

Article

# Group Contribution Revisited: The Enthalpy of Formation of Organic Compounds with “Chemical Accuracy” Part III

Robert J. Meier 

Pro-Deo Consultant, 52525 Heinsberg, North-Rhine Westphalia, Germany; r.meier@planet.nl

**Abstract:** Group contribution (GC) methods to predict thermochemical properties are eminently important to process design. We report on a GC parametrization for the heat of formation of organic molecules exhibiting chemical accuracy, i.e., a maximum 1 kcal/mol (4.2 kJ/mol) difference between experimental and model values, whilst having a minimum number of parameters to avoid overfitting. We report an extension of recent findings to chloro-alkanes, fluoro-hydrocarbons, benzylhalides, nitro-alkanes, and acetals. Compared to the existing literature, we obtained a superior model exhibiting chemical accuracy, with exceptions when the inherent GC assumption on linearity and additivity is not valid. Moreover, to have a reliable method and not only a low absolute average deviation as reported in most publications, we accepted no or exceptionally few outliers. The example of the 1,3-dioxolane acetals revealed that by adopting the appropriate size of a group representing the acetal leads to a model showing good accuracy. The overall conclusion of the three papers on this topic is that it is feasible to achieve chemical accuracy when using high-quality experimental data and the judicious definition of chemical groups. Despite the GC method being old, the present work shows substantial and necessary increase in performance can still be achieved.

**Keywords:** enthalpy of formation; thermodynamics; process design; physico-chemical property prediction; group-contribution method; chemical accuracy; quantum chemistry



**Citation:** Meier, R.J. Group Contribution Revisited: The Enthalpy of Formation of Organic Compounds with “Chemical Accuracy” Part III. *AppliedChem* **2022**, *2*, 213–228. <https://doi.org/10.3390/appliedchem2040015>

Academic Editor: Jason Love

Received: 21 October 2022

Accepted: 9 November 2022

Published: 11 November 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

For the purpose of the evaluation of the enthalpy of formation  $\Delta H_f$  of organic molecules from their molecular structure, the group contribution (GC) approach is one of the most important and applied methods. In this work, the  $\Delta H_f$  (in this paper also indicated as  $dH_f$ ) is the heat of formation for the gaseous phase at 298.15 K. The original GC method [1] is based on the assumption that a molecule can be decomposed in molecular fragments that are in essence mutually independent, and the molecular property of interest is the sum of the individual properties of the molecular fragments  $j$ :

$$\Delta H_f = \sum_{j=1,N} N_j \cdot \Delta H_f(j) \quad (1)$$

Equation (1) is the general formula for a GC method with chemical groups present and where only the first order term  $\sum_i N_i C_i$  has been retained.  $\Delta H_f(j)$  is the contribution to the heat of formation associated with Group  $j$ , and  $N_j$  is the number of times this group is present in the molecule. We hereby essentially follow the original Van Krevelen–Chermin approach [1].

In the course of time, further studies employing the GC methodology to evaluate the heat of formation of organic molecules have been reported. These include the works by Benson and co-workers [2,3], Joback and Reid [4], Constantino and Gani [5], Marrero and Gani [6], Hukkerikar et al. [7], and Kadda et al. [8]. Recently, an AI study based on neural networks was reported [9]. Although there are previous parametrizations with a rather good averaged absolute difference between model and experimental values, sometimes claiming chemical accuracy [7], there were very clear and sometimes very substantial

outliers (see [10,11] for explicit examples). The latter is directly linked to reliability of a method and therefore highly undesired in process design. Quite a few of the more recent publications share the feature of overfitting, which arises from a comparatively large number of parameters compared to the amount of experimental data [6–9]; this reduces reliable predictivity and is a most definite issue. Finally, the choice for higher-order GC parameters [6–8] was generally not driven by a physico-chemical understanding but an attempt to obtain better values, which is most likely the reason why serious outliers are rather common. Finally, in most cases, a mix of experimental data from a larger variety of sources was used, leading to a non-consistent data set with relative errors larger than the chemical accuracy we are after. For example, the experimental values taken from the DIIIPR data base and employed in a recent AI study [9] deviate substantially from the experimental data from the highly consistent studies due to Rossini employed in [11].

We recently reported a renewed attempt to parametrize the GC, i.e., determine the numerical values for the  $\Delta H_f(j)$  [10,11]. The reason for this was multifold. First of all, our aim was to establish a GC approach with chemical accuracy, i.e., 1 kcal/mol (4 kJ/mol). This is quite a challenge and mostly unrealized up until now, but the criterion is not unrealistically tight because the practical goal should be the prediction of chemical equilibrium  $\Delta G = \Delta H - T\Delta S$  of a chemical reaction such as  $A + B \rightarrow C + D$ , where an error of 1 kcal/mol in each of the individual heats of formation can lead to a cancellation of errors but could also add up to a total error of up to 4 kcal/mol. As there are quite some reactions for which  $\Delta G \sim 0$ , we do require highly accurate predictions within chemical accuracy. It needs to be accepted, however, that experimental data also have an error, and this error often goes up to several kJ/mol and is therewith comparable to the required chemical accuracy. To avoid the problems mentioned in earlier works [6–9], we decided to place specific emphasis on the use of highly reliable experimental data only important with respect the 1 kcal/mol criterion. Furthermore, we take into account common well-known chemical knowledge such as ring strain and geminal effects that cannot be considered as simple linear terms as in a GC model. Overfitting was avoided by carefully analyzing *individual* experimental data and only allowing for an absolute minimum number of additional GC parameters exclusively related to specified (next-)nearest neighbor interactions.

In two previous papers [10,11], we accounted for the n-alkanes; mono-, di-, tri- tetra-, and penta-substituted methyl-alkanes; ethyl alkanes; various classes of alkenes; alkynes; primary and secondary alcohols; n-alkylamines; n-aldehydes; methyl- and di-alkyl-ethers; 2-alkanones; mono- and di-carboxylic acids; dienes; mono- and di-nitriles; alkyl-substituted benzenes; alkyl-substituted naphthalenes; and alkyl-substituted cycloalkanes. Not only averaged absolute deviations but also individual results were within chemical accuracy, except for some more heavily alkyl-substituted molecules for which the group contribution approach breaks down. In this third and as envisioned last paper, we extend our previous work with chloro-alkanes, fluoro-hydrocarbons, benzylhalides, nitro-alkanes, and 1,3-dioxolane acetals.

## 2. Methods

### *Experimental Data and Computational Methods*

As the GC method is a so-called data-driven model, with experimental data being used to parametrize the model, our self-imposed requirement on chemical accuracy makes the quality and accuracy of the experimental data preeminent. Consequently, we almost exclusively used literature experimental data from a few sources involving high-quality expertise, and always the same measuring equipment, measuring protocol and data processing. Experimental errors are commonly around 1–1.5 kJ/mol (see, e.g., [12]), although for some species, the error is indicated as being larger. Therefore, we need to take into account that an experimental value could be off the true value up to half the value of the chemical accuracy we want to achieve. It goes without saying that this will have a certain impact on the quality of parametrization but it is, unfortunately, unavoidable. As we have argued before, we decided to determine the numerical values of the group contributions

by hand and group by group. As we also see in the present paper, this leads to very good models that are generally within chemical accuracy. Performance of the parameter estimation is verified by calculating the differences between model and experimental values, and in addition calculating the averaged absolute differences (AAD) per class of molecules expressed by

$$\text{AAD} = (1/N) \sum_{j=1,N} (\text{model-experiment}) \quad (2)$$

Whenever referred to, heats of formation and group contributions from the Marrero–Gani method [6] and the method of Joback and Reid [4] were obtained from the implementation in the ICAS23 software suite [13].

