

共同利用(産業利用トライアルユース: 先端研究施設共用促進事業『みんなのスパコン』TSUBAME によるペタスケールへの飛翔) 成果報告書 平成 21 年度 課題種別 i07qd

CONFLEX を用いた配座探索および結晶多形解析
Conformation and Crystal Polymorphism Analysis Using CONFLEX

大田 一男
Kazuo Ohta

コンフレックス株式会社
CONFLEX Corporation
www.conflex.co.jp

分子性結晶の結晶多形現象の解析や結晶多形スクリーニングを実施するために、分散処理技術を用いて結晶計算と配座空間探索法を改良し、高速な結晶構造予測法を開発した。並列化した結晶計算は、127 workers を利用した結晶構造最適化計算において 123 倍の高速化を実現した。また、タンパク質の安定コンホメーションを可能にする配座空間探索法の高速化は、31 workers を利用して 30 倍にまで到達した。これらの改良を行った結晶構造予測法を医薬品化合物であるマレイン酸ヒドラジドに適用したところ、結晶多形として存在する可能性の高い、2 種類の新しい結晶構造を得た。

この研究では、410 種類の化合物について 910 種類の結晶多形構造の結晶構造を最適化し、その結晶構造の座標や分子力場に基づく結晶エネルギー値などをデータベースとして蓄積した。

The method of parallelized crystal calculation and crystal structure prediction were developed for performing analysis of crystal polymorphism and polymorph screening. This parallelized method showed 123 times speed up at 127 workers in crystal structure optimization. Polymorph screening simulation of maleic hydrazide was performed using the prediction method, and two newly crystal structures were predicted theoretically. And database, which was contained optimized crystal structures and its crystal energies of 910 known polymorphs, was constructed.

And, parallelized conformation search method was developed. This method showed 30 times speed up at 31 workers in conformation search.

Keywords: drug, polymorphism, polymorph screening, crystal structure prediction, protein

背景と目的

新規医薬品開発では、疾患に関係するタンパク質の働きを阻害する化合物の発見や創製、選別が十分に行われるが、市販製剤の原薬や投与時の化合物の状態は、多くの場合、結晶である。一般に、医薬品に用いられる化合物は複雑な化学構造を持つことから、複数の結晶形、すなわち結晶多形が頻繁に出現することが知られている。この結晶多形に見られる分子の充填様式の違いは、融点、溶解度、溶解速度、吸湿性、熱容量、自由エネルギーなど、結晶が与える物理的、および熱力学的特性に大きく影響する。そのため医薬品開発において、結晶多形現象の解析や制御は、薬剤の安定性や生産性、有効性の鍵を握る重要な研究課題になっている。ある医薬品候補化合物の有意に存在する結晶構造や、それが示す物性の予測において、シミ

ュレーションによる解決の糸口が付けられ、データベース化できれば、課題の解決に対して強力なツールとなり得る。

また近年、糖尿病の治療薬であるインスリンや抗ウイルス薬であるインターフェロンなど、特定の機能を持つタンパク質を利用したタンパク質医薬品の産業が、バイオテクノロジーの発展に伴い、急速に成長し続けている。タンパク質医薬品は、低分子有機化合物を利用した医薬品では得ることのできない薬効を示すことが期待でき、かつ副作用も少ないため非常に注目されている。こうしたタンパク質医薬品の開発では、生体内で有意に存在するタンパク質の立体構造とその変位過程(以下、フォールディング過程)を明らかにし、タンパク質の構造と薬効を解明することが重要な研究課題として挙げられる。医薬品候補タンパク質の生体内におけ

る立体構造やフォールディング過程を明らかにするためには、そのタンパク質における配座解析が必要不可欠であり、高速な配座空間探索法の開発が望まれる。

本プロジェクトでは、高速な結晶多形解析を実施可能とする、並列結晶計算法を開発した。一般に、結晶を対象としたシミュレーションでは、膨大な数の分子間相互作用の計算を含むため、多くの計算時間を必要とする。一方、医薬品候補化合物の有意に存在する結晶構造を予測するといった、結晶多形スクリーニングでは、多数の試行結晶構造をすべて評価する必要がある。そのため、各試行結晶構造を高速に評価することができる結晶計算法の開発が必要であった。本プロジェクトより開発した並列結晶計算法は、127 Workers を利用した結晶構造最適化において、総計算時間で 123 倍の性能向上を実現した。また本手法を用いて、医薬品化合物 76 種、その結晶多形 180 構造を含む、有機化合物全 419 種、910 結晶構造で構成した結晶多形データセットについて、並列結晶構造最適化を行い、それらの最適化結晶構造や、結晶多形間の安定性評価に重要な結晶エネルギー等をデータベース化した。さらに、結晶多形スクリーニングを実施可能とする結晶構造予測法を開発し、医薬品化合物であるマレイン酸ヒドラジドに適用した。その結果、結晶多形として存在する可能性の高い、新しい 2 種類の結晶構造を予測した。これらの結晶構造データは、データベースに収録した。

