Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 171 (2004) 1–3
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Addendum

Addendum to "Using neural networks for ^{13}C NMR chemical shift prediction—comparison with traditional methods" [J. Magn. Reson. 157 (2002) 242–252]

Jens Meiler[a,*], Brent Lefebvre[b], Antony Williams[b], Michel Hachey[b]

[a] University of Washington, Box 357350, Seattle, WA 98195-7350, USA
[b] Advanced Chemistry Development, Inc., 90 Adelaide St. W., Suite 600, Toronto, Ont. Canada M5H 3V9

Available online 21 August 2004

In the above-referenced paper, nine methods for predicting ^{13}C NMR chemical shifts of small organic molecules are compared using the example of the drug agent Taxol. These methods include artificial neural networks (two versions of C_Shift [1]), incremental methods (ChemDraw [2], SpecTool [3]), quantum chemical calculation (Gaussian [4], Cosmos [5]), and HOSE code fragment based prediction tools (SpecInfo [6], ACD/CNMR [7], PredictIt NMR [8]).

To provide a fair comparison of the quality of each prediction method for an unknown and new class of organic substances, the authors required that all Taxol derivatives be removed from program internal databases as used by SpecInfo, ACD/CNMR, and PredictIt NMR and indirectly from C_Shift, ChemDraw, and SpecTool during training in these methods. As discussed in the original paper, it was at this time impossible for the authors to meet this requirement for the ACD/CNMR prediction tool.

In this addendum we comment on the chemical shift values for ACD/CNMR 6.12 published in the original paper, publish the ^{13}C NMR spectrum of Taxol as predicted with ACD/CNMR 6.12 *after* removal of Taxol derivatives from the internal database, and compare the results with those of other methods.

After the publication of the original paper it was shown that the large deviations reported for ACD/CNMR 6.12 around C-27 might be caused by an ambiguous drawing of the double bond C-23=C-27 in the structure editor. A representation of the structure similar to the one shown in Fig. 1 leads to a misinterpretation of the *cis*/*trans* isomery and reproduces the chemical shift data published previously. The grade of ambiguity in the original drawing remains unknown because the structure as drawn for the experiment was not preserved. When a structure is prepared in this structure drawing package and a possible ambiguity is present in the *E*/*Z* isomerism, it is important to resolve this by optimizing the structure in three dimensions. This was not apparent to the authors but has now been corrected.

The chemical shifts obtained with an unambiguous drawing using ACD/CNMR6.12 (the same version as in the original publication) yield a standard deviation of 0.8 ppm and an average deviation of 0.5 ppm. However, these very low error values are expected if Taxol and its derivatives are not removed from the internal database and do compare nicely with similar small deviations obtained with SpecInfo and PredictIt NMR under the same circumstances. Chemical shift predictions in these cases utilize a combination of database molecular fragment lookups and application of appropriate corrections based on the distribution of similarity of related fragments. The low deviations of the resulting chemical shifts likely reflect experimental variations in the chemical shift measurement, solvent effects, etc.

In the course of these experiments, B.L. performed the chemical shift prediction for Taxol after removing relevant derivatives from the database. The resulting chemical shift values are reported in Table 1. The deviations of the experimental chemical shift values are represented graphically by the area of the circle next to the numerical value. This allows fast comparison of the

DOI of original article: 10.1006/jmre.2002.2599.
* Corresponding author. Fax: +1 206 6851792.
E-mail address: jens@jens-meiler.de (J. Meiler).

1090-7807/$ - see front matter
doi:10.1016/j.jmr.2004.07.012

Table 1
Comparison of experimental and predicted ^{13}C chemical shifts for Taxol

