NJC

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/njc

NJC

A Theoretical Study of the Gas Phase (Proton Affinity) and Aqueous (pK_a) Basicity of a Series of 150 Pyrazoles†

Marta Marín-Luna,‡ Ibon Alkorta* and José Elguero

Instituto de Química Médica, CSIC, Juan de la Cierva, 3, E-28006 Madrid, Spain

* Author to whom correspondence should be addressed:
Email: ibon@iqm.csic.es
URL: http://are.iqm.csic.es/
FAX: +34 91 564 48 53

† Electronic supplementary information (ESI) available. For ESI see DOI: .....

‡ On leave from Departamento de Química Orgánica, Facultad de Química, Universidad de Murcia, Campus de Espinardo, 30100 Murcia, Spain

This paper deals with the basicity, both in the gas phase and in aqueous solution, of a set of 150 pyrazoles covering a range of about 200 kJ·mol^{-1} in proton affinity and 10-15 pK_a units. There are 63 NH-pyrazoles, in many cases, with two different tautomers, and 87 N-substituted pyrazoles. The gas phase results are well reproduced by DFT theoretical calculations when the less stable tautomers are removed. The pK_a values to be reproduced adequately by the calculations need not only to remove some tautomers but to use dummy variables accounting for the protonation on exocyclic amine groups and the conformation of 5-phenyl groups. Using these restrictions a large number of unknown pK_as can be predicted and a few errors corrected.

1. Introduction

Amongst the physicochemical properties of drugs (water solubility, partition coefficient, crystal structure, stereochemistry), related to ADMET (Absorbtion, Distribution, Metabolism, Excretion)[1] and to SAR (Structure Activity Relationships)[2] and QSAR (Quantitative Structure Activity Relationships),[3] one of the most important is their acid-base characteristics. The Lipinski's "rule of five" state that an orally active drug should have no more than five hydrogen bond (HB) donors (HBD) and no more than ten HB acceptors (HBA).[4] Since HB can be considered as the initial step of protonation,[5] these rules are related to the acidity and the basicity of drugs. Besides, the great majority of drugs contain ionizable groups, most being basic while only some are acidic.

The pyrazole skeleton is very common in a great variety of drugs as has been reported in several books and reviews.[6] Furthermore, some of us have been the main contributors to the determination of the pK_a,[7,8,9,10,11] and to the gas-phase basicity of pyrazoles, both experimental[12,13,14,15,16,17,18,19,20,21,22] and theoretically.[23,24,25,26] Therefore, we decided to carry out a theoretical study of the basicity of a large family of pyrazoles (150) both PA and pK_a (note that the collection is incomplete, only 68 PAs and 156 pK_as).

2. Computational methods

The geometries of pyrazoles (Pz) and pyrazolium cations (PzH^+) have been fully optimized using the functional B3LYP[27] and the 6-311++G(d,p) basis set[28] as implemented in the Gaussian 09 package (the coordinates of all the optimized geometries are gathered in the Supporting Information).[29] The minimum energy structure of all compounds was characterized by frequency analysis. The solvent effects have been evaluated by reoptimizing the structures at B3LYP/6-311++G(d,p) level and the Self-Consistency Reaction Field (SCRF) method,[30] based on the Polarized Continuum Model (PCM) of Tomasi and co-workers,[31] in water as solvent using the standard parameters provided by the Gaussian-09 program.

Scheme 1 shows the acid dissociation of a protonated pyrazole (PzH^+) into its conjugated base, neutral pyrazole (Pz), and a free proton (H^+). The pK_a value of an acid is defined as shown in the equation 1. Here, ΔG_{aq} was directly obtained as the difference of the free Gibbs energies of the products and reactants, as indicated in equation 2. Thus, the G_{aq} value of each species can be obtained by adding the solvation energy (ΔG_s) to the free Gibbs energy in gas phase (G_{gas}) (3).

$$PzH^+ \xrightarrow{\Delta G_{aq}} Pz + H^+$$

Scheme 1. The species involved.

$$pK_a = \frac{\Delta G_{aq}}{RT \ln 10} \quad (1)$$

$$\Delta G_{aq} = G_{aq}(Pz) + G_{aq}(H^+) - G_{aq}(PzH^+) \quad (2)$$

$$G_{aq} = G_{gas} + \Delta G_s \quad (3)$$

The proton free energy in aqueous solution is uncertain. Thus, we decided to use experimental values for this variable. The experimental estimations of the solvation energy of the free proton range from –1062.7 to –1112.5 kJ·mol^{-1}. Most programs for calculating solvent effects (SMD,[32] COSMO[33]) use the last value.[34]

With the aim of minimize possible errors we calculated two pK_a values for each Pz by using different ΔG_s data according to literature, *i.e* the highest calculated value ΔG_s = –1112.5 kJ·mol^{-1} and the mean value of the range, ΔG_s = –1085.7 kJ·mol^{-1}. [35]

$$pK_a(Pz) = \frac{G_{aq}(Pz) + \Delta G_s(H^+) - G_{aq}(PzH^+)}{RT \ln 10} \quad (4)$$

Table 1 contains all the data for the following discussion and Table S1 of the Electronic supplementary information (ESI) contains supplementary columns with intermediate data. For some substituents we have used simplified codes: they are reported in Fig. 1.

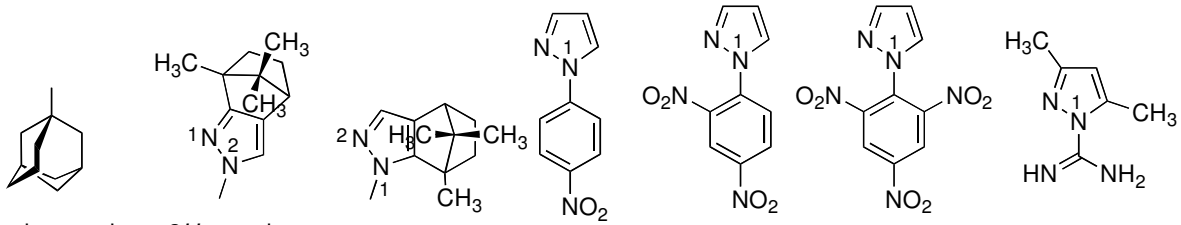


Fig 1 Structure of some substituents.

Table 1 Collection of data related to pyrazoles **1** to **150**. All pK_a values from reference [10] save compounds **55** and **56** (see footnote k). ΔH Gas = *H* Gas (Pz) – H Gas (PzH^+); ΔG Gas = *G* Gas (Pz) – *G* Gas (PzH^+); ΔH Aqueous = *H* Aqueous (Pz) – *H* Aqueous (PzH^+); ΔG Aqueous = *G* Aqueous (Pz) – *G* Aqueous (PzH^+). **c** compounds (**62**, **98**, **117**, **132**, **148**) correspond to protonation on the amino group.

No	Name	R^1	R^3	R^4	R^5	ΔH Gas (kJ·mol^{-1})	ΔG Gas (kJ·mol^{-1})	ΔH Aqueous (kJ·mol^{-1})	ΔG Aqueous (kJ·mol^{-1})	Calc.[i,j] pK_a	PA NIST corr. (kJ·mol^{-1})	Exp. pK_a	Fitted and predicted
1	Parent	H	H	H	H	891.03	890.67	1106.14	1105.37	3.44	894.1	2.48	1.83
2a	3-Me	H	CH_3	H	H	913.28	917.07	1116.87	1118.68	5.77	906.0	3.32	3.07
2b	5-Me	H	H	H	CH_3	914.75	917.84	1116.84	1119.46	5.91	906.0	3.32	3.14
3a	3-Et	H	C_2H_5	H	H	917.81	920.21	1117.16	1118.49	5.74		3.30	3.05
3b	5-Et	H	H	H	C_2H_5	919.54	922.14	1116.25	1114.78	5.09		3.30	2.71
4a[a]	3-*t*-Bu	H	*t*-C_4H_9	H	H	927.31	928.31	1114.83	1120.98	6.17	922.8	3.30	3.28
4b	5-*t*-Bu	H	H	H	*t*-C_4H_9	928.93	931.39	1115.48	1115.01	5.13	922.8	3.30	2.73
5	4-Me	H	H	CH_3	H	906.77	910.60	1112.78	1112.96	4.77	906.8	3.09	2.54
6[b]	4-Adamantyl[c]	H	H	1-Ad	H	927.21	926.65	1120.72	1116.94	5.47	913.1		2.91
7	3,5-diMe	H	CH_3	H	CH_3	934.66	942.65	1125.82	1126.47	7.14	933.5	4.12	3.80
8a	3,4-diMe	H	CH_3	CH_3	H	927.05	930.80	1123.23	1124.55	6.80	927.3	3.91	3.62
8b	4,5-diMe	H	H	CH_3	CH_3	930.39	931.49	1124.28	1125.95	7.04	927.3	3.91	3.74
9	3,4,5-triMe	H	CH_3	CH_3	CH_3	947.04	952.09	1132.45	1139.18	9.36	949.3	4.63	4.98
10a	3,4-diMe,5-Et	H	CH_3	CH_3	C_2H_5	949.81	954.54	1132.41	1134.56	8.55		4.60	4.55
10b	3-Et,4,5-diMe	H	C_2H_5	CH_3	CH_3	951.41	953.25	1133.77	1136.93	8.97		4.60	4.77
11	3,5-diEt,4-Me	H	C_2H_5	CH_3	C_2H_5	955.78	957.76	1133.91	1135.99	8.80	952.8	4.51	4.68
12a	3,4-triCH_2	H	CH_2	$(CH_2)_2$	H	927.61	929.95	1120.63	1122.05	6.36		3.61	3.38
12b	4,5-triCH_2	H	H	$(CH_2)_2$	CH_2	933.40	935.36	1125.15	1126.05	7.06		3.61	3.76
13a	3,4-tetraCH_2	H	$(CH_2)_2$	$(CH_2)_2$	H	935.43	937.63	1124.74	1126.61	7.16		4.01	3.81
13b	4,5-tetraCH_2	H	H	$(CH_2)_2$	$(CH_2)_2$	936.02	938.16	1124.70	1126.58	7.15		4.01	3.80
14a	3,4-pentaCH_2	H	$(CH_2)_2$	$(CH_2)_3$	H	937.89	940.27	1124.68	1125.86	7.03		3.96	3.74
14b	4,5-pentaCH_2	H	H	$(CH_2)_3$	$(CH_2)_2$	938.64	940.52	1124.22	1125.49	6.96		3.96	3.70
15a	2*H*-campho[d]	H	H	--	--	943.29	945.15	1121.93	1123.64	6.64		3.35	3.53
15b	1*H*-campho[d]	H	--	--	H	955.10	957.05	1132.16	1133.89	8.44		3.35	4.489
16a	3,4-triCH_2,5-Me	H	CH_2	$(CH_2)_2$	CH_3	947.95	950.77	1130.68	1131.40	8.00		4.38	4.26
16b	3-Me,4,5-triCH_2	H	CH_3	$(CH_2)_2$	CH_2	951.39	956.13	1134.42	1136.03	8.81		4.38	4.69
17a	3-Me,4,5-tetraCH_2	H	CH_3	$(CH_2)_2$	$(CH_2)_2$	953.61	957.93	1133.27	1134.81	8.60		4.65	4.57
17b	3,4-tetraCH_2,5-Me	H	$(CH_2)_2$	$(CH_2)_2$	CH_3	955.23	957.56	1134.16	1137.28	9.03		4.65	4.80
18a	3-Me,4,5-pentaCH_2	H	CH_3	$(CH_2)_3$	$(CH_2)_2$	955.77	959.44	1133.56	1136.50	8.89		4.57	4.73
18b	3,4-pentaCH_2,5-Me	H	$(CH_2)_2$	$(CH_2)_3$	CH_3	956.93	959.42	1134.01	1138.23	9.20		4.57	4.89
19	3,5-di*t*-Bu	H	*t*-C_4H_9	H	*t*-C_4H_9	957.62	961.57	1127.03	1127.05	7.24	952.7	3.99	3.85
20a	3-*t*-Bu,5-Me	H	*t*-C_4H_9	H	CH_3	946.87	949.53	1127.69	1130.13	7.78	946.2		4.14
20b	3-Me,5-*t*-Bu	H	CH_3	H	*t*-C_4H_9	946.99	948.59	1126.06	1128.89	7.56	946.2		4.02
21[k]	3,5-di*t*-Bu,4-Me	H	*t*-C_4H_9	CH_3	*t*-C_4H_9	967.90	970.29	1134.01	1135.85	8.78	967.5	4.25	4.67
22	3,5-diMe,4-*t*-Bu	H	CH_3	*t*-C_4H_9	CH_3	950.11	951.87	1130.08	1134.20	8.49		4.28	4.52
23a	5-NO_2	H	H	H	NO_2	814.17	815.83	1054.64	1055.73	–5.26	820.8	–4.66	–2.80
23b	3-NO_2	H	NO_2	H	H	815.41	816.30	1040.79	1041.59	–7.73	820.8	–4.66	–4.11

