

# プレスリリース



ISM2022-08  
2022年11月9日



## 全原子古典分子動力学法による高分子物性計算を全自動化する ソフトウェア RadonPy をリリース

**RadonPy** ～高分子材料物性大地図の作成に向けた第一歩～

### 研究成果のポイント

- 全原子古典分子動力学法による高分子物性計算の全自動化に成功した初のオープンソースソフトウェア [RadonPy](#) を公開
- スーパーコンピュータ「富岳」等の計算資源を活用し、10万種類以上の分子骨格を包含する高分子物性データベースを開発するプロジェクトを始動
- RadonPy およびデータベースの共同開発を推進する産学協働コンソーシアムを設立

### 概要

材料データとデータ科学・計算科学を融合した材料研究の新しい形態“マテリアルズインフォマティクス” (Materials Informatics: 以下、MI という。) がいま大きな注目を集めています。MI では、機械学習を適用して、データのパターンを読み解き、広大なデザインスペースから革新的な特性を有する新材料とその作製方法を予測します。ここ数年間で MI の技術は材料研究の様々な領域に広まり、実際に多くの新材料が発見されてきました。しかしながら、高分子材料<sup>(1)</sup>の MI の普及は、他の材料系に比べると大幅に遅れています。MI における最も重要な学術資源はいうまでもなくデータです。しかしながら、高分子物性のデータベースに関しては、国立研究開発法人物質・材料研究機構 (以下、NIMS という。) が開発している [PoLyInfo](#)<sup>(2)</sup> (論文等の実験値を網羅的に収集したデータベース) を除き、データ駆動型研究に資する体系的なオープンデータを創出しようという動きは極めて低調です。このままでは高分子材料の MI に携わる多くの研究者にとって、大学の研究室や一企業で生産可能なデータが標準になってしまう可能性があります。

このような現状を打破するために、統計数理研究所 (以下、統数研という。) の林慶浩助教と吉田亮教授は、東京工業大学の森川淳子教授と東京大学の塩見淳一郎教授、NIMS の PoLyInfo 開発チームの協力を得て、全原子古典分子動力学法<sup>(3)</sup> (以下、MD 計算あるいは MD シミュレーションという。) による高分子物性計算を全自動化するソフトウェア RadonPy を開発しました。高分子物性の MD 計算は計算条件の設定により大きくゆらぎ、計算量も膨大であるため、自動計算による大規模データベースの構築は技術的に困難であると考えられてきました。RadonPy は、高分子材料の繰り返し単位 (モノマー) の化学構造と重合度<sup>(4)</sup>、温度等の計算条件を入力とし、アモルファスポリマー<sup>(5)</sup>や高分子溶液等の系に対し、熱物性、機械特性、光学特性を含む様々な物性を自動計算します。今回リリースしたバージョンは、15 種類の物性の自動計

# プレスリリース



算機能を実装しています。RadonPy には、多数の煩雑な手続きからなる MD 計算の全工程を自動化し、パイプラインを簡便に構築するための諸機能が実装されています。RadonPy の特色の一つは、様々な骨格の高分子材料に適用可能な検証済みパラメータセットと計算条件（プリセット）を標準搭載していることです。プリセットの設定では、NIMS の協力を得て PoLyInfo の実験データをベンチマークとして使用しました。また、MD 計算の物性値と複雑な現実系（実験値）の間のバイアスやばらつきを補正するために、転移学習<sup>(6)</sup>という機械学習の解析技術を用いました。

同グループは、RadonPy を用いて 10 万種類以上の分子骨格を包含する高分子物性オープンデータベースを創出するプロジェクトを推進しています。現在、統数研に加えて 3 大学と 19 企業がプロジェクトに参画し、多くの研究者が産学の垣根を越えて RadonPy とデータベースの共同開発に取り組んでいます。この取り組みは文部科学省「富岳」成果創出加速プログラムの支援を受け、参画者は、スーパーコンピュータ「富岳」<sup>(7)</sup>の計算資源を最大限に活用して日々膨大なデータを生産・蓄積しています。プロジェクトの目標は、高分子材料物性の大地図を作成することです。データを生産していく過程で、複数物性の同時分布や物性間のトレードオフが生み出すパレートフロンティア<sup>(8)</sup>、さらにフロンティアを形成する新材料が明らかになっていきます。材料合成等の膨大なコストを伴う実験的アプローチのみでは、このような網羅的な観測は実現不可能です。世界最大級の二つの高分子物性データベース、PoLyInfo と RadonPy データベースを統合的に活用することで、不確かで複雑な現実系と不完全な計算モデルの壁を乗り越えられるかもしれません。

