EPA Public Access

Author manuscript

Environ Sci Technol. Author manuscript; available in PMC 2019 October 03.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Sci Technol. 2017 May 02; 51(9): 5008–5016. doi:10.1021/acs.est.6b05412.

Prediction of hydrolysis products of organic chemicals under environmental pH conditions

Caroline Tebes-Stevens^{†,*}, Jay M. Patel[‡], W. Jack Jones[†], Eric J. Weber[†]

[†]U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, Georgia 30605, United States

[‡]ORISE Fellow, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, Georgia 30605, United States

Abstract

Cheminformatics-based software tools can predict the molecular structure of transformation products using a library of transformation reaction schemes. This paper presents the development of such a library for abiotic hydrolysis of organic chemicals under environmentally relevant conditions. The hydrolysis reaction schemes in the library encode the process science gathered from peer-reviewed literature and regulatory reports. Each scheme has been ranked on a scale of one to six based on the median half-life in a data set compiled from literature-reported hydrolysis rates. These ranks are used to predict the most likely transformation route when more than one structural fragment susceptible to hydrolysis is present in a molecule of interest. Separate rank assignments are established for pH 5, 7 and 9 to represent standard conditions in hydrolysis studies required for registration of pesticides in Organisation for Economic Co-Operation and Development (OECD) member countries. The library is applied to predict the likely hydrolytic transformation products for two lists of chemicals, one representative of chemicals used in commerce and the other specific to pesticides, to evaluate which hydrolysis reaction pathways are most likely to be relevant for organic chemicals found in the natural environment.

Graphical Abstract

*Corresponding Author: Phone: (706) 355-8218; stevens.caroline@epa.gov.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Description of ChemAxon Metabolizer algorithm; table of 58 literature-reported activation energies for hydrolysis reaction schemes; full list of schemes in the abiotic hydrolysis reaction library, with example transformations; table of summary statistics for literature-reported hydrolysis rate data measured at neutral pH; table of 187 literature-reported hydrolysis half-lives adjusted to pH 5, 7 and 9 and temperature of 25°C; box-and-whisker plots showing distribution of 25°C-corrected half-lives at pH 5 and at pH 9; table of molecules in the external validation set (PDF)

INTRODUCTION

Regulatory decisions on the usage and disposal of agricultural and industrial chemicals are based on minimizing exposure of humans and vulnerable ecological species to hazardous chemicals. One important exposure pathway is through ingestion or absorption of dissolved organic contaminants, which may be introduced into waterways through agricultural and urban runoff, wastewater discharges and accidental spills. To assess the exposure risk associated with dissolved chemicals accurately, tools are needed to predict the distribution and fate of organic contaminants in water, including the potential formation of products. Hydrolysis, defined as the cleavage of chemical bonds by water, has long been recognized as a potentially important transformation process for dissolved organic contaminants in the environment.

Several comprehensive reviews have summarized and compiled published literature on hydrolysis of synthetic organic chemicals intentionally or accidentally released into the environment^{1,2,3,4}. These reviews include compilations of observed transformation rate constants for selected chemicals to characterize the overall likelihood of hydrolysis under neutral, acidic and basic conditions for various chemical classes of environmental concern (e.g., organophosphorus esters, carbamates, etc.).

Much of the data on hydrolytic transformations in the environment comes from studies on the environmental fate of agricultural chemicals. For pesticide registration, studies characterizing the hydrolysis rates and degradation products of the pesticide in water are required in most Organisation for Economic Co-Operation and Development (OECD) member countries^{5,6}. These hydrolysis studies are conducted in sterile water in the absence of light to distinguish hydrolytic transformations from other environmental transformation processes, including biodegradation and phototransformation. Hydrolysis reactions can occur even in dark environments with limited microbial activity, such as may be found in the deep subsurface; therefore, hydrolysis rates provide a lower limit on the transformation rates that may be expected to occur when a chemical is introduced into the environment. Predictive models used to assess exposure potential to pesticides, such as PRZM⁷ and EXAMS⁸, generally include hydrolysis as a first-order degradation process. These models

are also capable of simulating the fate of a limited number of transformation products; however, the fate of transformation products is generally not modeled unless a particular product is of toxicological concern.

Software tools that predict likely transformation products in environmental and biological systems can support chemical exposure and risk assessment by identifying potential products that should be considered in the assessment. A number of software tools have been developed in recent years to predict transformation products of organic chemicals resulting from mammalian metabolism (e.g., METEOR⁹ and TIMES¹⁰) and microbial degradation (e.g., UM/EAWAG-PPS¹¹ and enviPath¹²). There are also a few examples of software tools to predict abiotic transformation processes, such as the META expert system, which has been expanded to predict phototransformation products¹³. Additionally, the Zeneth¹⁴ software package predicts degradation of pharmaceuticals under the extreme conditions used in stability tests. Hydrolysis reactions may be included as incidental pathways in these software tools for predicting transformation products; however, to our knowledge, there is no software tool available for predicting hydrolytic transformation products in water, without the catalytic effects of enzymes and/or light.

At the core of these tools for predicting transformation products is a library of reaction schemes (also known as rules), which define how a particular structural fragment within a molecule will be modified by the specified transformation process. For example, the reaction scheme for hydrolysis of an anhydride structural fragment will show cleavage of the C – O bond of the molecule and formation of two carboxylic acids. The reaction schemes are encoded and implemented using cheminformatics software tools. For a chemical of interest, these tools search the molecule for each structural fragment in the reaction library and then modify the fragment according to the associated reaction scheme to predict the molecular structure of potential transformation products.

This paper presents the development of an abiotic hydrolysis reaction library which will be implemented in the Chemical Transformation Simulator¹⁵, a web-based software tool under development in the United States Environmental Protection Agency (U.S. EPA) Office of Research and Development. For each chemical class that is known to undergo abiotic hydrolysis under environmentally relevant conditions, one or more reaction schemes have been encoded to define how the structural fragment that is susceptible to hydrolysis may be modified by reaction with water. These schemes have been ranked using reported hydrolysis rates to enable qualitative prediction of the most likely transformation route when more than one structural fragment susceptible to hydrolysis is present in the molecule of interest.

