Supplementary Material S4: overview of the GC parameter values developed in the present and the previous [7-9] three papers

This document holds the formula and the individual parameters which should allow the evaluation of the heat of formation of gaseous organic molecules at 298 K using the Group Contribution method and the GC parameter values developed in the present and the previous [7-9] three papers. Up till now, we have reparametrized the GC parameters to cover a certain limited range of molecules, but for those essentially chemical accuracy can be achieved (= the prediction is within 1 kcal/mol from the experimental value). Still, it is highly recommended to consult the published papers in order to understand the limitations of the current version of the approach, and to apply the current parameters to molecules for which chemical accuracy can be achieved.

We have been using the simplest possible form of the GC approach which reads

$$\Delta H_f \ = \ \sum_{j=1,N} N_j . \Delta H_f (j) \qquad\qquad (1)$$

in which ΔH_f , the heat of formation of the molecule, is the sum (j runs from 1 till N if N different Groups are present, and N_j is the number of times Group j is present in the molecule. When we take the linear n-alkanes as an example, equation 1 becomes

$$\Delta H_f \text{ (n-alkanes)} = 2 * GC_{CH3} + N_{CH2} * GC_{CH2}$$

as there are two CH_3 end-groups and N_{CH2} Groups in between these end-groups.

We have determined values for a variety of different Groups, which are listed further below. Formula 1 can be applied to evaluate the ΔH_f for molecules consisting of (a subset) these Groups. However, and in a way we repeat ourselves now, this is absolutely crucial and practically nowhere stated this explicitly in other papers on this topic, one should only apply the formula for the type of molecules for which we have shown they work well. We have seen they do not work well for highly congested alkanes, but we have not yet investigated whether it works well for molecules with neighbouring C=O and NH_2 groups as in an amide. For some Groups we found a difference in parameter values for terminal versus non-terminal position, e.g. the OH Group.

It is also to be recognised that there are exceptions, e.g. due to the presence of germinal interactions as in malononitrile and butanedinitrile. These species should be treated as individual Groups by themselves in order to achieve chemical accuracy. The same holds for the chloromethane family, but as the experimental values are available there is no problem whatsoever. So despite that we have listed the individual Groups parameter values below, it is highly recommended to check in the original publications whether a molecule is a special case, and for good reasons.

CH3	CH2	CH	C	terminal OH	non-terminal OH	NH2	O=C aldehyde	O= keto	COOH
-42.36	-20.63	-4	-1	-171	-183.5	13	-124	-133	-391

C=C-	C≡C	C≡N	Cl non-terminal	(H3)COC-R	phenyl	naphthalene
62.5	229	116	-57	-175	84.5	151.8

C=C=C-	trans R-C=C-R'	C=C(C)-R	Me-O-Me	R'-COC-R	R-COC-R'R"	RR'-COC-R"R'"
205	73.5	70	-184.1	-168	-146	-149

cis R-C=C-R'	MeOCRR'	MeOC(t-Bu)	RRR-COC-RRR	NO2 terminal	NO2 non-terminal
78	-168	-155	-107	-38	-50

adjacent C: Cl-C-C-Cl	terminal Cl	terminal CHCl2	CCl3	F terminal	F non-terminal	C=C in ring
-120.5	-50.5	-86	-100	-218.5	-232.5	71.5

1,3-dioxolane	pyridine	quinoline
-301.6	142	197

Next to formula 1 we have assigned a number of additional parameters which are basically nearest or next-nearest neighbour interaction based. Nearest and next-nearest neighbour interaction parameters are very common in physics and intend to cover real interactions and not just for the purpose of fitting data to achieve a good result. Examples include methyl-methyl interactions when attached to an alkyl:
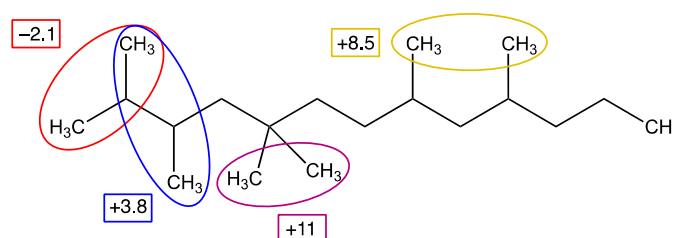
For 2,3-dimethylhexane this leads to