24	4-NO_2	H	H	NO_2	H	814.25	816.12	1058.82	1060.76	–4.38	822.2	–1.96	–2.33
25	3,5-diNO_2	H	NO_2	H	NO_2	749.13	750.48	993.93	995.15	–15.87	759.4		–8.44
26	4-F	H	H	F	H	860.65	862.01	1089.01	1090.15	0.77	863.0		0.41
27	4-Cl	H	H	Cl	H	865.73	865.29	1088.09	1087.38	0.29	868.5	0.60	0.15
28	4-Br	H	H	Br	H	868.64	868.22	1088.36	1087.75	0.35		0.64	0.19
29a	3-Cl	H	Cl	H	H	868.03	869.99	1079.56	1081.63	–0.72		–0.48	–0.38
29b	5-Cl	H	H	H	Cl	874.68	876.33	1091.96	1093.43	1.35		–0.48	0.72
30a	3,4-diBr	H	Br	Br	H	857.91	859.85	1066.29	1067.21	–3.25		–1.83	–1.73
30b	4,5-diBr	H	H	Br	Br	860.73	862.33	1075.47	1076.06	–1.70		–1.83	–0.90
31a	4-NO_2,5-Me	H	H	NO_2	CH_3	839.04	840.94	1068.05	1067.13	–3.26		–1.25	–1.73
31b	3-Me,4-NO_2	H	CH_3	NO_2	H	840.60	843.17	1070.90	1072.40	–2.34		–1.25	–1.24
32a	3-Me,4-Cl	H	CH_3	Cl	H	889.20	891.88	1098.97	1100.68	2.62		1.42	1.39
32b	4-Cl,5-Me	H	H	Cl	CH_3	890.50	890.15	1098.68	1100.31	2.55		1.42	1.36
33a	3-Me,4-Br	H	CH_3	Br	H	891.58	894.05	1098.81	1100.73	2.63		1.46	1.40
33b	4-Br,5-Me	H	H	Br	CH_3	892.85	893.15	1098.61	1101.94	2.84		1.46	1.51
34a	3-Br,4-Me	H	Br	CH_3	H	888.29	891.22	1087.06	1089.06	0.58		0.24	0.31
34b	4-Me,5-Br	H	H	CH_3	Br	893.73	893.87	1097.88	1100.61	2.61		0.24	1.39
35a	4-Cl,5-Et	H	H	Cl	C_2H_5	896.48	897.74	1099.92	1099.74	2.45		1.50	1.30
35b	3-Et,4-Cl	H	C_2H_5	Cl	H	897.05	897.83	1099.71	1099.57	2.42		1.50	1.29
36a	4-Br,5-Et	H	H	Br	C_2H_5	898.76	899.48	1099.17	1100.59	2.60		1.53	1.38
36b	3-Et,4-Br	H	C_2H_5	Br	H	899.29	899.98	1100.21	1100.08	2.51		1.53	1.34
37a	3-Cl,5-Me	H	Cl	H	CH_3	891.01	896.04	1089.59	1094.03	1.45		0.30	0.77
37b	3-Me,5-Cl	H	CH_3	H	Cl	896.36	901.97	1102.50	1106.37	3.61		0.30	1.92
38a	3-Br,5-Me	H	Br	H	CH_3	895.69	902.26	1091.35	1094.32	1.50		0.45	0.80
38b	3-Me,5-Br	H	CH_3	H	Br	899.16	906.58	1102.01	1104.13	3.22		0.45	1.71
39	3,5-diMe,4-NO_2	H	CH_3	NO_2	CH_3	863.07	865.97	1078.16	1079.13	–1.16		–0.46	-0.62
40	3,5-diMe,4-Cl	H	CH_3	Cl	CH_3	911.71	914.99	1107.95	1107.30	3.78		2.22	2.0
41	3,5-diMe,4-Br	H	CH_3	Br	CH_3	913.59	915.38	1107.78	1108.28	3.95		2.30	2.101
42a	3,4-DiBr,5-Me	H	Br	Br	CH_3	880.37	881.84	1075.49	1075.44	–1.80		–0.96	–0.96
42b	3-Me,4,5-diBr	H	CH_3	Br	Br	882.30	885.46	1085.06	1085.26	–0.08		–0.96	–0.04
43a	3-Cyclopropyl	H	c-C_3H_5	H	H	926.75	929.26	1118.70	1121.90	6.34		3.10	3.37
43b	5-Cyclopropyl	H	H	H	c-C_3H_5	928.38	929.44	1121.35	1124.35	6.76		3.10	3.60
44a	3-Cycloprop.,4-Me,5-Et	H	C_2H_5	CH_3	c-C_3H_5	952.76	954.31	1134.57	1134.58	8.56		4.30	4.55
44b	3-Et,4-Me,5-cyclopropyl	H	c-C_3H_5	CH_3	C_2H_5	955.10	956.00	1127.69	1123.72	6.65		4.30	3.54
45a	3-Ph	H	C_6H_5	H	H	920.97	920.00	1105.55	1108.09	3.92	914.2	2.13	2.08
45b	5-Ph	H	H	H	C_6H_5	922.68	924.85	1108.83	1111.37	4.49	914.2	2.13	2.02
46a	4-Cl,5-Ph	H	H	Cl	C_6H_5	898.75	900.57	1090.28	1091.42	1.00		0.26	0.16
46b	3-Ph,4-Cl	H	C_6H_5	Cl	H	901.53	903.42	1089.99	1091.13	0.94		0.26	0.50
47a	4-Br,5-Ph	H	H	Br	C_6H_5	900.63	902.36	1090.75	1091.90	1.08		0.30	0.20
47b	3-Ph,4-Br	H	C_6H_5	Br	H	903.71	905.75	1090.72	1093.54	1.37		0.30	0.73
48a	4-Me,5-Ph	H	H	CH_3	C_6H_5	933.66	935.02	1113.44	1114.24	4.99		2.68	2.28
48b	3-Ph,4-Me	H	C_6H_5	CH_3	H	934.84	937.12	1113.01	1113.97	4.95		2.68	2.63
49a	3-Ph,5-Me	H	C_6H_5	H	CH_3	940.18	939.05	1114.97	1120.28	6.05	932.1	2.92	3.22
49b	3-Me,5-Ph	H	CH_3	H	C_6H_5	940.76	943.56	1118.59	1125.29	6.93	932.1	2.92	3.32

