Proteome-wide reverse molecular docking reveals folic acid receptor as a mediator of PFAS-induced neurodevelopmental toxicity

Ally Xinyi Kong[1*], Maja Johnson[1*], Aiden F Eno[1], Khoa Pham[1], Ping Zhang[1], Yijie Geng[1#]

[1]Department of Environmental and Occupational Health Sciences, Seattle, WA 98105, USA.

*These authors contributed equally

[#]Corresponding author. Email: yjgeng@uw.edu

ABSTRACT

Per- and polyfluoroalkyl substances (PFAS) are a class of long-lasting chemicals with widespread use and environmental persistence that have been increasingly studied for their detrimental impacts to human and animal health. Several major PFAS species are linked to neurodevelopmental toxicity. For example, epidemiological studies have associated prenatal exposure to perfluorooctanoate (PFOA) and perfluorononanoate (PFNA) with autism risk. However, the neurodevelopmental toxicities of major PFAS species have not been systematically evaluated in an animal model, and the molecular mechanisms underlying these toxicities have remained elusive. Using a high-throughput zebrafish social behavioral model, we screened six major PFAS species currently under regulation by the Environmental Protection Agency (EPA), including PFOA, PFNA, perfluorooctane sulfonate (PFOS), perfluorohexanesulfonic acid (PFHxS), perfluorobutane sulfonate (PFBS), and hexafluoropropylene oxide dimer acid ammonium salt (GenX). We found that embryonic exposure to PFNA, PFOA, and PFOS induced social deficits in zebrafish, recapitulating one of the hallmark behavioral deficits in autistic individuals. To uncover protein targets of the six EPA-regulated PFAS, we screened a virtual library containing predicted binding pockets of over 80% of the 3D human proteome through reverse molecular docking. We found that folate receptor beta (FR-β, encoded by the gene *FOLR2*) interacts strongly with PFNA, PFOA, and PFOS but to a lesser degree with PFHxS, PFBS, and GenX, correlating positively with their *in vivo* toxicity. Embryonic co-exposure to folic acid rescued social deficits induced by PFAS. The folic acid pathway has been implicated in autism, indicating a novel molecular mechanism for PFAS in autism etiology.

INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition marked by social communication deficits. Over recent decades, the prevalence of autism has risen significantly, with reported cases in the United States increasing from 1 in 150 children diagnosed in 2000 to 1 in 36 in 2020, according to the Centers for Disease Control and Prevention (CDC). This increase is only partially accounted for by improved diagnostic criteria and awareness[1,2]. Meanwhile, the ever-growing chemical industry continues to introduce new chemicals into our environment, most of which with unknown toxicity to humans. An estimated 10 million new chemicals are synthesized annually, and within them, ~ 700 become mass produced for commercial use[3]. Emerging research indicates that environmental factors contribute substantially to autism risk, with recent estimates attributing up to 40% of ASD etiology to environmental influences[4,5]. Understanding how environmental factors contribute to ASD is critical to mitigating the disease prevalence and could inform targeted prevention and treatment strategies.

However, environmental contributors to ASD remain poorly understood, creating a gap in research methods and assessment approaches to identify and evaluate these risk factors comprehensively. Among environmental toxins with known associations with ASD, per- and polyfluoroalkyl substances (PFAS) have become a particular focus of recent research[6,7]. PFAS is a class of synthetic chemicals known for their persistence in the environment and biological systems. They are widely present in nonstick cookware, stain-resistant textiles, and various industrial applications. Their pervasive use has led to widespread human exposure, with the CDC reporting PFAS detected in the blood of 97% of Americans. The health risks associated with PFAS exposure are substantial, including links to kidney and

liver disease, immune and thyroid dysfunction, cancer, and, critically, developmental challenges[8]. Based on these pioneering studies, beginning in April 2024, the U.S. Environmental Protection Agency (EPA) established legally enforceable levels for six PFAS in drinking water, including perfluorooctanoic acid (PFOA), perfluorooctane sulfonic acid (PFOS), perfluorononanoic acid (PFNA), perfluorohexanesulfonic acid (PFHxS), perfluorobutane sulfonate (PFBS), and hexafluoropropylene oxide dimer acid ammonium salt (GenX). Despite these progresses, however, with over 7 million PFAS species reported in PubChem to date[9], a comprehensive understanding of their biological impact remains elusive, particularly regarding how prenatal exposures to various PFAS species contribute to ASD risk. For example, although epidemiology studies linked PFNA and PFOA to increasing ASD risk[6,7], whether the other four EPA-regulated PFAS contribute equally to ASD remains unclear. Mechanisms underlying their neurodevelopmental toxicity are also largely unknown.

To address these gaps, we developed a novel model and methodology using zebrafish. Zebrafish offer a unique advantage as a model organism for studying social behavior—a core behavioral deficit in ASD. They share approximately 70% of their genome with humans, begin social interactions at an early developmental stage, and are amenable to high-throughput chemical screening and behavioral assays. In our previous work, we designed the Fishbook social behavioral assay[10] to quantitatively assess social behavioral deficits in zebrafish, a behavior analogous to social deficits in ASD. In this assay, a test fish is exposed to a social stimulus in a controlled environment, allowing us to measure a social score based on proximity behaviors. Notably, the Fishbook assay has observed social deficits in zebrafish following ASD-related gene knockouts and exposure to ASD-related environmental toxicants[10,11], indicating its relevance for examining environmental toxicants' influence on ASD. The Fishbook system enables us to screen different PFAS compounds in parallel using the same assay platform for their impacts on sociality, alongside genetic knockout models to elucidate their molecular mechanisms.

Moreover, to uncover potential molecular targets of PFAS, we conducted virtual screening through reverse molecular docking using the HPRoteome-BSite database[12] to explore interactions between PFAS and the predictable 3D human proteome. HProteome-BSite is a database of predicted binding pockets covering over 80% of the UniProt entries in the AlphaFold human proteome database[13,14]. The reverse molecular docking approach, unlike traditional docking methods which typically screen a large number of chemicals against one or a few specific proteins, allows us to comprehensively evaluate how different PFAS species interact with a range of proteins involved in neurodevelopmental pathways.

In this study, we leverage our unique zebrafish model and assay with advanced virtual screening techniques to evaluate the impacts of the six EPA-regulated PFAS on social behavioral development and identified their molecular targets in the human proteome. Among these molecular targets, folic acid receptors were found to differentially bind to the six PFAS species, with the binding affinities positively correlating with their toxicity on social development. Co-exposure to folic acid rescued social deficits caused by PFAS. Previous research found an inverse association between folic acid and PFAS levels in human red blood cells and serum samples[15], indicating that PFAS exposure in humans likely impeded folic acid uptake. Given the importance of folic acid in early neurodevelopment[16-18] and its implications for autism[19], our findings strongly suggest that PFAS' interaction with folic acid receptors likely contribute to their ASD-related neurodevelopmental toxicity and point to a potential strategy for preventive interventions.

RESULTS

The six EPA-regulated PFAS showed differential impacts on social behavioral development in zebrafish.