ΔH_f (2,3-dimethylhexane) = 4 * GC $_{CH3}$ + 2 * GC $_{CH2}$ + 2 * GC $_{CH}$ -2.1 + 3.8 kJ/mol

For benzenes we have determined correction parameters for mono- (+6 kJ/mol), di- (+18.5 kJ/mol), tri-substituted (+30 kJ/mol) and tetra- (+40 kJ/mol) substituted alkylbenzenes. Thus, for example

ΔH_f (1,3-dimethylbenzene) = 2* GC $_{CH3}$ + 1 * GC $_{phenyl}$ +18.5 kJ/mol.

For the presence of an ethyl Group added onto an alkyl chain an interaction energy contribution of +6 kJ/mol needs to be added, so for 3-ethylhexane we have

ΔH_f (3-ethylhexane) = 3 * GC $_{CH3}$ + 4 * GC $_{CH2}$ + 1 * GC $_{CH}$ + 6 kJ/mol

When a benzene is alkyl substituted we need to add a correction depending on the degree of substitution. Correction factors alkyl substituted benzenes:

AromMonoalkyl	AromDialkyl	AromTrialkyl	AromTetraalkyl
6	18.5	30	40

The cycloalkanes and cycloalkenes suffer from ring strain which is a non-additive quantity.
The general formula to evaluate the GC heat of formation reads

ΔH_f (cycloalkane) = N $_{CH2}$ * GC $_{CH2}$ + ring strain

For cyclopropanes, also the methyl-substituted ones, a ring strain related energy correction of 115 kJ/mol is to be added.

For cyclobutanes, add a ring strain of 102.5 kJ/mol for the substituted cyclobutanes, and 110 kJ/mol for cyclobutane.

For the ring strain in cyclopentanes it is sufficient to adopt, as for the cyclopropanes, a single value for all cyclopentanes, for the cyclopentanes its value is 26.5 kJ/mol.

For the cyclohexanes a constant ring strain of -2 kJ/mol (note it is indeed a minus sign!) is to be added. Agreement within chemical accuracy between model and experiment for di-and tri-methylsubstituited cyclohexanes is obtained when adopting -3.6 kJ/mol for the ring strain.

For the chloromethanes one should adopt the experimental values. For the mono-chloro GC parameters were given in the Table above. However, for 1,1-dichloro and 2,2-dichlroro we need to add found an additional correction of 16.6 kJ/mol and for 1,2-dichloro a correction (= nearest neighbour interaction energy parameter) of +6 kJ/mol. As an example, hexachloroethane has two CCl$_3$ groups, and within each of these groups we find three 1,1-interaction, and between the two CCl$_3$ groups we have 3 x 3 = 9 1,2-interactions. Herewith we can now calculated the heat of formation of hexachloroethane

ΔH_f (hexachloroethane) = 2 * GC $_C$ + 6 * GC $_{Cl\ terminal}$ + 6 * 16.6 +9 * 6   (kJ/mol)
leading to a difference between model and experiemental values of 3.2 kJ/mol.

For fluoro hydrocarbons we find similar trends as for the chloroalkanes. For the mono-substituted it is all straightforward, but for e.g. 1,1,-difluoro alkanes an additional 1,1-interaction energy of -17.5 kJ/mol was established. For 1,1,1-trifluoro we have three times a 1,1-interaction and thus 3 x -17,.5 = -52.5 kJ.mol correction due to neighbour interactions, and this leads to very good agreement with the experimental heat of formation (difference 2.7 kJ/mol).

