

CHEMISTRY



On the Number of Phthalates

MATTHEW S. MACLENNAN¹

1. University of British Columbia

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:

msmaclennan@gmail.com

DATE RECEIVED:

August 15, 2015

DOI:

10.15200/winn.144348.80403

ARCHIVED:

September 28, 2015

KEYWORDS:

phthalates

CITATION:

Matthew S. MacLennan, On the Number of Phthalates, *The Winnower* 2:e144348.80403, 2015, DOI: 10.15200/winn.144348.80403

© MacLennan This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



ON THE NUMBER OF PHTHALATES

Recently, [four cover stories of C&EN magazine](#) were dedicated to the class of molecules called phthalates. Phthalate esters of fatty alcohols are generally used in industry as plasticizers, turning PVC into a more malleable form which comprises many toys for young children. There have been studies of the links between exposure to phthalate plasticizers and antiandrogenic effects in humans ([do a PubMed search for phthalates](#)). Reports are coming out about its [reproductive effects on females](#). So, the race is on to develop the next "healthier" plasticizer.

In the article "[Regulators And Retailers Raise Pressure On Phthalates](#)", I read the line:

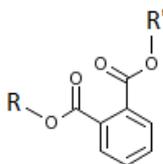
"The number of alcohol-acid combinations that are possible and their resulting properties are as endless as the number of applications."

I understood this line as a (friendly) challenge and took it upon myself to count the number of phthalates theoretically possible.

My parents always told me I took things too literally.

General Structure of Industrial Phthalates

The general structure of phthalates appears straightforward: They are hydrocarbon (limited here to C1 to C13) diesters of phthalic acid.^{1,2} I will only consider linear and branched alkyl chains here.



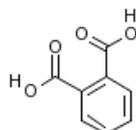
R, R' = C1-C13

COUNTING PHTHALATES

Enumerating phthalates begins with enumerating the phthalate scaffolds, then counting the alcohols that react with the scaffolds and then finding the combinations of alcohols on a phthalate.

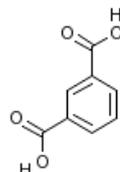
There are three phthalate scaffolds

There are only three major phthalate scaffolds in use: ortho-, meta-, and para-phthalic acid--or in other words: phthalic, isophthalic and terephthalic acids.^{1,2} They are shown below along with SMILES strings written from the functional perspective.



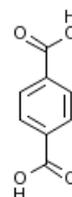
phthalic acid

c1c2c3ccc1.C2(=O)O.C3(=O)O



isophthalic acid

c1c2cc3cc1.C2(=O)O.C3(=O)O



terephthalic acid

c1c2ccc3c1.C2(=O)O.C3(=O)O

Counting the alcohol chains is much more difficult. Luckily, it has already been done.

These number of uniquely connected alcohols can be extracted from the set of all $C_nH_{2n+2}O$ isomers in SMILES notation³⁻⁵(which includes alcohols and ethers) using the regular expression `".*O\\).*|. *O$"` (double backslash is the regex escape character for R programming language).⁶ This regular expression extracts the full SMILES strings which represent alcohols.

The same numbers are described by the series of the number of n -carbon alkyl radicals C_nH_{2n+1} ignoring stereoisomers and can be found at [A000598](#).

If one considers stereochemistry, then the number of unique alcohols increases. For x chiral centres, there are 2^x stereoisomers. The total number of unique alcohols generated here is described by the series at this link [A000625](#).

The number of *chiral* alcohols is another important number to consider. These numbers have also been described by the series of "chiral planted trees" found at [A005628](#).

Table 1 shows the numbers.⁷

Table 1. All unique alcohols and some subclasses.

Carbon Number	Number of alcohols ignoring stereochemistry A000598	Total number of unique alcohols A000625	Number of chiral alcohols A005628
1	1	1	0
2	1	1	0
3	2	2	0
4	4	5	2
5	8	11	6
6	17	28	20
7	39	74	60

8	89	199	176
9	211	551	510
10	507	1553	1484
11	1235	4436	4314
12	3057	12832	12624
13	7639	37496	37126

Now we need to see what happens when we take these alcohols and react them with phthalic acid: how many unique combinations are mathematically possible?

Number of phthalates with identical alcohol chains

The possible number of molecules in this class will be equivalent to the total possible number of alcohol chains. There are 57,189 molecules in this class. This is also the number of possible singly esterified phthalic acids--likely by-products.