To compare the impacts of major PFAS species on social behavioral development, we tested the six EPA-regulated PFAS, namely PFNA, PFOA, PFOS, PFHxS, PFBS, and GenX (Figure 1A), using the Fishbook assay (Figure 1B). After an initial embryonic toxicity testing, a range of doses were determined for each PFAS to assess their dose responses. Specifically, PFNA was evaluated at 0.78 μM to 200 μM, PFOA at 3.125 μM to 1600 μM, PFOS at 0.25 μM to 12 μM, PFHxS at 3.125 μM to 400 μM, PFBS at 3.125 μM to 1600 μM, and GenX at 1.56 μM to 1600 μM. Dose responses were evaluated at every 2-fold increment from the lowest to the highest dose. Embryos were exposed to a single dose of PFAS at various concentrations listed above, from 3 hours post fertilization (hpf) to 3 days post fertilization (3 dpf).

Following the embryonic exposure, embryos were raised to the juvenile stage for Fishbook testing. We observed a clear dose-response for PFNA (Figure 1C), PFOA (Figure 1D), and PFOS (Figure 1E) but not for PFHxS (Figure 1F), PFBS (Figure 1G), and GenX (Figure 1H) (Supplementary Figure S1). The IC_{50} values were determined to be 1.378 μM for PFNA (Figure 1C), 40.15 μM for PFOA (Figure 1D), and 0.1892 μM for PFOS (Figure 1E). These results affirm PFNA and PFOA's neurodevelopmental toxicity, which are aligned with previous epidemiology findings linking PFNA and PFOA to ASD risk. Our results further demonstrate the toxicity of PFOS in social behavioral development and suggest its potential implication for ASD risk.

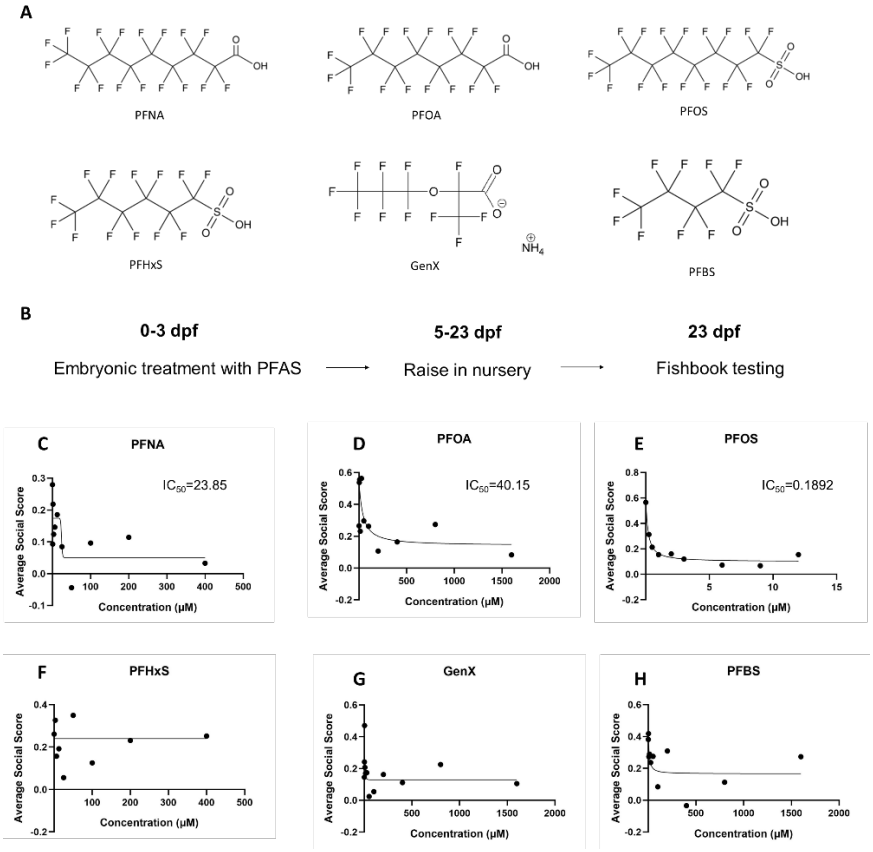

Figure 1. Fishbook assay revealed differential neurodevelopmental toxicity of six EPA-regulated PFAS on social behavioral development in zebrafish. (**A**) Chemical structures of the six EPA-regulated PFAS. (**B**) A schematic of the dosing and testing workflow. (**C-H**) Dosage curves generated by plotting average social scores (y-axis) against PFAS concentrations (x-axis). PFAS tested include PFNA (C), PFOA (D), PFOS (E), PFHxS (F), GenX (G), and PFBS (H). IC_{50} values are shown for PFNA (C), PFOA (D), and PFOS (E).

Discovering protein binding targets of PFAS via reverse molecular docking against the human proteome.

To discover the protein targets of PFAS, we conducted a reverse molecular docking screen against predicted binding pockets[12] of the human 3D proteome based on the AlphaFold database[13,14] (Figure 2A). A single domain was used for pocket prediction for most of the proteins, while multiple domains were used for predictions for a subset of proteins. For each domain, pockets were predicted based on both sequence and 3D structure. The highest-ranking sequence-based and 3D-structure-based pocket predictions were selected for each prot ein domain for the screen. A total of 33446 predicted pockets were screened. Each pocket was screened once via molecular docking for each PFAS. The combined screening results are shown in Supplementary Table S1, screening results for each PFAS are separately shown in Supplementary Tables S2-S7.

We set a cutoff at binding energy of -10 kcal/mol. Using this cutoff, no hit was identified for PFBS and GenX, and only 1 hit was identified for PFHxS (Supplementary Table S5). A number of hits were identified for PFNA, PFOA, and PFOS (Figure 2B & Supplementary Figure S2A). We noticed that folate receptor beta, or FOLR2, consistently appeared among the top hits for the three PFAS (Figure 2B), ranking #2 among the predicted targets of PFNA with a binding energy of -10.9 kcal/mol (Figure 2C), #3 for PFOA

at -10.2 kcal/mol (Figure 2D), and #1 for PFOS at -10.73 kcal/mol (Figure 2E). The other two folate receptor subtypes, folate receptor alpha (FOLR1) and folate receptor gamma (FOLR3) were also targeted by PFAS, although at lower affinities with FOLR1 binding to PFNA at an energy of -8.6 kcal/mol, to PFOA at -8.9 kcal/mol, and to PFOS at -9.4 kcal/mol, and with FOLR3 binding to PFNA at an energy of -9.9 kcal/mol, to PFOA at -9.4 kcal/mol, and to PFOS at -10 kcal/mol (Supplementary Tables S2-S4). Another two proteins involved in the folate metabolic pathway, folate hydrolase 1 (FOLH1) and methenyltetrahydrofolate synthetase (MTHFS), were also predicted to be bound by PFAS (Figure 2B & Supplementary Tables S2-S4), with FOLH1 binding to PFNA at an energy of -10.1 kcal/mol, to PFOA at -10.1 kcal/mol, and to PFOS at -10.1 kcal/mol, and with MTHFS binding to PFNA at -10 kcal/mol, to PFOA at -9.5 kcal/mol, and to PFOS at -6.5 kcal/mol. The combined outcome of the screening is summarized in Figure 2F and Supplementary Figure S2B.