また本プロジェクトでは、タンパク質のような、膨大な数の配座異性体を持つ物質の配座解析を可能にする、並列配座空間探索法を開発し、31 Workers を利用した配座空間探索において、30 倍の性能向上を実現した。

概要

配座空間探索

独自の配座空間探索法である CONFLEX 法の基本的なアルゴリズムは次のとおりである。まず、初期構造を選択し、これを局所的に微小変形させ、様々な試行構造を発生させる。それぞれの試行構造に対して構造最適化を行い、新しい配座が創出された場合は配座データベースに保存する。そして保存された配座データベースから新しい初期構造を選択して同様の操作を行

う。CONFLEX の配座空間探索において、構造最適化処理は全探索時間の 9 割以上を占めている。このため、この処理を並列分散化し、探索性能の向上を図った。ここでは、総原子数 138 の人口タンパク質 Chignolin (PDB ID: 1UAO) の並列配座空間探索を行い、並列化効率を調べることで性能向上率を検証した。

結晶多形データセット

結晶多形に関する解析を行うため、ケンブリッジ構造データベース(以下、CSD)より、C, H, N, O, F, Cl, S 原子から構成される分子で、結晶多形を持ち、 R 値が 5% 以下の良質な結晶構造データが存在する化合物を抽出した。その後、これらの化合物について、特殊点を持たず、三斜晶系、単斜晶系、斜方晶系のいずれかの晶系で、かつ原子座標情報をもつ結晶構造データを CSD より得た。さらに、これらの結晶構造データを入念に検討することで、解析に適したデータを選び出し、最終的に 419 種の化合物の結晶多形 910 構造(結晶多形ペアとして 488 対)を含む結晶多形データセットを作成した。

結晶計算法

CONFLEX による結晶計算は、非対称単位に含まれるすべての分子(以下、オリジナル分子)と、空間群で定められている対称操作と結晶の並進対称性にしたがって有効結晶半径 D_{\max} まで展開した非対称単位(以下、レプリカ単位)内の分子(以下、レプリカ分子)で構成される球状の結晶を構築する。そして、この球状結晶の結晶エネルギーを次のように評価する: 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、オリジナル分子と、 D_{\max} 以内に展開されたレプリカ分子との間に生じる相互作用エネルギー $E_{\text{inter}}^{\text{crystal}}$ 、および非対称単位内の分子間相互作用エネルギー $E_{\text{inter}}^{\text{a-unit}}$ の総和として定義した。ただし、非対称単位内の分子数が 1 ならば、 $E_{\text{inter}}^{\text{a-unit}} = 0$ である。またここで、 $E_{\text{inter}}^{\text{crystal}}$ と $E_{\text{inter}}^{\text{a-unit}}$ の和は、格子エネルギーと呼ぶ。

$$E_{\text{crystal}} = E_{\text{intra}} + E_{\text{lattice}} \quad (1)$$

$$E_{\text{lattice}} = E_{\text{inter}}^{\text{a-unit}} + E_{\text{inter}}^{\text{crystal}} \quad (2)$$

$$E_{\text{inter}}^{\text{crystal}} = \sum_i^N \sum_S^M \sum_J^N E_{\text{inter } ij} \{ \text{Min}(R_{ij}) \leq D_{\text{max}} \} \quad (3)$$

