A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP4555.

A Chemical Category-Based Prioritization Approach for Selecting 75 Per- and Polyfluoroalkyl Substances (PFAS) for Tiered Toxicity and Toxicokinetic Testing

Grace Patlewicz,[1] Ann M. Richard,[1] Antony J. Williams,[1] Christopher M. Grulke,[1] Reeder Sams,[1] Jason Lambert,[2] Pamela D. Noyes,[3] Michael J. DeVito,[4] Ronald N. Hines,[5] Mark Strynar,[6] Annette Guiseppi-Elie,[6] and Russell S. Thomas[1]

[1]National Center for Computational Toxicology, Office of Research and Development (ORD), U.S. Environmental Protection Agency (EPA), Research Triangle Park, North Carolina, USA
[2]National Center for Environmental Assessment (NCEA), ORD, U.S. EPA, Cincinnati, Ohio, USA
[3]Integrated Risk Information System Division, NCEA, ORD, U.S. EPA, Washington, District of Columbia, USA
[4]National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, USA
[5]National Health and Environmental Effects Research Laboratory, ORD, U.S. EPA, Research Triangle Park, North Carolina, USA
[6]National Exposure Research Laboratory, ORD, U.S. EPA, Research Triangle Park, North Carolina, USA

Summary: Per- and polyfluoroalkyl substances (PFASs) are a group of fluorinated substances of interest to researchers, regulators, and the public due to their widespread presence in the environment. A few PFASs have comparatively extensive amounts of human epidemiological, exposure, and experimental animal toxicity data (e.g., perfluorooctanoic acid), whereas little toxicity and exposure information exists for much of the broader set of PFASs. Given that traditional approaches to generate toxicity information are resource intensive, new approach methods, including *in vitro* high-throughput toxicity (HTT) testing, are being employed to inform PFAS hazard characterization and further (*in vivo*) testing. The U.S. Environmental Protection Agency (EPA) and the National Toxicology Program (NTP) are collaborating to develop a risk-based approach for conducting PFAS toxicity testing to facilitate PFAS human health assessments. This article describes the construction of a PFAS screening library and the process by which a targeted subset of 75 PFASs were selected. Multiple factors were considered, including interest to the U.S. EPA, compounds within targeted categories, structural diversity, exposure considerations, procurability and testability, and availability of existing toxicity data. Generating targeted HTT data for PFASs represents a new frontier for informing priority setting. https://doi.org/10.1289/EHP4555

Introduction

Per- and polyfluoroalkyl substances (PFASs) are a group of fluorinated substances that have generated increased public attention due to their potential health hazard and widespread presence in the environment (Wang et al. 2017; Xiao 2017; Ross et al. 2018). The U.S. Environmental Protection Agency (EPA) Office of Research and Development (ORD) in partnership with the National Toxicology Program (NTP) are currently engaged in producing toxicity information to facilitate human health assessments for PFASs. A few PFASs have comparatively extensive amounts of toxicity data (e.g., perfluorooctanoic acid), but little toxicity information exists for much of the broader set of PFASs identified from preliminary exposure studies that capture potential occurrence in the environment. The hundreds of untested PFASs provide a scenario in which traditional one-by-one toxicity testing would consume tremendous resources and useful toxicity information would not be available for decades. The U.S. EPA's ToxCast program and the multi-federal agency Tox21 program (which includes the NTP and the U.S. EPA as major partners) have developed the capacity to screen hundreds to thousands of chemicals for bioactivity through *in vitro* high-throughput toxicity (HTT) testing. Data generated from these assays are already being used to inform hazard identification and prioritize chemicals for further *in vivo* testing (U.S. EPA 2012, 2014a, 2014b, 2015; Judson et al. 2010, 2015; Kleinstreuer et al. 2017). Within the U.S. EPA, generating such data to inform agency and partner decision making regarding potential human health hazard and risk across the broad landscape of PFASs represents a real-world challenge that HTT coupled with cheminformatic approaches is uniquely designed to address.

