Supporting Information

Machine Learning-Based Correction for Spin-Orbit Coupling Effects in NMR Chemical Shift Calculations

Julius B. Kleine Büning^a, Stefan Grimme^{*a}, and Markus Bursch^{*b}

 ^a Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, University of Bonn, Beringstr. 4, 53115 Bonn, Germany
 ^b Max-Planck-Institut f
ür Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 M
ülheim an der Ruhr, Germany

> grimme@thch.uni-bonn.de bursch@kofo.mpg.de

Contents

Li	st of	Figure	S	S	53
Li	st of	Tables		S	53
1	Stat	tistical	Quantities	S	54
2	Fur	ther Te	echnical Details	S	55
	2.1	Struct	ures in the data set	 S	35
	2.2	Struct	ure and division of the data set \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	 S	35
	2.3	Input	feature vector	 S	37
3	Sup	plemer	itary Data	S	10
	3.1	Analy	sis of the data set	 S1	10
	3.2	Bench	mark studies	 S1	15
		3.2.1	Organotin and organolead compounds	 S1	15
		3.2.2	17HAC test set	 S1	17
4	Ref	erences		S2	24

List of Figures

S1	Color code of the distance categories
S2	Complete data set for ¹ H NMR chemical shifts
S3	Data set of $^{13}\mathrm{C}$ NMR connected to a HA via 1 bond
S4	Data set of $^{13}\mathrm{C}$ NMR connected to a HA via 2 bonds $\ .$
S5	Data set of $^{13}\mathrm{C}$ NMR connected to a HA via ≥ 3 bonds $\hdots \hdots \$
S6	Data set of ¹ H NMR connected to a HA via 1/2 bonds \ldots
S7	Data set of ¹ H NMR connected to a HA via 3 bonds $\dots \dots \dots$
S8	Data set of ¹ H NMR connected to a HA via ≥ 4 bonds $\ldots \ldots \ldots$
S9	Evaluation of $^1\mathrm{H}$ NMR data for the organotin and -lead compounds $~.~.~.$ S16

List of Tables

S1	Appearances of the HAs in the data set	S6
S2	List of all descriptors in the input feature vector of the $\Delta_{\rm SO}$ -ML model $$.	S8
S3	List of all compounds included in the 17HAC benchmark set	S18

1 Statistical Quantities

The following quantities were used for the statistical evaluation of the chemical shift data of the benchmark studies with n data points. δ_x denotes the predicted value (e.g., the purely SR chemical shift) and δ_r the reference, (e.g., the SO one).

Chemical shift deviation:

$$\Delta \delta = \delta_x - \delta_r \tag{1}$$

Mean deviation (MD):

$$MD = \frac{1}{n} \sum_{i}^{n} (\delta_{x_i} - \delta_{r_i})$$
(2)

Mean absolute deviation (MAD):

$$MAD = \frac{1}{n} \sum_{i}^{n} (|\delta_{x_i} - \delta_{r_i}|)$$
(3)

Mean squared deviation (MSD):

$$MSD = \frac{1}{n} \sum_{i}^{n} \left(\delta_{x_i} - \delta_{r_i}\right)^2 \tag{4}$$

Root mean square deviation (RMSD):

$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i}^{n} \left(\delta_{x_i} - \delta_{r_i}\right)^2} \tag{5}$$

Standard deviation or the errors (SD):

$$SD = \sqrt{\frac{1}{n-1} \sum_{i}^{n} \left(\delta_{x_i} - \delta_{r_i} - MD\right)^2}$$
(6)

2 Further Technical Details

2.1 Structures in the data set

The geometrical distortion procedure is used to generate structures with variations in bond lengths, angles and dihedral angles from an optimized parent structure. The process was used in the same way for the Δ_{corr} -ML method and is described in more detail in the respective Supporting Information.^{S1} For the data set of the Δ_{SO} -ML model, three distorted structures were generated from each parent structure. They were chosen automatically from random generations to select one per energy window of 2.5–5.0 kcal·mol⁻¹, 10.0–15.0 kcal·mol⁻¹, and 30.0–40.0 kcal·mol⁻¹ higher than the parent structure.

