Electronic Supplementary Material (ESI) for New Journal of Chemistry.

Electronic Supporting Information

Substituent effect on the basicity (pK_a) of aryl guanidines and 2-(arylimino)imidazolidines: correlations of pH-metric and UV-metric values with predictions from gas-phase ab initio bond lengths

Christophe Dardonville[a*], *Beth A. Caine*[b,c], *Marta Navarro de la Fuente*[a], *Guillermo Martín Herranz*[a], *Beatriz Corrales Mariblanca*[a], *and Paul L. A. Popelier*[b,c]

Table of contents:

Table S1. Bond lengths for 2-(phenylimino)imidazolidines optimised at the B3LYP/6-311G(d,p) level of theory in the gas-phase given to 5 decimal places, predicted pK_a values using appropriate linear equation given in Table S3, experimental pH-metric values and absolute errors, with those exceeding 0.5 units highlighted in bold. MAE and standard deviation values also given, in pK_a units. ChemAxon values are obtained using the more stable imino tautomer as the input structure.

Compound	r(C-N)	Exp (pH)	AIBLHiCoS	AE	ChemAxon	AE
1a	1.38836	10.24	10.57	0.33	11.21	**0.97**
2a	1.38751	10.49	10.34	0.15	10.10	0.39
3a	1.38681	10.29	10.16	0.13	9.80	0.49
4a	1.38745	10.42	10.33	0.09	11.04	**0.62**
5a	1.38805	10.50	10.49	0.01	11.05	**0.55**
6a	1.38824	10.44	10.54	0.10	10.91	0.47
7a	1.38906	10.78	10.76	0.02	10.96	0.18
8a	1.38872	10.62	10.67	0.05	10.96	0.34
9a	1.38683	10.17	10.16	0.01	8.97	**1.2**
10a	1.38211	9.11	8.90	0.21	10.32	**1.21**
11a	1.38079	8.52	8.54	0.02	9.75	**1.23**
12a	1.37608	7.29	7.28	0.01	8.93	**1.64**
13a	1.37607	8.07	7.28	**0.79**	9.25	**1.18**
14a	1.37727	7.41	7.60	0.19	8.64	**1.23**
15a	1.37572	7.80	7.18	**0.62**	9.21	**1.41**
16a	1.38435	8.99	9.50	**0.51**	10.45	**1.46**
17a	1.38605	9.37	9.95	**0.58**	10.26	**0.89**
18a	1.38857	10.33	10.63	0.30	10.63	0.30
19a	1.38312	9.45	9.17	0.28	9.83	0.38
20a	1.38421	9.47	9.46	0.01	9.82	0.35
21a	1.38469	9.68	9.59	0.09	9.63	0.05
22a	1.38067	8.97	8.51	0.46	9.30	0.33
23a	1.38187	9.08	8.83	0.25	10.03	**0.95**
			MAE	0.23		0.77
			sd	0.22		0.47

Figure S1. Three tautomers considered during HiCoS formation using compounds in Table S2. Candidate active bonds considered during model calibration are labelled a-g. Within tautomer T3, the C-N bond lengths labelled "a" are used to form a predictive model. This is referred to as "C-N(im)" in the text and the bonding distance is marked as such below. The bond distances labelled "d" are also used to construct a separate model, which is referred to as "N-C(Ph)" in the text and is also labelled below.

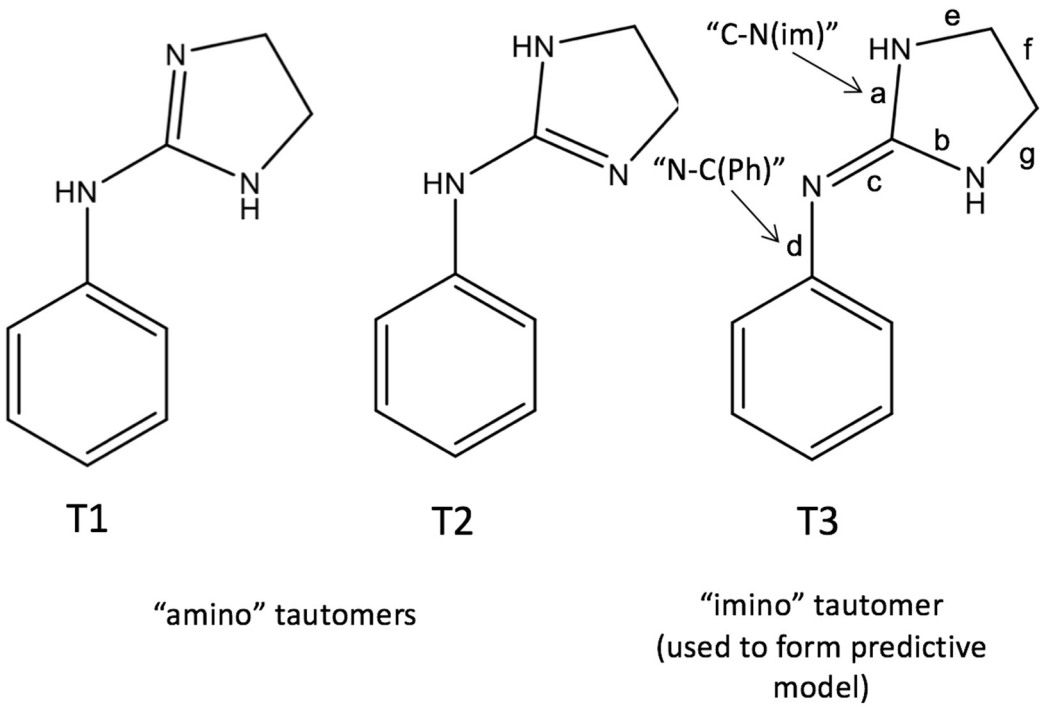


Table S2. Bond lengths for 12 phenylguanidines **1b-27b** optimized at the B3LYP/6-311G(d,p) level of theory in the gas-phase given to five decimal places, with experimental pH-metric values, predicted pK_a values using AIBLHiCoS using appropriate linear equation given in Table S4 and absolute errors of prediction for each method of prediction. Values exceeding 0.5 units highlighted in bold. MAE and standard deviation values also given, in pK_a units. ChemAxon values are obtained using the more stable imino tautomer as the input structure. Additional compounds alluded to in Section 2.3.2, 1-(2-pyridinyl)guanidine and 1-(6-methyl-2-pyridinyl)guanidine have pK_a values 10.17[a] and 10.26[a], and calculated C=N bond lengths 1.27606 and 1.27552.

Compound	r(C=N)	Exp pH	AIBLHiCoSA	AE pH	ChemAxonC	AE pH
1b	1.27459	10.90	11.01	0.11	11.45	**0.55**
2b	1.27506	11.36	11.22	0.14	10.40	**0.96**
3b	1.27478	10.98	11.10	0.12	10.10	**0.88**
6b	1.27490	10.98	11.15	0.17	11.18	0.20
7b	1.27511	11.35	11.24	0.11	11.22	0.13
17b	1.27366	10.05	10.60	**0.55**	10.54	0.49
18b	1.27405	11.01	10.77	0.24	10.91	0.10
22b	1.27571	10.47	11.52	**1.05**	9.58	**0.89**
24b	1.27524	12.01	11.31	**0.70**	11.76	0.25
25b	1.27632	11.68	11.79	0.11	10.02	**1.66**
26b	1.27432	11.03	10.89	0.14	10.84	0.19
27b	1.27462	11.01	11.03	0.02	10.85	0.16
			MAE	0.29		0.54
			sd	0.31		0.48

[a]P. J. Taylor, A. R. Wait, J. Chem. Soc. Perkin Transactions 2: Physical Organic Chemistry (1972-1999), **1986**, 1765-1770

Figure S2. Conformers of the three tautomers considered for the formation of the predictive model using compounds in Table S1. Bonds considered are labelled i-v. The C=N (ii) bond lengths of the form labelled A are used to construct the predictive model.