This article describes, in brief, the development of the PFAS screening library and the process by which a subset of 75 PFAS substances were selected for HTT screening and tiered toxicity testing, along with mention of the toxicity and toxicokinetic experiments currently underway.

Discussion

Development of the PFAS Screening Library

Since there are no specific chemical catalogs for PFASs, an initial scoping for potentially procurable PFAS substances relied on the use of candidate PFAS structure lists generated from the U.S. EPA's Distributed Structure–Searchable Toxicity (DSSTox) chemical database. DSSTox currently exceeds 760,000 substances, each of which has undergone some level of chemical structure curation prior to registration (Williams et al. 2017). The largest registered list of PFAS chemicals available at the time this study was initiated was the KEMI PFAS list in DSSTox (named PFASKEMI and available for download at https://comptox.epa.gov/dashboard/chemical_lists/pfaskemi). Approximately 1,200 structures from this list were provided to the chemical contractor for scoping purposes, from which approximately 600 substances were identified as potentially procurable but likely to require on-demand synthesis and exceed standard costs. Based on this preliminary scoping, U.S. EPA funds were secured for the purchase and processing of approximately 400 substances to create a PFAS testing library.

The first procurement phase considered the feasibility of procuring substances of interest to the U.S. EPA. A U.S. EPA workgroup was formed to identify PFASs of interest to U.S. EPA programs and regions and to include PFASs with associated toxicity data that would inform human health risk assessment. The final set of 31 PFASs recommended for further study by this workgroup (list denoted as EPA PFAS WG 31) identified PFASs whose review may support risk evaluation. Also included in the

Address correspondence to G. Patlewicz, National Center of Computational Toxicology (NCCT), 109 T.W. Alexander Dr., Mail Code: D143-02, Research Triangle Park, NC 27709 USA. Telephone: (919) 541-1540. Email: patlewicz.grace@epa.gov

The authors declare they have no actual or potential competing financial interests.

Received 9 October 2018; Revised 27 November 2018; Accepted 3 December 2018; Published 11 January 2019.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

request list for the first phase of PFAS procurements were PFASs that spanned a wider range of U.S. EPA research activities. This larger list of potentially procurable PFASs initially consisted of 89 unique substances (inclusive of EPA PFAS WG 31), which we denote here as EPA-PFAS (note, an updated, expanded U.S. EPA research list, inclusive of salts and anions, is titled EPAPFASRL and is available for download at https://comptox.epa.gov/dashboard/chemical_lists/epapfasrl).

The second phase of PFAS procurements considered a query of the expanded contents of the DSSTox database for chemicals satisfying a range of PFAS-defining criteria, including: >3 fluorines, no aromatic rings (i.e., aliphatic), and molecular weight (MW) <500. This list was reviewed manually and additional filters were applied to reduce the size of the list from >4,700 substances to <800 PFAS candidates for possible procurement, for example, excluding heavy metals, halogen salts, low-MW compounds (<100 amu) and compounds for which the ratio of F to C was less than 2:1. The initial set of compounds from this sub-list, for which procurement sources were identified, underwent manual expert DSSTox curation review. The resulting compound list, after confirmation of procurement feasibility, formed the remainder of the initial structure library considered in the present prioritization exercise. This final set of 271 DSSTox-registered substances is referred to herein as the PFAS-Landscape. This set of substances bounded the range of PFASs considered in the below analysis, which was used to identify the candidate 75 subset for tiered toxicity testing.

Categorization of the PFAS Screening Library

Although there are many ways to systematically select a representative subset of structures from a library using different structure-based cheminformatic approaches, in this study predefined, expert-based structural categories were relied upon to characterize the PFAS screening library. The structural categories initially proposed were informed by the work by Buck et al. (2011), who described a systematic terminology for naming and categorizing PFAS substances. The structural categories were used to manually assign each substance in the 271 PFAS library into a respective structural category. To maintain a practical and pragmatic number of structural categories, after this initial assignment was completed, some of the structural categories were combined, for example, *n*:1 fluorotelomer alcohols and *n*:2 fluorotelomer alcohols were collapsed into a single category of fluorotelomers. The linkages between the general categories and more specific categories (i.e., subcategories) were retained to offer additional flexibility in the selection of substances for testing. Retaining this layered category information could be particularly useful should specific activity trends within categories be identified and subsequently investigated. Overall, 53 unique structural categories were assigned. Some of the categories contained many more substances (members) than others. Categories containing only one member were referred to as singletons.