ここで、 N は非対称単位内の分子を構成する総原子数、 M は結晶空間に展開されたレプリカ単位の数を表す。また、 $\text{Min}(R_{ij})$ はオリジナル分子とレプリカ分子との間の最近接原子間距離であり、 R_{ij} が D_{max} 以内にあるレプリカ分子に関わるすべての非結合相互作用を $E_{\text{inter}}^{\text{crystal}}$ に含めた。有効結晶半径 D_{max} は結晶計算の精度を決定する主要因の一つである。そこで結晶多形データセット内の 61 結晶構造に対して、有効結晶半径 10~200Å の大きさの異なる結晶モデルを構築し、それぞれに対して結晶構造最適化を行い得られた最適化構造のエネルギー変化、および結晶構造変化を調べることで最適な有効結晶半径値を決定した。なお、結晶構造変化は式 4 により評価した。

$$F = \left(\frac{100\Delta a}{a}\right)^2 + \left(\frac{100\Delta b}{b}\right)^2 + \left(\frac{100\Delta c}{c}\right)^2 + \Delta\alpha^2 + \Delta\beta^2 + \Delta\gamma^2 + (10\text{rms}\Delta X_{\text{trans}})^2 + \left(\frac{\text{rms}\Delta\theta}{2}\right)^2 + (10\text{rms}\Delta X_{\text{atom}})^2 \quad (4)$$

ここで、 Δa , Δb , Δc は、それぞれ格子長 a , b , c の変化を示し、 $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$ は、それぞれ格子角 α , β , γ の変化を示す。また、 $\text{rms}\Delta X_{\text{trans}}$, $\text{rms}\Delta\theta$, $\text{rms}\Delta X_{\text{atom}}$ は、それぞれオリジナル分子の重心位置、配向、および全重原子位置の最小二乗差である。

並列結晶計算法

CONFLEX による結晶計算において最も計算量の多い処理は、分子間相互作用の計算であり、結晶構造最適化を行った場合、全計算時間の 9 割以上を占める。そのため、この分子間相互作用計算を並列分散処理することで大幅な計算時間の短縮が期待できる。

効率的に並列分散処理を行うためには、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、結晶が単位格子を最小単位として周期的に展開された構造として表現できること、およびオリジナル分子の原子座標、空間群、格子並進ベクトルに

より、結晶内のすべての原子座標を表現することができるという特徴を考慮した Master-Worker 型の並列分散処理技術を開発し導入した。

結晶構造予測法

分子性結晶の結晶多形スクリーニングを実現する化学計算手法を開発し CONFLEX に導入した。本手法は、単位格子内の全分子、あるいは任意の分子をランダムに回転させ試行結晶構造を生成し、これらをすべて並列結晶構造最適化することで、未知の結晶構造を予測する。ここでは、既知の X 線結晶構造を用いて、これを空間群 P1 の結晶へと変換し、単位格子内の 1 分子に対して x 軸、 y 軸、 z 軸、それぞれで 30° 刻みにより回転させ、試行結晶構造を得た。分子の回転は式 5 により行った。生成した試行結晶構造は、まず回転操作を行った分子の空間位置、および配向を固定した状態で、他の単位格子内の分子位置、配向、および格子定数を最適化した。その後、単位格子内のすべての分子の空間配置、配向、および格子定数について最適化を行った。なお、結晶構造最適化の際は、分子構造を固定している。

$$\mathbf{V}'_{\text{OM}} = \mathbf{e} \mathbf{V}_{\text{OM}} \quad (5)$$

$$\mathbf{e} = \begin{pmatrix} \cos\theta_2 \cos\theta_3 & -\sin\theta_3 \cos\theta_2 & \sin\theta_2 \\ \cos\theta_1 \sin\theta_3 + \sin\theta_1 \sin\theta_2 \cos\theta_3 & \cos\theta_3 \cos\theta_1 - \sin\theta_1 \sin\theta_2 \sin\theta_3 & -\sin\theta_1 \cos\theta_2 \\ \sin\theta_1 \sin\theta_3 - \cos\theta_1 \sin\theta_2 \cos\theta_3 & \cos\theta_3 \sin\theta_1 + \cos\theta_1 \sin\theta_2 \sin\theta_3 & \cos\theta_1 \cos\theta_2 \end{pmatrix} \quad (6)$$

ここで、 \mathbf{V}_{OM} は、分子の重心から、その分子を構成する原子へのベクトルを示す。また \mathbf{e} は、オイラー角による回転行列を示す。

力場

分子内ポテンシャルとして、ここでは CONFLEX の標準力場である Merck Molecular Force Field (以下、MMFF94s) を用いた。また、分子間ポテンシャルには、MMFF94s の非結合相互作用関数とパラメータを用いた。

