Employing the SeqAPASS tool to inform bioaccumulation potential of per- and polyfluorinated alkyl substances across species

†Carlie A. LaLone, ‡Jon A. Doering, § Weixiao Cheng, ¶ Hui Peng, and § Carla Ng

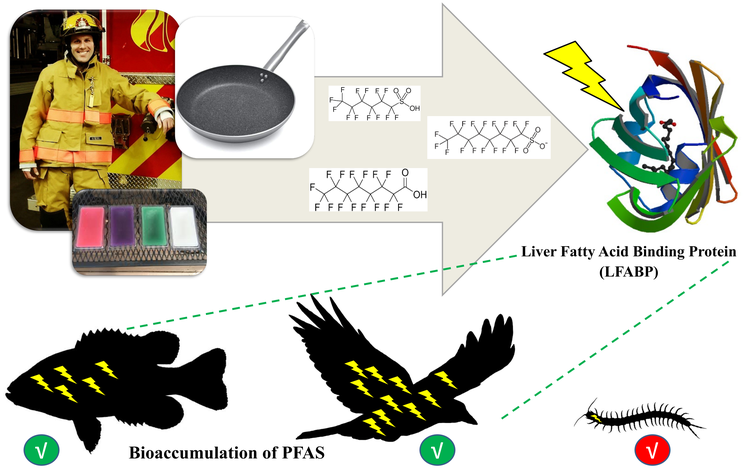
† U.S. Environmental Protection Agency, Center for Computational Toxicology and Exposure, Duluth, MN 55804; ‡National Research Council, Duluth, MN 55804; § University of Pittsburgh, Department of Civil & Environmental Engineering, Pittsburgh, PA 15261; ¶ University of Toronto, School of the Environment, Toronto, ON, Canada

Carlie A. LaLone | LaLone.Carlie@epa.gov | 218-529-5038

www.epa.gov

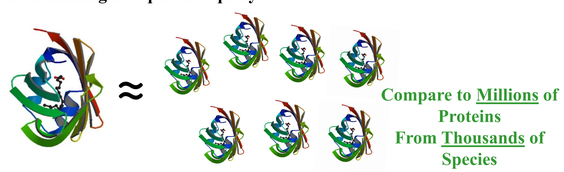
Introduction

Per- and polyfluorinated alkylated substances (PFAS) are synthetic chemicals used in a variety of industrial applications and consumer products, notably fire-fighting foams and stain and oil repellents. Due to the ubiquitous nature of PFAS in the environment, they have been measured in tissues from species as diverse as whales, birds, fish, and even invertebrates, covering a range of trophic levels. The ability of these chemicals to bioaccumulate is largely due to protein binding, with both serum albumin in the blood and fatty acid binding proteins in the liver capable of important interactions. Due to the involvement of proteins in bioaccumulation, the US Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool was used to evaluate protein conservation and predict similarities and differences in bioaccumulation potential across species.[1] Results from SeqAPASS were then used to guide molecular homology modeling and molecular dynamics simulations to further evaluate species similarities and differences to predict potential for bioaccumulation of PFAS across species. Cross-species chemical proteomic studies will be used to confirm and expand upon in silico results.



Sequence Alignment to Predict Across Species Susceptibility

SeqAPASS v4.0 (https://seqapass.epa.gov/seqapass/)
Computational Assessment of Protein Similarity
Necessary information for submitting a SeqAPASS query:



Greater similarity = Greater likelihood that chemical can act on the protein
Line of Evidence: Predict Potential Chemical Bioaccumulation Potential Across Species

