

高分子材料開発におけるシミュレーションと MI の活用

坂 下 竜 一*¹

Simulation and Materials Informatics on Polymer Research

Ryuichi SAKASHITA

In spite of the industrial importance of polymer materials, the majority of practical examples of materials informatics are for inorganic compounds and low molecular weight compounds. The multi variability and complex correlations based on hierarchical structures seem to make informatics on polymer materials difficult. Computational simulations such as molecular dynamics and ab initio calculation are able to provide homogeneous data for informatics, thus they are useful as data generator or descriptor generator for polymer materials. Herein several examples of synergistic research between conventional simulations and informatics are reported.

1. はじめに

機械学習等のインフォマティクスを化学分野の研究に応用する試みは、主に小分子を対象とした薬学（ケモインフォマティクス）で大きな発展を見たのち、続いて一般の材料開発に展開されてきた（マテリアルズ・インフォマティクス）。しかしながら、マテリアルズ・インフォマティクスにおいても、多くの実施対象は無機化合物や低分子化合物であり、高分子を中心としたソフトマターへの適用は高分子材料の工業的重要性のわりにそれほど進んでいないのが実情である。

この原因は多岐にわたるであろうが、大局的には高分子特有の階層的複雑性に由来したものであると考えられる（Fig. 1）。高分子材料の階層構造においては、それぞれのオーダーで物性に相関する変数を持つため、高分子材料の設計変数は一般に多数にわたる。例えば、固さ（弾性率）という比較的単純な物性であっても、一次構造に加えて分子鎖の絡み合いの影響を受けるために、分子量とその分布、特性比、熱履歴といった変数が存在する。さらに、結晶性高分子の場合は、

これに加えて結晶構造や結晶サイズ、結晶化度、なども変数となりうる。より高次の構造として、ブロック共重合体やポリマーブレンドの場合は相分離構造、フィラー等を含有するナノコンポジットの場合はフィラー配合や界面組成、分散材などが考慮に入れられるだろう。その上、これらの設計変数は物性に対して非線形性や多重共変性を持つことが多く、相対的な実験データの少なさも相まって、高分子材料の統計的分析や機械学習は困難であることが多い。

この問題に対する一つのアプローチとして、シミュレーションの活用があげられる。従来から高分子研究に用いられてきた、第一原理計算や粗視化分子動力学、有限要素法など各種のシミュレーション技術は、特定のスケールに着目した網羅的なデータを大量に生成可能であり、このようなデータは機械学習や統計解析に親和性が高い。以下の事例で紹介するように、機械学習モデルを用いて広域スクリーニングを行うほかにも、シミュレーション値を高次構造の説明変数（記述子）として用いたり、シミュレーションデータで学習させたニューラルネットワークを転移学習の転移元として活用したりすることができる。

