Instructions for using the new Excel file to compute the DP4+ probability

- In order to simplify the new DP4+ calculations, an Excel file is provided from the authors which can be downloaded from:

http://www.sarotti-NMR.weebly.com http://www.iquir-conicet.gov.ar/Sarotti-NMR

- The file contains 7 sheets (Figure S1).

- Main: contains the selection of the level of theory, and enabled cells to introduce the experimental and calculated NMR data.
- Unscaled: shows the unscaled chemical shifts, computed according to $\delta_{calc} = \sigma_{TMS} \sigma_x$
- Scaled: shows the scaled chemical shifts, computed according to $\delta_s = (\delta_{calc} b)/m$
- Unscaled Errors: shows the unscaled errors, computed according to $e_u = \delta_{calc} \delta_{exp}$
- Scaled Errors: shows the scaled errors, computed according to $e_s = \delta_s \delta_{exp}$
- Detailed Results: shows the scaled, unscaled, and full DP4+ results (sDP4+, uDP4+, and DP4+, respectively).
- *Parameters Estimation*: contains the experimental data of the 8 selected molecules shown in Figure 5, and enabled cells to introduce the calculated NMR shifts required to estimate the $[\mu,\sigma]$ terms.

The "Main" sheet

- The "Main" sheet contains the enabled cells to introduce the experimental and calculated NMR data, and provides the results of the DP4+ calculations. It is divided in 4 different regions: A, B, C, and D.





Region A: user selection (drop down lists):

- Settings: settings to be employed to compute the DP4+ probability (Default or Custom). The choice will be reflected in Region B.

- "Default" option allows to use the $[\mu,\sigma,\nu]$ terms calculated for a large set of 77 molecules at the 24 levels of theory evaluated in the original publication (JOC **2015**, *80*, 12526). If this option is selected, a new list is displayed with the following options:

- Functional: functional used to perform the GIAO NMR calculations (B3LYP or mPW1PW91).

- Solvent?: inclusion of solvent effect in the GIAO NMR calculations (PCM or Gas Phase).

- Basis Set: basis set used in the GIAO NMR calculations (6-31G*, 6-31G**, 6-31+G**, 6-311G*, 6-311G** or 6-311+G**).

After the selection is done, the corresponding $[\mu,\sigma,v]$ values and the isotropic shielding values for the reference standard (Tetramethylsilane, TMS) will be displayed in Region B.

- "Custom" option allows to set the values of the $[\mu,\sigma,\nu]$ terms, which could be estimated in different ways. For example, using the full set of 77 molecules evaluated in the original publication at different levels of theory, using a large set of different molecules, or using the reduced set of 8 molecules shown in Figure 5 for preliminary calculations. If this option is selected, the 16 values of the $[\mu,\sigma,\nu]$ set and the isotropic shielding values for the reference standard (TMS) should be provided in the enabled cells of Region B.

Region B: settings to compute the DP4+ probability

Contains the $[\mu,\sigma,\nu]$ parameters required to perform DP4+ calculations. These values depend upon the level of theory employed during the calculations, and can be introduced in two different ways:

- If "Default" option is selected in Region A, the Default settings of the selected level of theory will be shown in Region B.

- If "Custom" option is selected in Region A, the Custom settings will be highlighted in Region B. These settings can be changed freely depending on the user preferences (for example, those obtained by using other levels of theory and/or a different set of test molecules). Warning: to obtain meaningful results, the modified parameters must be estimated at the desired level of theory using a large set of known compounds, or the recommended set of 8 molecules for preliminary calculations.



Region C: enabled cells to enter the data:

- *Nuclei:* enter the letter "C" (without quotes) for ¹³C data or the letter "H" (without quotes) for ¹H data.

- *sp2?:* enter the letter "x" (without quotes) if the carbon is sp-sp² hybridized (or if the proton is attached to an sp-sp² hybridized carbon).

- *Experimental:* enter the experimental ¹³C and/or ¹H chemical shifts.