A complete detailed list of all 1597 compounds would be out of scope for this Supporting Information. However, all structures are provided in the supporting .zip archive as Cartesian coordinate (.xyz) files including combined files with a collection of optimized structures. All molecules consist of at least one carbon atom in a typical organic environment with at least one of the heavy atoms. The size of the molecules range between 3 and 46 atoms with a maximum number of 24 carbon or 27 hydrogen atoms per compound. The number of structures and chemical shifts that contain a certain element or group of elements is listed in Table S1. A more detailed list with extra information for every compound is provided in the supplementary raw data (.xlsx) file.

2.2 Structure and division of the data set

Like the ML architecture, the structure of the data set is also adapted from the previous Δ_{corr} -ML model. It consists of 1597 compounds, each of which features four structures (one optimized and three distorted ones). During the data acquisition process, each structure is screened for carbon or hydrogen atoms that make up a data point. There are three ways to shuffle the data before it is divided into training and test set:

- compounds mode: Shuffles only the compounds and it is ensured that all atoms from all structures of one compound are in the same data subset.
- structures mode: Shuffles the structures and it is ensured that all atoms of one structure are in the same data subset. However, different structures of the same compound can be part of the training and the test set.
- atoms mode: Shuffles all atoms, disregarding their structural origin. Atoms from one structure can be found in the training and the test set.

For the Δ_{corr} -ML method, we have shown that the data bias that stems from including different degrees of geometric distortion of the same compound in both data subsets (structures mode) is reasonable. Switching to compounds mode does lead to an expected

Element(s)	# structures	$\#^{13}\mathrm{C}$	$\#^1H$
Zn	308	1592	2900
Cd	304	1564	2860
Hg	304	1564	2860
Group 12	868	4408	7852
Ga	288	2000	3824
In	288	2000	3824
Tl	288	2000	3824
Group 13	804	5484	10452
Ge	592	4948	8780
Sn	592	4948	8780
Pb	592	4948	8780
Group 14	1680	13872	24372
As	424	2712	5384
Sb	424	2712	5384
Bi	424	2712	5384
Group 15	1260	8088	16008
Se	816	4372	6432
Te	788	4248	6264
Group 16	1440	7640	11424
Cl	888	4320	4952
Br	776	3856	4516
Ι	808	3984	4608
Group 17	2264	11004	13084
All	6388	38740	64436

Table S1: Numbers of structures (optimized and distorted) and chemical shifts that are included in the complete data set and consist of at least one atom of the respective element of group of elements.

slight decrease of the generality of the model (exhibited by a worse performance on the test data set). However, this is a small effect rendering the **structures** mode reasonable for the statistical investigations. Using the **atoms** mode is inherently questionable as all information of the molecular origin of the atoms is lost as the data points form a molecule are spread over both data subsets. Hence, the **structure** mode was applied in all investigations presented in this work (as it was the case for the Δ_{corr} -ML method). Note that for this reason, a somewhat reduced performance of the Δ_{SO} -ML method is expected for structures with low similarity to the training data set as it is the case for the investigated organotin and -lead compounds and for the compounds in the 17HAC test set. For more information on the data set structure and shuffling modes, we refer to the investigations in the Supporting Information for the Δ_{corr} -ML method.^{S1}

2.3 Input feature vector

The relative NMR chemical shift δ of a compound (c), that can be measured in an experiment relative to an external reference compound (ref.), is also computed as the difference of the respective absolute shielding constants σ :

$$\delta_{\rm c} = \sigma_{\rm ref.} - \sigma_{\rm c} \tag{7}$$

In all NMR shift calculations of this work, a tetrymethylsilane (TMS) molecule was treated in the exact same way as the actual compound. This holds both for the structures in the data set and for the conformer ensembles for the comparison with experimental data.

A complete list of the descriptors that have been included in the input feature vector of the neural network of the Δ_{SO} -ML scheme is given in Table S2. In addition to the list, definitions for the *magnetic* input features are given in the following. All descriptors in this group are derived from the chemical shielding tensor that is obtained from a standard (*low-level*) NMR shielding calculation. The definitions for the *magnetic* quantities are taken from the Δ_{corr} -ML model and are recapitulated below.

The total NMR shielding tensor σ is defined as

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$
(8)

with

$$\sigma_{11} \le \sigma_{22} \le \sigma_{33} \quad . \tag{9}$$

The total isotropic shielding constant σ_{iso} is the trace of the shielding tensor:

$$\sigma_{\rm iso} = \frac{\sigma_{11} + \sigma_{22} + \sigma_{33}}{3} \quad . \tag{10}$$

The span Ω , skew κ , asymmetry η and anisotropy Δ result from σ and σ_{iso} as defined in Table S2. Furthermore, the dia- and paramagnetic parts $\sigma_{iso,dia}$ and $\sigma_{iso,para}$ of the total isotropic shielding constant are also included in the input feature vector.