METHODS

Development of the hydrolysis reaction library began with an initial set of reaction schemes that were obtained from published reviews on hydrolysis reactions in the environment^{1,2,3,4,16}. For each scheme, scientific journals and government regulatory documents were searched for examples of specific molecules observed to be transformed according to the scheme. The goal was to find at least 10 example chemicals with a diversity of molecular structure to assess whether the reaction scheme was representative of the

transformation process. If predictions from the execution of the reaction scheme did not match the products reported in the literature or regulatory reports, the schemes within the library were refined and/or additional schemes were added as needed to capture the observed formation of products as a result of the transformation process.

The hydrolysis reaction library was built using ChemAxon's Reactor application and tested and implemented using ChemAxon's Metabolizer software (ChemAxon LLC, Cambridge, MA, U.S.A.). The entry for each reaction scheme includes the scheme itself, example transformations for molecules that have been observed to be transformed according to the scheme, citations for both the reaction scheme and examples of the transformation pathway, and the rank assigned to the scheme. The rank is essentially a relative reaction rate, defined on a scale of one to six, with six being assigned to the fastest reaction schemes. Some of the schemes also include "reactivity" and "selectivity" reaction rules defined using ChemAxon's Chemical Terms scripting language¹⁷. The purpose of a reactivity rule is to constrain the reaction to occur only if the structure of the molecule meets specified conditions (e.g., a particular atom within the reaction center is not part of a ring). The purpose of a selectivity rule is to indicate a preferred reaction site when more than one structural fragment within the molecule matches the reaction center for the scheme.

ChemAxon's Metabolizer application uses the rank of each scheme to calculate an approximate percentage production of each potential transformation product. First, potential transformation pathways are identified on the basis of the presence of molecular structural fragments that are susceptible to a transformation scheme within the library and the application of any reactivity or selectivity rules associated with the scheme. An explanation of the Metabolizer algorithm used to estimate the percent production of each product is provided in the Supporting Information. Briefly, for each potential transformation pathway, a unitless "formation" is calculated as an exponential function of the rank assigned to the reaction scheme. The difference in the calculated "formation" values at two different ranks is proportional to the difference in the midpoint of the range of first-order transformation rate constants associated with the rank assignments in the hydrolysis reaction library. The percentage production of the product formed by a specific pathway is approximated as the formation associated with that pathway divided by the sum of formation values for all potential transformation pathways for the molecule. In "fast enumeration mode", as well as the default settings for batch command-line execution, the Metabolizer application options are set so that products are ignored (i.e., excluded from the output file) when their percent production is less than 10% of the highest percent production for all possible transformation pathways of the parent compound. Metabolizer also allows the user to select an "exhaustive" mode with output of all possible products, regardless of formation likelihood. Metabolizer predicts successive generations of products by treating products from one generation as the parent compounds for the next level of products.

To support the formulation and ranking of hydrolysis reaction schemes, a database of reported hydrolysis products and measured rate constants or half-lives was compiled from a survey of peer-reviewed scientific literature and reports by government regulatory agencies. Reported rates were only included in the compilation if the measurements were made at roughly constant pH and temperature, in the dark, and within the environmentally relevant

pH range of 4 to 9. Publications that report rates of hydrolysis in the presence of catalysts or high concentrations of cosolvents (> ~10%) were not included. Finally, if a publication reported hydrolysis of a molecule that is subject to more than one hydrolytic reaction pathway, the hydrolysis rates were only included in the compilation if one of the reaction pathways was clearly the dominant hydrolytic transformation. Specifically, the rate constant was included in the data set for a particular reaction pathway if the reported product distribution indicated that more than 90% of the parent was transformed to the predicted products for that reaction scheme.

If a rate constant recorded in the database was measured at a non-standard temperature, the rate constant was adjusted to a temperature of 25°C using the Arrhenius equation, which requires an activation energy specific to the molecule and transformation pathway (*i.e.*, reaction scheme) of interest. If a published value was available for the activation energy of the molecule undergoing transformation by the pathway of interest, that value was used in the Arrhenius equation. If the activation energy was not available for the molecule, the temperature correction was performed with an average reported activation energy for other molecules undergoing transformation according to the same reaction scheme. Table SI-1 in the Supporting Information provides a compilation of 58 literature-reported activation energies for hydrolysis reactions.

Hydrolysis may occur due to nucleophilic attack by H₂O (*i.e.*, neutral hydrolysis), or the process may be catalyzed by H⁺ or OH⁻. To capture differences in these three mechanisms, rate constants measured under neutral, acidic and basic conditions were recorded as separate entries in the database. Current regulatory guidelines for pesticides^{6,18} require registrants to measure hydrolysis rates at pH values of 4, 7, and 9; however, earlier guidelines specified a measurement pH of 5 instead of 4 as representing hydrolysis under acidic conditions. Much of the historical rate data on acid-catalyzed hydrolysis was measured at pH 5, and this pH is more likely to be encountered in the environment; therefore, a pH of 5 was selected as the standard acidic pH for the database.

Hydrolysis rates are commonly reported as effective first-order rate constants or half-lives at a specific pH; however, acid- and base-catalyzed hydrolytic transformations are actually second-order reactions. The overall rate of hydrolysis (R_{Hyd}) at any given pH is given by the following equation:

$$R_{Hyd} = k_{A2}[H^+]C + k_N C + k_{B2}[OH^-]C \quad (1)$$

where k_{A2} is the second-order rate constant for acid-catalyzed hydrolysis, k_N is the first-order rate constant for neutral hydrolysis, k_{B2} is the second-order rate constant for base-catalyzed hydrolysis, and C is the concentration of the chemical undergoing hydrolysis. Ideally, three values of R_{Hyd} would be measured at separate pH values, and three equations in the form of (1) could be combined and rearranged to solve for k_{A2} , k_N , and k_{B2} . Unfortunately, there are numerous hydrolysis studies for which measured half-lives are available at only two pH values, with the molecule reported as being “stable” or having a half-life greater than a certain number of days at the third pH value. Furthermore, when

hydrolysis rates are available at three pH values, experimental error in one or more of the measured rates can result in an unrealistic solution with a negative value for one of the rate constants. Because it was not possible to reliably estimate k_{A2} , k_N , and k_{B2} for each molecule in the data set, rankings were assigned on the basis of overall hydrolysis rate constants at pH 5, 7 and 9.