50a	3-Ph,5-Et	H	C_6H_5	H	C_2H_5	943.85	943.65	1114.95	1118.23	5.69	935.6		3.03
50b	3-Et,5-Ph	H	C_2H_5	H	C_6H_5	944.10	947.02	1118.31	1121.55	6.27	935.6		2.97
51a	3,4-diMe,5-Ph	H	CH_3	CH_3	C_6H_5	950.04	952.97	1122.18	1124.3	6.76		3.47	3.23
51b	3-Ph,4,5-diMe	H	C_6H_5	CH_3	CH_3	953.76	956.42	1121.73	1123.74	6.66		3.47	3.54
52a	3-Me,4-Br,5-Ph	H	CH_3	Br	C_6H_5	919.70	921.99	1099.99	1102.23	2.89		1.20	1.17
52b	3-Ph,4-Br,5-Me	H	C_6H_5	Br	CH_3	923.66	923.83	1099.68	1100.84	2.65		1.20	1.41
53	4-Phenyl	H	H	C_6H_5	H	906.42	908.08	1102.06	1102.86	3.00	906.0	1.64	1.60
54	3,5-diPh	H	C_6H_5	H	C_6H_5	944.72	948.48	1106.36	1108.46	3.98	946.3	1.43	1.75
55a	3-CF_3,5-Me	H	CF_3	H	CH_3	863.85	867.90	1073.29	1081.59	–0.73	866.8	–0.50	–0.30
55b[K]	3-Me,5-CF_3	H	CH_3	H	CF_3	868.22	872.83	1086.94	1091.87	1.07	866.8	–0.50	0.57
56[K]	3,5-bisCF_3	H	CF_3	H	CF_3	798.34	799.21	1031.09	1031.48	–9.51	808.7	–8.23	–5.06
57	4-CO_2Et	H	H	CO_2Et	H	878.61	880.81	1086.64	1088.02	0.40	880.7		0.21
58a	3-Me,5-CO_2Et	H	CH_3	H	CO_2Et	905.86	909.68	1097.26	1102.78	2.99	902.6		1.59
58b	3-CO_2Et,5-Me	H	CO_2Et	H	CH_3	912.53	916.99	1093.33	1095.06	1.63	902.6		0.87
59	3,5-CO_2Et	H	CO_2Et	H	CO_2Et	884.44	887.67	1062.36	1066.22	–3.42	881.6		–182
60a	3-Ph,5-CO_2Et	H	C_6H_5	H	CO_2Et	908.81	910.00	1084.07	1083.51	0.51	899.7		0.27
60b	3-CO_2Et,5-Ph	H	CO_2Et	H	C_6H_5	921.16	924.10	1086.32	1088.65	–0.39	899.7		–0.58
61a	3-NH_2	H	NH_2	H	H	927.37	929.37	1125.48	1128.50	7.49	921.5	4.11	3.98
61b	5-NH_2	H	H	H	NH_2	938.35	939.26	1131.18	1133.44	8.36	921.5	4.11	4.45
62	4-NH_2	H	H	NH_2	H	913.88	914.35	1114.84	1116.47	5.38	907.6	5.57	2.86
62c	4-NH_2	H	H	NH_2	H	864.08	864.80	1103.60	1103.14	3.05		5.57	5.92
63	3,5-diPh,4-OH	H	C_6H_5	OH	C_6H_5	944.29	945.72	1103.56	1108.47	3.98		2.01	1.75
64	1-Me	CH_3	H	H	H	912.83	912.85	1113.05	1113.02	4.78	912.0	2.09	1.74
65	1,3-diMe	CH_3	CH_3	H	H	932.53	934.60	1123.06	1124.37	6.77	933.9	2.82	2.80
66	1,5-diMe	CH_3	H	H	CH_3	932.60	937.49	1121.94	1122.87	6.51	934.3	2.89	2.66
67	1,4-diMe	CH_3	H	CH_3	H	927.56	928.31	1119.61	1121.13	6.20	928.4	2.48	2.49
68	1,3,5-triMe	CH_3	CH_3	H	CH_3	951.07	955.32	1131.47	1129.83	7.72	949.3	3.80	3.30
69	1,3,4-triMe	CH_3	CH_3	CH_3	H	944.95	947.18	1127.90	1127.96	7.40	941.5		3.13
70	1,4,5-triMe	CH_3	H	CH_3	CH_3	945.83	946.97	1128.64	1130.28	7.80	941.5		3.34
71	1-Me,3-*t*-Bu	CH_3	*t*-C_4H_9	H	H	944.82	943.05	1177.83	1183.88	7.56	944.4		3.22
72	1-Me,5-*t*-Bu	CH_3	H	H	*t*-C_4H_9	942.28	943.95	1121.39	1123.46	6.61	939.2		2.71
73	1-Me,3,5-di*t*-Bu	CH_3	*t*-C_4H_9	H	*t*-C_4H_9	969.17	971.47	1132.77	1137.85	9.13	970.8		4.05
74	1,3,4,5-tetraMe	CH_3	CH_3	CH_3	CH_3	960.86	964.96	1136.02	1138.02	9.16	962.1	4.27	4.07
75	1,4-diMe-3,5-di*t*-Bu	CH_3	*t*-C_4H_9	CH_3	*t*-C_4H_9	976.75	978.91	1136.28	1137.19	9.01	979.6		3.99
76	1-Me,3,5-diNO_2	CH_3	NO_2	H	NO_2	773.55	777.06	1000.15	999.75	–15.07	788.8		–8.82
77	1-Me,4-Br	CH_3	H	Br	H	890.16	890.48	1094.63	1095.68	1.74		0.18	0.11
78	1,3-diMe,4-Br	CH_3	CH_3	Br	H	910.51	912.11	1104.72	1106.76	3.68		0.87	0.12
79	1,5-diMe,4-Br	CH_3	H	Br	CH_3	908.61	909.70	1102.90	1104.78	3.34		0.91	0.97
80	1,3,5-triMe,4-Br	CH_3	CH_3	Br	CH_3	927.75	928.30	1112.34	1113.67	4.89		1.78	1.80
81	1,3-diMe,5-Br	CH_3	CH_3	H	Br	917.44	919.25	1108.19	1111.17	4.46		1.20	1.57
82	1,3-diMe,4,5-diBr	CH_3	CH_3	Br	Br	899.37	901.02	1091.73	1093.86	1.42		–0.60	–0.05
83	1,5-diMe,3,4-diBr	CH_3	Br	Br	CH_3	895.35	896.49	1080.67	1082.22	–0.62		–1.52	–1.13
84	1,3-diMe,5-CO_2Et	CH_3	CH_3	H	CO_2Et	920.78	925.79	1100.44	1101.01	2.67	924.9		0.62
85	1,5-diMe,3-CO_2Et	CH_3	CO_2Et	H	CH_3	929.40	931.41	1098.25	1102.22	2.89	933.4		0.73

86	1-Me,3,5-diCO_2Et	CH_3	CO_2Et	H	CO_2Et	902.42	907.71	1070.36	1073.83	–2.09	913.4		–1.92
87	1-Me,3-NO_2	CH_3	NO_2	H	H	839.78	836.58	1048.78	1051.41	–6.01	850.3	–4.58	–4.00
88	1-Me,5-NO_2	CH_3	H	H	NO_2	836.78	840.05	1060.95	1062.41	–4.09	847.6	–2.35	–2.98
89	1-Me,4-NO_2	CH_3	H	NO_2	H	837.46	838.63	1064.79	1065.80	–3.49		–2.18	–2.66
90	1,3,5-TriMe,4-NO_2	CH_3	CH_3	NO_2	CH_3	879.52	882.04	1082.57	1084.18	–0.27		–0.95	–0.95
91	1-Me,3-Ph	CH_3	C_6H_5	H	H	937.51	933.71	1110.63	1113.49	4.86	932.6		1.78
92	1-Me,5-Ph	CH_3	H	H	C_6H_5	939.67	941.36	1115.46	1117.76	5.61	932.4		1.81
93	1,3-diMe,5-Ph	CH_3	CH_3	H	C_6H_5	956.43	959.76	1124.49	1127.04	7.24	956.6		2.68
94	1,5-diMe,3-Ph	CH_3	C_6H_5	H	CH_3	954.29	953.77	1118.65	1122.14	6.38	954.3		2.59
95	1-Me,3,5-diPh	CH_3	C_6H_5	H	C_6H_5	958.71	959.82	1112.48	1116.11	5.32	958.9	1.26	1.66
96	1-Me,3-NH_2	CH_3	NH_2	H	H	942.81	945.28	1129.19	1133.65	8.39	937.4	3.81	3.66
97	1-Me,5-NH_2	CH_3	H	H	NH_2	951.47	954.26	1133.48	1134.52	8.55	949.5	4.23	3.74
98	1-Me,4-NH_2	CH_3	H	NH_2	H	933.11	936.58	1120.05	1119.02	5.83		5.52	2.30
98c	1-Me,4-NH_2	CH_3	H	NH_2	H	877.75	881.81	1107.19	1107.61	3.83		5.52	5.53
99	1-Me,3,5-diPh,4-OH	CH_3	C_6H_5	OH	C_6H_5	958.30	961.65	1109.39	1111.89	4.58		1.96	1.26
100	1,3-diMe,5-OEt	CH_3	CH_3	H	OC_2H_5	965.61	969.83	1133.29	1135.96	8.80		3.51	3.88
101	1,5-diMe,3-OEt	CH_3	OC_2H_5	H	CH_3	967.24	970.17	1125.99	1128.66	7.52		2.05	3.20
102	1-Et	C_2H_5	H	H	H	919.72	921.26	1113.50	1115.45	5.21		2.00	1.97
103	1-Et,3-NO_2	C_2H_5	NO_2	H	H	847.92	848.04	1048.43	1049.76	-6.30		–4.71	–4.16
104	1-Et,5-NO_2	C_2H_5	H	H	NO_2	846.90	849.38	1062.04	1060.91	–4.35		–2.32	–3.12
105	1-Et,4-NO_2	C_2H_5	H	NO_2	H	845.68	848.49	1064.03	1064.35	–3.75		–2.13	–2.80
106	1-Bu	*n*-C_4H_9	H	H	H	926.30	926.72	1113.34	1113.66	4.89	928.8		1.80
107	1-*t*-Bu	*t*-C_4H_9	H	H	H	933.83	936.79	1115.61	1115.21	5.16		1.95	1.94
108	1-Adamantyl	1-Ad[c]	H	H	H	948.26	950.80	1118.93	1119.34	5.89	954.5		2.33
109	1-Phenyl (Ph)	C_6H_5	H	H	H	922.77	925.26	1099.19	1102.41	2.92		0.44	0.75
110	1-Ph,3,5-diMe	C_6H_5	CH_3	H	CH_3	964.57	974.98	1121.89	1124.14	6.73		2.65	2.77
111	1,5-diPh,3-Me	C_6H_5	CH_3	H	C_6H_5	966.07	971.92	1115.76	1119.17	5.86		1.51	1.94
112	1,3,5-Triphenyl	C_6H_5	C_6H_5	H	C_6H_5	968.83	968.72	1105.78	1106.68	3.67		0.39	0.78
113	1-Phenyl,4-Cl	C_6H_5	H	Cl	H	899.68	902.24	1082.93	1085.87	0.02		–1.40	–0.79
114	1-Phenyl,4-Br	C_6H_5	H	Br	H	901.93	904.10	1082.21	1085.31	–0.08		–1.32	–0.85
115	1-Phenyl-3-NH_2	C_6H_5	NH_2	H	H	946.74	951.88	1115.69	1118.39	5.72		2.91	2.24
116	1-Phenyl-5-NH_2	C_6H_5	H	H	NH_2	968.47	977.58	1128.33	1133.54	8.37		3.09	3.65
117	1-Phenyl-4-NH_2	C_6H_5	H	NH_2	H	938.97	940.69	1104.66	1106.77	3.68		4.73	1.15
117c	1-Phenyl-4-NH_2	C_6H_5	H	NH_2	H	873.39	876.80	1098.15	1099.75	2.45		4.73	4.80
118	1-Ph,3-Me,5-NH_2	C_6H_5	CH_3	H	NH_2	982.53	990.91	1135.11	1135.54	8.73		4.76	3.84
119	1-Ph,3-Me,5-OEt	C_6H_5	CH_3	H	OC_2H_5	971.99	980.31	1119.75	1123.13	6.55		2.34	2.68
120	1-Ph,3-MeO,5-Me	C_6H_5	OCH_3	H	CH_3	946.64	955.65	1103.60	1111.86	4.58		1.17	1.63
121	1Ph,3,4-diMe,5-OEt	C_6H_5	CH_3	CH_3	OC_2H_5	979.55	986.23	1124.54	1127.44	7.31		2.55	3.08
122	1-pNP	pNP[e]	H	H	H	878.30	880.42	1081.42	1084.03	–0.30		–0.65	–0.96
123	1-pNP,3-Me	pNP[e]	CH_3	H	H	895.22	899.97	1090.14	1089.56	0.67		0.14	–0.45
124	1-pNP,5-Me	pNP[e]	H	H	CH_3	906.47	913.89	1097.65	1100.91	2.66		0.62	0.61
125	1-pNP,4-Me	pNP[e]	H	CH_3	H	889.85	893.83	1085.55	1087.28	0.27		–0.16	–0.66
126	1-pNP,3,4-diMe	pNP[e]	CH_3	CH_3	H	904.84	904.56	1093.33	1097.13	1.99		0.51	0.25
127	1-pNP,4,5-diMe	pNP[e]	H	CH_3	CH_3	918.41	924.91	1102.42	1106.75	3.68		1.08	1.15