結果および考察

並列配座空間探索と並列化効率

Chignolin の配座空間探索に要した総計算時間、全試行構造の最適化に要した時間、およびそれらの並列

(様式第 20) 成果報告書

化効率について Figure 1 に示した。Figure 1 から明らかのように、1 台の Worker ホストによる配座空間探索において、試行構造の最適化に要する時間は、総計算時間の 98%にもなる。そのため、試行構造の最適化処理に対して施した並列化は非常に効果的で、31 並列計算において 30 倍の高速化を実現しており、十分な並列化効率を示した。一方、総計算時間では、並列化効率の著しい低下がみられた。これは Worker ホストを増やしたことで、全探索時間における構造最適化時間の割合が減少したことが主な原因として考えられる。

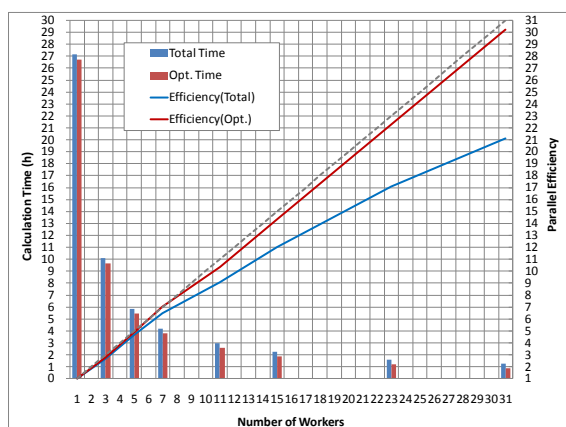


Figure 1 Calculation time and parallelization efficiency of conformational search for chignolin

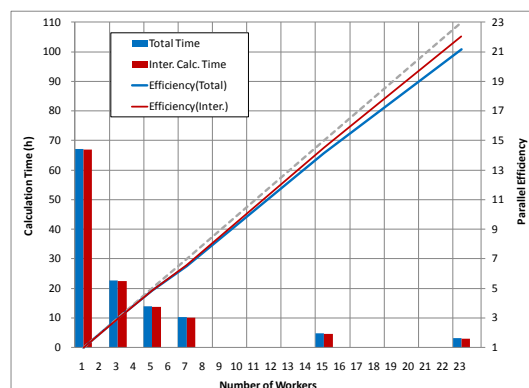
並列結晶計算と並列化効率

結晶半径 200Å のアスピリン結晶 (CSD Refcode: ACSALA01) の結晶構造最適化に要した総計算時間、分子間相互作用の計算時間、およびそれらの並列化効率について Figure 2a に示した。また同様に、非対称単位内の総原子数 228 のペプチドの結晶 (CSD Refcode: XERWOH) について Figure 2b に示した。

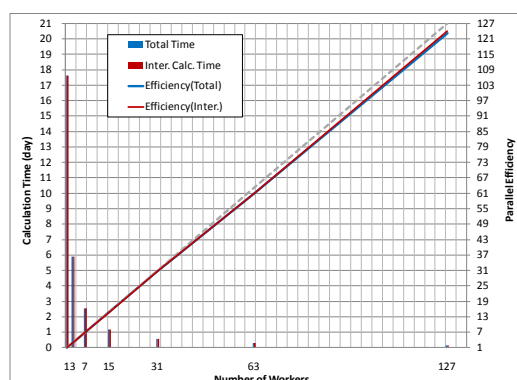
Figure 2a より、アスピリン結晶の構造最適化では、23 並列計算において、総計算時間で 21 倍、分子間相互作用計算時間で 22 倍の高速化を実現した。また、Figure 2b より、ペプチドの結晶の構造最適化では、127 並列計算において、総計算時間で 123 倍、分子間相互作用計算時間で 124 倍の高速化を実現した。

Figure 2 より、並列結晶計算における並列化効率の減少率がほぼ線形であることから、Worker ホスト用に利用可能な Core 数と演算性能に対して十分な分子間

相互作用の計算量があれば、確実に高い並列化効率が望めることが示唆される。



a.



b.

Figure 2 Calculation time and parallelization efficiency of parallelized crystal calculation, a. aspirin crystal, b. peptide crystal.