本研究成果は 2022 年 11 月 8 日 10 時（日本時間 8 日 19 時）に [npj Computational Materials](#) 誌（Nature Publishing Group）にて発表されました。

## 研究の背景

データ駆動型研究における最も重要な学術資源はいうまでもなくデータです。しかしながら、材料研究において利用できるデータ資源は、量的にも質的にも非常に乏しいという難しさがあります。特に、高分子材料のデータ資源の乏しさは際立っています。PoLyInfo を除く既存の高分子物性データベースは、いずれもデータ量が非常に少なく、データを自動抽出・成形・加工するためのツールが整備されていません（[図 1](#)）。また、試料の作製条件や高次構造等の情報も十分には記録されていません。データ駆動型高分子材料研究に資する体系的かつ包括的なオープンデータの創出が喫緊の課題となっています。

オープンデータの整備に遅れが生じている原因として、次の三点が考えられます。

- 材料合成、試料作製、物性測定、シミュレーション（計算機実験）のコストが高い。
- 研究対象の設計変数（材料種、プロセス条件等）が多様であるがゆえ、研究分野の共通基盤となるようなオープンデータを共創しようという機運が起きにくい。
- 競合相手に対する情報秘匿の意識が高く、データを公開するインセンティブが研究者に働きにくい。

このような背景からコミュニティが協調してオープンデータを創出していこうという動きが

# プレスリリース



低調になっています。これらの要因の多くは文化的な問題であるため、短中期的には、大学の研究室や一企業で生産可能なレベルのデータが、MI の標準的な解析対象になっていくことが予想されます。

そこで MI では、データの不足を補うためにシミュレーションから得られるデータを統合的に解析することが重要になってきます。現在、世界各国にて様々な材料を対象に大規模な計算物性データベースが開発されています。特に無機化合物や低分子化合物については、数万から数百万種類以上の物質を包含する第一原理計算データベース ([Materials Project](#)、[AFLOW](#)、[OQMD](#)、[QM9](#) 等) が開発されたことで、MI の技術的進歩と実践展開が飛躍的に進みました。一方、高分子材料については、物性計算の自動化の技術的な難しさと分子シミュレーションの膨大な計算コストが壁となり、計算物性データベースの開発はほとんど進んでいません。

## 研究内容

RadonPy は、全原子 MD シミュレーションによる高分子物性計算を全自動化する世界初のオープンソースソフトウェアです。RadonPy は、高分子の繰り返し単位とその化学構造、重合度、温度等の計算条件を入力とし、分子モデリング、電荷計算、平衡・非平衡 MD 計算、平衡化完了の自動判定、収束に失敗した場合の再スタートのスケジューリング、ポストプロセス段階での物性計算等、MD シミュレーションの全工程を完全に自動実行します ([図 2](#))。ファーストリリースで公開されたバージョンは、熱物性、機械特性、光学特性を含む 15 種類の物性を自動計算するアルゴリズムを実装しています。計算可能な系は、線状高分子<sup>(9)</sup> (ホモポリマー<sup>(10)</sup>、コポリマー<sup>(11)</sup>) のアモルファス状態<sup>(5)</sup>や溶液中の高分子等を含みます。

今回の研究は、大規模データベース構築に向けたパイロットスタディとして実施されました。本研究では、1,000 種類以上のアモルファスポリマーを対象に 15 種類の物性を計算し、PoLyInfo の実験値を用いた検証実験が実施されました。重合度等の計算条件や分子骨格の種類が計算値に与える影響や、物性毎の予測精度や性能限界が徹底的に調べられました。このような規模で高分子物性 MD 計算の性能が系統的に検証された事例はありませんでした。

