

1 **High-throughput protein target mapping enables accelerated** 2 **bioactivity discovery for ToxCast and PFAS compounds**

3 Diwen Yang^{1,2}, Xiaoyun Wang^{1,*}, Jiabao Liu^{1,3}, Yufeng Gong¹, Pranav Nair¹, Jianxian Sun^{1,4}, Xing
4 Qian¹, Claire Cui¹, Hong Zeng⁴, Aiping Dong⁴, Rachel J. Harding^{4,5,6}, Nicola Burgess-Brown⁷,
5 Tyler S. Beyett⁸, Datong Song¹, Henry Krause³, Miriam L. Diamond^{9,10}, Derek L. Bolhuis¹¹,
6 Nicholas G. Brown¹², Cheryl H. Arrowsmith^{4,13,14}, Aled M. Edwards⁴, Levon Halabelian^{4,5}, Hui
7 Peng^{1,4,9*}

8 ¹Department of Chemistry, University of Toronto, Toronto, ON, Canada

9 ²Department of Physical and Environmental Sciences, University of Toronto, Toronto, ON,
10 Canada

11 ³Terrence Donnelly Centre for Cellular & Biomolecular Research, University of Toronto, Toronto,
12 ON, Canada.

13 ⁴Structural Genomics Consortium, University of Toronto, Toronto, ON, Canada

14 ⁵Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada

15 ⁶Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada

16 ⁷Structural Genomics Consortium, University College London, London, United Kingdom

17 ⁸Department of Pharmacology and Chemical Biology, Emory University, Atlanta, GA, United
18 States

19 ⁹School of the Environment, University of Toronto, Toronto, ON, Canada

20 ¹⁰Department of Earth Sciences, University of Toronto, Toronto, ON, Canada

21 ¹¹Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel
22 Hill, NC, United States

23 ¹²Pharmacology Department, University of North Carolina at Chapel Hill, Chapel Hill, NC, United
24 States

25 ¹³Princess Margaret Cancer Centre, University of Toronto, Toronto, ON, Canada

26 ¹⁴Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

27 *Corresponding authors:

28 **Xiaoyun Wang**, xyun.wang@utoronto.ca; Department of Chemistry, University of Toronto

29 **Hui Peng**, hui.peng@utoronto.ca; Department of Chemistry, University of Toronto

30

Abstract

Chemical pollution is a global threat to human health, yet the toxicity mechanism of most contaminants remains unknown. Here, we applied an ultrahigh-throughput affinity-selection mass spectrometry (AS-MS) platform to systematically identify protein targets of prioritized chemical contaminants. After benchmarking the platform, we screened 50 human proteins against 481 prioritized chemicals, including 446 ToxCast chemicals and 35 per- and polyfluoroalkyl substances (PFAS). Among 24,050 interactions assessed, we discovered 35 novel interactions involving 14 proteins, with fatty acid-binding proteins (FABPs) emerging as the most ligandable protein family. Given this, we selected FABPs for further validation, which revealed a distinct PFAS binding pattern: legacy PFAS selectively bound to FABP1, whereas replacement compounds, PFECAs, unexpectedly interacted with all FABPs. X-ray crystallography further revealed that the ether group enhances molecular flexibility of alternative PFAS, to accommodate the binding pockets of FABPs. Our findings demonstrate that AS-MS is a robust platform for the discovery of novel protein targets beyond the scope of the ToxCast program and highlight the broader protein-binding spectrum of alternative PFAS as potential regrettable substitutes.

Introduction

The global production and use of chemicals have expanded significantly, with over 350,000 chemicals currently registered worldwide¹. Many of these chemicals ultimately enter human body, leading to undesired adverse health effects. In 2015, chemical contaminant exposure was estimated to contribute to approximately 9 million premature deaths globally². However, the sheer scale of chemical contaminants far exceeds the capacity of traditional toxicity testing³, resulting in a limited understanding of their toxicity mechanism. This knowledge gap has often led to the widespread use of chemicals that were later proven to be hazardous and regrettable⁴⁻⁶. However, identification of the toxicity mechanism (*i.e.*, protein targets) of toxicants is extremely challenging and is the bottleneck for chemical management. Over the past decades, extensive efforts have been invested to assess the toxicity mechanism of chemical contaminants, by employing *in vitro* high-throughput screening (HTS) as a key strategy⁷.

The U. S. Environmental Protection Agency (EPA) launched the ToxCast program in 2007, which aims to systematically assess the molecular toxicity of ~4,500 priority chemicals across ~900 bioassays⁸. As the largest HTS campaign to date, the ToxCast program focuses on ~300 protein targets, including common targets of toxins such as cytochromes and nuclear receptors (NRs). However, this subset of human proteins represents only 1.5% of ~20,000 human proteins, leaving many potentially involved in pathways that contribute to toxicity unexplored. For instance, thalidomide was included in the ToxCast, yet its primary target cereblon (CRBN) identified in 2010⁶, was absent from the screening program. The broader expansion of the ToxCast screening protein targets has proven challenging due to the limited availability of bioassays for most human proteins. Indeed, traditional HTS assays are often designed specifically for one protein at a time, which do not scale across the full proteome. To address these gaps, it is important to develop new

approaches that have the potential to screen protein targets at an unbiased, and potentially at a proteome-wide level.

Protein affinity selection mass spectrometry (AS-MS) is an unbiased approach for screening protein-ligand interactions⁹. Unlike traditional HTS bioassays, AS-MS is a direct binding assay that enables rapid and unbiased identification of protein-ligand interactions. Indeed, AS-MS has been successfully used to identify toxicants from environmental mixtures in our previous studies^{10–14}, primarily focusing on a few proteins (*e.g.*, nuclear receptors). This study aimed to advance the AS-MS platform for scalable discovery of novel human protein targets for environmental chemicals, achieving an unprecedented throughput beyond traditional HTS bioassays. To test this, we decided to focus on 446 ToxCast chemicals and 35 perfluoroalkyl and polyfluoroalkyl substances (PFAS) that were prioritized by the U.S. EPA^{15,16}. By employing a representative set of 50 human proteins not covered in the ToxCast program, we assessed 24,050 interactions within three weeks, leading to the identification of 35 novel protein-toxicant interactions. Among these, fatty acid-binding proteins (FABPs) emerged as the most ligandable protein family were selected for further validation. Notably, legacy PFAS selectively bound to FABP1, while alternative PFAS interacted with all seven FABP isoforms.

Results

Benchmarking the AS-MS platform

The AS-MS platform adapts our previous affinity purification with nontargeted analysis (APNA) method^{10–14}, as illustrated in Fig. 1a. In brief, all prioritized chemicals including 446 ToxCast compounds and 35 PFAS (Table S5) were mixed in one pool for AS-MS screening to increase the throughput. 100 nM of pooled compounds were incubated with 1 μ M His-tagged proteins and Ni-NTA magnetic beads in 96-well plate format. The ligands binding to target proteins were isolated

from free chemicals and proteins through affinity purifications for liquid chromatography–mass spectrometry (LC–MS) analysis. Putative ligands for each protein were identified by comparing the LC-MS signals (*i.e.*, fold change > 10, $p < 0.05$) against other proteins in the same batch. The ligands were further confirmed by orthogonal assays, such as the fluorescence displacement assay or protein-ligand co-crystal structure by X-ray crystallography.

To pilot the AS-MS platform for large-scale screening, we first conducted a double-blind experiment by spiking a known ligand ($K_D = 60$ nM) of WDR5 protein to ~16,000 compounds. This known compound was identified as the top hit within one week, with a high enrichment fold of 934 compared to negative control proteins (Fig. 1b). Encouraged by the promising results, we further benchmarked the platform using eight human proteins (WDR5, WDR91, HDAC6, LRRK2, USP21, CRBN, DCAF1, and GID4) representing diverse biological functions were selected. A total of 101 verified ligands for these proteins were mixed as a chemical pool, to evaluate the sensitivity and selectivity of our AS-MS platform. The high selectivity of the assay was demonstrated by the strong affinity ligands ($K_D < 1$ μ M) detected for WDR5, GID4, DCAF, CRBN, and HDAC6, which exhibited high fold changes while showing little to no binding to other proteins (Fig. S1). Collectively, 100%, 83%, and 15% of strong- ($K_D < 1$ μ M), moderate- (1 μ M < $K_D < 10$ μ M) and weak-affinity ($K_D > 10$ μ M) ligands were detected by our AS-MS method, respectively (Fig. 1c). Strong-affinity binding in the nanomolar (nM) range is typically essential for environmental contaminants to cause adverse health effects (*e.g.*, $K_D = 1$ nM for dioxins on AhR¹⁷), at environmentally relevant concentrations. Therefore, the high sensitivity (100%) of our AS-MS platform for strong-affinity ligands is well suited for protein target profiling of environmental toxicants.