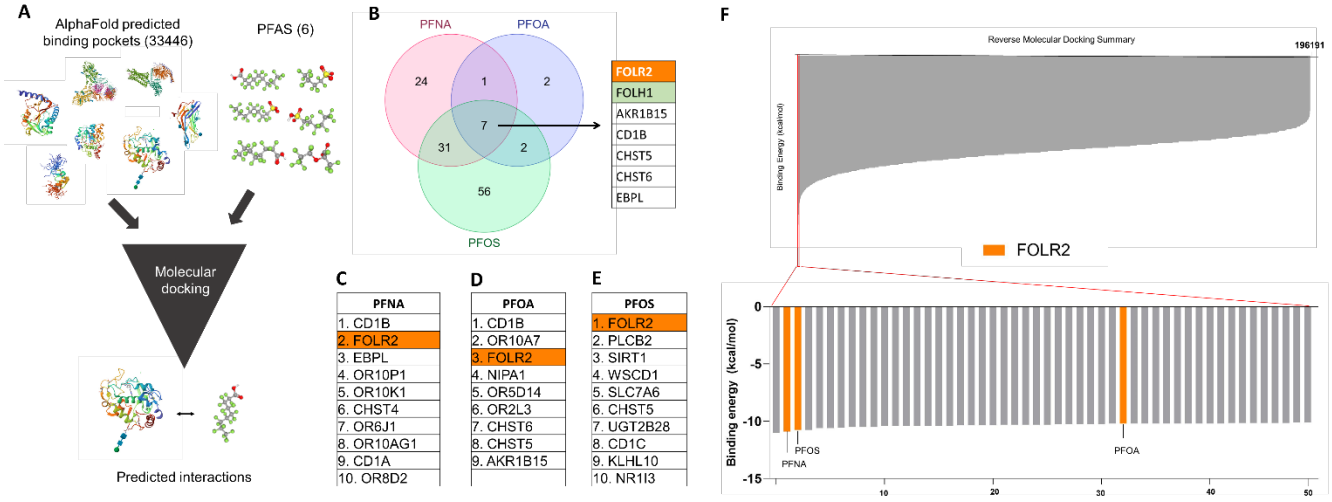

Figure 2. Reverse molecular docking discovered molecular targets of PFAS in the human proteome. (**A**) A schematic drawing demonstrating the reverse molecular docking approach. (**B**) A Venn diagram showing overlaps between the predicted protein targets of PFNA, PFOA, and PFOS. Hits were identified by selecting protein-PFAS interactions with predicted binding energy < -10 kcal/mol. The 7 proteins predicted to bind to all three PFAS are listed in the table pointed by the arrow. (**C-E**) Lists of the top 10 hits with predicted binding energy <-10 kcal/mol for PFNA (C), PFOA (D), and PFOS (E). (**F**) Summary of the reverse molecular docking screen. All molecular docking scores of six EPA-regulated PFAS compounds against predicted binding pockets in 80% of human proteome are shown as a bar chart at the top. The top 50 highest affinity binding interactions (the lowest scores) are zoomed in and shown as a bar chart at the bottom. Interactions between PFAS (PFNA, PFOS, and PFOA) and FOLR2 are highlighted in orange. Ranking is based on average binding scores for each PFAS-domain interaction, calculated from binding energies between each PFAS and binding pockets predicted by structural-based prediction and sequence-based prediction for each protein domain.

Consistent with our observations, enrichment analyses identified the folate pathway in PFNA, PFOA, and PFOS targeted proteins among the Gene Ontology (GO) pathways (Figure 3A & Supplementary Figures S3A-S3C) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Figure 3B & Supplementary Figures S3D-S3F). Protein-protein interaction (PPI) analysis using stringDB for the top hits (with predicted binding energy < -10 kcal/mol) of PFNA (Figure 3C), PFOA (Figure 3D), and PFOS (Figure 3E) consistently identified an interaction between FOLR2 and FOLH1. We did not find additional PPIs that consistently connect to FOLR2.

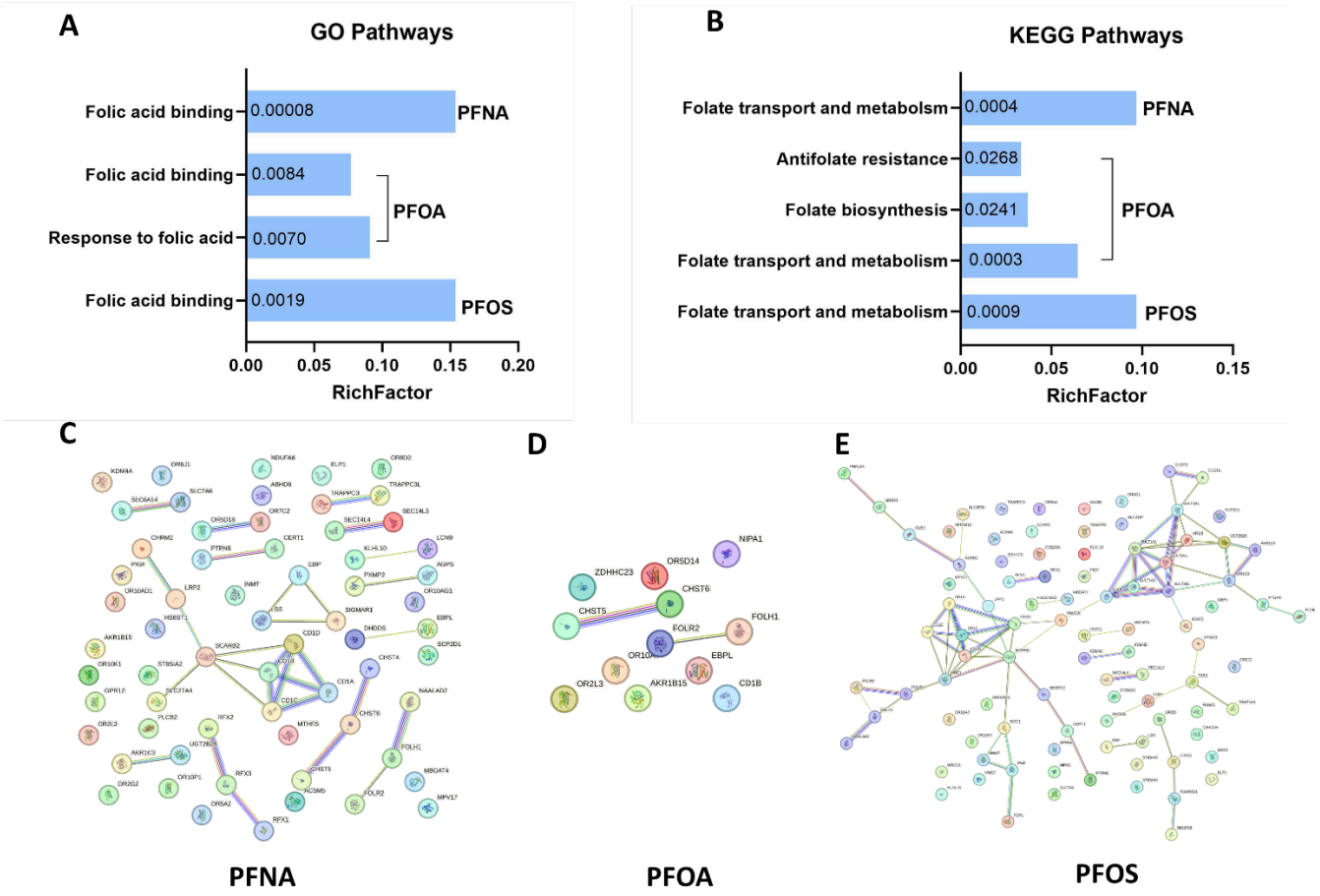

Figure 3. Enrichment analysis and protein-protein interaction analysis consistently identified the folate pathway as a target of PFAS. (**A-B**) Enrichment analyses for the GO (A) and KEGG (B) pathways consistently identified the folate pathway for PFNA, PFOA, and PFOS. Values inside each bar represents adjusted *p*-value. (**C-F**) stringDB analysis for the top hits of PFNA (C), PFOA (D), and PFOS (E) found no consistent PPI networks except a connection between FOLR2 and FOLH1.