Table S2: List of variables included in the input feature vector of the $\Delta_{\rm SO}$ -ML model for ¹H and ¹³C chemical shift SO contribution prediction in three groups. The term $n^{\rm th}$ bonding sphere refers to all atoms within a distance of n covalent bonds to the origin atom. Details of the equations in the magnetic properties group can be found above.

Designation	$^{1}\mathrm{H}$	$^{13}\mathrm{C}$		
Target: $\Delta_{\rm SO}\delta = \delta_{\rm SO} - \delta_{\rm SR}$ with $\delta = \sigma_{\rm ref.} - \sigma$	\checkmark	\checkmark		
geometric properties				
Coordination number as defined in D3 model $(CN(D3))$				
CN(D3) of the neighboring atom				
Distance H–X to neighboring atom X	\checkmark			
Number of neighboring atoms of type:				
H, C, N, O, S, Cl, Zn, Ga, Ge, As, Se, Br, Cd, In, Sn, Sb, Te, I, Hg, Tl, Pb, Bi		\checkmark		
Number of atoms of type in the second bonding sphere:				
H, C, N, O, S, Cl, Zn, Ga, Ge, As, Se, Br, Cd, In, Sn, Sb, Te, I, Hg, Tl, Pb, Bi	\checkmark	\checkmark		
Atom-centered symmetry functions (ACSF) G^1 with cutoff 5.0 angstroms including:				
H, C, N, O, Si, P, S, Cl, Zn, Ga, Ge, As, Se, Br, Cd, In, Sn, Sb, Te, I, Hg, Tl, Pb, Bi	\checkmark	\checkmark		
Each for 1^{st} and 2^{nd} bonding sphere:				
Sum of atomic numbers	\checkmark	\checkmark		
Sum of atomic masses	\checkmark	\checkmark		
Each for 1^{st} , 2^{nd} , 3^{rd} , and 4^{th} (and 5^{th} for ¹ H) bonding sphere:				
Number of HAs	\checkmark	\checkmark		
Average of atomic masses	\checkmark	\checkmark		
For every HA in direct vicinity: [*]				
Atomic number of the HA	\checkmark	\checkmark		
CN(D3) of the HA	\checkmark	\checkmark		
electronic properties				
Each for Mulliken and Löwdin model:				
Atomic charge	\checkmark	\checkmark		
s- and p-orbital populations	\checkmark	\checkmark		
d-orbital population		\checkmark		
Standard deviation of p-orbital populations (p_x, p_y, p_y)		\checkmark		
Each for Löwdin and Mayer model:				
Sum of bond orders	\checkmark	\checkmark		
Average of bond orders		\checkmark		
Mayer's total valence	\checkmark	\checkmark		

Table S2: Continued.						
Designation 1						
For every HA in direct vicinity:*						
Mulliken atomic charge of the HA						
Mulliken s -, p -, and d -orbital occupations in the valence shell of the HA	\checkmark	\checkmark				
magnetic properties						
Chemical shift δ_{low} from <i>low-level</i> (non-/SR-relativistic) DFT calculation	\checkmark	\checkmark				
Span: $\Omega = \sigma_{33} - \sigma_{11}$	\checkmark	\checkmark				
Skew: $\kappa = \frac{3(\sigma_{\rm iso} - \sigma_{22})}{\Omega}$	\checkmark	\checkmark				
Asymmetry: $\eta = \frac{\sigma_{22} - \sigma_{11}}{\sigma_{33} - \sigma_{iso}}$	\checkmark	\checkmark				
Anisotropy: $\Delta = \sigma_{33} - \frac{\sigma_{11} + \sigma_{22}}{2}$	\checkmark	\checkmark				
Dia- and paramagnetic part of the shielding constant $\sigma_{\rm iso,dia}$, $\sigma_{\rm iso,para}$	\checkmark	\checkmark				

^{*}For ¹H: HA bound directly to ¹H or via two bonds (max. 3 HAs), for ¹³C: HA bound directly to ¹³C (max. 4 HAs).