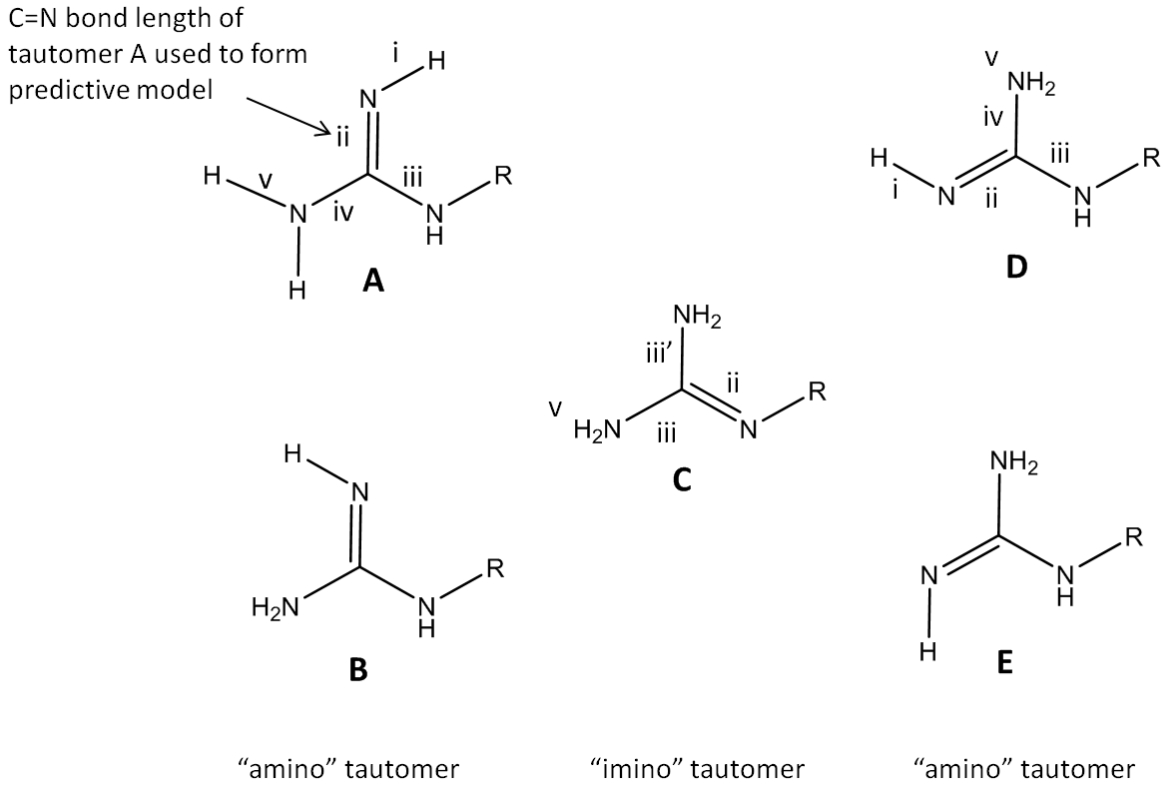


Table S3. 2-(phenylimino)imidazolidine derivatives and corresponding pK_a values and calculated gas-phase bond lengths used to obtain the linear equation for predictions (pK_a = 268.08*r(C-N) - 361.6). Squared correlation coefficient, leave-one-seventh-out q^2 and RMSEE values for model are also given.

ID	Compound	r(C-N(im))	pK_a [a]
1i	2,4,6-Br_3	1.37802	7.46
2i	2,4,6-Cl_3	1.37808	7.75
3i	2,4,6-Me_3	1.38904	10.78
4i	2,6-Br_2	1.37950	7.80
5i	2,6-Cl_2,4-Br	1.37806	7.72
6i	2,6-Cl_2,4-Me	1.38033	8.29
7i	2,6-Cl_2,4-NO_2	1.37407	6.86
8i	2,6-Cl_2,4-OMe	1.38108	8.57
9i	2,6-Cl_2	1.37968	8.05
10i	2,6-Et_2	1.38888	10.61
11i	2,6-Me_2,4-Br	1.38641	10.21
12i	2,6-Me_2,4-Cl	1.38645	10.25
13i	2,6-Me_2	1.38862	10.53
14i	2,3-Cl_2	1.38040	8.55
15i	2,4-Cl_2	1.38041	8.73
16i	2,4-Me_2	1.38824	10.56
17i	2,5-Cl_2	1.37981	8.50
18i	2-Cl,4-Me	1.38289	9.41
19i	2-Cl	1.38222	9.15
20i	2-Me,4-Cl	1.38502	9.99
21i	2-Me	1.38755	10.23
22i	H	1.38807	10.05
	r^2	0.97	
	q^2	0.96	
	RMSEE	0.23	

[a]Timmermans, P. B.; van Zwieten, P. A. Dissociation constants of clonidine and structurally related imidazolidines. *Arzneimittelforschung* **1978**, *28*, 1676-1681.

Table S4. Guanidine-type compounds and corresponding pK_a values and calculated gas-phase bond lengths used to obtain the linear equation for predictions (pK_a = 448.9*r(C=N) - 561.15) Squared correlation coefficient, leave-one-seventh-out q^2 and RMSEE values for model are also given.

ID	Compound	C=N	pK_a
g1	Guanidine	1.27828	13.60[a]
g2	Phenylbiguanide	1.27312	10.71[b]
g3	Ethylenebisbiguanide	1.27514	11.76[b]
g4	Famitodine	1.26497	6.78[c]
g5	Arginine	1.27846	12.50[d]
g6	N,N-Dimethylguanidine	1.28092	13.40[e]
g7	N,N',N''-Trimethylguanidine	1.28165	13.90[f]
g8	Hydroxyguanidine	1.26704	7.96[g]
g9	N''-Benzoylcarbonohydrazonic diamide	1.26910	7.94[g]
g10	N''-Phenylcarbonohydrazonicdiamide	1.26886	8.26[g]
g11	N'-(Diaminomethylene)benzamide	1.26524	6.98[g]
g12	Ethoxycarbonylguanidine	1.26569	7.03[g]
g13	Methoxyguanidine	1.26779	7.46[h]
	r^2	0.98	
	q^2	0.97	
	RMSEE	0.46	

[a]J. W. Shaw, D.H. Grayson, I. Rozas, *Topics in Heterocyclic Chemistry, Springer Berlin Heidelberg, Berlin, Heidelberg,* **2015**, pp. 1-51, [b]J. Bjerrum, G. Schwarzenbach, L. G. Sillen, *Stability Constants*, The Chemical Society, London, **1958**, p. 8., [c]T. Degim, V. Zaimoglu, C. Akay, Z. Degim, *Il Farmaco*, **2001**, *56*, 659-663, [d]L. M. Carrick, A. Aggeli, N. Boden, J. Fisher, E. Ingham, T. A. Waigh, *Tetrahedron*, **2007**, *63*, 7457-7467, [e]T. Davis, R. C. Elderfield, *J. Am. Chem. Soc.* **1932**, *54*, 1499-1503, [f]IUPAC Dissociation Constants of Organic Bases in Aqueous Solution, D.D. Perrin, **1972**, Butterworth & Co Ltd, pp446, [g]P. Taylor, A. R. Wait, *J. Chem. Soc. Perkin. Trans. 2*, **1986**, *11,* 1765-1770, [h]M. Charton, *J. Org. Chem.*, **1965**, *30*, 969.

Figure S3. Plots of V_{xc}^{CN} values vs bond lengths for bonds a-d, as labelled in Fig. S1. Training set compounds are shown in blue, the two m-halo,p-nitro compounds are shown in purple, and the three 3-pyridyl in red. R^2 values correspond to the training set only.