After benchmarking the AS-MS platform, we proceeded to design a library of prioritized chemicals for screening. We decided to focus on the ToxCast (1,040 compounds) and PFAS (74

compounds) chemical libraries prioritized by the U.S. EPA, as they represent a broad spectrum of commercially used substances with widespread human exposure^{15,16,18,19}. These compounds span more than ten categories, including pharmaceuticals and pesticides, among others (Fig. 1d). PFAS compounds were selected as they are well-known ‘forever chemicals’ with concerns regarding their high persistence, bioaccumulation and toxicity^{20–22}. The U.S. EPA recently established a library of 74 PFAS for toxicity research according to a chemical category-based prioritization approach^{16,23}. As high-resolution MS is used for detecting isolated ligands, we further prioritize compounds with high LC-MS sensitivity for AS-MS screening. Among 1,114 compounds tested, 481 compounds (43.2 %), including 446 ToxCast compounds and 35 PFAS, were detectable by LC-MS at 100 nM (Fig. 1e). To assess the potential coverage of AS-MS for bioactive compound screening, we compared the LC-MS detection of 1,114 compounds to their reported bioactivity towards nuclear receptors²⁴. Among them, 85 compounds showed strong bioactivity with AC₅₀ below 10 µM (Fig. 1f), while the majority were inactive. Notably, 74 of the 85 potent compounds (87.1%) were detectable by LC-MS, demonstrating that most bioactive toxicants are LC-MS active and are covered by AS-MS. This might not be surprising, because H-bonds are typically important for both protein binding and LC-MS ionization^{25,26}. In total, 481 LC-MS active compounds were selected for AS-MS screening in subsequent studies.

Application of the AS-MS platform to 481 ToxCast/PFAS chemicals and 50 human proteins

The benchmarked AS-MS platform was then applied to 50 human proteins (Table S3) to discover novel targets of the select 481 compounds. We selected proteins with a range of molecular weights, protein structures and biological functions associated with various diseases (Fig. 2a). While these proteins have been proposed as important targets for disease treatment and drug development,^{27,28} none of these proteins have been previously assessed in the ToxCast program, probably due to the

lack of bioassays. For instance, as mentioned earlier, CRBN was identified as the primary target of thalidomide in 2010,⁶ three years after the launch of the ToxCast program. Human fatty acid-binding proteins (FABP1-7) are another protein family of interests, primarily responsible for the transport of hydrophobic ligands, such as fatty acids, within cells²⁹. Among these, FABP1 has been thoroughly studied for its role in regulating the toxicokinetic of PFAS^{17,22,23}. Additionally, leucine-rich repeat kinase 2 (LRRK2), a protein associated with Parkinson's disease (PD), has been the subject of over 3,000 publications, with three ligands progressed to preclinical trails^{32,33}.

The screens were completed within three weeks, assessing 24,050 interactions (50 proteins × 481 compounds). We identified 35 compound-protein pairs targeting 14 proteins, representing a 10-fold enrichment compared to other proteins screened in the same batch (Fig. 2b). This leads to an average hit rate of 0.15% at the compound level which is 2.9-times lower than the hit rate of NRs covered by the ToxCast program (0.44%)²⁴. This is expected, as many proteins including NRs covered by the ToxCast program, are among the most studied and ligandable proteins. Among the 50 proteins studied here, FABPs are the most detected targets, with 1-16 putative ligands discovered for each of FABP1-7 (Fig. 2c). This might not be surprising as FABPs are major transporters of fatty acids, and they have been reported to bind to xenobiotics including 'forever compounds' PFAS^{25,30,31,34-39}. Among 35 compound-protein interactions, 8 interactions have been reported in previous studies, such as the binding of perfluorooctanesulfonic acid (PFOS) to FABP1³⁶. The detected ligands showed diverse structures distributed across the chemical space of compound library (Fig. 2c). These include clusters of compounds with 21 chemical scaffolds such as multiple aromatic rings, fluorinated compounds, and those with primarily aliphatic structures (Fig. S3). Additionally, these ligands were primarily enriched in pharmaceuticals and PFAS (Fig. 2d), highlighting their potential health concerns targeting these novel protein targets.

Discovery of ligands with new chemotypes binding to human FABPs

Among 14 novel protein targets, FABPs are of particular interest, not only due to the number of ligands but also their high AS-MS enrichment fold. As a result, we decided to focus on FABPs and used an orthogonal fluorescence displacement method²⁵ to validate the bindings. Among the 26 interactions for FABPs detected by AS-MS, five compounds with background fluorescence were excluded. Among the remaining interactions, 16 were successfully validated by the fluorescence displacement method, with K_D ranging from 0.02 to 149 μ M, leading to a validation rate of 76.2%. The lack of validation for 5 interactions was probably due to two reasons. 1) Hydrophobic compounds with low water solubility might not be validated by the fluorescence assay; 2) Some chemicals might be degraded during the storage in DMSO at -80 °C, leading to overestimated concentrations for validation experiments. Thus, the validation rate of AS-MS results might be underestimated. However, the limitations of fluorescence displacement assays, which are common in traditional HTS bioassays, highlight key advantages of AS-MS: 1) the high sensitivity of MS enables the use of very low chemical concentrations, eliminating solubility issues; and 2) AS-MS is not sensitive to absolute chemical concentrations, and thus potential degradation in large chemical libraries during storage does not impact screening. These additional strengths make AS-MS well-suited for protein target discovery.

Among seven FABPs, more ligands were detected for FABP1 (15) and FABP6 (5), albeit with lower binding affinities (K_D between 0.05-9.6 μ M and 3.32-149 μ M, respectively) compared to the other FABPs (Fig. 3a). Principal component analysis (PCA) further revealed that FABP1 and FABP6 cluster closely based on their ligand interaction patterns, suggesting similarities in their binding profiles (Fig. 3b). Given these similarities, we investigated whether the accessibility of compounds to protein binding pockets plays a key role in determining ligand interactions⁴⁰. To test this, the solvent-accessible volumes of seven FABP isoforms in their ligand binding conformations

were calculated using the CASTp (Fig. 3c)⁴¹. Among the seven FABPs, the binding cavity sizes of FABP1 and FABP6 were calculated as 361 Å³ and 393 Å³, respectively, bigger than the five other FABPs (268-304 Å³). This may explain the larger number of ligands detected for FABP1 and FABP6, as their larger binding pockets can accommodate more ligands. Further PCA analysis showed that PFOS and perfluorononanoic acid (PFNA) clustered closely with FABP1, indicating a selective binding of these PFAS to FABP1 (Fig. 3b). Due to their structural similarity to fatty acids, it is expected that PFAS would be prominent among the detected ligands, binding to FABP1 in the low micromolar range^{25,30,31,34-39}. This is consistent with FABP1's primary role as a transporter of endogenous fatty acids in the liver. For example, PFOS and perfluorooctanoic acid (PFOA) were found to bind to FABP1 with K_D of 8.10 and 18.8 μM, respectively²⁵. However, the other novel ligands of FABP1 discovered by AS-MS showed very different chemotypes compared to known ligands with alkyl chains. Indeed, eight of 16 FABP1 ligands are aromatic compounds showing even stronger binding affinities than alkyl ligands, such as FK-739 (K_D = 0.56 μM) and DTXCID5027304 (K_D = 4.45 μM) (Fig. 3d).

Discovery of structure-specific binding of legacy and alternative PFAS to FABPs

Building on our findings of diverse compound classes targeting FABPs, we observed distinct binding patterns among different PFAS groups. Both legacy PFAS and perfluoroether carboxylic acids (PFECAs) were enriched by FABPs in AS-MS, yet their interactions varied. Due to the environmental persistence, bioaccumulation, and toxicity⁴²⁻⁴⁶, legacy PFAS (commonly known as “forever chemicals”) have been phased out and replaced with alternative PFAS⁴⁷⁻⁵⁰. Among these alternatives, PFECAs were designed as replacement for legacy PFAS with the similar structures⁵¹ (see representative structures in Fig. 4a). Notably, while legacy PFAS were only enriched by FABP1, PFECAs (*e.g.*, PFO3TDA) were enriched by other FABPs in addition to FABP1. Inspired

by this observation, we systematically assessed the FABP binding of six legacy and six alternative PFAS (Fig. 4a). Consistent with the AS-MS results, the fluorescence displacement assay revealed that legacy PFAS strongly bind to FABP1, while showing minimal binding to other FABPs (Fig. 4b). In contrast, long-chain PFECAs bound to all seven FABPs with greater affinities than legacy PFAS. For instance, PFO3TDA, previously identified as an FABP1 ligand²⁵, strongly bound to all seven FABPs with K_D values ranging between 0.02 and 4.35 μ M (Table S1). Similarly, HFPO-TeA and HFPO-TA bound to all seven FABP isoforms with the K_D values ranging from 0.02-3.32 μ M and 0.25-69.0 μ M, respectively (Fig. 4c and Fig. S5). In contrast, shorter-chain PFECAs showed no detectable FABP binding (Fig. S5), underscoring the importance of the hydrophobic interactions of fluorocarbon tails^{25,31}. Collectively, despite their structural similarities, legacy and alternative PFAS showed structure-specific binding patterns to FABPs.

Molecular basis for the selective binding of PFO3TDA to FABP7

To understand the mechanism underlying selective binding, we selected PFO3TDA and FABP7 for X-ray co-crystallization, as this interaction exhibited the strongest affinity among those tested in fluorescence displacement assays ($K_D = 0.02 \mu$ M) (Table S2). The co-crystal structure (1.9 Å) revealed that PFO3TDA adopts a U-shape conformation within the binding pocket (Fig. 4d). However, the precise ligand conformation was only partially resolved due to the ambiguous electron density map around the ligand, likely a result of its inherent flexibility within the binding site. Despite this limitation, the U-shape conformation is clearly distinct from that of the endogenous ligand, palmitic acid, which was previously co-crystallized with FABP7 (PDB: 7E25)⁵². Analysis of the electron density map suggests that the observed U-shaped conformation is stabilized by a conformational shift in the ether linkages from the ‘trans’ to the ‘gauche’ conformation, allowing for a better fit within the binding pocket and enhanced interactions with

surrounding hydrophobic residues. Additionally, the hydrogen bonds between the carboxylic acid group of PFO3TDA and residues R127 and Y129 further stabilizes the interaction, reinforcing its accommodation in FABP7.