128	1-pNP,3,5-diMe	pNP[e]	CH_3	H	CH_3	922.32	933.98	1105.40	1107.25	3.77		1.50	1.20
129	1-pNP,3,4,5-triMe	pNP[e]	CH_3	CH_3	CH_3	931.38	943.72	1110.00	1117.74	5.61		1.87	2.18
130	1-pNP,3-NH_2	pNP[e]	NH_2	H	H	902.81	907.02	1099.32	1108.37	3.96		1.93	1.30
131	1-pNP,5-NH_2	pNP[e]	H	H	NH_2	925.14	933.04	1111.31	1112.81	4.74		2.11	1.72
132	1-pNP,4-NH_2	pNP[e]	H	NH_2	H	893.19	893.95	1084.64	1086.33	0.1		4.05	–0.75
132c	1-pNP,4-NH_2	pNP[e]	H	NH_2	H	839.00	840.76	1087.11	1087.04	0.23		4.05	3.62
133	1-DNP	DNP[f]	H	H	H	875.64	876.58	1078.16	1076.97	–1.54		–1.40	–1.62
134	1-DNP,3-Me	DNP[f]	CH_3	H	H	893.82	896.01	1086.16	1087.35	0.28		–0.28	–0.66
135	1-DNP,3-Et	DNP[f]	C_2H_5	H	H	898.03	899.65	1085.46	1089.17	0.60		–0.42	–0.49
136	1-DNP,3-*t*-Bu	DNP[f]	*t*-C_4H_9	H	H	906.49	906.24	1086.41	1083.38	-0.41		–0.60	–1.02
137	1-DNP,5-Me	DNP[f]	H	H	CH_3	893.45	899.27	1088.04	1099.03	2.33		–0.13	0.43
138	1-DNP,4-Me	DNP[f]	H	CH_3	H	886.73	888.95	1081.62	1081.76	-0.70		–0.90	–1.18
139	1-DNP,3,4-diMe	DNP[f]	CH_3	CH_3	H	903.41	905.99	1089.08	1089.02	0.57		–0.10	–0.50
140	1-DNP,4,5-diMe	DNP[f]	H	CH_3	CH_3	904.78	909.72	1089.56	1096.79	1.94		0.20	0.23
141	1-DNP,3,5-diMe	DNP[f]	CH_3	H	CH_3	909.19	919.11	1095.55	1106.27	3.60		0.55	1.11
142	1-DNP,3,4,5-triMe	DNP[f]	CH_3	CH_3	CH_3	917.65	925.75	1100.55	1102.90	3.01		0.70	0.80
143	1-DNP,2*H*-campho[d]	DNP[f]	---	---	H	917.69	918.13	1089.43	1091.69	1.04		–1.00	–0.25
144	1-DNP,1*H*-campho[d]	DNP[f]	H	---	---	933.05	938.00	1102.18	1106.08	3.56		0.36	1.09
145	1-TNP	TNP[g]	H	H	H	857.16	862.63	1066.04	1067.45	-3.20		–2.00	–2.51
146	1-Acetyl	$COCH_3$	H	H	H	865.79	867.77	1075.68	1077.59	-1.43		2.94	–1.57
147	1-NO_2	NO_2	H	H	H	834.43	837.72	1047.96	1049.75	-6.30		–4.15	–4.16
148	1-NH_2	NH_2	H	H	H	894.13	895.17	1100.76	1099.84	2.47		0.00[l]	0.51
148c	1-NH_2	NH_2	H	H	H	799.31	801.83	1047.65	1048.75	-6.48		0.00[l]	–4.25
149	1-Chloro	Cl	H	H	H	866.32	867.80	1075.59	1077.60	–1.43			–1.57
150	1-C(NH)NH_2,3,5-diMe	[h]	CH_3	H	CH_3	927.85	933.20	1103.20	1103.90	3.18			0.89
150c	1-C(NH)NH_2,3,5-diMe	[h]	CH_3	H	CH_3	994.31	997.37	1167.66	1169.23	14.63			6.98

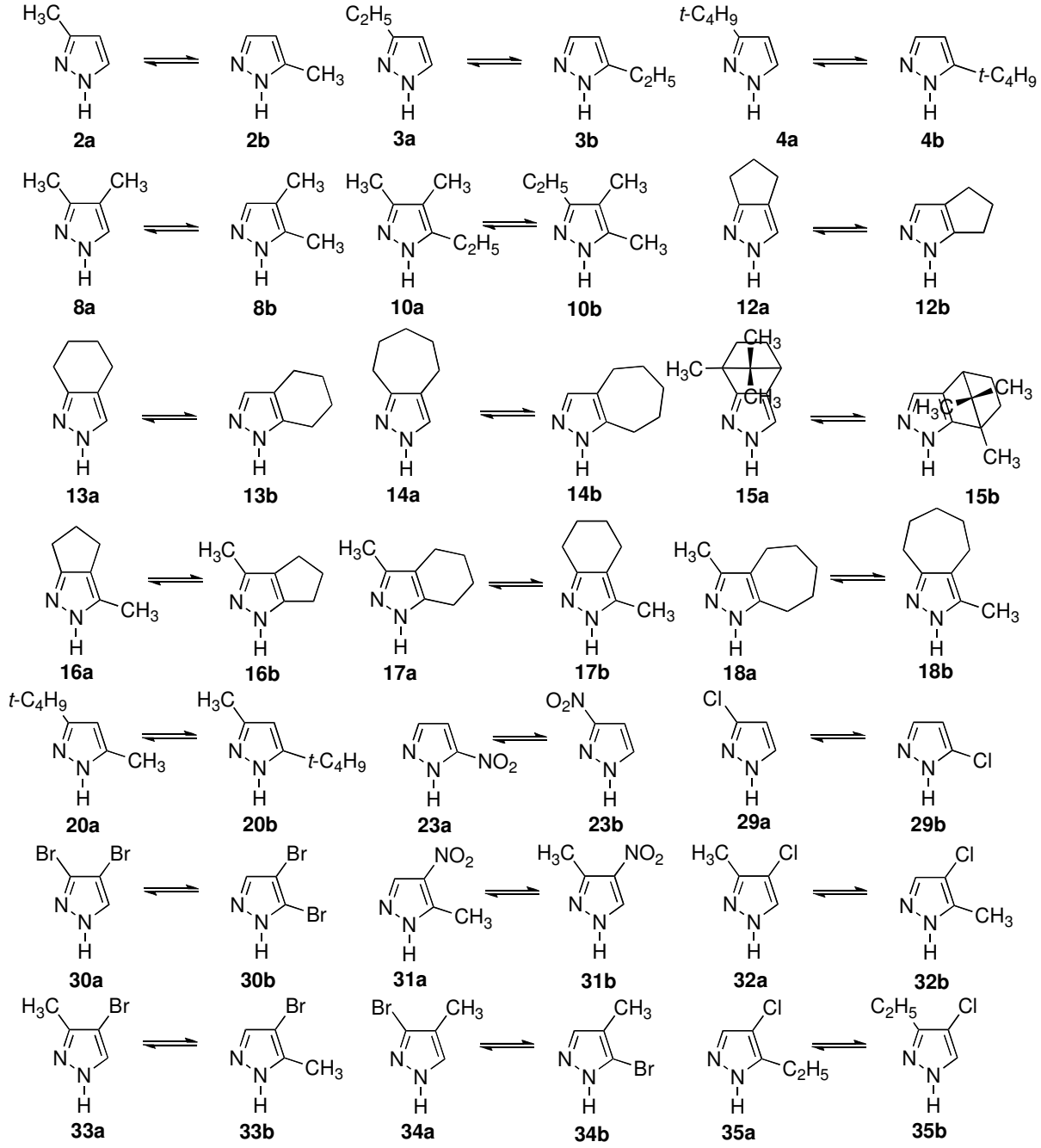
[a] Tautomer **4a** with solvent effect has an imaginary frequency (*i*17cm^{-1}); [b] Compound **6** with solvent effect has an imaginary frequency (*i*52cm^{-1}); [c] 1-Ad; [d] (4*S*,7*R*)-7,8,8-trimethyl-4,5,6,7-tetrahedro-4,7-methano-2(1)*H*-indazole; [e] pNP; [f] DNP; [g] TNP. [h] 3,5-dimethyl-1H-pyrazole-1-carboximidamide; **150c**, 1-carboximidamidinum cation. [i] pK_a = 0.1752*[ΔGaq + ΔGs(H^+)]; where ΔGs(H^+) = –1112.5 kJ·mol^{-1}, these values differ from those calculated according to note j by 4.69 pK_a units and are only reported in Table S1; [j] pK_a 0.1752*[ΔGaq + ΔGs(H^+)]; where ΔGs(H^+) = –1085.7 kJ·mol^{-1}. [k] From reference 18; [l] Experimental result <0.4, assumed to be 0.00.

