



国プロジェクトと産学連携のご紹介

アドバンスソフト技術セミナー

イノベーション基盤シミュレーションソフトウェアの作成
量子バイオシミュレーションシステムの作成
アドバンスソフト株式会社
主事研究員

日野 理 1.

講演内容

- 創薬シミュレーションとはどのようなものか
- 量子化学計算の特徴
- 生体分子量子化学計算ソフトウェアBioStation
- イノベーションPJにおけるアドバンスソフトの役割
- 密度汎関数理論(DFT)の導入
- RI(Resolution of Identity)法の導入
- 計算化学ソフトウェア開発への協力体制

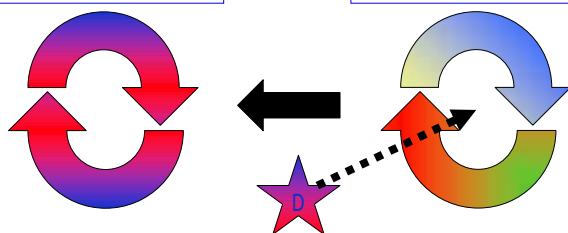


創薬の考え方

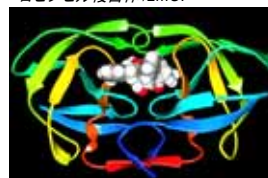
「薬」という視点から生命(人間)を考える
 主として有機化合物と水から成る自律的な化学反応サイクル
 病気とは
 化学反応サイクルが“不安定”になること
 薬とは
 サイクルに投入し、安定性を回復させる化学物質
 良く効く薬
 化学反応を支配する生体分子によく“くっつく”化学物質

健康=安定なサイクル

病気=不安定なサイクル



HIV-1プロテアーゼダイマー
 ロビナビル複合体:1MU1



HIVの複製をブロック

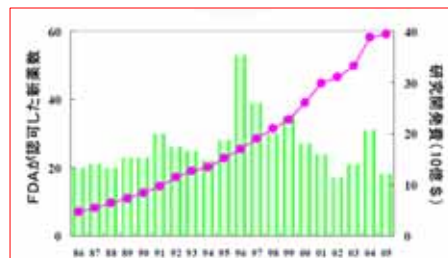


創薬の問題点

現代的な薬の発見 (Drug Discovery)

- 病気の原因の解明(仮説の構築) 関連生体分子(ターゲット)の特定
- ターゲットと結合する分子(リード化合物)の探索・合成
- リードの最適化(活性向上、毒性除去)と臨床試験ののち市場へ

創薬における問題点
 リード探索が困難化
 開発コストの上昇
 製薬企業の巨大化(合併)
 ブロックバスター志向



- 薬価の上昇
- “儲からない”薬の敬遠



創薬プロセスの効率化を促進するための
 シミュレーション技術発展は極めて重要

創薬シミュレーションの役割

生命はあらゆる階層の化学が集約したシステム 様々なシミュレーション技術

- タンパク質構造モデリング(ホモロジーモデリング)
- タンパク質-化学物質複合体予測(ドッキング)
- バイオ、ケモインフォマティクス
- ADMET(吸収、分布、代謝、排出、毒性)解析
- タンパク質-各種化学物質相互作用解析(分子力学)



微細な分子修飾による結合能の大きさを予測することが必要
水素結合、ファンデルワールス相互作用等の精密な評価

経験的パラメータを用いない第一原理的量子化学シミュレーション



量子化学シミュレーション

量子力学による分子の理論

E. Schrödinger



Emilio Segre Visual Archives

◆ 分子のシュレディンガー方程式

$$\hat{H} = \sum_{i=1}^N \left(-\frac{\hbar^2}{2m} \Delta - \sum_{A=1}^M \frac{Z_A e^2}{|\mathbf{R}_A - \mathbf{r}_i|} \right) + \sum_{i < j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}$$

$$\hat{H}\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N) = E\Psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)$$