- Isomer 1, Isomer 2, etc: enter the computed data for the candidate structures (either as isotropic shielding values, unscaled chemical shifts or scaled chemical shifts).

Region D: displays the DP4+ probabilities once the data are entered in Regions A, B and C.

Tutorial 1: DP4+ calculations using the "Default" option.

Calculated NMR data: compounds **T1-T8** at the PCM/mPW1PW91/6-31+G**//B3LYP/6-31G* level. Experimental NMR data from compound **T1** (correct isomer). *Synthesis* **2019**, *51*, 1545.



Step 1: perform a full conformational search using a suitable package (Hyperchem, Spartan, MacroModel, etc). It is recommended to keep all conformations within 5 kcal/mol (~21 kJ/mol) from the global minimum.

Step 2: optimize all conformations found at the B3LYP/6-31G* level (gas phase). Remove duplicates if necessary.

Step 3: perform the NMR GIAO calculations at the selected level of theory (in this case, at the PCM/mPW1PW91/6-31+G**) for all representative conformations. Compute the Boltzmann-averaged shielding values for each compound.

Step 4: open the Excel file, and select the "Default" and "Shielding tensors" options in Region A, and select the level of theory that was used to perform the NMR calculations. In this case, Functional > mPW1PW91; Solvent > PCM; Basis set > 6-31+G**.



Step 5: Introduce the experimental data. Important: mark the sp² nuclei with "x". Otherwise, improper distributions will be used for those nuclei, with potentially negative effects in the DP4+ values.

	A	2	1		1		G	H		1		1	-M	N.	Ó	2	Ó.	ŧ	5	÷Ť	0		w	. 9	¥	2	AA .
1		Sculege		- 14	pc of dits (sh	ns)		TMS III	31.560	TIMS \$3C	196.600		1942-224		786 190												
4		peans			weiding trins	161		Delait		9.246	F 164				_												
								1304.001	2,909	1.500	6,269																
1	Failer	tinut .	Sei	eent?	Los	s feet		181.007	0.547	0.138	4411		Manual														- 1
£.,	m9W1	PWBL		CM .	6.21-	(ad ₁ 2)		18-u.qull	-0014	0.112	0.451		120.004														
r.								190		1.553	6.227																
-		Internet MP			1 1		1.4	Line C	-	0.104	3.89		1. 10		12	4.8	14	*1	16	17	18	10	20	21	22	22	- 14
10		in the second	H data			0	-			1							- 14										
	DP4	(4)	C date				-		-				-		1	-		-	-							-	
12			All Jole																								
13	Type	5927	Exp	1	2	3	4	5	6	7.	1	9	10	11	12	15	14	15	16	17	18	19	20	21	22	23	24
14	c		164.7																								
15	6		121.5																								_
17	e .	-	29.5																								_
18	c		75.2																								_
19	c		40.6																								
in i	c		21.6																								
15	c		40.8																								
22	C		60.1																								
15	e .		41.4																								
-	6		139.6																								_
76	è	- Q - 1	1301																								_
:9	c		104.5																								
28	c		\$26.7																								
29	C		1.28.8																								_
50	c		128.0																								_
11																											
2			805																								_
			7.40																								
15	34		4.67																								
16	- 14		2.04																								_
57	H		373																								
18	24		1.50																								
22	. 64		2.04																								
-11			391																								
12			1.5/																								_
10	44		401																								
4	- 14		619																								
45	111		650																								
0	**	*	/58																								
47	н	ж	751																								
	H	×	724	1																							-
	all same	100000000000000000000000000000000000000	a contraction	A designed and the	Andrea 7 of mark	and Products	Company States of the				1.0						100										1000

Step 6: Introduce the calculated averaged isotropic shielding values for each possible candidate structure (8 in this case). The data can be inserted manually or copy-pasted. The DP4+ values are automatically calculated and displayed for each isomer above the number of isomer. Important: the NMR data must be assigned (know which shift corresponds to which nuclei). Using unassigned or misassigned NMR data might lead to erroneous results.