Process for Selection of a Subset of 75 PFASs

Using the expert-assigned structural categories described above, we constructed a step-wise workflow to guide efforts to prioritize and weigh various factors for chemical selection within categories. The workflow is graphically illustrated in Figure 1 and consists of an initial PFAS-Landscape characterization step (0), followed by a series of five steps, described in more detail below, to balance the somewhat competing goals of creating a data set that would support read-across within categories while also capturing structural diversity aspects of the PFAS landscape.

Characterizing the PFAS Library (Step 0)

Structural diversity of the full PFAS-Landscape can be represented both in terms of overall chemical counts within categories and labeled by chemical membership in one of the three main groupings, ordered by level of U.S. EPA interest, that is, the 31 PFASs with associated data to inform human health risk assessment (EPA PFAS WG 31), additional PFASs of interest to U.S. EPA researchers (EPA-PFAS), and the remaining PFAS Landscape (PFAS-Landscape) initially identified as procurable. This is illustrated in Figure 2a.

The U.S. EPA's ToxVal database is a database of source-referenced human health reference or toxicity values collected from *in vivo* studies that are available through the U.S. EPA CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard). Figure 2b provides a representation of record(s) from ToxVal as surrogates

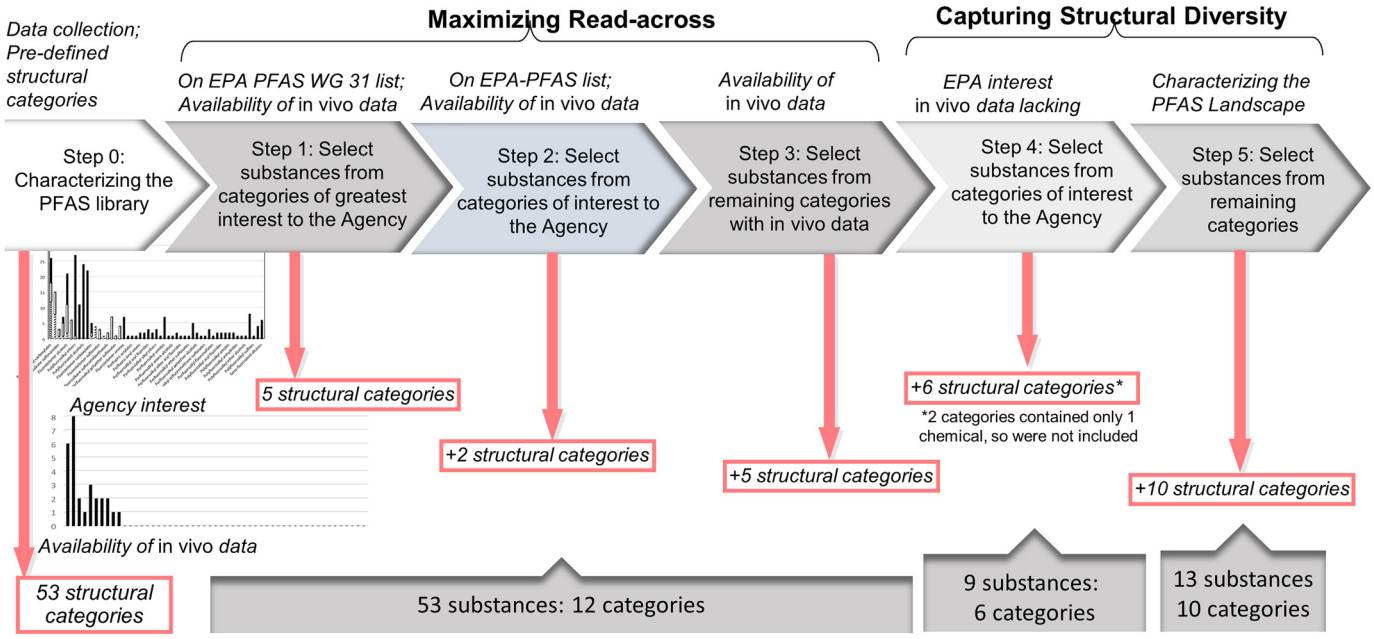

Figure 1. Workflow for selection of structural categories to identify the subset of 75 per- and polyfluoroalkyl substances (PFAS).

