**ADME/Tox WEB *in silico* predictions of longer chain
perfluoroalkyl carboxylic acid pK_a values are more accurate
than other computational methods**

Sierra Rayne^{1*} and Kaya Forest²

¹Department of Chemistry, University of Winnipeg, Winnipeg, Manitoba, Canada

²Department of Chemistry, Okanagan College, Penticton, British Columbia, Canada

*Address correspondence to Sierra Rayne, Department of Chemistry, University of Winnipeg,
Winnipeg, Manitoba, Canada; Phone: +1 204 786 9265, Fax: +1 204 775 2114; E-mail:
rayne.sierra@gmail.com

Because of their environmental and toxicological relevance,^{1,2,3,4,5,6,7,8,9} the acidity constants (pK_a values) of the perfluoroalkyl carboxylic acids (PFCAs; Fig. 1) have been the subject of much recent interest and debate in the literature.^{10,11,12,13,14,15,16} Given the wide variability in pH values within both biological and environmental systems, the accurate knowledge of PFCA pK_a values is required for reliable toxicological and environmental fate and partitioning studies, as well as for designing and optimizing treatment methods for these ubiquitous contaminants.¹⁶ Perfluoroalkyl and carboxyl groups are also widely employed in medicinal chemistry for the creation of new pharmaceuticals. Thus, accurate prediction of carboxylate acidities coupled to perfluoroalkyl moieties is an important goal in the area of drug design.

In previous work, we have shown that the widely used semiempirical PM6 method,^{17,18} SPARC,^{19,20} and COSMOtherm²¹ computational tools all underestimate the pK_a values of longer straight chain PFCAs by up to several pK_a units when compared to the limited experimental database.¹⁶ As well, both the PM6 method and SPARC approaches predict that branching along the perfluoroalkyl chain will influence the acidity of the carboxylate head group. This effect is particularly evident in the PM6 method, which suggested that the branched isomers will be less acidic than their more linear counterparts by up to several pK_a units.¹⁶ By comparison, the SPARC method suggests intrahomologue differences of up to one pK_a unit, but with no discernible structural patterns due to branching on the perfluoroalkyl chain (Table 1).

In contrast to the other computational methods used in the literature to date, we report here that the AB/ pK_a method from the ADME/Tox WEB²² package in the ALOGPS 2.1 program^{23,24,25,26} using the SMILES chemical structure input format^{27,28} accurately predicts the pK_a values of the C_1 (i.e., trifluoroacetic acid; TFA) through C_{10} straight chain PFCAs, including the well known straight chain C_7

PFCA **39** (perfluorooctanoic acid; PFOA) (Table 2), within the range of quoted experimental and computational errors. As well, only a modest difference is observed between the experimental pK_a for the straight chain C_{11} PFCA ($pK_{a,exp}=3.1$) and the upper limit of the AB/pKa range ($pK_{a,est}=2.9$). Unlike the PM6 and SPARC methods, however, the AB/pKa method does not estimate any branching influences on the acidity constants. To resolve whether branching of the perfluoroalkyl chain does appreciably influence the pK_a of the carboxylate head group, authentic standards of representative highly branched PFCA congeners - which are currently unavailable due to synthetic challenges - will need to be experimentally tested. At present, however, it appears that the AB/pKa method is the most reliable general approach for estimating acidity constants of straight chain perfluoroalkyl carboxylic acids.

Acknowledgments

S.R. thanks the Natural Sciences and Engineering Research Council (NSERC) of Canada for financial support.