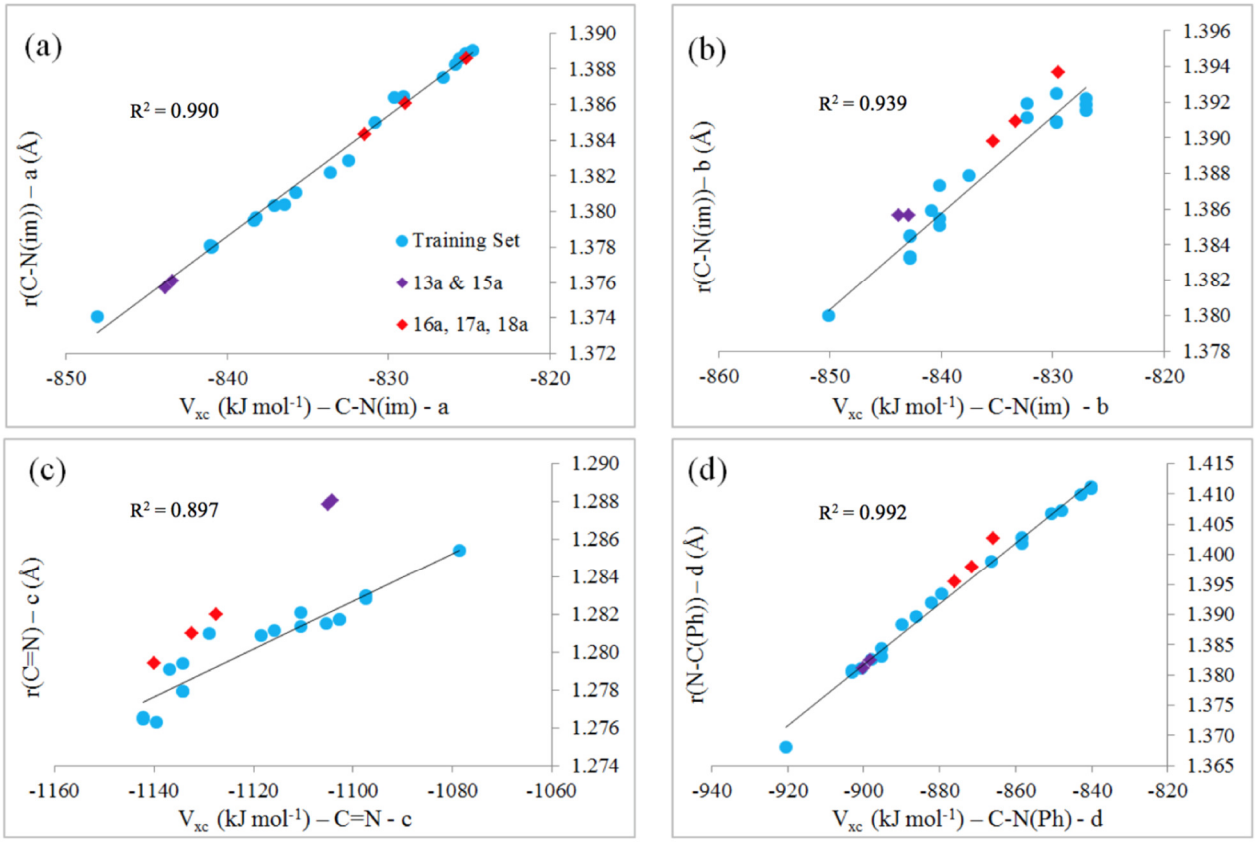


Figure S4. V_{xc}^{CN} vs pK_a plots for C-N(im) (a) and N-C(Ph) (b), and bond length vs pK_a plots for C-N(im) (c) and N-C(Ph) (d).

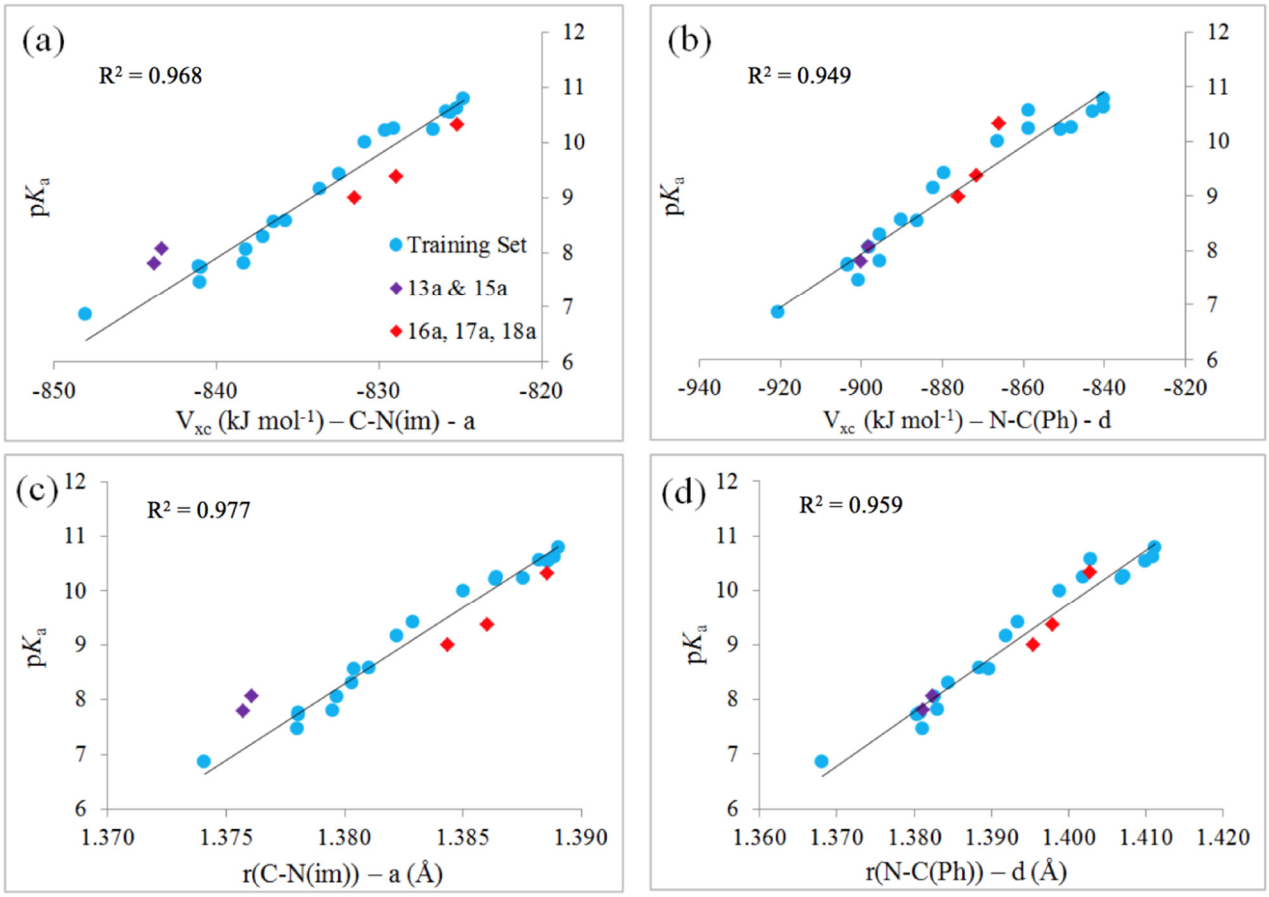
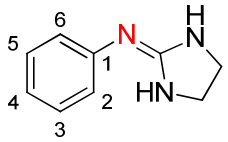


Table S5. Data used to construct plots in Fig S3 and S4: V_{xc}^{CN} values, given in kJ mol^{-1}, for bonded atom interactions (a-d in Fig. S1), corresponding calculated bond lengths, given in Angstroms. pK_a values for imidazolidine training set compounds taken from Table S3 are denoted by "i", test set compounds are denoted by "a".