ID	exp.	Neural networks		Increments		HOSE code predictings			Quantum chemical	
		C_Shift (new)	*C_Shift (old)*	*CS Chem Draw Pro*	*SpecTool*	*SpecInfo*	*PredictIt NMR 1.3*	*CNMR 6.12*	*Cosmos 4.5*	*Gaussian 98*
1	167.0	166.1	167.2	167.9	167.6	166.8	167.6	166.6	157.7	190.6
3	133.6	*136.0*	*133.9*	*133.5*	*133.5*	*133.2*	*134.2*	*136.3*	*121.7*	*135.7*
4	127.0	127.8	127.4	127.3	127.3	127.6	127.7	127.2	126.8	133.6
5	129.0	128.6	129.0	128.6	128.6	128.1	128.4	128.1	130.8	124.5
6	131.9	*131.9*	*132.3*	*131.9*	*131.9*	*131.6*	*132.0*	*131.7*	*129.6*	*132.9*
7	129.0	128.6	129.0	128.6	128.6	128.1	128.4	128.1	131.3	125.6
8	127.0	127.8	127.4	127.3	127.3	127.6	127.7	127.2	128.1	133.6
10	55.0	55.0	59.9	52.4	51.6	54.6	54.9	55.8	54.6	59.1
11	138.0	*136.7*	*138.1*	*142.4*	*138.8*	*136.7*	*140.7*	*136.1*	*135.8*	*138.7*
12	127.0	126.8	128.0	127.1	128.3	127.2	126.9	128.3	128.0	130.1
13	128.7	128.3	130.0	128.3	128.6	128.6	128.4	129.2	130.1	127.8
14	128.3	*127.9*	*127.7*	*126.5*	*125.8*	*127.7*	*127.4*	*126.2*	*130.6*	*128.2*
15	128.7	128.3	130.0	128.3	128.6	128.6	128.4	129.2	129.4	127.6
16	127.0	126.8	128.0	127.1	128.3	127.2	126.9	128.3	129.9	130.7
17	73.2	73.3	70.8	85.2	85.2	72.9	74.0	74.0	75.4	76.0
19	172.7	171.4	171.7	172.0	172.0	172.8	172.1	172.3	163.8	178.6
22	72.3	71.5	70.8	71.1	73.1	72.0	75.8	64.3	70.5	69.2
23	142.0	140.6	130.4	132.9	132.2	139.4	134.0	132.8	142.3	137.0
24	35.7	33.9	33.2	34.9	34.9	35.8	41.1	36.3	35.8	38.6
25	79.0	79.3	82.6	81.3	81.3	78.7	79.7	77.4	77.4	73.1
26	43.2	43.7	44.3	33.4	35.4	42.8	41.5	44.8	47.2	41.0
27	133.2	133.2	131.0	138.5	141.6	134.6	137.1	142.7	118.7	133.8
28	14.8	16.9	18.4	10.9	12.9	16.3	16.6	19.8	-	18.5
29	26.9	24.1	23.1	16.3	16.3	24.0	21.4	29.3	22.4	26.7
30	21.8	24.1	23.1	16.3	16.3	24.0	21.4	26.8	21.9	24.5
31	75.5	77.9	73.2	79.6	81.6	76.0	74.0	66.9	76.9	80.0
32	203.6	204.7	207.3	212.8	213.9	204.4	210.2	205.4	-	206.4
33	58.6	59.6	56.9	39.6	41.6	57.9	55.3	58.7	54.5	54.3
34	45.6	44.3	43.4	27.7	29.7	44.3	52.4	45.6	48.4	44.7
35	74.9	74.1	72.7	70.0	70.0	74.2	80.7	75.6	73.4	71.9
36	72.2	74.9	66.8	64.9	66.6	73.3	75.4	73.2	69.8	65.5
37	35.6	33.1	31.9	28.9	30.9	36.6	36.5	36.0	37.4	30.6
38	84.4	82.0	74.9	86.7	86.7	81.9	74.0	84.2	75.4	82.9
39	81.1	82.3	80.5	79.2	79.6	81.1	77.0	81.2	76.5	79.8
41	171.2	170.3	169.6	171.0	171.0	170.6	170.1	168.9	174.9	176.7
42	20.8	21.5	20.3	27.4	17.4	21.7	21.0	21.0	22.4	26.2
47	167.0	166.5	167.7	167.0	167.0	166.5	165.8	166.8	159.4	173.0
48	129.1	129.2	128.4	130.5	130.5	129.4	130.7	129.3	128.7	128.1
49	130.2	129.8	129.0	129.7	129.7	129.5	129.7	130.3	124.9	130.4
50	128.7	128.3	129.1	128.4	128.4	128.9	129.1	128.7	130.2	126.1
51	130.2	129.8	129.0	129.7	129.7	129.5	129.7	130.3	124.5	135.3
52	128.7	128.3	129.1	128.4	128.4	128.9	129.1	128.7	130.2	127.8
53	133.7	132.8	133.4	132.8	132.8	132.8	132.8	133.9	130.4	136.9
57	76.5	75.7	72.6	79.6	79.2	76.0	65.0	76.7	69.2	73.1
58	170.4	171.1	170.0	171.0	171.0	170.5	170.2	170.5	171.7	175.8
59	22.6	21.7	21.2	17.9	17.9	22.0	21.6	22.6	21.4	26.9
61	9.5	12.9	16.6	10.3	10.3	11.6	17.0	10.7	-	21.2
Std.dev.:		1.3	3.1	5.6	5.3	1.0	3.7	1.3	4.3	5.1
Aver.dev.:		1.0	2.1	3.5	3.3	0.8	2.3	0.9	3.3	3.9
Max.dev.:		3.4	11.6	19.0	17.0	3.0	11.5	4.4	14.5	23.6