The reported rate constants or half-lives at pH 5, 7, and 9 were first compared to evaluate whether the molecule was transformed through acid-catalyzed, neutral and/or base-catalyzed hydrolysis. Neutral hydrolysis is likely the dominant mechanism if rate constants or half-lives are the same order of magnitude across all three pH values. A shorter half-life at pH 5 relative to the half-life at pH 7 (or equivalently, larger rate constant at pH 5 relative to pH 7) provides evidence that the hydrolysis reaction is catalyzed by H^+ . In this case, to calculate an effective half-life at pH 5 ($t_{1/2, pH 5}$) from literature-reported rate constants at pH 5, the assumption was made that neutral and base-catalyzed hydrolysis were both negligible and that acid-catalyzed hydrolysis was the dominant process at pH 5 or lower:

$$t_{1/2, pH 5} = \frac{\ln(2)}{k_{A1, pH 5}} \quad (2)$$

where $k_{A1, pH 5}$ is the effective first-order rate constant for acid-catalyzed hydrolysis at pH 5. While most publications report effective first-order rate constants, some do report second-order rate constants; in those cases, the effective first-order rate constant at pH 5 was calculated as the product of k_{A2} and the concentration of $[H^+]$ (10^{-5} mol/L). If the hydrolysis rate under acidic conditions was measured at a pH value other than 5, the half-life was adjusted to pH 5 as follows:

$$t_{1/2, pH 5} = \frac{\ln(2)}{k_{A1, pH_{meas}} 10^{pH_{meas} - 5}} \quad (3)$$

Similarly, a shorter half-life at pH 9 relative to the half-life at pH 7 (or equivalently, larger rate constant at pH 9 relative to pH 7) provides evidence that the hydrolysis reaction is catalyzed by OH^- . In this case, to calculate an effective half-life at pH 9 ($t_{1/2, pH 9}$) from literature-reported rate constants at pH 9, the assumption was made that base-catalyzed hydrolysis was the dominant process near pH 9:

$$t_{1/2, pH 9} = \frac{\ln(2)}{k_{B1, pH 9}} \quad (4)$$

where $k_{B1, pH 9}$ is the effective first-order rate constant for base-catalyzed hydrolysis at pH 9. For those publications that report second-order rate constants for base-catalyzed hydrolysis, the effective first-order rate constant at pH 9 was calculated as the product of k_{B2} and the concentration of $[OH^-]$, which is equal to 10^{9-pK_w} mol/L, where pK_w is the negative

logarithm of the dissociation constant of water. If the hydrolysis rate under basic conditions was measured at a pH value other than 9, the half-life was adjusted to pH 9 as follows:

$$t_{1/2, pH\ 9} = \frac{\ln(2)}{k_{B1, pH_{meas}} 10^{9 - pH_{meas}}} \quad (5)$$

For each reaction scheme in the library, the mean, median, standard deviation, and interquartile range (defined as the difference between the 25% percentile and 75% percentile) were calculated for the transformation half-lives at pH values of 5, 7 and 9 over all molecules included in the database for that scheme. Large differences between the mean and the median suggest that there may be outliers for a given scheme. Half-lives that differ from the median by more than 2.2 times the interquartile range were flagged as potential outliers¹⁹ and excluded from the data set when assigning a rank to the reaction scheme.

Each reaction scheme was assigned a pH-specific rank based on the median half-life for hydrolytic transformation according to the scheme at pH values of 5, 7 and 9. Table 1 defines the six ranking levels, which span residence times of environmental relevance. These ranks are similar to those used in a survey conducted by Boethling *et al.*²⁰ which asked experts to rate the biodegradability of 200 organic chemicals on a semi-quantitative scale with five levels: hours, days, weeks, months and longer. Reaction schemes with a rank of “one” would likely only be relevant for the fate of organic chemicals in groundwater, which can have hydrologic residence times of hundreds or even thousands of years^{21,22}. Reaction schemes with ranks of two or three may be significant for lakes, with typical residence times of less than one year²³, and estuaries, with residence times from a couple of days to more than 100 days^{24,25}; however, these schemes will be of lower relevance for rivers, with residence times up to two weeks, and atmospheric water, with residence times on the order of eight to ten days^{21,26}.

RESULTS AND DISCUSSION

Encoding of the Process Science Underlying Abiotic Hydrolysis

The hydrolysis reaction library includes 24 reaction schemes representing hydrolytic transformations that have been widely observed under environmentally relevant conditions. Figure 1 shows selected schemes from the library; the full list of schemes is provided in Supporting Information. As is illustrated in Figure 1, the schemes are written in generic terms to provide complete coverage of organic chemicals containing the functional group that is susceptible to hydrolysis. The schemes are written to capture the formation of observable products; therefore, they do not capture the formation and disappearance of unstable intermediates. The library also includes a scheme for the spontaneous dehydration of a geminal diol to form a carbonyl. This dehydration scheme is not a hydrolysis reaction; however, it must be included in the library to arrive at the correct product distribution for some hydrolysis studies.