3 Supplementary Data

3.1 Analysis of the data set

Complementarily to the analyses of the data set in the manuscript (see Fig. 3), more visualisations are presented herein. In the following figures, the color codes corresponds to the distance of the ${}^{13}C/{}^{1}H$ nuclei to the closest heavy atom (HA) in terms of numbers of covalent bonds. For a clearer picture, the color code is depicted in Figure S1.

A representation of the complete data set for ¹H NMR is shown in Figure S2 (analogous to Fig. 3). Separate and more detailed analyses of the data set for all three distance categories are given in Figures S3 to S5 for ¹³C and in Figures S6 to S8 for ¹H.



Figure S1: Example molecule showing the color code of the distance categories used throughout. (a) ¹³C, division in one (blue), two (yellow), and three or more (gray) covalent bonds between ¹³C and the HA (= Se, Te), (b) ¹H, division in one or two (blue), three (yellow), and four or more (gray) covalent bonds between ¹H and the HA. The indifferent H/C atoms are shown in white. (a) is equivalent to Fig. 3(a).



Figure S2: Complete data set with 64436 ¹H NMR chemical shifts showing the relation between the purely scalar-relativistic NMR shift δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the relativistic reference $\delta_{SO} = \delta_{SR} + \Delta_{SO}\delta$ with the SO contribution calculated at the PBE0/TZ2P level. For the color code, see Figure S1(b).



Figure S3: NMR shifts from ¹³C nuclei connected to a HA via one bond showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.



Figure S4: NMR shifts from ¹³C nuclei connected to a HA via two bonds showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.



Figure S5: NMR shifts from ¹³C nuclei connected to a HA via three or more bonds showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.



Figure S6: NMR shifts from ¹H nuclei connected to a HA via one or two bonds showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.



Figure S7: NMR shifts from ¹H nuclei connected to a HA via three bonds showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.



Figure S8: NMR shifts from ¹H nuclei connected to a HA via four or more bonds showing the correlation between δ_{SR} calculated with the *low-level* PBE0/ZORA-def2-TZVP method and the target value $\Delta_{SO}\delta$ calculated at the PBE0/TZ2P level.

3.2 Benchmark studies

3.2.1 Organotin and organolead compounds

For the evaluations of the organotin and -lead compounds, two benchmark sets were investigated. The organotin compounds were taken from the SnS51 set S2 and for the organolead compounds, a successor study is in progress during the publication process of this work. The benchmark study for ²⁰⁷Pb NMR chemical shifts contains a similar number of experimentally accessible organolead structures, for which conformer ensembles were generated. In both cases (Sn and Pb), only the conformer lowest in Gibbs free energy was taken from each ensemble and the $Sn(CH_3)_4$ and $Pb(CH_3)_4$ compounds were included as well. Then, the reference value of $\Delta_{SO}\delta$ was calculated at the level used throughout the study (SO-ZORA-PBE0/TZ2P). Some compounds were omitted due to technical issues including convergence problems of the SO-PBE0 calculation and the presence of atoms that were not included in any training data structure (such as transition metal atoms). In the case of the SnS51 set, the following compounds were not included here: 1, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 18, 22, 29, 41, 47, 48. In total, this yields 34 Sncontaining compounds with a total number of 817 calculated ¹³C and 1170 ¹H NMR shifts (without averaging of chemically equivalent atoms). For the Pb test set, some compounds were excluded for the same reason as for the Sn-containing compounds. For a detailed analysis of the structures, we refer to the upcoming publication. In total, 48 compounds were included with 1415 calculated ¹³C and 2059 ¹H NMR shifts (without averaging of chemically equivalent atoms). All structures are included as Cartesian coordinate files (.xyz) in the supplementary .zip archive.

In addition to the results presented in Fig. 8, the evaluation of the ¹H NMR shift data for the SO-HALA effects of Sn and Pb atoms is presented in Figure S9.



Figure S9: Comparison of ¹H NMR metrics for the H–Sn and H–Pb test structures without any SO contribution (PBE0/ZORA-def2-TZVP, full colors) and with the MLpredicted $\Delta_{SO}\delta$ values (brighter colors). The data is averaged over ¹H nuclei bound to a HA via one/two (H-X_{0/1}-HA), three (H-X₂-HA), and four or more bonds (H-X₃-HA) and over the full data (all).

