf 2>&1
5. 自然軌道によるMullikenポピュレーション解析
mpirun-lamd O huds_mpi_pop_mulliken > zzlog_uds_mpi_pop_mulliken 2>&1
6. LAM/MPIの解除
wipe -v ~/lamhost-4
```



ワークフローへの展開を検討中

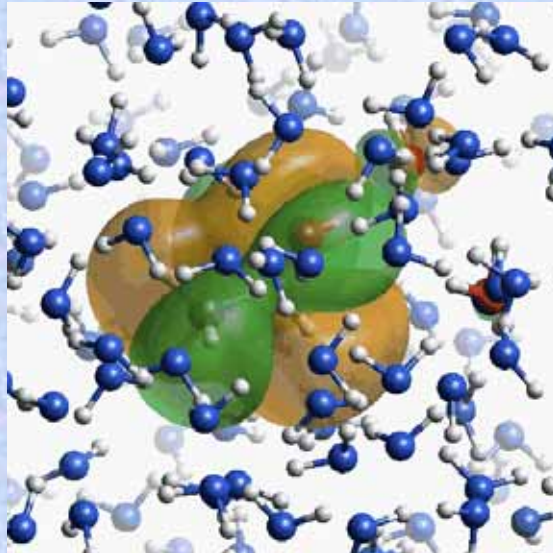
## コンポーネントプログラミングの利点

1. 新計算機能の付加が容易  
必要なD - XMLのみの情報で、機能付加が実現
2. 計算機能の更新が容易  
コンポーネントが疎結合のため、修正箇所が限定
3. 複合した計算機能の実現が容易  
コンポーネントの組み合わせを変えるだけで、異なる機能を実現
4. システムのメンテナンスが容易  
プログラム単位が小さく、修正箇所が明瞭
5. GRIDやWEBサービスなど、分散処理に適合  
プログラム単位が小さく、GRIDなどの新技術に対応可能

## QM / MM法によるエタノール + 水系のMD計算

**Ethanol in  
water (TIP4P)  
at 283 K**

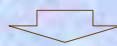
**HOMO in  
the QM is  
displayed.**



出典: [www.biogrid.jp](http://www.biogrid.jp)

## GRID ASPビジネスの創出に向けて

データベース検索、立体構造予測、シミュレーション結果  
など科学技術情報のインターネットによる販売

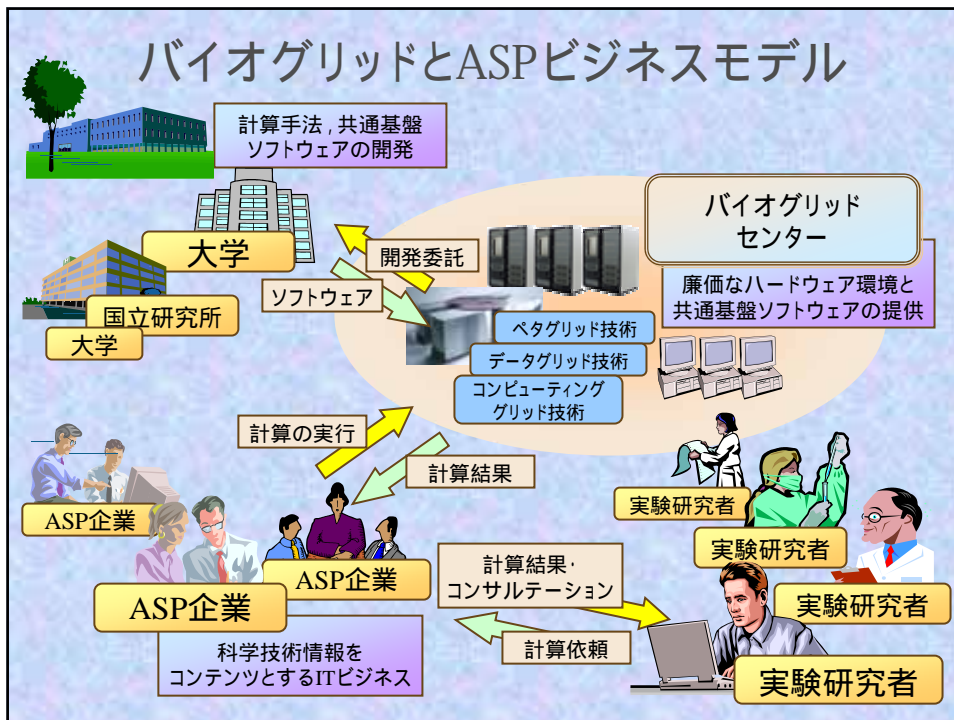


科学技術情報をコンテンツとした、新IT(ASP)ビジネス

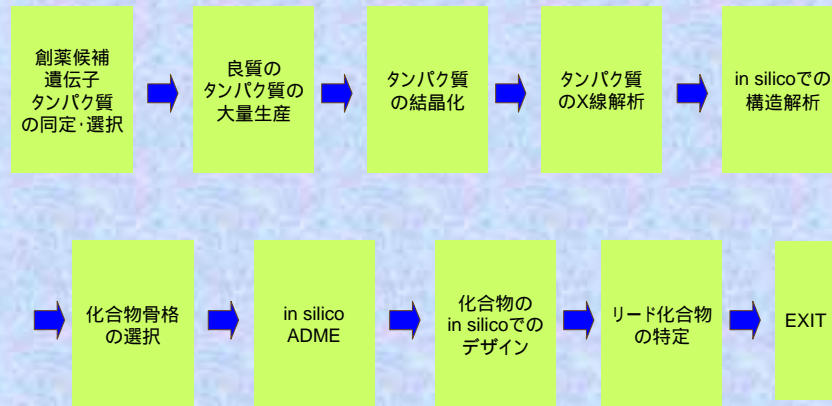
- ・ 物流を伴わない情報販売ビジネス
- ・ インターネットによる世界市場の形成

## バイオIT(ASP)におけるユーザの利点

1. コンピュータやプログラムの購入の必要がない
2. コンピュータの使い方を覚える必要がない
3. 初期投資がなく、負担がすくない
4. オンラインコンサルテーションが受けられる
5. 計算規模に対するフレキシビリティが高い
6. 最新のコンピュータとプログラムが使える