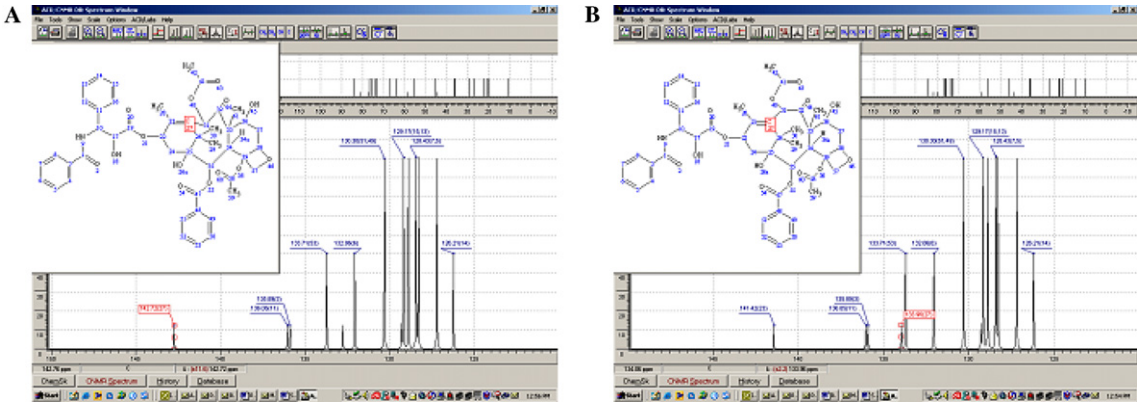

Fig. 1. ^{13}C NMR spectrum as predicted with ACD/CNMR 6.12 for Taxol with the ambiguously defined configuration (A) and the correct configuration (B).

different procedures (columns) and of the individual structural features (rows). Standard and mean deviations of the ACD/CNMR method (1.3 ppm/0.9 ppm) fall in the same range as those of the SpecInfo HOSE code prediction and also with the neural network approach C_Shift.

References

[1] J. Meiler, M. Will, R. Meusinger, J. Chem. Inf. Comput. Sci. 40 (2000) 1169.
[2] Cambridge Soft Corporation, CS ChemDraw Pro 4.5, 1985–1997.
[3] Upstream Solutions GMBH, NMR Prediction Products (Spec Tool), 1998–2001.
[4] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J.J. Dannenberg, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian, Pittsburgh, PA, 2001.
[5] U. Sternberg, F. Koch, P. Losso, Cosmos 4.52, 2001.
[6] BASF AG, SpecInfo Database of 13 NMR spectra (BASF in-house version), 2002.
[7] Advanced Chemistry Development, ACD/CNMR 6.12, 1996–2001.
[8] Bio-Rad, Preditcit NMR 1.3, 2002.