To better understand the ligand's conformation within the binding pocket, given the limited resolution of the crystallization results, we assessed the energetic cost of its conformational change. Density Functional Theory (DFT) calculations were performed to evaluate the torsional energy of PFO3TDA and PFOA (Fig. 4f). The energy changes for the CF₂-X-X-CF₂ dihedral angles were comparable between the two compounds, ranging from 0.52 to 6.68 kcal/mol for PFOA and 0.71 to 4.81 kcal/mol for PFO3TDA. However, the energy changes associated with the O-CF₂-CF₂-X dihedral angles were significantly lower for PFO3TDA (0.01-0.34 kcal/mol), suggesting that the ether linkage in PFECAs enhances conformational flexibility. Collectively, the co-crystal structure and DFT calculations demonstrate that the enhanced conformational flexibility of PFECAs allows them to better adapt to the binding pockets of FABPs, contributing to their selective binding.

Discussion

We benchmarked the AS-MS platform to identify novel protein targets of prioritized chemical contaminants, with a 100% detection rate for strong-affinity ligands. Out of 24,050 assessed interactions (50 proteins × 481 compounds), AS-MS identified 35 compound-protein interactions. Among the 35 identified interactions, 27 were previously unreported, including 16 FABP-ligand interactions with high fold-change enrichment, which were successfully validated using a fluorescence displacement assay. Compared to the HTS bioassays used for the ToxCast program, our AS-MS platform has two major advantages: 1) it enables screening of a wide variety of proteins including novel proteins for which bioassays are not available, facilitating protein target scouting for environmental compounds at the proteome-wide level; 2) its throughput is much higher because

thousands of compounds can be assayed simultaneously. Identifying protein targets often unlocks exciting research opportunities, but this process can be lengthy and challenging. For example, it took approximately 40 years to identify CRBN as the target of thalidomide⁶ and 30 years to discover aryl hydrocarbon receptor (AhR) as the target of dioxins⁵³. Our AS-MS approach significantly accelerates this process by enabling scalable target discovery at both the protein and chemical levels. At the current rate, the screening of all ToxCast chemicals can be completed for 50 proteins within one week, and ~1,500 proteins within one year, when purified His-tagged proteins are available. While production of proteins is a major limiting factor, our team at Structural Genomics Consortium (SGC) are the key contributors to the Target2035 program⁵⁴ which seeks identify a chemical probe to every human protein by 2035. This ambitious program necessitates the purification of thousands of proteins, which could in turn also be used to perform ToxCast-like screens. We believe this will be a unique opportunity for target discovery for chemical contaminants approaching a proteome-wide level. However, like any technology, the AS-MS platform has its limitations: 1) it assesses protein bindings only, and additional functional experiments are needed to confirm functional consequences; 2) good MS sensitivity is essential for AS-MS screening, and thus complementary MS (*e.g.*, GC-MS) is needed to increase the coverage which might sacrifice throughput.

Of the 50 human proteins assessed beyond the scope of the ~300 proteins covered by the ToxCast program, AS-MS binding was detected for 14 of them. This highlights the need for proteome-wide discovery of novel protein targets to achieve comprehensive insights into toxicity mechanisms. Particularly, the novel interactions discovered in this study for PFAS ‘forever chemicals’ revealed unexpected chemotype and structural selectivity. PFECAs were found to bind all 7 FABPs due to their more flexible conformation, whereas legacy PFAS bound only to FABP1. FABP1 has been well documented as a key regulator of the toxicokinetics of legacy PFAS⁵⁵, which can disrupt its

280 fatty acid transport function and hinder delivery to PPARs, a pathway well known to be linked to
 281 metabolic dysregulation and obesity^{35,56,57}. Unfortunately, the broader FABP binding spectrum of
 282 PFECAs suggests an even greater toxicological impact by interrupting the FABP-PPAR pathways
 283 in multiple tissues beyond liver (Fig. 4g). Indeed, previous studies reported a stronger hepatic
 284 accumulation and toxicity of HFPO-TA compared to PFOA,^{61,62} consistent with its stronger
 285 binding affinity to liver FABP (FABP1). Collectively, this raises concerns that alternative PFAS
 286 may be regrettable substitutes for legacy compounds. Future studies are needed to evaluate their
 287 potential toxicity through the newly identified FABP pathways.

Methods

Chemical and reagents.

The 1,040 ToxCast and 74 PFAS compounds were gifts from Dr. Keith A. Houck from the U.S. Environmental Protection Agency (EPA) (Table S5). The PFECAs include perfluoro-2,5,8-trimethyl-3,6,9-trioxadodecanoic acid (HFPO-TeA), 2,3,3,3-tetrafluoro-2-(heptafluoropropoxy)propanoic acid (HFPO-DA), perfluoro-2,5-dimethyl-3,6-dioxanonanoic acid (HFPO-TA), perfluoro-3,6,9-trioxatridecanoic acid (PFO3TDA), perfluoro(4-methoxybutanoic) acid (PFMOBA), and perfluoro-3,6-dioxaheptanoic acid (PFO2HpA), in which the first three were purchased from the Toronto Research Chemicals (Toronto, ON, CA), while the others were provided by the Evotec SE (Branford, CT) under contract to the U. S. Environmental Protection Agency (EPA) through a Material Transfer Agreement. The legacy PFAS include perfluorononanoic acid (PFNA), n-ethylperfluorooctanesulfonamide (N-EtFOSA), perfluorooctanesulfonic acid (PFOS), perfluorooctanesulfonamide (PFOSA), perfluorooctanoic acid (PFOA), and perfluorohexanesulfonate (PFHxS), all of which were provided by the Evotec SE (Branford, CT). Human FABP cDNAs were purchased from Origene, including FABP1 (NM_001443), FABP2 (NM_000134), FABP3 (NM_004102), FABP4 (NM_001442), FABP5 (NM_001444), FABP6 (NM_001445), and FABP7 (NM_001446). pET28-MHL vector was kindly provided by Structural Genomics Consortium (SGC). BL21(DE3) competent *Escherichia coli* was purchased from BioLabs Inc. (Whitby, ON, CA). 1-anilinonaphthalene-8-sulfonic acid (1,8-ANS), 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid dipotassium salt (bisANS) and Ni-NTA Agarose were purchased from Thermo Fisher Scientific (Ottawa, ON, CA). Desoxycholic acid (CAS 83-44-3) is one types of bile acids purchased from Toronto Research Chemicals Inc (Toronto, ON, CA). Methanol (HPLC grade), ultrapure water (HPLC grade), acetonitrile (HPLC grade), DMSO, glycerol, and formic acid were purchased from Fisher Scientific (Ottawa, ON,

CA). HIS-Select nickel magnetic agarose beads and Triton X-100 were obtained from Sigma-Aldrich (St. Louis, MO, USA). Tris base, sodium chloride (NaCl), Tris(2-Carboxyethyl)phosphine (TCEP) were purchased from BioShop Canada Inc. (Burlington, ON, CA). Imidazole was provided by Bio Basic Canada Inc. (Markham, ON, CA).

Protein Overexpression and Purification.

Yang et al. have previously described the FABP protein overexpression process²⁵. In brief, each individual human FABP was expressed in BL21 (DE3) *E. coli* cells as TEV-cleavable 6×histidine-tagged fusion proteins. BL21-Codon Plus cells were heat-shock transformed with isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible His-tagged FABPs ligand-binding domain (LBD) constructs. The transformed *E. coli* was plated on kanamycin-antibiotic plates and incubated at 37 °C overnight. The single colony of each plate was inoculated into LB media for incubation overnight, followed by overexpression by the addition of 1 mM IPTG. The bacteria cells were lysed by sonication in the lysis buffer as described before²⁵. The FABP proteins were affinity-purified with Ni-NTA agarose according to the procedures from Cronet et al⁶³. Different from our previous study²⁵, size exclusion fast protein liquid chromatography (FPLC) (HiLoad 16/60 Superdex 200, GE Healthcare, Life Sciences) was used to further purify the protein with 150 mM NaCl, 50 mM HEPES (pH 8.2), and 0.5 mM DTT, at a flow rate of 1 ml/min. The fractions of the elution were concentrated using Amicon centrifugal filters (10K, Sigma-Aldrich) and reconstituted with the stocking buffer (50 mM tri-HCl, 150 mM NaCl, 1 mM dithiothreitol, pH 8.0). The further purification was conducted because the impurities were found to impact the K_D values of 1,8-ANS binding to FABP1 by ~5 times. Among 50 proteins, the information of 34 proteins was reported in our previous study⁹, the protein purification information of the rest of proteins was provided in the supporting information (Table S4).