More than half of the schemes within the library include reactivity or selectivity reaction rules defined using ChemAxon's Chemical Terms scripting language. For example, reactivity reaction rules are used to distinguish between cyclic and acyclic structures containing the same structural fragment. This distinction is needed to allow the assignment of different ranks to the schemes associated with cyclic and acyclic structures; for example, the schemes for lactones and lactams are up to three ranks higher than the schemes for their acyclic counterparts, carboxylic acid esters and amides, respectively. Selectivity rules are used to avoid duplication of transformation products when the reaction center is symmetrical, as is the case for anhydrides, carbonates, and ureas. Additionally, the four schemes for hydrolysis of halogenated aliphatics include a selectivity reaction rule to indicate that, for molecules with multiple different halogen substituents, the order of removal of the leaving halogen is inverse to its atomic number, e.g., bromine will be removed before chlorine.

Ranking of Schemes in the Hydrolysis Reaction Library

The ranks of all schemes in the hydrolysis reaction library at pH values of 5, 7 and 9, listed in order of fastest to slowest for hydrolysis at neutral pH, are provided in Table 2. The fastest reaction schemes, with a median half-life less than 0.1 days, are those associated with anhydride hydrolysis and acid halide hydrolysis. Under neutral conditions, the hydrolysis of lactones, nitriles, epoxides and imides are also observed to be fast, with median half-lives less than one day. On the other hand, amide hydrolysis and nucleophilic substitution reactions for di- and tri-substituted halogenated aliphatics are observed to be quite slow, with median reported half-lives greater than one year.

The pH-specific ranks in Table 2 were assigned based on the median half-life for all molecules included in the database of literature-reported rate data for each scheme in the library. The full data set of 187 rate constants adjusted to 25°C and standard pH values of 5, 7, and 9 is provided in Table SI-3 in the Supporting Information. For some reaction schemes, there was limited rate constant data available under acidic and basic pH values; in these cases, ranks (marked with a superscript in the table) were assigned on the basis of the rate data at neutral pH.

Figure 2 shows the distribution of observed temperature-corrected rate constants measured at a pH of 7. The half-life range associated with each of the six ranks is designated with colored shading; for example, median half-lives greater than one year are assigned a rank of "one" and fall within the purple-shaded range of the figure. Similar plots are provided in the Supporting Information for observed rate constants at pH 5 and pH 9. For some schemes, notably acid halide hydrolysis, elimination of halogenated aliphatics, and epoxide hydrolysis, there is a large spread in the range of observed half-lives. Additionally, for a number of other schemes, the first quartile and/or third quartile of observed half-lives lie within adjacent ranks. The variability in reported half-life values highlights the uncertainty and qualitative nature of the predicted formation rates of hydrolysis products based on the ranking scheme in this library.

Ranks are used to estimate the predicted product distribution when a chemical contains more than one functional group that is susceptible to hydrolysis; therefore, deviations of actual

hydrolysis rate constants from the defined range of rate constants for the rank assigned to a particular scheme may affect the predicted product distribution. For example, if a molecule is transformed by a particular scheme at a substantially faster rate than the average rate constant in our data set for that scheme, then the transformation rate constant may fall within the range associated with a higher rank than the one assigned to the scheme. Whether or not this incorrect rank assignment affects the product distribution depends upon the actual rate constants of the other potential transformation products of the molecule. If the rate constants for the other potential transformation schemes fall within the expected range for the rank assignments of those schemes, then the library will underpredict the percentage production for the faster-than-expected transformation pathway. On the other hand, if the rates are elevated relative to the rank assignments for all schemes that may transform the molecule, then the predicted product distribution may be minimally affected by the incorrect rank assignments.

When Metabolizer is executed using the default settings, rank assignments may affect the number of products that are provided to the user. Specifically, with default settings, a product will be removed from the Metabolizer output when its percent production is less than one-tenth of the highest percent production of all predicted products from the same parent molecule. In contrast, the exhaustive mode of Metabolizer generates all potential products, even those that are predicted to form in low amounts relative to the other products. Given the uncertainty in rank assignments, this more conservative approach may be preferred.

Despite the spread in observed rate constants, the relative ranking of hydrolysis schemes within the library is generally consistent with organic chemistry principles. For example, acid halide, acid anhydride, carboxylic acid ester, lactone, amide, and lactam hydrolysis transformations all belong to a general reaction type known as Nucleophilic acyl substitution reactions (Scheme 1): where $:Q^-$ and $:Nu^-$ are the leaving group and nucleophile respectively. Their relative reactivities vary inversely with the basicity of the leaving group. One way to compare the relative strength of the base is to consider the pK_a of its conjugate acid; the lower the pK_a , the stronger the conjugate acid, the weaker the corresponding base, and the more readily the leaving group will be cleaved from the molecule. Thus, we expect the highest rates with acid halide leaving groups (Cl^- , Br^- and I^-) and the anhydride leaving group, ($RCOO^-$); intermediate rates with the carboxylic acid ester leaving group (RO^-); and the lowest rates with the amine leaving group (NH_2^-). We do, in fact, see this order of reactivity in Table 2 and the box plot in Figure 2. Another trend supported by organic chemistry principles is that the lactams (cyclic amides) and lactones (cyclic esters) undergo hydrolysis at higher rates than their acyclic counterparts. This relative ranking is intuitive considering that, in addition to leaving group strength, ring strain makes lactones and lactams more reactive. As a related observation, ethers generally do not undergo base-catalyzed hydrolysis, but epoxides are an exception, mainly as a result of ring strain.

It should be noted that the ranks at pH 5, 7 and 9 are not equivalent to ranks for acid-catalyzed, neutral and base-catalyzed hydrolysis, respectively; however, some mechanistic insight can be gleaned from the relative rankings at different pH values. Schemes that have a higher rank at pH 9 relative to pH 5 are likely to be promoted by base catalysis, and schemes that have a higher rank at pH 5 are likely to be promoted by acid catalysis. For example, the carboxylic acid ester hydrolysis scheme is assigned a higher rank at pH 9 than at pH 5 and 7. This leads us to correctly conclude that base-catalyzed hydrolysis, or saponification, is faster than acid-catalyzed hydrolysis. Indeed, such an assertion is widely stated in organic chemistry textbooks²⁹, because saponification, an irreversible reaction, tends to go faster than the reversible acid-catalyzed hydrolysis reaction.