	(C-N(im)) a		(C-N(im)) b		(C-N(im)) c		(N-C(Ph))		
ID	V_{xc}^{CN}	r(CN)	V_{xc}^{CN}	r(CN)	V_{xc}^{CN}	r(CN)	V_{xc}^{CN}	r(NC)	pK_a
1i	-836.44	1.380403	-840.91	1.385953	-1110.46	1.282117	-886.17	1.389742	8.55
2i	-840.98	1.378017	-842.79	1.383306	-1097.46	1.282850	-900.55	1.381169	7.46
3i	-841.04	1.378080	-842.79	1.383332	-1097.46	1.283019	-903.17	1.380802	7.75
4i	-824.78	1.389038	-827.03	1.392221	-1142.09	1.276443	-840.16	1.411237	10.78
6i	-825.83	1.388236	-829.66	1.392496	-1136.84	1.279115	-858.54	1.402844	10.56
8i	-838.32	1.379501	-842.79	1.384466	-1102.71	1.281715	-895.30	1.383064	7.80
9i	-840.95	1.378055	-842.79	1.383270	-1097.46	1.283023	-903.17	1.380461	7.72
10i	-837.12	1.380333	-840.16	1.385095	-1105.34	1.281514	-895.30	1.384474	8.29
11i	-848.04	1.374073	-850.15	1.380012	-1078.63	1.285438	-920.45	1.368077	6.86
12i	-835.75	1.381075	-840.16	1.385501	-1110.59	1.281379	-890.04	1.388429	8.57
13i	-838.19	1.379678	-842.79	1.384561	-1102.71	1.281742	-897.92	1.382647	8.05
14i	-825.20	1.388883	-827.03	1.391536	-1139.47	1.276315	-840.16	1.410930	10.61
15i	-829.62	1.386407	-829.66	1.390890	-1134.22	1.277923	-850.66	1.406864	10.21
16i	-829.09	1.386446	-829.66	1.390937	-1134.22	1.277928	-848.04	1.407207	10.25
17i	-825.57	1.388616	-827.03	1.391851	-1142.09	1.276568	-842.79	1.409939	10.53
18i	-832.45	1.382892	-837.53	1.387923	-1118.46	1.280884	-879.54	1.393399	9.41
19i	-833.64	1.382220	-840.16	1.387365	-1115.84	1.281133	-882.17	1.391926	9.15
20i	-830.85	1.385017	-832.28	1.391174	-1128.97	1.280977	-866.42	1.398802	9.99
21i	-826.62	1.387545	-832.28	1.391911	-1134.22	1.279420	-858.54	1.401926	10.23
13a	-843.37	1.376070	-843.83	1.385670	-1104.97	1.287830	-898.36	1.382350	8.07
15a	-843.83	1.375720	-842.88	1.385670	-1104.23	1.288030	-900.15	1.381120	7.80
16a	-831.48	1.384347	-835.36	1.389785	-1127.44	1.282001	-876.12	1.395386	8.99
17a	-828.93	1.386050	-833.33	1.390930	-1132.37	1.280990	-871.61	1.397860	9.37
18a	-825.16	1.388570	-829.46	1.393656	-1140.03	1.279435	-866.01	1.402724	10.33

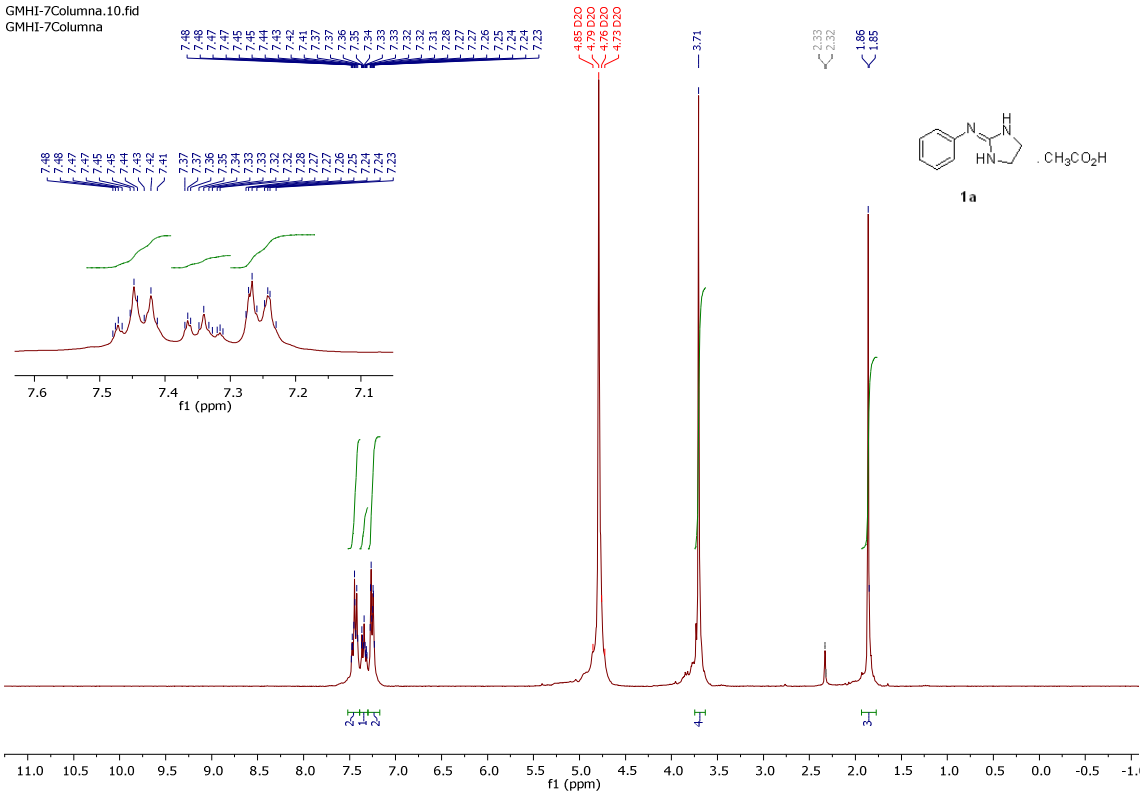
Table S6. Effect of substituents on the pK_a of 2-(phenylimino)imidazolidine.



