

分子物性推算データベース: CPDB - 分子設計支援システム: MolWorksの開発 -

田島 澄恵^a, 八木 徹^a, 福田 朋子^b, 長嶋 雲兵^{b*}

^a 株式会社ベストシステムズ, 〒 305-8568 つくば市梅園 1-1-1

^b 産業技術総合研究所グリッド研究センター, 〒 305-8568 つくば市梅園 1-1-1

*e-mail: u.nagashima@aist.go.jp

(Received: March 29, 2005; Accepted for publication: November 10, 2005; Published on Web: February 27, 2006)

原子団寄与法による分子の物性推算データを集約した分子物性推算データベース:CPDB の開発を開始した。産業技術総合研究所 RIODB(<http://www.aist.go.jp/RIODB/cpdb/>) で一部の公開を開始したので概要を報告する。

CPDB は、原子団寄与法の一つである Joback 法を用いて、計算機を用いて機械的に生成された 12 種類の物性データをまとめたものであり、現在 258,907 分子のデータが公開されている。検索方法は、推算物性の範囲を指定し、それにより候補分子を検索するというものであり、候補物質を入力して物性値を得ることばかりでなく、逆に物性値からの物質候補のスクリーニングを可能とするデータベース構築の試みである。

キーワード: Property Estimation, Joback Method, Molecular Screening, Property Data Base

1 はじめに

新規材料開発を行なう際、合成段階物質や目的物質などの構造や物性情報が必要となる。そのため、計算機化学・化学工学計算・物性推算・データ解析・データベースシステムなどが利用されている。これらを独立的に利用するだけでは不十分であり、統合的に利用できることが望ましい。しかし、そのようなシステムは皆無である。そこで、我々は、統合的な分子設計支援システム: MolWorks[1] の開発を行なっている。

MolWorks は、分子軌道計算、物性推算、化学工学計算などに対応している。物性推算は、Joback 法 [2] をベースとした原子団寄与法を用いている。原子団寄与法は分子の化学構造から物性を推算する手法である。予測対象となる物性をいくつかの因子の関係式として表し、それぞれの因子を分子構造の中に含まれる原子団からの寄与 (原子団パラメーター) の和として計算する。関係式や原子団パラメーターは実測データの

解析により求められることから、化学構造と物性値を関連づけたデータベースが効果的に利用できる手法である。

MolWorks では分子を原子団に自動分割し、沸点・融点・臨界定数など 13 種類の推算を行なう。例として手元に沸点の実測データのある 1727 分子の Joback 法による推算結果と実測との相関を Figure 1 に示した。Joback 推算法は特に低温部・高温部での誤差が大きく、200-600 °K の限られた範囲で実測との相関がよいことが分かるが、実測のない任意の分子の物性が非常に簡便にえられるという意義がある。

MolWorks はフリーで公開されているが、そのプログラムをダウンロードしインストールすることなく、推算された物性を参照したいというユーザからの要望をいくつかいただいた。そこで我々は、MolWorks に実装された原子団寄与法による分子の物性推算プログラムによって生成されたデータを集約した分子物性推算データベース:CPDB の開発を開始した。産業技術