Image: Proper term Image: Propertee term Image: Propropertee term Image: Propertee t															The In Decem 200,007 200,00000000		5.354 6.269 4.911 3.551 6.227 3.893	0 1745 1000 0118 0112 1557 0108	4.000 4.000 5.347 4.038	EXCarpt EXCarpt EXcarpt EXcarpt Excarp	Deletit Deletit Solverst? Solverst? NV92 PCM 6-33-rC(dp)							
bit bit <th>2.5</th> <th>ž I</th> <th>22</th> <th>23</th> <th>20</th> <th>19</th> <th>15</th> <th>17</th> <th>16</th> <th>15</th> <th>14</th> <th>12</th> <th>12</th> <th>11</th> <th>90</th> <th>,</th> <th></th> <th>7</th> <th>6</th> <th>5</th> <th>. 4</th> <th></th> <th>1</th> <th>. L</th> <th></th> <th>isomer N1</th> <th></th>	2.5	ž I	22	23	20	19	15	17	16	15	14	12	12	11	90	,		7	6	5	. 4		1	. L		isomer N1		
0000-1000 10000			-			1.4	- 14 C										0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	rs darta	100000	0.000	
b b b b b b b b b c		-	-	-								-					1150%	0.00%	DOON	C.00%	0.00%	0.00N	3.00%	88.50%	Cdata	1951	D94+	
ype upp upp <th>-</th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th>100</th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th>0.00%</th> <th>0.00%</th> <th>0.00%</th> <th>CON</th> <th>0.000</th> <th>0.00%</th> <th>3,005</th> <th>100.00%</th> <th>All data</th> <th>201</th> <th>- 11</th>	-									100							0.00%	0.00%	0.00%	CON	0.000	0.00%	3,005	100.00%	All data	201	- 11	
C N Dial Dial <thdial< th=""> Dial Dial<!--</th--><th>25</th><th>2</th><th>. 12</th><th>21</th><th>20</th><th>1.5</th><th>13</th><th>17.</th><th>16</th><th>15</th><th>14</th><th>11</th><th>12</th><th>- 11</th><th>\$0</th><th>3</th><th>0</th><th>7</th><th>6</th><th>5</th><th>4</th><th>3</th><th>2</th><th>1</th><th>Dep</th><th>5p21</th><th>type</th></thdial<>	25	2	. 12	21	20	1.5	13	17.	16	15	14	11	12	- 11	\$0	3	0	7	6	5	4	3	2	1	Dep	5p21	type	
L 4 143 843 1433 8438 1438 <td></td> <td>15.2799</td> <td>35 3068</td> <td>351703</td> <td>35,3031</td> <td>33.6334</td> <td>33.0152</td> <td>35.1097</td> <td>15.2715</td> <td>164.7</td> <td>×</td> <td>C</td>																	15.2799	35 3068	351703	35,3031	33.6334	33.0152	35.1097	15.2715	164.7	×	C	
C A M																	10.0452	10.0370	170432	10.0018	79.6905	79.5460	70,1691	10.5215	241.3		C	
C 71 1000 7000 1000																	48.3170	165 4041	48.9189	49.0018	49.049/	169 1019	49.21.19	48.8133	349.7		6	
C M																	122.0142	120 1148	128.8308	121 2085	110.1741	110 2607	1107386	110.6271	21.2		c	
C																	152.7645	331.5483	133.8534	152 8521	154.8932	103.4703	133 0527	154.0191	43.8		c	