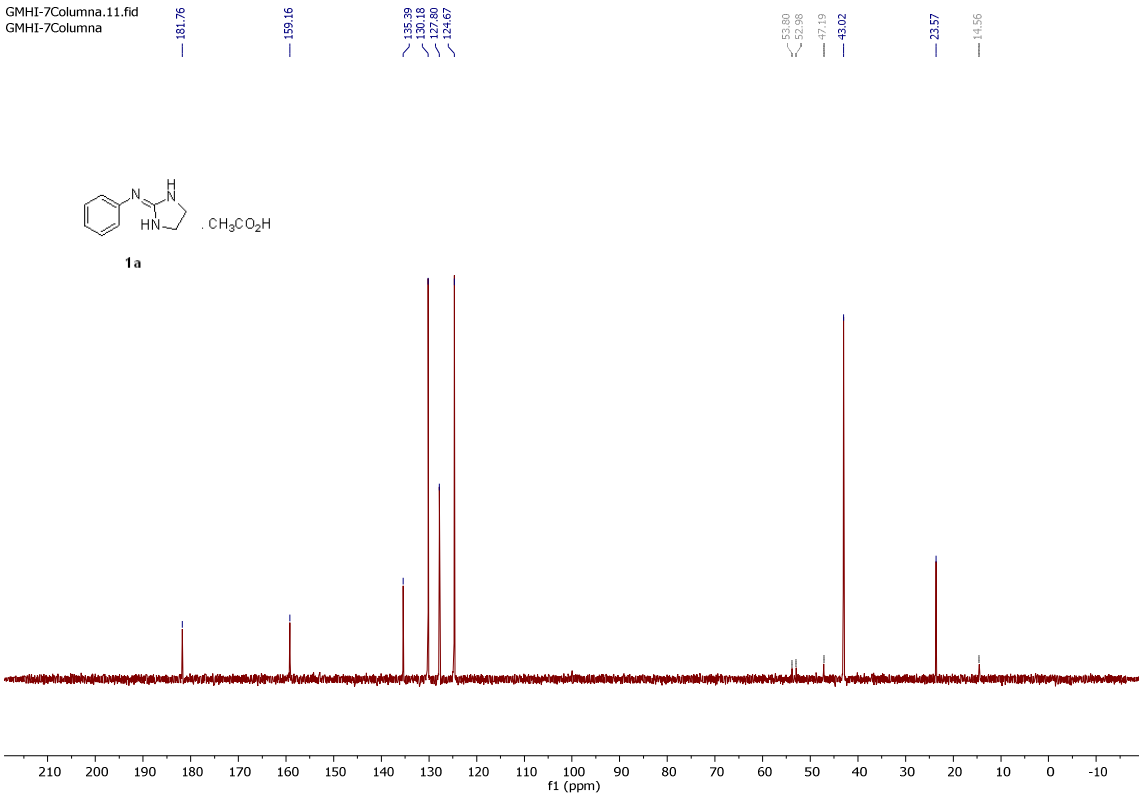
Substituent	pK_a (20 ºC)[a]	pK_a (25 ºC)[b]	ΔpK_a[c]
H	10.05	10.24	0
4-Bn		10.29	+0.05
2,6-di-Me, 6-Br	10.21		+0.16
4-Et		10.42	+0.18
2-Me	10.23		+0.18
2,6-di-Me, 4-Cl	10.25		+0.20
3,4-cyclohexyl (fused)		10.44	+0.20
4-PhNH		10.49	+0.25
3,4-di-Me		10.50	+0.25
4-OMe		10.62	+0.38
2,6-di-Me	10.53		+0.48
2,4-di-Me	10.56		+0.51
2,4,5-tri-Me		10.78	+0.53
2,6-di-Et	10.61		+0.56
2,4,6-tri-Me	10.78		+0.73
2-Me, 4-Cl	9.99		-0.06
4-PhNHCO		10.17	-0.07
2,4-di-Me, 6-Cl	9.59		-0.46
2-Cl, 4-Me	9.41		-0.64
2-Cl, 6-Me	9.40		-0.65
2-Cl	9.15		-0.90
2,4-di-Cl, 6-Me	9.03		-1.02
4-Ac		9.11	-1.14
2,4-di-Cl	8.73		-1.32
2,6-di-Cl, 4-OMe	8.57		-1.48
2,3-di-Cl	8.55		-1.50
2,5-di-Cl	8.50		-1.55
4-NO_2		8.52	-1.73
2,6-di-Cl, 4-Me	8.29		-1.76
2,6-di-F	8.18		-1.87
2,6-di-Cl	8.05		-2.00
2-Cl, 6-F	8.01		-2.04
2-Br, 6-Cl	7.93		-2.12
3-Cl, 4-NO_2		8.07	-2.17
2,6-di-Br	7.80		-2.25
2,4,6-tri-Cl	7.75		-2.30
3-F, 4-NO_2		7.80	-2.45
2-F, 4-NO_2		7.41	-2.84
2-Cl, 4-NO_2		7.29	-2.95

[a] Determined by potentiometric titrations at 20 ºC: Timmermans, P. B.; van Zwieten, P. A. Dissociation constants of clonidine and structurally related imidazolidines. *Arzneimittelforschung* **1978**, *28*, 1676-1681. [b] Determined by potentiometric titrations using Sirius T3 apparatus. [c] ΔpK_a is the difference in pK_a of the substituted compound with respect to the unsubstituted compound.

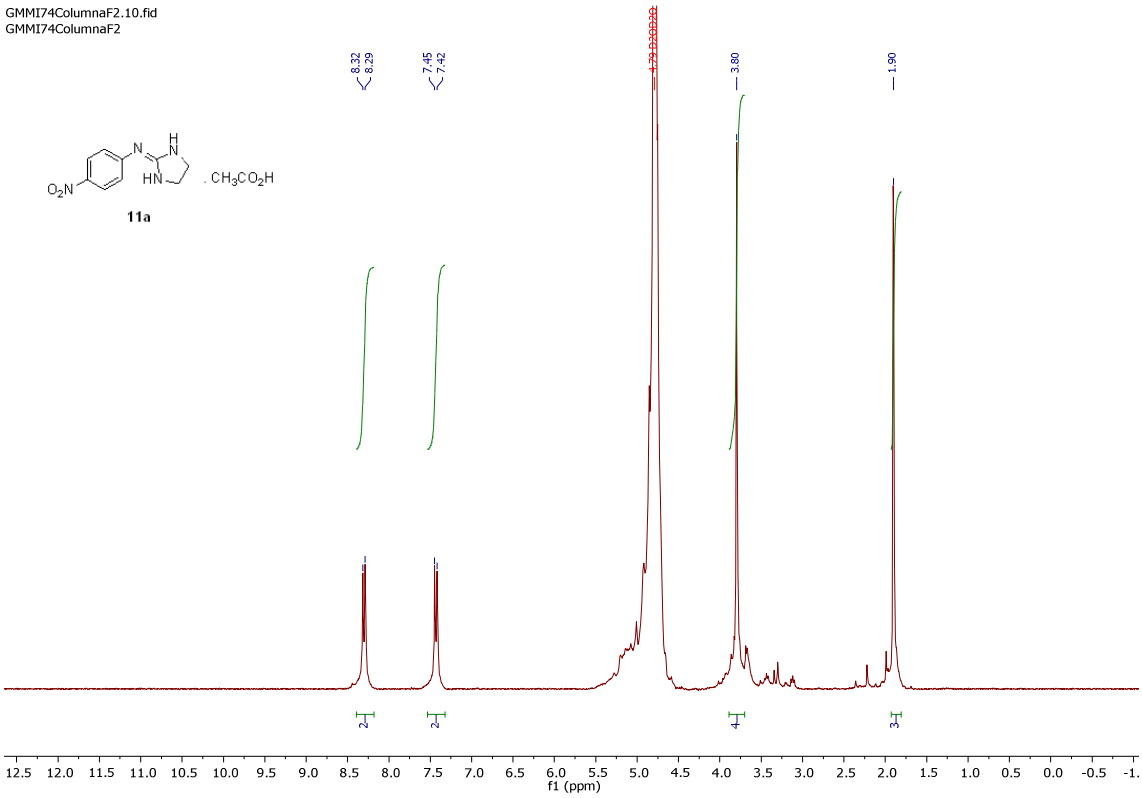
^{1}H NMR of compound **1a** (300 MHz, D_2O)