3.2.2 17HAC test set

The 17HAC test set was compiled to address the heavy atom effect on 13 C (HAC) for all 17 HAs included in this study. For this, 13 C NMR data for 63 (metal-)organic compounds were collected yielding a total number of 236 experimental 13 C NMR shifts. For each heavy element, at least three compounds that contain it are included in the 17HAC set. The 13 C NMR shifts originate from atoms at different distances from the HA (all categories shown in Figure S1(a) are represented) and the subsequent computational investigation revealed that the magnitude of the HALA effect to 13 C varies throughout the range of compounds. A complete list of the compounds included in the 17HAC set and some further details is given in Table S3. The chemical shift data itself and the experimental references are provided in the supplementary raw data (.xlsx) file.

The workflow used for the *low-level* NMR chemical shift calculations is a commonly applied way to incorporate solutions to the typical error sources of *conformational flexibility* and solvation effects into the computation. For each compound, a conformer ensemble was generated and refined (here with CREST and CENSO, see the manuscript) and the results of the subsequent NMR shielding calculations applying the implicit CPCM solvation model were Boltzmann averaged (at 298 K) to obtain the chemical shielding constant for each atom. The reference was tetramethylsilane (TMS) for all compounds and was treated in the same way in the respective solvent. The final computed NMR chemical shift was obtained via equation (7) and the values for all chemically equivalent ${}^{13}C$ nuclei were averaged to be comparable to the experimental value. To disentangle the (ML) correlation and SO contributions to the NMR shift from the solvation contribution, the ML models are always applied to chemical shift values from calculations without a solvent model before the contributions are added to the values including solvation. Therefore, all chemical shifts were recalculated without applying the CPCM model before both Δ -ML methods were applied (separately) to each chemical shift value of every conformer. In this way, the final chemical shift value can be divided into the *low-level* value (including conformational and solvation effects) and the correlation as well as the spin-orbit contribution for the results obtained by the Δ_{corr} - and the Δ_{SO} -ML methods, respectively. For this purpose, all chemical shift calculations for the data set of the Δ_{corr} -ML method were performed at the PBE0/ZORA-def2-TZVP level of theory and this new low-level data was used to retrain the Δ_{corr} -ML model. Finally, after the presented workflow was applied to every compound in the 17HAC set, the statistics were obtained that can be found in Table 3.

Table S3: List of the compounds of the 17HAC data set including the solvent used in the NMR measurement, the number of conformers obtained from the workflow described above, the number of experimentally obtained ¹³C NMR chemical shifts, and their respective reference.

#	Compound	Solvent	#Conformers	$\#\delta_{\mathrm{exp}}(^{13}\mathrm{C})$	Ref.
1		CHCl ₃	5	3	S3
2		CHCl ₃	1	3	S3
3	SiMe ₃ ⁷ 'SiMe ₃ ⁷ 'SiMe ₃ ⁷ 'SiMe ₃ ^N ^N ^N ^N ^N ^N	benzene	4	4	S4
4		CHCl ₃	30	3	S3
5		CHCl ₃	22	3	S3
6	SiMe 3 Cd Me 3Si Me 3Si	benzene	8	4	S4
7	-Hg-===	CHCl_3	1	3	S5
8	⟨	CHCl_3	1	6	S5
9	Hg-Cl	DMSO	1	4	S6
10	Hg-Cl	CHCl_3	2	2	S7
11	Hg	CHCl ₃	21	2	S7

Table S3: Continued.					
#	Compound	Solvent	#Conformers	$\# \delta_{\mathrm{exp}}(^{13}\mathrm{C})$	Ref.
12	CI CI Ga CI	CHCl ₃	1	6	S8
13	Br Ga Br	CHCl ₃	8	6	S8
14	F Ga F	CHCl ₃	3	4	S8
15	Ga	CHCl ₃	2	5	S8
16		benzene	3	7	S9
17	N N N	benzene	29	9	S9
18		benzene	1	3	S10
19	Br TI	DMSO	144	3	S11
20	Br TI	DMSO	7	2	S11
21	Br TI	DMSO	15	3	S11

