Comparison of small molecules biotransformation half-lives between activated sludge and soil: Opportunities for read-across?

Kathrin Fenner^{*,1,2,3}, Claudio Screpanti⁴, Peter Renold⁴, Marwa Rouchdi⁴, Bernadette Vogler¹,
Stephanie Rich^{1,5}

¹Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland.

²Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, 8092 Zürich, Switzerland.

³Department of Chemistry, University of Zürich, 8057 Zürich, Switzerland. ⁴Syngenta Crop Protection
AG, Chemical Research, Schaffhauserstrasse 101, CH-4332 Stein, Switzerland. ⁵Karlsruhe Institute of
Technology, Karlsruhe, Germany.

*Corresponding author (email: kathrin.fenner@eawag.ch)

This document is the accepted manuscript version of the following article:
Fenner, K., Screpanti, C., Renold, P., Rouchdi, M., Vogler, B., & Rich, S. (2020).
Comparison of small molecule biotransformation half-lives between activated sludge and
soil: opportunities for read-across? Environmental Science and Technology, 54(6),
3148-3158. <https://doi.org/10.1021/acs.est.9b05104>

Abstract

Compartment-specific degradation half-lives are essential pieces of information in the regulatory risk assessment of synthetic chemicals. However, their measurement according to regulatory testing guidelines is laborious and costly. Despite the obvious ecological and economic benefits of knowing environmental degradability as early as possible, its consideration in the early phases of rational chemical design is therefore challenging. Here, we explore the possibility to use half-lives determined in highly time- and work-efficient biotransformation experiments with activated sludge and mixtures of chemicals to predict soil half-lives from regulatory simulation studies. We experimentally determined half-lives for 52 structurally diverse agrochemical active ingredients in batch reactors with three concentrations of the same activated sludge. We then developed bi- and multivariate models for predicting half-lives in soil by regressing the experimentally determined half-lives in activated sludge against average soil half-lives of the same chemicals extracted from regulatory data. The models differed in how we accounted for sorption-related bioavailability differences in soil and activated sludge. The best performing models exhibited good coefficients-of-determination (R^2 of around 0.8), low average errors ($<$ factor of 3 in half-life predictions) and were robust in cross-validation. From a practical perspective, these results suggest that it may indeed be possible to read across from half-lives determined in highly efficient biotransformation experiments in activated sludge to soil half-lives, which are obtained from much more work- and resource-intensive regulatory studies, and that these predictions are clearly superior to predictions based on the output of the publicly available BIOWIN QSBR model. From a theoretical perspective, these results suggest that soil and activated sludge microbial communities, although certainly different in terms of taxonomic composition, may be functionally similar with respect to the enzymatic transformation of environmentally relevant concentrations of a diverse range of chemical compounds.

1 Introduction

2 Compartment-specific degradation half-lives are essential pieces of information in the risk assessment
3 of synthetic chemicals, namely in hazard assessment to evaluate contaminant persistence, and in risk
4 assessment as input into exposure models ¹⁻⁴. Experimental determination of half-lives currently
5 requires significant investments in terms of time and money, most notably because the different risk
6 assessment frameworks often request testing in so-called simulation studies for several relevant
7 compartments (e.g., soil, sediment, activated sludge, surface water), and because these laboratory
8 experiments should be carried out with radioactively labeled compounds and for each individual
9 compound separately to be acceptable in a regulatory context. Both from an environmental protection
10 and economic standpoint, it is highly desirable to consider the biodegradability of new synthetic
11 compounds more up-stream in the discovery processes and to include it as prioritization criterion in
12 the selection process during rational chemical design ⁵. However, during the lead exploration stage,
13 which aims at identifying leads over one or more chemical classes, any degradation testing would have
14 to be fast (i.e., a few days to weeks) and simultaneously applicable to a broad range of analogs
15 resulting from multiple synthesis campaigns. This cannot be achieved with the current type of
16 regulatory simulation studies, highlighting the need for new approaches to include benign-by-design
17 considerations early into the chemical discovery process.

18 Given that degradation by microbial communities is generally considered the mass balance-wise most
19 relevant environmental loss process for most organic synthetic chemicals ⁶, it has been suggested that
20 models predicting microbial biodegradation, i.e., quantitative structure-biodegradability relationships
21 (QSBRs), could be used as alternatives to experimental testing ⁷. These models would be applicable at
22 the early stages of the discovery process, and use of structure-based prediction models to evaluate fate
23 is even accepted in some regulatory frameworks ⁸. However, there are currently very few QSBRs
24 available that allow predicting compartment-specific biotransformation half-lives (rather than just
25 ready biodegradability outcomes) and across a wide range of compound classes ^{9, 10}. The model that is
26 most readily usable for that purpose is the BIOWIN primary survey model (PSM), which is part of the
27 US EPA EPI Suite package and which aims to assess primary biotransformation ¹¹. Its predicted scores
28 can be translated into half-life categories in water, which, as the developers suggest, can be further
29 extrapolated to compartment-specific primary biotransformation half-lives for soil and sediment by
30 applying correction factors ^{11, 12}. However, validation of BIOWIN PSM models against measured
31 biotransformation half-lives has typically yielded unsatisfactory results ¹²⁻¹⁴. A few more relationships
32 for prediction of primary biotransformation half-lives have emerged over the years (e.g., ^{13, 15}) but they
33 have so far not been externally validated and their predictive accuracy therefore remains unclear.

34 With the lack of high-quality QSBRs to predict compartmental half-lives, yet the availability of large
35 amounts of regulatory data from simulation studies, the question arises whether more could be learned
36 from a meta-analysis of the latter data. For instance, improved compartment-specific QSBRs might be

1 derived from large collections of half-life data produced according to consistent testing guidelines,
2 which could be used both for prediction and for rational design of derivatives with improved
3 biotransformation behavior ^{16, 17}. Also, comparison of half-life data across environmental
4 compartments would inform about opportunities to extrapolate biotransformation information between
5 compartments. Read-across between environmental compartments is well accepted in ecotoxicology
6 where experimentally determined aquatic endpoints are used to obtain estimates for toxicity in
7 sediments or soils ^{18, 19}. This is typically done using the so-called equilibrium partitioning approach,
8 which hinges on two major assumptions, i.e., (i) that only substances freely dissolved in pore water are
9 available to be taken up and affect the resident organisms, and (ii) that sensitivities towards the
10 toxicant are similar for aquatic and terrestrial organisms. Similar considerations could be applicable
11 when attempting to extrapolate microbial biotransformation half-lives between environmental
12 compartments, but only few attempts to do so have been reported.

13 For the specific substance class of sulfonylurea herbicides, Berger et al. ²⁰ found a good correlation
14 between reaction rates in native sediments and soils and concluded that microbial transformation in
15 soil and sediment, in principle, followed similar mechanisms. They also noted that, particularly for
16 half-lives in soil, consideration of differences in adsorption properties between sulfonylureas lead to
17 improved half-life predictions. Even earlier, Leoni et al. ²¹ had correlated half-lives in activated sludge
18 with known soil half-lives for nine pesticides and found an exponential relationship that was consistent
19 with theoretical considerations, but rather uncertain across the nine compounds. Around that same
20 time, Blok ²² derived a theoretical framework to predict how biotransformation half-lives for a given
21 substance would compare between sewage treatment plants and the receiving river environment, but
22 were not able to support these considerations with data. More recently, Honti et al. ²³ studied the
23 behavior of three substances across four OECD 308 and 309 setups and showed that substance-
24 specific biomass- and bioavailability-normalized biotransformation rate constants could be estimated
25 that were valid across all four systems.