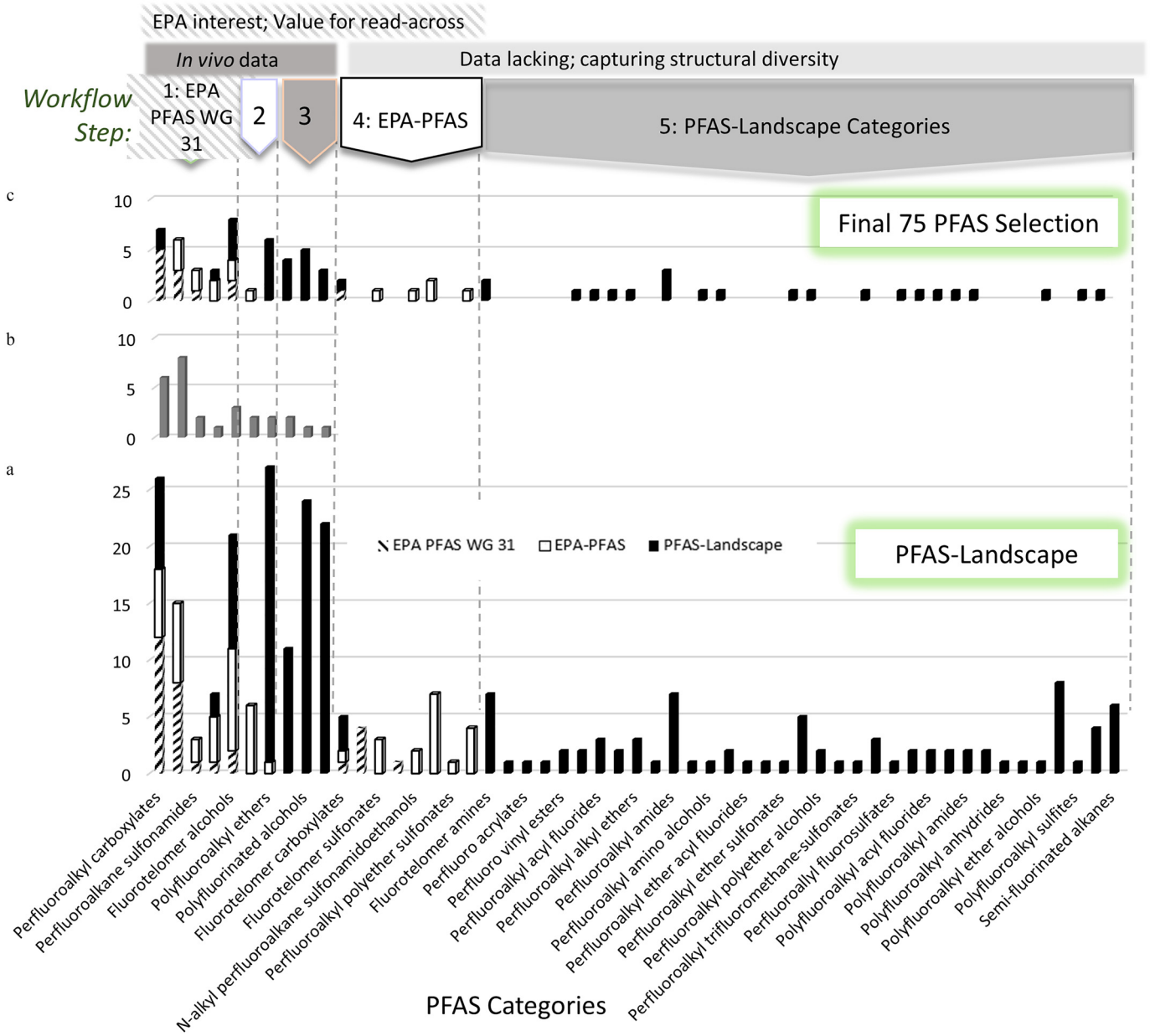

Figure 2. Per- and polyfluoroalkyl substances (PFAS) Library characteristics by category aligned with workflow steps shown in Figure 1: (*a*) Total chemical count plot showing the total number of PFASs in a specific structural category color-coded by "status," that is, membership on a single list (EPA PFAS WG 31 > EPA-PFAS >PFAS-Landscape); (*b*) ToxVal record count plot showing the number of PFASs with *in vivo* toxicity studies available within the category, as sourced from the U.S. EPA ToxVal database; (*c*) Same count plot as in *a*, showing only the final selected 75 PFAS substances by list membership and category representation. Color figures are available at https://doi.org/10.23645/epacomptox.7479866.

for *in vivo* toxicity information per chemical (1, yes; 0, no). These records are mapped onto the same category legend as shown in Figure 2a and demonstrate that toxicity data are available for only a limited subset of categories of interest to the U.S. EPA.

Selection of Structural Categories (Steps 1–5)

Strategies for selecting a subset of PFAS substances for HTT from the PFAS-Landscape focused on two main objectives: 1) maximizing information to support read-across within structure-based groupings, and 2) capturing the structural diversity of the PFAS landscape of interest to the U.S. EPA. Steps 1–3 of the workflow in Figure 1 identified structural categories addressing Objective 1 (maximizing read-across), whereas Steps 4–5 identified structural categories addressing Objective 2 (capturing structural diversity).

In Step 1 of the workflow, structural categories were identified that were both of high interest to the U.S. EPA (EPA PFAS WG 31) and had a record in ToxVal indicating availability of *in vivo* data. Step 2 identified structural categories that were of broader interest to the U.S. EPA (EPA-PFAS and PFAS-Landscape) and had a record in ToxVal. Structural categories that had a record in ToxVal but were not captured in Steps 1–2, that is, of lesser interest to the agency, were identified in Step 3. Step 4 identified structural categories of interest to the U.S. EPA irrespective of having a record in ToxVal. The remaining structural categories in PFAS-Landscape not captured in Steps 1–4 were considered as part of Step 5.