AS-MS platform

The prioritizing library is composed of 1,114 compounds, including 1,040 ToxCast chemicals and 74 per- and polyfluoroalkyl substances (PFAS). All 1,114 compounds were mixed together into one pool with final concentration of 10 μ M of each compound. We selected 50 His-tagged human proteins (or domains thereof) from the SGC's collection (Table S3), using approximately 0.03mg of each protein divided into two batches on two 96-well plates. The chemical pool was exposed to all 55 proteins with 3 replicates with the binding buffer (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM TCEP, 0.5% glycerol, 0.01% Triton X-100, pH 7.5). 5 μ L of Ni-NTA magnetic beads was added to the binding buffer, then adjusted to a final volume of 250 μ L per sample. The final concentrations were set to 1 μ M for the proteins and 100 nM for each chemical. The compounds and proteins were incubated on a rotary mixer at 4 $^{\circ}$ C for 30 minutes. After incubation, the 96-well plate was placed on a magnetic plate to immobilize the beads to help to remove the binding buffer. The beads were then washed twice with 100 μ L of washing buffer 1 (150 mM NaCl, 50 mM Tris-HCl, 0.1 mM TCEP, 0.5% glycerol, 0.01% Triton X-100, 5 mM imidazole, and pH 7.5), followed by one wash with 100 μ L of washing buffer 2 (150 mM NaCl, 50 mM Tris-HCl, 5 mM imidazole, and pH 7.5). When the beads were washed with washing buffer 2, they were resuspended and transferred to a new plate. After removing the washing buffer 2, 60 μ L of methanol was added to each well to denature proteins and release chemical ligands. 50 μ L of methanol extracts was aliquoted from each well and directly subjected to LC-MS analysis.

Identification of protein hits using LC-MS.

The AS-MS samples were analyzed using an Q Exactive Orbitrap mass spectrometer equipped with a Vanquish UHPLC system (Thermo Fisher Scientific, CA, USA). Chromatographic separation was conducted on an Accucore Vanquish C18 column (50 mm 2.1 mm \times 1.5 μ m, Thermo Scientific) with a flow rate of 0.2 mL/min. 1 μ L of each sample was injected for analysis. Two sets of mobile phases were used to facilitate the compound ionization: (1) Ultrapure water

with 0.1% formic acid (A) and methanol with 0.1% formic acid (B) were used as the one set of mobile phases; and (2) Ultrapure water with 1 mM ammonium formate, pH 8 (A) and pure methanol (B) were used as the other set of mobile phases. The separation gradient elution started with 5% B for 1 minute, increased to 80% at 3 min, then reached to 100% at 4 min, held for 4 minutes, finally returned to 5% B over 1 min and equilibrated for 1 minute. The column temperature was maintained at 40 °C, and the sample compartment at 4 °C. Data acquisition was performed in full MS¹ scan mode (100 - 800 *m/z*) with a resolution of *R* = 70,000 (at *m/z* 200) and a maximum of 3×10⁶ ions collected within 160 ms, in both positive and negative ionization modes.

Fluorescence Displacement Assay.

The binding of PFAS to recombinant FABPs was determined by the displacement of 1,8-ANS from proteins as described in our previous study except for FABP6²⁵. Bis-ANS was used to determine the binding affinities of PFAS with FABP6 because of its stronger electrostatic and hydrophobic interactions with amino acid residues within the binding pocket of FABP6^{64,65}.

The interactions between the legacy PFAS and FABPs (three replicates) were determined by displacing 1,8-ANS or bisANS by the PFAS from FABPs. For 1,8-ANS fluorescence probe, 40 μM of 1,8-ANS was incubated with 1 μM of each FABP for 30 min, followed by adding a fixed concentration (100 μM) of legacy PFAS individually to the mixture (*n* = 3). After another 30 min incubation, the fluorescence signal was measured by an Infinite 200 PRO multimode plate reader (Tecan Group Ltd., Switzerland) with 350 nm excitation wavelength and emission spectra between 420 nm and 600 nm. The same assay was used to determine the binding affinities for the PFECAs. Before determining the dissociation constant (*K_D*) of PFECAs, the *K_D* of 1,8-ANS to different FABPs was required. These *K_D*'s were experimentally derived from measuring maximal fluorescence intensities of bound 1,8-ANS at various concentrations (1.875-60 μM, *n* = 3) with different FABPs. The one-site binding model based in the Scatchard equation was fit to the

displacement of protein-bound 1,8-ANS to calculate K_D to each FABP. Then, the K_D of six PFECAs was determined from the decreasing fluorescence signal induced by the displacement of 1,8-ANS from the FABP. Two concentration gradients were selected for the six PFECAs; the higher concentration gradient was selected for HFPO-TeA, HFPO-TA and HFPO-DA (0.41, 1.23, 3.70, 11.11, 33.33, and 100 μ M); the lower concentration gradient was for PFO3TDA, PFO2HpA and PFMOBA (0.37, 1.11, 3.33, 10, 30, and 50 μ M). Like the legacy PFAS, the different concentrations of PFECAs were added into 1,8-ANS (40 μ M) and FABP (1 μ M) mixture and the inhibition dose-response curves were plotted based on the relative fluorescence change. The K_D of PFAS to seven FABPs were determined through fitting to a sigmoidal model. The calculation method was as same as used in our previous study²⁵.

The process for bisANS was similar to 1,8-ANS except for the determination of K_D value of the fluorescence probe. Six concentrations of FABP6 with 3-fold serial dilution (1.23, 3.7, 11.1, 33.3, 100, and 300 μ M) were added into 1 μ M of bisANS individually to determine the fluorescence signal intensity produced when the probes bind to more proteins. The fluorescence signal was also measured by an Infinite 200 PRO multimode plate reader with 390 nm excitation wavelength and emission spectra between 420 nm and 600 nm. Regression analysis was used to calculate the K_D of bisANS⁶⁴. To determine the binding affinities of compounds to FABP6, the same concentration gradients of compounds were added to compete with 1 μ M of bisANS bound to 10 μ M of protein. The calculation steps were the same as 1,8-ANS. Principal component analysis (PCA) was conducted to analyze the FABP-ligand interaction behaviors based on the K_D values measured by the fluorescence displacement assay. The K_D of oleic acid (OA) to seven FABPs were obtained from literature^{29,66,67}. Before analysis, the K_D values were transformed to $-\log_{10}(K_D)$ to ensure that higher values indicate stronger protein-ligand interactions. In PCA biplot, the blue dots represent different ligands and the arrows of FABPs reflect their contribution to the principal components.

Calculating Protein Binding Cavity.

The volume of the binding cavity of seven FABPs was obtained by calculating solvent-accessible (SA) volume of proteins through the Computed Atlas of Surface Topography of proteins (CASTp), an online resource that computes the geometric and topological properties of proteins⁴¹. The crystal structures of proteins were downloaded from Protein Data Bank, with the resolution of X-ray diffraction less than 2.0 Å were uploaded to CASTp to measure the SA volume. The radius of the probe was set to 1.4 Å as a default. The selected protein structure IDs with the measured SA volumes are listed in SI (Table S1).

Protein Crystallization.

Perfluoro-3,6,9-trioxatridecanoic acid (PFO3TDA) was co-crystallized with FABP7 protein. A mixture of 2 mg/mL of compound and 9.7 mg/mL of TEV-cleavable 6×histidine-tagged FABP7 was prepared in a crystallization buffer containing 12.5% (w/v) PEG 1000, 12.5% (w/v) PEG 3350, 12.5% (v/v) MPD, 0.03 M of each halide, 0.1 M MES/imidazole and pH 6.5. The protein-ligand mixture was mixed with the crystallization buffer in a 1:1 ratio, with 200 nL of each component, and set up in a sitting-drop crystallization plate at 20 °C. The crystal was harvested and flash-cooled in liquid nitrogen without additional cryoprotectant. Diffraction data were collected using a Rigaku FR-E SuperBright source. The data were processed and scaled using HKL-3000, and the structure was solved by molecular replacement in PHASER using the starting coordinates 6L9O. Refinement was performed using BUSTER (v2.10.3). The final coordinates of the protein and ligand were deposited in the PDB under accession code 9NIU.

DFT Calculations.

The changes of torsional energies with bond angles of legacy PFAS (PFOA) and PFECAs (PFO3TdA) were calculated using the GAUSSIAN 16 software. The molecular geometries of PFO3TDA and PFOA were optimized using the B3LYP functionals and 6-311G basis set. No

solvent correction was applied to best emulate the solvent-free conditions of the hydrophobic binding pockets of FABP proteins. The geometry-optimized structures were used to perform bond angle scans of CF₂-O-CF₂ in PFO₃TDA and CF₂-CF₂-CF₂ in PFOA with a step size of 15°. Similarly, dihedral angle scans were performed for CF₂-CF₂-CF₂-CF₂ in PFO₃TDA and PFOA, as well as O-CF₂-CF₂-X in PFO₃TDA, with a step change of 10° (Fig. S6, Fig. S7).

Code availability

The source code and demo datasets can be found via GitHub at <https://github.com/huiUofT/EAS-MS>.