有効結晶半径の検討

結晶多形データセットに含まれる 61 結晶構造について、有効結晶半径 10~200Å の範囲で大きさの異なる結晶モデルを構築し、それぞれに対して結晶構造最適化を行い得られた最適化構造の E_{intra} , $E_{lattice}$, および $E_{crystal}$ の変化を Figure 3 に示した。ただし、各エネルギーは有効結晶半径 200Å 時の最適値からの差として示した。

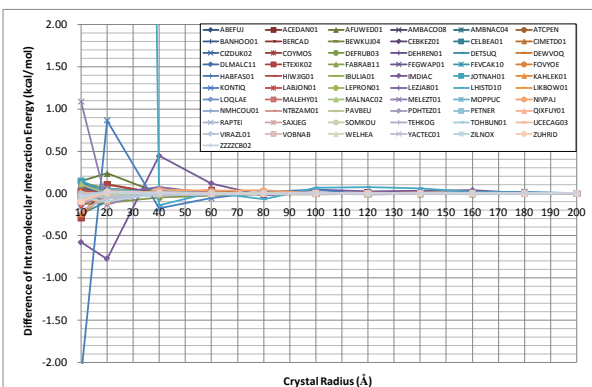
Figure 3 より、結晶モデルの規模に対する $\Delta E_{crystal}$ の変化は $\Delta E_{lattice}$, すなわち分子間相互作用エネルギーの寄与が大きいことが分かる。一方、 ΔE_{intra} の変化は、多くの化合物において、有効結晶半径 20 Å 付近から徐々に収束に向かい、60~70 Å 以上では $\Delta E_{crystal}$ の変化にほとんど寄与していない。一方、 $\Delta E_{lattice}$ の変化は 70~80 Å 付近から収束に向かい、180 Å まで結

(様式第 20) 成果報告書

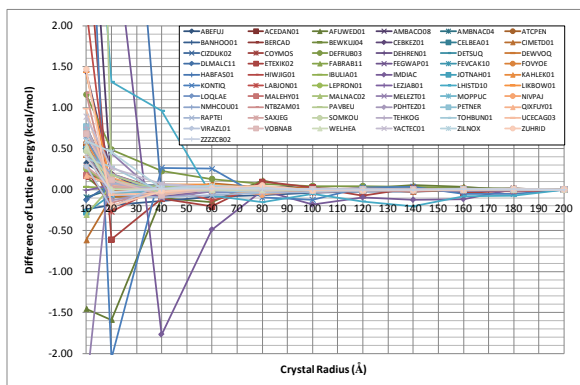
晶を大きくすると、200 Å との差は非常に小さくなる。

各最適化結晶構造における X 線結晶構造との再現誤差を式 4 より求め、Figure 4 にまとめた。ただし、再現誤差は有効結晶半径 200Å 時の値からの差として示した。結晶モデルの規模に対する結晶構造の変化は、有効結晶半径 40 Å 付近から徐々に収束に向かい、120 Å 以上ではほとんど変化は見られない。

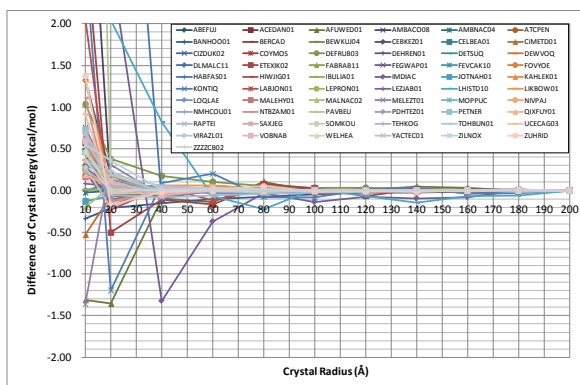
本プロジェクトでは、結晶多形データセットに対する結晶計算に要すると予想される総計算時間を考慮し、結晶多形解析のための有効結晶半径として、分子内相互作用エネルギーで 10⁻³kcal/mol 以内、格子エネルギー、および結晶エネルギーで 10⁻²kcal/mol 以内、結晶構造変化で 10⁻¹ 以内の精度での定量的な解析が可能で、80Å を採用することにした。



a.



b.



c.

Figure 3 Energy differences depending on crystal radius, a. ΔE_{intra} , b. $\Delta E_{lattice}$, c. $\Delta E_{crystal}$.