In silico and *In vivo* validations suggest folic acid receptor to be the biological target of PFAS.

To verify FOLR2 as a binding target for PFNA, PFOA, and PFOS, and compare their binding affinities with the endogenous ligand of FOLR2, folic acid (or folate), in its natural binding pocket, we conducted molecular docking using a 3D structure of FOLR2 from the Protein Data Bank (PDB) in complex with folate (4KMZ). We found that the natural binding pocket of folate (Figure 4A) can be effectively bound by PFNA (Figure 4B), PFOA (Figure 4C), and PFOS (Figures 4D). In fact, an average of 5 random docking trials showed that PFNA, PFOA, and PFOS all bind to FOLR2 at a higher potency (-11.04±0.2191 kcal/mol, -10.34±0.1949 kcal/mol, and -10.28±0.04472 kcal/mol, respectively) compared to folate (-9.68±0.04472 kcal/mol), while PFHxS, PFBS, and GenX bind to FOLR2 at a lower potency compared to folate (-9.3±0.0 kcal/mol, -8.82±0.2049 kcal/mol, and -8.1±0.0 kcal/mol, respectively) (Figure 4E). We also conducted molecular docking with FOLR1 using a 3D structure of FOLR1 from the PDB in complex with folate (4LRH) (Supplementary Figures S4A-S4D) and found a similar pattern of differential binding affinity among the PFAS species: PFOS and PFNA bind to FOLR1 at a potency higher than folate (PFOS: -10.08±0.08357 kcal/mol, PFNA: -9.8±0.2236 kcal/mol, folate: -9.58±0.1095 kcal/mol), whereas PFOA, PFHxS, PFBS, and GenX bind to FOLR1 at a lower potency compared to folate (-9.3±0.0 kcal/mol, -8.7±0.0 kcal/mol, -8.4±0.0 kcal/mol, and -7.7±0.0 kcal/mol, respectively) (Supplementary Figure S4E). Consistent with a previous report which associated PFAS exposure to lower folate levels in human blood[15], our findings suggest that exposure to PFNA, PFOA, and PFOS may out-compete FOLR2's natural ligand, folate, and lead to reduced folate uptake *in vivo*.

We then examined whether folate deficiency contributes to social behavioral deficits in our zebrafish social behavior model. The zebrafish only possess one gene encoding the folate receptor, *folr*. To evaluate how *folr* loss-of-function affects sociality, we knocked out the zebrafish *folr* using CRISPR/Cas9

by injecting fertilized eggs with a mixture of *folr* gRNAs and Cas9 protein. The resulting F_0 mutants show significant social deficits compared to wild-type controls (Figure 4F), demonstrating that folate pathway is required for social behavioral development in zebrafish.

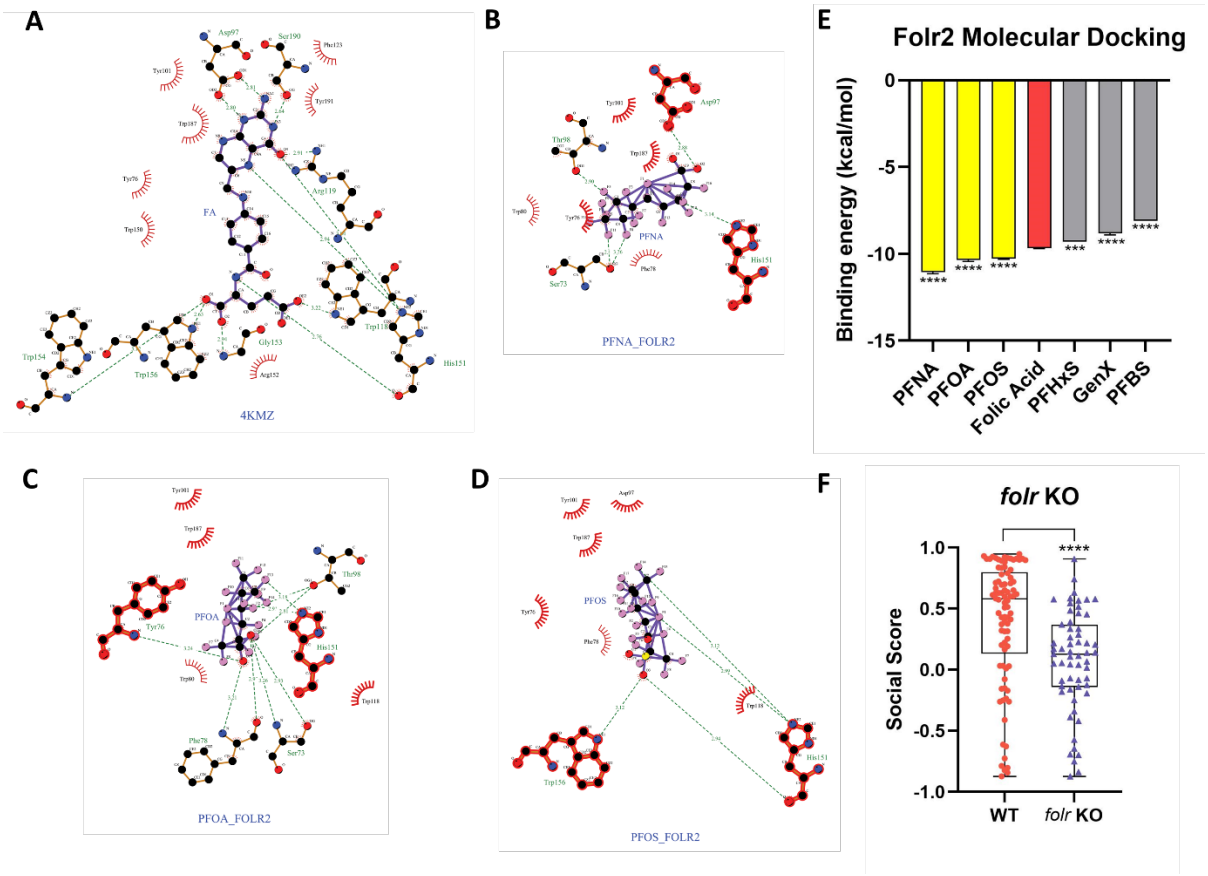

Figure 4. *In silico* and *In vivo* assays validated folic acid receptor as the biological target of PFAS. (**A-D**) *In silico* validation via molecular docking. Showing docking images for FOLR2 against folic acid (FA) (A), PFNA (B), PFOA (C), and PFOS (D). FOLR2 structure was acquired using the PDB identifier 4KMZ. (**E**) Comparing the binding energies for folic acid and the six EPA-regulated PFAS against FOLR2. Binding energies are shown as average and standard deviation of 5 independent docking attempts. Significance values are calculated to compare PFAS data with folic acid. ***: $p < 0.001$, ****: $p < 0.0001$. (**F**) CRISPR-Cas9 induced F_0 knockout of the zebrafish folic acid receptor *folr* induced social deficits in zebrafish. ****: $p < 0.0001$.