Acknowledgements

We extend our deepest gratitude to the Structural Genomics Consortium-Toronto (SGC-Toronto) and Structural Genomics Consortium-Frankfurt (SGC-Frankfurt) for their generous contribution of proteins, which were invaluable to this study. From SGC-Toronto, we acknowledge the provision of GID4, SLC9C1 and C2, DNMT3A, SKP1, SKP1FBXO22, MSH6, USP37, SETDB1, HTT/HAP40, WDR55, SEC31A, DCAF7, HDAC6, CRBN, TLE4, DDB1, COPB2, WDR5, FBXW7, RFWD3, DCAF1, WDR91, LRRK2, HAT1, PIAS1, DDX1, HUWE1WWE, YTHDF2, YTHDC1, CHIP, and USP21. From SGC-Frankfurt, we acknowledge the provision of 1TIF1A, 2TIF1A, TRIM28, TRIM7, TRIM21, BIR1, BUCK, and SPSB2. We also appreciate Wyatt Yue from Newcastle University for providing AASSLKR and AASSSDH proteins. We are especially grateful to Elisa Gibson, Pegah Ghiabi, Carla Kharadjian, Serah Kimani, Renu Chandrasekaran, Taraneh Hajian, Andreas C. Joerger, Krishna Saxena, Yeojin Kim, and Madhushika Silva for their dedicated efforts in protein preparation. Additionally, we sincerely acknowledge Albina Bolotokova, Vijayaratnam Santhakumar, and Suzanne Ackloo for their

invaluable contributions in preparing benchmarking chemicals and coordinating the delivery of chemicals and proteins. We also extend our gratitude to Keith A. Houck from the U.S. Environmental Protection Agency (EPA) for providing ToxCast and PFAS compound data. Furthermore, we would like to thank Dafydd Owen (Pfizer), Judith Guenther (Bayer), Scott Johnson (Bristol Myers Squibb), Oliver Kraemer (Boehringer Ingelheim), Julian Schmid (Merck), Juncai Meng (Janssen), Lawrence Szewczuk (Janssen), Xidong Feng (Pfizer), Wenyi Hua (Pfizer), Anja Giese (Bayer), Matteo Aldeghi (Bayer), Nidhi Arora (Takeda), James Kiefer (Genentech), Ingo Hartung (Merck), and Matthew Troutman (Pfizer) for their valuable advice and discussions on method development. The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funding from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, the EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking ([EUbOPEN grant 875510]), Janssen, Pfizer, and Takeda. This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, an Ontario Early Researcher Award, the Canadian Institutes of Health Research (CIHR), and the National Institutes of Health (NIH) (U54AG065187). The authors also acknowledge the support of instrumentation grants from the Canada Foundation for Innovation, the Ontario Research Fund, and an NSERC Research Tools and Instruments Grant. AME holds the Temerty Nexus Chair of Health Technology and Innovation at the University of Toronto.

Author contributions

DY, XW, JL, JS, YG: AS-MS method development

DY, XW, XQ, CC: Chemical library screening

RJH, NB, TSB, HZ, DLB, NGB: Proteins preparation and coordination

PN, DS: DFT calculation

480 HZ, AD: Protein Crystallization

481 HP, LH, HK, MLD, AME, CHA: Supervision, review and editing

482 DY, XW, HP: Writing, reviewing and editing with input from all

483

484 **Competing interests**

485 The authors declare no competing interests.

486

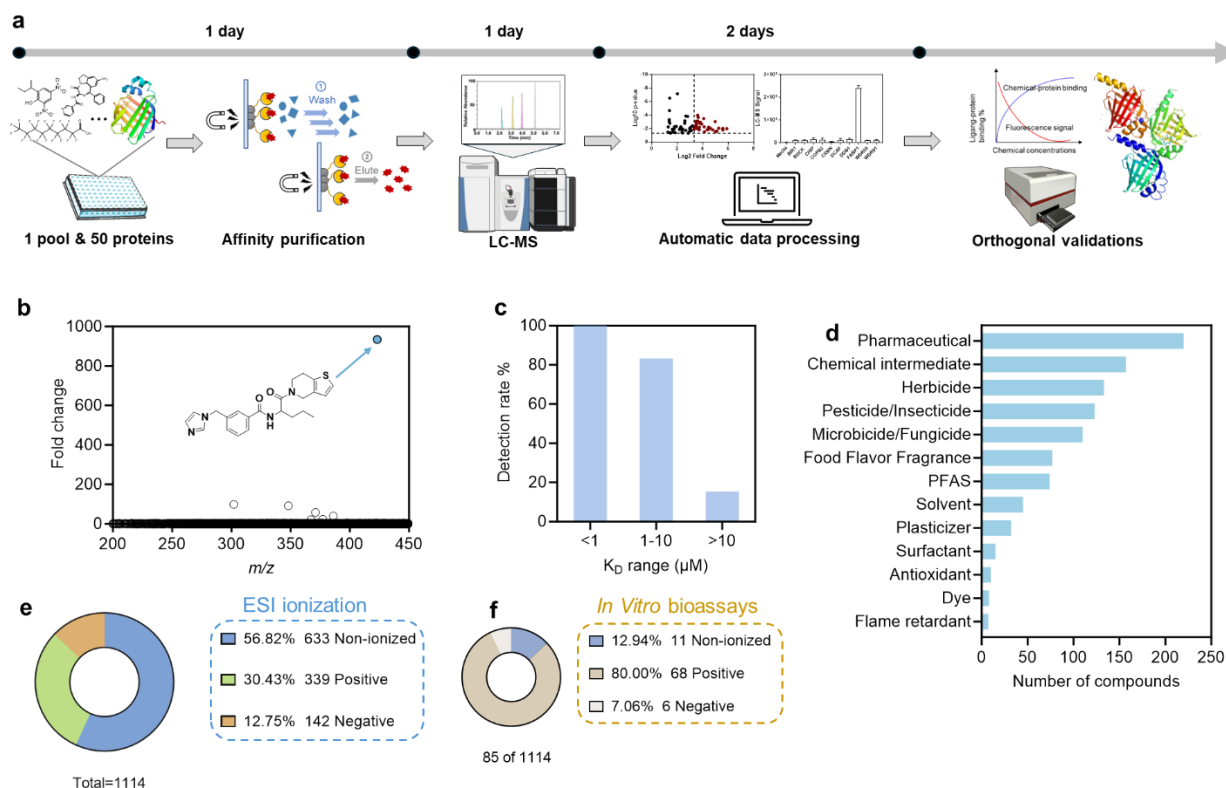


Fig. 1 | Benchmarking the AS-MS platform for scalable protein target discovery.

(a) Workflow of AS-MS platform. (b) Benchmarking the AS-MS platform using a double-blind experiment. The AS-MS platform successfully identified a known ligand (blue dot, $K_D = 60$ nM) of WD repeat-containing protein 5 (WDR5) from a library of ~16,000 compounds. (c) Detection rate of compounds according to their dissociation constants (K_D). (d) Top 10 compound categories in the 1,114-chemical library based on the Chemical Products Categories (CPCat) database^{68,69}. (e) The fraction of MS activity compounds in the 1,114-chemical library. (f) A subset of 85 bioactive compounds from the library exhibited $AC_{50} < 10$ μ M against nuclear receptors.

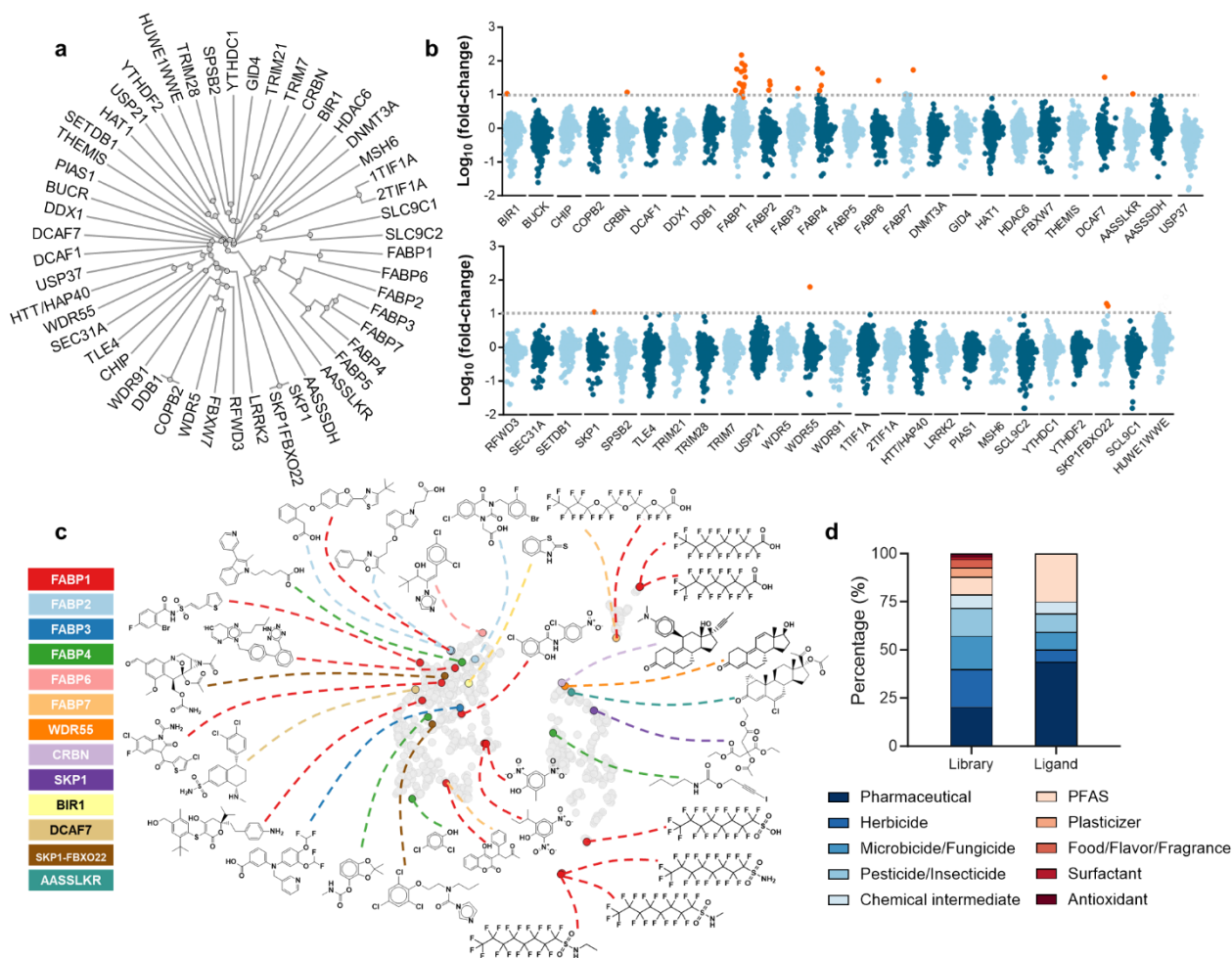


Fig. 2 | Mapping interactions between 481 selected chemicals and 50 human proteins.