C 4.8 5.8 5.8 5.98 5.988 C 7.8 1.8 1.901 1.912 1.913 <th< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>1/5.5518</td><td>327,4020</td><td>12/ 95/8</td><td>L28.7540</td><td>131.5305</td><td>178.1555</td><td>1192275</td><td>1142528</td><td>/1.0</td><td></td><td>C</td></th<>																	1/5.5518	327,4020	12/ 95/8	L28.7540	131.5305	178.1555	1192275	1142528	/1.0		C	
C 01.1 15.82 12.02 12.08 <th12.08< th=""> 12.08 12.08</th12.08<>																	153,9809	157.3574	154.3799	155.9638	156.3156	158.0217	1540617	1545143	42.8		с	
C 41 8 80 8 80 8 80 8 80 8 80 8 80 8 80 8 80 8 80 8 80 8 80 8 80																	126.8877	129.7386	128.7138	128.6133	128.5496	119.9876	1102075	1165342	65.1		c	
C 73 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 10 40 10 40 C × 10 40 10 40 10 40 10 40 10 40 10 40 10 40 V 10 40 10 40 10 40 </td <td></td> <td>152,5145</td> <td>157.1082</td> <td>155.0534</td> <td>195.3461</td> <td>107.4142</td> <td>157.1225</td> <td>1567364</td> <td>1138013</td> <td>41.4</td> <td></td> <td>C</td>																	152,5145	157.1082	155.0534	195.3461	107.4142	157.1225	1567364	1138013	41.4		C	
L A L03 B-L03 B-L04 B-L04 <td></td> <td>1179665</td> <td>321.7159</td> <td>121.0756</td> <td>121.5499</td> <td>121 1009</td> <td>110 3567</td> <td>110 6269</td> <td>117/8454</td> <td>75.3</td> <td></td> <td>c</td>																	1179665	321.7159	121.0756	121.5499	121 1009	110 3567	110 6269	117/8454	75.3		c	
C A 164 1.74 81.48 8.6.74 <th< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>\$8.8314</td><td>68 19/9</td><td>07.1246</td><td>01.2012</td><td>59.5105</td><td>67.0011</td><td>05.5005</td><td>10.12/3</td><td>149.5</td><td></td><td>C.</td></th<>																	\$8.8314	68 19/9	07.1246	01.2012	59.5105	67.0011	05.5005	10.12/3	149.5		C.	
c x 1144 114																	61 0003	61 7320	619/04	61 6176	6' 4MB	62 4551	61 6601	61.7748	116.8	- C	c	
C A 113 103 713 <th713< th=""> <th713< th=""> <th713< th=""></th713<></th713<></th713<>																	72 1677	12 2106	22.0180	71 8008	7: 5645	21 2206	72 14:0	12 0885	116.7		6	
C A 100																	20.2471	10.0987	705046	70,4125	70.2138	72,3651	70.0340	70.3057	178.8		c	
1 2 6.00 3.01 3.01 3.01 3.00<																	71.3772	710125	71.5129	71.4770	7.2991	71.2468	71.2808	11.3482	138.0		с	
N N O TAM																												
N N 080 74.00 24.00 <td></td> <td>25.8475</td> <td>25.4477</td> <td>25.4654</td> <td>23.4450</td> <td>25.4150</td> <td>25.3759</td> <td>25.4171</td> <td>25.4392</td> <td>603</td> <td>: X</td> <td>H</td>																	25.8475	25.4477	25.4654	23.4450	25.4150	25.3759	25.4171	25.4392	603	: X	H	
x 34 54.2																	24,2553	14,2505	242303	24,2547	24.2738	24.3791	24.2187	24.2301	6.89	×	н	
N C																	10 8247	10 1615	291120	20.2163	24/0612	22.0328	29.15.55	20:1424	246		**	