Folic acid rescued social deficits induced by PFAS through embryonic co-exposure.

Finally, to evaluate whether PFAS induces social deficit through inhibiting the folate pathway, we attempted to rescue social deficits caused by PFAS exposure by co-exposing with folic acid (Figure 5A). We hypothesized that certain doses of folic acid may outcompete PFAS binding, restoring folate metabolism in the developing embryos, without causing developmental toxicity. We co-exposed the three PFAS species including 400 μM PFNA, 1600 μM PFOA, and 10 μM PFOS with a range of folic acid doses (7 doses ranging from 3.125 μM to 200 μM in 2-fold increments) during the first 3 days of embryonic development. Folic acid co-exposure effectively rescued social deficits in all three PFAS treatments. PFNA was rescued by folic acid at the lowest applied doses of 3.125 μM to 25 μM (Figure 5B & Supplementary Figure S5A). Both PFOA and PFOS showed a bell-shaped rescue curve when co-exposed to folic acid: PFOA was most effectively rescued by folic acid at 6.25 μM to 12.5 μM (Figure 5C & Supplementary Figure S5B), whereas PFOS was most effectively rescued by 12.5 μM to 25 μM folic acid treatment (Figure 5D & Supplementary Figure S5C). These results indicate that PFAS likely induces social deficit by inhibiting the folate pathway, and this inhibition can be prevented by embryonic co-exposure to folic acid.

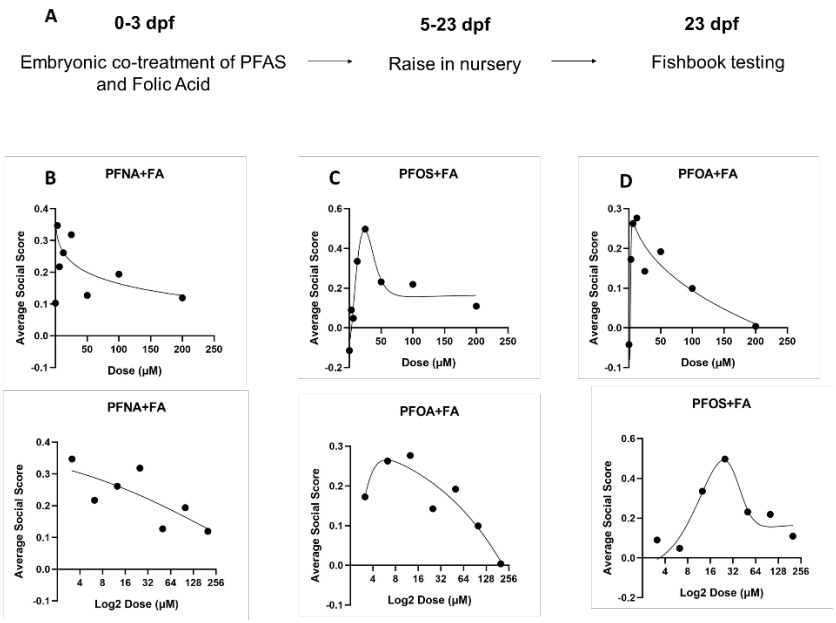

Figure 5. Embryonic co-exposure to folic acid rescued social deficits induced by PFAS. (**A**) A schematic of the co-treatment and testing workflow. (**B-D**) Dosage curves generated by plotting average social scores (y-axis) against folic acid (FA) concentrations (x-axis). Showing both linear (top) and log2 (bottom) scales for the x-axis. PFAS were exposed to embryos at 400 μM for PFNA (B), 1600 μM for PFOA (C), and 10 μM for PFOS (E).

DISCUSSION

Our study provides critical insights into the neurodevelopmental toxicity of six EPA-regulated PFAS compounds—PFNA, PFOA, PFOS, PFHxS, PFBS, and GenX—on social behavioral development in zebrafish. Using the Fishbook assay, we observed distinct dose-dependent responses for PFNA, PFOA, and PFOS, which displayed neurodevelopmental impacts on social behavior in the exposed embryos, while PFHxS, PFBS, and GenX exhibited no apparent dose-dependent effects on social behavior within the tested concentration range. IC_{50} values were established for PFNA, PFOA, and PFOS, reinforcing epidemiological links between PFNA, PFOA, and autism spectrum disorder (ASD) risk, while suggesting a potential contribution of PFOS to ASD risk. This specificity in behavioral effects aligns with our molecular docking results, where PFNA, PFOA, and PFOS displayed stronger binding affinities to the folate receptor FOLR2, suggesting that these PFAS may disrupt folate metabolism and thereby influence neurodevelopmental pathways implicated in social behavioral development.

Recent studies have revealed a potential role of folic acid in ASD[19]. Low maternal folate levels have been associated with a higher incidence of ASD in offspring[20-22]. A number of studies have linked maternal folic acid supplementation with reduced risk of ASD in offspring[22-28]. For example, maternal folic acid supplementation at 600 μg has been associated with a significantly reduced risk of ASD in offspring[23], suggesting that adequate folate levels during pregnancy may be beneficial for disease prevention. However, the dosage of folic acid supplementation appears to affect its protective effects, with higher or lower levels of supplementation associated with increased risk of ASD[25]. Supplementation with folic acid and/or folinic acid in children diagnosed with ASD improved their neurological and behavioral symptoms[29-33]. Several studies have reported an elevated prevalence of serum auto-antibodies against folate receptor alpha in children with ASD[34-36]. Finally, polymorphisms in the methylenetetrahydrofolate reductase (MTHFR) gene have been associated with ASD risk[37-43]. These findings align with our observations that PFAS-induced folate pathway disruption can lead to social behavioral deficits.

The folate pathway may contribute to ASD through several potential mechanisms. Folic acid is required for DNA methylation, a critical process during brain development: it is a source of the one-carbon group used to methylate DNA and required for the production of S-adenosylmethionine, the primary methyl group donor for DNA methylation. Folic acid is also essential for neural tube formation. It is worth noting that the optimal protective effects of folic acid against ASD was observed when folic acid was taken during preconception and the first trimester of pregnancy[23,28], the latter of which overlaps with the period of neural tube formation. Both DNA methylation and neural tube formation are critical processes for brain

development, pointing to potential mechanisms of action for folic acid's involvement in ASD. Additionally, folic acid is essential for the production of neurotransmitters that affect mood, motivation, stress, and cognitive performance, including serotonin, norepinephrine, and dopamine[44,45], influencing neural signaling pathways that are often disrupted in ASD.

Some studies indicate a possible link between elevated maternal folate levels and an increased risk of ASD, although findings are inconsistent and complex due to variations in individual metabolism, timing, and dosage of supplementation[46]. These findings suggest that while folic acid is crucial for brain development, there is a need for more nuanced guidance on optimal intake levels to balance benefits with potential risks, especially for populations at risk for ASD. This is aligned with our finding that while moderate to low doses of folic acid exerted beneficial rescue effects on social behavior when co-exposed to PFAS, higher doses showed significantly reduced rescue potencies.