26 One reason for the lack of studies exploring options for read-across in biotransformation half-lives
27 might have been the lack of sufficiently large sets of biotransformation data of good quality and
28 measured under consistent conditions. However, recent developments in mass spectrometry affording
29 increased mass resolution and sensitivity have opened up new opportunities to measure
30 biotransformation of large numbers of low concentrated substances in mixtures, and to thus quickly
31 generate large and consistent sets of biotransformation data ^{24, 25}. Furthermore, we have recently
32 published a large database of regulatory pesticide soil half-lives including metadata on soil type and
33 experimental conditions ¹⁶. In this study, we aimed to exploit these novel opportunities to evaluate the
34 possibility to use half-lives obtained in biotransformation experiments with mixtures of chemicals in
35 activated sludge to predict soil half-lives, which would otherwise have to be determined in much more
36 work- and resource-intensive regulatory simulation studies.

1 **Methods**

2 *Compound selection*

3 Compounds initially selected for developing and testing a read-across approach between activated
4 sludge and soil half-lives include 52 agrochemical active ingredients classified as fungicides,
5 herbicides, or insecticides. The 52 compounds are listed in Table 1, along with information on
6 substance class, half-lives in laboratory studies with different soils as extracted from the draft
7 assessment reports from pesticide registration ¹⁶, typical organic carbon-water partition coefficient
8 (K_{oc}) values, and median calculated freely dissolved fractions in the soils tested. Additionally, a
9 complete list of all 52 compounds along with their respective CAS numbers and chemical structures is
10 given in the Supporting Information (SI), Table S1. To test the robustness of the read-across approach
11 towards a wide variety of chemical structures, compounds were selected from diverse agrochemical
12 classes, and to cover a range of soil biotransformation half-lives (i.e., $DT_{50,soil}$ from 0.09 to 690 d) and
13 organic carbon-water partitioning behavior (i.e., K_{oc} from 1.6 to 35'000 L/kg). Transformation rate
14 constants for 19 of the selected compounds have previously been measured in activated sludge ²⁵,
15 whereas for the other 33 compounds rate constants in activated sludge have not previously been
16 determined to the best of our knowledge.

1 **Table 1:** 52 active ingredients initially selected for testing of read-across approach with information on their substance class, half-lives in soil ($DT_{50,soil}$) from
2 regulatory testing, the total number of half-life entries, typical organic carbon-water partition coefficients (K_{oc}), and calculated freely dissolved fraction in soil
3 ($f_{aq,soil}$).

Compound	Substance Class ^a	$DT_{50,soil}$ (d) ^b	n_{DT50}	K_{oc} (L/kg) ^a	$f_{aq,soil}$ (%) ^c
Amisulbrom	Sulfonamide fungicide	97 (22;348)	6	284	2.9 (2.7;3.2)
Asulam	Carbamate herbicide	4.7 (2.7;10.3)	5	20*	32 (28;39)
Azoxystrobin	Strobilurin fungicide	93 (56;153)	8	589	1.5 (0.64;4.6)
Benzovindiflupyr	Amide/pyrrazole fungicide	639 (349;1000)	7	5578	0.30 (0.19;0.45)
Bixafen	Amide/pyrrazole fungicide	> 365 (n.a.)	4	3869*	0.44 (0.39;0.56)
Carbendazim	Benzimidazole fungicide	34 (26;44)	6	223*	6.3 (4.9;12.1) [§]
Carbetamide	Carbamate herbicide	9.0 (3.3;35)	6	89*	6.4 (5.1;8.0)
Chlorantraniliprole	Diamide insecticide	219 (20;886)	24	362	2.7 (0.84;7.9)
Clomazone	Isoxazolidinone herbicide	39 (21;86)	6	300	7.3 (4.3;11.9) [§]
Clothianidin	Neonicotinoid insecticide	N/A	N/A	123	N/A
Cyantraniliprole	Diamide insecticide	38 (9.2;136)	6	241	6.0 (4.4;8.3)
Cyclaniliprole	Diamide insecticide	840 (444;1730)	9	790*	1.9 (1.0;3.2)
Dicamba	Benzoic acid herbicide	4.9 (2.9;8.0)	6	12.4*	46 (37;58)
Dimethenamid	Chloroacetamide herbicide	14 (7.8;43)	5	69*	7.0 (2.4;13.1)
Dimoxystrobin	Amide fungicide/Strobilurin fungicide	271 (88;1203)	10	486*	1.2 (0.30;3.8)
Diuron	Phenylamide herbicide	79 (20;372)	7	813	2.9 (1.6;4.0)
Fenhexamid	Hydroxyanilide fungicide	0.29 (0.08;0.74)	8	475	2.6 (2.2;3.0)
Fenoxycarb	Carbamate insecticide	1.3 (0.25;4.0)	7	1816*	0.60 (0.44;0.71)
Fipronil	Phenylpyrazole insecticide	233 (42;747)	8	727*	1.7 (0.77;2.5)
Flonicamid	Pyridine insecticide/aphicide	1.2 (0.7;2.4)	5	1.6	90 (88;95)
Florasulam	Sulfonanilide herbicide	3.5 (0.46;45)	21	22	30 (20;41)
Fluopicolide	Benzamide fungicide	380 (196;1148)	11	321*	2.0 (1.4;2.8)
Fluopyram	Benzamide fungicide	332 (162;746)	12	279*	5.3 (3.2;7.3)
Flupyradifurone	Butenolide insecticide	65 (33;242)	16	98.4	12 (8.4;16)
Flutianil	Thiazolidine fungicide	336 (284;378)	5	35492*	N/A
Fluxapyroxad	Anilide fungicide	237 (69;810)	6	728*	0.93 (0.62;1.2)

Imidacloprid	Neonicotinoid insecticide	186 (106;365)	6	225*	3.1 (1.4;5.2)
Iprovalicarb	Carbamate fungicide	5.8 (5.1;6.7)	4	106	11 (8.8;12.8)
Isofetamid	Amide fungicide	44.1 (22;97)	6	489*	2.5 (1.6;3.8)
Isoproturon	Urea herbicide	11.5 (7.1;30)	14	122	15 (11;18)
Isopyrazam	Pyrazolecarboxamide fungicide	258 (107;976)	8	2416	N/A
Kresoxim-methyl	Strobilurin fungicide	0.46 (0.36;0.56)	3	308*	2.6 (1.4;5.2)
Mandipropamid	Amide fungicide	53 (12.6;159)	23	847*	1.2 (0.39;2.5)
Mesosulfuron-methyl	Sulfonylurea herbicide	42.4 (7.62;155)	9	92*	12 (7.8;20)
Mesotrione	Benzoylcyclohexanedione herbicide	11.7 (3.8;34)	19	122	9.7 (2.3;18)
Metaflumizone	Semicarbazone insecticide	256 (186;423)	5	30714	0.037 (N/A)
Napropamide	Alkanamide herbicide	343 (120;670)	8	839	2.0 (1.1;3.7)
Oxathiapiprolin	Oxazole, pyrazole, thiazole fungicide	76 (18;135)	12	9674	0.19 (0.13;0.28)
Picoxystrobin	Methoxyacrylate strobilurin fungicide	23 (19;29)	9	965	1.2 (0.92;1.7)
Pinoxaden	Phenylpyrazole herbicide	0.32 (0.08;1.8)	16	349*	3.2 (2.3;9.2)
Proquinazid	Quinazolinone fungicide	136 (40;449)	7	12870*	0.070 (0.0090;0.24)
Pyroxsulam	Sulfonamide herbicide	3.9 (0.8;17)	33	33.2	30 (18;57)
Sedaxane	Pyrazolecarboxamide fungicide	83 (58;150)	9	534*	2.5 (1.4;9.4)
Spirodiclofen	Tetronic acid acaricide	6.4 (1.1;13)	5	31037	0.2 (0.038;1.1) [§]
Spiromesifen	Tetronic acid acaricide	5.3 (2.6;18)	6	30900	0.1 (0.04;0.67) [§]
Spirotetramat	Tetramic acid insecticide	0.38 (0.08;2.2)	6	289	5.5 (4.8;7.0)
Sulfoxaflor	Sulfoximine insecticide	0.087 (0.04;0.26)	5	41	17 (11;28)
Tembotrione	Benzoylcyclohexanedione herbicide	12.4 (4.1;42)	9	66	12 (6.7;18)
Terbutylazine	Triazine herbicide	110 (58;455)	22	231*	4.5 (2.7;6.8)
Topramezone	Benzoylpyrazole herbicide	201 (86;421)	12	171	7.2 (4.3;10.4)
Trinexapac-ethyl	Cyclohexanecarboxylate derivative	0.12 (0.03;0.32)	11	280*	16 (1.4;33) [§]
Valifenalate	Acylamino acid fungicides	0.13 (0.04;0.36)	4	753	2.3 (1.9;2.8)

