Principal properties (PPs) as solvent descriptors for multivariate optimization in organic synthesis: specific PPs for ethers

Francesco P. Ballistreri, Cosimo G. Fortuna, Giuseppe Musumarra,* Didier Pavone, and Salvatore Scirè

Dipartimento di Scienze Chimiche, Università di Catania,
Viale Andrea Doria 6, 95125 Catania, Italy
E-mail: gmusumarra@dipchi.unict.it

Dedicated to Professor Domenico Spinelli on his 70th birthday
(received 07 Oct 02; accepted 21 Nov 02; published on the web 29 Nov 02)

Abstract
Principal properties (PPs) for solvents were extended to 113 solvents by the addition of ten ethers. Specific "PPs" for ethers, suitable for solvent optimization in Grignard reactions, were also derived and their physico-chemical significance discussed.

Keywords: Solvent descriptors, principal properties, Grignard reaction

Introduction

The application of multivariate strategies provides great advantages in optimizing desired properties (yield, regio- or stereoisomeric ratios, etc.) by simultaneous variation of experimental conditions—which may be continuous (temperature, time, concentrations, etc.) and discrete (solvent, catalyst, etc.) variables: the latter are parameterized by the so-called principal properties (PPs). Intrinsic properties suitable for experimental design need to be orthogonal to each other. In order to fulfil the above requirement, the concept of principal properties (PPs), *i.e.*, of intrinsic properties representative of experimentally observable macroscopic descriptors, has recently been introduced.[1]

PPs, calculated as the principal component analysis (PCA) scores from several experimentally measured properties, are now available for solvents[1], aldehydes and ketones,[1] amines,[1] Lewis acids,[1] aromatic substituents[2] and amino acids.[3–5] First and second generation principal properties for heteroaromatics moieties, based on aromaticity[6] and on 3D-GRID structural parameters[7] respectively, have been reported by our group.

A typical example of the potentialities of optimization in PPs is provided by the results of a multivariate experimental design, based on the PPs of Lewis acid catalysts and solvents, which

give a better understanding of the effects of the above parameters on the isomeric distribution in the reaction of phenylhydrazones derived from unsymmetrical ketones (the well known Fischer indole synthesis). Multivariate optimization achieved not only regiospecific synthesis of single indole regioisomers,[8] but almost quantitative yields in a single step (one-pot) reaction.[9] It is striking that the above result was achieved for a reaction, such as the Fischer indole synthesis[10] known for more than a century, and to which an entire book has been devoted.[11]

In addition to statistical orthogonality, PPs derived from PCA multivariate characterizations have the advantage of being less influenced by measurement errors and system-specific variations than are single descriptors. Moreover, reliable PPs can be obtained from original data matrices even with missing data.

However, as PPs are derived by an empirical statistical method, such as PCA, carried out on a data matrix with a given number of objects and variables, the updating of such descriptors is needed as new properties (variables) become available for "old" and "new" objects. In particular, solvent PPs first derived by Carlson and co-workers based on eight parameters for 82 solvents[12] have been extended by the same author[1] to 103 solvents and integrated by addition of a 9th variable (water-solubility).

For specific organic syntheses such as Grignard reactions, however, the use of solvent is chemically limited to a specific class, *i.e*., ethers.

In this context, we extend the solvent set by addition of ten solvents, all ethers, to the 103 considered by Carlson[1] in order to derive PPs from a larger data set containing 113 solvents.

Furthermore, we also consider an "ether" class model in order to derive "ether PPs" which might be more informative in characterizing the above solvents specifically for the Grignard reaction.

Results and Discussion

Table 1 reports the data set considered to derive PPs in ref. 1, integrated by corrected or new descriptors for the original 103 solvents and expanded by including descriptors for ten new solvents (all ethers). PCA of the above data matrix, including 113 solvents and 9 descriptors (cf. Table 1) provided a 2- principal-components (PC) model explaining 69.4% of the variance. The scores of such a model, new PPs for all 113 solvents, are also recorded in Table 1.

Table 1. Descriptors (1–9) and principal properties (t_1 and t_2) for solvents 1–113[a]