化学に関する自然法則は、これによってほとんど尽くされている
残されているのは計算技術の問題である

P.A.M Dirac



Emilio Segre Visual Archives

Hartree-Fock法 最も基本的な近似法

$$\left(-\frac{\hbar^2}{2m} \Delta - \sum_{A=1}^M \frac{Z_A e^2}{|\mathbf{R}_A - \mathbf{r}|} \right) \varphi_p(\mathbf{r}) + \sum_{i=1}^N \int \varphi_i(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \varphi_i(\mathbf{r}') d\mathbf{r}' \varphi_p(\mathbf{r}) - \sum_{i=1}^N \int \varphi_i(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \varphi_p(\mathbf{r}') d\mathbf{r}' \varphi_p(\mathbf{r}) = \varepsilon_p \varphi_p(\mathbf{r})$$

小さな分子(~ 100原子程度) は計算できるが...



D. R. Hartree



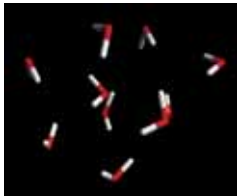
V. A. Fock

Emilio Segre Visual Archives



量子化学シミュレーションのコスト

- 量子化学計算にかかるコスト(時間) vs. 分子サイズ
- 実験的計算: 水分子クラスター-6-31G基底

H ₂ O × 1	H ₂ O × 10	H ₂ O × 100
		
HF 0.16 sec MP2 0.07 sec	HF 59.8 sec MP2 33.8 sec	HF 48201.3 sec MP2 3.7 days

- ◆ 量子化学計算のスケールリングは非常に大きい
- ◆ 生体分子を量子化学で扱うのはきわめて非現実的

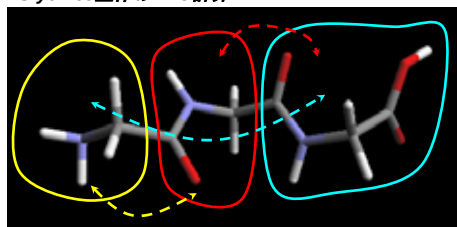


フラグメント分子軌道法

フラグメント分子軌道(FMO)法 北浦和夫博士(1998)

- 巨大分子を分割して個々に量子化学計算を行う
- フラグメントペア(ダイマー)計算による補正
- FMO法は、なぜ機能するのか
- 巨大分子を複数のフラグメントの集合体(クラスター)とみなす
- 3クラスター間の多体相互作用がダイマー計算によって全て考慮される

Glycine3量体のFMO計算



- ◆ フラグメント計算の手順はほぼ同じ。
 - ◆ フラグメントペアについて量子化学計算を行い、各フラグメントとフラグメントペアの電子密度から全電子密度を再構成する。
- | | |
|---------|----------------|
| HF | -696076.054mEh |
| FMO-HF | -696075.026mEh |
| MP2 | -697442.951mEh |
| FMO-MP2 | -697441.407mEh |



~ 200残基程度のタンパク質の量子化学計算が1日程度で計算可能になる

BioStation開発の背景

- 戦略的基盤、革新的シミュレーションソフトウェアプロジェクトでFMO法に基づく量子化学計算のタンパク質・化学物質相互作用解析を行うためのソフトウェア **BioStation**を開発
- **大規模生体分子の量子化学解析**を専門に行う世界でも例のないソフトウェア
- BioStationは**創薬研究支援を目的**としたソフトウェア。アカデミアや産業界への貢献も現実化
- バイオ分野の量子シミュレーションは、進歩も急速である。欧米も国策として量子計算ソフトウェアの強化を図っている
- BioStationを維持発展させるためには、最近の研究成果に基づきBioStationの機能を強化することが必要不可欠