3. Results and discussion

3.1. The gas phase.

3.1.1. The tautomerism of pyrazoles. Neutral molecules.

We have represented in Fig. 2 the pyrazole tautomers when they are different, i.e., when $R^3 \neq R^5$. We have labeled **a** the most abundant tautomer in the gas phase according to our calculations, even if in some cases they differ from those reported in the literature. However, note that most literature results come from condensed phases (solution and solid state)[36] and will be discussed in section 3.2 (water solution).



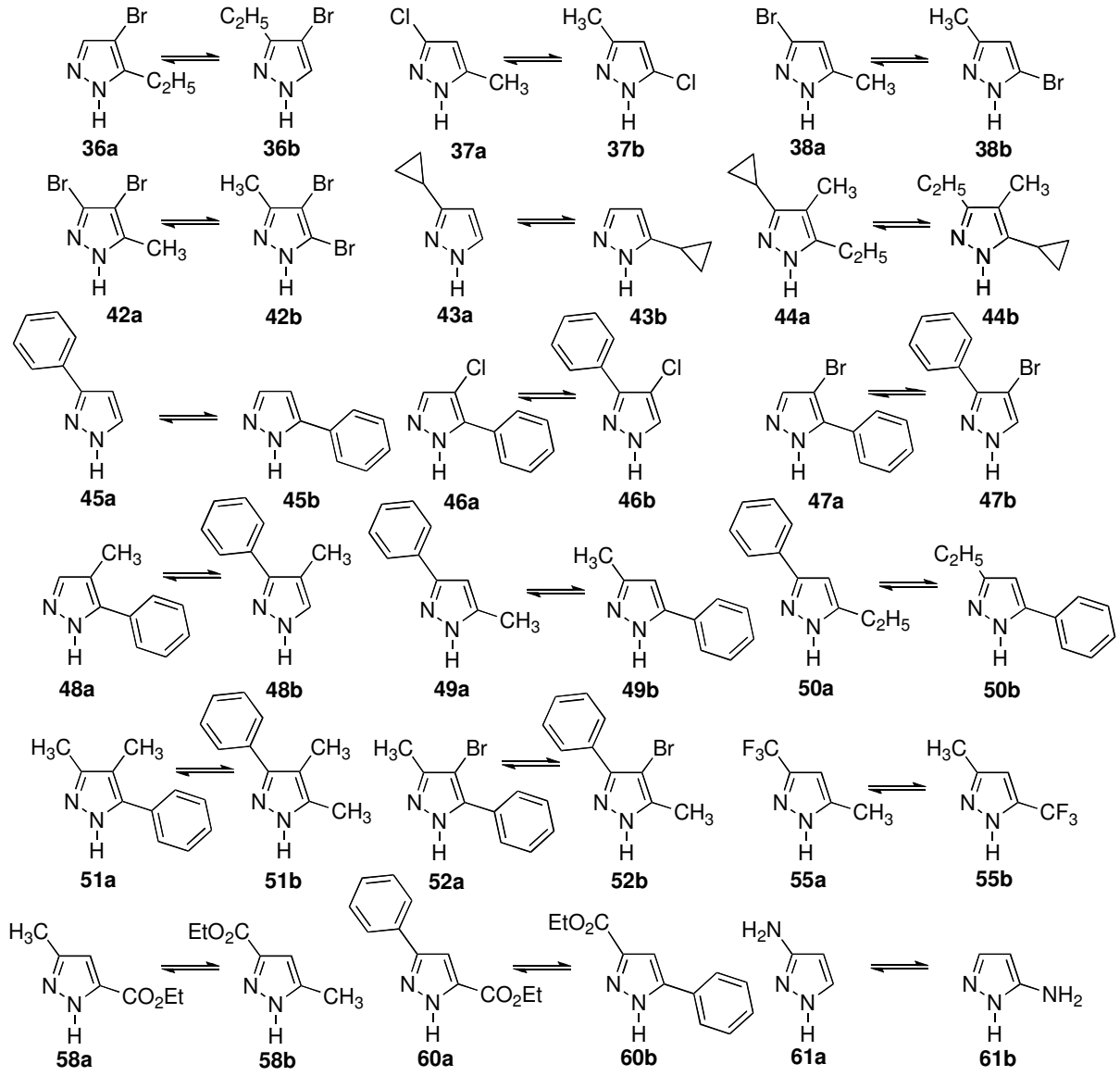


Fig. 2 Pairs of tautomeric pyrazoles.

We have reported in Table 2 the differences in energy between tautomers **a** and **b** as well as other properties related to the tautomerism of pyrazoles.

Table 2 Differences in energy including the zero-point vibrational energy (ZPVE) correction (kJ·mol^{-1}) between tautomers **a** and **b** as well as the effect of the water solvation defined as ΔΔE (water–gas). Columns: 1st is the compound number of Table 1; 2nd is the name of the pyrazole tautomer **a**; 3rd specifies the substituent at position 3 of tautomer **a**; 4th is the difference in energy between the **a** (0 by definition) and **b** tautomers in the gas phase; 5th is the corresponding values for the water solution (PCM) but here **a** is not always the most stable tautomer (negative values correspond to the **b** tautomer being more stable); 6th is the difference in energy, ΔΔE, when tautomer **a** is the most stable in the gas phase and in water while the 7th is the difference in energy (actually a sum), ΔΔE, when tautomer **a** is the most stable in the gas phase and tautomer **b** in water; 8th and 9th reports the populations of tautomers **a** and **b** in the gas phase and in water, respectively; 10th summarizes the calculated results of columns 6 and 7; finally, 11th and 12th report the experimental results. 11th also contains a comment about the agreement or not between columns 10 and 11.

1	2	3	4	5	6	7	8		9		10	11	12
No	pyrazole	**a** taut.	gas	water	ΔΔE	ΔΔE	gas		water		tautomer	Exp.	Cond.
		3-subs	**b**	**b**	**a**	**b-a**	**% a**	**% b**	**% a**	**% b**			
2	3-methyl	Me	1.36	0.24	–1.12		63.4	36.6	52.4	47.6	**2a≈2b**	**2a≈2b**[a]	H_2O, HMPA
3	3-ethyl	Et	1.74	–1.84		–3.58	66.9	33.1	32.2	67.8	**3b**		
4	3-*t*-butyl	*t*-Bu	1.8	–2.03		–3.83	67.4	32.6	30.6	69.4	**4b**		
8	3,4-dimethyl	Me	2.94	1.32	–1.62		76.6	23.4	63.0	37.0	**8a**	**8b**[b]	X-ray
10	3,4-dimethyl,5-ethyl	Me	1.16	1.47	+0.31		61.5	38.5	64.4	35.6	**10a**		
12	3,4-trimethylene	CH_2	5.72	4.36	–1.36		91.0	9.0	85.3	14.7	**12a**		
13	3,4-tetramethylene	CH_2	0.59	0.02	–0.57		55.9	44.1	50.2	49.8	**13a≈13b**	**13≈13b**[a]	$CDCl_3$, X-ray
14	3,4-pentamethylene	CH_2	0.66	–0.39		–1.05	56.6	43.4	46.1	53.9	**14a≈14b**		
15	camphopyrazole	2*H*	12.00	10.42	–1.58		99.2	0.8	98.5	1.5	**15a**	**15a**[a]	AcOH, X-ray
16	3,4-trimethylene,5-methyl	CH_2	3.74	3.95	+0.21		81.9	18.1	83.1	16.9	**16a**	**16a≈16b**	$CDCl_3$, DMSO
17	3-methyl,5,4-tetramethylene	Me	1.29	1.38	+0.09		62.7	37.3	63.6	36.4	**17a**		
18	3-methyl,5,4-pentamethylene	Me	0.94	0.85	–0.09		59.4	40.6	58.5	41.5	**18a≈18b**		
20	3-*t*-butyl,5-methyl	*t*-Bu	0.22	–1.70		–1.92	52.2	47.8	33.5	66.5	**20b**	**20b**[a]	$CDCl_3$
23	5-nitro	H	1.18	–14.22		–15.40	61.7	38.3	0.3	99.7	**23b**	**23b**[a]	H_2O, X-ray
29	3-chloro	Cl	6.53	12.07	+5.54		93.3	6.7	99.2	0.8	**29a**		
30	3,4-dibromo	Br	2.72	9.02	+6.30		75.0	25.0	97.4	2.6	**30a**		
31	5-methyl,4-nitro	H	1.74	3.39	+1.65		66.9	33.1	79.7	20.3	**31a**	**31a>31b**[a]	Toluene, X-ray
32	3-methyl,4-chloro	Me	0.98	–0.31		–1.29	59.8	40.2	46.9	53.1	**32a≈32b**		
33	3-methyl,4-bromo	Me	1.01	0.22	–0.79		60.1	39.9	52.2	47.8	**33a≈33b**	**33b**[a]	DMSO, X-ray
34	3-bromo,4-methyl	Br	5.01	10.96	+5.95		88.3	11.7	98.8	1.2	**34a**		
35	5-ethyl,4-chloro	H	0.46	–0.42		–0.88	54.6	45.4	45.8	54.2	**35a≈35b**		
36	5-ethyl,4-bromo	H	0.46	0.45	–0.01		54.6	45.4	54.5	45.5	**36a≈36b**		
37	3-chloro,5-methyl	Cl	5.32	12.63	+7.31		89.6	10.4	99.4	0.6	**37a**		
38	3-bromo,5-methyl	Br	3.52	10.28	+6.76		80.6	19.4	98.4	1.6	**38a**	**38a**[a]	H_2O
42	3,4-dibromo,5-methyl	Br	2.14	9.48	+7.34		70.4	29.6	97.9	2.1	**42a**		
43	3-cyclopropyl	Cy	1.56	2.70	+1.14		65.2	34.8	74.8	25.2	**43a**		
44	3-cyclopropyl,4-methy,5-ethyl	Cy	2.35	4.50	+2.15		72.1	27.9	86.0	14.0	**44a**		
45	3-phenyl	Ph	1.89	3.25	+1.36		68.2	31.8	78.8	21.2	**45a**	**45a**[a]	HMPA, acetone

46	5-phenyl,4-chloro	H	2.81	–0.30		–3.11	75.7	24.3	47.0	53.0	**46a≈46b**		
47	5-phenyl,4-bromo	H	3.13	0.39	–2.74		78.0	22.0	53.9	46.1	**47a≈47b**	**47b**	X-ray
48	5-phenyl,4-methyl	H	1.43	–0.44		–1.87	64.0	36.0	45.6	54.4	**48a≈48b**		
49	3-phenyl,5-methyl	Ph	0.93	3.94	+3.01		59.3	40.7	83.1	16.9	**49a**	**49a≈49b**	HMPA, X-ray
50	3-phenyl,5-ethyl	Ph	0.52	4.26	+3.74		55.2	44.8	84.8	15.2	**50a**		
51	3,4-dimethyl,5-phenyl,	Me	3.78	–0.44		–4.22	82.1	17.9	45.6	54.4	**51a≈51b**		
52	3-methyl,4-bromo,5-phenyl	Me	3.61	–0.54		–4.15	81.1	18.9	44.6	55.4	**52a≈52b**		
55	3-trifluoromethyl,5-methyl	CF_3	4.35	13.14	+8.79		85.3	14.7	99.5	0.5	**55a**	**55a**[a]	$CDCl_3$
58	3-methyl,5-ethoxycarbonyl	Me	11.77	–0.53		–12.30	99.1	0.9	44.7	55.3	**58b**	**58b**[a]	DMSO, X-ray
60	3-phenyl,5-ethoxycarbonyl	Ph	12.37	–0.36		–12.73	99.3	0.7	46.4	53.6	**60a**	**60a**[a]	DMSO, X-ray
61	3-amino	NH_2	10.68	5.40	–5.28		98.7	1.3	89.9	10.1	**61a**	**61a**[a]	DMSO

[a] Agree; [b] Disagree.