(a) Overview of 50 selected human proteins. **(b)** Scatter plot of hits interacting with 50 proteins.

The gray line represents a fold change cutoff of 10. Red dots indicate hits with a fold change greater than 10 and a *p-value* less than 0.05. The fold change for each compound-protein interaction was

calculated by dividing the average LC-MS peak intensity of the target protein by the average signal across all other proteins on the plate and *p-values* were determined using Student's t-tests. **(c)**

Chemical space of the 1,114 compounds clustered based on their molecular fingerprints and visualized using uniform manifold approximation and projection (UMAP). Detected ligands are

highlighted in different colors according to their target proteins, with chemical structures labeled on the plot. **(d)** Comparison of the top 10 compound categories in the 481 selected compounds

508 (left column) with the categories of ligands that showed protein interactions in the AS-MS
509 screening (right column).

510

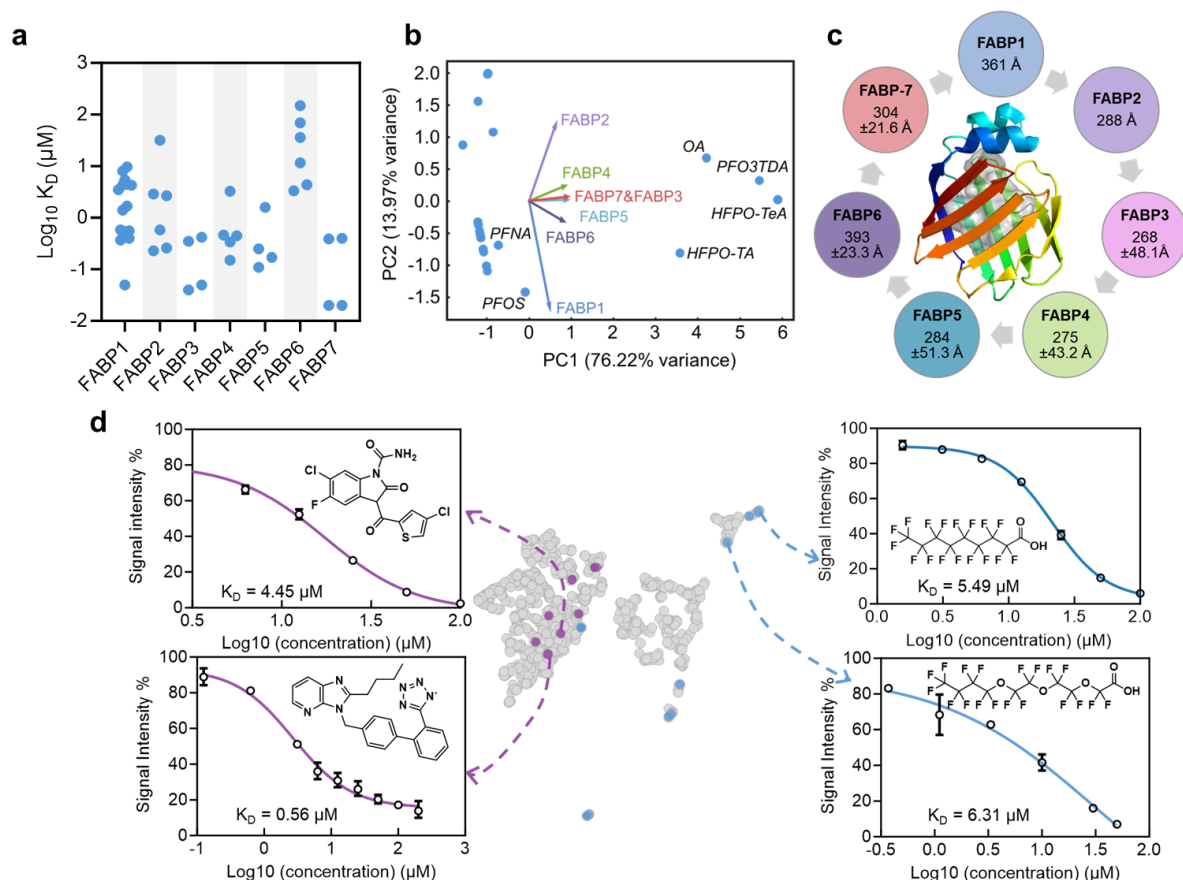


Fig. 3 | Distinct and novel interactions between ToxCast/PFAS chemicals and FABPs.

(a) Overview of the binding affinity range ($\text{Log}_{10} K_D$, μM) for ligands interacting with seven FABPs measured by fluorescence displacement assay. Each blue dot represents a single ligand. (b) Principal component analysis (PCA) of ligand binding affinities ($\text{Log}_{10} K_D$, μM) across seven FABPs. Fluorescence displacement assays were used to determine binding affinities, with 1,8-ANS as the probe for FABP1, 2, 3, 4, 5, and 7, and bisANS for FABP6, while oleic acid (OA) K_D values for the seven FABPs were obtained from the literature^{29,66,67}. (c) Solvent-accessible volumes (light gray) of FABPs. FABP1 (PDB: 3STM) is shown as a representative structure for the seven FABPs. For FABP isoforms with multiple available crystal structures, mean and standard deviation values were calculated to determine solvent-accessible volume variations. (d) Distribution of known ligands (blue dots) and newly discovered ligands (purple dots) for FABP1

within the chemical space of the 1,040 ToxCast compounds and 74 PFAS compounds. Dose-response results from the fluorescence displacement assay validation are shown, along with the measured dissociation constants (K_D). DTXCID5027304 (CAS, 135080-03-4; top left panel) was tested with six concentrations (3.12, 6.25, 12.5, 25, 50 100 μ M); FK-739 (CAS, 136042-19-8; bottom left panel) was tested with nine concentrations (0.12, 0.62, 3.12, 6.24, 12.5, 25, 50, 100, 200 μ M); PFNA (CAS, 375-95-1; top right panel) was tested with seven concentrations (1.78, 3.12, 6.25, 12.5, 25, 50 100 μ M); PFO3TDA (CAS, 330562-41-9; bottom right panel) was tested with six concentrations (0.37, 1.11, 3.33, 10, 30, 50 μ M).

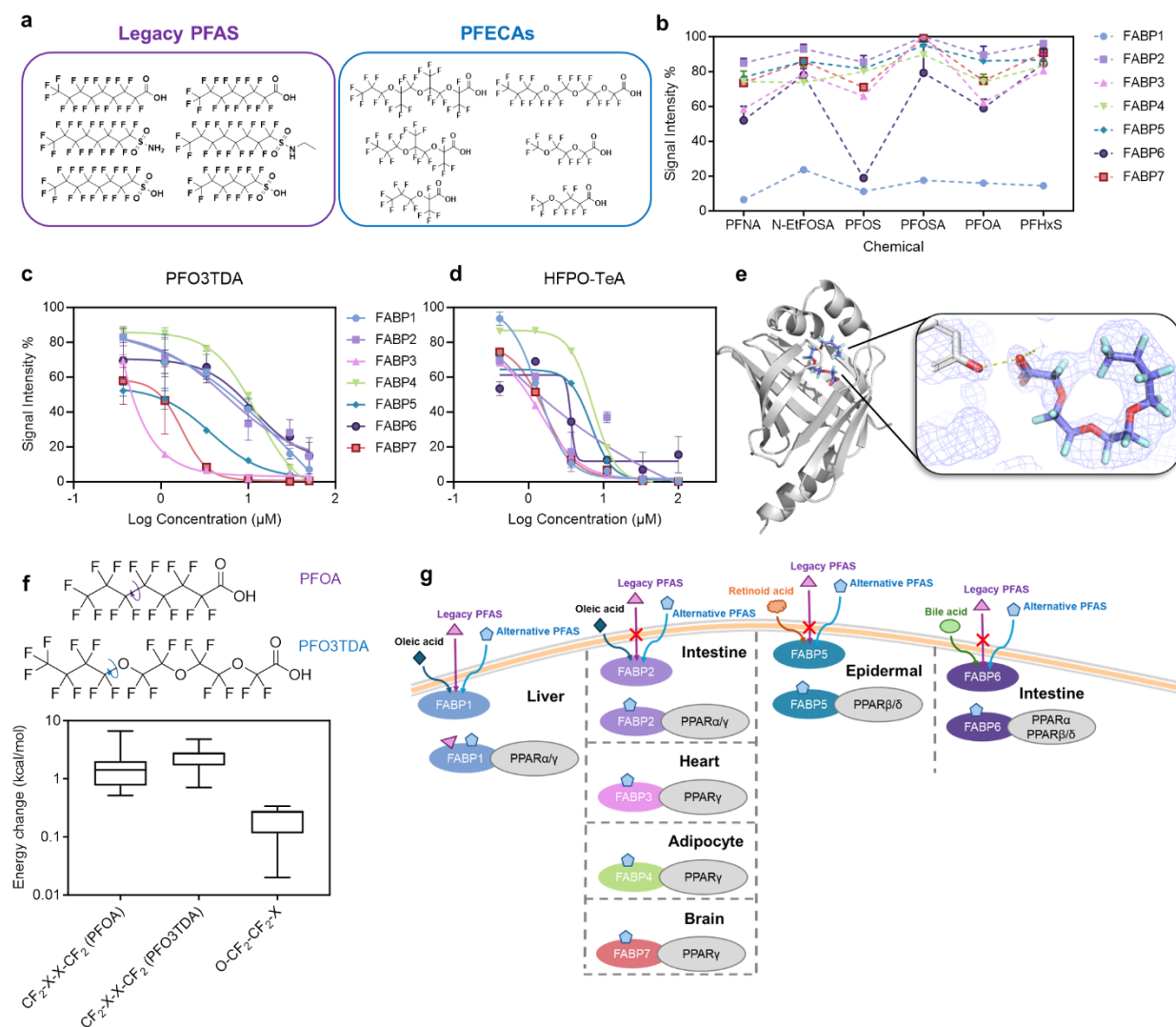


Fig. 4 | Selective interactions between alternative PFAS and FABPs.