統合シミュレータ OCTA¹⁾ は高分子の各階層構造に対応したエンジンを取り揃えており、python による

*1 ファンクショナルポリマー研究所 光学材料グループ

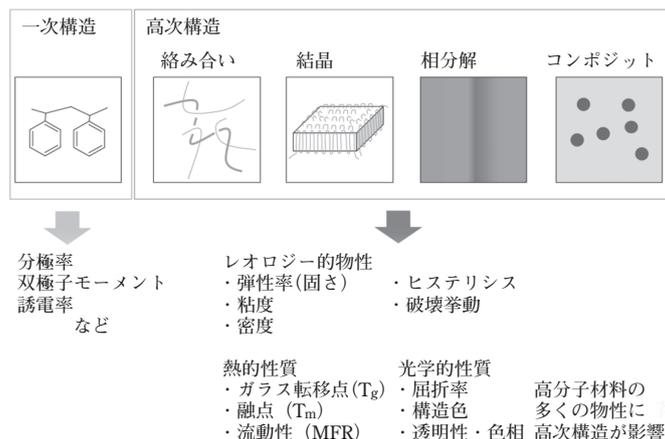


Fig. 1 高分子の階層構造と物性の関係

入出力 API を備えることから、同じく python を用いた機械学習と一体のコーディングも可能である。

2. シミュレーションデータの機械学習とスクリーニング

[1] ブロック共重合体の弾性率推算

ブロック共重合体は、各モノマーが分離するために微細な規則構造（マイクロ相分離構造）を形成する。この構造はマクロ物性に大きな影響を与えるため実用上重要であり、相分離構造及びその物性のシミュレーションは広範に研究されてきた²⁾。濃度場を用いた自己無撞着場（Self-Consistent Field, SCF）法・全原子モデルや粗視化モデルに基づく分子動力学（Molecular Dynamics, MD）・鎖のすり抜けを許すことで MD より高速に再安定構造を計算できる散逸粒子動力学（Dissipative Particle Dynamics, DPD）³⁾などは、マイクロ相分離構造を良好にシミュレーション可能である。

得られた構造を有限要素法（Finite Element Method, FEM）などで解析することで、弾性率や光線透過率、熱物性など各種の物性も推算できる。

しかし、共重合体の配列やモノマーの組み合わせは莫大に存在するにもかかわらず、シミュレーション資源は限られていることから、関心のある領域を網羅的にスクリーニングするのは一般的に困難とされてきた。

そこで、ブロック共重合体の一次構造（モノマー配列・相互作用パラメータ）とシミュレーションされた弾性率の相関を学習したモデルを用いて、弾性率を高速に推算可能か検討した⁴⁾。スキームを Fig. 2 に示す。

ランダムに生成したブロック共重合体について、DPD によりブロック共重合体の相分離構造を生成し、FEM によりそれらの弾性率（体積弾性率 K 、ずり弾性率 G ）を求めて教師データとした。DPD 及び FEM シミュレーションには OCTA システムを用いた¹⁾。

機械学習には RBF カーネルを用いた非線形サポー

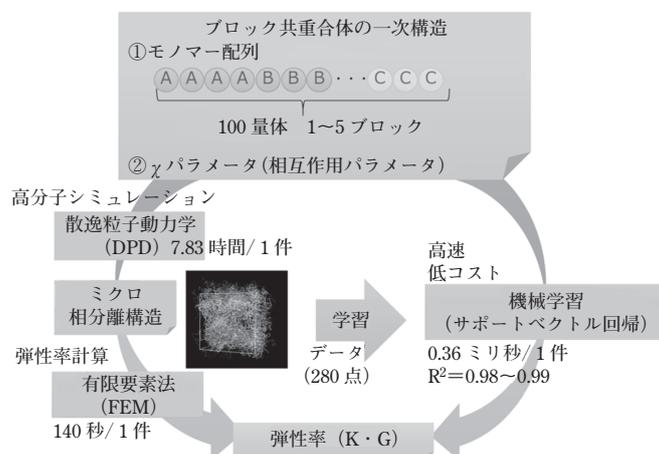


Fig. 2 機械学習によるブロック共重合体の物理推算

Table 1 実行時間

工程	計算時間
DPD	7.83 時間 / 1 組成 (4 コア)
FEM	140 秒 / 1 組成 (1 コア)
SVR	0.36 秒 / 1000 組成 (1 コア)

トベクトル回帰 (Support Vector Regression, SVR) ⁵⁾ を用いた。SVR によるモデルはシミュレーション結果を $R^2 = 0.99$ (体積弾性率 K) と、ほぼ同等精度で予測できることが確認された。また、実行速度を比較したところ、シミュレーションに対して SVR による推算は極めて高速であり (Table 1)、1000 組成の推算を 1 秒以下で実行できた。これにより任意の弾性率を満たす共重合配列の候補を容易に見つけることができる。

[2] 分岐高分子の伸長粘度推算

ラミネート加工やブロー成形等の成形過程において、樹脂は伸長変形される。この際、分岐構造を持つ高分子では絡み合いの寄与から粘度が大きく上がることが知られており (ひずみ硬化性)、この効果を利用して安定した成形が可能となる。このため、伸長ひずみに対する粘度変化 (伸長粘度) はポリオレフィン等の高分子材料の成形加工性を特徴づける重要なレオロジー物性である。