Testing and Validation of Hydrolysis Reaction Library

The validity of the reaction schemes included in the CTS hydrolysis reaction library was assessed against 114 example transformations that were collected from the literature. Testing of a reaction scheme requires a comparison of predicted and reported hydrolysis products for molecules that undergo transformation according to the scheme. This comparison must be performed manually, because Metabolizer output provides the structure of predicted transformation products in the form of a SMILES string, while literature sources generally report transformation products by name or with a picture of the structure of the product. The reaction schemes correctly predicted all 114 products in the data set of example transformations included in the Supporting Information. The majority of these examples can be considered as an external validation set for the schemes; however, in some cases, the schemes were adjusted after an example transformation was found that was not covered by the scheme. (For example, the sulfonylurea scheme (Scheme S20) was modified to allow the nitrogen atom not bound to the sulfone group to be part of a heteroaromatic ring after the Thiencarbazone-Methyl example was found.)

To test the ranks assigned to the schemes in the hydrolysis reaction library, an external validation set was compiled from the Pesticide Manual³⁰ published by the British Crop Protection Council (BCPC). Hydrolysis half-lives were reported for 121 of the 862 pesticides included in the manual; however, the manual did not always identify the products of hydrolytic transformations. A total of 42 of these chemicals were removed from the validation set, because hydrolysis rate constants for these chemicals had been included in the data set used for rank assignments. Another 15 of the chemicals were removed from the list because they were not predicted to hydrolyze according to any of the schemes in the hydrolysis reaction library. An examination of the hydrolysis data for these 15 chemicals revealed that the majority of the hydrolysis products listed in the manual were due to heterocyclic ring opening reactions. Appropriate reaction schemes will be added to a future version of the hydrolysis reaction library to capture these transformations.

Of the remaining 64 pesticides in the validation set, 39 contained only one functional group that was susceptible to hydrolysis, and the remaining 25 contained more than one functional group susceptible to hydrolysis. Table SI-4 in the Supporting Information provides the full list of pesticides in the external validation set. Among the 39 molecules with only one hydrolytic transformation pathway, the reported half-lives fell within the range of half-lives

associated with the rank assigned to the corresponding scheme for 41% of the molecules and within the range of half-lives associated with the rank immediately above or below the assigned rank for another 26% of the molecules. For 28% of the molecules, the hydrolysis half-life was less than the lower limit of the rank directly above the rank assigned to the scheme, and for 5%, the hydrolysis half-life was greater than the upper limit of the rank directly below the rank assigned to the scheme.

Comparison of the assigned ranks to reported half-lives was more challenging for the 25 pesticides in the BCPC Pesticide Manual³⁰ that could be transformed by more than one hydrolysis scheme. If the half-life was measured based on the disappearance of the parent chemical, the reported half-life should reflect the fastest of the possible hydrolysis pathways; therefore, the reported half-lives were compared to the half-life range for the scheme with the maximum rank. This evaluation revealed that 20% of the half-lives fell within the range of half-lives associated with the maximum rank and another 32% fell within the range of the rank immediately above or below the maximum rank. For the remaining 48% of the molecules, the half-life deviated more substantially from the range that would be predicted based on the rank assignment of the scheme. Unexpectedly, nine of these molecules had half-lives that were much longer than would be expected based on the rank assignment. Many of these slower-than-expected transformations were for molecules that contained cyano groups, which can be transformed into amide groups according to the relatively fast nitrile hydrolysis scheme. The BCPC Pesticide Manual³⁰ does not provide any details about the experimental conditions used to measure the rate constants; however, hydrolysis studies are often conducted with labeled chemicals to detect the formation of a particular product. It is possible that the hydrolysis half-lives reported for these molecules reflect the kinetics of a slower hydrolytic reaction that results in cleavage of the molecule (e.g. carboxylic acid ester hydrolysis), which may be of greater interest than the hydrolysis of the nitrile group on the molecule.

The external validation of the rank assignments revealed that reported hydrolysis rate constants fell outside of the expected range for the relevant hydrolysis scheme for more than half of the molecules in the validation set. Significant deviations between observed hydrolysis rate constants and the expected values based on the rank assignments should not be surprising given the spread in observed rate constants shown in Figure 2. Work is currently underway to examine whether quantitative structure-activity relationship (QSAR) models can be used to better characterize the effects of molecular structure on the estimated percent production of each product; however, this approach will also significantly increase the computational burden over the current approach that uses a constant rank for all molecules transformed according to a particular hydrolysis scheme. It should be noted that the ranks are only used to ascertain the relatively likelihood of product formation when more than one transformation pathway is possible. If only one transformation pathway is possible, then the production will be 100% for that pathway, regardless of the rank assigned to the scheme. For a molecule that may be transformed by multiple hydrolytic pathways, if the observed transformation rates are equally higher (or lower) than would be expected from the rank assignments of each scheme, the deviations between observed and expected rates may have a minimal effect on the product distributions. On the other hand, if the observed hydrolysis rate of one of the pathways deviates from the expected rate, while the other rates

do not, the predicted distribution of products may be significantly different than the observed distribution.

Predicted Hydrolysis Pathways for Environmental Chemicals

To examine which hydrolysis reaction pathways are most likely to be relevant for organic chemicals found in the environment, hydrolysis transformation products were predicted through batch command-line execution of ChemAxon's Metabolizer application for two lists of chemicals. The first list consisted of 636 organic chemicals regulated under EPA's Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA); this list was generated by removing inorganics, organometallics, and duplicates (after deconstructing salts into component molecules) from the list of pesticides registered under FIFRA as compiled by Scorecard³¹. The second list, containing more than 32,000 molecules, was the prediction set from the CERRAP modeling project³². The CERRAP list was assembled to include a large fraction of the synthetic organic chemicals to which humans may be exposed in the home or workplace.