Although prior work has utilized reverse molecular docking on selected proteins to predict binding targets of the environmental toxicant toluene[47], to our knowledge, this work represents the first proteome-wide reverse molecular docking in toxicology to discover *in vivo* protein targets for environmental toxicants. In reverse molecular docking, a small molecule drug or toxicant is virtually docked into the potential binding pockets of a large set of proteins. The binding characteristics of these interactions are then analyzed to rank the protein targets based on their binding affinities to the small molecules. This approach is different from traditional molecular docking which typically screens a large number of small molecules against a predetermined protein target, with the goal of identifying a ligand with high binding affinity to this protein. The successful application of proteome-wide reverse molecular docking in toxicology demonstrates a rapid and systematic *in silico* approach to predict protein targets for key environmental toxicants, which is expected to significantly promote the advance of mechanistic toxicology.

In summary, this study indicates that specific PFAS, notably PFNA, PFOA, and PFOS, may contribute to social behavioral deficits through their interaction with folate receptors, an effect mitigated by co-exposure to folic acid. Our findings align with existing evidence on the association between folate pathway disruptions and ASD risk, contributing to a deeper understanding of the biochemical mechanisms by which environmental toxicants may influence neurodevelopmental and disease outcomes. Furthermore, the ability of folic acid to rescue PFAS-induced deficits highlights the potential of folate supplementation to counteract PFAS toxicity in developing embryos. Together, these results underscore the need for further research into folate-based interventions and the long-term implications of exposure to various species of PFAS on neurodevelopmental disorders, particularly ASD. In addition, by incorporating the innovative proteome-wide reverse molecular docking approach, we provide a comprehensive framework for advancing the field of mechanistic toxicology through the identification of novel toxicological targets.

MATERIALS AND METHODS

Zebrafish husbandry

Zebrafish were housed at 26°C–27°C on a 14-hour light, 10-hour dark light cycle. Wild-type AB strain was used for all experiments. All zebrafish experiments were approved by the Institutional Animal Care and Use Committee at the University of Washington.

Chemical exposure

Fertilized embryos were sorted and transferred into 60 mm diameter and 20 mm deep glass petri dishes at 30 embryos per dish. Each dish was filled with 10 mL of HEPES-buffered E3 medium. Compound stocks were prepared in their specified solvents to maintain stability and stored at -20°C. Each compound was added to duplicate dishes at various concentrations, with negative controls receiving an equal volume of the corresponding solvent. To prevent media contamination, dead embryos were removed at 1 and 2 days post-fertilization (dpf). At 3 dpf, all viable larvae were rinsed with E3 medium and transferred into clean petri dishes containing fresh E3 medium. At 5 dpf, larvae from each petri dish were transferred to separate nursery tanks and raised to 23 dpf for Fishbook assays.

The following per- and polyfluoroalkyl substances (PFAS) were obtained from Cayman Chemical (Ann Arbor, MI, USA): Perfluorooctanesulfonic Acid (PFOS, CAS# 1763-23-1, Item no. 37233), Perfluorooctanoic Acid (PFOA, CAS# 335-67-1, Item no. 37232), Perfluorononanoic Acid (PFNA, CAS# 375-95-1, Item no. 37250), Perfluorohexanesulfonic Acid (PFHxS, CAS# 355-46-4, Item no. 37248),

Perfluorobutanesulfonic Acid (PFBS, CAS# 375-73-5, Item no. 37243), and Folic Acid (CAS# 59-30-3, Item no. 20515). Ammonium perfluoro(2-methyl-3-oxahexanoate) (GenX, CAS# 62037-80-3, Item no. C23990) was obtained from Sigma-Aldrich. PFOS came as a solution in 100% ethanol. PFBS came in liquid form. To prepare stock solutions, PFOA, PFNA, and folic acid were dissolved in HEPES-buffered E3 media, GenX and PFHxS were dissolved in dimethyl sulfoxide (DMSO).

Fishbook assay

The Fishbook assay was run as described previously[10]. Briefly, the test arena consists of a total of 44 3D printed, 10-mm-deep, 8.5-mm-wide, and 80-mm-long rectangular chambers grouped together in parallel with each other. Each chamber is divided into three compartments by two transparent acrylic windows (1.5 mm thick): a 60-mm-long middle testing chamber to place the test subject and two 8.5-mm-long end chambers to place the social stimulus fish or remain empty, respectively. Test subjects (PFAS treated 23 dpf fish or controls) were each placed inside an individual test chamber using a plastic transfer pipette with its tip cut off to widen the opening. A 3D printed white comb-like structure was placed in front of the social stimulus compartment to block the test subject's visual access to social stimulus fish before testing begins. After test subjects were placed inside the chambers, the arena was placed inside the imaging station, and the combs were removed to visually expose the social stimulus fish to the test subjects. Following a brief acclimation period, a 10-min test session was video recorded.

Videos were streamed through the software Bonsai[48]. Videos were analyzed in real time during recording, and the frame-by-frame x and y coordinates of each fish relative to its own test compartment were exported as a CSV file. Data were analyzed using custom Python scripts to calculate social scores and generate tracking plots. Social score was defined as a fish's average y-axis position for all frames. The middle of each test chamber was designated as the origin of the y axis, with an assigned value of zero. A value of 1 was assigned to the end of the chamber next to the social stimulus fish and a value of −1 to the other end of the chamber next to the empty control compartment. In this coordinate system, all social scores have values between −1 and 1. A higher social score demonstrates a shorter average distance between a test subject and a social stimulus fish during a test, which suggests a stronger social preference.

CRISPR-Cas9 gene knockout

We designed 3 sets of sgRNAs targeting the zebrafish *folr* gene to maximize efficiency of F_0 knockout[49-51]. CHOPCHOP[52] was used for the sgRNA design. The *folr* sgRNA sequences designed are as follows: gRNA1 "ATTTAGGTGACACTATATGAGGGACAGTTATATCAGCGTTTTAGAGCTAGAAATAGCAAG", gRNA2 "ATTTAGGTGACACTATATGTATCCAGGGCCCCAGGTGGTTTTAGAGCTAGAAATAGCAAG", and gRNA3 "ATTTAGGTGACACTATAGTCCATGTTCTGCTCAGGTGGTTTTAGAGCTAGAAATAGCAAG". sgRNAs targeting early coding exons were selected to introduce nonsense mutations as early as possible and increase likelihood of loss-of-function mutations.

The SP6 promoter sequence was added before the target RNA sequence, followed by an overlap adapter, which is complementary to the 5' end of an 80 bp constant oligo according to the published protocol[53]. Oligonucleotides were synthesized by Eurofins Scientific. Double-stranded DNAs were generated by Phusion Hot-Start Flex DNA Polymerase (New England Biolabs) using the gene specific oligos and the constant oligo. The double-stranded DNA was cleaned up using the Zymo Clean and Concentrator-5 kit (Zymo Research). *In vitro* sgRNA transcription was conducted using the MEGAscript SP6 transcription kit (Thermo Fisher Scientific) and cleaned up using the Zymo RNA Clean and Concentrator-5 kit (Zymo Research).