(a) Chemical structures of representative legacy PFAS and PFECAs (alternative PFAS). (b) The relative fluorescence signals were measured when different legacy PFAS (100 μ M) were added to compete with the FABP bound fluorescence probes; 1,8-ANS were used for FABP 1, 2, 3, 4, 5, and FABP7, and bisANS was used for FABP6. Dose-response curves showing the competition between increasing compound concentrations and fluorescent probes for binding to the seven FABPs. The three compounds are displayed as follows: PFO3TDA(c) and HFPO-TeA(d). PFO3TDA (CAS, 330562-41-9) was tested with six concentrations (0.37, 1.11, 3.33, 10, 30, 50 μ M) across seven FABPs; HFPO-TeA (CAS, 65294-16-8) was tested with six concentrations

542 (3.12, 6.25, 12.5, 25, 50 100 μ M) across seven FABPs. (e) PFO3TDA co-crystalized with FABP7
 543 (PDB: 9NIU). The carboxylate head group of PFO3TDA formed two hydrogen bonds with
 544 residues arginine 127 and tyrosine 129. (f) The Gibbs free energy difference between ‘trans’ and
 545 ‘gauche’ conformations for PFOA, PFO3TDA (CF₂-X-X-CF₂), and PFO3TDA (O-CF₂-CF₂-X).
 546 (g) The diagram showing the interactions of endogenous ligands²⁹ (oleic acid, retinoid acid and
 547 bile acid), legacy PFAS, and alternative PFAS (long-chain PFECAs) with FABPs and their
 548 corresponding peroxisome proliferator-activated receptors (PPARs), distributed across different
 549 body tissues^{29,70–75}.

References

1. Wang, Z., Walker, G. W., Muir, D. C. G. & Nagatani-Yoshida, K. Toward a Global Understanding of Chemical Pollution: A First Comprehensive Analysis of National and Regional Chemical Inventories. *Environ. Sci. Technol.* **54**, 2575–2584 (2020).
2. Landrigan, P. J. *et al.* The Lancet Commission on pollution and health. *The Lancet* **391**, 462–512 (2018).
3. Krewski, D. *et al.* Toxicity Testing in the 21st Century: A Vision and a Strategy. *J. Toxicol. Environ. Health Part B* **13**, 51–138 (2010).
4. Blum, A. *et al.* Organophosphate Ester Flame Retardants: Are They a Regrettable Substitution for Polybrominated Diphenyl Ethers? *Environ. Sci. Technol. Lett.* **6**, 638–649 (2019).
5. Folkerson, A. P., Schneider, S. R., Abbatt, J. P. D. & Mabury, S. A. Avoiding Regrettable Replacements: Can the Introduction of Novel Functional Groups Move PFAS from Recalcitrant to Reactive? *Environ. Sci. Technol.* **57**, 17032–17041 (2023).
6. Ito, T. *et al.* Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **327**, 1345–1350 (2010).
7. Escher, B. I., Stapleton, H. M. & Schymanski, E. L. Tracking complex mixtures of chemicals in our changing environment. *Science* **367**, 388–392 (2020).
8. US EPA. *EPA's Toxicity Forecaster (Toxcast) Website*. <https://www.epa.gov/chemical-research/toxicity-forecasting>.
9. Wang, X. *et al.* Enantioselective Protein Affinity Selection Mass Spectrometry (EAS-MS). 2025.01.17.633682 Preprint at <https://doi.org/10.1101/2025.01.17.633682> (2025).
10. Yang, D. *et al.* Widespread formation of toxic nitrated bisphenols indoors by heterogeneous reactions with HONO. *Sci. Adv.* **8**, (2022).

11. Gong, Y. *et al.* Disclosing Environmental Ligands of L-FABP and PPAR γ : Should We Re-evaluate the Chemical Safety of Hydrocarbon Surfactants? *Environ. Sci. Technol.* **57**, 11913–11925 (2023).
12. Gong, Y., Yang, D., Barrett, H., Sun, J. & Peng, H. Building the Environmental Chemical-Protein Interaction Network (eCPIN): An Exposome-Wide Strategy for Bioactive Chemical Contaminant Identification. *Environ. Sci. Technol.* (2023) doi:10.1021/acs.est.2c02751.
13. Peng, H. *et al.* Ligand interaction landscape of transcription factors and essential enzymes in *E. coli*. *Cell* S0092867425000327 (2025) doi:10.1016/j.cell.2025.01.003.
14. Liu, J. *et al.* Diindoles produced from commensal microbiota metabolites function as endogenous CAR/Nr1i3 ligands. *Nat. Commun.* **15**, 2563 (2024).
15. Sipes, N. S. *et al.* Profiling 976 ToxCast Chemicals across 331 Enzymatic and Receptor Signaling Assays. *Chem. Res. Toxicol.* **26**, 878–895 (2013).
16. Patlewicz, G. *et al.* A Chemical Category-Based Prioritization Approach for Selecting Per- and Polyfluoroalkyl Substances (PFAS) for Tiered Toxicity and Toxicokinetic Testing. *Environ. Health Perspect.* **127**, 014501 (2019).
17. Giani Tagliabue, S., Faber, S. C., Motta, S., Denison, M. S. & Bonati, L. Modeling the binding of diverse ligands within the Ah receptor ligand binding domain. *Sci. Rep.* **9**, 10693 (2019).
18. Kavlock, R. *et al.* Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol.* **25**, 1287–1302 (2012).
19. Dix, D. J. *et al.* The ToxCast Program for Prioritizing Toxicity Testing of Environmental Chemicals. *Toxicol. Sci.* **95**, 5–12 (2007).
20. Cousins, I. T. *et al.* The High Persistence of PFAS is Sufficient for their Management as a Chemical Class. *Environ. Sci. Process. Impacts* **22**, 2307–2312 (2020).

21. Fenton, S. E. *et al.* Per- and Polyfluoroalkyl Substance Toxicity and Human Health Review: Current State of Knowledge and Strategies for Informing Future Research. *Environ. Toxicol. Chem.* **40**, 606–630 (2021).
22. Brunn, H. *et al.* PFAS: forever chemicals—persistent, bioaccumulative and mobile. Reviewing the status and the need for their phase out and remediation of contaminated sites. *Environ. Sci. Eur.* **35**, 20 (2023).
23. Han, J. *et al.* A Roadmap to the Structure-Related Metabolism Pathways of Per- and Polyfluoroalkyl Substances in the Early Life Stages of Zebrafish (*Danio rerio*). *Environ. Health Perspect.* **129**, 77004 (2021).
24. Martin, M. T. *et al.* Impact of Environmental Chemicals on Key Transcription Regulators and Correlation to Toxicity End Points within EPA’s ToxCast Program. *Chem. Res. Toxicol.* **23**, 578–590 (2010).
25. Yang, D. *et al.* Nontarget Screening of Per- and Polyfluoroalkyl Substances Binding to Human Liver Fatty Acid Binding Protein. *Environ. Sci. Technol.* **54**, 5676–5686 (2020).
26. Bieber, S., Letzel, T. & Krue, A. Electrospray Ionization Efficiency Predictions and Analytical Standard Free Quantification for SFC/ESI/HRMS. *J. Am. Soc. Mass Spectrom.* **34**, 1511–1518 (2023).
27. Arrowsmith, C. H., Bountra, C., Fish, P. V., Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* **11**, 384–400 (2012).
28. Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nat. Rev. Drug Discov.* **18**, 949–963 (2019).
29. Smathers, R. L. & Petersen, D. R. The human fatty acid-binding protein family: Evolutionary divergences and functions. *Hum. Genomics* **5**, 170–191 (2011).

