

結晶多形解析のための並列結晶計算法の開発 (豊橋技科大工) 小畑繁昭, 中山尚史, 鎌倉寿行, 後藤仁志

【はじめに】

分子のコンフォメーションの自由度が大きいフレキシブルな有機分子は、結晶多形が多く存在し、また、これらは密度や融点などの物性が異なる。そのため、有機材料設計や医薬品開発において、分子シミュレーションによる結晶多形の解析や評価に対する期待は大きい。一方、分子シミュレーションによる結晶計算は、扱う結晶の大きさを大きくするほど、より信頼性の高い結果を得られると考えられるが、結晶の大きさに依存して分子間相互作用数が増加するため、計算量は膨大になる。そこで、本研究では、大規模な結晶構造モデルの計算を可能とするため、結晶の高い対称性に基いた効率的な Master-Worker 型の並列結晶計算法を開発し、CONFLEX に導入した。そして、結晶半径 0.1 μm のアスピリン結晶を用いて並列化効率を調べ、結晶計算の並列計算処理に対する本手法の有効性について検討した。

【結晶計算の並列化】

結晶構造は、構成分子が単位格子を基本とし、並進対称性を維持するように、三次元に規則正しく配列している (図 1)。そのため、基本となる単位格子内の分子の原子座標と格子並進ベクトルより、結晶内のすべての原子座標を表現できる。また、単位格子内の分子配置は空間群が定義する対称性により決められるため、単位格子内分子の原子座標は、非対称単位中の分子の原子座標と空間群から生成することが可能である。

CONFLEX の結晶計算は、このような結晶の持つ高い対称性を仮定しており、非対称単位中の分子 (以下、「オリジナル分子」と呼ぶ) と周囲に展開した分子 (以下、「レプリカ分子」と呼ぶ) の構造変化は同期し、結晶の並進対称性は維持される。そのため、結晶構造の安定性を決定する結晶エネルギー (E_{crystal}) は、オリジナル分子の分子内相互作用エネルギー (E_{intra}) と格子エネルギー (E_{lattice}) の和として定義した (式 1)。

$$E_{\text{crystal}} = E_{\text{intra}} + E_{\text{lattice}}, \quad E_{\text{lattice}} = \sum_i^N \sum_j^{N(M-1)} E_{\text{inter}_{ij}} \quad (\text{式 1})$$

ここで、 N はオリジナル分子を構成する原子数を示し、 M は計算する結晶に含まれる分子数を示す。 E_{lattice} はオリジナル分子の原子 i と、レプリカ分子の原子 j との間に働く分子間相互作用エネルギーの総和である。CONFLEX の結晶計算に要する時間の 9 割以上は、分子間相互作用の計算であるため、この計算を並列分散化することで大幅な計算時間の短縮が期待できる。

効率的に並列分散化を行うためには、データの通信量および通信回数の削減と適切な負荷分散が重要となる。そこで、あらかじめ結晶構造を構築するために必要なデータ (空間群や格子並進ベクトルなど) を各プロセスに保持させるようにした (図 2: B-B')。そして、分子間相互作用の計算を行う際、Master プロセスはオリジナル分子の原子座標と格子定数、および三次元空間に展開する全単位格子に付加した通し番号 (以下、「Cell ID」と呼ぶ) のみを Worker プロセスに渡す (図 2: C-C')。Worker プロセスは、保持する結晶構造情報を元に、Cell ID で指定された単位格子内の原子座標を生成し、分子間相互作用ポテンシャルに関するエネルギーや、一次微分係数、

あるいは二次微分行列の計算を行う (式 2, 図 2: D)

$$E_{\text{worker}} = \sum_{l \in L} E_l, E_l = \sum_i^N \sum_j^{NK_l} E_{\text{inter}_{ij}} \quad (\text{式 2})$$

$$E_{\text{lattice}} = \sum_p^P E_{\text{worker}_p} \quad (\text{式 3})$$

ここで, L は Master プロセスから渡される Cell ID の集合であり, K_l は Cell ID l の単位格子に含まれる分子数を示す. 渡される Cell ID の数は, 各 Worker プロセスにおいて等しくなるようにした. Worker プロセスは計算が終了した後, 求めた値を Master プロセスに返信する (図 2, E-E'). そして, Master プロセスは, Worker プロセスから受け取った値を用いて, 全分子間相互作用に関するエネルギー, 一次微分係数, 二次微分行列を求める (式 3, 図 2: F). Master-Worker 間のデータ通信には Message-Passing Interface (MPI) を使用した.

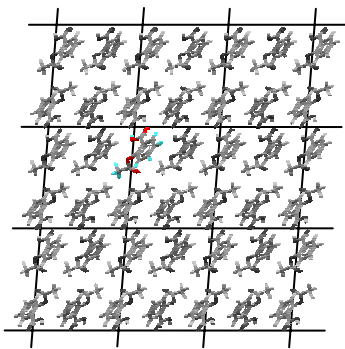


図 1 結晶構造

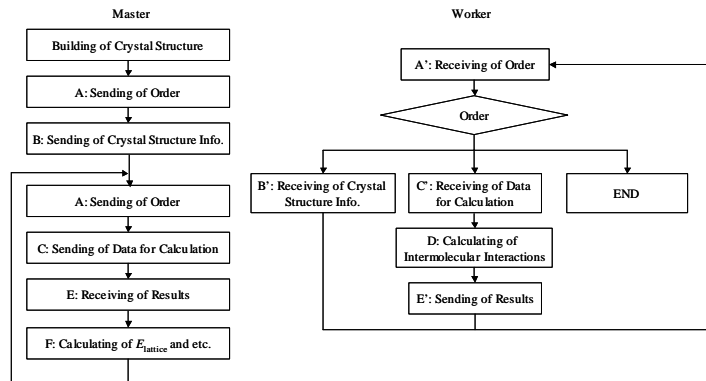


図 2 結晶計算の並列分散アルゴリズム

【結果と考察】

本手法を CONFLEX に導入し, 結晶半径 $0.1 \mu\text{m}$ のアスピリン結晶を用いて, 分子間相互作用ポテンシャルに関するエネルギー, 一次微分係数, 二次微分行列をそれぞれ 3 度求めた際の計算時間とその並列化効率について調べた (図 3). 図 3 から分るとおり, 本手法における並列化効率はとても高く, 31Workers までは理想値と一致している. 31Workers 以降に見られる並列化効率の低下は, Worker プロセスが増えることで, 1 プロセスの計算負荷が小さくなったためであり, 構築する結晶の半径を大きくし, 計算負荷を増やすことで, 理想値にほぼ重なるグラフが得られると考えられる.

【謝辞】

本研究の一部は, 科学研究費補助金の支援を受けて行われています. また, 並列計算は東京工業大学の西川武志特任助教授のもと, スーパーコンピュータ TSUBAME を用いて行いました.

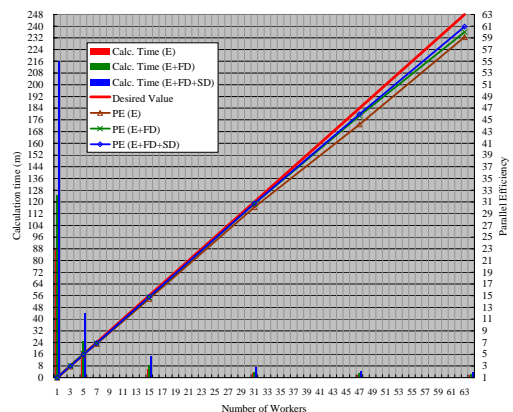


図 3 計算時間と並列化効率