Table 1. Considerations for selection of the 75 PFAS.

Aspect name	Scoring
1) Structural diversity within a category	Approximated by category size, with score ranging from 1 (≥20 members) to 0 (1 member).
2) Data availability	Availability of *in vitro* ToxCast data (score = 0.5) or ToxVal *in vivo* data (score = 0.75) or both (score = 1).
3) Data quantity	Number of ToxVal records for a substance indicating a stronger source–analog for read-across, with scores ranging from 0.15 (for 1 record) to 1 (for ≥20 records).
4) Read-across category-level weight	Value of substance for anchoring read-across trends within a category (e.g., chain length), serving as a source analog (score = 0.5) or target analog (score = 0.25), or as a target analog for capturing structural diversity (score = 0.15).
5) Numerical indicator of U.S. EPA interest	EPA PFAS WG31 (score = 1), other U.S. EPA-PFAS (score = 0.75), only in PFAS-Landscape (score = 0.5).
6) Physicochemical indicators of testability	Both logKow and vapor pressure favorable (score = 0.75), one favorable (score = 0.5), both unfavorable (score = 0). For example, logKow <4.5, vapor pressure $<10^3$mmHg considered favorable. LogKow and vapor pressure properties relied on OPERA model predictions as available from the U.S. EPA CompTox Chemicals Dashboard. It is recognized that there are issues surrounding the validity of predictions for PFAS substances. The predictions here were used in relative terms within a structural category.
7) Figure 1 workflow step	Step 1 (score = 1), Step 2 (score = 0.75), Step 3 (score = 0.5), Step 4 (score = 0.25), Step 5 (score = 0).
Total score	Summation of scores from the preceding considerations used to rank each PFAS substance.