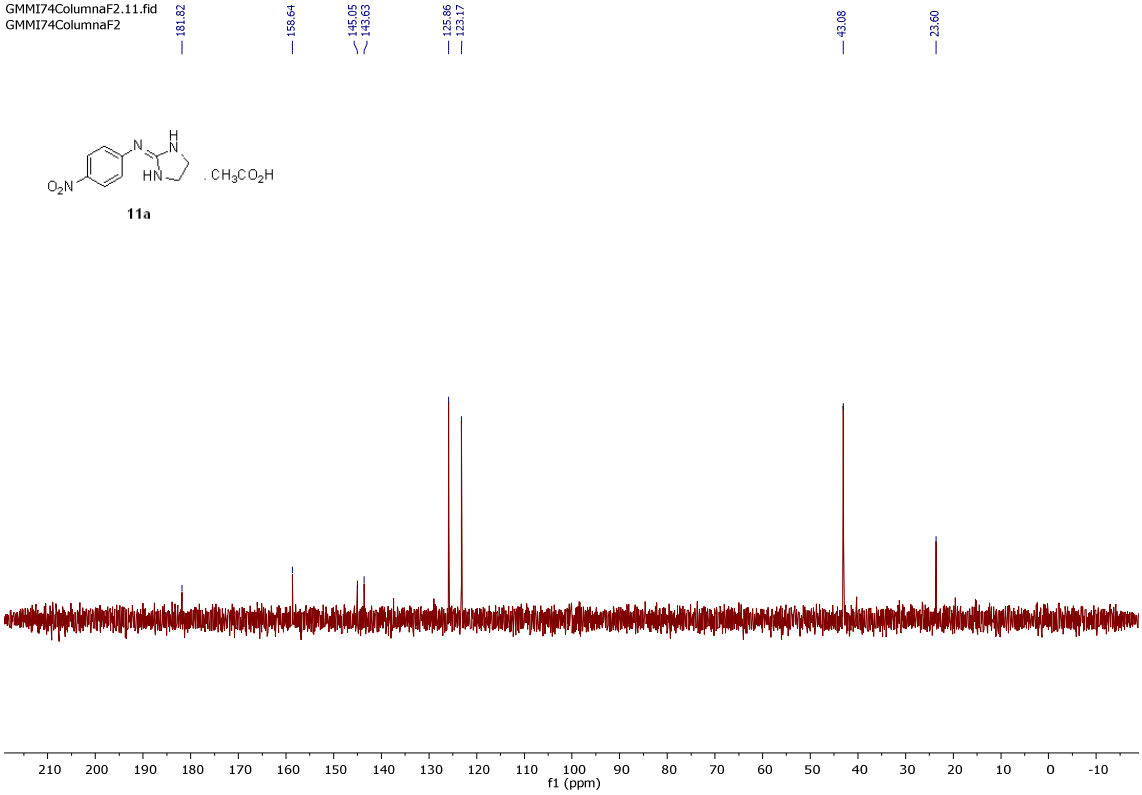
^{13}C NMR of compound **1a** (75 MHz, D_2O)



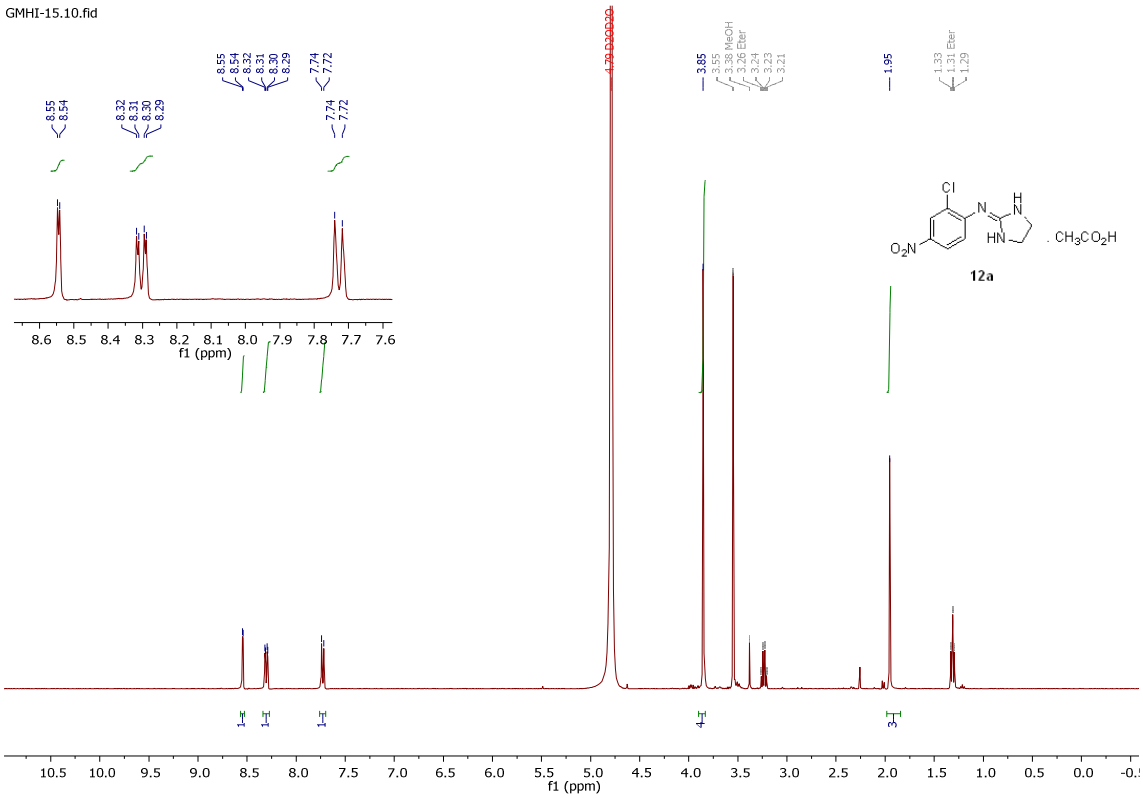
^{1}H NMR of compound **11a** (300 MHz, D_2O)



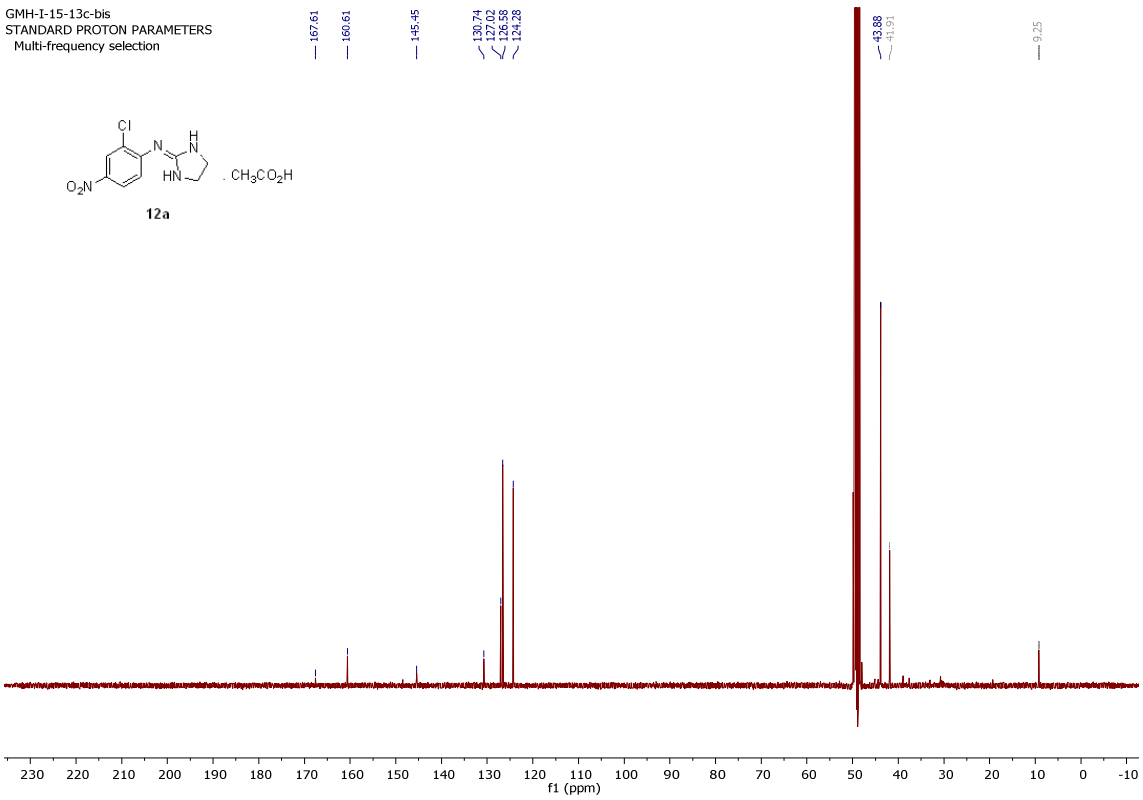
^{13}C NMR of compound **11a** (75 MHz, D_2O)