¹ ^a Substance classes and typical K_{oc} values were collected from the Pesticide Properties Database of the University of Hertfordshire³⁰, or directly from harmonized classification and labeling (CLH) dossiers from the European Chemicals Agency (i.e., for benzovindiflupyr). For K_{oc} , where possible linear partition coefficients are given, but in some cases only Freundlich adsorption coefficients were available (the latter marked with *). ^b Half-lives were used as collected in the Eawag-Soil database¹⁶, and are tabulated as geometric mean, minimum and maximum values. ^c Freely dissolved fractions were calculated for each soil half-life according to the metadata reported in Eawag-Soil and as given in eqs. 3 and 4. Fractions marked with [§] were calculated with alternative eq. S1 given in SI section S4. Freely dissolved fractions are tabulated as geometric mean, minimum and maximum values.

Activated sludge biotransformation experiments

Seventeen of the compounds were provided as pure solid standards by Syngenta (Crop Protection, Switzerland), and the other 35 compounds already existed as stock solutions dissolved in organic solvents (ethanol, methanol, acetone, or acetonitrile) at a concentration of 1.0 g/L in our laboratories. New stock solutions were created from the 17 pure solid standards at a concentration of 1.0 g/L. A spike mix with all 52 pesticides at an individual concentration of 6 mg/L was created before the start of the experiment by adding 60 μ L of each stock solution to an empty 10 mL amber glass vial. Details on the stock solutions as well as theoretical compound concentrations in the spike mix are given in Table S2 in the SI.

The experimental setup for the biotransformation experiments with activated sludge was adopted from ²⁶, except that the sludge was adjusted to three different total suspended solids (TSS) concentrations to obtain kinetic estimates for a wide range of compounds, i.e., including compounds with very fast or slow degradation, and more strongly sorbing compounds (see Section S2 in SI for details). The experiments were carried out from August 8th to August 10th 2018 with activated sludge sampled from the nitrification basin of the municipal wastewater treatment plant of Dübendorf, Switzerland (ARA Neugut, serving approx. 105'000 population equivalents, hydraulic retention time (without settling tank) of 18 h, sludge retention time in biological treatment of 13 d). The three final TSS concentrations in the biomass stocks were 13.14 ± 0.08 g TSS/L, 3.13 ± 0.03 g TSS/L and 0.52 ± 0.03 g TSS/L, and used to set up high biomass (HB), normal biomass (NB), and dilute biomass (DB) experiments, respectively. Initial pH, ammonium concentrations and oxygen uptake rates of the biomass stocks are given in Table S3 in the SI. The stocks were then used to set up three different types of batch reactors, namely biotransformation experiments (BEs), sorption control experiments (SEs), and abiotic control experiments (AEs). Altogether 20 reactors were prepared, i.e., nine BE reactors with freshly sampled activated sludge (triplicates for each biomass concentration), nine SE reactors with autoclaved sludge (20 mins, 103 kPa, 121 °C, triplicates for each biomass concentration), and two AE reactors with autoclaved sludge filtrate. Reactors were prepared by plating aliquots of the spike mix before sludge addition to yield initial concentrations of 6 μ g/L for each individual compound. 1.5 mL subsamples were taken at different time points over a 2 d period and analyzed with LC-HRMS. Further details on reactor setup, sampling and LC-HRMS analysis are given in Section S2 in SI.

For comparison purposes, a second experiment with activated sludge from a different source was run at only one sludge concentration. Details on this experiment are given in the SI, Section S5.

Derivation of half-lives from activated sludge experiments

R 3.5.0 coupled with RStudio 1.0.143 was used to further process the concentration-time series for each compound. To obtain pseudo-first-order biotransformation rate constants in activated sludge, linear regression analysis was applied to the natural log-transformed concentration-time series. For each biomass concentration and compound, an average rate constant, $k_{\text{deg,sludge},i,j}$, was obtained by simultaneously fitting across the concentration-time series from the triplicate reactors, where i stands for the individual compounds and j stands for the different biomass concentrations (j =HB, NB, DB) in the activated sludge experiments. Concentrations below 300 ng/L (i.e., 5% of the initial spiked concentration at which for some compounds no reasonable chromatographic peak, i.e., less than three full-scans, and deviation from first-order kinetics was observed) were not included for linear regression. The validity of the first-order kinetic assumption was evaluated for each individual compound considered in this study by assessing residual trends and R^2 as goodness-of-fit parameter (Table S4 and Figure S6 in the SI). Half-lives were calculated from the rate constants thus obtained as

$$DT_{50,\text{sludge},i,j} = \ln(2)/k_{\text{deg,sludge},i,j}.$$

To obtain average, compound-specific sludge half-lives for regression modeling, we further normalized the half-lives for differences in biomass and bioavailability at the different biomass concentrations according to eq. 1.

$$DT_{50,\text{sludge,norm},i,j} = DT_{50,\text{sludge},i,j} \cdot B_j \cdot f_{\text{aq,sludge},i,j} \quad \text{eq. 1}$$

where $DT_{50,\text{sludge,norm},i,j}$ stands for the biomass- and bioavailability-normalized half-life of compound i determined for experiment j (j =HB, NB, DB) and has dimensions of d*g TSS/L, B_j stands for the biomass concentration in experiment j as approximated by the measured TSS concentration in g TSS/L, and $f_{\text{aq,sludge},i,j}$ stands for the observed dissolved fraction of compound i in experiment j , which is used here as a proxy for the concentration that is bioavailable to the bacteria. The dissolved fractions were calculated for each of the three experiments (HB, NB, DB) as the ratio of the aqueous phase concentrations in the biotransformation experiment at time 0, $[C_{0,i}]_{\text{BE}}$, versus the abiotic control concentrations at time 0, $[C_{0,i}]_{\text{AE}}$, for each compound i (eq. 2). The latter was assumed to be a better measure of the spiked concentration than the nominal spike concentration. Note that measured $[C_{0,i}]_{\text{BE}}$ values might slightly overestimate actual freely dissolved and hence bioavailable concentrations since they include the substance pool bound to dissolved organic carbon in the sludge reactors. However, for substances with K_{oc} values $< 10^4$ L/kg, which effectively is the case for those substances used in model development later, the relative error is estimated to be less than 10% (assuming a DOC of 10 mg/L and, to obtain a worst-case estimate, the lowest used sludge concentration of 0.5 g TSS/L).

$$f_{\text{aq,sludge},i} = \frac{[C_{0,i}]_{\text{BE}}}{[C_{0,i}]_{\text{AE}}} \quad \text{eq. 2}$$

In our original study design we had included sorption control experiments to assess the dissolved fraction relative to the abiotic controls. However, upon data analysis, we found, first, that $[C_{0,i}]_{\text{BE}}$ was consistent with the aqueous concentrations in the sorption control experiments for most compounds (see, e.g., fipronil and fluopyram in Figure S6), but, second, that this was not the case for a few compounds without own surrogate standards, likely due to changes in matrix composition upon autoclaving (see, e.g., flonicamid and iprovalicarb in Figure S6). We therefore concluded that measured concentrations in the BE reactors at time 0, i.e., $[C_{0,i}]_{\text{BE}}$, were a more accurate representation of dissolved fractions in the biotransformation experiments and consistently used eq. 2 across all compounds. Note that, in doing so, we assumed that the dissolved fraction was constant throughout the experiment. This might not be entirely correct, e.g., if sorption constants were changing with decreasing substance concentrations.