Table 3 shows the number of chemicals that are susceptible to transformation by the CTS library of hydrolysis schemes for the FIFRA and CERRAP chemical lists. Overall, 46% of the chemicals on the FIFRA list and 40% of the chemicals on the CERRAP list contained one or more structural fragments subject to hydrolysis. For both lists, the most common hydrolyzable structural fragment is the carboxylic acid ester group, which was present in slightly more than 13% of the molecules on each list. For the FIFRA list, the second and third most commonly predicted hydrolysis reactions were associated with organophosphorus esters (6%) and amides (5%). For the CERRAP list, the second and third most commonly predicted hydrolysis reactions were associated with amides (8%) and lactams (3%).

All 25 schemes in the proposed hydrolysis reaction library were predicted to occur for multiple chemicals on the CERRAP list. In contrast, there were five schemes (*i.e.*, dehydration of geminal diols and hydrolysis of carbonates, cyclic carbonates, and cyclic anhydrides) that were not predicted to occur for any molecules on the FIFRA list, and another two schemes (*i.e.*, hydrolysis of anhydrides and cyclic ureas) were only predicted to occur for a single molecule on this list. This result suggests that there is greater diversity in molecular structure on the CERRAP list in comparison to that on the FIFRA list; however, for all of the schemes with no or only one representative molecule on the FIFRA list, less than 1% of the molecules on the CERRAP list were hydrolyzed by the scheme. In fact, the percentage of molecules on each list that are susceptible to a particular hydrolytic transformation pathway was generally similar for most schemes in the library. Exceptions included the hydrolysis schemes for amides, lactams and cyclic ureas, which were found to affect more chemicals on the CERRAP list. By comparison, molecules on the FIFRA list were more likely to be affected by the hydrolysis schemes associated with halogenated aliphatics, organophosphorus esters, carbamates, thiocarbamates, acyclic ureas, sulfonyleureas, and cleavage of the nitrogen-sulfur bond.

Much of the hydrolysis data that was compiled for the development of the abiotic hydrolysis reaction library came from studies on the environmental fate of pesticides. Under FIFRA, pesticide registrants are required to conduct laboratory studies on hydrolysis, and the data

from many of these studies is readily available, in either journal publications or regulatory reports. On the other hand, hydrolysis data for pharmaceuticals and industrial chemicals is often more difficult to find in the literature. A comparison of the functional group representations in the FIFRA and CERRAP lists suggests that there may be unpublished literature in company archives that could be mined to test and refine the ranking of some schemes in the library. For example, the CERRAP list contains numerous molecules that are predicted to be susceptible to hydrolysis according to the schemes for ureas, carbonates, lactams and imides, all of which were characterized by relatively few rate constants in our data set.

The reaction library for abiotic hydrolysis encodes the process science underlying hydrolysis pathways for 25 structural fragments contained in chemicals that are susceptible to hydrolysis under environmental conditions. The reaction schemes within the library have been ranked using reported hydrolysis rates to enable a qualitative prediction of the most likely route of transformation when more than one structural fragment susceptible to hydrolysis is present in the molecule of interest. Execution of the reaction library through the CTS will provide the user with predicted hydrolysis products for organic chemicals containing one or more of the hydrolyzable structural fragments defined in the reaction library. Future research is planned to expand the reaction library for hydrolysis to account for the effects of other dissolved species typically found in natural waters and to develop additional reaction libraries for other environmental transformation processes. Additionally, for hydrolysis reaction schemes with an adequate amount of kinetic data available, QSARs are under development to evaluate whether the incorporation of more sophisticated approaches for estimation of transformation rate constants will improve our ability to predict the likely transformation products over the current approach using ranked reaction schemes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency (EPA). Mention of trade names or products does not convey, and should not be interpreted as conveying, official EPA approval, endorsement or recommendation. This research was supported in part by an appointment to the Internship/Research Participation Program at the National Exposure Research Laboratory administered by the Oak Ridge Institute for Science and Education through Interagency Agreement No. DW-922983301-01 between the U.S. Department of Energy and the U.S. EPA.

REFERENCES

- (1). Ellington JJ Hydrolysis rate constants for enhancing property-reactivity relationships; EPA/600/3-89/063; U.S. Environmental Protection Agency: Washington, DC, 1989.
- (2). Kollig HP, Ellington JJ, Karickhoff SW, Kitchens BE, Long JM, Weber EJ and Wolfe NL Environmental fate constants for organic chemicals under consideration for EPA's Hazardous Waste Identification Rule (HWIR); EPA 601R92006; U.S. Environmental Protection Agency: Washington, DC, 1992.
- (3). Katagi T Abiotic hydrolysis of pesticides in the aquatic environment. Rev. Environ. Contam. Toxicol 2002, 175: 79-261. [PubMed: 12206055]