Using the Boltzmann equation (T = 298 K) we have calculated the populations of tautomers. We will comment here Table 2 gas phase results concerning tautomerism in what concerns some inversion in stability of the phenyl group at positions 3(5). Tautomer **a** (3-phenyl) is the most stable for compounds **45**, **49** and **50** without substituents at position 4; on the other hand, tautomer **b** (5-phenyl) is the most stable for compounds **46**, **47**, **48**, **51** and **52**. This is related to the conformation of the 3-phenyl group in **a** tautomers: for the first three compounds the dihedral angle between rings ranges from 0 to 4º while for the last five ones it ranges from 29 to 33º (see Table S2). A planar phenyl group prefers the 3-position while a twisted one prefers the 5-position.

3.1.2. The tautomerism of pyrazoles. Protonated molecules.

Concerning the structure of the cations, for all of them we have calculated the structures resulting from the protonation on the N2 nitrogen atom (pyrazolium salts); besides, for six of them, we have also calculated the cations (pyrazoles **c**) resulting from the protonation of the functional group (Fig. 3), mainly of the amino group (ammonium salts).

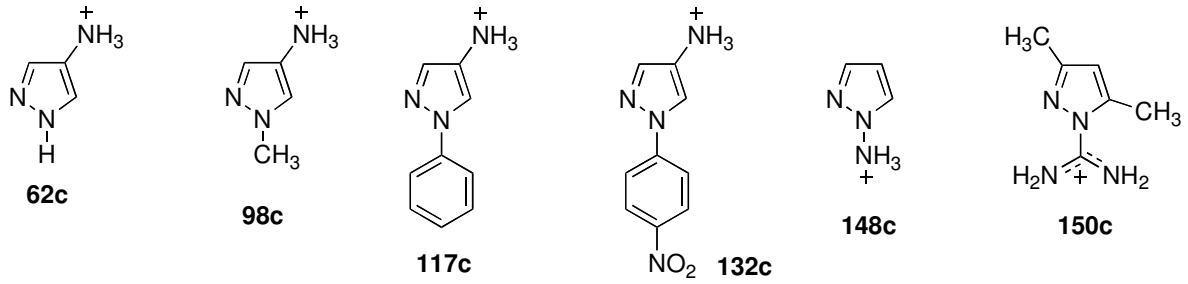


Fig. 3 Cations protonated on the amino and amidine groups.

In Table 3 are reported the differences in stability ($kJ \cdot mol^{-1}$) of the cations corresponding to the pyrazoles of Fig. 3. The 7th column indicates the tautomer (**c** or **a**) found experimentally.

Table 3 Relative stabilities of N_2 *vs*. functional group cations including the ZPVE correction. The values of Δ*H* Gas for **c** cations of Table 1 are also reported.

No	Name	Δ*H* Gas ($kJ \cdot mol^{-1}$)	gas phase		water		ΔEE **a**[a]	Exp.	Ref.
			N_2 **a**	subs. **c**	N_2 **a**	subs. **c**			
62	4-NH_2	864.08	0	50.9	0	12.7	–38.2	**62c**	[37]
98	1-Me,4-NH_2	877.75	0	56.0	0	12.9	–43.1	**98c**	[37]
117	1-Phenyl,4-NH_2	873.39	0	66.0	0	7.0	–59.0	**117c**	[37]
132	1-pNP,4-NH_2	839.00	0	54.7	1.6	0	–56.3	**132c**	[37]
148	1-NH_2	799.31	0	94.6	0	52.6	–42.0	**148a**	[38]
150	1-Carboximidamide	994.31	66.2	0	65.0	0	–1.2	---	
---	aniline	871.65[b]						---	

[a] Water effect on the stability of tautomer **a** calculated as **c** – **a**; [b] Experimental value, 876.55 $kJ \cdot mol^{-1}$.[36]

In the gas phase, excluding the very basic amidine derivative **150**, all aminopyrazoles, either 4-NH_2 or 1-NH_2, protonate on the ring. The protonation on the C-amino group lead to compounds having similar basicities than aniline (Table 3); the *N*-amino group decreases considerably the basicity of the pyrazole ring compared with the effect of the 4-amino one.

3.1.2. Proton affinities.

For proton affinities (PAs), we have used all the NIST values that differ from those published in our papers by being corrected for temperature effects.[39] There are some values that were not reported in the NIST database (Table S1). We have transformed these data into PA NIST corr. empirically by means of equation 5:

PA NIST corr. = –(63±35) + (1.07±0.04) PA, n = 14, R^2 = 0.985 (5)

We have then compared the PA NIST corr. with the calculated Δ*H* values:

PA NIST corr. = (67±13) + (0.92±0.01) Δ*H*, n = 69, R^2 = 0.984 (6)

This is already acceptable but the correlation improves eliminating the less stable **b** tautomers:

PA NIST corr. = (65±11) + (0.93±0.01) Δ*H*, n = 57, R^2 = 0.990 (7)

Fig. 4 is the plot corresponding to eq. 7.

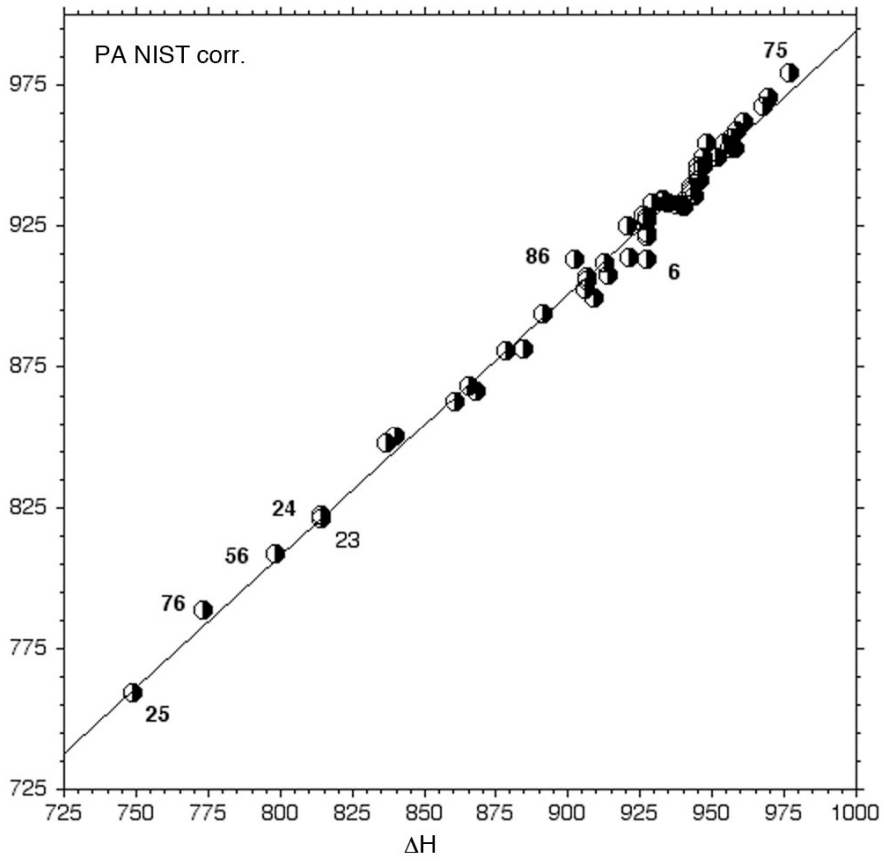


Fig. 4 Plot corresponding to eq. (7) with worse points (**6**, **75**, **86**) and less basic pyrazoles (**23**, **24**, **25**, **56**, **76**).

We have repeated this analysis using only the NH-derivatives:

PA NIST corr. = (79±15) + (0.91±0.02) Δ*H*, n = 42, R^2 = 0.986 (8)

PA NIST corr. = (71±14) + (0.91±0.02) Δ*H*, n = 30, R^2 = 0.992 (9)

The elimination of the less abundant tautomers improves the correlations (compare eqs. 7 *vs*. 6 and 9 *vs*. 8) but the quality of the regressions is not sufficient to determine the proportions of both tautomers (the tautomeric equilibrium constant, K_T) since the largest deviations are found for compounds, like **6** (4-adamantyl-1*H*-pyrazole-3.5-dicarboxylic acid diethyl ester) and **86** (diethyl ester of 1-methyl-1*H*-pyrazole-3,5-dicarboxylic acid), that are not involved in tautomerism. There is no anomaly in our gas phase calculations but there is a significant intercept of about 65 kJ·mol^{-1}, that it is related to the less basic pyrazoles (bottom-left of Fig. 4); removing them the intercept decreases to 19 kJ·mol^{-1}.

3.2. The aqueous phase.

We will now discuss the basicity of pyrazoles in water at 25 ºC (thermodynamic values) of Table 1. When the pK_as have been determined at another temperature, for instance at 20 ºC, or in the presence of small amounts of a co-solvent, for instance 5-10% of methanol or ethanol, these values have been empirically corrected.[9,10]

3.2.1. The tautomerism of pyrazoles. Neutral molecules.

On going from the gas phase to water solution (Table 2) in most cases (26) the predominant tautomer **a** remains unchanged but in several other (13) an inversion is observed and tautomer **b** becomes the most stable. Considering only the effects > |2| kJ·mol^{-1}, with some substituents, the stability of tautomer **a** increases, this is the case of Cl (**29**, **37**), Br (**34**, **38**, **42**), planar phenyl (**49**, **50**) and CF_3 (**55**) while for other it decreases: ethyl (**3**), *t*-butyl (**4**), twisted phenyl (**46**, **47**, **51**, **52**), cyclopropyl (**44**), ethoxycarbonyl (**58**, **60**), amino (**61**) and specially nitro (**23**).