総合研究所 RIODB[3] で CPDB の一部の公開を開始したので、その概要を報告する。CPDB の URL は、

<http://www.aist.go.jp/RIODB/cpdb/>である

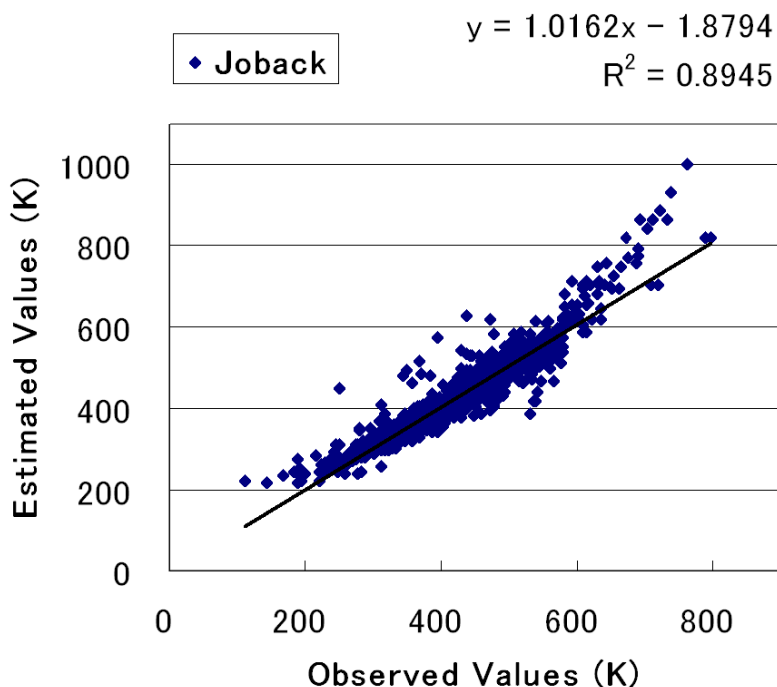


Figure 1. Correlation between Estimated and Observed Boiling Point Values

Table 1. Parameters included in a record of CCDB

- | | |
|-----|---|
| 1) | Molecular Structure |
| 2) | Molecular Weight |
| 3) | Joback Method Boiling Point (K) / (): 沸点 |
| 4) | Joback Method Critical Temperature T_c (K) / (): 臨界温度 |
| 5) | Joback Method Critical Pressure P_c (bar) / (atm) : 臨界圧力 |
| 6) | Joback Method Critical Volume V_c (cm ³ /mol) : 臨界体積 |
| 7) | Edmister Method omega(From Joback Parameter Acentric Factor) : 偏心係数 |
| 8) | Lee-Kesler Method omega(From Joback Parameter Acentric Factor) : 偏心係数 |
| 9) | Critical Compressibility Factor Z_c : 臨界圧縮係数 |
| 10) | Yen-Woods Method Density (g/cm ³) : 密度 |
| 11) | Reidel Method Vapor Pressure (mmHg) : 蒸気圧 |
| 12) | Vetre Method Heat of Vaporization (at boiling point) (cal/mol) : 蒸発熱 |
| 13) | L-J Collision Diameter for Chemkin* (Å) : Chemkin*入力用衝突断面積 |
| 14) | L-J Well Depth for Chemkin* (J/mol) : Chemkin*入力用井戸型ポテンシャル |

*<http://www.cdaj.co.jp/product/chemkin/chemkin.html>

2 検索

現在 RIODB で公開されている CPDB に含まれるレコード数は 258,907 である。レコードは分子構造に加え Table 1 に示す 13 の項目から構成される。3)-6) は、それぞれ Joback 法により得られた沸点、臨界温度、臨界圧力、臨界体積である。7) は Edmister 法 [4] によって得られた偏心係数である。計算には Joback 法のパラメータがもちいらている。Edmister 法は、一番精度の高い方法であると知られているが、小さな分子の精度は良くない。8) は Lee-Kesler 法 [5] の偏心係数である。9) は、臨界圧縮係数 Z_c 、10) は Yen-Woods 法から得られる密度である。11) は Reidel 法 [6] によって得られる蒸気圧、12) は、Vetre 法による沸点における蒸発熱である。最後の 13) 衝突断面積および 14) 井戸型ポテンシャルの 2 つは、素反応ベースの化学反応解

析ソフトウェアである Chemkin[7] 用のパラメータである。

Figure 2 に検索画面を示した。

検索画面では、物性値を範囲で指定することが可能である。ここで、必要な物性量を指定して画面下方の Submit ボタンを押して検索を実行すると検索結果一覧 (Figure 3) が示される。検索結果一覧は新しい画面が表示される。またこの検索画面では検索結果の表示件数を選択できる。