そのため、伸長粘度を予測することは応用上重要であり、シミュレータが盛んに研究されてきた。スリップリンクモデルは、絡み合いに関係した量のみで系の状態を記述する為に計算負荷が小さく、長い緩和挙動までシミュレーションできる。中でも、増漕らが開発した PCN (Primitive Chain Network) モデル ⁶⁾ は、絡み合い点間分子量に相当するセグメントからなるフ

ントム鎖と、それらからなるネットワークを用いて系を記述している点の特徴である。この手法の利点として、pompom 型 (一点より広がる多数の側鎖を持つ) や櫛型など多様な分岐高分子を表現できることが挙げられる ⁷⁾。

しかしながら、分岐高分子の分岐パターンは、そこに分子量やその分布、分岐様式自体の分布等の諸要素まで含むと莫大な組み合わせが考えられ、スリップリンクモデルの高速性をもってしても網羅的に計算検討を行うことは容易ではなかった。

そこで、PCN モデルによるシミュレーション結果を機械学習することで、分岐構造と伸長粘度の関係について大規模な検討を試みた (Fig. 3) ⁸⁾。

ポリオレフィン樹脂の改質問題のモデルとして、広い分子量分布を持つ直鎖高分子に対して少量の櫛形高分子をブレンドした系を用い、様々な分岐様式 (主鎖長・側鎖長・分岐数等) を持つ分岐高分子を少量 (~5vol%) 添加したとき、どのような分岐様式が最も伸長粘度を向上させるか調べた。

以下に探索範囲を示す。分子量は下記式(1)で無次元化されている (M は分子量、 M_e は絡み合い点間分子量を表す)。

$$Z = M/M_e \times 1.6 \tag{1}$$

A. 直鎖

主鎖分子量: $Z = 1 \sim 100$

($\log Z$ に対して正規分布 $\sigma = 0.4$)

B. 櫛形鎖

形状: 櫛形分岐高分子 (pompom 型も含む)

総分子量 (totalZ) : $Z = 20 \sim 100$

側鎖分子量 (armZ) : $Z = 1 \sim 98$

分岐点あたり側鎖数 (armN) : $1 \sim 2$

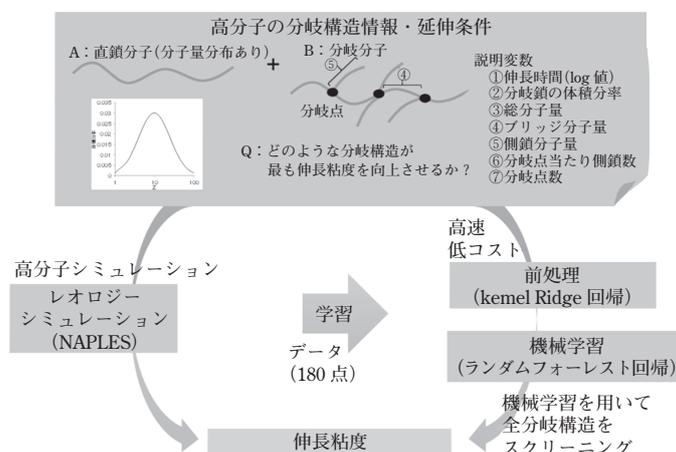


Fig. 3 分岐高分子の伸長粘度推算

分岐点数 (junctionN) : 2 ~ 10

分岐鎖の体積分率 (volume ratio) : 0.005 ~ 0.05

PCN シミュレーションには OCTA システム¹⁾ の NAPLE⁹⁾ を用いた。

PCN シミュレーション結果を前処理 (時間・粘度双方の対数を取ったのちスムージング) の後、ランダムフォレスト回帰を行った。説明変数には伸長時間、分岐鎖の体積分率、分岐構造パラメータ (総分子量・ブリッジ分子量・側鎖分子量・分岐点当たりの側鎖数・分岐点数) を用いた。系の広さに対して小さい 180 点のデータを用いたにもかかわらず、ランダムフォレスト回帰は $R^2 = 0.94$ の十分な精度のモデルを与えた。

機械学習モデルを用いてスクリーニングを行うことにより、伸長粘度を効率的に向上させる分岐様式は添加量と相関することが明らかになった (Fig. 4)。分岐鎖が直鎖に対して少量の場合、分岐鎖同士の絡み合いが無視でき、最も緩和時間が長くなる短・多側鎖の