The experimental data on some of these compounds come from rather different conditions, solution in solvents as different as chloroform, DMSO, water and others, as well as solid state (NMR and crystallography, usually only one tautomer). Even though, the agreement (Table 2) is satisfactory, with very few exceptions. The most evident is **8** where the calculations predict tautomer **a** to be the most stable by 1.32 kJ·mol^{-1}, whereas the experimental result corresponds to the **b** tautomer but it corresponds to a crystal structure.

Using ΔG instead of ΔE does not change the main conclusions because both parameters are correlated:

ΔG = (0.99±0.04) ΔE, n = 39, R^2 = 0.955 (10)

3.2.2. The tautomerism of pyrazoles. Protonated molecules.

As reported in Table 3, water causes a truly enormous decrease (between 38 and 59 kJ·mol^{-1}) of the stability of tautomer **a** in the case of aminopyrazoles. The calculations reproduce adequately the behavior of 1-aminopyrazole (**148**), but in the case of C-aminopyrazoles, the calculations reproduce the greater stability of the ammonium salt **132c** but not those of the remaining compounds, where differences of 7-13 kJ·mol^{-1} in favor of **a** tautomers are calculated.

3.2.3. Thermodynamic pK_a values.

The first and most obvious observation of Table S1 is the effect of 4.69 pK_a units due to the value used for the free energy of the H^+. The following points have to be removed **62**, **98**, **117** and **132** all of them 4-aminopyrazoles because it has been determined experimentally that they do not protonate on the pyrazole ring but on the amino group;[37] the values for the protonation on the amino group are consistent with the experimental pK_a. In the case of 1-aminopyrazole (**148**) it protonates on the pyrazole ring,[38] as the agreement with the experimental pK_a confirms. After several attempts (see S3 in the ESI), we found that the best equation is the following one (using pK_a^i instead of pK_a^j adds an intercept related to the 4.69 difference):

Exp. pK_a = (0.53±0.01) Calc. pK_a^j – (0.81±0.05) NR – (0.37±0.12) 5-Ph + (4.30±0.21) NH_3^+, n = 148, R^2 = 0.979 (11)

i. All **b** tautomers clearly less stable than the **a** ones have been removed in equation (11): **15b** (camphopyrazole tautomer 2*H* is much more stable than 1*H*), **29b**, **30b**, **34b**, **37b**, **38b**, **42b** (all of them 5-halo substituted) and **55b** (5-CF_3 substituted).

ii. All other effects being equal, experimental values are in average 53% of the calculated ones.

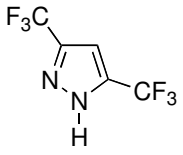
iii. NR is a dummy variable that has the value 0 for the NH and the value 1 for the NR pyrazoles. The *N*-unsubstituted pyrazoles are weaker bases than the corresponding *N*-substituted ones. The decrease of 0.81 pK_a units is related to the statistical effect that amounts to 2 for NH-pyrazoles having identical substituents at positions 3 and 5.[9,10,40]

iv. 5-Ph is a dummy variable that has the value 0 for all pyrazoles save those having a 5-phenyl substituent. The origin of the small but significant effect, –0.37±0.12, is probably related to the conformation of the phenyl ring that could not be well reproduced by the calculations in water both for the neutral and the protonated forms.

v. The very big effect necessary to account for the pK_a of the 4-aminopyrazoles (**62**, **98**, **117** and **132**) is related to the NH_3^+ cation, not only the statistical factor of 3 but a different kind of structure. They are more related to aniline that we have calculated at the same level (equation 12).

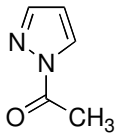
Exp. pK_a = (4.4±0.3) + (0.39±0.08) Calc. pK_a^j – (0.53±0.25) NR, n = 5, R^2 = 0.947 (12)

Two outlier points remain (Fig. 5):



56

Exp. pK_a = –8.23
Fitted pK_a = –5.06



146

Exp. pK_a = 2.94
Fitted pK_a = –1.57

Fig. 5 Outliers.

The pK_a value of **56** has not been measured being too weak but estimated from its acid (proton loss) pK_a; compound **146** is an azolide[41] and therefore very easy to hydrolize: both experimental values are probably wrong.

This equation predicts for **150**, an important compound,[42] a pK_a of 0.92 if protonates on N2 and 6.96 if protonates on the amidine (the N1 atom of pyrazole has its lone pair

involved in the aromatic sextet, thus compound **150** is not a guanidine). Although it is clear that it will behave as a guanidine, the second value should be considered with care in case **150** does not follow eq. (11) similarly to what happens to 4-aminopyrazoles.

Eq. 11 is useful to predict values and we have added them to the last column of Table 1 but it is not entirely satisfactory since it has a slope of 0.53, that is, the calculated pK_as are almost twice the experimental ones. However, an examination of the literature in what concerns pK_a calculations show that the reported results are not better than ours, for instance, the mean absolute deviation (MAD) for 30 organic molecules[43] and for 55 neutral organic and inorganic acids[35] using the direct method (like us) is between 2.5 and 3 pK_a units, and for our much larger set is 1.9 pK_a units. Using linear correlations, other authors have found R^2 values ranging from 0.720 to 0.908.[44] Only for small sets (13 heterocyclic compounds) bearing only H, CH_3 and Cl substituents better correlations have been found.[45]

4. Concluding remarks

We have studied a large variety of compounds that although homogeneous in structure, being all pyrazoles, differ in their protonation sites and in their tautomeric stability. The range of experimental PA is 208.1 kJ·mol^{-1} (from 759.4 to 967.5) and that of pK_as is 10.28 pK_a units (from –4.71 to +5.57), range that increases if predicted values are included (15.80, from –8.82 to +6.98).

We have chosen the more demanding approach, leaving aside methods, like the isodesmic and other thermodynamic cycles, that use a molecule, generally the parent compound, as reference.[34,46,47] Our results prove that this restriction is not absolutely necessary.

Although most authors prefer for the proton solvation free energy the value of –1112.5 kJ·mol^{-1},[32,33,34] we have obtained better results with –1085.7 kJ·mol^{-1}, further works are necessary to decide between both values.

Even if the main purpose of this paper is the calculation of acid-base properties of pyrazoles, it also provides information to the related topic of NH-pyrazoles tautomerism (remember that both tautomer share the same conjugated acid, the pyrazolium cation). Tautomerism of pyrazoles is still a matter of some confusion (see, ref.[48] for a paper where only one pyrazole tautomer is considered in the ground state without any proof).

Acknowledgements

This work has been supported by the Spanish Ministerio de Economía y Competitividad (CTQ2012-35513-C02-02) and Comunidad Autónoma de Madrid (S2013/MIT-2841, Fotocarbon). Computer, storage and other resources from the CTI (CSIC) are gratefully acknowledged. One of us (M.M.-L.) benefits from a contract from the Ministerio de Economía y Competitividad.

References

1 (*a*) K. Tsaioun and S. A. Kates, Eds. *ADMET for Medicinal Chemists: A Practical Guide*, Wiley, Hoboken, 2011; (*b*) J. Wang and L. Urban, Eds. *Predictive ADMET. Integrated Approaches in Drug Discovery and Development*, Wiley, Hoboken, 2014.

2 Atta-ur-Rahman and M. I. Choudhary, Eds. *Structure Activity Relationships Studies in Drug Development by NMR Spectroscopy*, Bentham Books, Amsterdam, 2011.

3 (*a*) J. G. Topliss, Ed. *Quantitative Structure–Activity Relationships of Drugs*, Medicinal Chemistry, Vol. **19**, Academic Press, New York, 1983; (*b*) C. Cativiela, J. I. García, J. Elguero, D. Mathieu and R. Phan Tan Luu, *Quant. Struct. Activ. Relat.*, 1987, **6**, 173–178; (*c*) C. Cativiela, J. Elguero, J. I. García and D. Mathieu "The Search of Heterocyclic Descriptors: Free-Wilson Models" in QSAR in Drug Design an Toxicology, Proc. of the 6th Europ. Sympos. on Quantit. Structure-Activity Relationships, Portoroz, Sept. 22–26, 1986. D. Hadži and B. Jerman-Blažič Eds., Elsevier Sci. Publ., Amsterdam 1987, XII, 376 pages; (*d*) H. Kubinyi, *QSAR: Hansch Analysis and Related Approaches*. Vol. **1**. VCH, Weinheim, 1993; (d) T. Puzyn, J. Leszzynski and M. T. Cronin, Eds. *Recent Advances in QSAR Studies, Theoretical and Computational Chemistry, Challenges and Advances in Computational Chemistry and Physics*, Vol. **8**, 2010.

4 (*a*) C. A. Lipinski, F. Lombardo, B. W. Donimy and P. J. Feeny, P.J. *Advanced Drug Delivery Reviews*, 1997, **23**, 3–25; (*b*) D. F. Veber, S. R. Johnson, H. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, *J. Med. Chem.*, 2002, **45**, 2615–2623.

5 (a) T. Steiner, *Angew. Chem. Int. Ed.*, 2002, **41**, 46–76; (*b*) B. Chan, J. E. Del Bene, J. Elguero and L. Radom, *J. Phys. Chem. A*, 2005, **109**, 5509–5517; (*c*) N. V. Belkova, L. M. Epstein and E. S. Shubina, *ARKIVOC*, 2008, **iv**, 120–138; (*d*) M. Chan-Huot, A. Dos, R. Zander, S. Sharif, P. M. Tolstoy, S. Compton, E. Fogle, M. D. Toney, I. Shenderovich, G. S. Denisov and H.-H. Limbach, *J. Am. Chem. Soc.*, 2013, **135**, 18160–18175.