Final compound-specific sludge half-lives, $DT_{50,\text{sludge,norm},i}$, were determined for all compounds that showed noticeable degradation (i.e., $k_{\text{deg,sludge},i,j}$ values $> 0.11 \text{ d}^{-1}$, see results section for rationale for this value) for at least one biomass concentration (HB, NB, DB). For compounds meeting this criterion in only one experiment, $DT_{50,\text{sludge,norm},i}$ was set equal to $DT_{50,\text{sludge,norm},i,j}$. For the others with noticeable degradation at two or all three biomass concentrations, $DT_{50,\text{sludge,norm},i}$ was calculated as the geometric mean of the $DT_{50,\text{sludge,norm},i,j}$ values (see Table S4 for $k_{\text{deg,sludge},i,j}$ values used for the calculation of $DT_{50,\text{sludge,norm},i,j}$ and for final $DT_{50,\text{sludge,norm},i}$ values used in regression analysis).

Calculation of half-lives and freely dissolved concentrations in soil

Soil half-lives for each compound ($DT_{50,\text{soil},i}$) and metadata on study conditions were exported from the Eawag-Soil package contained in enviPath (envipath.org²⁷) on November, 14, 2018 (complete list of exported half-lives and metadata provided as electronic SI). For each soil half-life, the freely dissolved fraction of compound i under the given study conditions for which the respective half-life was reported, $f_{\text{aq,soil},i}$, was calculated according to eq. 3. In using eq. 3, we assume that sorption is dominated by sorption to soil organic carbon, which is a reasonable approximation for neutral and anionic substances²⁸. While this is not necessarily the case for cationic or zwitterionic substances, no such substances were included in our study. Five compounds, namely mesotrione, pyroxsulam, tembotrione, topramezone, and trinexapac-ethyl, were acids with pK_a values in the range of the pH values of the studied soils. For those compounds, pH-dependence of K_{oc} was likely and therefore considered in the calculation of $f_{\text{aq,soil},i}$ according to the models and equations given in the SI (Section S4).

$$f_{\text{aq,soil},i} = \frac{1}{1 + f_{\text{oc}} K_{\text{oc},i} r_{\text{sw}}} \quad \text{eq. 3}$$

where f_{oc} is the study condition-specific organic carbon content of the soil in kg OC/kg soil solids, $K_{\text{oc},i}$ is the compound-specific organic-carbon water partition coefficient in L water /kg OC and r_{sw} is the study condition-specific ratio of soil solids to volume of water in kg soil solids/L water and was calculated for each soil and study condition as given in eq. 4.

$$r_{\text{sw}} = \frac{1}{\sigma / 100 \cdot \text{WSC}} \cdot 100 \quad \text{eq. 4}$$

where σ is the experimental soil moisture content given in % of the water storage capacity (WSC , in g water/100 g soil solids) reported for a given suction tension. Alternatively, if WSC was not reported, r_{sw} was estimated from soil bulk density as explained in the SI (Section S4).

Bi- and multivariate regression analyses

Different bi- and multivariate models were developed with the aim of predicting soil half-lives from normalized sludge half-lives and are given in Table 2. Models differed mostly in how differences in bioavailability in soil between compounds were accounted for. In model 1, bioavailability differences were not accounted for at all. In contrast, in deriving models 2-5, the assumption was made that only the freely dissolved concentration of a compound was available for uptake and biotransformation by members of the microbial communities, and that, hence, observed biodegradation kinetics in soil were directly related to the freely available fraction of a compound in soil. These assumptions only hold if biotransformation kinetics are limited by biological transformation rather than desorption from the matrix. While this is known to not always be valid for hydrophobic compounds and in actual heterogeneous and dry soils²⁹, we suggest that it is a reasonable simplification for the mostly polar agrochemicals studied here²⁸ and given the fact that soil half-lives in this study are retrieved from regulatory biotransformation tests that are run with sieved soils adjusted to field capacity.

In the following, model 2 is formulated as a bivariate model with the freely dissolved fraction explicitly calculated from $K_{\text{oc},i}$ and soil properties as given in eqs. 3-4, and models 3-5 are formulated as multivariate models including more qualitative sorption proxies (i.e., $\log P$, $\log D$ and $\log K_{\text{oc},i}$). To benchmark the read-across approach against a QSBR-based approach, scores from BIOWIN 4 for primary biotransformation as implemented in the EPI SuiteTM package¹¹ were obtained for all compounds and used directly in regression analysis against soil half-lives.

For all bi- and multivariate analyses, we calculated average compound-specific soil half-lives ($DT_{50,soil,i}$) as the geometric mean of all soil half-lives of the specific compound reported for different study conditions (Table 1). For derivation of model 2, average compound-specific freely dissolved fractions ($f_{aq,soil,i}$) were calculated as the geometric mean of the freely dissolved fractions for the different study conditions (Table 1). The geometric rather than arithmetic mean was chosen for the latter because each freely dissolved fraction is associated with a specific half-life for a given set of study conditions. In this way, average freely dissolved fractions were consistent with the calculation of geometric mean half-lives.

R 3.5.1 coupled with RStudio 1.1.463 was used to develop the models and to calculate coefficients of determination (R^2), root-mean-squared errors (RSE) and prediction intervals as measures of goodness-of-fit. The two best performing models were additionally subjected to leave-one-out cross-validation.

Results

Biotransformation in activated sludge

Based on the concentration-time series data, five of the 52 spiked compounds (i.e., amisulbrom, metaflumizone, pinoxaden, spiroadiclofen, and spiromesifen) were excluded from further analysis because they were not at all or only sporadically detected in the aqueous phase of the sludge reactors, even at the early experimental time points (see Figure S6 in SI). Pinoxaden, spiroadiclofen and spiromesifen are known to undergo rather fast abiotic hydrolysis (i.e., with half-lives typically below 7 days at slightly acidic or basic pH)³⁰, which might explain why they were not detected in the sludge reactors. This explanation is also consistent with their mostly low soil half-lives, i.e., < 1 d for pinoxaden, and < 7 d for the more strongly sorbing compounds spiroadiclofen and spiromesifen. Metaflumizone exhibits one of the highest K_{oc} values, i.e. > 30'000 L/kg, amongst the compounds studied, which results in an estimated freely dissolved fraction of less than 10% even in the low biomass reactor (DB). This might explain why metaflumizone was hardly observable in the aqueous phase of the sludge reactors, and is also consistent with its rather high average soil half-life of 256 d, which can be rationalized by its very low availability in soil ($f_{aq,soil} < 0.1$ %). For the less strongly sorbing amisulbrom, we did not find any plausible explanation for its lack of detectability in the sludge experiments that would be consistent with its rather high average soil half-life of 97 d.

For the remaining 47 compounds, aqueous phase concentrations were well detectable and quantifiable, and degradation rate constants could hence be estimated. Exemplary concentration-time series for biotransformation and control experiments, including first-order

kinetic fits, are shown for picoxystrobin and fluopyram in Figure 1, and for the other compounds in Figure S6 in the SI. First-order degradation rate constants and R^2 for all combinations of reactors and compounds for which they could be derived are given in Table S4 in the SI. It was found that R^2 values for the first-order fits were above 0.6 for 71 out of 73 combinations of compounds and experiments where more than 20% loss relative to the initial concentration was observed during the experiment, corresponding to a $k_{\text{deg,sludge},i,j}$ value of 0.11 d^{-1} . For combinations of compounds and experiments exhibiting lower degrees of degradation, R^2 values were lower due to the fact that experimental and analytical uncertainties dominated the observed concentration-time series data.

10

Figure 1: Example concentration-time series for two compounds (i.e., picoxystrobin (open symbols) and fluopyram (closed symbols)). A: average concentrations and standard deviations from triplicate experiments are given and abiotic control data are additionally shown. B: logarithmic concentrations and linear fits according to first-order kinetics are shown.