In specific cases, we used the density functional theory (DFT)-type quantum chemical (QC) calculations. Quantum chemical calculations were performed using the Spartan 10 program suite [14] involving full geometry optimization involving the B3LYP functional. This functional is well known in the quantum chemistry world as one that describes structures of standard organic molecules well, and also relative energies are well-accounted for [15]. Herewith, we can verify energy differences between similar species in order to verify whether the experimental heat of formation differences can be substantiated independently. In addition, we can investigate whether certain effects, seen in the experimental data set, are realistic and not due to an error in single experimental data points (which is in essence a consistency check). Thus, the ab initio results are part of a check and do not directly influence the value of the GC parameters as determined in this work.

### 3. Results

#### 3.1. Chloro Hydrocarbons

In the next sections, we discuss fluorinated and chlorinated species and thus determine group parameters for F and Cl. Of course, this has been done before (see e.g., [4,7]), but we need to reevaluate because on the one hand, we want to achieve chemical accuracy, and on the other hand, the new parameters must be used in combination with the parameter values for other groups, particularly alkyl-related groups, we established in our previous works.

When we first look at the chloromethanes, and in particular when we look at  $\text{CCl}_4$ , this is a molecule on itself and cannot be not part of a larger entity, whereas  $\text{CCl}_3$  can, and we do have an experimental value for its heat of formation. Therefore, a model does not necessarily have to predict its  $\text{dH}_f$  very well, particularly when some sort of difficulty arises, which was indeed the case. When we look at the chlorinated methanes, i.e.,  $\text{CH}_3\text{Cl}$ ,  $\text{CH}_2\text{Cl}_2$ ,  $\text{CHCl}_3$ , and  $\text{CCl}_4$ , and recognizing that a typical value for the group contribution for a Cl atom is, somewhat roughly, around  $-50$  kJ/mol, the experimental values [16] for these four species (see Table 1) immediately reflect that we have an issue here: tetrachloromethane ( $\text{CCl}_4$ ) has an almost identical (at least within experimental accuracy) heat of formation as dichloromethane, so it is no surprise if a GC model does not predict very well. At least qualitatively, this result is not surprising, as chlorine is a large atom, and steric and electronic effects can be the origin of non-linear behavior. In this case, the quantitative effect is very significant. At this point, it is relevant to question the reason as to why we need a GC method. The obvious answer would be to predict property values, which, however, is only useful when we do not have reliable experimental values. For the chloromethanes, we do have such values, and therefore it is not a real problem that the GC approach does not work for the chloromethanes and we will simply adopt the individual experimental values.

**Table 1.** Experimental values for the heat of formation of chloromethanes in the gas phase.

Chloromethanes	Manion [16]
$\text{CH}_3\text{Cl}$	−81.9
$\text{CH}_2\text{Cl}_2$	−95.1
$\text{CHCl}_3$	−102.9
$\text{CCl}_4$	−95.6

We now turn to the monochlorinated alkanes, i.e., species with isolated chlorine atoms. For the monochloroalkanes, we unfortunately had but a single experimental value from Manion [16], and we therefore adopted additional values from various literature sources (see Table 2). Because the values for chloroethane taken from Manion [16] and Fletcher and Pilcher [17] are very close ( $-111.2$  and  $-112.3$  kJ/mol respectively), we felt confident in adopting the values from Fletcher and Pilcher for the chloropropanes. The equation to evaluate the heat of formation now reads

$$\Delta H_f (\text{Monochloroalkanes}) = N_{\text{CH}_3} * GC_{\text{CH}_3} + N_{\text{CH}_2} * GC_{\text{CH}_2} + N_{\text{Cl terminal}} * GC_{\text{Cl terminal}} + N_{\text{Cl non-terminal}} * GC_{\text{Cl non-terminal}} \quad (3)$$

in which the  $N_X$  are the number of times a group  $X$  occurs in the molecular structure, and  $GC_X$  is the group contribution parameter value associated with the heat of formation associated with the Group  $X$ . In order to achieve the required accuracy, which we had set at 1 kcal/mol (chemical accuracy), it was found that we need different group contributions for terminal and non-terminal chlorines. This was also what we commonly found for other groups, including the OH and C=O (aldehyde versus ketone) and C=C groups [10,11]. Despite having to take data from four different sources, we obtained good agreement (Table 2) between experimental and model values with all individual values within chemical accuracy, and the AAD read as 1.09 kJ/mol. The group values for Cl were  $-57$  kJ/mol for a non-terminal Cl and  $-50.5$  kJ/mol for a terminal Cl. The difference between terminal and non-terminal Cl, as derived from a fit to the experimental data, is in essence confirmed by ab initio calculations: 2-chlorobutane was calculated as being 8.6 kJ/mol more stable than 1-chlorobutane (B3LYP//6-311++G(2df,2p)). Of course, the group values were related to dHf, which include zero-point energy corrections and an internal energy term calculated for  $T = 298.15$  K, terms which are small but also a reason why the calculated energy difference by quantum method cannot exactly equal the difference in dHf values presented in Table 2. When we compare the model values from the Marrero–Gani method as implemented in the ICAS23 software, the results look satisfactory, although the value for chloroethane was much more off and the absolute averaged deviation was almost three times as high compared to our model results.

**Table 2.** Experimental and model values for mono-chloroalkanes. All values in kJ/mol. The origin of the experimental values is given in column 1. Values for the Marrero–Gani method evaluated through the ICAS23 software are shown for comparison.

Chloroalkanes	Exp. Value	Model dHf	Model-Exp	ABS (Model-Exp)	ABS (MG-Exp)
Chloroethane (Manion [16])	-112.1	-113.49	-1.39	1.39	9.30
1-Chloropropane (Fletcher and Pilcher [17])	-132	-134.12	-2.12	2.12	0.10
1-Chlorobutane (Stridth et al. [18])	-154.6	-154.75	-0.15	0.15	2.00
1-Chloropentane (Stridth et al. [18])	-175.2	-175.38	-0.18	0.18	1.90
1-Chlorooctane (Stridth et al. [18])	-238.9	-237.27	1.63	1.63	3.50
1-Chlorododecane (Stridth et al. [18])	-322	-319.79	2.21	2.21	3.90
2-Chloropropane (Fletcher and Pilcher [17])	-145	-145.72	-0.72	0.72	0.40
2-Chlorobutane (He et al. [19])	-166.7	-166.35	0.35	0.35	3.30
Averaged absolute difference				1.09	3.05

The dichloroalkanes and higher substituted chloroalkanes were the next challenge, as well as our model values shown in Table 3. In our model, we initially only used the two previously assessed Cl groups value, viz., Equation (3), and the numerical model

values are shown in column 3. The Marrero–Gani [6] and Joback and Reid [4] model values were evaluated using the ICAS23 software [13]. From Table 3, we corroborate that the dichlorobutanes were well-described by our model, and all three were within chemical accuracy. Whereas we adopted the two dichloroethane values from Manion [16], we note that the experimental values for these two species as reported by Larcher et al. [20] were 5–7 kJ/mol less negative (−127.6 and −125.4 kJ/mol) than those from Manion. This could suggest that Larcher’s value for 1,2-dichloropropane should be lowered by some 6 kJ/mol to make a fair comparison, thereby reducing the difference between this corrected value and our model to some 5 kJ/mol. For the 1,1-dichloroethane and 1,2-dichloroethane, we had deviations of 15 and 10 kJ/mol, respectively. Joback and Reid [4] also used single Cl parameters, but for the dichloroethane, they did not perform better than our present model, whereas for the dichlorobutanes, they performed very much beyond chemical accuracy (26–32 kJ/mol difference, see the seventh column in Table 3). The Marrero–Gani model as implemented in ICAS23 showed a very good agreement for 1,1-dichlorobutane, but this is to be attributed to  $\text{CHCl}_2$  as an additional group and thus one additional adjustable parameter. However, for the other two 1,2-dichloroalkanes, the result was clearly beyond chemical accuracy (6.80 and 7.45 kJ/mol off), whereas agreement with experimental values was within chemical accuracy for the dichlorobutanes, although less beneficial than our current model. Thus, whereas none of the three models performed well overall for the dichloroalkanes, one may say that the present model at least described the latter two species in which the chlorines were spatially clearly separated best (0.12 kJ/mol difference). We return to the dichlorinated alkanes further below and show the much-improved results.