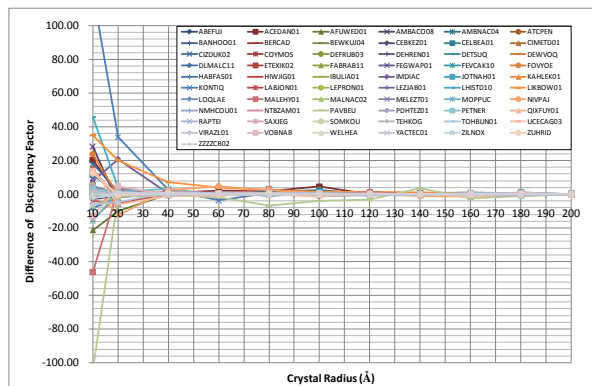


Figure 4 Differences of discrepancy factor for crystal structure depending on crystal radius.

結晶多形データセットにおける結晶多形解析

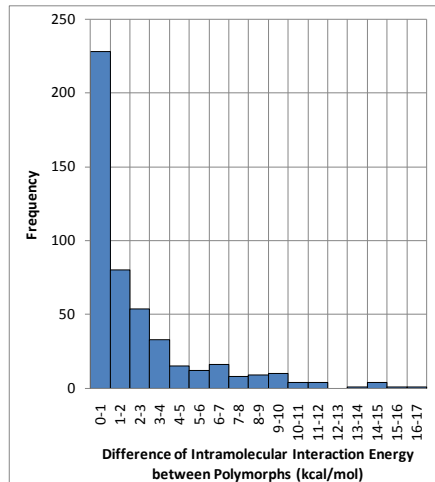
結晶多形データセットに含まれる結晶構造 910 に対し、結晶構造最適化を行い得られた最適化構造について、多形間の $|\Delta E_{intra}|$, $|\Delta E_{lattice}|$, $|\Delta E_{crystal}|$, $|D|$ (D は密度を示す。)を調べ、頻度分布を作成した (Figure 5)。ただし、比較する結晶多形は、互いに z' が一致するもので行い、最安定な結晶エネルギーを示す結晶構造からの差を求めている。

Figure 5c より、 $\Delta E_{crystal}$ において、3kcal/mol 以内のエネルギー差を示す結晶多形は、全体の 6 割を占め、10kcal/mol 以内では全体の 9 割を示す。また Figure 5d より、 $|D|$ において、2% 以内の密度差を示す結晶多形は、全体の 6 割を占め、5% 以内では 9 割を占めている。つまり、結晶多形として有意に存在する結晶構造は、最安定結晶構造の結晶エネルギーから 10kcal/mol 以内の差を持ち、かつ密度差が 5% 以内であり、それを超えることは稀であることが分かった。また $\Delta E_{crystal}$ と $|\Delta E_{lattice}|$ は、ほぼ同じ頻度分布であり、 $|\Delta E_{intra}|$ に比べ、ブロードな分布を示している。これは、結晶内の分子の空間配置、および配向の自由度が非常に多いことに由来すると考えられ、分子性結晶ではこのようなブロードな頻度分布が得られるものと思われる。

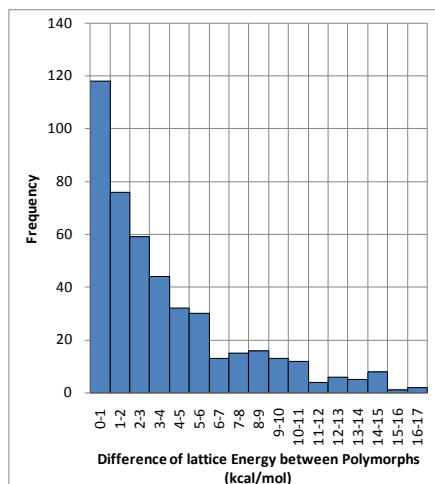
一方、 $|\Delta E_{intra}|$ は、0~1kcal/mol 差において局所的なピークが見られ (Figure 5a)、1kcal/mol 以内のエネルギー差を示す結晶多形は全体の 5 割を占めている。ま

(様式第 20) 成果報告書

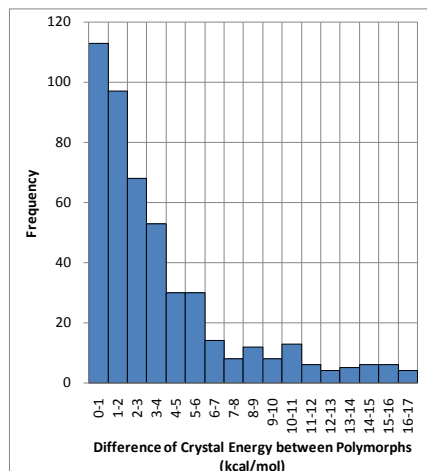
た、 $|\Delta E_{\text{intra}}|$ が 4kcal/mol 以内である結晶多形では、全体の 8 割を示す。これは、結晶内における分子内構造の自由度が非常に小さく、約 4kcal/mol 程度の構造変化しか許容できないこと、またエネルギー的に安定な配座異性体が結晶化するため、結晶内に見られる配座異性体間のエネルギー差は非常に小さいことを示していると考えられる。



a.