To perform zebrafish embryonic microinjections, pairs of male and female adult zebrafish were kept in a mating cage overnight while separated by a divider. 1-cell-stage embryos were collected the next day early in the morning immediately before injection by pulling out the divider to allow breeding. Injection solution was prepared by combining sgRNA with 2 μM Spy Cas9 NLS and 1X NEBuffer (New England Biolabs).

Reverse molecular docking and molecular docking

All *in silico* calculations were performed using a Digital Storm workstation with AMD Ryzen Threadripper PRO 5955WX (16-Core) processor (Up to 4.0 GHz), 128GB DDR4 memory, and NVIDIA RTX A2000

6GB graphics card. Both reverse molecular docking screen and molecular docking validation were conducted using Autodock Vina[54] and Open Babel[55]. Ligand structures were downloaded from PubChem as sdf files and converted to pdbqt. Binding pocket predictions were acquired from the HPRoteome-BSite database[12] which covers over 80% of the UniProt entries in the AlphaFold human proteome database[13,14]. In the HPRoteome-BSite database, a binding pocket was predicted for a single domain for most of the proteins, whereas a subset of proteins had a binding pocket predicted for multiple domains. For each domain, pockets were predicted based on either the peptide sequence or the 3D structure of the domain; both types of predictions were used for the screen. The highest-ranking sequence-based predictions and the highest-ranking 3D-structure-based pocket predictions were respectively selected for each protein domain, resulting in a total of 33446 predicted pockets selected for the screen. The domain structures were downloaded from the HPRoteome-BSite database and prepared for molecular docking by removing all water molecules, ions, and other substructures. The docking coordinates of each predicted binding pocket were provided by the HPRoteome-BSite database. Pocket size of 60 × 60 × 60, energy range of 3, number of modes of 9, and an exhaustiveness of 8 were employed for the molecular docking screen and validation. Each pocket was screened once for each PFAS.

Protein-protein interaction analysis

Protein-protein interaction (PPI) analysis was conducted using STRING[56] (https://string-db.org/).

Gene enrichment analysis

Molecular targets with a binding energy of less than -10 kcal/mol from the reverse molecular binding virtual screen were selected for gene enrichment analysis. Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were conducted using the clusterProfiler package in R (version 4.12.5). The GO analysis included Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories, with the org.Hs.eg.db database used for annotation (v3.19.1). Results were ranked based on RichFactor and adjusted *p*-value. Adjusted *p*-value was calculated using the Benjamini-Hochberg (BH) method with a cutoff of 0.05 for significance. RichFactor is defined as the ratio of input genes that are annotated under a certain term to all genes annotated under this term. Results were visualized as bar plots using the ggplot2 package (v3.5.1).

Statistical analysis

Graphs were generated using GraphPad Prism. Two-tailed Student's *t* test was used to analyze data when there were two groups. For analysis of multiple groups, normal distribution of datasets was confirmed using Shapiro-Wilk test, and if more than half the data was normally distributed and standard deviations fell within the variance ratio, one-way analysis of variance (ANOVA) assuming gaussian distribution and equal standard deviation was performed, using Dunnett's multiple comparison test. *P* values less than 0.05 were considered significant.

ACKNOWLEDGEMENTS

We thank the University of Washington Office of Comparative Medicine for providing zebrafish husbandry support. This work was supported by the National Institute of Environmental Health Sciences (NIEHS) of the NIH under the award number R00ES031050. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

AUTHOR CONTRIBUTIONS

M.J. and A.X.K. designed and conducted the experiments and analyzed data. A.E. and K.P assisted in conducting experiments. P.Z. performed pathway enrichment analysis. Y.G. conceived the study and interpreted the data. M.J., A.X.K., and Y.G. wrote the manuscript. All authors contributed meaningful insights during discussions and reviewed and approved the final version of the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

DATA AND MATERIALS AVAILABILITY

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

REFERENCES

1 Hansen, S. N., Schendel, D. E. & Parner, E. T. Explaining the increase in the prevalence of autism spectrum disorders: the proportion attributable to changes in reporting practices. *JAMA Pediatr* **169**, 56-62 (2015).

2 King, M. & Bearman, P. Diagnostic change and the increased prevalence of autism. *Int J Epidemiol* **38**, 1224-1234 (2009). PMC2800781

3 Muir, D. C. G., Getzinger, G. J., McBride, M. & Ferguson, P. L. How Many Chemicals in Commerce Have Been Analyzed in Environmental Media? A 50 Year Bibliometric Analysis. *Environ Sci Technol* **57**, 9119-9129 (2023). PMC10308830

4 Gaugler, T. *et al.* Most genetic risk for autism resides with common variation. *Nat Genet* **46**, 881-885 (2014). PMC4137411

5 Sandin, S. *et al.* The familial risk of autism. *JAMA* **311**, 1770-1777 (2014). PMC4381277

6 Ames, J. L. *et al.* Prenatal Exposure to Per- and Polyfluoroalkyl Substances and Childhood Autism-related Outcomes. *Epidemiology* **34**, 450-459 (2023). PMC10074577

7 Oh, J. *et al.* Prenatal exposure to per- and polyfluoroalkyl substances in association with autism spectrum disorder in the MARBLES study. *Environ Int* **147**, 106328 (2021). PMC7856021

8 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). PMC7906952

9 Schymanski, E. L. *et al.* Per- and Polyfluoroalkyl Substances (PFAS) in PubChem: 7 Million and Growing. *Environ Sci Technol* **57**, 16918-16928 (2023). PMC10634333

10 Geng, Y. *et al.* Top2a promotes the development of social behavior via PRC2 and H3K27me3. *Sci Adv* **8**, eabm7069 (2022). PMC9683714

11 Geng, Y., Yates, C. & Peterson, R. T. Social behavioral profiling by unsupervised deep learning reveals a stimulative effect of dopamine D3 agonists on zebrafish sociality. *Cell Rep Methods* **3**, 100381 (2023). PMC9939379

12 Sim, J., Kwon, S. & Seok, C. HProteome-BSite: predicted binding sites and ligands in human 3D proteome. *Nucleic Acids Res* **51**, D403-D408 (2023). PMC9825455

13 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). PMC8371605

14 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022). PMC8728224

15 Zhang, Y. *et al.* Folate concentrations and serum perfluoroalkyl and polyfluoroalkyl substance concentrations in adolescents and adults in the USA (National Health and Nutrition Examination Study 2003-16): an observational study. *Lancet Planet Health* **7**, e449-e458 (2023). PMC10901144

16 Balashova, O. A., Visina, O. & Borodinsky, L. N. Folate action in nervous system development and disease. *Dev Neurobiol* **78**, 391-402 (2018). PMC5867258

17 Irvine, N., England-Mason, G., Field, C. J., Dewey, D. & Aghajafari, F. Prenatal Folate and Choline Levels and Brain and Cognitive Development in Children: A Critical Narrative Review. *Nutrients* **14** (2022). PMC8778665

18 Naninck, E. F. G., Stijger, P. C. & Brouwer-Brolsma, E. M. The Importance of Maternal Folate Status for Brain Development and Function of Offspring. *Adv Nutr* **10**, 502-519 (2019). PMC6520042

19 Hoxha, B. *et al.* Folic Acid and Autism: A Systematic Review of the Current State of Knowledge. *Cells* **10** (2021). PMC8394938

20 Gao, Y. *et al.* New Perspective on Impact of Folic Acid Supplementation during Pregnancy on Neurodevelopment/Autism in the Offspring Children - A Systematic Review. *PLoS One* **11**, e0165626 (2016). PMC5119728

21 DeVilbiss, E. A., Gardner, R. M., Newschaffer, C. J. & Lee, B. K. Maternal folate status as a risk factor for autism spectrum disorders: a review of existing evidence. *Br J Nutr* **114**, 663-672 (2015).