**Table 3.** Experimental and model values for multiple chlorosubstituted alkanes. All values are in kJ/mol. The origin of the experimental values is given in column 1. The final model of the present work (last column) values are based on Equation (3), complemented with neighbor interactions of the type 1,1-dichloro and 1,2-dichloro (for an explanation, see the text).

Dichloroalkanes	Exp. Value	Model dHf	Model-Exp	ABS (Model-Exp)	MG-Exp	Joback and Reid-Exp	Final Model Present Work-Exp
1,1-dichloroethane (Manion [16])	−132.5	−147.36	−14.86	14.86	0.50	11.5	1.74
1,2-dichloroethane (Manion [16])	−132	−142.26	−10.26	10.26	6.80	16	−4.26
1,2-dichloropropane (Lacher et al. [20])	−162.8	−174.49	−11.69	11.69	7.45	20.8	−5.69
1,2-dichlorobutane (He et al. [19])	−191.2	−195.12	−3.92	3.92	0.84	28.6	2.08
1,3-dichlorobutane (He et al. [21])	−195	−195.12	−0.12	0.12	4.00	32.40	−0.12
1,4-dichlorobutane (An et al. [21])	−183.4	−183.52	−0.12	0.12	3.25	26.4	−0.12
averaged absolute difference				6.83	3.81	22.6	2.33
Tri-, tetra-, hexachloroethane	exp. value [16]	model dHf	model-exp	ABS (model-exp)	MG-exp	Joback and Reid-exp	Final model present work-exp
1,1,1-trichloroethane	−144.6	−194.86	−50.26	50.26	2.3	4	−0.46
1,1,2-trichloroethane	−148	−176.13	−28.13	28.13	13.00	11	0.47
1,1,1,2-tetrachloroethane	−152.3	−222.63	−70.33	70.33	17.70	−3.7	−2.53
1,1,2,2-tetrachloroethane	−156.7	−210	−53.30	53.3	4.70	−1.3	3.90
pentachloroethane	−155.9	−257.5	−101.60	101.6	31.10	−21.1	0.80
hexachloroethane	−148.2	−305	−156.80	156.8	52.80	−48.4	−3.20
averaged absolute difference				76.7	20.3	14.9	1.89

When we now look at the second half of Table 3 comprising the tri- up to hexachloro-substituted species, for all species except 1,1,2-trichloroethane, we found very significant deviations for our present model. Regarding the two other named models, MG ICAS23 revealed good agreement for two species, and Joback and Reid for three species; all other

species revealed significant deviations. It is remarkable that the MG approach with a dedicated  $\text{CCl}_3$  group parameter showed, as expected, good agreement for 1,1,1-trichloroethane, but still truly worse agreement for hexachloroethane. It is even more remarkable that the MG ICAS23 tool showed only good agreement for two species of the higher substituted alkanes, despite the fact that three additional parameters were introduced compared to our approach and that of Joback and Reid:  $\text{CHCl}_2$ ,  $\text{CCl}_3$ , and a third-order contribution in the form of a parameter for Cl-C-Cl. We did make the attempt to introduce a  $\text{CCl}_3$  group in our present model, which clearly and significantly improved the results, but most were still far beyond the experimental values; for example, for hexachloroethane, the difference between model and experimental values reduced from  $-156.60$  down to  $-53.80$  kJ/mol.

In our previous papers on the GC approach [10,11], we introduced nearest or next-nearest neighbor interactions, as this is an approach that is quite common in chemistry and physics. By doing that, we were able to obtain very good results for, e.g., the methylalkanes [11] regarding the heats of formation. When we adopted an interaction parameter of  $+16.6$  kJ/mol for neighboring Cl atoms attached to the same carbon atom, e.g., 1,1-dichloro or 2,2-dichloro, and a  $+6$  kJ/mol interaction parameter for 1,2 dichloro interactions, e.g., as in 1,2-dichloropentane, we obtained excellent results, as shown in the last column of Table 3 (final model present work). Note that both these corrections had positive values, therefore suggesting that it is likely that we are dealing with steric repulsion effects. When we had a  $\text{CCl}_3$  terminal group, this contained three different Cl-Cl interactions all attached to the same carbon atom, and thus the interaction parameter reads three times  $+16.6$  kcal/mol =  $49.8$  kJ/mol. The results, with only these two well-defined additional neighbor interaction parameters, shown in the last column of Table 3, showed very satisfactory agreement between the experimental and our final model values. In fact, only the value for 1,2-dichloropropane ( $-5.69$  kJ/mol) was somewhat beyond chemical accuracy, but as previously indicated, the value for 1,2-dichloropropane (Lacher et al. [20]) can be subject to discussion and the result might well fall within chemical accuracy. With the annotation that the error in the experimental value was around 4 kJ/mol, the results shown in the last column of Table 3 were very satisfactory indeed.

In conclusion, our final model with only four parameters (terminal and non-terminal Cl and two dichloro interaction parameters) led to excellent results that have not been achieved previously.