References

1. Prevedouros, K. et al. Sources, fate and transport of perfluorocarboxylates. *Environ. Sci. Technol.* **40**, 32-44(2006).
2. Fromme, H. et al. Perfluorinated compounds – Exposure assessment for the general population in western countries. *International Journal of Hygiene and Environmental Health* **in press**, 10.1016/j.ijheh.2008.04.007
3. Armitage, J. et al. Modeling global-scale fate and transport of perfluorooctanoate emitted from direct sources. *Environ. Sci. Technol.* **40**, 6969-6975(2006).
4. Washburn, S.T. et al. Exposure assessment and risk characterization for perfluorooctanoate in selected consumer articles. *Environ. Sci. Technol.* **39**, 3904-3910(2005).
5. Trudel, D. et al. Estimating consumer exposure to PFOS and PFOA. *Environ. Sci. Technol.* **28**, 251-269(2008).
6. Houde, M. et al. Biological monitoring of polyfluoroalkyl substances: A review. *Environ. Sci. Technol.* **40**, 3463-3473(2006).
7. Kennedy, G.L. et al. The toxicology of perfluorooctanoate. *Crit. Rev. Toxicol.* **34**, 351-84
8. Conder, J.M. et al. Are PFCAs bioaccumulative? A critical review and comparison with regulatory criteria and persistent lipophilic compounds. *Environ. Sci. Technol.* **42**, 995-1003(2008).
9. Rayne, S., Forest, K. & Friesen, K.J. Congener-specific numbering systems for the environmentally relevant C4 through C8 perfluorinated homologue groups of alkyl sulfonates, carboxylates, telomer alcohols, olefins, and acids, and their derivatives. *J. Environ. Sci. Health A* **43**, 1391-1401(2008).
10. Moroi, Y. et al. Determination of acidity constants of perfluoroalkanoic acids. *Bull. Chem. Soc. Jpn.* **74**, 667-672(2001).
11. Goss, K. The pKa values of PFOA and other highly fluorinated carboxylic acids. *Environ. Sci.*

Technol. **42**, 456-458(2008).

12. Goss, K. The pKa values of PFOA and other highly fluorinated carboxylic acids - Correction. *Environ. Sci. Technol.* **42**, 5032(2008).
13. McMurdo, C.J. et al. Aerosol enrichment of the surfactant PFO and mediation of the water-air transport of gaseous PFOA. *Environ. Sci. Technol.* **42**, 3969-3974(2008).
14. Mader, B.T. Comment on "Aerosol enrichment of the surfactant PFO and mediation of the water-air transport of gaseous PFOA". *Environ. Sci. Technol.* **43**, 1232-1233(2009).
15. Ellis, D.A. & Webster, E. Response to Comment on "Aerosol enrichment of the surfactant PFO and mediation of the water-air transport of gaseous PFOA". *Environ. Sci. Technol.* **43**, 1234-1235(2009).
16. Rayne, S., Forest, K. & Friesen, K.J. Computational approaches may underestimate pKa values of longer-chain perfluorinated carboxylic acids: Implications for assessing environmental and biological effects. *J. Environ. Sci. Health A* **44**, 317-326(2009).
17. Stewart, J. MOPAC2009. *MOPAC2009* at <<http://openmopac.net/>>
18. Stewart, J.J.P. Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry* **10**, 209-220(1989).
19. Hilal, S.H., Karickhoff, S.W. & Carreira, L.A. A Rigorous Test for SPARC's Chemical Reactivity Models: Estimation of More Than 4300 Ionization pKas. *Quantitative Structure-Activity Relationships* **14**, 348-355(1995).
20. Sparc On-Line Calculator. at <<http://ibmlc2.chem.uga.edu/sparc/>>
21. COSMOlogic- COSMO-RS, COSMOtherm, COSMOfrag and Turbomole. at <<http://www.cosmologic.de/>>
22. ADME/Tox WEB. at <<http://www.pharma-algorithms.com/webboxes/>>
23. Tetko, I.V. et al. Virtual computational chemistry laboratory--design and description. *J. Comput.*

Aided Mol. Des. **19**, 453-63(2005).

24. Tetko, I.V. Computing chemistry on the web. *Drug Discov. Today* **10**, 1497-500(2005).
25. Mannhold, R. et al. Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *J. Pharm. Sci.* **98**, 861-93(2009).
26. On-line Lipophilicity(logP), Water Solubility(LogS) & pKa Prediction. at
<<http://www.vcclab.org/lab/alogps/start.html>>
27. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comp. Sci.* **28**, 31-36(1988).
28. Weininger, D., Weininger, A. & Weininger, J.L. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inf. Comp. Sci.* **29**, 97-101(1989).
29. Henne, A.L. & Fox, C.J. Ionization constants of fluorinated acids. *J. Amer. Chem. Soc.* **73**, 2323-2325(1951).
30. Brace, N.O. Long chain alkanoic and alkenoic acids with perfluoroalkyl terminal segments. *J. Org. Chem.* **27**, 4491-4498(1962).
31. Burns, D.C. et al. Experimental pKa determination for perfluorooctanoic acid (PFOA) and the potential impact of pKa concentration dependence on laboratory-measured partitioning phenomena and environmental modeling. *Environ. Sci. Technol.* **42**, 9283-9288(2008).
32. Steinle-Darling, E. & Reinhard, M. Nanofiltration for trace organic contaminant removal: Structure, solution, and membrane fouling effects on the rejection of perfluorochemicals. *Environ. Sci. Technol.* **42**, 5292-5297(2008).