量子バイオシミュレーションシステム

量子化学計算ソフトウェア BioStation

FMO法による量子化学計算に特化したソフトウェア

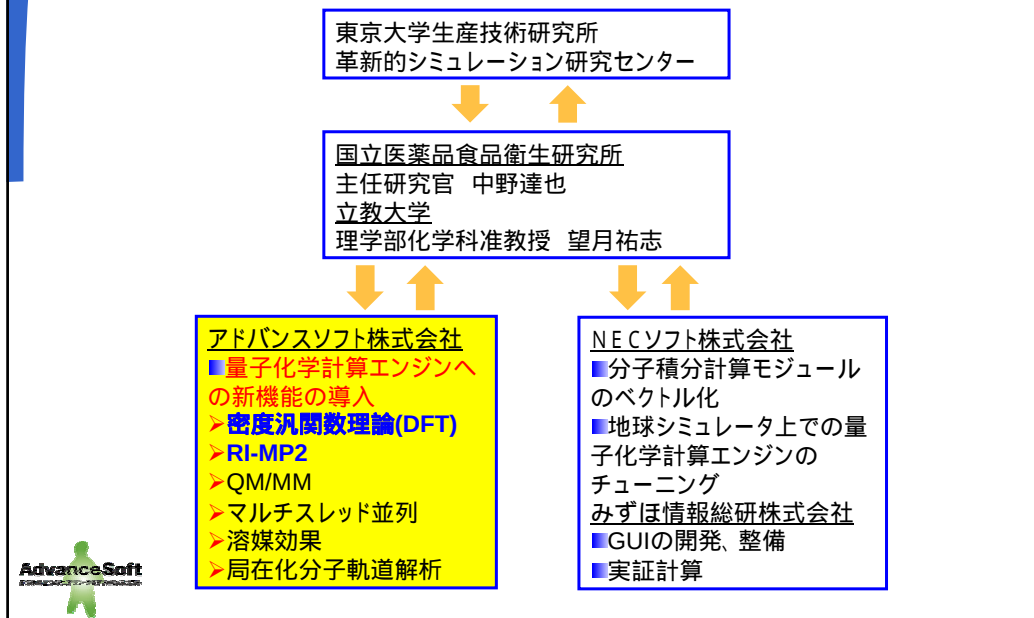
- 戦略的基盤ソフトウェアの開発 (平成14年度～平成17年度)
タンパク質化学物質相互作用解析
- 革新的シミュレーションソフトウェアの研究開発 (平成17年度～平成19年度)
タンパク質・化学物質相互作用マルチスケールシミュレーション

主要な計算機能

- FMO-Hartree-Fock計算
- FMO-MP2計算
- 多層化FMO計算 (活性部位をMP2計算、その他をHF計算)
- XUFF (eXtended Universal Force Field) による高精度古典力場計算
- FMO-QM/MM (QM:FMO/MM: XUFF) 計算
- PDBフォーマットから直接FMO計算が可能
- フラグメント間相互作用エネルギー解析機能
- 高効率な並列計算 (並列化率～99.95%)



アドバンスソフトの役割



FMO-DFTの導入

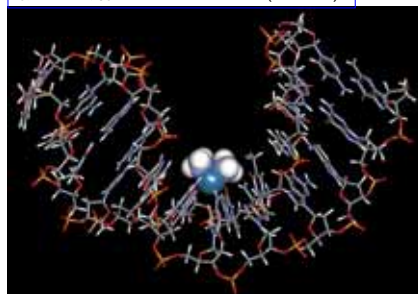
DFTの特長

- 量子化学計算のde facto standard (Gaussian, Jaguar, etc.)
- 低計算コストで高精度な計算が可能 ($\sim O(N^3)$)
- 有機分子のみならず、金属を含む分子でも高精度計算可能

高血圧症薬エナラプリル (1UZE)