		Table S3	S: Continued.	(12 c)	
#	Compound	Solvent	#Conformers	$\#\delta_{\exp}(^{13}\mathrm{C})$	Ref.
22	Ge	CHCl_3	7	6	S12
23	Ge Ge	CHCl ₃	5	5	S13
24	Ge-Ge-	CHCl ₃	17	4	S14
25	Sn Sn	CHCl_3	7	6	S12
26		CHCl ₃	10	5	S13
27	Sn-Sn-	CHCl ₃	31	4	S14
28	Pb I	CHCl_3	8	6	S12
29	Pb C	CHCl ₃	14	5	S13
30	Pb-Pb	CHCl ₃	40	4	S14
31	 _As _	toluene	3	2	S15
32	∣ ∠ ^{As}	CHCl ₃	2	1	S15

m 11 a α 1

	Table 53: Continued.							
#	Compound	Solvent	#Conformers	$\#\delta_{\exp}(^{13}\mathrm{C})$	Ref.			
33		CHCl ₃	5	4	S15			
34	As	CHCl ₃	1	3	S15			
35	-Sb.	CHCl ₃	3	1	S16			
36	Br Sb Br	CHCl ₃	3	4	S16			
37		CHCl_3	7	5	S16			
38	Bi—Bi	toluene	7	1	S17			
39	Bi	toluene	31	3	S17			
40	Bi-Bi	toluene	53	3	S17			
41		CHCl ₃	27	3	S17			
42	Bi-Bi	benzene	5	2	S17			

-	#	Compound	Solvent	#Conformers	$\#\delta_{\exp}(^{13}\mathrm{C})$	Ref.
-	43	Se	THF	2	1	S18
	44	SeSeSe	THF	4	1	S18
	45	Se Se	CHCl ₃	1	4	S19
	46	Se	$\mathrm{CH}_{2}\mathrm{Cl}_{2}$	1	9	S19
	47	_Te_	benzene	2	1	S18
	48	_TeTe_	CHCl_3	5	1	S18
	49	$F \xrightarrow{Te} F$ $F \xrightarrow{Te} F$	CHCl ₃	1	1	S18
	50	Те	CHCl ₃	1	4	S19
	51		-	1	2	S20
	52	CI	-	1	2	S20

Table S3: Continued.					
#	Compound	Solvent	#Conformers	$\#\delta_{ m exp}(^{13} m C)$	Ref.
53	≡− CI	-	1	2	S20
54	CI	-	1	4	S20
55	Br Br Br	-	1	1	S20
56	$=$ C $\stackrel{\text{Br}}{\underset{\text{H}}{\leftarrow}}$	-	1	3	S20
57	HN Br Br	-	1	8	S20
58	Br Br Br Br	-	1	5	S20
59		-	1	2	S20
60	I— <u>—</u> —I	-	1	1	S20
61		-	1	7	S20
62		-	55	9	S20
63	Br	benzene	1	6	S21

4 References

- (S1) Kleine Büning, J. B.; Grimme, S. Computation of CCSD(T)-Quality NMR Chemical Shifts via Δ-Machine Learning from DFT. J. Chem. Theory Comput. 2023, 19, 3601–3615.
- (S2) Stückrath, J. B.; Gasevic, T.; Bursch, M.; Grimme, S. Benchmark Study on the Calculation of ¹¹⁹Sn NMR Chemical Shifts. *Inorg. Chem.* **2022**, *61*, 3903–3917.
- (S3) Atwood, J. L.; Berry, D. E.; Stobart, S. R.; Zaworotko, M. J. Aspects of Organocadmium Chemistry. 3. Chelate-Supported Alkyls and Aryls of Zinc, Cadmium, and Mercury: Crystal and Molecular Structure of Bis[2-((dimethylamino)methyl)phenyl]mercury(II). *Inorg. Chem.* **1983**, 22, 3480–3482.
- (S4) van den Ancker, T. R.; Engelhardt, L. M.; Henderson, M. J.; Jacobsen, G. E.; Raston, C. L.; Skelton, B. W.; White, A. H. Syntheses and structures of highly hindered N-functionalised alkyl and amido group 12 complexes MR₂ (M=Zn, Cd, and Hg), [MRCl]₂ (M=Zn and Hg). J. Organomet. Chem. **2004**, 689, 1991–1999.
- (S5) Sebald, A.; Wrackmeyer, B. ¹³C and ¹⁹⁹Hg Nuclear magnetic resonance study of alkynyl mercury(II) compounds. Spectrochim. Acta Part A 1982, 38A, 163–173.
- (S6) Michel, E.; Perie, J.; Lattes, A. NMR investigation of aryl- and benzyl-mercuric compounds; experimental evidence for σ-π conjugation involving the carbonmercury bond. J. Organomet. Chem. 1981, 204, 1–12.
- (S7) Browning, J.; Goggin, P. L.; Goodfellow, R. J.; Hurst, N. W.; Mallinson, L. G.; Murray, M. Preparation and Vibrational Spectra of Tetra-n-propylammonium Carbonyldichloro-organoplatinate(II) Salts, and a Comparison of their Carbon-13 and Platinum-195 Nuclear Magnetic Resonance Properties with those of Organomercury Compounds. J. Chem. Soc., Dalton Trans. 1978, 872–876.
- (S8) Freeman, W. H.; Miller, S. B.; Brill, T. B. Metal-Carbon Interactions in Organogallium and Organoindium Compounds. A ¹³C NMR and ^{69,71}Ga and ¹¹⁵In NQR Study. J. Magn. Reson. **1975**, 20, 378–387.
- (S9) Schumann, H.; Hartmann, U.; Wassermann, W. Synthesis and Characterization of Organogallium and Organoindium Compounds with Tridentate 2,6-Bis[(dialkylamino)methyl]phenyl Ligands. Chem. Ber. 1991, 124, 1567–1569.