6 (*a*) J. Elguero, "Pyrazoles and Benzo Derivatives" in Comprehensive Heterocyclic Chemistry, A. R. Katritzky and C. W. Rees, Eds., Pergamon, London, Vol. 5, 167–304, 1984; (*b*) J. Elguero, "Pyrazoles and Benzo Derivatives" in Comprehensive Heterocyclic Chemistry II, A. R. Katritzky, C. W. Rees and E. F. V. Scriven, Eds., Pergamon, London, Vol. 3, 1–76, 1996; (*c*) L. Yet, "Pyrazoles" in Comprehensive Heterocyclic Chemistry III, A. R. Katritzky, C. A. Ramsden, E. F. V. Scriven and R. J. K. Taylor, Eds., Elsevier, Cambridge, Vol. 4, 1–141, 2008; (*d*) J. Elguero, P. Goya, N. Jagerovic and A. M. S. Silva, "Pyrazoles as Drugs: Facts and Fantasies", Targets in Heterocyclic Systems, Italian Society of Chemistry, Roma, Vol. **6**, 52-98, 2002; (*e*) A. Chauhan, P. K. Sharma and N. Kaushik, *Int. J. ChemTech Res.*, 2011, **3**, 11–17; (*f*) V. Kumar, K. Kaur, G. K. Gupta and A. K. Sharma, *Eur. J. Med. Chem.*, 2013, **69**, 735–753; (*g*) R. Pérez-Fernández, P. Goya and J. Elguero, *ARKIVOC*, 2014, **ii**, 233–293; (*h*) S. Fustero, A. Simón-Fuentes, O. Delgado and R. Román, "Fluorinated Pyrazoles and Indazoles" in V. Nenajdenko, Ed., *Fluorine in Heterocyclic Chemistry*, Vol. 1: 5-Membered Heterocycles and Macrocycles, Springer International Publishing, Switzerland, 2014.

7 R. Jacquier and G. Maury, *Bull. Soc. Chim. Fr.*, 1967, 295–297.

8 J. Elguero, E. Gonzalez and R. Jacquier, *Bull. Soc. Chim. Fr.*, 1968, 707–713.

9 J. Elguero, E. Gonzalez and R. Jacquier, *Bull. Soc. Chim. Fr.*, 1968, 5009–5017.

10 J. Catalán, J. L. M. Abboud and J. Elguero, *Adv. Heterocycl. Chem.*, 1987, **41**, 187–274.

11 C. Foces-Foces, A. L. Llamas-Saiz, M. Menéndez, N. Jagerovic and J. Elguero, *J. Phys. Org. Chem.*, 1997, **10**, 637–645.

12 R. W. Taft, F. Anvia, M. Taagepera, J. Catalán and J. Elguero, *J. Am. Chem. Soc.*, 1986, **108**, 3237–3239.

13 J. Catalán, R. M. Claramunt, J. Elguero, J. Laynez, M. Menéndez, F. Anvia, J. H. Quian and R. W. Taft, *J. Am. Chem. Soc.*, 1988, **110**, 4105–4111.

14 J. Catalán, M. Sánchez-Cabezudo, J. L. G. de Paz, J. Elguero, R. W. Taft and F. Anvia, *J. Comput. Chem.*, 1989, **10**, 426–433.

15 J. Catalán, J. L. G. de Paz, M. Yáñez, R. M. Claramunt, C. López, J. Elguero, F. Anvia, J. H. Quian and R. W. Taft, *J. Am. Chem. Soc.*, 1990, **112**, 1303–1312.

16 J.-L. M. Abboud, R. Notario, M. Berthelot, R. M. Claramunt, P. Cabildo, J. Elguero, M. J. El Ghomari, W. Bouab, R. Mokhlisse and G. Guihéneuf, *J. Am. Chem. Soc.*, 1991, **113**, 7498–7493.

17 L. Z. Chen, R. Flammang, A. Maquestiau, R. W. Taft, J. Catalán, P. Cabildo, R. M. Claramunt and J. Elguero, *J. Org. Chem.*, 1991, **56**, 179–183.

18 J. Elguero, G. I. Yranzo, J. Laynez, M. Menéndez, J. Catalán, J. L. G. de Paz, F. Anvia and R. W. Taft, *J. Org. Chem.*, 1991, **56**, 3942–3947.

19 J.-L. M. Abboud, P. Cabildo, T. Cañada, J. Catalán, R. M. Claramunt, J. L. G. de Paz, J. Elguero, H. Homan, R. Notario, C. Toiron, G. I. Yranzo, *J. Org. Chem.*, 1992, **57**, 3938–3946.

20 R. Notario, M. Herreros, A. El Hammadi, H. Homan, J.-L. M. Abboud, I. Forfar, R. M. Claramunt and J. Elguero, *J. Phys. Org. Chem.*, 1994, **7**, 657–662.

21 J.-L. M. Abboud, C. Foces-Foces, M. Herreros, H. Homan, L. Infantes, R. Notario, A. W. Krebs, J. Neubauer, J. Elguero and N. Jagerovic, *J. Phys. Org. Chem.*, 1996, **9**, 79–86.

22 J.-L. M. Abboud, R. Notario, M. Yáñez, O. Mó, R. Flammang, N. Jagerovic, I. Alkorta and J. Elguero, *J. Phys. Org. Chem.*, 1999, **12**, 787–795.

23 J. Catalán and J. Elguero, *J. Chem. Soc., Perkin Trans. 2*, 1983, 1869–1874.

24 J. Catalán, O. Mó, J. L. G. De Paz, M. Yáñez and J. Elguero, *J. Org. Chem.*, 1984, **49**, 4379–4386.

25 A. El Hammadi, M.E. Mouhtadi, R. Notario, A. Werner and J. Elguero, *J. Chem. Soc., Perkin Trans. 2*, 1995, 379–383.

26 A. El Hammadi, M. El Mouhtadi, R. Notario, J.-L. M. Abboud and J. Elguero, *J. Chem. Res.*, 1995, *S*, 172–173; *M*, 1080–1096.

27 *(a)* R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, NY, 1989; *(b)* L. J. Bartolotti and K. Fluchichk, in *Reviews in Computational Chemistry*; K. B. Lipkowitz and D. B. Boyds, Eds.; VCH: New York, NY, 1996; Vol. 7, pp 187–216; *(c)* W. Kohn, A. D. Becke and R. G. Parr, *J. Phys. Chem.* 1996, 100,12974–12980; *(d)* T. Ziegler, *Chem. Rev.*, 1991, **91**, 651–667.

28 (*a*) R. Ditchfield, *Mol. Phys.*, 1974, **27**, 789-807. (*b*) M. J. Frisch, J. A. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265-3269.

29 Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, 2013.

30 (*a*) J. Tomasi and M. Persico, *Chem. Rev.* 1994, **94**, 2027–2094; (*b*) B. Y. Simkin and I. Sheikhet, *Quantum Chemical and Statistical Theory of Solutions: A Computational Approach*; Ellis Horwood: London, UK, 1995, pp. 78–101.

31 (*a*) S. Miertus, E. Scrocco and J. Tomasi, *J. Chem. Phys.*, 1981, **55**, 117–129; (*b*) R. Cammi and J. Tomasi, *J. Chem. Phys.*, 1994, **100**, 7495–7502; (c) V. Barone, M. Cossi and J. Tomasi, *J. Chem. Phys.*, 1997, **107**, 3210–3221. (*c*) J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999-3093. (*d*) F. Florsi and J. Tomasi *J. Comp. Chem.*, 1989, **10** 616-627. (*e*) F. Florsi, J. Tomasi and J. L. Pascual-Ahuir *J. Comp. Chem.*, 1991, **12**, 784-791.

32 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.

33 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.

34 J. Ho, *Aust. J. Chem.*, 2014, **67**, 1441–1460.

35 J. Ho and M. L. Coote, *Theor. Chem. Acc.*, 2010, **125**, 3–21 and references cited therein.

36 V. I. Minkin, A. D. Garnovskii, J. Elguero, A. R. Katritzky and O. V. Denisko, *The Tautomerism of Heterocycles: Five-Membered Rings with Two or More Heteroatoms*, *Adv. Heterocycl. Chem*. 2000, **76**, 157–323.

37 M. Bruix, J. de Mendoza, R. M. Claramunt and J. Elguero, *Magn. Reson. Chem.*, 1985, **23**, 367–374.

38 R. M. Claramunt, D. Sanz, J. Catalán, F. Fabero, N. A. García, C. Foces-Foces, A. L. Llamas-Saíz and J. Elguero, *J. Chem. Soc., Perkin Trans. 2*, 1993, 1687–1699

39 (*a*) P. J. Linstrom and W. G. Mallard, Eds., *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, National Institute of Standards and Technology, Gaithersburg MD, 20899, http://webbook.nist.gov, (retrieved November 6, 2014); (*b*) Most values came from E. P. Hunter and S. G. Lias, *J. Phys. Chem. Ref. Data*, 1998, **27**, 413–656.

40 (*a*) D. M. Bishop and K. J. Laidler, *J. Chem. Phys.*, 1965, **42**, 1688–1691 (*b*) J. Catalán and J. Elguero, *J. Chem. Soc., Perkin Trans. 2*, 1983, 1869–1874; (*c*) C. Laurence and J.-F. Gal, Lewis Basicity and Affinity Scales: Data and Measurement, John Wiley,& Sons, Chichester, 2010.

41 (*a*) H. A. Staab, H. Bauer and K. M. Schneider, *Azolides in Organic Synthesis and Biochemistry*, Wiley-VCH, Weinheim, 1998; (*b*) J. Elguero, C. Foces-Foces, D. Sanz and R. M. Claramunt, *Azolides: Structural Aspects*, in *Advances in Nitrogen Heterocycles*, JAI Press, Stamford, 2000, **46**, 295–367.

42 (*a*) Y. Zhangh and A. J. Kennan, *Org. Lett.*, 2001, **3**, 2341–2344; (*b*) J. A. Castillo-Meléndez and B. T. Golding, *Synthesis*, 2004, 1655–1663; (*c*) *Superbases for Organic Synthesis*, T. Ishikawa, Ed., John Wiley & Sons, Chichester 2009. Chapter 4, T. Ishikawa, *Guanidines in Organic Synthesis*, where compound **150** is discussed (pp. 95–97).

43 Y. Takano and K. N. Houk, *J. Chem. Theor. Comput.*, 2005, **1**, 70–77.

44 B. Kallies and R. Mitzner, *J. Phys. Chem. B,* 1997, **101**, 2959–2967.

45 S. Hwang, Y. H. Jang and D. S. Chung, *Bull. Korean Chem. Soc.*, 2005, **26**, 585–588.

46 N. Sadlej-Sosnowska, *Theor. Chem. Acc.*, 2007, **118**, 281–293.

47 T. B. Lee and M. L. McKee, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10258–10269.

48 P. O. Hubin, A. D. Laurent, D. P. Vercauteren and D. Jacquemin, *Phys. Chem. Chem. Phys.*, 2014, **16**, 25288–25295.