- (4). Mabey W and Mill T Critical review of hydrolysis of organic compounds in water under environmental conditions. J. Phys. Chem. Ref. Data 1978, 7(2): 383–415.
- (5). Data Requirements for Pesticide Registration in OECD Member Countries: Survey Results; OECD Environmental Monographs No. 77; Organization for Economic Co-operation and Development: Paris, 1994; [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=ocde/gd\(94\)47](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=ocde/gd(94)47).
- (6). Fate, Transport and Transformation Test Guidelines: OPPTS 835.2120 Hydrolysis; EPA 712-C-08-012; U.S. Environmental Protection Agency: Washington, DC, 2008; <http://www.regulations.gov/contentStreamer?documentId=EPA-HQ-OPPT-2009-0152-0009>.
- (7). Suárez LA PRZM-3, A Model for Predicting Pesticide and Nitrogen Fate in the Crop Root and Unsaturated Soil Zones: User's Manual for Release 3.12.2; EPA/600/R-05/111; U.S. Environmental Protection Agency: Washington, DC, 2005.
- (8). Burns LA, Cline DM and Lassiter RP Exposure Analysis Modeling System (EXAMS): User Manual and System Documentation; EPA/600/R-00/081; U.S. Environmental Protection Agency: Washington, DC, 2000.
- (9). Marchant CA, Briggs KA and Long A In silico tools for sharing data and knowledge on toxicity and metabolism: DEREK for windows, METEOR, and VITIC. Toxicol. Mech. Methods 2008, 18(2–3): 177–187. [PubMed: 20020913]
- (10). Mekenyan O, Dimitrov S, Pavlov T, Dimitrova G, Todorov M, Petkov P, Kotov S Simulation of Chemical Metabolism for Fate and Hazard Assessment. V. Mammalian Hazard Assessment, SAR QSAR Environ Res. 2012, 23(5–6): 553–606. [PubMed: 22536822]
- (11). Gao J, Ellis LB and Wackett LP The University of Minnesota Pathway Prediction System: multi-level prediction and visualization. Nucl. Acids Res 2011, 39(Suppl. 2): W406–W411. [PubMed: 21486753]
- (12). Wicker J, Lorschach T, Gütlein M, Schmid E, Latino D, Kramer S, and Fenner K enviPath – The environmental contaminant biotransformation pathway resource. Nucl. Acids Res 2016, 44: D502–D508. [PubMed: 26582924]
- (13). Sedykh A, Saiakhov R and Klopman G META V. A model of photodegradation for the prediction of photoproducts of chemicals under natural-like conditions. Chemosphere 2001, 45: 971–981. [PubMed: 11695620]
- (14). Kleinman MH, Baertschi SW, Alsante KM, Reid DL, Mowery MD, Shimanovich R, Foti C, Smith WK, Reynolds DW, Nefliu M and Ott MA In Silico Prediction of Pharmaceutical Degradation Pathways: A Benchmarking Study. Mol. Pharmaceutics 2014, 11(11): 4179–4188.
- (15). Wolfe K, Pope N, Parmar R, Galvin M, Stevens C, Weber E, Flaishans J and Purucker T Chemical Transformation System: Cloud Based Cheminformatic Services to Support Integrated Environmental Modeling. In: Proceedings of the 8th International Congress on Environmental Modelling and Software; Sauvage S, Sánchez-Pérez JM, Rizzoli AE (Eds.); Toulouse, France., 7 10–14, 2016; ISBN:978-88-9035-745-9.
- (16). Larson RA and Weber EJ Reaction Mechanisms in Environmental Organic Chemistry. CRC Press, Inc.: Boca Raton, FL, 1994.
- (17). Pirok G, Máté N, Varga J, Szegezdi J, Vargyas M, Dóránt S, and Csizmadia F Making “real” molecules in virtual space. J. Chem. Inf. Model 2006, 46: 563–568. [PubMed: 16562984]
- (18). OECD Guidelines for the Testing of Chemicals Nr. 111: Hydrolysis as a Function of pH; Organization for Economic Co-operation and Development: Paris, 2004; http://www.oecd-ilibrary.org/environment/test-no-111-hydrolysis-as-a-function-of-ph_9789264069701-en.
- (19). Hoaglin DC, and Iglewicz B Fine-Tuning Some Resistant Rules for Outlier Labeling. J. Am. Stat. Assoc. 1987, 82(400): 1147–1149.
- (20). Boethling RS, Howard PH, Meylan W, Stiteler W, Beauman J and Tirado N Group contribution method for predicting probability and rate of aerobic biodegradation. Environ. Sci. Technol 1994, 28(3): 459–485. [PubMed: 22165881]
- (21). Freeze RA; Cherry JA Groundwater. Prentice-Hill, Inc.: Englewood Cliffs, New Jersey, 1979; pp. 1–13.
- (22). Washington JW Hydrolysis rates of dissolved volatile organic compounds: Principles, temperature effects and literature review. Ground Water 1995, 33(3): 415–424.

- (23). Brooks JR, Gibson JJ, Birks SJ, Weber MH, Rodecap KD and Stoddard JL Stable isotope estimates of evaporation: inflow and water residence time for lakes across the United States as a tool for national lake water quality assessments. *Limnol. Oceanogr.* 2014, 59(6): 2150–2165.
- (24). Camacho RA and Martin JL Hydrodynamic modeling of first-order transport timescales in the St. Louis Bay Estuary, Mississippi. *J. Environ. Eng* 2013, 139(3): 317–331.
- (25). De Brauwere A, de Brye B, Blaise S and Deleersnijder E Residence time, exposure time and connectivity in the Scheldt Estuary. *J. Mar. Syst.* 2011, 84(3–4): 85–95.
- (26). Newson M *Hydrology and the River Environment*. Oxford University Press, Inc.: New York, 1994; pp. 3–9.
- (27). Penning H, Cramer CJ and Elsner M Rate-dependent carbon and nitrogen kinetic isotope fractionation in hydrolysis of isoproturon. *Environ. Sci. Technol* 2008, 42(21): 7764–7771. [PubMed: 19031858]
- (28). Draft Assessment Report (DAR): Initial risk assessment provided by the Member State the Netherlands for the existing active substance Pencycuron of the third stage (part A) of the review programme referred to in Article 8(2) of Council Directive 93/41/EEC, Volume 3, Annex B, B.8.; EFSA (European Food Safety Authority): Parma, Italy, 2006; <http://dar.efsa.europa.eu/dar-web/provision>.
- (29). McMurry JE *Organic Chemistry*, Eighth Edition, Cengage Learning: Boston, 2011; pp 814–852.
- (30). Turner JA, Ed. *The Pesticide Manual*, 17th, ed.; British Crop Protection Council: Hampshire, U.K., 2015.
- (31). List of Registered Pesticides (Federal Insecticide, Fungicide, and Rodenticide Act); http://scorecard.goodguide.com/chemical-groups/one-list.tcl?short_list_name=pest
- (32). Mansouri K, Abdelaziz A, Rybacka A, Roncaglioni A, Tropsha A, Varnek A, Zakharov A, Worth A, Richard AM, Grulke CM, Trisciuzzi D, Fourches D, Horvath D, Benfenati E, Muratov E, Wedebye EB, Grisoni F, Mangiatordi GF, Incisivo GM, Hong H, Ng HW, Tetko IV, Balabin I, Kancherla J, Shen J, Burton J, Nicklaus M, Cassotti M, Nikolov NG, Nicolotti O, Andersson PL, Zang Q, Politi R, Beger RD, Todeschini R, Huang R, Farag S, Rosenberg SA, Slavov S, Hu X, and Judson RS CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. *Environ. Health Perspect* 2016, 124(7): 1023–1033. [PubMed: 26908244]