さらに、転移学習という機械学習の解析技術が、計算物性の系統バイアスやばらつきを補正する上で強力な性能を発揮することが明らかになりました ([図 3](#))。様々な高分子材料に対して普遍的に適用できる計算条件は存在しません。したがって、全自動計算で大量生産されるデータには、必ずバイアスやばらつきが生じます。また、実験条件や試料に関する非観測因子や計測系の特性により、実験値にもバイアスとばらつきが生じます。このような複雑な現実系と不完全な計算モデルの間のギャップを機械学習で埋めることができます。

さらに今回の研究により、高分子材料の複数物性の同時分布やパレートフロンティアの位置、そこに存在する高分子材料の構造的特徴が明らかになってきました ([図 4](#))。特に本研究は高分子材料の熱伝導率に着目しました。近年のモバイル機器の小型化・高性能化に伴う発熱量の増加により、モバイル機器の絶縁樹脂、成形樹脂、接着剤、コーティング剤等に应用できる材料へのニーズが高まり、成形性に優れた高分子材料の高熱伝導化の研究に注目が集まっています。今回の計算では、熱伝導率が  $0.4 \text{ W}/(\text{m} \cdot \text{K})$  を超える 8 種類のアモルファスポリマーが同定されまし

# プレスリリース



た (図5)。通常のアモルファスポリマーで到達できる熱伝導率は高々0.2~0.3 W/(m・K) 程度といわれていますが、同定された材料はこの水準を大きく上回ります。また、RadonPy に実装されている熱伝導成分分割の解析手法を適用し、アモルファスポリマーの高熱伝導化のメカニズムと設計指針を明らかにしました。その結果、水素結合ユニットを高密度に含む高分子では、分子間の水素結合や双極子相互作用を介して熱伝導性が向上すること、あるいは高い硬直性と線形性を持つ高分子では共有結合を介して熱伝導率が向上することが明らかになりました。

## 今後の展望

現在、同グループは RadonPy を用いて 10 万種類以上の分子骨格を包含する世界最大級の高分子物性データベースを創出するプロジェクトを推進しています。2022 年 10 月には、RadonPy とデータベースの共同開発を目的とする産学協働コンソーシアムを正式に発足しました（非公式の活動は 2021 年 4 月に始動）。コンソーシアムには現在、統数研以外に 3 大学と企業 19 社が参画しており、約 90 名の研究者が産学の垣根を越えて RadonPy とデータベースの共同開発を推進しています。この取り組みは文部科学省「富岳」成果創出加速プログラムの支援を受け、産学の多くの研究者が「富岳」の計算資源を最大限に活用し、日々膨大なデータを生産・蓄積しています。プロジェクトが掲げるスローガンは、高分子材料物性大地図の作成です。データを生産・蓄積する過程で、複数物性の同時分布とパレートフロンティア、さらにフロンティアを形成する新材料を明らかにしていきます。特に、生分解性プラスチックや高熱伝導性高分子材料に関する体系的なデータを創出し、脱炭素・循環型社会やサーマルマネジメントに資する新材料の創製を目指します。また、産学の組織の垣根を超えたデータの共創というモデルケースを社会に発信していくこともコンソーシアムの活動目的の一つです。

本事業がマイルストーンに到達した暁には、PoLyInfo と RadonPy データベースという実験・計算の世界最大級の高分子物性データベースが利用可能になります。全ての計算モデルは不完全です。計算資源の制約により、実材料開発における混練や加熱、フィルター導入等のプロセス情報を MD 計算に取り組むことは難しく、必ず計算結果には現実系との乖離が生じます。同様に、実験データも不完全であり不確かです。実験と計算という車の両輪は二つが揃ってこそ意味があり、データ科学が両輪を回すエンジンの役割を担います。

## 掲載論文

論文題目: RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics

著者: Yoshihiro Hayashi<sup>1</sup>, Junichiro Shiomi<sup>2</sup>, Junko Morikawa<sup>3</sup>, Ryo Yoshida<sup>1,4,5</sup>

雑誌: [npj Computational Materials](#)

DOI: 10.1038/s41524-022-00906-4

掲載日時: 2022 年 11 月 8 日 19 時 (英国時間 10 時)

# プレスリリース



- 2 Department of Mechanical Engineering, The University of Tokyo
- 3 Department of Materials Science and Engineering, School of Materials and Chemical Technology, Tokyo Institute of Technology
- 4 Graduate University for Advanced Studies, Department of Statistical Science
- 5 Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS)