Note: OPERA, OPEn structure-activity/property Relationship App; PFAS, per- and polyfluoroalkyl substances.

Substances to address read-across (Objective 1) were drawn from the 10 categories identified in Steps 1–3. Substances for capturing structural diversity were drawn from the remaining 43 categories.

Selection of PFASs within the Prioritized Structural Categories

Test substances were selected on a structural category basis. The process was guided by a quantitative scoring scheme that aimed to capture and rank considerations such as structural variation and physical property information (logKow, vapor pressure), as well as availability of a record in the ToxCast and/or ToxVal databases. The scheme comprised seven aspects described in Table 1.

An initial selection of 75 substances was made based on these scoring considerations. Backup alternatives were also selected based on the same scoring considerations. However, technical challenges were encountered as procurement orders were processed. Of specific note was the physical form of the test substance received (i.e., gas vs. solid), hazmat considerations (e.g., flammability), evidence for volatility/sublimation of stored neat samples, and insolubility in dimethyl sulfoxide (DMSO). These technical considerations resulted in further adjustment of the selection of substances from specific categories, which in turn impacted the degree of structural diversity reflected in the final PFAS procured library as well as the extent to which categories for maximizing read-across were represented. The final set of 75 substances for which DMSO solutions were prepared, and test samples submitted for HTT, is represented graphically by category and status assignment in Figure 2c. The final DSSTox list of 75 PFAS substances, with associated structural information, is labeled EPAPFAS75S1 and is available for download at https://comptox.epa.gov/dashboard/chemical_lists/epapfas75S1.

HTT: Tiered Toxicity and Toxicokinetic Testing

The final set of 75 PFAS substances is currently undergoing targeted and tiered HTT in partnership with NTP. Tier 1 HTT includes *in vitro* assays focused on multiple end points, including hepatotoxicity, immunotoxicity, developmental toxicity, mitochondrial toxicity, and developmental neurotoxicity along with assays to estimate *in vivo* toxicokinetics. The assays selected for the Tier 1 toxicity and toxicokinetic characterization were based on both the known *in vivo* adverse responses of previously tested PFASs and the anticipated effects of a broader range of PFAS. In general, the proposed strategy is to utilize data generated from new approach methods (e.g., HTT toxicological and toxicokinetic assays) in combination with human exposure information (measured and/or predicted) to derive a biological exposure ratio (BER; Thomas et al. 2013). BERs will serve as a measure of potential risk and will be used to prioritize subsequent Tier 2 *in vivo* testing and inform human health risk assessment. Data generated from *in vitro* assays will also be used with existing *in vivo* information to support read-across efforts (Patlewicz et al. 2013).

Summary

A PFAS screening library was constructed and categorized by expert review into 53 structural categories. The final PFAS 75 (EPAPFAS75S1) comprised 46 substances representing 10 of the structural categories with some existing *in vivo* toxicity information and 29 substances covering a further 24 structural categories (Figure 2c). New PFAS candidates will be selected from the U.S. EPA's complete, DMSO-solubilized PFAS inventory (EPAPFASINV), now totaling 430 unique DSSTox substances. This complete inventory and the list of 43 PFASs that were procured but found to be DMSO-insoluble (EPAPFASINSOL) as well as the final EPAPFAS75S1 list are available for download at https://comptox.epa.gov/dashboard/chemical_lists/pfasmaster.

Acknowledgments

Multiple individuals and groups throughout EPA provided information that contributed to the development of this article. EPA list contributions and literature review by Linda Gaines, and Lynn Flowers are gratefully acknowledged. Special thanks to DSSTox curators, Inthirany Thillainadarajah and Brian Meyers, for expert review and structure-annotation of the various PFAS chemical lists compiled for this study. The PFAS Screening Library was procured through Evotec (US) Inc., (EPA Contract No. 68HE0D18D0001). This work was supported (in part for MDV) by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

References

Buck RC, Franklin J, Berger U, Conder JM, Cousins IT, de Voogt P, et al. 2011. Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins. Integr Environ Assess Manag 7(4):513–541, PMID: 21793199, https://doi.org/10.1002/ieam.258.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. 2010. *In vitro* screening of environmental chemicals for targeted testing prioritization: the ToxCast project. Environ Health Perspect 118(4):485–492, PMID: 20368123, https://doi.org/10.1289/ehp.0901392.