櫛形が有効である。他方、大量添加系では分岐鎖の広がりが大きく分岐鎖同士の絡み合いの寄与が支配的であるため、ネットワークを形成しやすい pompom 型が有効である。

このように、機械学習により少数のシミュレーションデータから短時間での網羅的スクリーニングが可能となり、膨大な組み合わせの中から傾向を読み解くことができた。

3. シミュレーションと遺伝的アルゴリズムの併用

遺伝的アルゴリズム (GA) とシミュレーションを併用することで、任意の複数の物性の最大化あるいは最小化を行うこともできる (多目的最適化)。GA は世代ごとと繰り返しの評価を必要とするが、これを簡便に実施できる点でシミュレーションと相性が良い。

Python のケモインフォマティクス用ライブラリ

	少量添加(0.5vol%)	大量添加(0.5vol%)
最も伸長粘度を上昇させる分岐構造	総分子量 : 98 側鎖分子量 : 2 分岐点当たり側鎖数 : 1 分岐点数 : 7  多分岐数・長い主鎖	総分子量 : 99 側鎖分子量 : 23 分岐点当たり側鎖数 : 2 分岐点数 : 2  小分岐数・短い主鎖≒pompom 型
想定されるメカニズム	分岐鎖同士の絡み合いはない →最も緩和時間が長い構造 	分岐鎖同士の絡み合いが支配的 →最も効率的に絡合う pompom 型 

Fig. 4 伸長粘度向上に最も効果のある分岐構造

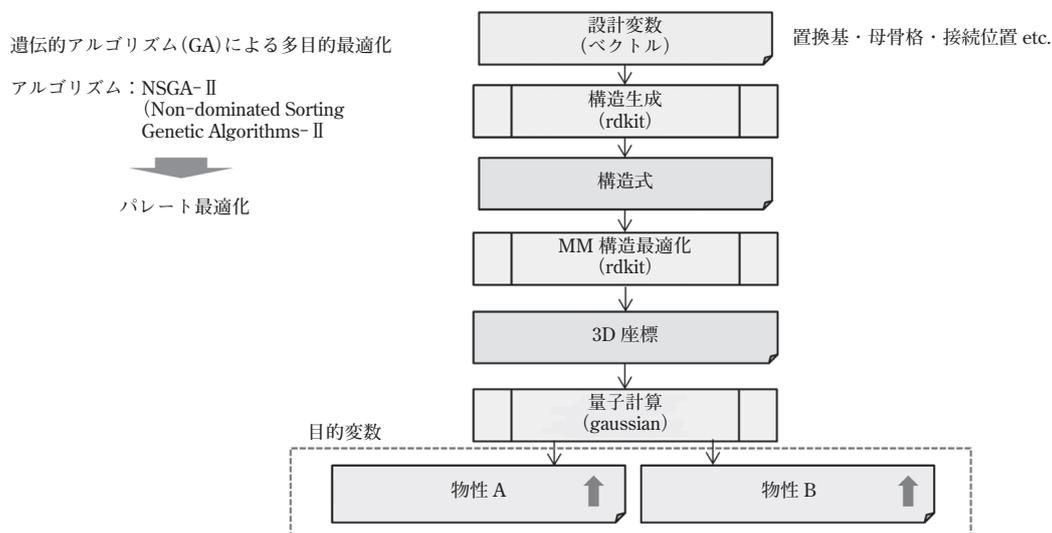


Fig. 5 遺伝的アルゴリズム (GA) と量子計算による多目的最適化

RDKit¹⁰⁾ を用いた構造生成プログラムと、Gaussian¹¹⁾ による量子計算を組み合わせ、モノマーのある物性 A と B についてパレート解の探索を行った (Fig. 5)。アルゴリズムとしては NSGA-II (Non-dominated Sorting Genetic Algorithms-II)¹¹⁾ を用いた。NSGA-II は高速非優越ソートにより、パレート解に近くかつ混雑距離の小さい (周りに似た物性の少ない) 個体を優先的に生き残らせることで多様性を維持するアルゴリズムである (Fig. 6)。50 世代の計算後、パレート前線が押し上げられ、多目的最適化が良好に進行した (Fig. 7)。