Figure 3 に分子量を 100-110 と指定して検索を実行した結果一覧画面の一部を示した。検索結果は 214 件であり、そのうち番号が若い順からの最初の 20 件が示されている。ボタンをクリックすれば次のページに移動可能であり、ページを指定すれば直接移動が可能である。

Search

| | |
|--|--|
| Molecular Structure | <input type="text"/> |
| Molecular Weight | <input type="text"/> to <input type="text"/> |
| Boiling Point(K) | <input type="text"/> to <input type="text"/> |
| Critical Temperature(K) | <input type="text"/> to <input type="text"/> |
| Critical Pressure(bar) | <input type="text"/> to <input type="text"/> |
| Critical Volume(cm^3) | <input type="text"/> to <input type="text"/> |
| Edmister Method omega | <input type="text"/> to <input type="text"/> |
| Lee-Kesler Method omega | <input type="text"/> to <input type="text"/> |
| Critical Constant | <input type="text"/> to <input type="text"/> |
| Yen-Woods Method Density (g/cm^3) | <input type="text"/> to <input type="text"/> |
| Reidel Method Vapor P (mmHg) | <input type="text"/> to <input type="text"/> |
| Vetre Method Heat of Vaporization (at BP) (cal/mol) | <input type="text"/> to <input type="text"/> |
| L-J Collision Diameter for Chemkin (\AA) | <input type="text"/> to <input type="text"/> |
| L-J Well Depth for Chemkin (J/mol) | <input type="text"/> to <input type="text"/> |
| Display <input type="text" value="20"/> results per page | |

Figure 2. Retrieval Window of CPDB on RIO-DB at AIST. (<http://www.aist.go.jp/RIODB/cpdb/>)

Search results

Number of results : 214
1 to 20 of results is displayed.

Go to Page

| Serial No | Molecular Structure | Molecular Weight | Boiling Point(K) | Critical Temperature (K) | Critical Pressure (bar) | Critical Volume (cm ³) | Edmister Method omega | Lee-Kesler Method omega | Critical Constant |
|-----------|---------------------|------------------|------------------|--------------------------|-------------------------|------------------------------------|-----------------------|-------------------------|-------------------|
| 4 | CF3.Cl: | 1.04457E+02 | 2.5429E+02 | 4.0352E+02 | 4.647E+01 | 1.835E+02 | 2.13330774E-01 | 2.1484799E-01 | 2.54144836E-01 |
| 23 | F.CCl2: | 1.0292E+02 | 2.9597E+02 | 4.7095E+02 | 4.925E+01 | 2.015E+02 | 2.22645126E-01 | 2.23002762E-01 | 2.53417788E-01 |
| 35 | CHF2.CHF2: | 1.02029E+02 | 2.4136E+02 | 3.6988E+02 | 4.036E+01 | 2.075E+02 | 2.87993186E-01 | 2.92030535E-01 | 2.72344008E-01 |
| 40 | CHF2.C(CH3)3: | 1.08126E+02 | 3.0867E+02 | 4.6707E+02 | 3.225E+01 | 3.345E+02 | 2.54995443E-01 | 2.63161403E-01 | 2.77754546E-01 |
| 54 | CH(CH3)2.C(CH3)3: | 1.00196E+02 | 3.5589E+02 | 5.3472E+02 | 2.881E+01 | 4.105E+02 | 2.39986566E-01 | 2.50092477E-01 | 2.66029272E-01 |
| 76 | CHCl.F.Cl: | 1.0292E+02 | 2.9597E+02 | 4.7095E+02 | 4.925E+01 | 2.015E+02 | 2.22645126E-01 | 2.23002762E-01 | 2.53417788E-01 |
| 103 | CHCl.CH3.CH(CH3)2: | 1.06589E+02 | 3.5035E+02 | 5.3142E+02 | 3.384E+01 | 3.525E+02 | 2.63556187E-01 | 2.70890198E-01 | 2.69986157E-01 |

