

1. はじめに

『化学物質の審査及び製造等の規制に関する法律』(化審法)は、化学物質の環境経路による人健康へ悪影響を未然に防ぐために制定された法律である。本法律において、事業者は、新たな化学物質(新規化学物質)を製造又は輸入をする際には、各種有害性試験データを国へ提出し、審査を受けることが義務付けられている。一方、本法律施行以前に製造または輸入が行われていた化学物質(約2万物質)は、既存化学物質と称され、国が有害性試験を実施しており、現在まで約1500物質の評価が終了している。化審法における現行の有害性試験は、多大な費用と期間を必要とすることから、構造活性相関を活用した迅速かつ低コストな有害性評価手法の導入が期待されている。

化審法既存物質の試験データは公開されており、これらをトレーニングセットとした構造活性相関モデルは数多く開発されている。特に生分解性試験及び生物濃縮性試験のデータは、同一の試験方法で行われた世界最大の試験データ集となっているため、世界の主要な予測ソフトウェアのほとんどは、本データ集をトレーニングセットとして用いており、これらのソフトウェアは、欧米諸国の一部の当局により規制への活用がなされている。

一方、新規化学物質は既存化学物質と同一の方法により試験がなされているが、その試験データは非公開であるため、これまで構造活性相関モデルの開発には利用されてこなかった。製品評価技術基盤機構は、化審法新規化学物質の生分解性試験及び生物濃縮性試験の審査業務を担当しているため、これらの試験報告書を保管しているが、近年、これら新規化学物質の試験データを構造活性相関モデルの開発に役立てるための検討を開始した。

新規化学物質は、近年における化学物質の開発傾向が反映されており、その化学構造は既存化学物質と比較すると極めて多様なものとなっている。従って、既存化学物質をトレーニングセット

とした構造活性相関モデルを新規化学物質の試験データを用いて検証することにより、各モデルの適用限界や今後改良すべき点を、より明確にすることができるものと思われる。本発表では、化審法既存化学物質の試験データをトレーニングセットとした代表的な4つの生分解性予測モデル(BIOWIN5、BIOWIN6、生分解性予測システム、CATABOL)の予測性を、化審法新規化学物質の生分解性試験データを用いて検証した結果を報告する。

2. バリデーションセット

化審法では標準の生分解性試験として、OECD301C法が採用されている。本試験法では、好気性微生物の混合体を添加した水溶液中における化学物質の生分解性を生物化学的酸素要求量(BOD)によって測定する。28日後のBODが60%以上の場合、易分解性と判定され、60%未満の場合、難分解性と判定される。

本研究では、平成14年度までに届出られた化審法新規化学物質の生分解性試験の報告書を調査し、標準法で試験が行われた有機化合物のうち、構造が特定可能な物質(高分子や混合物でないもの)の試験データ(易分解性313物質、難分解性808物質、合計1121物質)をバリデーションセットとして選択した。

3. 検証した予測モデル

3.1 BIOWIN 5

42の部分構造及び分子量を記述子とし、下式により予測がなされる[1]。

$$Y_j = a_0 + a_1 f_1 + a_2 f_2 + \dots + a_n f_n + a_m MW + e_j \quad (1)$$

ここで、 Y_j は化学物質Jが易分解性となる確率、 f_n は化学物質Jにおける部分構造nの数、 a_0 は切片、 a_n は部分構造nの回帰係数、MWは分子量、 a_m は分子量の回帰係数、 e_j は誤差項を表す。 Y_j が0.5以上で易分解性、0.5未満で難分解性と判定される。

トレーニングセットとして589物質の試験データが用いられており、易分解性物質の Y_j を1、難分解性の Y_j を0とし、最小二乗法により各部分構

*sakuratani-yuki@nite.go.jp

造の回帰係数が算出されている。

3.2 BIOWIN 6

記述子及びトレーニングセットは BIOWIN5 と同じであるが、非線形の予測式 (式 2) が用いられている [1]。

$$Y_j = \frac{\exp(a_0 + a_1 f_1 + a_2 f_2 + \dots + a_n f_n + a_m MW)}{1 + \exp(a_0 + a_1 f_1 + a_2 f_2 + \dots + a_n f_n + a_m MW)} \quad (2)$$