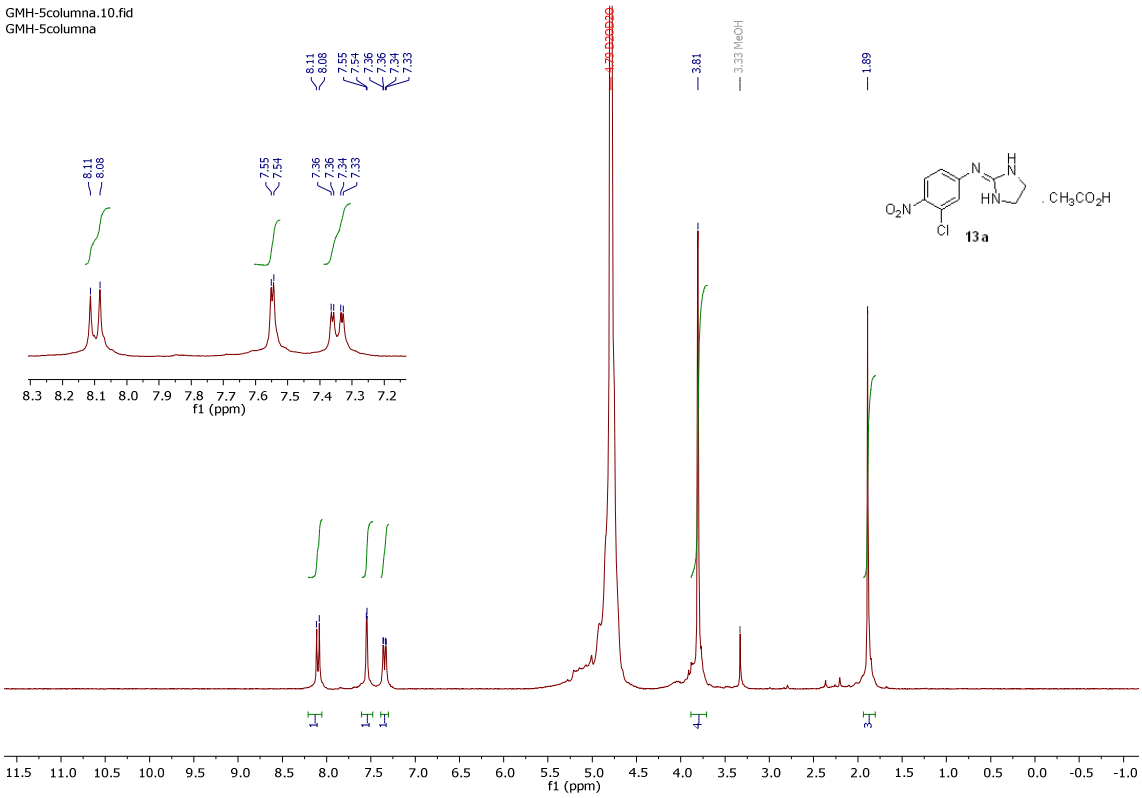
^{1}H NMR of compound **12a** (400 MHz, D_2O)



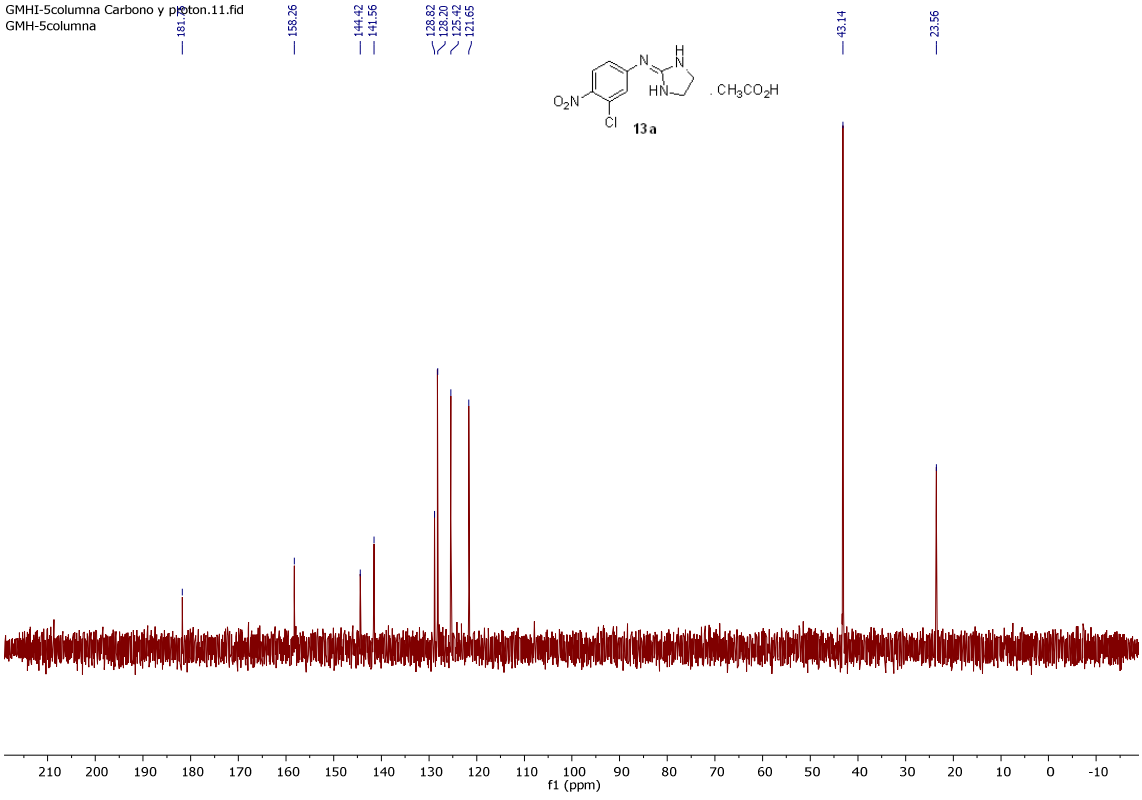
^{13}C NMR of compound **12a** (100 MHz, CD_3OD)