The current model does not (yet) provide acceptable results for aromatic fluoro species. Aromatic chloro compounds were not included yet at all, and therefore the current method should

Also for dinitro compounds with the nitro groups as neighbours (e.g. 1,1-, 1,2- and 2,2-dinitro) an additional interaction energy correction needs to be added. Due to a lack of sufficient data, the magnitude could, thus far, only be set for the 1,1-dinitro species (for further details see Ref. 9).

For the methyl-alkylethers and di-alkylethers (see Table below) we found that an ether group GC parameter value can be established but the actual numerical value depends on the COC valence angle as illustrated in the Tables below. For other similar ethers one needs to evaluate the COC valence angle using a quantum mechanical calculation, which is unfortunately unavoidable according to our current model, but this leads to excellent results within chemical accuracy.

Methyl-alkyl-ethers	Verevkin 2002	model dHf	model-exp	ABS (model-exp)	ether group constitution	GC value ether group	COC valence angle
dimethylether	-184.1	-184.1	0.00	0.00	Me-O-Me	-184.1	112.7
methyl ethyl ether	-216.4	-217.36	-0.96	0.96	Me-O-C-R	-175	113.1
methyl propyl ether	-238.4	-237.99	0.41	0.41	Me-O-C-R		113.1
methyl n-butyl ether	-258.3	-258.62	-0.32	0.32	Me-O-C-R		113.1
methyl decyl ether	-381.1	-382.4	-1.30	1.30	Me-O-C-R		113.1
methyl isopropyl ether	-252	-252.72	-0.72	0.72	Me-O-CRR'	-168	115.1
methyl t-butylether	-283.4	-282.08	1.32	1.32	Me-O-CRR'R''	-156	118.4
methyl t-amylether	-301.1	-302.71	-1.61	1.61	Me-O-CRR'R''		118.7
averaged absolute difference				0.83			

Di-alkyl ethers	Verevkin 2002	model dHf	model-exp	ABS (model-exp)	ether group constitution	GC value ether group	COC valence angle
diethylether	-252.1	-252.72	-0.62	0.62	R-COC-R'	-168	113.5
ethyl propyl ether	-272.4	-273.35	-0.95	0.95	R-COC-R'		113.5
ethyl butyl ether		-293.98			R-COC-R'		113.5
di-n-propylether	-293.1	-293.98	-0.88	0.88	R-COC-R'		113.5
di-n-butylether	-332.9	-335.24	-2.34	2.34	R-COC-R'		113.7
di-n-pentylether	-380.4	-376.5	3.90	3.90	R-COC-R'		113.4
ethyl t-amylether	-333.5	-336.07	-2.57	2.57	R-COC-R'R''R'''	-146	119.4
butyl t-amylether	-375.7	-377.33	-1.63	1.63	R-COC-R'R''R'''		119.1
ethyl t-butylether	-316.8	-315.44	1.36	1.36	R-COC-R'R''R'''		118.8
propyl t-butylether	-339.3	-336.07	3.23	3.23	R-COC-R'R''R'''		118.7
n-butyl t-butylether	-360.1	-356.7	3.40	3.40	R-COC-R'R''R'''		118.6
amyl t-butylether	-380.6	-377.33	3.27	3.27	R-COC-R'R''R'''		118.6
di-i-propylether	-319.4	-318.44	0.96	0.96	RR'-COC-R''R'''	-149	116
di-sec-butylether	-361.3	-359.7	1.60	1.60	RR'-COC-R''R'''		116.5
t-butyl s-butylether	-379	-381.43	-2.43	2.43	RR'-COC-R''R'''R''''	-149	119.8
t-butyl i-propylether	-360.1	-360.8	-0.70	0.70	RR'-COC-R''R'''R''''		119.9
t-butyl i-butylether	-367.9	-364.8	3.10	3.10	RR'-COC-R''R'''R''''		119
di- t-butylether	-361.2	-361.16	0.04	0.04	tBU-COC-tBu	-107	128
averaged absolute difference				1.94			