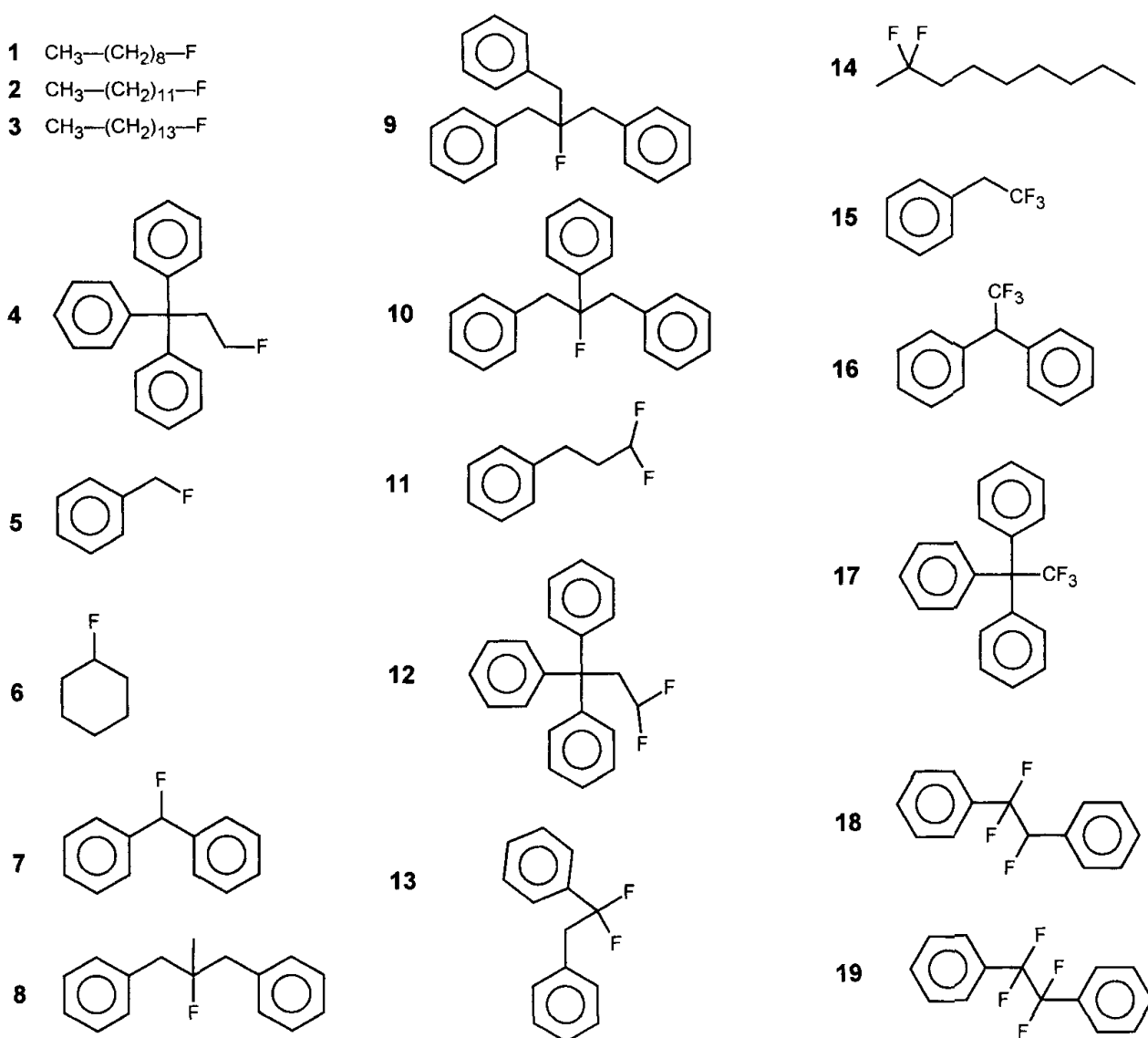
### 3.2. Fluoro Hydrocarbons

Polyfluoroalkyls are part of the class of PFAS (per- and polyfluoroalkyl substances) that are a highly relevant class of compounds, as PFAS do not easily break down and some types have been shown to accumulate in the environment and in our bodies. It must be considered well known that exposure to some PFAS has been linked to serious health effects.

We will see that the fluorohydrocarbons form a somewhat difficult group for establishing a group contribution approach. This can be due to various reasons, wherein the lack of sufficient reliable and consistent experimental data, in particular for smaller, non-aromatic molecules, is a definite problem. Although for the chloroalkanes, steric effects were likely an issue due to the bulky chlorine atoms, this would not be the case for the smaller fluorine atoms. However, with fluor being the most electronegative element, effects on neighboring atoms are expected, as we know for instance from nearest and next-nearest neighbor effects on the  $\text{C}_{1s}$  core level energies in XPS spectroscopy [22], and therefore there are measurable effects, even on the most inner electrons of carbon atoms due to the presence of fluorine atoms. In addition, several of the molecules for which experimental data are available clearly show steric interactions due to phenyl group congestion, i.e., Van der Waals overlap, which disrupts the wider application of pure group additivity.

Table 4 contains a set of experimental data mostly originating from Schaffer et al. [23]. The structures considered by Schaffer et al. are shown in Scheme 1, which was adopted from their paper with permission. Schaffer et al. also reported significant strain enthalpies, which

are enthalpy differences that reflect a deviation from the additive GC contribution (we refer to the original publication to how these strain energies were calculated). These strain energies can, however, be the cause of steric effects (ring strain, Van der Waals overlap), or electronic effects, a combination of these, or something we cannot yet explain. In addition, heats of formation for some smaller fluoroalkanes were taken from [24], whereas the first three entries were adopted from [25] and are the result from G3 quantum chemical calculations. This level of calculation, G3, normally yields reliable results for the heat of formation property of these small molecules. It is realized that it is far from ideal to adopt data from various different sources, but the lack of data on fluoroalkanes leaves us without a realistic alternative, whereas the small species will facilitate the parametrisation of more simple mono-fluoroalkanes, wherein we have a good start to properly analyze the other species.



**Scheme 1.** Fluor containing structures considered here as taken from Schaffer et al. [23] and reproduced with permission from Wiley-VCH GmbH.

**Table 4.** Experimental and model values for fluoroalkanes. All values in kJ/mol. Most experimental values were taken from [23] and some more from [24,25]; for comments, see the text. The corrections (corr.) in column 5 only apply to species involving 1,1- or 1,2-difluoro interactions; for an explanation, see the text.

Fluoroalkanes	Exp. Value	Model dHf	Model-Exp	ABS (Model-Exp) + Corr.	ABS (ICAS23-Exp)	Number in [23], See Scheme 1
Fluoroethane [25]	-278	-281.49	-3.49	3.49	13.60	
2-Fluoropropane [25]	-315.7	-317.22	-1.52	1.52		
2-Fluoro-2-methylpropane [25]	-360	-360.58	-0.58	0.58		
1-Fluorononane [23]	-423.5	-425.9	-2.40	2.40	14.30	1
1-Fluorododecane [23]	-489.2	-487.8	1.41	1.41	18.20	2
1-Fluorotetradecane [23]	-533.0	-529.1	3.95	3.95	20.00	3
2,2-Difluorononane [23]	-671.4	-674.5	-3.10	3.10		14
(1,3-Diphenyl)-2-methyl-2-fluoropropane [23]	-136.7	-136.1	0.58	0.58		8
Fluorocyclohexane [23]	-336.6	-341.9	-5.28	5.28	103.22	6
1,1-Difluoroethane [24]	-497	-483.36	13.64	3.86	88.40	
1,1-Difluoro-3-phenyl-propane [23]	-414.4	-392.9	21.5	4.0	6.04	11
1,1,1-Trifluoroethane [24]	-748.7	-698.9	49.84	2.7	43.04	
1-Chloro-1-fluoroethane [24]	-313.4	-315.36	-1.96	1.96	291.96	
Averaged absolute difference				2.68	66.53	
1,1,2-Trifluoroethane [24]	-691	-659.50	31.50	0.0		
1,1-Diphenyl(fluoromethane) [23]	-42.6	-55.5	-12.90	12.9	-69.40	7
1,3-Diphenyl-(2-methylphenyl)-2-fluoropropane [23]	-14.1	-23.9	-9.8	9.8		9
1,3-Diphenyl-2-phenyl-2-fluoropropane [23]	15.0	-3.3	-18.3	18.3	-42.00	10
Fluoromethylbenzene [23]	-126.3	-148.6	-22.4	22.4	-64.09	5
Trifluoroethylbenzene [23]	-623.9	-586.6	37.2	15.3	31.32	15
1,1-Diphenyl(trifluoroethane) [23]	-516.1	-479.5	36.6	15.9	24.42	16
Triphenyl(3-fluoropropane) [23]	57.5	10.7	-46.8	46.8	-10.69	4
1,2-Diphenyl-1,1-difluoroethane [23]	-260.5	-305.6	-45.1	62.6	-110.62	13
1,1,1-Triphenyl-3,3-difluoropropane [23]	-157.6	-191.1	-33.5	51.0	-31.99	12
1,1,1-Triphenyl-2,2,2-trifluoroethane [23]	-364.9	-386.0	-21.1	73.6	-8.85	17
1,2-Diphenyl-1,1,2-trifluoroethane [23]	-462.2	-521.5	-59.3	90.8	-169.30	18
1,2-Diphenyl-1,1,2,2-tetrafluoroethane [23]	-689.0	-751.0	-62.0	125.0	-195.84	19

When we look at the first nine entries in Table 4, in which the model involved new group parameters for a terminal (-218.5 kJ/mol) and a non-terminal fluorine (-232.5 kJ/mol) only and where we evaluated the heats of formation by applying

$$\Delta H_f = N_{CH_3} * GC_{CH_3} + N_{CH_2} * GC_{CH_2} + N_{F_{terminal}} * GC_{F_{terminal}} + N_{F_{non-terminal}} * GC_{F_{non-terminal}}, \quad (4)$$

we observed good performance of our GC model with all model values within chemical accuracy, except for a somewhat larger deviation for fluorocyclohexane (in view of experimental accuracies, the latter value is not necessarily beyond chemical accuracy). For structure 8 (1,3-diphenyl-2-methyl-2-fluoropropane), we incorporated twice the correction



term for a monosubstituted benzene (phenyl) ring (AromMonoalkyl, see [10]). The difference between terminal and non-terminal fluor group parameter values was 14 kJ/mol, which was close to the values in the range 18 kJ/mol, which we calculated for 1- and 2-fluoroalkanes using B3LYP/6-311+G\*\* quantum chemical calculations, with the 2-position being the energetically more stable form compared to the 1-position.