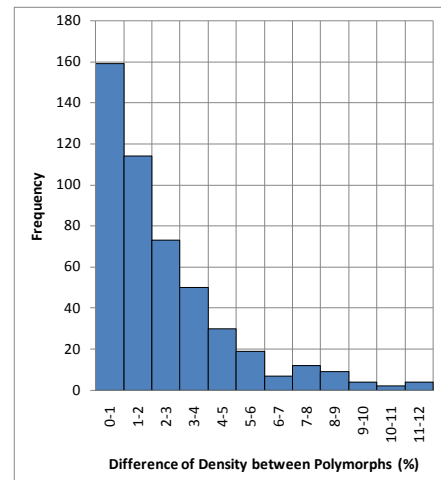


b.



c.

Figure 5 Frequency distributions for energy and density differences between polymorphs, a. $|\Delta E_{\text{intra}}|$, b. $|\Delta E_{\text{lattice}}|$, c. $\Delta E_{\text{crystal}}$, d. $|\Delta D|$.



d.

Figure 5 continued.

結晶多形スクリーニング

配座異性体の存在しないマレイン酸ヒドラジド (Maleic Hydrazide:MH, Figure 6)を用いた結晶多形スクリーニングシミュレーションを実施した。MH は、生物学的特性を示す物質であり、プリン誘導体、あるいはピリミジン誘導体として振る舞うことができる。また、MH には 3 種類の結晶多形 MH1(空間群: $P\bar{1}$, $z'=1$, $a=5.8181(10)$, $b=5.800(1)$, $c=7.309(1)$, $\alpha=78.80(3)$, $\beta=99.36(3)$, $\gamma=107.13(3)$), MH2(空間群: $P2_1/c$, $z'=1$, $a=6.891(2)$, $b=9.674(2)$, $c=6.946(1)$, $\alpha=90$, $\beta=100.07(2)$, $\gamma=90$), MH3(空間群: $P2_1/n$, $z'=1$, $a=6.6070(15)$, $b=6.9070(7)$, $c=10.539(3)$, $\alpha=90$, $\beta=104.00(4)$, $\gamma=90$)が存在することが分かっている。

Figure 6 Structural formula of maleic hydrazide.

MH の結晶多形スクリーニングを行うため、MH3 の結晶を空間群 $P1$ の結晶へと変換し、また単位格子内

(様式第 20) 成果報告書

の 1 分子を x 軸, y 軸, z 軸, それぞれで 30° 刻みにより回転させ, ユニークな 744 の試行結晶構造を得た. これらの試行結晶構造に対し並列結晶構造最適化を行った. その後, 得られた最適化構造において最安定結晶エネルギーから 10kcal/mol 以内の結晶構造について, PLATON により空間群を再定義し, これらの結晶モデルについて再度, 結晶構造最適化を行った. この際, 単位格子内の分子の空間配置, 配向, および格子定数以外に, 分子構造も同時に最適化した.

Table 1 は, 結晶多形スクリーニングにより創出された結晶構造についてまとめた. ここで, 多形スクリーニングにより得られた結晶構造について, 3 種類の既知の結晶構造と比較したところ, 2 は MH1, 5 は MH2, 4, 6 は MH3 の結晶構造とよく一致した. つまり, 本多形スクリーニングにおいて, 試行結晶構造を生成するため

に利用した MH3 に限らず, MH のすべての既知の結晶構造を得ることができた. また興味深いことに, 既知の結晶構造とは異なり, かつ最安定結晶エネルギーを示す空間群 $P2_1/c$ の結晶構造 1 が得られた (Table 1). さらに, MH2, MH3 の結晶構造を示す, 4, 5, 6, よりも低い結晶エネルギーを示す, 空間群 $P\bar{1}$ の結晶構造 3 が得られた (Table 1).