Figure 3. A part of the Result Window for a query of Molecular Weight 100-110

3 CPDB の利用例

– Molecular Mining ツールとして –

CPDB の利用例として CPDB を用いた Molecular Mining の例を紹介する。我々は、物性推算式の応用として、Molecular Mining を提案している。通常の分子軌道計算や物性推算は、物質の構造を入力情報とし、物性値を出力情報とするものである。一方、Molecular Mining は、物性値を入力情報とし、物質を出力する。

CPDB を用いた Molecular Mining の一例として、パラキシレンの代替物質を探索してみた。パラキシレンの物性値は、沸点:411.4 [K]、臨界温度:617 [K]、臨界圧力:35 [bar]、臨界体積:378 [cm³/mol] である。以上の条件のもと、Molecular Mining を行なった。CPDB に含まれる探索対象物質 258,907 分子からパラキシレンの物性値に対応した分子を抽出すると、7 分子にまで絞り込むことができた。

CPDB の探査時間は、カリフォルニアの San Diego から通常のインターネットを経由して約 10 分であった。この時間はネットワークの混雑具合により大きく変動するため探査時間の一つの目安を与えるにすぎないが、作業量を見積もるための指標となる。

Molecular Mining をプログラムを用いておこなうとする。分子の生成を Java を用いた独自プログラムで行い、そして物性推算には MolWorks を用いる事とする。これを OS : Linux kernel-2.4.7、CPU : Pentium 1GHz、Memory : PC133 512MB の計算機上で行くと、258,907 分子の生成に約 4 分、258,907 分子の物性推算に約 20 分、合計約 25 分必要であった。

今後急速な計算機性能の改善が予測され、計算時間の大幅な短縮が実現されるとはいえ、CPDB 探査の有効性はなお大きい。

このように、Molecular Mining を利用することによって、短時間で大量の候補物質から少量の有力候補物質に絞り込むことができる。

本データベースがそういった研究に広く役立つよう、今後とも開発を継続していく。

本研究の一部は、産業技術総合研究所先端情報計算センター研究情報公開データベース RIODB 開発プロジェクトの支援を受けている。

参考文献

- [1] <http://www.molworks.com>
- [2] Joback, K. G., Reid, R. C., *Chem. Eng. Comm.*, **57**, 233 (1987).
- [3] <http://www.aist.go.jp/RIODB/riohomej.html>
- [4] Edmister, W. C. and Lee, B. I., *Applied Hydrocarbon Thermodynamics*, vol. 1, Gulf Publishing Company Book Division, Houston (1983).
- [5] Lee, B. I. and Kesler, M. G., A Generalized Thermodynamic Correlation Based on Three-Parameter Corresponding States, *AIChE J.*, **21(3)**, 510-527 (1973).
- [6] Riedel, L, *Chemie-Ing., Techn.*, **26**, 83 (1954).
- [7] <http://www.cdaj.co.jp/product/chemkin/chemkin.html>

Computed Property Data Base: CPDB – Development of an Integrated Software Tool for Molecular Design: MolWorks –

Sumie TAJIMA^a, Tohru YAGI^a, Tomoko FUKUDA^b and Umpei NAGASHIMA^{b*}

^aBest Systems Inc., 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

^bGrid Technology Research Center, National Institute of Advanced Industrial Science and Technology
1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

**e-mail: u.nagashima@aist.go.jp*

We have started to develop Computed Property Data Base: CPDB. A part of the CPDB has been opened to the public at RIO-DB (Research Information Database) of the National Institute of Advanced Industrial Science and Technology, Tsukuba from January 2005.

(<http://www.aist.go.jp/RIODB/hrmsdb/index.html>)

Thirteen molecular properties (Table 1) of 258,907 molecules were stored in CPDB. All of the properties are estimated using mainly the Joback method which is implemented in an Integrated Software Tool for Molecular Design: MolWorks. It is possible to retrieve not only molecular property of input but also molecules which have inputted properties.

CPDB is an attempt to develop a system for efficient molecular screening using estimated molecular property.

Keywords: Property Estimation, Joback Method, Molecular Screening, Property Data Base