Figure 1. General structure of perfluoroalkyl carboxylic acids (PFCAs).

Table 1. SPARC (v.4.2) estimated pK_a values for the C_1 through C_8 perfluoroalkyl carboxylic acids (PFCAs). Isomer numbering system from ref. ⁹. HG=homologue group in the perfluoroalkyl chain.

HG	Substitution	Isomer	pK_a
C_1	methyl	C1-1	1.14
C_2	ethyl	C2-1	0.08
C_3	n-propyl	C3-1	-0.20
	isopropyl	C3-2	-0.14
C_4	1,1'-dimethylethyl	C4-1	-0.18
	1-methylpropyl	C4-2	-0.18
	2-methylpropyl	C4-3	0.11
	n-butyl	C4-4	-0.06
C_5	1-ethylpropyl	C5-1	-0.22
	1,1'-dimethylpropyl	C5-2	-0.22
	1,2-dimethylpropyl	C5-3	-0.15
	2,2'-dimethylpropyl	C5-4	0.24
	1-methylbutyl	C5-5	-0.32
	2-methylbutyl	C5-6	-0.04
	3-methylbutyl	C5-7	-0.11
	n-pentyl	C5-8	-0.13
C_6	1-ethyl-1'-methylpropyl	C6-1	-0.26
	1-ethyl-2-methylpropyl	C6-2	-0.19
	1,1',2-trimethylpropyl	C6-3	-0.19
	1,2,2'-trimethylpropyl	C6-4	-0.02
	1-ethylbutyl	C6-5	-0.36
	2-ethylbutyl	C6-6	-0.18
	1,1'-dimethylbutyl	C6-7	-0.36
	1,2-dimethylbutyl	C6-8	-0.29
	1,3-dimethylbutyl	C6-9	-0.37
	2,2'-dimethylbutyl	C6-10	0.10
	2,3-dimethylbutyl	C6-11	-0.09
	3,3'-dimethylbutyl	C6-12	0.04
	1-methylpentyl	C6-13	-0.38
	2-methylpentyl	C6-14	-0.10
	3-methylpentyl	C6-15	-0.17
	4-methylpentyl	C6-16	-0.15

HG	Substitution	Isomer	pK _a
C ₇	n-hexyl	C6-17	-0.15
	1,1'-diethylpropyl	C7-1	-0.30
	1-ethyl-1',2-dimethylpropyl	C7-2	-0.23
	1-ethyl-2,2'-dimethylpropyl	C7-3	-0.06
	1-isopropyl-2-methylpropyl	C7-4	-0.16
	1,1',2,2'-tetramethylpropyl	C7-5	-0.07
	1-ethyl-1'-methylbutyl	C7-6	-0.40
	1-ethyl-2-methylbutyl	C7-7	-0.33
	1-ethyl-3-methylbutyl	C7-8	-0.33
	2-ethyl-1-methylbutyl	C7-9	-0.44
	2-ethyl-2'-methylbutyl	C7-10	-0.04
	2-ethyl-3-methylbutyl	C7-11	-0.23
	1,1',2-trimethylbutyl	C7-12	-0.33
	1,1',3-trimethylbutyl	C7-13	-0.41
	1,2,2'-trimethylbutyl	C7-14	-0.17
	1,2,3-trimethylbutyl	C7-15	-0.34
	1,3,3'-trimethylbutyl	C7-16	-0.22
	2,2',3-trimethylbutyl	C7-17	0.05
	2,3,3'-trimethylbutyl	C7-18	0.07
	1-isopropylbutyl	C7-19	-0.33
	1-propylbutyl	C7-20	-0.51
	1-ethylpentyl	C7-21	-0.51
	2-ethylpentyl	C7-22	-0.24
	3-ethylpentyl	C7-23	-0.24
	1,1'-dimethylpentyl	C7-24	-0.42
	1,2-dimethylpentyl	C7-25	-0.35
	1,3-dimethylpentyl	C7-26	-0.43
	1,4-dimethylpentyl	C7-27	-0.40
	2,2'-dimethylpentyl	C7-28	0.04
	2,3-dimethylpentyl	C7-29	-0.15
	2,4-dimethylpentyl	C7-30	-0.12
	3,3'-dimethylpentyl	C7-31	-0.02
	3,4-dimethylpentyl	C7-32	-0.20
	4,4'-dimethylpentyl	C7-33	-0.05