Herewith, we found a sound basis for discrimination between group contribution parameter values for the terminal and non-terminal F atoms. For 1,1-difluoroethane, we observe a clear difference of 13.6 kJ/mol between the experimental value and initial model value (column 4 in Table 4), whereas for 1,1-difluoro-3-phenyl-propane, this difference was as high as 21.5 kJ/mol. These effects can be attributed to the additional effect of two F atoms at the terminal positions. B3LYP QC calculations confirmed that this double substitution at the same C-atom resulted in a more stable species, e.g., 1,1-difluorooctane was more stable than 1,8-difluorooctane. The value of the correction accounting for the 1,1-interaction was established as  $-17.5$  kJ/mol, whose value was accounted for in the values in column 5 in Table 4 for those species comprising a 1,1-interaction. The value of  $-17.5$  kJ/mol was the result from minimizing the difference between experimental and model values for the 1,1- and 1,1,1-substituted species in the upper part of Table 4 on the one hand, but at the same time looking at a minimum number of deviations beyond chemical accuracy (4.2 kJ/mol). By selecting this value for the 1,1-interaction, all three values were within chemical accuracy from the corresponding experimental values. As we did not see any steric hindrance effects in these three structures, the fact that this correction term was now negative (for chlorine it was positive) could be indicative of electronic effects related to the strong electronegativity of the F-atom. Similarly to the chlorine case, and even though in the chlorine case, we were dealing with steric effects, whereas for fluorine, these are primarily electronic effects, we adopted three times the 1,1-interaction for 1,1,1-trifluoroethane and thus a correction of  $-52.5$  kJ/mol, which led to a good result.

In summary, for the data in the upper part of Table 4, we arrived at an averaged absolute deviation between model and experimental values of 2.68 kJ/mol, which was very satisfactory. In comparison, the MG ICAS23 method did not allow for the evaluation of the dHf for several species (e.g., with a fluorine at the 2-position, see Table 4), whereas the averaged absolute deviation for all other structures was as high as 66 kJ/mol. Still, it should also be admitted that the result should be considered with care: the 1,1-interaction parameter as we have determined it involved three structures only, and these originated from two different publications.

When we now turn to the lower section of Table 4, first of all, it is to be noted that all structures in the second half of Table 4 exhibiting a larger difference between experimental and model values were exactly the structures for which Schaffer et al. (Table 2 in [23]) have reported strain energies. We added the data for 1,1,2-trifluoroethane to determine the 1,2-difluoro interaction parameter from a structure definitely not subject to steric effects. Its value was determined as  $-7$  kJ/mol (note that in 1,1,2-trifluoroethane, this contribution needs to be accounted for twice). The large(r) differences between the experimental and model values could not be resolved, something we elaborate on below. For readers not interested in the details, we suggest skipping to Section 3.3.

The structures in the lower part of Table 4 with the smallest deviation were 1,1-diphenyl (fluoromethane) (Structure 7) and 1,3-diphenyl-(2-methylphenyl)2-fluoropropane (Structure 9). These were relatively open structures where we had some steric effects, which can be corroborated both from building space filling models as well as from the fact that the experimental values were less negative than the corresponding model value. For comparison, Structure 8 was more open and was subject to negligible steric effects that were reflected by the good agreement between the model and the experimental values. Thus, the deviations for Structures 7 and 9 can most likely be attributed to moderate steric effects but cannot be accounted for by the GC method up to the level of chemical accuracy.

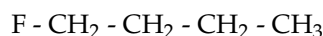
For fluoromethylbenzene (Structure 5) and trifluoroethylbenzene (Structure 15), there was no reason whatsoever to assume steric hindrance, but still the difference between model

and experiment was significant at 22 and 15 kJ/mol, respectively. It may seem somewhat peculiar when we compared it to 1,1-difluoro-3-phenyl-propane (Structure 11), wherein we also did not see steric hindrance but observed good agreement with the final model value and the experimental value (4.0 kJ/mol difference). However, as the experimental values were more negative for fluoromethylbenzene (Structure 5) and trifluoroethylbenzene (Structure 15) than the initial model values (column 3), this indicated that electronic effects are more likely to be the cause for the discrepancy for these two species. Due to the fact that for Structure 11, there are three carbon atoms between the fluorine and the phenyl group, this possibly prohibits an electronic interaction between the very electronegative fluorines and the phenyl group  $\pi$ -electrons.

When we compared columns 4 and 5 in the lower part of Table 4, we saw large differences between the values in these columns, in particular for the species comprising tri- or tetra-fluorine substitution. The 1,1-interaction parameter could only be determined (upper part of Table 4) from very few structures from two different literature sources. This was even more so for the 1,2-interaction parameter. In particular, the latter strongly influenced the structures 18 and 19 exhibiting the largest deviations. For a possible 2,2- or 3,3- interaction parameter, we can only refer to the experimental value for 2,2-difluorononane, which does not necessarily suggest the need for such an additional parameter. Furthermore, steric effects will have a clear effect in various structures comprising three phenyl rings.

Because of the discrepancies we noted, and as we have suggested that the strong electronegativity of fluorine as an element could play an essential role not accounted for by the group contribution terms, we performed QC calculations using the B3LYP functional and the 6-311+G\*\* basis set on fluorosubstituted alkanes. Whatever the precise explanation, we found for instance that whereas according to our current GC model 1,1,9,9-tetrafluorononane was 31 kJ/mol less stable than 2,4,6,8-tetrafluorononane, the quantum calculations suggested a negligible energy difference (2 kJ/mol). On the other hand, 2,2,8,8-tetrafluorononane was predicted, by the model, to be 58 kJ/mol more stable than 1,1,9,9-tetrafluorononane, which did not compare that badly with the QC B3LYP difference of 49 kJ/mol. A significant difference was noted when looking at 4,4,5,5-tetrafluorononane, wherein our model predicted a difference of 58 kJ/mol compared to 1,1,9,9-tetrafluorononane, whereas the QC B3LYP calculations revealed a difference of only 14 kJ/mol. Another example was 1,1,1-trifluorooctane, for which the GC model predicted a 19 kJ/mol higher stability compared to 2,4,6-trifluorooctane, whereas the QC calculations suggested 30 kJ/mol. At the same time, according to the GC model, 2,3,4-trifluorooctane had the same stability as 2,4,6-trifluorooctane, wherein the QM calculations suggest it to be 36 kJ/mol less stable. These results indicate that we were facing a complex situation in which, most likely, multiple factors play a role, including electronegativity and steric effects. Additional experimental data, or heats of formation calculated using high level quantum methods such as G3, are needed to further analyze the situation.