We initially used the kinetic data to test the validity of the sludge half-life normalization procedure as described in eq. 1. Given that the three experiments should only differ in biomass concentration but not composition, we expected $DT_{50,\text{sludge, norm},i,j}$ to be very similar across all three experiments for a given compound. To test the validity of this expectation, we analyzed the ratios of DT_{50} values between HB and NB, and between NB and DB experiments for all compounds exhibiting k_{deg} values $> 0.11 \text{ d}^{-1}$ in at least two of the three experiments ($n=22$). We found average ratios of half-lives between NB and HB experiments of $3.0(\pm 1.7)$ and $1.0(\pm 0.52)$, and between DB and NB experiments of $3.6(\pm 2.1)$ and $0.73(\pm 0.45)$ when using directly observed $DT_{50,\text{sludge}}$ values versus biomass- and bioavailability normalized $DT_{50,\text{sludge, norm}}$ values, respectively. These results demonstrate that normalization for biomass

and bioavailability yields consistent half-lives between experiments, which therefore can be readily averaged to obtain robust estimates of half-lives derived from activated sludge testing for regression analysis.

Regression analysis between sludge and soil half-lives

Three compounds did not show noticeable degradation in any experiment and were therefore excluded from model development (i.e., fluopicolide, florasulam, and pyroxsulam). Two more compounds, i.e., bixafen and clothianidine, were also excluded from model development because their half-lives in soil were either only given indicatively as >365 d (bixafen), or were not at all available (clothianidine). With the five compounds that were excluded from further evaluation due to a general lack of detectability in the activated sludge experiments, this left 42 compounds to be included in model development.

Table 2 gives an overview of the developed bi- and multivariate regression models and their quality metrics. For model 2, $f_{aq,soil}$ values could not be calculated for two more compounds due to a lack of data on the experimental soil moisture content (σ), leaving 40 compounds to be included in model development. For each model, results for a version with all compounds included, and one with strong outliers removed (defined as residuals $> 2 \cdot RSE$) are given in Table 2. Regressions for models 2 and 5, including confidence and prediction intervals, are plotted in Figure 2. In these plots, also the three recalcitrant compounds and bixafen, which were not considered in model development, have been included for validation purposes.

Figure 2: Final regressions for model 2 (A) and model 5 (B). Outlier compounds removed for final model development are given in green. Compounds that were persistent in either sludge or soil are included with their respective lower limit half-lives for model validation purposes and are given in red. Arrows indicate that symbols represent lower limits of either soil or sludge half-lives. The grey shaded area and the red dotted lines indicate the 95% confidence and prediction intervals on the regression, respectively.

1 The results presented in Table 2 for model 1 clearly show that there is only a very weak direct
2 relationship between observed half-lives in soil and activated sludge. Similarly bad results are
3 obtained if using the scores from the BIOWIN PSM model (model 6). In contrast, models 2-5,
4 which are based on sludge half-lives and additionally include correction terms for
5 bioavailability, all yield significant relationships with R^2 values above 0.36 with all
6 compounds included. The best performing model (model 5 with three outliers removed) has
7 an R^2 of 0.82, and an RSE of 0.41, which translates into an average error in half-life
8 prediction of a factor of 2.6. These results demonstrate that biotransformation in activated
9 sludge has the potential to predict biotransformation in soil. Notably, both the sludge half-life
10 and the bioavailability correction terms are highly significant in all of these models,
11 indicating that both pieces of information are needed to properly predict soil half-lives.

12 Several points are worth noting about models 2-5. First, the best model (i.e., model 5) is
13 obtained if bioavailability correction is not done with an explicit $f_{aq,soil}$ factor, but only
14 qualitatively with a $\log K_{oc}$ -based term. This introduces some additional degree of freedom
15 into the model, which is obviously useful for compensating partially wrong assumptions in
16 model 2 and/or uncertainty in data. On the practical side, model 5 is less data-greedy than
17 model 2 (i.e., it does not require any explicit information on the soils and experimental
18 conditions used to determine the soil half-lives) and thus makes for an easier application.
19 Again, from a practical perspective, it is interesting to note that models 3 and 4, which use $\log$
20 P and $\log D$ values (see Table S1 for values used), respectively, perform better than the
21 BIOWIN-based model 6, and that the $\log D$ -based model is superior to the $\log P$ -based model.
22 The latter was expected since the speciation-corrected $\log D$ values are generally accepted to
23 be better proxies of sorption to soil for acidic compounds than $\log P$ ²⁸. While $\log D$ values
24 are rather easy to determine both experimentally and computationally, $\log K_{oc}$ values are more
25 difficult to obtain. Therefore, particularly in the early phases of compound development,
26 where soil half-life prediction based on biotransformation data in activated sludge could be
27 highly valuable but K_{oc} values have typically not yet been determined, half-life predictions
28 based on $\log D$ values might be feasible and useful.

29 A number of compounds consistently behave as outliers in the best performing models 2 and
30 5, i.e., fenhexamid, sulfoxaflo, and, exclusively in model 5, also valifenalate. When also
31 considering the compounds found to be persistent in activated sludge (i.e., compounds
32 highlighted in red in Figure 2), florasulam and pyroxsulam additionally appear to be outliers,
33 whereas bixafen and fluopicolide are well predicted. For all outliers, half-life estimates for
34 soil are too high, i.e., extrapolation from activated sludge data underestimates the speed of
35 transformation of these compounds in soil. In terms of their initial biotransformation reactions
36 in soil, these compounds fall into diverse categories, i.e., O-demethylation for florasulam and

pyroxsulam, ester hydrolysis for valifenalate, and hydrolysis of an aliphatic nitrile group for sulfoxaflor, whereas dimerization is the main transformation reaction reported for fenhexamid (envipath.org²⁷). Given the diversity of initial biotransformation reactions affected, we consider it unlikely that a general lack of certain transformation pathways explains the deviating behavior of the outlier compounds. Rather, what is common to all of them, with the exception of valifenalate, is the fact that they are weak acids with pK_a values in the range of 4.5 to 8. pH values of soils used for OECD 307 studies are around 6.6 on average and typically range from 4.5 to 8¹⁶, whereas pH in our activated sludge experiments varied between 7.9 (HB) and 8.6 (DB) (Table S3). Therefore, we assume that the higher fraction of dissociated and hence negatively charged weak acids in our activated sludge experiments relative to most soil studies, and the lower permeability of bacterial cell membranes for the charged species, is the more likely explanation for the overestimation of half-lives for these compounds in the activated sludge assay. Similar, but inverse pH-dependences of biotransformation in activated sludge have been described for aliphatic amines²⁶.

Generalization of relationship between sludge and soil half-lives

The best performing models 2 and 5 were found to be robust in cross-validation (see Table 2). This is consistent with the fact that even models built on data from individual experiments still perform satisfactorily (i.e., $R^2=0.65$, $RSE=0.68$, $n=19$ for DB, $R^2=0.77$, $RSE=0.44$, $n=20$ for NB, and $R^2=0.61$, $RSE=0.43$, $n=29$ for HB, respectively, for model 2 built on experimental data of individual experiments with outliers removed). If model 2 is trained on a combination of data from DB and HB experiments only, the obtained model and its performance is very comparable (w/o outliers, $n=38$, $R^2=0.75$, $RSE=0.52$, Intercept=-0.48 ± 0.19 , $c_1=0.96 \pm 0.09$) to when data from all three experiments is used. These results suggest that it would potentially be sufficient to run only two (DB and HB) of the experiments.

