Science of the Total Environment 497–498 (2014) 60–67

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Comparison of ready biodegradation estimation methods for fragrance materials

Robert Boethling *,†

United States Environmental Protection Agency, Office of Pollution Prevention and Toxics, 1200 Pennsylvania Ave, NW, Washington, DC, USA

HIGHLIGHTS

- Fragrances were used to evaluate no-cost ready biodegradability prediction methods.
- The comparison was based on accuracy, sensitivity, specificity and MCC.
- Overall accuracy (ca. 70%) varied little, but other measures more widely.
- Based on MCC, VEGA and Biowin3/6 performed best and are suitable for screening.
- Removing fragrances with one quaternary C improved predictivity.

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form 22 July 2014
Accepted 22 July 2014
Available online xxxx

Editor: D. Barcelo

Keywords:
Fragrance
Ready biodegradation
Estimation
Biowin©
QSAR
VEGA©

ABSTRACT

Biodegradability is fundamental to the assessment of environmental exposure and risk from organic chemicals. Predictive models can be used to pursue both regulatory and chemical design (green chemistry) objectives, which are most effectively met when models are easy to use and available free of charge. The objective of this work was to evaluate no-cost estimation programs with respect to prediction of ready biodegradability. Fragrance materials, which are structurally diverse and have significant exposure potential, were used for this purpose. Using a database of 222 fragrance compounds with measured ready biodegradability, 10 models were compared on the basis of overall accuracy, sensitivity, specificity, and Matthews correlation coefficient (MCC), a measure of quality for binary classification. The 10 models were VEGA© Non-Interactive Client, START (Toxtree©), Biowin©1–6, and two models based on inductive machine learning. Applicability domain (AD) was also considered. Overall accuracy was ca. 70% and varied little over all models, but sensitivity, specificity and MCC showed wider variation. Based on MCC, the best models for fragrance compounds were Biowin6, VEGA and Biowin3. VEGA performance was slightly better for the <50% of the compounds it identified as having "high reliability" predictions (AD index >0.8). However, removing compounds with one and only one quaternary carbon yielded similar improvement in predictivity for VEGA, START, and Biowin3/6, with a smaller penalty in reduced coverage. Of the nine compounds for which the eight models (VEGA, START, Biowin1–6) all disagreed with the measured value, measured analog data were available for seven, and all supported the predicted value. VEGA, Biowin3 and Biowin6 are judged suitable for ready biodegradability screening of fragrance compounds.
Published by Elsevier B.V.

1. Introduction

Fragrance materials are a group of structurally heterogeneous substances widely used in consumer products including cosmetics, detergents, fabric softeners, cleaning products, perfumes, skin creams, etc. (Salvito et al., 2002). As a result, the potential for human exposure to these substances is high, particularly for products that come into direct contact with the skin. In the consumer market, detergents and fabric softeners are the largest category by use in volume, and most of these materials are expected to be sewered and eventually enter domestic wastewater treatment. This creates the potential for widespread environmental occurrence and exposure for both humans and wildlife.

For most organic compounds, consideration of biodegradability is fundamental to the assessment of environmental exposure and thus risk. Predictive approaches, such as those based on quantitative structure–activity or structure–property relationships (QSAR or QSPR), can be used for untested compounds in order to help characterize their environmental persistence. In the European Union (EU), REACH (Registration, Evaluation, Authorization and restriction of CHemicals) requires that information be collected and registered for all chemical

* US Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, Washington, DC 20460, USA. Tel.: +1 202 564 8533 (office); fax: +1 202 564 8892.
E-mail addresses: Robert_Boethling@abtassoc.com, robertsboethling@gmail.com.
† Current address: Abt Associates, 4550 Montgomery Avenue, Suite 800 North, Bethesda, MD 20814, USA.

http://dx.doi.org/10.1016/j.scitotenv.2014.07.090
0048-9697/Published by Elsevier B.V.

substances manufactured or imported in an amount exceeding one ton per year, but alternative methods such as QSARs can be used for this purpose (Boethling et al., 2009) (http://www.orchestra-qsar.eu/regulatory-context/the-reach-requirements-for-qsar). Another use of QSARs for biodegradability involves screening chemical substances before their synthesis, with the aim of designing alternatives that are less persistent than existing ones. This approach has found its home in Green Chemistry, which encourages the design of chemicals and chemical processes that reduce or eliminate negative environmental impacts (Anastas and Warner, 1998). In silico approaches like QSARs and rational design strategies are tools that can help achieve these objectives, by highlighting potential hazards before time and money are invested in commercial development of a product.

Existing QSARs for biodegradation aim to predict biodegradability from molecular structure alone, and generally are based on molecular fragments (substructures) or structural alerts, with prediction implemented through multiple regression equations or decision trees, respectively. Many biodegradation QSARs have been developed over the years. Regulatory use of QSARs for biodegradation was reviewed by Pavan and Worth (2008); subsequently, modeling of ready biodegradability in the context of green chemistry applications was reviewed by Rücker and Kümmerer (2012). Recently, several in silico models were compared by Pizzo et al. (2013). Regardless of the application, regulatory and green chemistry objectives are met most effectively if software exists and the model is freely available, or easily implemented without special tools or knowledge.