22 Steenweg-de Graaff, J., Ghassabian, A., Jaddoe, V. W., Tiemeier, H. & Roza, S. J. Folate concentrations during pregnancy and autistic traits in the offspring. The Generation R Study. *Eur J Public Health* **25**, 431-433 (2015).

23 Schmidt, R. J., Iosif, A. M., Guerrero Angel, E. & Ozonoff, S. Association of Maternal Prenatal Vitamin Use With Risk for Autism Spectrum Disorder Recurrence in Young Siblings. *JAMA Psychiatry* **76**, 391-398 (2019). PMC6450282

24 Levine, S. Z. *et al.* Association of Maternal Use of Folic Acid and Multivitamin Supplements in the Periods Before and During Pregnancy With the Risk of Autism Spectrum Disorder in Offspring. *JAMA Psychiatry* **75**, 176-184 (2018). PMC5838577

25 Raghavan, R. *et al.* Maternal Multivitamin Intake, Plasma Folate and Vitamin B(12) Levels and Autism Spectrum Disorder Risk in Offspring. *Paediatr Perinat Epidemiol* **32**, 100-111 (2018). PMC5796848

26 DeVilbiss, E. A. *et al.* Antenatal nutritional supplementation and autism spectrum disorders in the Stockholm youth cohort: population based cohort study. *BMJ* **359**, j4273 (2017). PMC6168830

27 Nilsen, R. M. *et al.* Analysis of self-selection bias in a population-based cohort study of autism spectrum disorders. *Paediatr Perinat Epidemiol* **27**, 553-563 (2013). PMC3851582

28 Suren, P. *et al.* Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. *JAMA* **309**, 570-577 (2013). PMC3908544

29 Adams, J. B. *et al.* Comprehensive Nutritional and Dietary Intervention for Autism Spectrum Disorder-A Randomized, Controlled 12-Month Trial. *Nutrients* **10** (2018). PMC5872787

30 James, S. J. *et al.* Efficacy of methylcobalamin and folinic acid treatment on glutathione redox status in children with autism. *Am J Clin Nutr* **89**, 425-430 (2009). PMC2647708

31 Kaluzna-Czaplinska, J., Michalska, M. & Rynkowski, J. Vitamin supplementation reduces the level of homocysteine in the urine of autistic children. *Nutr Res* **31**, 318-321 (2011).

32 Moretti, P. *et al.* Cerebral folate deficiency with developmental delay, autism, and response to folinic acid. *Neurology* **64**, 1088-1090 (2005).

33 Sun, C., Zou, M., Zhao, D., Xia, W. & Wu, L. Efficacy of Folic Acid Supplementation in Autistic Children Participating in Structured Teaching: An Open-Label Trial. *Nutrients* **8** (2016). PMC4924178

34 Zhou, J. *et al.* High prevalence of serum folate receptor autoantibodies in children with autism spectrum disorders. *Biomarkers* **23**, 622-624 (2018).

35 Frye, R. E. *et al.* Folinic acid improves verbal communication in children with autism and language impairment: a randomized double-blind placebo-controlled trial. *Mol Psychiatry* **23**, 247-256 (2018). PMC5794882

36 Ramaekers, V. T., Blau, N., Sequeira, J. M., Nassogne, M. C. & Quadros, E. V. Folate receptor autoimmunity and cerebral folate deficiency in low-functioning autism with neurological deficits. *Neuropediatrics* **38**, 276-281 (2007).

37 Ismail, S. *et al.* Study of C677T variant of methylene tetrahydrofolate reductase gene in autistic spectrum disorder Egyptian children. *Am J Med Genet B Neuropsychiatr Genet* **180**, 305-309 (2019).

38 El-Baz, F., El-Aal, M. A., Kamal, T. M., Sadek, A. A. & Othman, A. A. Study of the C677T and 1298AC polymorphic genotypes of MTHFR Gene in autism spectrum disorder. *Electron Physician* **9**, 5287-5293 (2017). PMC5633227

39 Shaik Mohammad, N. *et al.* Clinical utility of folate pathway genetic polymorphisms in the diagnosis of autism spectrum disorders. *Psychiatr Genet* **26**, 281-286 (2016).

40 Guo, T., Chen, H., Liu, B., Ji, W. & Yang, C. Methylenetetrahydrofolate reductase polymorphisms C677T and risk of autism in the Chinese Han population. *Genet Test Mol Biomarkers* **16**, 968-973 (2012).

41 James, S. J. *et al.* A functional polymorphism in the reduced folate carrier gene and DNA hypomethylation in mothers of children with autism. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 1209-1220 (2010). PMC2943349

42 Goin-Kochel, R. P. *et al.* The MTHFR 677C-->T polymorphism and behaviors in children with autism: exploratory genotype-phenotype correlations. *Autism Res* **2**, 98-108 (2009).
43 Mohammad, N. S. *et al.* Aberrations in folate metabolic pathway and altered susceptibility to autism. *Psychiatr Genet* **19**, 171-176 (2009).
44 Zhou, Y., Cong, Y. & Liu, H. Folic acid ameliorates depression-like behaviour in a rat model of chronic unpredictable mild stress. *BMC Neurosci* **21**, 1 (2020). PMC6961331
45 Miller, A. L. The methylation, neurotransmitter, and antioxidant connections between folate and depression. *Altern Med Rev* **13**, 216-226 (2008).
46 Wiens, D. & DeSoto, M. C. Is High Folic Acid Intake a Risk Factor for Autism?-A Review. *Brain Sci* **7** (2017). PMC5704156
47 McCarthy, M. J., Chushak, Y. & Gearhart, J. M. Reverse molecular docking and deep-learning to make predictions of receptor activity for neurotoxicology. *Comput Toxicol* **24** (2022).
48 Lopes, G. *et al.* Bonsai: an event-based framework for processing and controlling data streams. *Front Neuroinform* **9**, 7 (2015). PMC4389726
49 Kim, B. H. & Zhang, G. Generating Stable Knockout Zebrafish Lines by Deleting Large Chromosomal Fragments Using Multiple gRNAs. *G3 (Bethesda)* **10**, 1029-1037 (2020). PMC7056962
50 Kroll, F. *et al.* A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *Elife* **10** (2021). PMC7793621
51 Wu, R. S. *et al.* A Rapid Method for Directed Gene Knockout for Screening in G0 Zebrafish. *Dev Cell* **46**, 112-125 e114 (2018).
52 Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* **42**, W401-407 (2014). PMC4086086
53 Gagnon, J. A. *et al.* Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, e98186 (2014). PMC4038517
54 Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-461 (2010). PMC3041641
55 O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J Cheminform* **3**, 33 (2011). PMC3198950
56 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613 (2019). PMC6323986