## RadonPy 公開ウェブサイト

<https://github.com/RadonPy/RadonPy>

データベース	概要
PoLyInfo (polymer.nims.go.jp)	学術文献から抽出したデータをまとめた世界最大の高分子物性データベース。約18,000種類のモノマーから重合されたポリマー群の約100物性のデータを収録。
Polymer Genome – Khazana (khazana.gatech.edu)	24の出版物から抽出した実験データと第一原理計算で算出した物性値を提供しているプラットフォーム。約1,400種類のポリマー・有機材料と約2,600種類の無機材料の特性データを収録。
Polymer Property Predictor and Database (pppdb.uchicago.edu)	文献から抽出した約260件のFlory-Huggins $\chi$ パラメータと約210件のガラス転移温度のデータを含む。
NanoMine (materialsmine.org)	ポリマーコンポジットの微細組織構造の組成、プロセス、電子顕微鏡データ、物性を含む。
CROW (polymerdatabase.com)	ポリマーの熱物性データを含むレポジトリ。文献から抽出した実験データや定量的構造活性相関解析から算出した計算物性のデータを含む。
Polymers: A Property Database (poly.chemnetbase.com)	Wiley社の書籍“Polymers: A Property Database”の付録として提供されている高分子物性データ

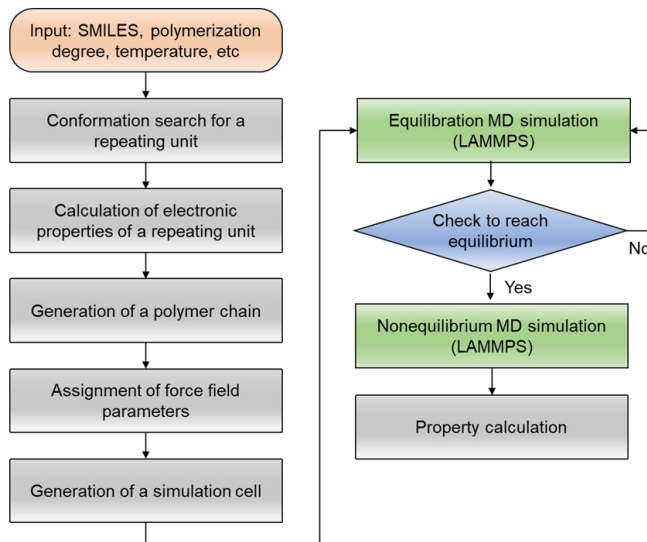
図1: 代表的な高分子物性オープンデータの一覧。データ科学の他の応用分野に比べるとデータ量は非常に少なく、データを自動抽出するためのツールも整備されていないケースがほとんどである。



# プレスリリース



Workflow of automated all-atom molecular dynamics simulations for polymeric materials



15 properties implemented (2022/10/01)

- ◆ Thermal conductivity
- ◆ Thermal diffusivity
- ◆ Density
- ◆ Radius of gyration
- ◆ Specific heat capacity  $C_p$
- ◆ Specific heat capacity  $C_v$
- ◆ Compressibility (isothermal)
- ◆ Isentropic compressibility
- ◆ Bulk modulus (isothermal)
- ◆ Isentropic bulk modulus
- ◆ Self-diffusion coefficient
- ◆ Thermal expansion coefficient
- ◆ Linear expansion coefficient
- ◆ Dielectric constant (static)
- ◆ Refractive index

Polymer systems

- ◆ Amorphous
- ◆ Crystalline states
- ◆ Oriented structures

図 2：高分子物性自動計算ソフトウェア RadonPy の概要。高分子の繰り返し単位とその化学構造、重合度、温度等の計算条件を入力とし、分子モデリング、電荷計算、平衡・非平衡 MD 計算、平衡化完了の自動判定、収束に失敗した場合の再スタートのスケジューリング、ポストプロセス段階での物性計算等、MD シミュレーションの全工程を完全に自動実行する。ファーストリリースで公開されたバージョンは、熱物性や光学特性を含む 15 種類の物性を自動計算できる。