Judson RS, Magpantay FM, Chickarmane V, Haskell C, Tania N, Taylor J, et al. 2015. Integrated model of chemical perturbations of a biological pathway using 18 *in vitro* high-throughput screening assays for the estrogen receptor. Toxicol Sci 148(1):137–154, PMID: 26272952, https://doi.org/10.1093/toxsci/kfv168.

Kleinstreuer NC, Ceger P, Watt ED, Martin M, Houck K, Browne P, et al. 2017. Development and validation of a computational model for androgen receptor activity. Chem Res Toxicol 30(4):946–964, PMID: 27933809, https://doi.org/10.1021/acs.chemrestox.6b00347.

Patlewicz G, Ball N, Booth ED, Hulzebos E, Zvinavashe E, Hennes C. 2013. Use of category approaches, read-across and (Q)SAR: general considerations. Regul Toxicol Pharmacol 67(1):1–12, PMID: 23764304, https://doi.org/10.1016/j.yrtph.2013.06.002.

Ross I, McDonough J, Miles J, Storch P, Kochunarayanan T, Kalve E, et al. 2018. A review of emerging technologies for remediation of PFASs. Remediation (N Y) 28(2):101–126, https://doi.org/10.1002/rem.21553.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, et al. 2013. Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. Toxicol Sci 136(1):4–18, PMID: 23958734, https://doi.org/10.1093/toxsci/kft178.

U.S. EPA (Environmental Protection Agency). 2012. FIFRA Scientific Advisory Panel; notice of public meeting. Docket No. EPA–HQ–QPP–2012–0818, FRL–9367–9. Fed Reg 77(222):68773–68775, https://www.gpo.gov/fdsys/pkg/FR-2012-11-16/pdf/2012-27816.pdf [accessed 5 December 2018].

U.S. EPA. 2014a. FIFRA Scientific Advisory Panel; notice of public meeting. Docket No. EPA–HQ–OPP–2014–0331, FRL–9910–22. Fed Reg 79(104):31111–31113. https://www.gpo.gov/fdsys/pkg/FR-2014-05-30/pdf/2014-12593.pdf [accessed 5 December 2018].

U.S. EPA. 2014b. FIFRA Scientific Advisory Panel; notice of public meeting. Docket No. EPA–HQ–OPP–2014–0614, FRL–9915–55. Fed Reg 79(179):55475–55477. https://www.gpo.gov/fdsys/pkg/FR-2014-09-16/pdf/2014-22044.pdf [accessed 5 December 2018].

U.S. EPA. 2015. Use of high throughput assays and computational tools; endocrine disruptor screening program; notice of availability and opportunity for comment. EPA–HQ–OPPT–2015–0305, FRL–9928–69. Fed Reg 80(118):35350–35355. https://www.gpo.gov/fdsys/pkg/FR-2015-06-19/pdf/2015-15182.pdf [accessed 5 December 2018].

Wang Z, DeWitt JC, Higgins CP, Cousins IT. 2017. A never-ending story of per- and polyfluoroalkyl substances (PFASs)? Environ Sci Technol 51(5):2508–2518, PMID: 28224793, https://doi.org/10.1021/acs.est.6b04806.

Williams AJ, Grulke CM, Edwards J, McEachran AD, Mansouri K, Baker NC, et al. 2017. The CompTox Chemistry Dashboard—a community data resource for environmental chemistry. J Cheminformatics 9(1):61, PMID: 29185060, https://doi.org/10.1186/s13321-017-0247-6.

Xiao F. 2017. Emerging poly- and perfluoroalkyl substances in the aquatic environment: a review of current literature. Water Res 124:482–495, PMID: 28800519, https://doi.org/10.1016/j.watres.2017.07.024.