HG	Substitution	Isomer	pK _a
	1-methylhexyl	C7-34	-0.41
	2-methylhexyl	C7-35	-0.12
	3-methylhexyl	C7-36	-0.20
	4-methylhexyl	C7-37	-0.17
	5-methylhexyl	C7-38	-0.17
	n-heptyl	C7-39	-0.16
C ₈	1-ethyl-1',2,2'-trimethylpropyl	C8-1	-0.11
	1-ethyl-1'-isopropylpropyl	C8-2	-0.27
	1-isopropyl-1',2-dimethylpropyl	C8-3	-0.20
	1-isopropyl-2,2'-dimethylpropyl	C8-4	-0.03
	1-tert-butylbutyl	C8-5	-0.21
	1,1'-diethylbutyl	C8-6	-0.44
	1,2-diethylbutyl	C8-7	-0.48
	2,2'-diethylbutyl	C8-8	-0.18
	1-isopropyl-1'-methylbutyl	C8-9	-0.37
	1-isopropyl-2-methylbutyl	C8-10	-0.31
	1-isopropyl-3-methylbutyl	C8-11	-0.38
	2-isopropyl-3-methylbutyl	C8-12	-0.28
	1-ethyl-1',2-dimethylbutyl	C8-13	-0.37
	1-ethyl-1,3-dimethylbutyl	C8-14	-0.45
	1-ethyl-2,2'-dimethylbutyl	C8-15	-0.21
	1-ethyl-2,3-dimethylbutyl	C8-16	-0.38
	1-ethyl-3,3'-dimethylbutyl	C8-17	-0.26
	2-ethyl-1,1'-dimethylbutyl	C8-18	-0.47
	2-ethyl-1,2'-dimethylbutyl	C8-19	-0.31
	2-ethyl-1,3-dimethylbutyl	C8-20	-0.49
	2-ethyl-2',3-dimethylbutyl	C8-21	-0.09
	2-ethyl-3,3'-dimethylbutyl	C8-22	-0.07
	1,1',2,2'-tetramethylbutyl	C8-23	-0.21
	1,1',2,3-tetramethylbutyl	C8-24	-0.38
	1,1',3,3'-tetramethylbutyl	C8-25	-0.25
	1,2,2',3-tetramethylbutyl	C8-26	-0.21
	1,2,3,3'-tetramethylbutyl	C8-27	-0.19
	2,2',3,3'-tetramethylbutyl	C8-28	0.21

HG	Substitution	Isomer	pK _a
	1-methyl-1'-propylbutyl	C8-29	-0.54
	2-methyl-1'-propylbutyl	C8-30	-0.48
	3-methyl-1'-propylbutyl	C8-31	-0.56
	1-ethyl-1'-methylpentyl	C8-32	-0.46
	1-ethyl-2-methylpentyl	C8-33	-0.40
	1-ethyl-3-methylpentyl	C8-34	-0.47
	1-ethyl-4-methylpentyl	C8-35	-0.44
	2-ethyl-1-methylpentyl	C8-36	-0.50
	2-ethyl-2'-methylpentyl	C8-37	-0.50
	2-ethyl-3-methylpentyl	C8-38	-0.29
	2-ethyl-4-methylpentyl	C8-39	-0.26
	3-ethyl-1-methylpentyl	C8-40	-0.49
	3-ethyl-2-methylpentyl	C8-41	-0.21
	3-ethyl-3'-methylpentyl	C8-42	-0.08
	3-ethyl-4-methylpentyl	C8-43	-0.26
	1-isopropylpentyl	C8-44	-0.40
	2-isopropylpentyl	C8-45	-0.29
	1,1',2-trimethylpentyl	C8-46	-0.39
	1,1',3-trimethylpentyl	C8-47	-0.47
	1,1',4-trimethylpentyl	C8-48	-0.44
	1,2,2'-trimethylpentyl	C8-49	-0.23
	1,2,3-trimethylpentyl	C8-50	-0.40
	1,2,4-trimethylpentyl	C8-51	-0.38
	1,3,3'-trimethylpentyl	C8-52	-0.28
	1,3,4-trimethylpentyl	C8-53	-0.45
	1,4,4'-trimethylpentyl	C8-54	-0.31
	2,2',3-trimethylpentyl	C8-55	-0.01
	2,2',4-trimethylpentyl	C8-56	0.01
	2,3,3'-trimethylpentyl	C8-57	-0.17
	2,3,4-trimethylpentyl	C8-58	-0.03
	2,4,4'-trimethylpentyl	C8-59	0.02
	3,3',4-trimethylpentyl	C8-60	-0.04
	3,4,4'-trimethylpentyl	C8-61	-0.10
	1-propylpentyl	C8-62	-0.57