BIOWIN 5,6 共に、米環境省が同国の新規化学物質の審査において活用している。

3.3 生分解性予測システム

フローチャートにより物質を分類し、予測を行う。生分解性は、物質の骨格構造と置換基の組み合わせにより決定されるとの経験則に基づいた予測がなされる [2]。フローチャートにおいて、予測対象物質はまず、分子量 500 以上、分子量 500 未満に分類され、分子量 500 未満の物質では、骨格構造を基に、鎖状脂肪族、単環ベンゼン誘導体等、9つのグループに分類される。それぞれグループには、置換基の種類、位置、数を基にした判定条件が設けられており、予測対象物質は、最終的に、易分解、難分解、予測困難のいずれかのグループに振り分けられる。本システムは、旧通産省プロジェクトのもと、(財)化学物質評価研究機構により開発され、同機構のホームページにおいて公開されている。現在、当 NEDO プロジェクトでは、本システムの改良版の開発を行っている。

3.4 CATABOL

612 階層の生分解反応式で構築された予測エンジンにより、予測対象物質の反応部位が順次検索され、変換されることにより、無機化までの生分解経路が予測される [3]。また、各反応式には、進行確率が与えられており、下式により親物質及び代謝物の残留率、BOD 値が算出される。

$$y = \frac{\Delta k_1}{k_{TOD}} P_1 + \frac{\Delta k_2}{k_{TOD}} P_1 P_2 + \dots + \frac{\Delta k_l}{k_{TOD}} P_1 P_2 P_3 \dots P_l \quad (3)$$

ここで P_i は i 番目の反応の進行確率、 y は BOD、 $\Delta k_i/k_{TOD}$ は理論的酸素要求量に対する各反応における酸素要求量の割合で、

$$k_{TOD} = \sum_{i=1}^l \Delta k_i \quad (4)$$

の関係がある。トレーニングセットの 723 物質について、エキスパートにより決定された無機化までの生分解経路を基に、式 3 が構築されており、非線形最小二乗法により反応進行確率 P が算出されている。本モデルは、カナダ環境省が同国の新規物質の審査に活用している。

4. 結果と考察

入力した物質のうち、易分解または難分解として予測の判定がなされた物質の割合は、生分解性

予測システムが 88%、その他の 3 つのモデルが 98%であった。

表 1 に検証した生分解性予測モデルのトレーニングセット (既存物質) 及びバリデーショナルセット (新規物質) に対する的中率を示す。的中率は、易分解又は難分解と予測された物質数に対する正しく予測された物質数の割合で表している。全てのモデルにおいて、バリデーショナルセットでは 75%以上の的中率が得られ、トレーニングセットに対する的中率との差は 10%以内に収まった。

表 1 検証した生分解性予測モデルの的中率

モデル	的中率 (%)	
	トレーニングセット (既存物質)	バリデーショナル セット(新規物質)
BIOWIN5	82	76
BIOWIN6	83	77
生分解性予測システム	82	81
CATABOL	90	81

また、全てのモデルにおける共通の傾向として、
① 難分解性と予測した物質の的中率(82-85 %)に対し、易分解性と予測した物質の的中率が低い (56-75 %)。
② 鎖状脂肪族に対する的中率が低い (66-72 %)。
③ エステル基等、水との反応性が高い官能基を有する物質の的中率が低い (69-77 %)。
などが見出された。

それぞれの予測モデルにおける長所・短所の詳細については、当日報告する。

謝辞

本研究は、NEDO 化学物質総合管理プログラム「既存物質点検の加速化プロジェクト」における研究の一環として行われた。

参考文献

- [1] J. Tunkel, P. H. Howard, R. S. Boethling, W. Stiteler, H. Loonen, Environ. Toxicol. Chem., 19, 2478-2485, (2000).
- [2] K. Hiromatsu, Y. Yakabe, K. Katagiri, T. Nishihara, Chemosphere, 41, 1749-1754 (2000).
- [3] J. Jaworska, S. Dimitrov, N. Nikolova, O. Mekenyan, SAR QSAR Environ. Res., 13, 307-323 (2002).