This leaves the question whether this performance is unique to the activated sludge tested, or whether it would generalize to sludge sourced from other WWTPs. We have addressed this question using two more data sets: One from an experiment with activated sludge sourced from the nitrification basin of a pilot wastewater treatment plant operated at Eawag with an initial TSS of 6.73 g SS/L and spiked with the same selection of compounds as used here (details on experiment and data given in SI Section S5, here abbreviated as AB for “alternate biomass” experiment), and one from experiments with activated sludge adapted to solids retention times of 10 and 15 d in 12 L sequencing batch reactors and spiked with a set of compounds that only partially overlaps with the selection of compounds used here (details on experiment and data given in²⁵, here abbreviated as RB for “retention time biomass” experiment). Model 2 has been trained on these data sets and results are presented in the SI (Tables S6 and S7). Model performance for the RB experiment is very comparable to the

1 results reported in Table 2 (i.e., w/o outliers, $R^2=0.74$, $RSE=0.49$, $n=21$), whereas model
2 performance for the AB experiment is clearly worse (i.e., w/o outliers, $R^2=0.44$, $RSE=0.61$,
3 $n=21$). We currently cannot give any obvious reason for the worse performance of the AB
4 experiment, except for the fact that half-lives were determined at one sludge concentration
5 only and thus might be more uncertain than those derived from independent experiments with
6 three biomass concentrations. Also, the AB biomass stems from a pilot plant at Eawag whose
7 sludge might not have been fully equilibrated at the time of sampling due to a major sludge
8 loss event two months before. Nevertheless, the results indicate that the validity of the
9 approach should be tested on more activated sludge samples and possibly be accompanied
10 with a taxonomic and functional analysis of the sludge biomass to further define its
11 applicability domain.

1

2 **Table 2:** Bi- and multivariate regression models to predict $DT_{50,soil}$ (with intercepts c_1 and c_2 indicating values of model coefficients, and including metrics of
3 model performance, i.e., R^2 : coefficient of determination, RSE: root-squared error).

	Model		R^2	RSE	Intercept	c_1	c_2	Outlier compounds
Model 1	$\log DT_{50,soil} \sim c_1 * \log DT_{50,sludge,norm}$	All (n=42)	0.20	0.99	1.15*** ^b (±0.17)	0.52** (±0.16)	NA	Sulfoxaflo
		W/o outliers (n=41)	0.24	0.91	1.20*** (±0.16)	0.54** (±0.15)	NA	
Model 2	$\log DT_{50,soil} \sim c_1 * \log \left(\frac{DT_{50,sludge,norm}}{f_{aq,soil}} \right)$	All (n=40)	0.67	0.64	-0.58* (±0.24)	1.01*** (±0.12)	NA	Fenhexamid, Sulfoxaflo
		W/o outliers (n=38)	0.77 (0.77) ^a	0.49 (0.54) ^a	-0.41* (±0.19)	0.97*** (±0.09)	NA	
Model 3	$\log DT_{50,soil} \sim c_1 * \log DT_{50,sludge,norm} + c_2 * \log P$	All (n=42)	0.39	0.88	-0.02 (±0.37)	0.76*** (±0.16)	0.39 (±0.11)	Fenhexamid, Sulfoxaflo
		W/o outliers (n=40)	0.55	0.69	-0.19 (±0.31)	0.80*** (±0.13)	0.49 (±0.10)	

Model 4	$\log DT_{50, \text{soil}} \sim c_1 * \log DT_{50, \text{sludge, norm}} + c_2 * \log D$	All (n=42) ^c	0.36	0.90	0.52* (±0.25)	0.72*** (±0.16)	0.24*** (±0.07)	Fenhexamid, Sulfoxaflo, Topramezone
		W/o outliers (n=39)	0.65	0.62	0.19 (±0.20)	0.82*** (±0.11)	0.39*** (±0.06)	
Model 5	$\log DT_{50, \text{soil}} \sim c_1 * \log DT_{50, \text{sludge, norm}} + c_2 * \log K_{oc}$	All (n=42)	0.71	0.61	-1.77*** (±0.37)	0.92*** (±0.11)	1.09*** (±0.13)	Fenhexamid, Sulfoxaflo, Valifenalate
		W/o outliers (n=39)	0.82 (0.78) ^a	0.41 (0.45) ^a	-1.39*** (±0.26)	0.82*** (±0.08)	1.00*** (±0.09)	
Model 6	$\log DT_{50, \text{soil}} \sim c_1 * S_{\text{PSM}}$	All (n=42)	0.15	1.02	5.21*** (±1.46)	-1.17 (±0.44)	NA	Fenhexamid, Sulfoxaflo
		W/o outliers (n=41)	0.22	0.89	5.72*** (±1.28)	-1.28* (±0.39)	NA	

- 1 ^a Results from leave-one-out cross-validation. ^b S_{PSM} : Score from BIOWIN primary survey model (PSM) as implemented in the US EPA EPI Suite package ¹¹. ^b
- 2 Asterisks next to model coefficients represent the p-values for each coefficient as shown from the summary(lm) function in R, where *** is p=0, ** is p <
- 3 0.001, and * is p < 0.01 is. ^c For neutral compounds, log D was set equal to log P .

Environmental significance

In this study, we have demonstrated that it may be possible to read across from half-lives determined in highly efficient biotransformation experiments in activated sludge to soil half-lives, which are obtained from much more work- and resource-intensive regulatory studies. While previous attempts in doing so have either only worked on one class²⁰ or a very limited number of compounds²¹, we found predictive relationships across a wide range of pesticide classes and many different types of initial biotransformation reactions, which are also clearly superior to predictions based on the output of the publicly available BIOWIN QSBR model. Notably, these results were obtained with testing the chemicals in mixture. While the validity of such an approach has been questioned, others have shown negligible differences between results from single-compound and mixture testing in lake water biotransformation experiments if chemicals are spiked in low enough (i.e., ng/L to µg/L) concentrations³¹. It should also be noted that the predictive relationships were obtained despite the fact that, due to only analyzing dissipation in the aqueous phase and hence an effective lack of a complete mass balance in the activated sludge reactors, sludge half-lives might potentially be confounded by additional removal processes (e.g., uptake in living biomass).

At first sight, the observation that bioavailability-normalized half-lives linearly correlate between soil and activated sludge is unexpected since it must be assumed that the composition and diversity of the microbial communities in the two environments differ due to completely different environmental influencing factors³². Yet, the results suggest that, despite differences in taxonomic composition, the two environments are rather similar in terms of the biotransformation reactions they can catalyze. Conformingly, it has been reported previously that microbial communities from strongly differing environments, such as gut and plant rhizospheres, can display remarkable functional similarity in terms of metabolic capacities^{33, 34}.

In our case, the similarity in terms of biotransformation capacity between soil and activated sludge could be explained if observed compound loss was mostly due to fortuitous biotransformation by unspecific enzymes that can readily catalyze the observed biotransformation reactions, and if both microbial communities in question were competent with respect to these types of biotransformation reactions. Both, in regulatory simulation studies with soil³⁵ and the type of activated sludge batch reactor studies described here (e.g.,^{36, 37}), typically no lag phases are observed in biotransformation experiments run at environmentally relevant substance concentrations, supporting the notion of co-metabolic biotransformation by readily available enzymes. Also, in a recent study with activated sludge microbial communities, we demonstrated saturation of functional richness in terms of enzyme classes at solids retention times above 7 days, which are typical for nitrifying wastewater

1 treatment plants³⁸. Since soils typically harbor an even higher taxonomic richness than
2 activated sludge communities³⁹, functional saturation with respect to initial biotransformation
3 reactions could also plausibly be expected in soil. Even so, functional saturation does not
4 address the question of the relative abundances of enzyme classes, which, given the very
5 different nutrient situation and organic matter composition between soil and activated sludge,
6 must be expected to differ. To the best of our knowledge, there is currently a lack of data sets
7 (e.g., metatranscriptomes) that would allow for a direct comparison of the functional
8 composition of microbial communities from agricultural soils and activated sludge, and to
9 thus further scrutinize the above speculations.