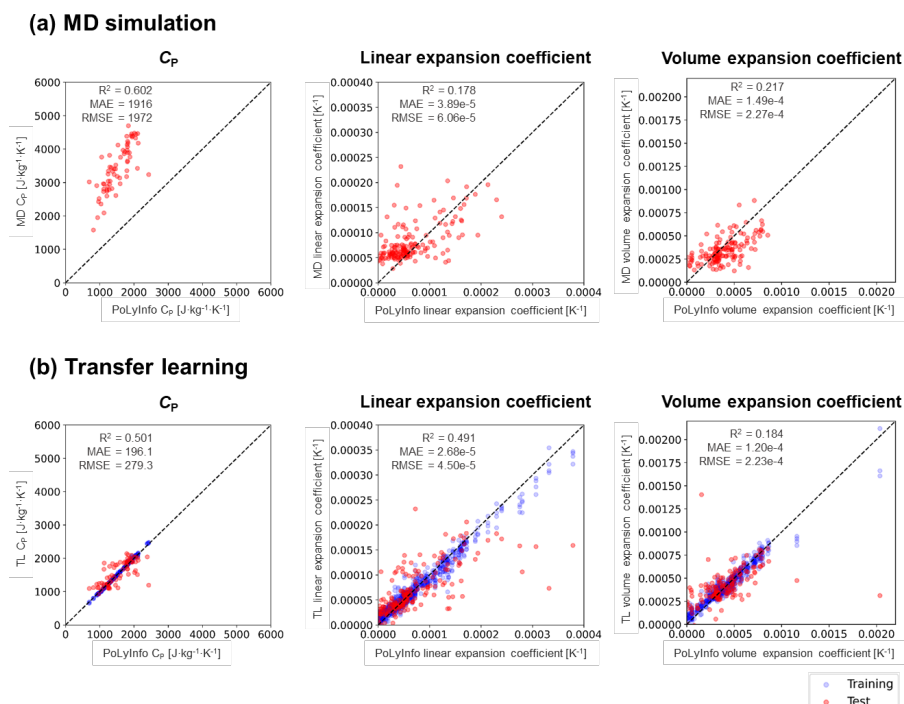


図 3：転移学習による MD 計算物性（定圧比熱容量 ( $C_p$ )、線膨張係数、体積膨張率）のキャリ

# プレスリリース



ブレーション。MD 計算値と実験値の間に存在するバイアスとばらつき（上）が転移学習を用いたキャリブレーションによって大幅に改善された（下）。これは実験と計算両方のバイアスとばらつきを転移学習は予測可能であることを意味する。