n ·																	28.5897	28.9458	267328	28.6582	24.7391	25.7627	28.8107	26.7877	4.67		**	
N 1.20 51.200 51.240 50.240 50.250 59.267 50.271																	17 5984	170875	222620	21 1874	27.0165	27 4445	27 4898	17 7242	373		**	
H 704 N 408 24.04 <th24.04< th=""> 24.04 24.04<</th24.04<>																	30,4731	199667	30,0189	30 1356	29.9573	33 4468	30,4326	30,4080	130		H	
N S0 74.47 74.68 72.49 72.49 73.40 73.41 74.69 74.41 1 1 7 50.32 50.42 50.40 50.44 50.40 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>79 8330</td><td>198253</td><td>500175</td><td>30.0868</td><td>24 9774</td><td>23,87.45</td><td>20 6048</td><td>20 7420</td><td>264</td><td></td><td>H</td></td<>																	79 8330	198253	500175	30.0868	24 9774	23,87.45	20 6048	20 7420	264		H	
I L17 20.02 20.476 20.020 20.476 20.024 20.235 20.000 I - 2.18 2.18 2.28 2.080 2.091 2.010																	17.5432	174025	272415	31.3534	27.8 N6	37.5006	27.4928	37.4875	2.02		84	
N 4.25 7.470 7.470 2.47																	30.3050	30 2 21 9	30.0041	30.0296	23.0704	30.62.92	29.5255	20.3232	137		11	
H 401 J7 498 30 577 32 408 20 895 12 208 27 881 K 6 3 37 73 34 40 20 208 24 991 32 895 17 208 17 881 K 6 3 37 73 34 40 20 208 24 991 32 892 27 891 44 91 K 6 3 37 73 34 40 32 893 32 891 33 891 33 891 <td></td> <td>29.5707</td> <td>198744</td> <td>29.9840</td> <td>29,9308</td> <td>29.5571</td> <td>29.5890</td> <td>29.4657</td> <td>29.6975</td> <td>31.5</td> <td></td> <td>#</td>																	29.5707	198744	29.9840	29,9308	29.5571	29.5890	29.4657	29.6975	31.5		#	
H x 613 24773 34841 24789 34321 24789 34809 34789 34809 34789 34809 34799 34809 34799 34809 34799 34809 34799 34809 34799 34809 34799 35809 35874 35829 35874 35829 35874 35829 35874 35829 <td></td> <td>27.4581</td> <td>17 2 3 9 6</td> <td>26.9153</td> <td>26.9560</td> <td>27.0385</td> <td>25.7093</td> <td>26.7547</td> <td>17.4898</td> <td>401</td> <td></td> <td>H</td>																	27.4581	17 2 3 9 6	26.9153	26.9560	27.0385	25.7093	26.7547	17.4898	401		H	
n s 50.8 32.50% 34.70% 24.70%																	34,7699	14 5312	247896	24.8388	24.193	24,6739	24.6461	34.7728	619	×	#	
m a 794 49.000 a0.770 49.000 20.00 20.000 10.000 30.000 10.0000 10.0000 10.0000 10.0000 10.0000 10.0000 10.0000 10.0000 10.0000																	24,7557	148190	247504	24.2719	24 m204	24 /192	24.7624	24.7565	0.58		-	
The second addres addres addres addres addres addres addres addres																	10.0000	120074	59/0103	21.8165	23.0340	23.8760	10.7770	13 1030	7.00			
N X 7/24 14.000 21.000 21.000 24.000 24.000 24.000 24.000																	10.1100	LIL MARKA	VALTER	14.18.28	20.03943	VALUESO	CT SWID	10.0070	1.74	- 2		