## in silico 創薬バリューチェーン



出典: [www.biogrid.jp](http://www.biogrid.jp)

## NPO法人バイオグリッドセンターのテストベッド事業

The screenshot shows the website for NPO法人バイオグリッドセンター (BioGrid Center). The header includes the BioGrid logo and the text '特定非営利活動法人 バイオグリッドセンター 関西'. Below the header is a navigation menu with links: top, 設立趣旨, 活動内容, 法人概要, and [ご入会/お問合せ](#). The main content area features a green banner for 'バイオグリッド・テストベッド実証実験' (BioGrid Test Bed Realization Experiment). Below this are links for '概要', 'サービスプロバイダのご案内', 'ご利用申請', and '各プロバイダの皆様へ'. A paragraph states: 'バイオグリッド・テストベッド実証実験の利用者にサービスを提供するサービスプロバイダをご案内します。' Below this is a green box with the text '■ NEC-BIOASP'. At the bottom right, there is a link '利用者のログインは[こちらから](#)' and the source '出典: [www.biogrid.jp](http://www.biogrid.jp)'.

## 計算実行及び解析のためのWEBサービスプログラム



## 結論

GRIDが、プログラミングスタイルを変更

連成プログラミング技術の発展

バイオでの、GRIDコンピューティングへの期待

科学技術情報をコンテンツとする、ITビジネスへの展開

最大の課題は、セキュリティの確保

### 阪大BioGirdプロジェクト関係者

大阪大学蛋白研 : 米澤康滋、中村春木  
大阪大学院理 : 山中秀介、山口兆  
日立製作所 : 黒澤隆、何希倫  
NEC情報システムズ : 佐久間俊広  
NECソフト : 中田一人、山本純一 (BMSML)  
日本電気 : 徳島大介 (BMSML)