Table 1 に示した 1~6 の結晶構造は, 最安定結晶構造の結晶エネルギー (E_{crystal}/z') との差が 3kcal/mol 以内であり, また密度差が 4% 以内に収まる. 結晶多形データセットを用いた結晶多形間の結晶エネルギー差, および密度差を考慮すると, 1 と 3 は結晶多形として存在する可能性が高いと考えられる.

Table 1 Predicted crystal structures of maleic hydrazide.

Crystal			Energy (kcal mol ⁻¹)		Density (mg m ⁻³)	Lattice parameters					
Idx.	Space group	z'	E_{crystal}	E_{crystal}/z'		$a(\text{\AA})$	$b(\text{\AA})$	$c(\text{\AA})$	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$
1	$P2_1/c$	1	-47.82	-47.82	1.471	7.542	9.587	7.824	90	116.59	90
2	$P\bar{1}$	1	-46.99	-46.99	1.518	5.844	6.121	7.404	77.85	79.74	72.48
3	$P\bar{1}$	1	-46.85	-46.85	1.498	5.965	6.309	7.441	89.64	70.43	71.29
4	$P2_1/c$	1	-46.11	-46.11	1.491	6.880	7.067	10.866	90	109.26	90
5	$P2_1/c$	1	-45.78	-45.78	1.471	7.093	10.477	9.423	90	133.75	90
6	$P\bar{1}$	2	-89.65	-44.83	1.491	6.921	7.067	10.806	88.02	71.08	86.67
7	P1	4	-178.87	-44.72	1.443	10.156	7.147	7.207	88.06	90.27	99.51
8	P1	4	-178.85	-44.71	1.483	9.941	7.086	10.796	88.45	71.36	132.15
9	$P\bar{1}$	2	-88.85	-44.42	1.436	6.935	9.074	8.987	89.96	77.88	109.97
10	$P\bar{1}$	2	-88.79	-44.40	1.476	7.064	7.148	10.769	79.07	74.82	75.69
11	P1	4	-176.60	-44.15	1.477	7.107	7.071	10.113	88.21	94.31	96.01
12	P1	4	-176.33	-44.08	1.466	7.328	7.112	12.172	53.21	87.89	90.22
13	P1	4	-175.00	-43.75	1.483	7.126	7.017	10.275	89.47	94.55	101.58

まとめ、今後の課題

本プロジェクトでは, 並列配座空間探索法を開発し, 31 Workers を利用した配座空間探索において, 30 倍の高速化を実現した. また, 並列結晶計算法を開発し, 127 Workers を利用した結晶構造最適化において 123 倍の高速化を実現した.

並列結晶計算法, および結晶多形データセットを用いた結晶多形解析により, 最安定結晶構造の結晶エネルギーから 10kcal/mol 以内の差を持ち, かつ密度差が 5% 以内である結晶構造が, 結晶多形として有意に存在することが分かった. 結晶内の分子内構造は, 約 4kcal/mol 程度の構造変化しか許容できず, またエネ

(様式第 20) 成果報告書

ルギー的に安定な配座異性体が結晶化するため、結晶内に見られる配座異性体間のエネルギー差は非常に小さいことが分かった。

医薬品化合物であるマレイン酸ヒドラジドの結晶多形スクリーニングシミュレーションを行った結果、結晶多形として存在する可能性の高い、既知の結晶構造とは異なる新しい 2 種類の結晶構造を予測した。

結晶多形データセットを構成する、医薬品化合物 76 種、その結晶多形 180 構造を含む、有機化合物全 419 種、910 結晶構造について行った並列結晶構造最適化により得られた、最適化結晶構造や結晶多形間の安定性評価に重要な結晶エネルギー等を収録したデータベースを生成した。また、本プロジェクトにより予測したマレイン酸ヒドラジドの新しい 2 種類の結晶構造データもまた、データベースに収録した。

今後、継続的に医薬品候補化合物を対象とした結晶多形スクリーニングシミュレーションを実施し、その結果を本データベースに収録することで、その価値を高めてゆく。本プロジェクトにより開発した、並列配座空間探索法、および並列結晶計算法、結晶多形スクリーニング法、結晶多形データベースが、医薬品開発において活用されることで、実験化学的、および計算化学的手法による結晶多形スクリーニングや物性予測等に要する費用の節約につながり、将来的に開発現場で数億円から数十億円規模の効率化が期待できると考えている。