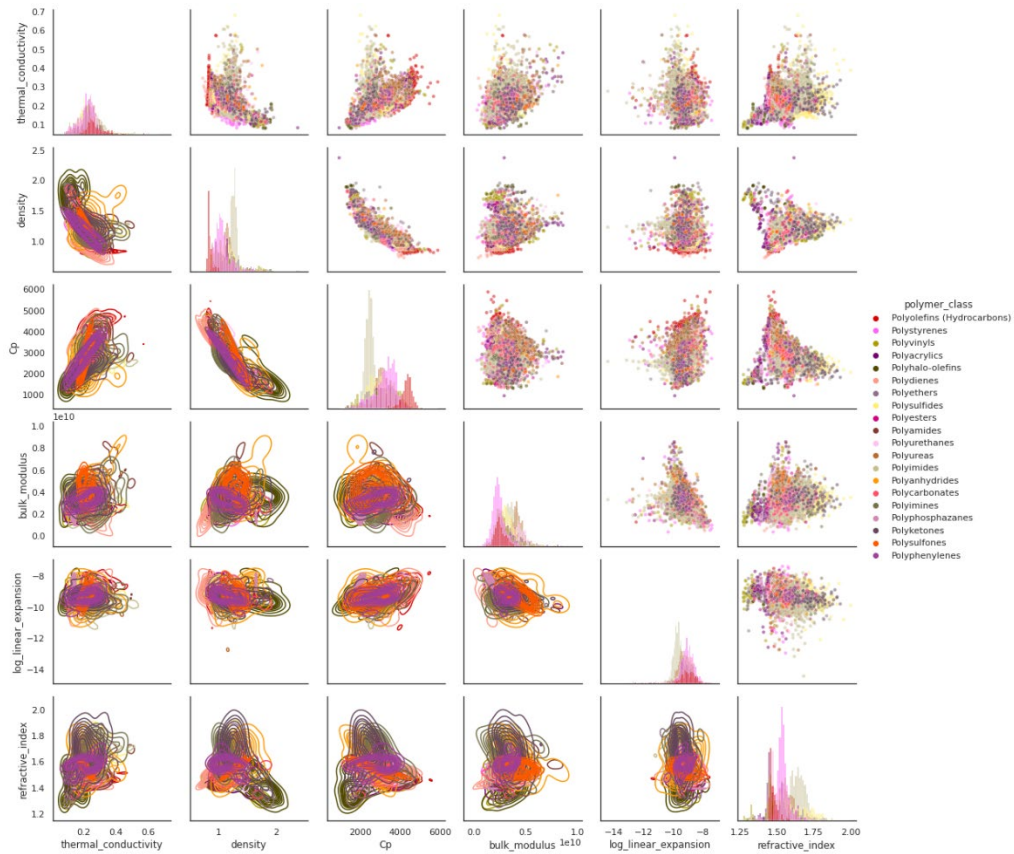


図 4: RadonPy によるハイスループット自動計算で明らかとなった高分子材料の複数物性（熱伝導率、密度、定圧比熱容量 ( $C_p$ )、体積膨張率、線膨張係数、屈折率) の同時分布とパレートフロンティア