4. おわりに

このように、各種のシミュレーションとインフォマティクスを組み合わせることで、単独で困難であった多様な応用が可能となってきている (Fig. 8)。

本論で紹介したように、①機械学習を用いると高分子シミュレーションの計算資源的な制約を取り払うことができ、高分子シミュレーションをスクリーニングや最適化に用いることができるほか、系を全体的に俯瞰することで新たな知見が得られる場合もある。また、②自動設計手法 (遺伝的アルゴリズムやベイズ最適化)

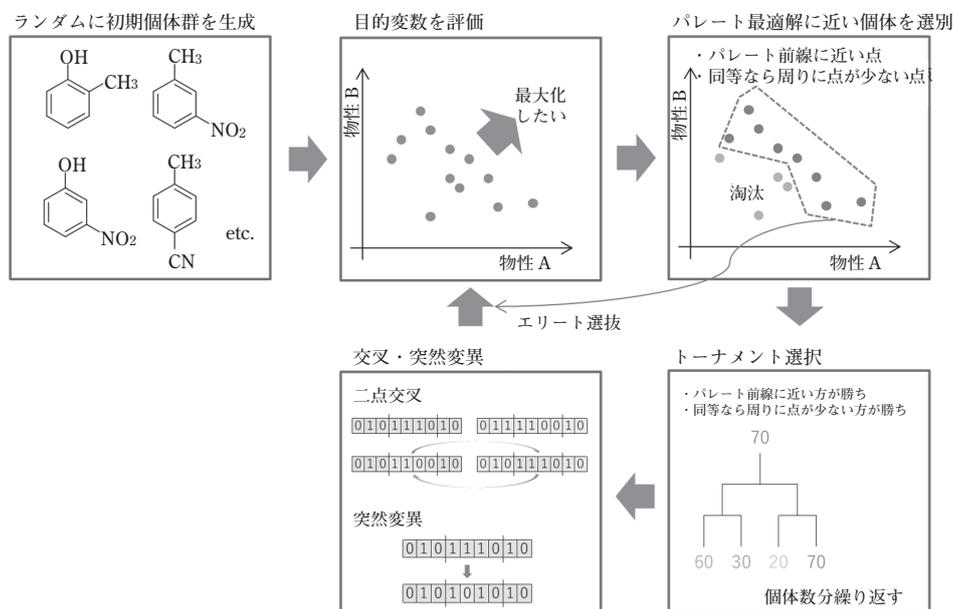


Fig. 6 NSGA-II の概略

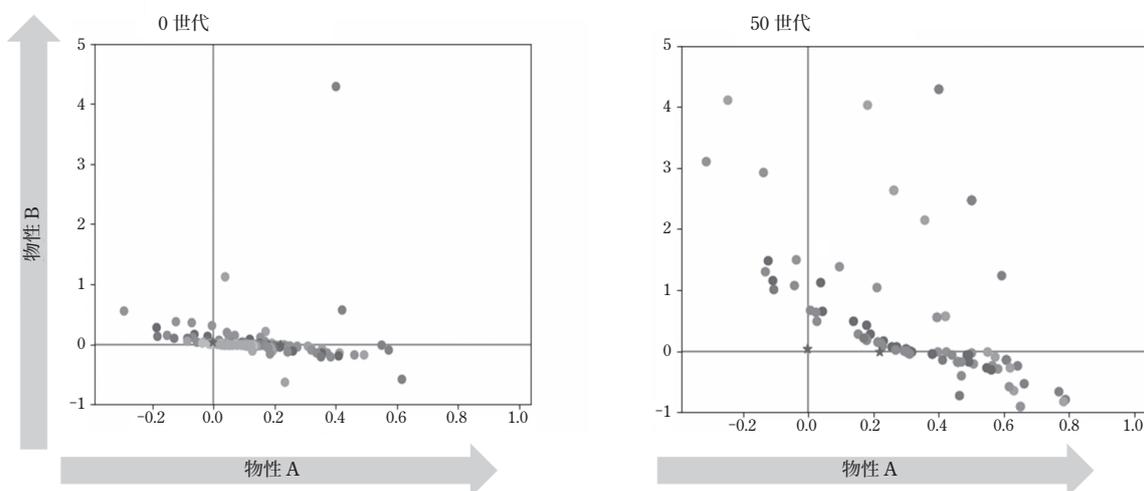


Fig. 7 NSGA-II による分子構造最適化前と後

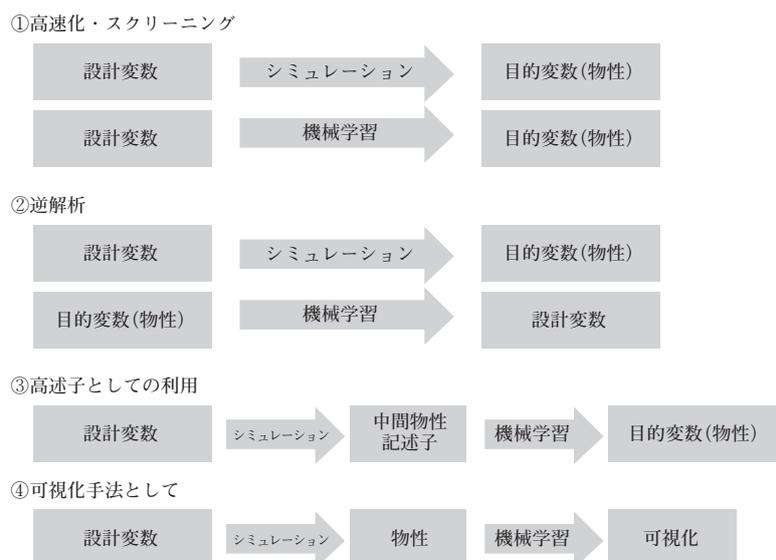


Fig. 8 高分子シミュレーションと機械学習のシナジー例

とシミュレーションを組み合わせることで、目的の物性を与える条件を効率的に探索して逆問題を解くことができる。

他の応用として、③シミュレーション結果を機械学習の説明変数(記述子)として用いて実測値の予測に用いる、あるいはそうして得たモデルを転移学習して他物性の予測に転用する、④シミュレーション結果が多次元であり可読性が乏しい場合に、クラスタリングや次元削減などを用いて整理・可視化・意味づけを行うことなども期待される。

ソフトマター開発においては説明変数の数が多くなりがちであるため、探索する空間に対して実験データは常に不足しがちであり、実験データのみからの汎用的なモデルの作成は困難である。今後はシミュレータ・実験データ・機械学習アルゴリズムをシームレスに統合した材料開発が必要とされると予想する。

5. 謝辞

本研究は(公社)新化学技術推進協会の高分子シミュレーション技術セミナーにおいて行われました。顧問の滝本 淳一先生(山形大院・有機)始めメンバーの皆様にご心より御礼申し上げます。また、PCN ソルバーをご提供いただきました増渕 雄一先生(名大院工)に厚く御礼申し上げます。

参考文献

- 1) OCTA 公式サイト (<http://octa.jp/jp/>)
- 2) T. Aoyagi et al., *J. Chem. Phys.*, **117** (17), 8153 - 8161 (2002)
- 3) R. D. Groot et al., *J. Chem. Phys.*, **107** (11), 4423 - 4435 (1997)
- 4) 坂下竜一、小峯拓也、東ソー研究・技術報告、**62**、43-49 (2018)
- 5) A. J. Smola et al., *Stat. and Compt.*, **14**, 199 - 222 (2004)
- 6) 増渕雄一、日本ゴム協会誌、**82** (11)、459-463 (2009)
- 7) K. Takeda et al., *Adv. Model. And Simul. in Eng. Sci.*, 2-11 (2015)
- 8) 坂下竜一、東ソー研究・技術報告、**63**、31-36 (2019)
- 9) NAPLES 公式サイト (<http://masubuchi.jp/NAPLESweb/index.html>)
- 10) RDKit 公式サイト (<http://www.rdkit.org>)
- 11) Gaussian 16, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams - Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A.

Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

- 12) Deb, K. et al., *IEEE Trans. on Evolutionary Computation*, **6** (2), 182-197 (2002)