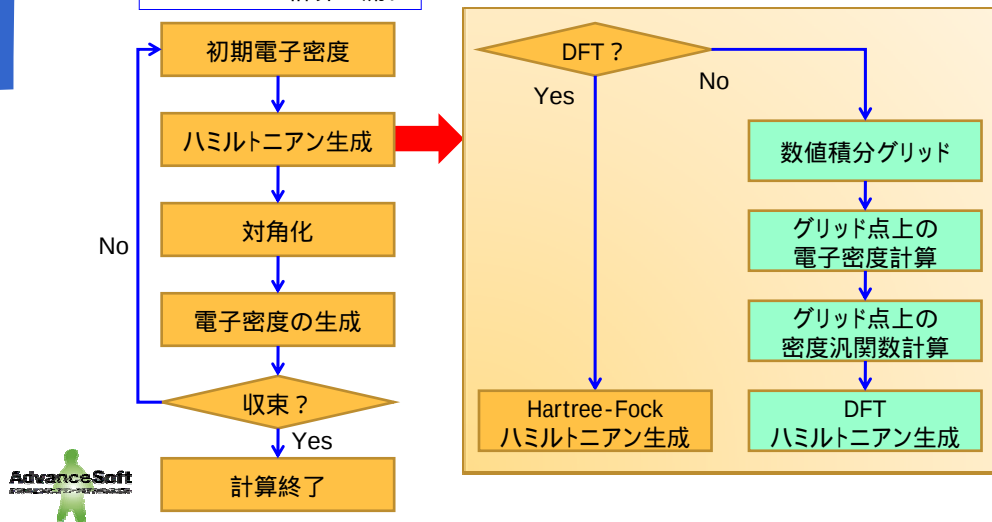
抗がん剤カルボプラチン (1A84)



FMO-DFTの導入

スムーズなDFT計算の導入: 開発時間、経費の圧縮

Hartree-Fock計算の流れ



FMO-DFTの導入

ABINIT-MPで使用可能な密度汎関数

汎関数	説明	Gaussian
SLATER	Slater交換汎関数	HFS
BECKE88	Becke 88交換汎関数	HFB
SVWN	Slater交換汎関数 + Vosko-Wilk-Nusair相関汎関数	SVWN5
SVWNrpa	Slater交換汎関数 + Vosko-Wilk-Nusair (RPA)相関汎関数	SVWN
BVWN	Becke88交換汎関数 + Vosko-Wilk-Nusair相関汎関数	BVWN5
BVWNrpa	Becke 88交換汎関数 + Vosko-Wilk-Nusair (RPA)相関汎関数	BVWN
SLYP	Slater交換汎関数 + Lee-Yang-Parr相関汎関数	SLYP
BLYP	Becke 88交換汎関数 + Lee-Yang-Parr相関汎関数	BLYP
B3LYP	$(0.2 \times \text{Hartree-Fock} + 0.8 \times \text{Slater} + 0.72 \times \text{Becke88})$ 交換汎関数 + $(0.19 \times \text{Vosko-Wilk-Nusair (RPA)} + 0.81 \times \text{Lee-Yang-Parr})$ 相関汎関数	B3LYP

平成21年度以降、長距離相互作用を考慮した密度汎関数も導入予定



生体分子の相互作用解析では長距離相互作用が重要になる



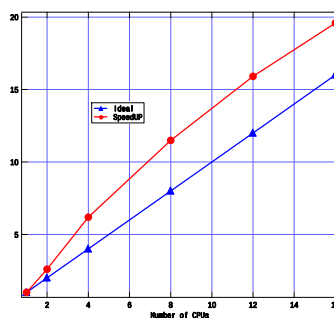
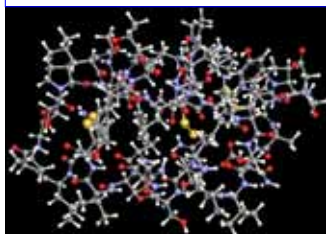
FMO-DFTの導入

プロトタイプモジュールのテスト計算結果

Glycine5量体/B3LYP/6-31G



Crambin/FMO/6-31G



方法	Monomer	Dimer	Total
MP2	25	371	396
B3LYP	106	333	439



RI法の導入

RI法とは

- 量子化学計算で最も計算負荷の大きな部分 **2電子積分計算**
- RI法は2電子積分計算を分割する方法 計算負荷を軽減
- MP2、DFTなどの高精度計算だけでなく、2電子積分を使った全ての計算部分に適用して計算の高速化が可能

$$(\mu\nu|\lambda\sigma) = \iint \varphi_\mu(\mathbf{r}_1)\varphi_\nu(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \varphi_\lambda(\mathbf{r}_2)\varphi_\sigma(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$$

$$(ai|bj) = \sum_{\mu\nu\lambda\sigma} (\mu\nu|\lambda\sigma) C_{\mu a} C_{\nu i} C_{\lambda b} C_{\sigma j}$$