The observation that we cannot develop a GC approach for this type of series lies in the nature of the GC approach: the assumption that a molecule can be decomposed in molecular fragments that are in essence mutually independent. As the strong electronegativity of the fluorine atom has a clear influence on the nearest and next-nearest carbon atoms [14], even though we did not have independent *quantitative* information on the magnitude regarding the heats of formation, this implies that in the context of a GC approach we have, e.g., for chlorobutane



comprising the Groups  $\text{Cl}_{\text{terminal}}$ ,  $\text{CH}_2$  and  $\text{CH}_3$ . For the fluorine equivalent



which now comprises the Groups  $\text{F}_{\text{terminal}}$ ,  $\underline{\text{CH}_2}$ ,  $\underline{\text{CH}_2}$  and  $\text{CH}_3$ . Note that whereas in the first case all  $\text{CH}_2$  are equivalent, in the second case  $\text{CH}_2 \neq \underline{\text{CH}_2} \neq \underline{\text{CH}_2}$  that now comprises the groups  $\text{F}_{\text{terminal}}$ ,  $\text{CH}_2$ ,  $\underline{\text{CH}_2}$ , and  $\text{CH}_3$ . Note that whereas in the first case all

CH<sub>2</sub> were equivalent, in the second case, CH<sub>2</sub> ≠ CH<sub>2</sub> ≠ CH<sub>2</sub>. Thus, we would need to introduce two further GC parameters. However, for multiple substituted fluoroalkanes, further parameters will be required, and not only for the CH<sub>2</sub> but also for the CH and the C group. Whereas this is possible on paper, in practice, we do not have the experimental data to develop this further. Finally, we have now been talking about the fluoroalkanes only, but the same applies to structures involving affected carbon atoms in phenyl rings such as in structures 5, 7, 10, 13, 18, and 19.

### 3.3. Benzylhalides

This short section on a few benzylhalides is introduced to further illustrate the essential difference between the chlorine and the fluorine atoms as we have seen in the previous sections. Experimental data on benzylhalides were reported by Verevkin et al. [26], from which we took the experimental heats of formation in the gas phase of the F- and Cl-containing species benzylfluoride, benzylchloride, (1-chloroethyl)benzene, and α,α-dimethylbenzyl (cumyl) chloride.

For benzylchloride, our model-evaluated value of +19.37 kJ/mol compared very favorably with the experimental value of +17.5 kJ/mol. Similarly, the model value for (1-chloroethyl) benzene of −11.9 kJ/mol was still in reasonable though not good agreement with the experimental value of −5.4 kJ/mol. For benzylfluoride thereagainst the model value −148.6 kJ/mol clearly deviates from the experimental value −126.4 kJ/mol by more than 20 kJ/mol, further confirming that fluorine is different though not due to steric effects. Finally, for cumyl chloride, the difference between the experiment (−35.9 kJ/mol) model (−52.2 kJ/mol) of 16.3 kJ/mol was large but very comparable to the difference for t-butylbenzene, namely, 17.9 kJ/mol, illustrating the influence of steric hindrance that could not be accounted for by the GC approach.

In summary, the results for the benzylhalides confirm our former conclusions on chloro and fluoro substituted compounds.

### 3.4. Nitro Compounds

Experimental data were taken from Verevkin et al. [27], who have noted that “Thermochemical properties of aliphatic nitroalkanes available in the literature are scarce and inconsistent”, something that is not totally uncommon in the field. Therefore, they re-evaluated the nitro compounds using both experimental data and the GC approach. Consequently, previous GC parametrizations should be used with care, depending on the experimental data set used. Experimental and model values were collected, as shown in Table 5. On the basis of the mononitroalkanes, we established the value −38 kJ/mol for a terminal NO<sub>2</sub> group and −50 kJ/mol for a non-terminal NO<sub>2</sub> group, and the heat of formation for the nitroalkanes can be evaluated from

$$\Delta H_f = N_{\text{CH}_3} * GC_{\text{CH}_3} + N_{\text{CH}_2} * GC_{\text{CH}_2} + N_{\text{NO}_2 \text{ terminal}} * GC_{\text{NO}_2 \text{ terminal}} + N_{\text{NO}_2 \text{ non-terminal}} * GC_{\text{NO}_2 \text{ non-terminal}} \quad (5)$$

For nitromethane, we saw a discrepancy between experimental and model values, which is not unexpected for the smallest species in a series, as this was also observed for another nitrogen-containing smallest species in a series, i.e., methylamine [10], as well as for methylfluoride (we did not include this in the previous section as we had an isolated experimental value from the NIST website [12]). Therefore, as the value is known, the experimental value should be adopted in this case.

For the mononitroalkanes, excluding 2,4,4-trimethyl-2-nitropentane, the averaged absolute difference was established as 2.78 kJ/mol (see Table 5). For 2-nitrobutane, the error in the experimental result was given as 3.1 kJ/mol [19], and for 2-nitrodecane as 3.3 kJ/mol [19], and therefore also the model value for the latter value was not necessarily beyond chemical accuracy. For the exception 2,4,4-trimethyl-2-nitropentane, we did not have the appropriate *Ansatz* on how to correct for a 2-nitro substitution where there was already a 2-Me substitution. Still, on the basis of the experience we obtained on the highly substituted methylalkanes [11], it was not surprising if the GC approach cannot be applied

here without an appropriate additional interaction parameter that we cannot analyze properly on the basis of a single experimental value.


**Table 5.** Experimental and model values for nitroalkanes. All values in kJ/mol. In the third column entitled “model dHf”, the corrections related to alkyl-substituted alkanes were accounted for, e.g., correction for 2-Me to terminal Me group interaction. In the final column, the interaction correction term for two nitro groups attached to the same carbon atom was added, with the value being 35 kJ/mol (see text).

Nitro Compounds	Exp. [27]	Model dHf	Model-Exp	ABS (Model-Exp)	ABS (Model-Exp + Corr.)
<b>Mononitrile Alkanes</b>					
Nitromethane	−71.5	−80.36	−8.86		
Nitroethane	−102.4	−100.99	1.41	1.41	
1-Nitropropane	−124.4	−121.62	2.78	2.78	
2-Nitropropane	−140	−138.72	1.28	1.28	
1-Nitrobutane	−145	−142.25	2.75	2.75	
2-Nitrobutane	−163	−159.35	3.65	3.65	
1-Nitropentane	−165	−162.88	2.12	2.12	
2-Nitrodecane	−278.1	−283.13	−5.03	5.03	
1,3-Dinitropropane	−135.5	−137.89	−2.39	2.39	
1,4-Dinitrobutane	−155.6	−158.52	−2.92	2.92	
Nitrocyclohexane	−159.2	−156.75	2.45	2.45	
2-Me-2-nitropropane	−186.1	−182.28	3.82	3.82	
Averaged absolute difference				2.78	
2,4,4-Trimethyl-2-nitropentane	−249.4	−279.73	−30.33	30.33	
<b>Dinitro alkanes</b>					
Dinitromethane	−38.4	−80	−41.6	41.6	6.6
1,1-Dinitroethane	−87.2	−122.36	−35.16	35.16	0.16
1,1-Dinitropropane	−109.5	−142.99	−33.49	33.49	1.51
1,1-Dinitrobutane	−132.1	−163.62	−31.52	31.52	3.48
1,1-Dinitropentane	−147.8	−184.25	−36.45	36.45	1.45
2,2-Dinitropropane	−135	−185.72	−50.72	50.72	
1,2-Dinitroethane	−96.7	−117.26	−20.56	20.56	
2,3-Dimethyl-2,3-dinitrobutane	−226.2	−271.44	−45.24	45.24	

Turning to the dinitroalkanes, after introducing a nitro–nitro interaction parameter for two nitro groups attached to the same carbon atom of a magnitude of 35 kJ/mol, we obtained very satisfactory results (see Table 5). This interaction energy, at least at a qualitative level, is justified when realizing the presence of geminal effects in dinitro compounds. For 2,2-di-nitroadamantane, the *destabilizing* interaction of the geminal substituents was evaluated as 59 kJ/mol [28]. This suggests that the magnitude of 35 kJ/mol we established (by fitting) was in the correct range to be attributed to the geminal effect. However, this was for the 1,1-dinitroalkanes, whereas for other alkanes, different values will apply, preventing a generic correction factor in the context of a GC approach as here we also only isolated experimental data (1,2-, 2,2-, and 2,3-dinitro), which prohibits a proper determination of the interaction parameters at this moment.