Conservation of Serum Albumin: SeqAPASS Results

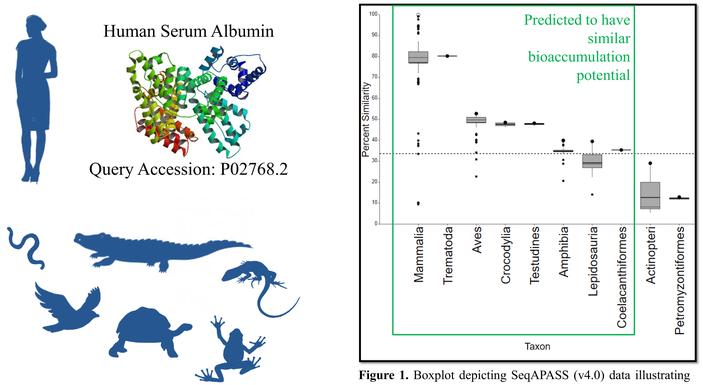


Figure 1. Boxplot depicting SeqAPASS (v4.0) data illustrating the percent similarity across species compared to the primary amino acid sequences for human (Homo sapiens) serum albumin

Characterization of PFAS-Liver Fatty Acid Protein Binding

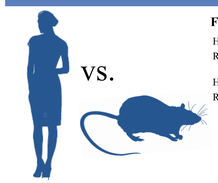
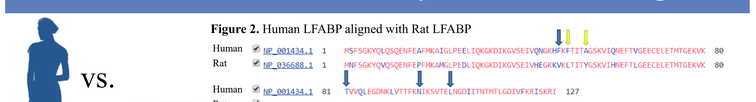


Figure 2. Human LFABP aligned with Rat LFABP



22 amino acid residues different between Human and Rat
Of those, **only 6 amino acids** are predicted by SeqAPASS to impact the protein-chemical interaction

	hLFABP		rLFABP	
ligands	H-bond interaction	largest energy contribution	H-bond interaction	largest energy contribution
PFBA	ARG 122, SER 124	ARG 122, SER 39, ILE 52	–	SER 57, LYS 58, LYS 32
PFPA	ARG 122, SER 124	ARG 122, VAL 83, PHE 50	–	ARG 122, TYR 55, ILE 53
PFHxA	ARG 122, SER 124	ARG 122, SER 39, SER 124	TYR 120	ARG 122, ILE 53, LYS 58
PFHpA	ARG 122, SER 124	ARG 122, SER 39, ILE 52	–	ARG 122, ILE 60, MET 74
PFOA	ARG 122, SER 124	ARG 122, SER 39, ILE 52	–	ARG 122, TYR 55, ILE 60
PFNA	ARG 122, SER 124	ARG 122, SER 39, ILE 52	–	ARG 122, ILE 60, ILE 53
PFBS	ARG 122, SER 39	ARG 122, SER 124, LEU 9	ARG 122, SER 39	ARG 122, SER 100, LEU 71
PFHxS	ARG 122	ARG 122, SER 124, SER 39	–	ARG 122, ASN 111, LEU 51
PFOS	ARG 122, SER 124	ARG 122, SER 124, ILE 52	TYR 120	ARG 122, ILE 60, ILE 53
EEA	ARG 122, SER 39	ARG 122, SER 39, ASN 111	–	ARG 122, MET 74, ILE 60
GenX	THR 102	ARG 122, ASN 111, THR 73	–	ARG 122, MET 74, ILE 53
ADONA	ARG 122, SER 124	ARG 122, SER 39, SER 124	–	ARG 122, MET 74, TYR 55
2m-PFOA	SER 100	ARG 122, SER 100, ASN 111	–	ARG 122, TYR 120, ILE 60
F-53	ARG 122, SER 124	ARG 122, PHE 50, SER 39	TYR 120	ARG 122, SER 124, ILE 53
F-53B	ARG 122, SER 124	ARG 122, SER 124, SER 39	TYR 120	ARG 122, TYR 55, ILE 60

Table 1. Amino Acids interacting with PFAS (From Cheng and Ng, 2018[3])

Characteristics for predicting differences using SeqAPASS[2]:

- Amino Acid Residue Classification (e.g., Aromatic, Basic, Hydroxylic)
- Molecular weight as surrogate for size (Difference of > 30g/mol)

1. Position 48 hLFABP Phenylalanine and rLFABP Valine
2. Position 50 hLFABP Phenylalanine and rLFABP Leucine
3. Position 54 hLFABP Alanine and rLFABP Tyrosine
4. Position 81 hLFABP Threonine and rLFABP Alanine
5. Position 97 hLFABP Asparagine and rLFABP Glycine
6. Position 104 hLFABP Leucine and rLFABP Phenylalanine