# プレスリリース

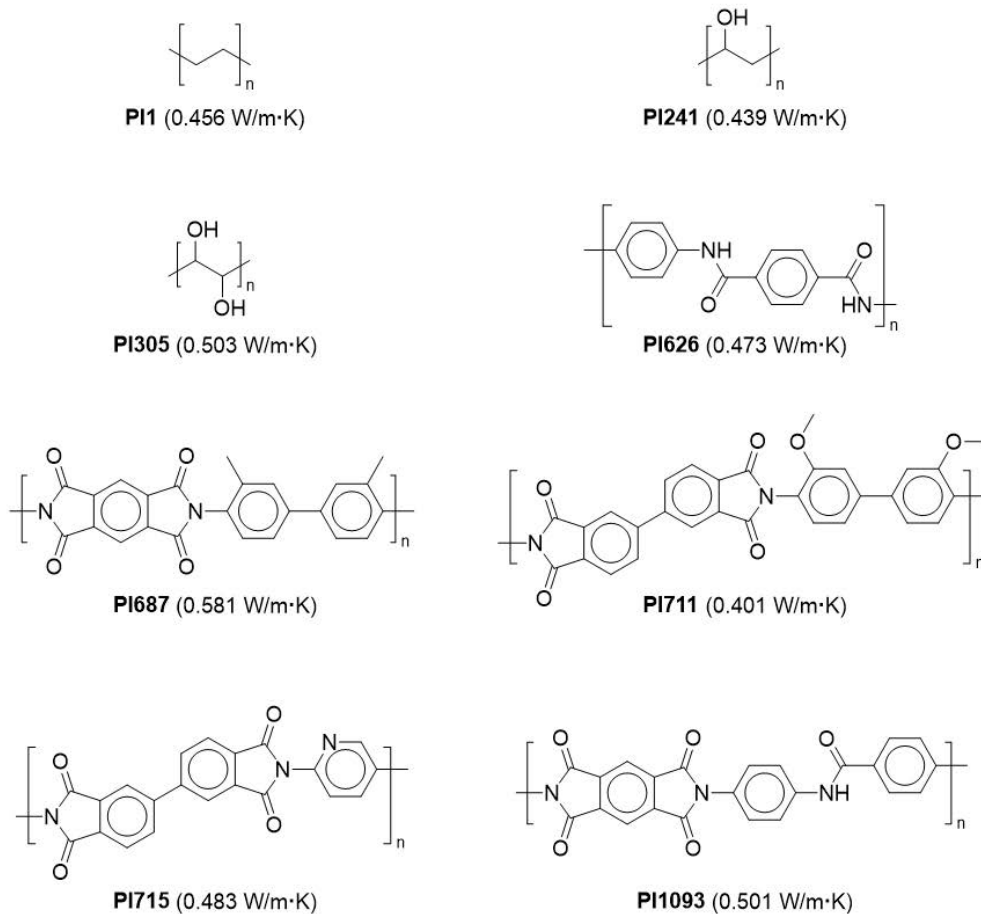


図5：アモルファス状態で0.4 W/(m・K)を超える熱伝導率を示した8種類の高分子の化学構造

## 謝辞

本研究は、文部科学省「富岳」成果創出加速プログラム「データ駆動型高分子材料研究を変革するデータ基盤創出」(JPMXP1020210314) および JST-CREST 研究課題「高分子の熱物性マテリアルズインフォマティクス」(JPMJCR19I3) の支援を受けて実施されました。また、本研究の一部は、スーパーコンピュータ「富岳」の計算資源（課題番号：hp210264, hp210213）と自然科学研究機構計算科学研究センターの計算資源（課題番号：21-IMS-C126, 22-IMS-C125）の提供を受けて実施されました。本研究を実施するにあたり、高分子物性データベース PoLyInfo を提供していただいた国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門 材料データプラットフォームセンターの桑島功氏（主幹エンジニア）ならびに石井真史氏（グループリーダー）に感謝の意を表します。

## 用語解説

(1) 高分子材料：通常は分子量が 10,000 以上の大きな分子のことを高分子あるいはポリマー分子という。高分子は、1 種または数種の構成要素（繰り返し単位）が数多く繰り返し連結した構造を持つ。高分子から構成される材料のことを高分子材料あるいはポリマーという。



# プレスリリース



大学共同利用機関法人 情報・システム研究機構  
統計数理研究所



東京工業大学  
Tokyo Institute of Technology



東京大学  
THE UNIVERSITY OF TOKYO



(2) PoLyInfo：国立研究開発法人 物質・材料研究機構が保有する高分子物性の世界最大級のデータベース。学術文献から収集した約 100 種類の物性（熱物性、電気的特性、力学的特性等）、化学構造、測定条件、重合方法等を収録している。

(3) 全原子古典分子動力学法：原子間に働く相互作用ポテンシャルのもとで、ニュートンの運動方程式（古典力学）を解き、計算された原子の動的な振る舞いから物性や構造を求める計算機実験。

(4) 重合度：高分子を構成するモノマー単位（繰り返し単位）の数。

(5) アモルファスポリマー：原子が不規則に配列し、結晶構造を持たない物質の状態をアモルファス状態といい、この状態のポリマーをアモルファスポリマーという。

(6) 転移学習：機械学習で予測モデルを構築したいが、データ量が少なくてフルスクラッチでの学習は難しい場合、関連する他のドメインからデータや事前に学習されたモデルを利用することで、データ量の不足を補うための統計的機械学習の方法論の総称を転移学習という。ここでは、MD シミュレーションの物性値から実験値への転移学習を実施した。

(7) スーパーコンピュータ「富岳」：スーパーコンピュータ「京」の後継機として理化学研究所に設置された計算機。令和 2 年 6 月から令和 3 年 11 月にかけてスパコンランキング 4 部門で 1 位を 4 期連続で獲得する等、世界トップレベルの性能を持つ。令和 3 年 3 月 9 日に本格運用開始。

(8) パレートフロンティア：複数物性の多目的最適化において、通常は一つの物性を最適水準に近づけると他の物性の性能が低下するというトレードオフの関係が存在する。現在得られているある解（例えば、材料）が、これ以上は複数の物性を同時に改善することが不可能な場合、これをパレート最適解という。パレート最適解の集合のことをパレートフロンティアという。

(9) 線状高分子：線状にモノマーが繰り返す高分子。

(10) ホモポリマー：1 種類のモノマーから構成される高分子。

(11) コポリマー：複数種類のモノマーから構成される高分子。

## 本件に関するお問い合わせ先

### 【研究内容について】

大学共同利用機関法人 情報・システム研究機構 統計数理研究所 ものづくりデータ科学研究センター 教授（センター長）

吉田 亮（よしだ りょう）

TEL: 050-5533-8534

E-mail: [yoshidar@ism.ac.jp](mailto:yoshidar@ism.ac.jp)

# プレスリリース



## 【報道・広報について】

大学共同利用機関法人 情報・システム研究機構 統計数理研究所 運営企画本部企画室 URA ステーション

〒190-8562 東京都立川市緑町 10-3

TEL: 050-5533-8580 FAX: 041-526-4348

E-mail: [ask-ura@ism.ac.jp](mailto:ask-ura@ism.ac.jp)