The objective in this work was to evaluate the performance of no-cost estimation programs with respect to prediction of ready biodegradability. For this purpose fragrance compounds with ready biodegradation data were used. Fragrance compounds are a good group for comparative study because they are structurally diverse, but their domain is limited to small molecules (molecular weight approximately 300 Da or less; Turin and Yoshii, 2014); molecular fragments or substructures are mostly limited to those that are well represented in small organic molecules and for which a substantial amount of test data exist; they consist almost exclusively of C, H, O and N atoms; and they are nondissociating—not ionized at typical environmental pH values (dissociation adds another layer of complexity and ionic compounds are less well treated by models).

An earlier and related work (Boethling, 2011) focused on a specific subgroup of fragrance compounds, musk odorants, where the incorporation of environmental attributes including biodegradability in chemical design was explored. In this work all fragrance compounds are used to compare estimation accuracy of eight models for which no-cost software exists and is downloadable to a user's personal computer (Biowin©1–6; VEGA©; START©); and two models based on inductive machine learning (IML1 and 2) that are easily run manually.

2. Methods

2.1. Data for evaluating estimation programs

A list of odorants with measured % degradation in one or more ready biodegradation tests was curated from various sources. The starting point was three files prepared under the US EPA's Design for Environment (DfE) program (http://www.epa.gov/dfe/) for use in developing DfE's Criteria for Fragrances (DfE, 2012a), a variation of the DfE Master Criteria (DfE, 2012b), which delineate the "low-concern" end of the ingredient hazard spectrum for DfE-labeled products. These sources were augmented by searches of the scientific literature and Internet resources, with the objective of finding additional fragrances with measured ready biodegradation data. These efforts yielded a file containing chemical structure and ready biodegradation data for 108 fragrance compounds.

The next step was to augment the 108-compound fragrance list by adding fragrances identified as new chemical substances in Premanufacture Notifications (PMNs) submitted under the Toxic Substances Control Act. This was accomplished by searching PMN files for the period from 1995 through mid-2013 for those with stated use as a fragrance or aroma chemical. Since fragrance compounds originally notified via PMN could (and in some cases, did) later enter commerce and the literature, subsequently being identified in the search for existing fragrance data, the file was searched for duplicate entries and these were identified and eliminated. A total of 114 substances had measured ready biodegradation data, and were added to the 108 existing fragrance compounds with ready biodegradation test data, to yield a file containing name, structure and other data for 222 compounds.

Biodegradation studies were carefully scrutinized for quality control purposes, in most cases by retrieving and reviewing the original report. No rigid criteria for purity of test substance were applied, but in general, test results had to be interpretable relative to a discrete or representative structure. Organization for Economic Cooperation and Development (OECD) ready biodegradation tests are 28-day screening tests with rigidly defined pass criteria (OECD, 2006). The pass level is set at 60% of theoretical for CO_2 and biological oxygen demand (BOD) tests, and 70% for tests based on disappearance of dissolved organic carbon (DOC). In regulatory applications a kinetics criterion also may be applied (called the '10-day window', by which the 60% or 70% pass level must be achieved within 10 days of degradation first meeting or exceeding 10%); but since 10-day window information often is not given in test reports, here substances are classified as readily biodegradable (RB) or not readily biodegradable (NRB) based only on the final pass level.

2.2. Biodegradation models and software

All programs were accessed or downloaded in 2013. The following is only a brief overview; details may be found in the cited source publications and several reviews (Pavan and Worth, 2008; Rücker and Kümmerer, 2012; Jaworska et al., 2003).

2.2.1. Biowin

The Biowin program is part of the Estimation Programs Interface (EPI) Suite© (Table 1) and contains six models whose purpose is to provide screening-level estimates of aerobic biodegradability for organic chemicals: Biowin1 = linear probability model; Biowin2 = nonlinear probability model; Biowin3 = expert survey ultimate biodegradation model; Biowin4 = expert survey primary biodegradation model; Biowin5 = Japanese Ministry of International Trade and Industry (MITI) linear model; and Biowin6 = MITI nonlinear model. Biowin1 and 2 constitute the original Biodegradation Probability Program and were the first two models developed (Howard et al., 1992). Biowin3 and 4 estimate the time required to achieve complete ultimate (Biowin3) or primary (Biowin4) biodegradation in a typical, or 'evaluative', aquatic environment. They were developed from estimates of biodegradability for 200 substances, provided by a panel of experts (Boethling et al., 1994). The output of the model for any given chemical is a number from one to five (continuous scale; fractional values permitted) that either can be used as it is, or if desired can be related to the words (e.g. 'days', 'weeks') used in the expert survey. In Biowin3 values >2.75 and <= 3.25 are assigned the term 'weeks'. Here 2.75 was used as the criterion for ready biodegradability, so that a calculated value >2.75 is a prediction of RB and is equivalent to a prediction of 'weeks' or an even shorter timeframe. For Biowin4 a criterion of 3.6 (i.e., <=3.6 is NRB) yielded a reasonable separation of RB and NRB compounds. Biowin5 and 6 are linear and nonlinear probability models developed exclusively from the MITI data (MITI, 2014). Biowin5/6 classify a substance as either RB or NRB based on the MITI ready biodegradation test (OECD 301C).

Characteristics of this test, as well as details of model development and validation, are given by Tunkel et al. (2000). The model estimates the likelihood that a chemical will degrade under the test conditions. Values of the dependent variable >0.5 usually are taken as predicting that the chemical will be RB in the OECD 301C test, but as with other