### 3.5. Acetals: 1,3-Dioxolane

Experimental data on acetals and other oxygen-containing groups were reported by Verevkin [29]. We attempted to establish groups to account for the experimental heats of formation. The groups were defined in various ways, including ether  $-C-O-C-$ , the “double ether”  $-C-O-C-O-C-$ , or the  $-O-$  as a group itself, but none of these led to an acceptable result. In fact, the differences were very similar to those reported in [29].

Now acetals such as di-methoxymethane as well as the corresponding ketals are not exceptional molecules and are without any specific steric effects or electronic conjugation effects. Therefore, for most of the species in [29], we can exclude that direct steric effects are the cause of the differences. On the other hand, also for the ethers we needed [11] dedicated GC parameters depending on the geometrical factors, more specifically the COC bond angle. Therefore, our next step was to adopt a larger chemical unit, e.g., 1,3-dioxolane  as a group. We can now see from the data in Table 6 that we achieved good agreement between the experimental and our model values for a further eight dioxolanes, with only the value for 2,2-di-iPr-1,3-dioxolane beyond somewhat beyond chemical accuracy. For 2-Me-2-iPr-1,3-dioxolane and 2,2-di-iPr-1,3-dioxolane, we included the Me-Me neighbor interaction parameters as established previously in our work on branched alkanes [11]. The averaged absolute difference between experimental and model values for the eight substituted 1,3-dioxolanes read as 2.50 kJ/mol. On the contrary, the model of Marrero and Gani and the model of Joback and Reid (both implementations in ICSAS23) revealed significant differences between the model and the experimental values.

**Table 6.** Experimental and model values for the gas phase heat of formation for a set of 1,3-dioxolanes. All values in kJ/mol. 1,3-Dioxolane was adopted as individual groups with a group-contribution-specific parameter value that was its experimental value. Consequently, the value “model-exp.” is given as zero.

1,3-Dioxolanes	Experimental Value [29]	Model	ABS (Model-Exp)	ABS (ICAS23-Exp)	ABS (Joback and Reid-Exp)
1,3-Dioxolane	-301.6		0	6.2	13.2
2-Me-1,3-dioxolane	-344.1	-343.96	0.14	14.8	14.7
2-nPr-1,3-dioxolane	-386.8	-385.22	1.58	11	16.1
2,2-di-Me-1,3-dioxolane	-389.4	-386.32	3.08	33.2	54.6
2-Me-2-Et-1,3-dioxolane	-409.7	-406.95	2.75	28.9	54.2
2-Me-2-nPr-1,3-dioxolane	-429.2	-427.58	1.62	27.8	53.1
2-Me-2-nPe-1,3-dioxolane	-468.8	-468.84	0.04	26	51.4
2-Me-2-iPr-1,3-dioxolane	-421.9	-417.78	4.12	40.3	40.5
2,2-di-iPr-1,3-dioxolane	-456	-449.24	6.76	48.9	28.1
Averaged absolute difference			2.51	28.86	39.09

## 4. Summary

Our aim was to establish a group contribution approach revealing “chemical accuracy”, i.e., a maximum 1 kcal/mol (4.2 kJ/mol) difference between experimental and model values. Moreover, to have a reliable method and not only a low absolute average deviation as propagated in most publications, we wanted no or exceptionally few outliers. Finally, we also needed to identify when the linear additive group contribution method fails, i.e., it cannot be applied because the assumption of linearity and additivity is invalid.

We primarily used experimental data sets from reliable and consistent sources only, which turned out to be essential to the quality of the results and to help in identifying when the group contribution approach truly fails. Secondly, as in our previous work, only by determining the GC group by group and verifying each individual result (rather

than a computer-automatic procedure), specific trends and individual outliers could be assessed and analyzed. We applied molecular modelling tools to build space-filling models (Van der Waals radii of the atoms) in order to confirm steric overlap and consequently identify why the GC approach is tempted to fail, at least with respect to the requirement of chemical accuracy.

In the present work, we obtained a model for the mono-chloroalkanes with an absolute averaged deviation (ADD) of 1.09 kJ/mol and all individual values within chemical accuracy. For the dichloroalkanes, an ADD of 2.33 kJ/mol was found, and for the tri-, tetra-, penta-, and hexa-chloroethanes, we obtained an ADD of 1.89 kJ/mol. For the fluoroalkanes, we obtained a model for which ADD = 2.68 kJ/mol, with only the value for fluorohexane slightly beyond chemical accuracy (5.3 kJ/mol). However, this series only comprised 1- and 2-mono and 1,1-di- and 1,1,1-trifluoroalkanes. Regarding the other fluoro-hydrocarbons that were considered, some had multiple phenyl rings leading to steric effects that are molecule specific and therefore cannot be accounted for by a limited set of GC parameters. However, there were also various other species for which steric effects cannot explain the observations. From density functional theory (DFT) quantum calculations, we obtained relative energies for different F-substitution patterns, and it had to be concluded that the overall situation regarding multiple substituted fluor-containing hydrocarbons is too complex to be elucidated at present. The results for some benzylhalides confirmed our findings for the other chloro- and fluoro-substituted molecules.

We could effectively model the heat of formation of nitroalkanes that do not have two close nitro groups—the AAD was 2.78 kJ/mol. Moreover, the 1,1-dinitroalkanes could be accounted for well after adopting a single additional (neighbor) interaction parameter representing, for this type of species, the known geminal effect. As only isolated experimental data were available for 1,2-, 2,2- and 2,3-dinitro alkanes, it was premature to establish specific interaction parameters.

For a set of acetals, we initially did not achieve a good model (results within chemical accuracy) but deviations between the experimental value and model comparable to those reported in [29]. First, after adopting 1,3-dioxolane as a group, we were able to establish an appropriate GC model with good performance.

## 5. Conclusions and Outlook

Despite the GC method being old, this work shows that a substantial and necessary increase in performance can still be achieved. The overall conclusion of the three papers on this topic [10,11], present work is that it is feasible to achieve heats of formation of organic molecules with chemical accuracy when using high-quality experimental data and the judicious definition of chemical groups. Moreover, to have a reliable method and not only a low absolute average deviation between experimental and model values as reported in most publications, our results reveal no or exceptionally few outliers. Our results herewith outperform previous parametrizations. It is important to recognize the limitations of the GC approach, something that is rarely pointed out explicitly in the literature, and something which specifically applies to a breakdown related to the conditions of linearity and additivity as they can show up as a result of steric hindrance, ring strain, geminal effects, or electronic conjugation effects.

Regarding the future perspective and to achieve wide applicability to a larger variety of organic molecules, further high-quality experimental data should be fostered. However, this will remain an issue because there is only a limited amount of experimental data. Alternatively, it is possible to evaluate the heat of formation by pure quantum chemical calculations, e.g., the so-called G3 or G4 composite quantum methods, which have been applied in this context in various papers [25,30]. The results of such calculations, such as in the first three entries in Table 4, led to good consistency with the other experimental data. Such additional data can enable the determination of interaction parameters for, e.g., the 1,2-, 2,2- and 2,3-dinitro alkanes, but also of many other systems for which we need in

essence some data on key groups. Such types of quantum calculations can also be applied to molecules for which the GC method is inapplicable.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/appliedchem2040015/s1>, Table S1. Quantum mechanics B3LYP calculated total energies (hartree) and relative energies (kJ/mol) for fluorinated alkanes.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article or the Supplementary Materials.

**Acknowledgments:** The author also gratefully acknowledges Sergey Verevkin for truly interesting and relevant discussions. The author also sincerely thanks Georgios Kontogeorgis and Gürkan Sin and Guoliang Wang (all Danish Technical University DTU) for allowing the use and providing a copy of the ICAS23 software suite, particularly the ProPred module, which was, in part, used in this study.