Warning: do not cut-paste information in this Excel file since it modifies the cell references for the hidden equations).

Once the Boltzmann-averaged isotropic shielding values are introduced, the calculated chemical shifts (both unscaled and scaled) are shown in the Unscaled and Scaled sheets, respectively. The corresponding errors can are shown as well in the Unscaled and Scaled errors sheets, respectively.



The probabilities computed using unscaled (*u*DP4+) and scaled (*s*DP4+), as well as using carbon and/or proton data, is shown in the sheet "Detailed Results"

	DIRA D	E.	0	E -		<u>u</u> .	H	1	1			N.	n .	0	۳.	<u>u</u>			1.	U.	V	W	х		£	- M -
1	and the second second second second					1 6	6	2			10			42		16	10		12		20	34	17	12	16	- x []
	+Dht- (H sists)	A 99.37%	0.00%	10.007	10.025	1 C.004	0.00%	8 0.00%	0.03%						-	-	-		-	-	-11					
i I	stand+ (C data)	24.93%	0.19%	0.00%	0.075	C.40%	1.15%	2.10%	21.1.06		1.54	12 4 7			-		-	1	1.4.1		-			-		1
5	sDP4+ (all data)	199.82%	0.00%	0.00%	0.02%	0.005	0.00%	0.00%	0.18%	-		3 4 3	-		1 = 0			1 2 3				5 a 1				100
6	uDN4+ (III data)	199.91%	0.41%	10.00%	JE 8.00%	1 C.00%	0.00%	0.00%	0.07%			8		-	2 8			(() i		() ()			(
	uDP4+ (Coata)	40,6405	0.00%	0.01%	1 0.00%	C.00%	0.00%	0.00%	31,54%			14 m - 1		1.12	$c \rightarrow 1$			1 x 1		1.1.1					1.1	
1	uDP4+ (all data)	99.97%	0.00%	0.00%	0.07%	C.00%	0.00%	0.00%	0.03%	1.6	- 24	N 161 I	-	1.14		(a)	5 - 64 - 1	1.44	1.00	1.10	 e 	1.1	1(6))	1.14	S 82 1	
1	DE4+ (Fidata)	120.00%	0.00%	0.00%	A60.6 1	10.001	0.00%	10.00%	d0.00%																	
13	DP4+ (C data)	4 80.505	0.80%	0.00%	0.02%	C.00%	0.00%	10,0006	11.50%	-		1		1.5	1 - 1			1			-			-		
11	DPS+ (al data)	130.00%	0.40%	0.00%	0.07N	C.00%	0.00%	0.00%	0.00%	-		1					1.1							-		<u> </u>
								lport a suit																		

Tutorial 2: DP4+ calculations using the "Custom" option.

Calculated NMR data: compounds Ta1-Ta2 at the PCM/ ω B97XD/6-311+G**//M06-2X/6-31+G** level.

Experimental NMR data from compound Ta1 (correct isomer). JOC 2015, 80, 12526



Step 1: The computational work should be done as described in the Tutorial 1, using the M06-2X/6-31+G^{**} level for the geometry optimization step, and the PCM/ ω B97XD/6-311+G^{**} for the GIAO NMR calculation step.

Step 2: open the Excel file, and select the "Custom" and "Shielding tensors" options in Region A. This will activate the edition of the $[\mu,\sigma,v]$ values in Region B.

Step 3: compute the isotropic shielding values for the reference standard (TMS) at the same level of theory employed for the test compounds (PCM/ ω B97XD/6-311+G**//M06-2X/6-31+G**), and introduce those values in the upper region of Region B.



Step 4: if the $[\mu,\sigma,\nu]$ parameters have been previously estimated using a large set of molecules at the selected level of theory, fill the data in the lower part of Region B. If not, for preliminary calculations the values can be estimated with the 8 molecules shown in Figure 5. The NMR chemical shifts of compounds **49-56** must be computed at the PCM/ ω B97XD/6-311+G**//M06-2X/6-31+G** level of theory, and the data must be filled in the "Parameters Estimation" sheet. Once all the σ_x values are introduced (along with the σ_0 values of TMS), the unscaled and scaled chemical shifts will be automatically computed, and the $[\mu,\sigma]$ values will be shown. Important: the NMR data must be assigned (know which shift corresponds to which nuclei). It is essential to respect the assignment indicated in the Sheet. Using other assignments will lead to erroneous results.



Step 5: complete the Region B with these $[\mu,\sigma]$ values estimated in the previous step. Regarding the degrees of freedom (v), for preliminary estimations can be used the averaged values from the 24 levels of theory previously evaluated (see Table S14). Those values are: $v_{13Cu,sp2} = 8$; $v_{13Cu,sp3} = 10$; $v_{1Hu,sp2} = 8$; $v_{13Hu,sp3} = 4$; $v_{13Cs} = 7$; and $v_{1Hs} = 4$.

Step 6: complete the Region C with the experimental and calculated NMR data. The DP4+ values are shown in Region C.