- Highlighted positions differ across vertebrates:
  - Positions 50F, 54A, 81T, 93T, 97N Evaluated using SeqAPASS

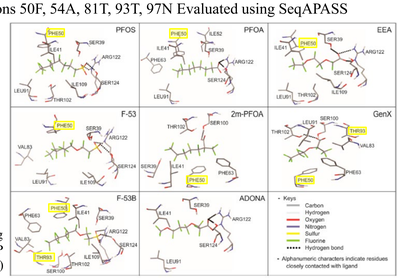


Figure 3. Molecular docking of PFAS with human LFABP (From Cheng and Ng, 2018[3])

Conservation of Liver Fatty Acid Binding Protein: SeqAPASS

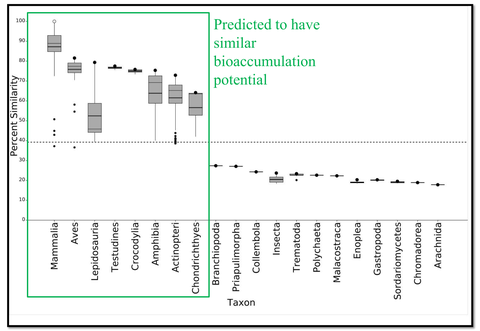


Figure 4. Boxplot depicting SeqAPASS (v4.0) data illustrating the percent similarity across species compared to the primary amino acid sequences for human LFABP

Position Amino Acid	Type 1 Primates, Ruminants, Whales/Dolphins	Type 2 Rodents and other mammals; Fish; Amphibians; Testudines	Type 3 Aves, Lepidosauria, Chondrichthyes	Type 4 Crocodylia	Mutation in DUET	Stability Change from DUET (ΔΔG, kcal/mol)
50	Phenylalanine (F)	Valine (V) Isoleucine (I) Leucine (L)	Valine Isoleucine Leucine	Phenylalanine	F50V F50I F50L	-1.196 (Destabilizing) -0.808 (Destabilizing) -0.893 (Destabilizing)
54	Alanine (A)	Threonine (T)	Threonine	Threonine	A54T	-0.195 (Destabilizing)
81	Threonine (T)	Alanine (A) Glycine (G)	Alanine	Threonine	T81A T81G	-0.749 (Destabilizing) -0.023 (Destabilizing)
93	Threonine (T)	Threonine (T) Valine (V)	Alanine	Alanine	T93A T93V	-1.004 (Destabilizing) 0.031 (Stabilizing)
97	Asparagine (N)	Glycine (G)	Glycine	Glycine	N97G	0.521 (Stabilizing)

Table 2. Amino acid differences across species compared to human LFABP predict different bioaccumulation

"Validate" Predictions With Molecular Modeling and Chemical Proteomics

Conclusions

SeqAPASS Results to Understand Conservation of Assay Target Across Species

- Proteins thought to be important in PFAS bioaccumulation are conserved in vertebrates
  - Serum Albumin (majority of vertebrate taxa conserved) and Liver Fatty Acid Binding Protein
- Proteins involved in bioaccumulation (or lack-thereof) in invertebrates likely not the same as vertebrates
- Different PFAS interact differently with LFABP
  - Different amino acid residues involved
- Amino acid residues that are important for binding PFAS in humans differ across species

References

1. LaLone, C. A., Villeneuve, D. L., Lyons, D., Helgen, H. W., Robinson, S. L., Swintek, J. A., Saari, T.W, Ankley, G. T. (2016) Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicological Sciences* 153(2), 228-245.
2. Doering, J.A., Lee, S., Kristiansen, K., Evenseth, L., Barron, M., Sylte, I., and LaLone, C.A. (2018). In silico site-directed mutagenesis informs species-specific predictions of chemical susceptibility derived from the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool. *Toxicological Sciences* Published online.
3. Cheng, W. and Ng, C.A. (2018) Predicting Relative Protein Affinity of Novel Per- and Polyfluoroalkyl Substances (PFASs) by An Efficient Molecular Dynamics Approach. Environmental Science & Technology 52,7972-7980.