- (S10) Kim, S.-J.; Yang, N.; Kim, D.-H.; Kang, S. O.; Ko, J. New Types of Base-Stabilized Alkyl Aluminum, Gallium, and Indium Complexes. Organometallics 2000, 19, 4036–4042.
- (S11) Brady, F.; Matthews, R. W.; Thakur, M. M.; Gillies, D. G. Carbon-13 and proton NMR parameters of mono- and di-organothallium(III) derivatives. J. Organomet. Chem. 1983, 252, 1–27.
- (S12) Adcock, W.; Cox, D. P.; Kitching, W. Metalloidalmethyl Substituent Effects by ¹³C and ¹⁹F NMR: Electron Release in the Neutral Ground State. J. Organomet. Chem. 1977, 133, 393–422.
- (S13) Charissé, M.; Roller, S.; Dräger, M. Tetra-p-tolyl-Verbindungen p-Tol₄Si und p-Tol₄Ge: Ein Beitrag zur Konfiguration der Tetraaryl-Methan-Analoga Ar₄M (M = C, Si, Ge, Sn, Pb). J. Organomet. Chem. **1992**, 427, 23–31.
- (S14) Vaickus, M. J.; Anderson, D. G. Investigation of Phenyl Derivatives of Group IV Elements using Carbon-13 NMR Spectroscopy. Org. Magn. Reson. 1980, 14, 278– 279.
- (S15) Tähtinen, P.; Saielli, G.; Guella, G.; Mancini, I.; Bagno, A. Computational NMR Spectroscopy of Organoarsenicals and the Natural Polyarsenic Compound Arsenicin A. Chem. Eur. J. 2008, 14, 10445–10452.
- (S16) Yanaga, M.; Miura, T.; Endo, K.; Nakahara, H.; Takeda, M. ¹²¹Sb Mössbauer and ¹H and ¹³C NMR Spectroscopic Studies of a Series of Organoantimony(V) Compounds, R₃SbX₂. Bull. Chem. Soc. Jpn. **1986**, 59, 3085–3089.
- (S17) Ashe III, A. J.; Ludwig Jr, E. G.; Oleksyszyn, J. Preparation and Properties of Dibismuthines. Organometallics 1983, 2, 1859–1866.
- (S18) Rusakov, Y. Y.; Rusakova, I. L. Relativistic heavy atom effect on ¹³C NMR chemical shifts initiated by adjacent multiple chalcogens. *Magn. Reson. Chem.* 2018, 56, 716–726.
- (S19) Rusakov, Y. Y.; Rusakova, I. L.; Krivdin, L. B. On the significant relativistic heavy atom effect on ¹³C NMR chemical shifts of β- and γ-carbons in seleno- and telluroketones. *Mol. Phys.* **2017**, *115*, 3117–3127.
- (S20) Casella, G.; Bagno, A.; Komorovsky, S.; Repisky, M.; Saielli, G. Four-Component Relativistic DFT Calculations of ¹³C Chemical Shifts of Halogenated Natural Substances. *Chem. Eur. J.* **2015**, *21*, 18834–18840.
- (S21) Bagno, A.; Rastrelli, F.; Saielli, G. Predicting ¹³C NMR Spectra by DFT Calculations. J. Phys. Chem. A 2003, 107, 9964–9973.