Figure 1:

Examples of reaction schemes included in the abiotic hydrolysis reaction library. The schemes are written using the notation and structural query features (L, ~L, L1-X, etc) from ChemAxon's Marvin tools. The Supporting Information provides definitions for the subset of Marvin query features used in the reaction schemes.

Figure 2:

Box-and-whisker plots showing the distribution of 25°C-corrected half-lives at neutral pH for all schemes included in the hydrolysis reaction library (bottom and top of the box are first and third quartiles, respectively; band inside the box is median; whiskers are max/min values, excluding outliers; + is mean; O is outlier)

Table 1:

Rank assignments for hydrolysis schemes according to median half-lives for hydrolytic transformation of organic chemicals in the environment

Rank	Range of Median Hydrolysis Half-Life
6	Less than 2.4 hours
5	2.4 hours to 24 hours
4	24 hours to 7 days
3	7 days to 60 days
2	60 days to 1 year
1	Greater than 1 year

Table 2:

pH-specific ranks assigned to reaction schemes based on the median half-life for all molecules included in the database of literature-reported hydrolysis rate data

Hydrolysis Scheme	pH 5 Rank	pH 7 Rank	pH 9 Rank
Dehydration of Geminal Diols	6	6	6
Acid Halide Hydrolysis	6 ^[1]	6	6 ^[1]
Anhydride Hydrolysis	6 ^[1]	6	6
Cyclic Anhydride Hydrolysis	6 ^[1]	6	6
Lactone Hydrolysis	1	5	5
Nitrile Hydrolysis	4	5	5
Epoxide Hydrolysis	6	5	1
Imide Hydrolysis	2	5	5
N-S Cleavage	5	5	5
Lactam Hydrolysis	4 ^[3]	4	4 ^[3]
Carbonate Hydrolysis	1	3	5
Cyclic Carbonate Hydrolysis	1	3	5
Carbamate Hydrolysis	3	3	5
Halogenated Aliphatics: Nucleophilic Substitution w/ no adjacent X	3 ^[3]	3	3 ^[3]
OP Triester Hydrolysis 1 (Base-Catalyzed)	2	3	4
OP Triester Hydrolysis 2 (Neutral or Acid-Catalyzed)	2	3	3
Sulfonylurea Hydrolysis	3	2	2
Urea Hydrolysis	3	2	2 ^[2]
Cyclic Urea Hydrolysis	3	2	2 ^[2]
Carboxylic Acid Ester Hydrolysis	2	2	5
Halogenated Aliphatics: Elimination	1	1	2
Thiocarbamate Hydrolysis	1 ^[3]	1	1 ^[3]
Amide Hydrolysis	1	1	3
Halogenated Aliphatics: Nucleophilic Substitution w/ geminal X	1 ^[3]	1	1
Halogenated Aliphatics: Nucleophilic Substitution w/ vicinal X	1 ^[3]	1	1

Notes:

^[1] Because acid halides and anhydrides undergo hydrolysis via the nucleophilic acyl substitution mechanism, and both RCOO^- and X^- are very good leaving groups, it is assumed that hydrolysis at pH 5 and pH 9 will be at least as fast as at pH 7.

^[2] The limited rate data available for urea hydrolysis indicates that hydrolysis rates under basic conditions are approximately equal to rates under near-neutral conditions^{27,28}, suggesting that neutral hydrolysis is the dominant mechanism for urea hydrolysis for pH values near and above pH 7.

^[3] The mechanism for these reactions can be assumed to follow the same neutral hydrolysis mechanism over the entire pH 5–9 range; therefore, the rank at pH 5 and/or pH 9 is assumed to be the same as the rank at pH 7.

Table 3:

Number of molecules that are susceptible to the hydrolysis schemes within the library for A) 636 pesticides regulated under FIFRA and B) 32,583 chemicals used in the prediction set of the CERRAP project.

Hydrolysis Scheme	FIFRA	CERRAP
Carboxylic Acid Ester Hydrolysis	121	4423
OP Triester Hydrolysis 1 (Base-Catalyzed)	56	334
OP Triester Hydrolysis 2 (Neutral or Acid-Catalyzed)	56	306
Amide Hydrolysis	44	2683
Carbamate Hydrolysis	31	270
Nitrile Hydrolysis	30	851
Halogenated Aliphatics: Nucleophilic Substitution w/ no adjacent X	29	871
Halogenated Aliphatics: Nucleophilic Substitution w/ geminal X	28	488
Halogenated Aliphatics: Elimination	22	602
Lactam Hydrolysis	18	1100
Sulfonylurea Hydrolysis	18	90
Urea Hydrolysis	17	300
Lactone Hydrolysis	16	673
Imide Hydrolysis	14	598
Halogenated Aliphatics: Nucleophilic Substitution w/ vicinal X	11	183
Thiocarbamate Hydrolysis	11	39
Epoxide Hydrolysis	9	279
N-S Cleavage	3	18
Anhydride Hydrolysis	1	36
Cyclic Urea Hydrolysis	1	66
Acid Halide Hydrolysis	0	173
Carbonate Hydrolysis	0	61
Dehydration of Geminal Diols	0	26
Cyclic Carbonate Hydrolysis	0	6