30. Zhang, L., Ren, X.-M. & Guo, L.-H. Structure-Based Investigation on the Interaction of Perfluorinated Compounds with Human Liver Fatty Acid Binding Protein. *Environ. Sci. Technol.* **47**, 11293–11301 (2013).
31. Sheng, N., Li, J., Liu, H., Zhang, A. & Dai, J. Interaction of perfluoroalkyl acids with human liver fatty acid-binding protein. *Arch. Toxicol.* **90**, 217–227 (2016).
32. Greggio, E. *et al.* The Parkinson Disease-associated Leucine-rich Repeat Kinase 2 (LRRK2) Is a Dimer That Undergoes Intramolecular Autophosphorylation*. *J. Biol. Chem.* **283**, 16906–16914 (2008).
33. Taymans, J.-M. *et al.* Perspective on the current state of the LRRK2 field. *Npj Park. Dis.* **9**, 1–9 (2023).
34. Wang, J. *et al.* RNA-sequencing analysis reveals the hepatotoxic mechanism of perfluoroalkyl alternatives, HFPO2 and HFPO4, following exposure in mice. *J. Appl. Toxicol.* **37**, 436–444 (2017).
35. Zhang, L., Ren, X.-M., Wan, B. & Guo, L.-H. Structure-dependent binding and activation of perfluorinated compounds on human peroxisome proliferator-activated receptor γ . *Toxicol. Appl. Pharmacol.* **279**, 275–283 (2014).
36. Luebker, D. J., Hansen, K. J., Bass, N. M., Butenhoff, J. L. & Seacat, A. M. Interactions of fluorochemicals with rat liver fatty acid-binding protein. *Toxicology* **176**, 175–185 (2002).
37. Woodcroft, M. W. *et al.* Experimental characterization of the mechanism of perfluorocarboxylic acids' liver protein bioaccumulation: The key role of the neutral species. *Environ. Toxicol. Chem.* **29**, 1669–1677 (2010).
38. Cao, H. *et al.* Screening of Potential PFOS Alternatives to Decrease Liver Bioaccumulation: Experimental and Computational Approaches. *Environ. Sci. Technol.* **53**, 2811–2819 (2019).

39. Han, J. *et al.* Quantitative Chemical Proteomics Reveals Interspecies Variations on Binding Schemes of L-FABP with Perfluorooctanesulfonate. *Environ. Sci. Technol.* **55**, 9012–9023 (2021).
40. Stank, A., Kokh, D. B., Fuller, J. C. & Wade, R. C. Protein Binding Pocket Dynamics. *Acc. Chem. Res.* **49**, 809–815 (2016).
41. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).
42. Moody, C. A. & Field, J. A. Determination of Perfluorocarboxylates in Groundwater Impacted by Fire-Fighting Activity. *Environ. Sci. Technol.* **33**, 2800–2806 (1999).
43. Nascimento, R. A. *et al.* Sulfluramid use in Brazilian agriculture: A source of per- and polyfluoroalkyl substances (PFASs) to the environment. *Environ. Pollut. Barking Essex 1987* **242**, 1436–1443 (2018).
44. Hill, P. J., Taylor, M., Goswami, P. & Blackburn, R. S. Substitution of PFAS chemistry in outdoor apparel and the impact on repellency performance. *Chemosphere* **181**, 500–507 (2017).
45. Begley, T. H. *et al.* Perfluorochemicals: Potential sources of and migration from food packaging. *Food Addit. Contam.* **22**, 1023–1031 (2005).
46. Giesy, John P; Kannan, K. Peer Reviewed: Perfluorochemical Surfactants in the Environment. *Environ. Sci. Technol.* **36**, 146A-152A (2002).
47. Giesy, J. P. & Kannan, K. Global Distribution of Perfluorooctane Sulfonate in Wildlife. *Environ. Sci. Technol.* **35**, 1339–1342 (2001).
48. Vestergren, R. & Cousins, I. T. Tracking the Pathways of Human Exposure to Perfluorocarboxylates. *Environ. Sci. Technol.* **43**, 5565–5575 (2009).

- 667 49. Wang, S. *et al.* First Report of a Chinese PFOS Alternative Overlooked for 30 Years: Its
668 Toxicity, Persistence, and Presence in the Environment. *Environ. Sci. Technol.* **47**, 10163–
669 10170 (2013).
- 670 50. D’Agostino, L. A. & Mabury, S. A. Identification of Novel Fluorinated Surfactants in
671 Aqueous Film Forming Foams and Commercial Surfactant Concentrates. *Environ. Sci.*
672 *Technol.* **48**, 121–129 (2014).
- 673 51. Pan, Y. *et al.* Worldwide Distribution of Novel Perfluoroether Carboxylic and Sulfonic
674 Acids in Surface Water. *Environ. Sci. Technol.* **52**, 7621–7629 (2018).
- 675 52. Nam, K. H. Crystal structure of human brain-type fatty acid-binding protein FABP7
676 complexed with palmitic acid. *Acta Crystallogr. Sect. Struct. Biol.* **77**, 954–965 (2021).
- 677 53. Okey, A. B. An Aryl Hydrocarbon Receptor Odyssey to the Shores of Toxicology: The
678 Deichmann Lecture, International Congress of Toxicology-XI. *Toxicol. Sci.* **98**, 5–38 (2007).
- 679 54. Müller, S. *et al.* Target 2035 – update on the quest for a probe for every protein. *RSC Med.*
680 *Chem.* **13**, 13–21 (2022).
- 681 55. Ng, C. A. & Hungerbühler, K. Bioconcentration of Perfluorinated Alkyl Acids: How
682 Important Is Specific Binding? *Environ. Sci. Technol.* **47**, 7214–7223 (2013).
- 683 56. Ishibashi, H., Hirano, M., Kim, E.-Y. & Iwata, H. In Vitro and In Silico Evaluations of
684 Binding Affinities of Perfluoroalkyl Substances to Baikal Seal and Human Peroxisome
685 Proliferator-Activated Receptor α . *Environ. Sci. Technol.* **53**, 2181–2188 (2019).
- 686 57. Qi, W., Clark, J. M., Timme-Laragy, A. R. & Park, Y. Per- and Polyfluoroalkyl Substances
687 and Obesity, Type 2 Diabetes and Non-alcoholic Fatty Liver Disease: A Review of
688 Epidemiologic Findings. *Toxicol. Environ. Chem.* **102**, 1–36 (2020).

58. Petrescu, A. D. *et al.* Impact of L-FABP and glucose on polyunsaturated fatty acid induction of PPAR α -regulated β -oxidative enzymes. *Am. J. Physiol.-Gastrointest. Liver Physiol.* **304**, G241–G256 (2013).
59. Terra, X. *et al.* FABP 4 is associated with inflammatory markers and metabolic syndrome in morbidly obese women. *Eur. J. Endocrinol.* **164**, 539–547 (2011).
60. Liu, R.-Z., Mita, R., Beaulieu, M., Gao, Z. & Godbout, R. Fatty acid binding proteins in brain development and disease. *Int. J. Dev. Biol.* **54**, 1229–1239 (2010).
61. Sheng, N., Pan, Y., Guo, Y., Sun, Y. & Dai, J. Hepatotoxic Effects of Hexafluoropropylene Oxide Trimer Acid (HFPO-TA), A Novel Perfluorooctanoic Acid (PFOA) Alternative, on Mice. *Environ. Sci. Technol.* **52**, 8005–8015 (2018).
62. Robarts, D. R., Dai, J., Lau, C., Apte, U. & Corton, J. C. Hepatic Transcriptome Comparative In Silico Analysis Reveals Similar Pathways and Targets Altered by Legacy and Alternative Per- and Polyfluoroalkyl Substances in Mice. *Toxics* **11**, 963 (2023).
63. Cronet, P. *et al.* Structure of the PPAR α and - γ ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family. *Struct. Lond. Engl.* **1993** **9**, 699–706 (2001).
64. Pastukhov, A. V. & Ropson, I. J. Fluorescent Dyes as Probes to Study Lipid-Binding Proteins. *Proteins Struct. Funct. Genet.* **53**, 607–615 (2003).
65. Schönbrunn, E., Eschenburg, S., Luger, K., Kabsch, W. & Amrhein, N. Structural basis for the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) with the antibiotic target MurA. *Proc. Natl. Acad. Sci.* **97**, 6345 LP – 6349 (2000).
66. Hughes, M. L. R. *et al.* Fatty Acid-binding Proteins 1 and 2 Differentially Modulate the Activation of Peroxisome Proliferator-activated Receptor α in a Ligand-selective Manner *. *J. Biol. Chem.* **290**, 13895–13906 (2015).

67. Siegenthaler, G. *et al.* Purification and characterization of the human epidermal fatty acid-binding protein: localization during epidermal cell differentiation in vivo and in vitro. *Biochem. J.* **302**, 363–371 (1994).
68. Dionisio, K. L. *et al.* Exploring consumer exposure pathways and patterns of use for chemicals in the environment. *Toxicol. Rep.* **2**, 228–237 (2015).
69. Wang, J. *et al.* High-throughput screening and chemotype-enrichment analysis of ToxCast phase II chemicals evaluated for human sodium-iodide symporter (NIS) inhibition. *Environ. Int.* **126**, 377–386 (2019).
70. Wolfrum, C., Borrmann, C. M., Börschers, T. & Spener, F. Fatty acids and hypolipidemic drugs regulate peroxisome proliferator-activated receptors α - and γ -mediated gene expression via liver fatty acid binding protein: A signaling path to the nucleus. *Proc. Natl. Acad. Sci.* **98**, 2323–2328 (2001).
71. Itoh, T. *et al.* Structural basis for the activation of PPAR γ by oxidized fatty acids. *Nat. Struct. Mol. Biol.* **15**, 924–931 (2008).
72. Gillilan, R. E., Ayers, S. D. & Noy, N. Structural basis for activation of fatty acid-binding protein 4. *J. Mol. Biol.* **372**, 1246–1260 (2007).
73. Wolf, G. Retinoic acid activation of peroxisome proliferation-activated receptor delta represses obesity and insulin resistance. *Nutr. Rev.* **68**, 67–70 (2010).
74. Kurz, M. *et al.* Insights into the bile acid transportation system: the human ileal lipid-binding protein-cholyltaurine complex and its comparison with homologous structures. *Proteins* **50**, 312–328 (2003).
75. Owada, Y. *et al.* Altered emotional behavioral responses in mice lacking brain-type fatty acid-binding protein gene. *Eur. J. Neurosci.* **24**, 175–187 (2006).