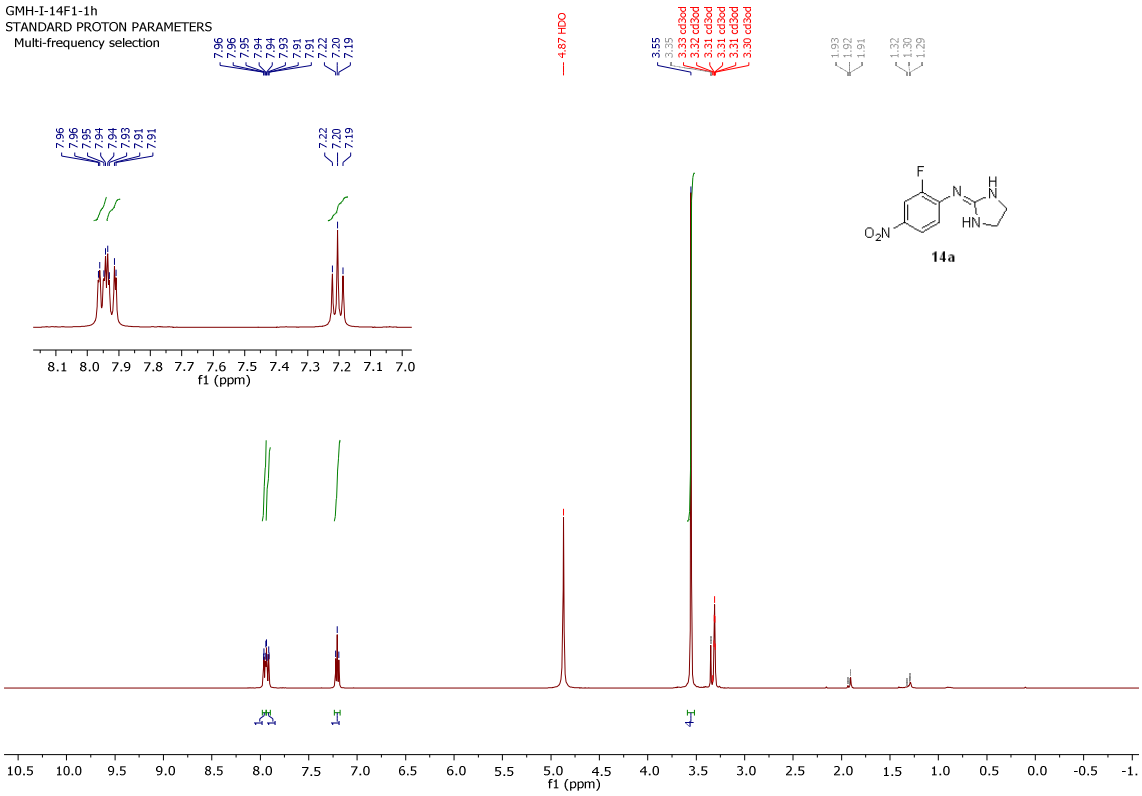
^{1}H NMR of compound **13a** (300 MHz, D_2O)



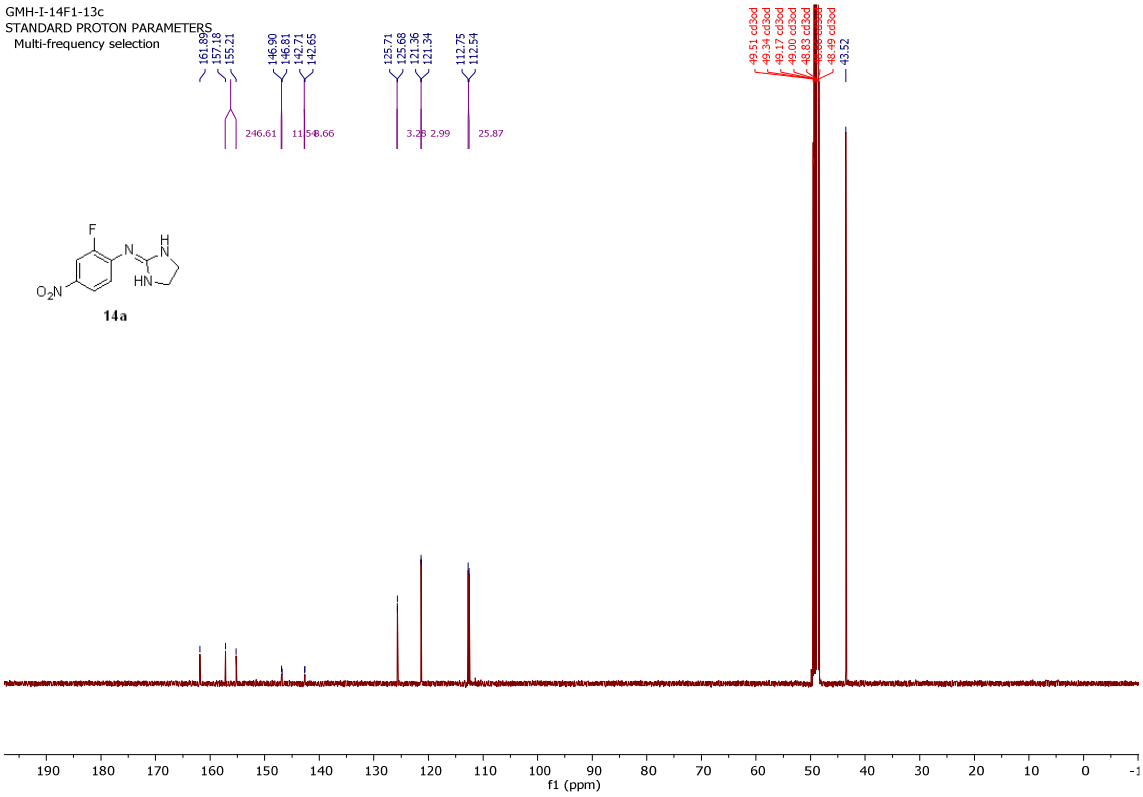
^{13}C NMR of compound **13a** (75 MHz, D_2O)



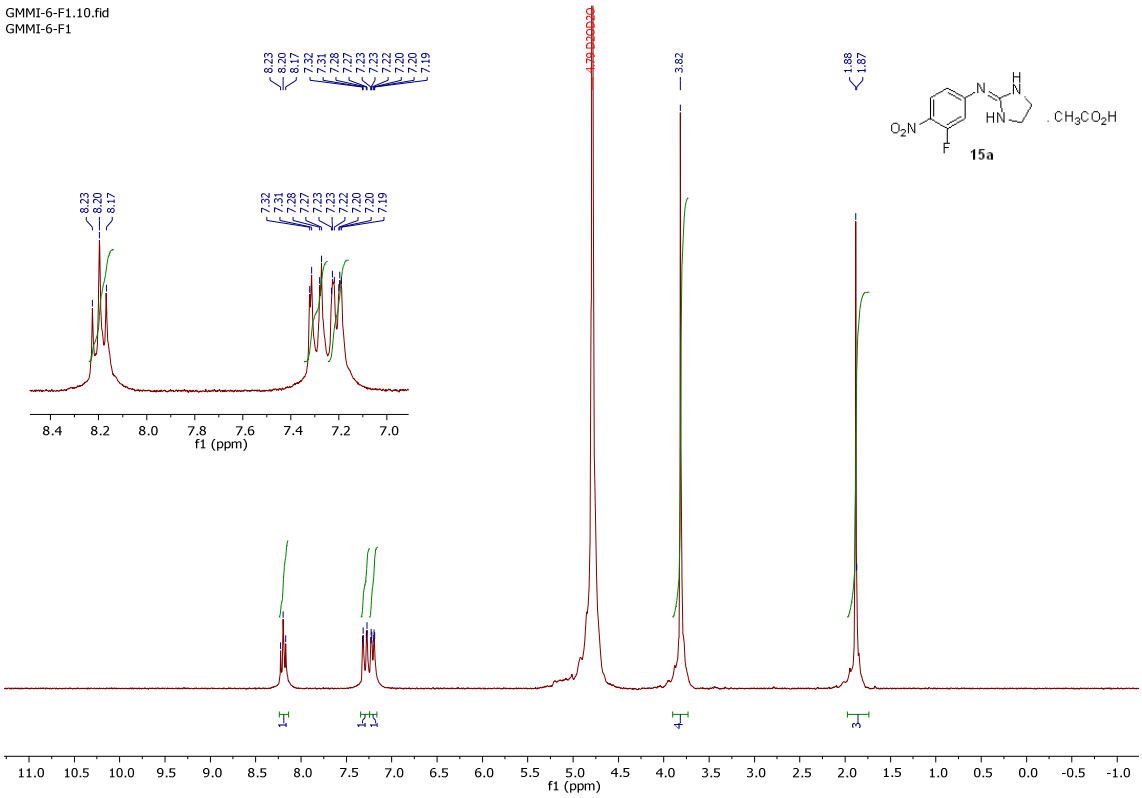
^{1}H NMR of compound **14a** (500 MHz, CD_3OD)



^{13}C NMR of compound **14a** (125 MHz, CD_3OD)



^{1}H NMR of compound **15a** (300 MHz, D_2O)



^{13}C NMR of compound **15a** (75 MHz, D_2O)