**Conflicts of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare no conflict of interest.

## References

1. Van Krevelen, D.W.; Chermin, H.A.G. Estimation of the free enthalpy (Gibbs free energy) of formation of organic compounds from group contributions. *Chem. Eng. Sci.* **1951**, *1*, 66–80, Erratum in **1952**, *1*, 238. [[CrossRef](#)]
2. Benson, S.W.; Cruickshank, F.R.; Golden, D.M.; Haugen, G.R.; O'Neal, H.E.; Rodgers, A.S.; Shaw, R.; Walsh, R. Additivity rules for the estimation of thermochemical properties. *Chem. Rev.* **1968**, *69*, 279–324. [[CrossRef](#)]
3. Cohen, N.; Benson, S. Estimation of the heats of formation of organic compounds. *Chem. Rev.* **1993**, *93*, 2419–2438. [[CrossRef](#)]
4. Joback, K.G.; Reid, R.C. Estimation of Pure-Component Properties from Group-Contributions. *Chem. Eng. Commun.* **1987**, *57*, 233–243. [[CrossRef](#)]
5. Constantinou, L.; Gani, R. New group contribution method for estimating properties of pure compounds. *AIChE J.* **1994**, *40*, 1697–1710. [[CrossRef](#)]
6. Marrero, J.; Gani, R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria* **2001**, *183–184*, 183–208. [[CrossRef](#)]
7. Hukkerikar, A.S.; Meier, R.J.; Sin, G.; Gani, R. A method to estimate the enthalpy of formation of organic compounds with chemical accuracy. *Fluid Phase Equilibria* **2013**, *348*, 23–32. [[CrossRef](#)]
8. Argoub Kadda, A.; Mustapha, B.A.; Yahiaoui, A.; Khalel, T.; Hadji, D. Enthalpy of Formation Modeling Using Third Order Group Contribution Technics and Calculation by DFT Method. *Int. J. Thermodyn. (IJoT)* **2020**, *23*, 34–41. [[CrossRef](#)]
9. Aouichaoui, A.R.N.; Fan, F.; Mansouri, S.S.; Abildskov, J.; Sin, G. Molecular Representations in Deep-Learning Models for Chemical Property Prediction. *Comput. Aided Chem. Eng.* **2022**, *49*, 1591–1596. [[CrossRef](#)]
10. Meier, R.J. Group contribution revisited: The enthalpy of formation of organic compounds with “chemical accuracy”. *ChemEngineering* **2021**, *5*, 24. [[CrossRef](#)]
11. Meier, R.J. Group contribution revisited: The enthalpy of formation of organic compounds with “chemical accuracy” Part II. *AppliedChem* **2021**, *1*, 111–129. [[CrossRef](#)]
12. Available online: <https://webbook.nist.gov/> (accessed on 20 October 2022).
13. Proped (Property Prediction) Module Version 4.7 in ICAS23. Danish Technical University, Kongens Lyngby. Available online: <https://www.kt.dtu.dk/english/research/kt-consortium/software> (accessed on 20 October 2022).
14. Spartan' 10. Wavefunction Inc., Irvine, CA, USA. Available online: [www.wavefun.com](http://www.wavefun.com) (accessed on 20 October 2022).
15. Tirado-Rives, J.; Jorgensen, W.L. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. *J. Chem. Theory Comput.* **2008**, *4*, 297–306. [[CrossRef](#)] [[PubMed](#)]
16. Manion, J.A. Evaluated Enthalpies of Formation of the Stable Closed Shell C1 and C2 Chlorinated Hydrocarbons. *J. Phys. Chem. Ref. Data* **2002**, *31*, 123–172. [[CrossRef](#)]
17. Fletcher, R.A.; Pilcher, G. Measurements of heats of combustion by flame calorimetry. Part 7.—Chloromethane, chloroethane, 1-chloropropane, 2-chloropropane. *Trans. Faraday Soc.* **1971**, *67*, 3191–3201. [[CrossRef](#)]
18. Stridth, G.; Sunner, S. Enthalpies of formation of some 1-chloroalkanes and the CH<sub>2</sub>-increment in the 1-chloroalkanes series. *J. Chem. Thermodyn.* **1975**, *7*, 161–168. [[CrossRef](#)]
19. He, J.; An, X.; Hu, R. Measurements of enthalpies of formation of 2-chlorobutane and 1,2-dichlorobutane in gaseous state. *Acta Chim. Sin.* **1992**, *50*, 961–966.

20. Lacher, J.R.; Amador, A.; Park, J.D. Reaction heats of organic compounds. Part 5.—Heats of hydrogenation of dichloromethane, 1,1- and 1,2-dichloroethane and 1,2-dichloropropane. *Trans. Faraday Soc.* **1967**, *63*, 1608–1611. [[CrossRef](#)]
21. An, X.; He, J.; Hu, R. Study on the electrostatic interaction in organic chlorocompounds. Enthalpies of compustion and formation of 1,3- and 1,4-dichlorobutanes. *Thermochim. Acta* **1990**, *169*, 331–337.
22. Pijpers, A.P.; Meier, R.J. Core Level Photoelectron Spectroscopy for Polymer and Catalyst Characterisation. *Chem. Soc. Rev.* **1999**, *28*, 233–238. [[CrossRef](#)]
23. Schaffer, F.; Verevkin, S.P.; Rieger, H.-J.; Beckhaus, H.-D.; Röchardt, C. Enthalpies of Formation of a Series of Fluorinated Hydrocarbons and Strain-Free Group Increments to Assess Polar and Anomeric Stabilization and Strain. *Liebigs Ann.* **1997**, *1997*, 1333–1344. [[CrossRef](#)]
24. Kolesov, V.P.; Papina, T.S. Thermochemistry of Haloethanes. *Russ. Chem. Rev.* **1983**, *52*, 425. [[CrossRef](#)]
25. Kormos, B.L.; Liebman, J.F.; Cramer, C.J. 298 K enthalpies of formation of monofluorinated alkanes: Theoretical predictions for methyl, ethyl, isopropyl and tert-butyl fluoride. *J. Phys. Org. Chem.* **2004**, *17*, 656–664. [[CrossRef](#)]
26. Verevkin, S.P.; Krasnykh, E.L.; Wright, J.S. Thermodynamic properties of benzyl halides: Enthalpies of formation, strain enthalpies, and carbon–halogen bond dissociation enthalpies. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2605–2611. [[CrossRef](#)]
27. Härtel, M.A.C.; Klapötke, T.M.; Emel'yanenko, V.N.; Verevkin, S.P. Aliphatic nitroalkanes: Evaluation of thermochemical data with complementary experimental and computational methods. *Thermochim. Acta* **2017**, *656*, 151–160. [[CrossRef](#)]
28. Fritzsche, K.; Dogan, B.; Beckhaus, H.-D.; Röchardt, C. Geminale substituenteneffekte: Teil I. Thermochemie von 1-nitro-, 2-nitro-, 2,2-dinitro- und 2-cyano-2-nitroadamantan. *Thermochim. Acta* **1990**, *160*, 147–159. [[CrossRef](#)]
29. Verevkin, S.P. Improved Benson Increments for the Estimation of Standard Enthalpies of Formation and Enthalpies of Vaporization of Alkyl Ethers, Acetals, Ketals, and Ortho Esters. *J. Chem. Eng. Data* **2002**, *47*, 1071–1097. [[CrossRef](#)]
30. Verevkin, S.P.; Konnova, M.E.; Turovtsev, V.V.; Riabchunova, A.V.; Pimerzin, A.A. Weaving a Network of Reliable Thermochemistry around Lignin Building Blocks: Methoxy-Phenols and Methoxy-Benzaldehydes. *Ind. Eng. Chem. Res.* **2020**, *59*, 22626–22639. [[CrossRef](#)]