Number of phthalates with identical length alcohol chains but possibly different identity

The possible number of molecules in this class will be equivalent to the possible combinations within each alcohol chain carbon number class. This will be equal to $\sum_{i=1}^n \left[\frac{n!}{(n-2)!2!} + n \right]$ where n equals the number of alcohols in each carbon number class.

Number of phthalates with any alcohol side chain

This class of phthalates will contain $\frac{n!}{(n-2)!2!} + n$ members, just like the situation for identical length alcohol chains, but in this case, n = 57,189.

Number of phthalate meso compounds

Because phthalates have a mirror plane which cuts between the ester linkages, it is possible to have meso compounds of phthalate esters. Meso compounds of phthalate esters are phthalate esters whose alcohol side chains bear one or more chiral centres, but the phthalate itself is not chiral. In other words, a meso compound phthalate ester is actually superposable on its mirror image, even though there are chiral centres within the molecule. This is in contrast with *enantiomers* whose mirror images are not superposable making them unique compounds.

Meso compounds of phthalate esters are interesting because they will not contribute much to the optical properties of phthalate mixtures. Therefore, from the analytical perspective, their presence is particularly important to consider. Phthalate ester meso compounds must have two of the same alcohol chains and must have opposite chirality at each chiral centre.

For every chiral alcohol chain, there is a match for opposite chirality. Therefore, the total number of meso compounds is half the total number of chiral alcohols. The series of meso compounds with n carbons in each linear or branched alkyl chain has been submitted by me and has been recently approved at <http://oeis.org/A261336>.

Summary - Some grand totals

It is important to note that even though many of these alcohol chains are mathematically possible, this does not suggest that these alcohol chain geometries actually exist! Incorporating *a priori* knowledge concerning reactants and products will filter out these numbers

to make them smaller. Therefore, one can interpret the numbers listed here as absolute upper bounds to the real number of phthalates.

0.0035% of all phthalates with identical length chains are meso compounds. 0.00172% of all possible phthalates with alcohol chains between C1 and C13 are meso compounds. Table 2 summarizes some numbers.

Table 2. Summary of the possible number of special classes of unique phthalates (stereochemistry included)

Number of possible phthalates with identical alcohol chains	57,189
OR	
Number of unique singly-esterified phthalates	
Number of possible phthalates with identical length alcohol chains but possibly different identity (isomeric differences, non-identical)	796,553,474
Number of possible phthalates with any alcohol side chain (all combinations)	1,635,319,455
Number of possible phthalate meso compounds	28,161

The number of singly esterified phthalates equals the number of number of possible phthalates with identical alcohol chains: 57,189.

If you were to store every phthalate as a SMILES string of 50 bytes (which is a little on the large side), it would take 76 GB of space without any compression.

I went back to the [C&EN article](#) and read the next line,

"Most phthalates, however, never make it into commercial products because of concerns about performance, cost, availability, or toxicity."

I thought two things:

1. "I take things too literally" and
2. "I guess I can reduce these numbers a little"

REFERENCES

1. "GIF/PNG-Creator for 2D Plots of Chemical Structures". <http://cactus.nci.nih.gov/gifcreator/>, NCI/CADD Group.
2. *J. Chem. Inf. Comput. Sci.* (1983) 23, 61-65. <http://dx.doi.org/10.1021/ci00038a002>
3. M. Shimizu, H. Nagamochi and T. Akutsu, **Enumerating tree-like chemical graphs with given upper and lower bounds on path frequencies**, *BMC Bioinformatics*, vol. 12, Suppl 14, S3, 2011. <http://dx.doi.org/10.1186/1471-2105-12-S14-S3>
4. Y. Ishida, Y. Kato, L. Zhao, H. Nagamochi and T. Akutsu, **Branch-and-bound algorithms for enumerating treelike chemical graphs with given path frequency using detachment-cut**, *Journal of Chemical Information and Modeling*, vol. 50, no. 5, pp. 934-946, 2010. <http://dx.doi.org/10.1021/ci900447z>
5. H. Fujiwara, J. Wang, L. Zhao, H. Nagamochi and T. Akutsu, **Enumerating treelike**

chemical graphs with given path frequency, *Journal of Chemical Information and Modeling*, vol. 48, no. 7, pp. 1345-1357, 2008. <http://dx.doi.org/10.1021/ci700385a>

6. R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

7. David B. Dahl (2014). xtable: Export tables to LaTeX or HTML. R package version 1.7-4. <http://CRAN.R-project.org/package=xtable>.