MP2計算に必要な行列要素
~ O(N⁵)、使用メモリサイズの低減

$$J_{\mu\nu} = \sum_{\lambda\sigma} (\mu\nu|\lambda\sigma) P_{\lambda\sigma}$$

DFT計算に必要な行列要素
~ O(N⁴)、計算スケーリングがO(N³)

$$E_{Coulomb} = \sum_{\mu\nu\lambda\sigma} (\mu\nu|\lambda\sigma) P_{\mu\nu} P_{\lambda\sigma}$$

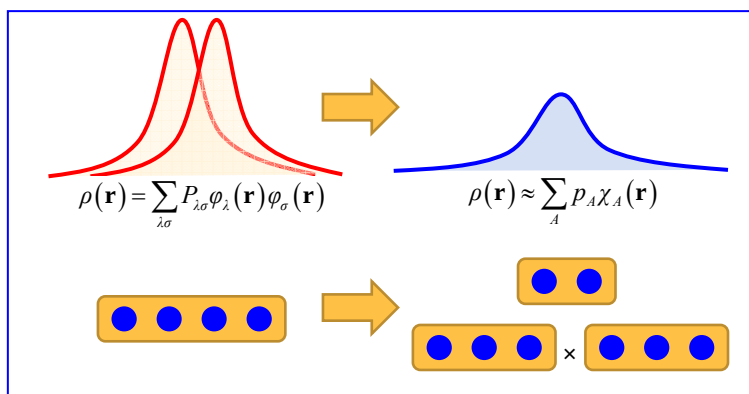
FMO計算に必要なクーロン反発エネルギー
~ O(N⁴)、計算スケーリングがO(N²)



RI法の導入

RI法の高速度の原理

- 電子密度分布関数 通常のガウス型関数で展開する
- 2電子積分 (4中心積分) 2中心積分、3中心積分に分解



RI法の導入

平成20年度の開発項目: 基盤モジュールの作成

開発項目	開発内容
補助基底関数の決定	適用範囲と計算精度の観点よりLCADを採用
補助基底関数設定モジュール	LCADに関する情報を設定するモジュールおよびサブルーチンの作成
2中心積分計算モジュール	LCADによる2中心積分計算モジュールの作成
3中心積分計算モジュール	LCADによる3中心積分計算モジュールの作成
3中心積分計算プロファイリング、チューニング	RI法による高速化の中心となる3中心積分計算のプロファイリングを行い、シリアル計算でのチューニングを行う
3中心積分計算並列化	RI法による高速化の中心となる3中心積分計算の並列化する
クーロン密度フィッティング計算モジュール	電子間クーロン相互作用をRI法で計算するモジュールの作成。計算精度のチェックとHartree-Fock、DFT計算への応用が可能

平成21年度以降、MP2、DFT等の計算の高速化を行う



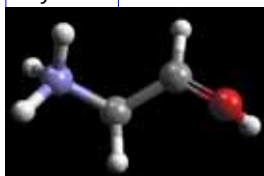
高速で高精度な、生体分子の量子化学計算が可能



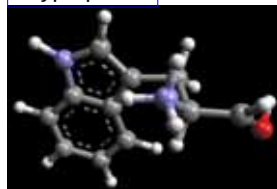
RI法の導入

プロトタイプモジュールの
テスト計算結果

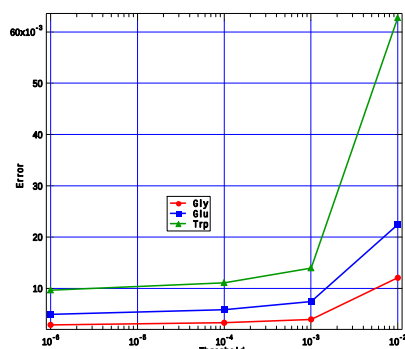
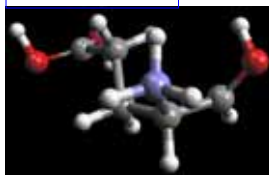
Glycine