HG	Substitution	Isomer	pK _a
	2-propylpentyl	C8-63	-0.30
	1-ethylhexyl	C8-64	-0.45
	2-ethylhexyl	C8-65	-0.27
	3-ethylhexyl	C8-66	-0.26
	4-ethylhexyl	C8-67	-0.20
	1,1'-dimethylhexyl	C8-68	-0.45
	1,2-dimethylhexyl	C8-69	-0.38
	1,3-dimethylhexyl	C8-70	-0.46
	1,4-dimethylhexyl	C8-71	-0.43
	1,5-dimethylhexyl	C8-72	-0.42
	2,2'-dimethylhexyl	C8-73	0.01
	2,3-dimethylhexyl	C8-74	-0.17
	2,4-dimethylhexyl	C8-75	-0.14
	2,5-dimethylhexyl	C8-76	-0.14
	3,3'-dimethylhexyl	C8-77	-0.05
	3,4-dimethylhexyl	C8-78	-0.22
	3,5-dimethylhexyl	C8-79	-0.21
	4,4'-dimethylhexyl	C8-80	-0.08
	4,5-dimethylhexyl	C8-81	-0.19
	5,5'-dimethylhexyl	C8-82	-0.13
	1-methylheptyl	C8-83	-0.42
	2-methylheptyl	C8-84	-0.14
	3-methylheptyl	C8-85	-0.21
	4-methylheptyl	C8-86	-0.19
	5-methylheptyl	C8-87	-0.18
	6-methylheptyl	C8-88	-0.17
	n-octyl	C8-89	-0.17

Table 2. Experimental and estimated pK_a values for the C₂ through C₁₂ perfluoroalkyl carboxylic acids (PFCAs). Note that the homologue designation in the compound names (i.e., C₁, C₆, etc.) designate the number of carbons in the perfluoroalkyl chain, and do not include the carbon in the carboxylate head group (e.g., C₇ PFCA is perfluorooctanoic acid).

Compound	Experimental	AB/pK _a (current work)	SPARC v.4.2 (current work)	PM6 ^f	SPARC ^{g,h}	SPARC ⁱ	COSMOtherm ^{g,h}
linear-C ₁ PFCA (TFA)	0.3 to 0.6 ^{a,b,c}	0.6±0.5	0.98	-0.2	1.4		0.9
linear-C ₂ PFCA	0.5±0.1 ^b	0.6±0.5	0.08	-0.2			
linear-C ₃ PFCA	0.2 to 0.5 ^{b,c}	0.6±0.5	-0.20	-0.1	0.4		0.7
linear-C ₄ PFCA	0.5±0.1 ^b	0.6±0.5	-0.06	-0.1		-0.1	
linear-C ₅ PFCA	0.8±0.1 ^b	0.6±0.5	-0.13	-0.1		-0.2	
linear-C ₆ PFCA	n/a	2.4±0.5	-0.15	-0.1			
linear-C ₇ PFCA (PFOA)	2.8 to 3.8 ^{d,e}	2.4±0.5	-0.16	-0.1	-0.1	-0.2	0.7
linear-C ₈ PFCA	n/a	2.4±0.5	-0.17	-0.1			
linear-C ₉ PFCA	2.6±0.1 ^b	2.4±0.5	-0.17	-0.1		-0.2	
linear-C ₁₀ PFCA	2.6 ^b	2.4±0.5	-0.17	-0.1		-0.2	
linear-C ₁₁ PFCA	3.1 ^b	2.4±0.5	-0.17	-0.1	0.2	-0.2	0.8

^a ref. 1^b ref. 10^c ref. 29^d ref. 30^e ref. 31^f ref. 16^g ref. 11^h ref. 12ⁱ ref. 32.

Nature Precedings : hdl:10101/npre.2009.2936