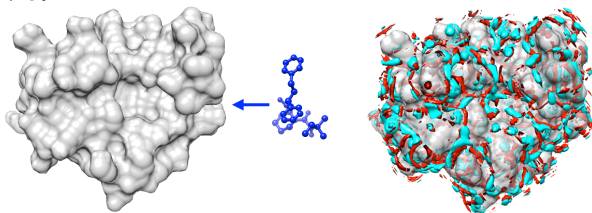


分子性液体の積分方程式理論による大規模生体分子系における高速な溶媒和自由エネルギー計算プログラムの開発



溶媒和自由エネルギー

溶媒和自由エネルギーは、分子が真空中から水などの溶媒に溶けた時の自由エネルギー差を表す。生体分子は一般に水溶液中にあり、反応の過程で水が結合したり(水和)、外れたり(脱水和)する。例えば、蛋白質に薬効分子が結合する過程では、蛋白質に結合していた水が脱水和し、薬効分子が結合する。



蛋白質と薬効分子及び蛋白質の周りの水の3次元分布(赤:O、水:H)

生体分子の機能の研究や創薬開発などで水和脱水和効果を正確に見積もるために、高精度で高速な溶媒和自由エネルギー計算法が求められている。

分子性液体の積分方程式理論

水などの分子性液体を取り扱う理論として3D-RISM理論がある。この理論は2つの方程式からなり、3D-FFTを用いて交互に計算を行なって3次元分布関数を収束させていく。そのため高速な3D-FFTを必要とする。分布関数を元に溶媒和自由エネルギーを計算する。

3D-RISM方程式

$$\tilde{h}_\gamma^{ub}(r) = \tilde{c}_x^{Ua}(r) * \left[\tilde{\omega}_{xy}^{ab}(r) + \rho^b \tilde{h}_{xy}^{ab}(r) \right] \quad \text{フーリエ空間}$$

クロージャ方程式

$$h_\gamma^{ub}(r) = \begin{cases} \exp(\chi) - 1 & \text{for } \chi \leq 0 \\ \chi & \text{for } \chi > 0 \end{cases}$$

$$\chi = -\beta u_\gamma^{Ub}(r) + h_\gamma^{ub}(r) - c_\gamma^{Ub}(r).$$



(従来の)溶媒和自由エネルギー式

$$\Delta\mu = \rho k_B T \sum_\gamma \int_{V_{all}} dr \left[\frac{1}{2} (h_\gamma(r))^2 \Theta(-h_\gamma(r)) - c_\gamma(r) - \frac{1}{2} h_\gamma(r) c_\gamma(r) \right]$$

3次元FFTの計算

一般にはメモリアクセス性能が重要

従来のSXシリーズは細粒度のメモリアクセスが可能 ⇒ Y方向やZ方向への不連続アクセスの効率が良い。

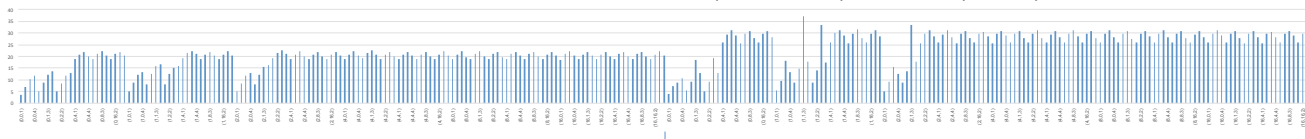
ACEは転送単位が増大 ⇒ GPU用に提案された手法(額田ら@SC'08)が有効であると考えられる。

	主記憶のデータ転送単位
SX-8	8 バイト
SX-9	8 or 16 バイト
SX-ACE	64 or 128 バイト
GPU	64 ~ 128 バイト

現在SX-ACE用の実装を行っている段階である。SX-ACEでは基本的に演算およびメモリアクセスのベクトル化が必須。

GPU	SX-ACE
各スレッドのレジスタのデータ交換が自由	ベクトルレジスタ内のデータの順序等を変えるのは容易ではない
各スレッドがレジスタにデータを持つ	ベクトルデータレジスタにデータを持つ
Shared memoryをスクラッチパッドメモリとして使える	ADBはWrite through

SX-ACE 1プロセッサによる3次元複素倍精度FFT(256×256×256)の演算性能(GFLOPS)



実部と虚部を交互に持つインターリーブ形式

実部と虚部を別々に持つセパレート形式

パラメータ設定

ASLライブラリのマルチスレッド版FFTライブラリを使用して計測。

パラメータは(X方向のパディング、Y方向のパディング、スレッド数)を表す。

パディングは{0, 1, 2, 4, 8, 16}、スレッド数はコア数に合わせて{1,2,3,4}から選択。

傾向

X方向のパディングが2以上あれば高性能で安定。

スレッド数による性能向上は鈍い⇒メモリアクセスがネック。

最適なパラメータは(X方向のパディング=1、Y方向のパディング=1、スレッド数=3)。

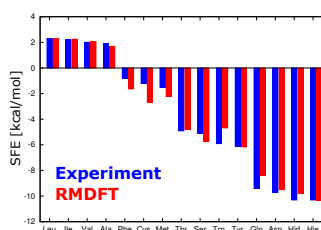
新しい溶媒和自由エネルギー表式(RMDFT)

従来の溶媒和自由エネルギー式では、絶対値にずれが生じる事が知られていた。そこで、溶媒和自由エネルギーの精度を向上させるために、密度汎関数理論に基づいて新しい溶媒和自由エネルギー表式を導出した。(式は複雑なので省略する。詳細は"墨、光武、丸山、J. Comp. Chem. DOI: 10.1002/jcc.23942"に記載されている。希望者には別刷りをお渡しする。)

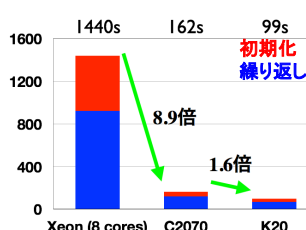
3D-RISM理論を使って3次元分布関数を計算し、それを元にRMDFTを使って溶媒和自由エネルギーを算出する。新しい表式は畳み込み積分の形をしており、計算に3D-FFTを使用する。

計算結果

理論の検証として、アミノ酸側鎖分子の溶媒和自由エネルギーを計算し、実験値と比較した。側鎖分子は疎水性、親水性など多様な性質を持っているが、その特徴を再現する結果を得られた。これにより蛋白質等の生体分子への適用時の信頼性が向上した。



アミノ酸側鎖の溶媒和自由エネルギー



CPU及びGPUにおける計算時間 (FKB蛋白質)

並列化

3D-RISM及びRMDFTの計算は3D-FFTを繰り返し実行する。実際、繰り返し部分において50%近くが3D-FFTとなっている。そのため計算はメモリ律速となる。また、Xeonと比較してメモリバンド幅の広いGPUで大きく速度向上している。一方、計算にはメモリを必要とするためGPU単体では、計算対象や精度に制限が生じる。

この制限を解消するため、SX-ACE及びGPUでの並列化を試みる。繰り返しになるが、この計算においては3D-FFTが重要であり、そのチューニングを行なっている。