10 Nevertheless, the evidence presented in this work suggests that the extrapolation of half-lives
11 from a simple experimental system to more environmentally relevant half-lives holds
12 potential for application, particularly (i) at the early phase of industrial research, more
13 specifically in the lead exploration stage, where access to labeled materials and time to run
14 full simulation studies is limited, and (ii) for the majority of industrial compounds registered
15 under REACH⁸ or legacy pharmaceuticals brought to market before 2006⁴⁰, for which any
16 compartment-specific biotransformation half-lives are missing completely. However, the
17 approach needs to be further investigated both in terms of its validity across a wider range of
18 activated sludge samples, but also in terms of its chemical applicability domain. Based on the
19 experimental design used in this study, we could reasonably assess compounds with half-lives
20 in soil between 0.1 and 300 days. In terms of persistence assessment, this range is clearly
21 relevant since regulatory persistence criteria lie in the range of 120 to 180 d. One limitation
22 of the current experimental design is that it only works for compounds that do not sorb
23 excessively to soil organic material (i.e., $K_{oc} < 30\,000$ L/kg) because above that value more
24 dilute experimental systems would be required to limit loss through sorption. This, however,
25 would slow down degradation too much to become discernible within the short time periods
26 for which these experiments are reasonably run. Another restriction of the set of chemicals
27 tested is that it does not include compounds with cationic speciation. Although they are rather
28 rare in agrochemical active ingredients, cation-specific mechanisms such as ion trapping in
29 activated sludge⁴¹ and sorption to soil constituents other than soil organic matter would need
30 to be considered in a model that includes such compounds.

31 From a microbiology perspective, before the proposed read-across approach can confidently
32 be applied in an industrial or regulatory context, it would need to be further underpinned with
33 molecular biology data characterizing the potential of soil and activated sludge microbial
34 communities to catalyze the wide range of observed initial contaminant biotransformation
35 reactions. Finally, it needs to be noted that neither the soils used for regulatory testing of
36 pesticides, nor, most likely, the activated sludge used in this study had seen prior exposure to

the compounds tested. It is currently not sufficiently understood to what extent prior exposure in any of the two environmental compartments, and hence the opportunity for adaptation, would impact the presented read-across approach.

Acknowledgments

We thank the operators and the staff of the WWTP ARA Neugut for providing activated sludge. We also thank Dr. Cresten Mansfeldt and Dr. Stefan Achermann for their experimental and analytical support, Mr. Tobias Plec for logistic support, and Dr. Myriem El Qacemi and Dr. Alain De Mesmaeker for the fruitful discussions in the early phase of the project, which helped to shape the experimental strategy. We acknowledge financial support from the European Research Council under the European Union's Seventh Framework Programme (ERC grant agreement 614768, PROduCTS).

Supporting Information

Information on test chemicals, including structures, stock solutions and internal standards used in quantification; information on main and alternate biotransformation experiments, including details on mass spectrometric analysis, concentration-time series and kinetic data evaluation; additional information on calculation of freely dissolved fractions; soil half-lives and metadata extracted from envipath.org as electronic SI (HalflivesConditions+ESI.csv).

Literature

1. EU, Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J. Europ. Union* **2006**, L 396, (49), 1-849.
2. EU, Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. *Official J. Europ. Union* **2009**, L 309, (52), 1-50.
3. EMEA *Guideline on the Environmental Risk Assessment of Medicinal Products for Human Use*; Committee for Medicinal Products for Human Use (CHMP); European Medicines Agency (EMA): London, 01 June 2006, 2006.
4. EMEA *Guideline on Environmental Impact Assessment for Veterinary Medicinal Products in Support of the VICH Guidelines GL6 and GL38*; Committee for Medicinal Products for Veterinary Use (CVMP); European Medicines Agency (EMA): London, 02 May 2007, 2007.

5. Voutchkova, A. M.; Osimitz, T. G.; Anastas, P. T., Toward a Comprehensive Molecular Design Framework for Reduced Hazard. *Chem. Rev.* **2010**, *110*, (10), 5845-5882.
6. Fenner, K.; Canonica, S.; Wackett, L. P.; Elsner, M., Evaluating Pesticide Degradation in the Environment: Blind Spots and Emerging Opportunities. *Science* **2013**, *341*, (6147), 752-758.
7. Rücker, C.; Kümmerer, K., Modeling and predicting aquatic aerobic biodegradation - a review from a user's perspective. *Green Chem.* **2012**, *14*, 875-887.
8. ECHA *Guidance on information requirements and chemical safety assessment, Chapter R.7b: Endpoint specific guidance*; European Chemicals Agency: Helsinki, Finland, June 2016, 2016.
9. Mamy, L.; Patureau, D.; Barriuso, E.; Bedos, C.; Bessac, F.; Louchart, X.; Martin-Laurent, F.; Miege, C.; Benoit, P., Prediction of the Fate of Organic Compounds in the Environment From Their Molecular Properties: A Review. *Crit. Rev. Environ. Sci. Tec.* **2015**, *45*, (12), 1277-1377.
10. Nolte, T. M.; Ragas, A. M. J., A review of quantitative structure–property relationships for the fate of ionizable organic chemicals in water matrices and identification of knowledge gaps. *Environ. Sci.-Proc. Imp.* **2017**, *19*, (3), 221-246.
11. U.S., E. Estimation Programs Interface Suite™ for Microsoft® Windows (EPI Suite), v 4.11. <https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface> (last accessed December 12, 2018).
12. Gouin, T.; Cousins, I.; Mackay, D., Comparison of two methods for obtaining degradation half-lives. *Chemosphere* **2004**, *56*, (6), 531-535.
13. Kuhne, R.; Ebert, R. U.; Schuurmann, G., Estimation of compartmental half-lives of organic compounds - Structural similarity versus EPI-Suite. *QSAR Comb. Sci.* **2007**, *26*, (4), 542-549.
14. Fenner, K.; Canonica, S.; Escher, B. I.; Gasser, L.; Spycher, S.; Tulp, H. C., Developing methods to predict chemical fate and effect endpoints for use within REACH. *Chimia* **2006**, *60*, (10), 683-690.
15. Nolte, T. M.; Pinto-Gil, K.; Hendriks, A. J.; Ragas, A. M. J.; Pastor, M., Quantitative structure–activity relationships for primary aerobic biodegradation of organic chemicals in pristine surface waters: starting points for predicting biodegradation under acclimatization. *Environ. Sci.-Proc. Imp.* **2018**, *20*, (1), 157-170.
16. Latino, D. A. R. S.; Wicker, J.; Gutlein, M.; Schmid, E.; Kramer, S.; Fenner, K., Eawag-Soil in enviPath: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data. *Environ. Sci.-Proc. Imp.* **2017**, *19*, (3), 449-464.
17. Corsi, C.; Lamberth, C., New Paradigms in Crop Protection Research: Registrability and Cost of Goods. In *Discovery and Synthesis of Crop Protection Products*, American Chemical Society: 2015; Vol. 1204, pp 25-37.
18. Menz, J.; Müller, J.; Olsson, O.; Kümmerer, K., Bioavailability of Antibiotics at Soil–Water Interfaces: A Comparison of Measured Activities and Equilibrium Partitioning Estimates. *Environ. Sci. Technol.* **2018**, *52*, (11), 6555-6564.
19. Golsteijn, L.; van Zelm, R.; Hendriks, A. J.; Huijbregts, M. A. J., Statistical uncertainty in hazardous terrestrial concentrations estimated with aquatic ecotoxicity data. *Chemosphere* **2013**, *93*, (2), 366-372.
20. Berger, B. M.; Muller, M.; Eing, A., Quantitative structure-transformation relationships of sulfonylurea herbicides. *Pest Manag. Sci.* **2002**, *58*, (7), 724-735.
21. Leoni, V.; Cremisini, C.; Giovinazzo, R.; Puccetti, G.; Vitali, M., Activated-sludge biodegradation test as a screening method to evaluate persistence of pesticides in soil. *Sci. Tot. Environ.* **1992**, *123*, 279-289.
22. Blok, J., Extrapolation of Biodegradability Test Data by Use of Growth Kinetic-Parameters. *Ecotox. Environ. Safe.* **1994**, *27*, (3), 306-315.
23. Honti, M.; Hahn, S.; Hennecke, D.; Junker, T.; Shrestha, P.; Fenner, K., Bridging across OECD 308 and 309 data in search of a robust biotransformation indicator. *Environ. Sci. Technol.* **2016**, *50*, 6865 - 6872.