Tryptophan



Glutamic acid



BioStationの高性能化・高機能化

イノベーション基盤シミュレーションソフトウェアの研究開発

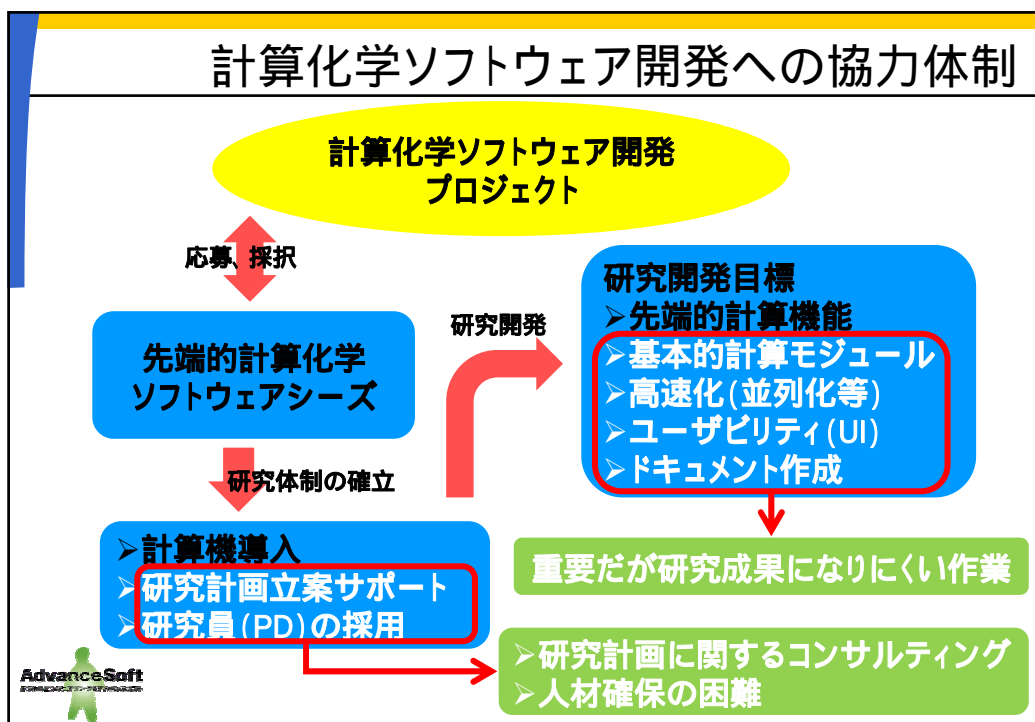
- 産業イノベーションへの貢献
- 我が国独自のシミュレーションソフトウェアの研究開発
- 開発されたシミュレーションソフトウェアの普及

量子化学計算エンジンに求められる機能

- 高精度な計算機能 密度汎関数理論(DFT) 計算機能
- 高速化(1) 計算アルゴリズム: RI法の導入
- 高速化(2) 計算アルゴリズム: QM/MMによる構造最適化
- 高速化(3) 並列計算スキームの改善: マルチスレッド並列の導入
- 生体内環境を考慮した計算機能 溶媒効果を考慮した計算機能
- 計算結果の詳細な解析機能 局在化分子軌道を用いた相互作用解析



計算化学ソフトウェア開発への協力体制



計算化学ソフトウェア開発への協力体制

戦略・革新・イノベーションPJ参加経験者(計算化学関連)

氏名	ソフトウェア
日野 理	BioStation, ProteinDF
小林 将人	BioStation
小川 哲司	BioStation
谷森 奏一郎	BioStation
小池 聡	ProteinDF
西村 民男	ProteinDF
佐藤 昌宏	ProteinDF
長谷川 浩司	BioStation

- 全員が学位保持者、複数のPD経験者
- 実務経験4年以上
- 理論、並列化、チューニングに関する知識
- 受託作業を通じて得た幅広い計算科学に関する知識
- 費用:2000万円/年