24. Birch, H.; Hammershoj, R.; Mayer, P., Determining Biodegradation Kinetics of Hydrocarbons at Low Concentrations: Covering 5 and 9 Orders of Magnitude of K_{ow} and K_{aw}. *Environ. Sci. Technol.* **2018**, *52*, (4), 2143-2151.
25. Achermann, S.; Falås, P.; Joss, A.; Mansfeldt, C. B.; Men, Y.; Vogler, B.; Fenner, K., Trends in Micropollutant Biotransformation along a Solids Retention Time Gradient. *Environ. Sci. Technol.* **2018**, *52*, (20), 11601-11611.
26. Gulde, R.; Helbling, D. E.; Scheidegger, A.; Fenner, K., pH-Dependent Biotransformation of Ionizable Organic Micropollutants in Activated Sludge. *Environ. Sci. Technol.* **2014**, *48*, (23), 13760-13768.
27. Wicker, J.; Lorschach, T.; Gütlein, M.; Schmid, E.; Latino, D.; Kramer, S.; Fenner, K., enviPath – The environmental contaminant biotransformation pathway resource. *Nucleic Acids Res.* **2016**, *44*, (Database issue), D502-D508.
28. Schwarzenbach, R. P., *Environmental organic chemistry*. Third edition ed.; Hoboken, New Jersey, Wiley, 2017.
29. Schamfuss, S.; Neu, T. R.; Harms, H.; Wick, L. Y., A Whole Cell Bioreporter Approach to Assess Transport and Bioavailability of Organic Contaminants in Water Unsaturated Systems. *Jove-J. Vis. Exp.* **2014**, (94), e52334.
30. Lewis, K. A.; Tzilivakis, J.; Warner, D. J.; Green, A., An international database for pesticide risk assessments and management. *Hum. Ecol. Risk Assess.* **2016**, *22*, (4), 1050-1064.
31. Hammershøj, R.; Birch, H.; Redman, A. D.; Mayer, P., Mixture Effects on Biodegradation Kinetics of Hydrocarbons in Surface Water: Increasing Concentrations Inhibited Degradation whereas Multiple Substrates Did Not. *Environ. Sci. Technol.* **2019**, *53*, (6), 3087-3094.
32. Wu, L.; Ning, D.; Zhang, B.; Li, Y.; Zhang, P.; Shan, X.; Zhang, Q.; Brown, M.; Li, Z.; Van Nostrand, J. D.; Ling, F.; Xiao, N.; Zhang, Y.; Vierheilig, J.; Wells, G. F.; Yang, Y.; Deng, Y.; Tu, Q.; Wang, A.; Acevedo, D.; Agullo-Barcelo, M.; Alvarez, P. J. J.; Alvarez-Cohen, L.; Andersen, G. L.; de Araujo, J. C.; Boehnke, K.; Bond, P.; Bott, C. B.; Bovio, P.; Brewster, R. K.; Bux, F.; Cabezas, A.; Cabrol, L.; Chen, S.; Criddle, C. S.; Deng, Y.; Etchebehere, C.; Ford, A.; Frigon, D.; Gómez, J. S.; Griffin, J. S.; Gu, A. Z.; Habagil, M.; Hale, L.; Hardeman, S. D.; Harmon, M.; Horn, H.; Hu, Z.; Jauffur, S.; Johnson, D. R.; Keller, J.; Keucken, A.; Kumari, S.; Leal, C. D.; Lebrun, L. A.; Lee, J.; Lee, M.; Lee, Z. M. P.; Li, Y.; Li, Z.; Li, M.; Li, X.; Ling, F.; Liu, Y.; Luthy, R. G.; Mendonça-Hagler, L. C.; de Menezes, F. G. R.; Meyers, A. J.; Mohebbi, A.; Nielsen, P. H.; Ning, D.; Oehmen, A.; Palmer, A.; Parameswaran, P.; Park, J.; Patsch, D.; Reginatto, V.; de los Reyes, F. L.; Rittmann, B. E.; Robles, A. N.; Rossetti, S.; Shan, X.; Sidhu, J.; Sloan, W. T.; Smith, K.; de Sousa, O. V.; Stahl, D. A.; Stephens, K.; Tian, R.; Tiedje, J. M.; Tooker, N. B.; Tu, Q.; Van Nostrand, J. D.; De los Cobos Vasconcelos, D.; Vierheilig, J.; Wagner, M.; Wakelin, S.; Wang, A.; Wang, B.; Weaver, J. E.; Wells, G. F.; West, S.; Wilmes, P.; Woo, S.-G.; Wu, L.; Wu, J.-H.; Wu, L.; Xi, C.; Xiao, N.; Xu, M.; Yan, T.; Yang, Y.; Yang, M.; Young, M.; Yue, H.; Zhang, B.; Zhang, P.; Zhang, Q.; Zhang, Y.; Zhang, T.; Zhang, Q.; Zhang, W.; Zhang, Y.; Zhou, H.; Zhou, J.; Wen, X.; Curtis, T. P.; He, Q.; He, Z.; Brown, M.; Zhang, T.; He, Z.; Keller, J.; Nielsen, P. H.; Alvarez, P. J. J.; Criddle, C. S.; Wagner, M.; Tiedje, J. M.; He, Q.; Curtis, T. P.; Stahl, D. A.; Alvarez-Cohen, L.; Rittmann, B. E.; Wen, X.; Zhou, J.; Global Water Microbiome, C., Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nature Microbiol.* **2019**, *4*, (7), 1183-1195.
33. Mendes, R.; Raaijmakers, J. M., Cross-kingdom similarities in microbiome functions. *ISME J.* **2015**, *9*, 1905.
34. Ramírez-Puebla, S. T.; Servín-Garcidueñas, L. E.; Jiménez-Marín, B.; Bolaños, L. M.; Rosenblueth, M.; Martínez, J.; Rogel, M. A.; Ormeño-Orrillo, E.; Martínez-Romero, E., Gut and Root Microbiota Commonalities. *Appl. Environ. Microbiol.* **2013**, *79*, (1), 2.
35. EFSA *Draft Assessment Reports (DAR)*; <https://dar.efsa.europa.eu/dar-web/provision> (last accessed June 30, 2019).
36. Gulde, R.; Meier, U.; Schymanski, E. L.; Kohler, H.-P. E.; Helbling, D. E.; Derrer, S.; Rentsch, D.; Fenner, K., Systematic Exploration of Biotransformation Reactions of

- 1 Amine-Containing Micropollutants in Activated Sludge. *Environ. Sci. Technol.* **2016**, *50*, (6),
2 2908-2920.
- 3 37. Helbling, D. E.; Hollender, J.; Kohler, H.-P. E.; Fenner, K., A structure-based
4 interpretation of biotransformation pathways of amide-containing compounds in sludge-
5 seeded batch reactors. *Environ. Sci. Technol.* **2010**, *44*, 6628-6635.
- 6 38. Mansfeldt, C.; Achermann, S.; Men, Y.; Walser, J.-C.; Villez, K.; Joss, A.; Johnson,
7 D. R.; Fenner, K., Microbial residence time is a controlling parameter of the taxonomic
8 composition and functional profile of microbial communities. *ISME J.* **2019**, *13*, (6), 1589-
9 1601.
- 10 39. Curtis, T. P.; Sloan, W. T.; Scannell, J. W., Estimating prokaryotic diversity and its
11 limits. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, (16), 10494-10499.
- 12 40. Burns, E. E.; Carter, L. J.; Snape, J.; Thomas-Oates, J.; Boxall, A. B. A., Application
13 of prioritization approaches to optimize environmental monitoring and
14 testing of pharmaceuticals. *J. Toxicol. Environ. Health, Part B* **2018**, *21*, (3), 115-141.
- 15 41. Gulde, R.; Anliker, S.; Kohler, H.-P. E.; Fenner, K., Ion Trapping of Amines in
16 Protozoa: A Novel Removal Mechanism for Micropollutants in Activated Sludge. *Environ.*
17 *Sci. Technol.* **2018**, *52*, (1), 52-60.
- 18