Entry	CAS[b]	Solvent	Descriptors									PPs	
			1	2	3	4	5	6	7	8	9	t_1	t_2
1	7732-18-5	Water	0	100	78.3	5.9	1.333	1	0.9282	-1.38	1.774	3.67	-1.78
2	75-12-7	Formamide	2.55	210.5	111	11.2	1.4475	0.8	1.159	-1.51	1.401	4.56	0.67
3	107-21-1	1,2-Ethanediol	-12.6	197.15	37.7	7.7	1.4318	0.79	1.1088	-1.36	1.06	2.93	0.16
4	67-56-1	Methanol	-97.7	64.5	32.66	5.5	1.3284	0.76	0.7914	-0.77	1.393	1.63	-3.04
5	123-39-7	N-Methylformamide	-3.8	182.5	182.4	12.9	1.4319	0.72	1.01	**-0.6**		5.84	-0.13
6	111-46-6	Diethylene glycol	-7.8	245.7	31.69	7.7	1.4475	0.71	1.109	**-1.51**		3.06	0.65
7	112-27-6	Triethylene glycol	-4.3	288	23.69	10	1.4558	0.7	1.125	-1.33	0.852	3.07	1.21
8	109-86-4	2-Methoxyethanol	-85.1	124.6	16.93	6.8	1.4021	0.67	0.965	-0.77	1.103	1.54	-1.45
9	79-16-3	N-Methylacetamide	30.6	206.7	191.3	14.2	1.4253	0.67	0.957	-1.05		6.48	0.15
10	64-17-5	Ethanol	**-114.5**	78.3	24.55	5.8	1.3614	0.65	0.785	-0.31	1.231	1.11	-2.71
11	141-43-5	2-Aminoethanol	10.5	170.95	37.72	7.6	1.4545	0.65	1.018	-1.31	1.216	2.66	0.27
12	64-19-7	Acetic acid	16.7	117.9	6.17	5.6	1.3719	0.65	1.0492	-0.17	1.002	1.56	-0.54
13	100-51-6	Benzyl alcohol	-15.3	205.45	13.1	5.5	1.5404	0.61	1.042	1.1	-0.992	0.71	1.67
14	71-23-8	1-Propanol	-126.2	97.15	20.45	5.5	1.3856	0.62	0.804	0.25	1.126	0.78	-2.35
15	71-36-3	1-Butanol	-88.6	117.7	17.51	5.8	1.3993	0.6	0.8098	0.88	0.125	0.52	-1.47
16	78-83-1	2-Methyl-1-propanol	-108	107.9	17.93	6	1.3959	0.55	0.794	0.76	0.176	0.36	-1.78
17	67-63-0	2-Propanol	-88	82.2	19.92	5.5	1.3772	0.55	0.786	0.05	1.117	0.81	-2.19
18	78-92-2	2-Butanol	-114.7	99.5	16.56	5.5	1.3971	0.51	0.808	0.61	0.176	0.19	-1.87
19	123-51-3	3-Methyl-1-butanol	-117.2	130.5	15.19	6.1	1.4072	0.57	0.8092	**1.22**	-0.518	0.07	-1.42
20	108-93-0	Cyclohexanol	25.25	161.1	15	6.2	1.4648	0.5	0.9642	1.23	-0.755	0.68	1.01
21	4427-89-8	4-Methyl-1,3-dioxol-2-one	-54.5	241.7	64.94	16.5	1.4215	0.49	1.204	**0.86**		3.14	0.76
22	6032-29-7	2-Pentanol	**-73**	119	13.71	5.5	1.4064	0.49	0.81	1.19	-0.612	-0.01	-1.05
23	75-52-5	Nitromethane	-28.55	101.2	35.94	11.9	1.3819	0.48	1.137	-0.35	0.235	1.80	-0.52
24	75-05-8	Acetonitrile	-43.8	81.6	35.94	11.8	1.3441	0.46	0.7857	-0.34	1.282	1.85	-1.94
25	584-02-1	3-Pentanol	-75	115.3	13.35	5.5	1.4104	**0.46**	0.821	1.21	-0.234	0.02	-1.08
26	67-68-5	Dimethyl sulfoxide	18.5	189	46.45	13.5	1.4793	0.44	1.101	-1.35		3.19	1.06
27	62-53-3	Aniline	-6	184.4	6.71	5	1.5863	0.42	1.0217	0.9	-0.37	0.39	1.91
28	126-33-0	Sulfolane	28.45	287.3	43.3	16	1.4816	0.41	1.262	-0.77		3.74	2.31
29	108-24-7	Acetic anhydride	-73.1	140	20.7	9.4	1.3904	0.41	1.082	**-0.48**	0.07	1.09	-0.74
30	75-65-0	2-Methyl-2-propanol	25.6	82.3	12.47	5.5	1.3877	0.39	0.789	0.35	-1.057	0.19	-0.55
31	68-12-2	N,N-dimethylformamide	-60.4	153	36.71	10.8	1.4305	0.4	0.945	-1.01	1.112	1.94	-0.65
32	127-19-5	N,N-dimethylacetamide	-20	166.1	37.78	12.4	1.4384	**0.38**	0.937	-0.77	1.11	2.23	-0.04
33	107-12-0	Cyanopropane	-92.8	97.35	28.86	11.7	1.3658	0.4	0.782	0.16		0.87	-1.96
34	872-50-4	1-Methyl-pyrrolidine-2-one	-24.4	202	32.2	13.6	1.47	0.36	1.026	**-0.39**		2.17	0.78
35	67-64-1	Acetone	-94.7	56.1	20.56	9	1.3587	0.36	0.79	-0.24	1.134	0.74	-2.46
36	98-95-3	Nitrobenzene	5.8	210.8	34.78	13.3	1.5562	0.32	1.204	1.85	-1.778	0.92	2.88
37	100-47-0	Benzonitrile	-12.75	191.1	25.2	13.4	1.5282	0.33	1.01	1.56	-1.013	0.93	1.79

Table 1. Continued

Entry	CAS[b]	Solvent	Descriptors									PPs	
			1	2	3	4	5	6	7	8	9	t_1	t_2
38	107-15-3	1,2-Diaminoethane	11.3	116.9	12.9	6.3	1.4568	0.35	0.899	-2.04		1.22	-0.14
39	107-06-2	1,2-Dichloroethane	-35.7	83.5	10.7	6.1	1.4448	**0.33**	1.235	1.48	-1.092	-0.38	0.40
40	75-85-4	2-Methyl-2-butanol	-8.8	102	5.78	5.7	1.405	0.32	0.806	0.89	0.039	0.09	-0.64
41	78-93-3	2-Butanone	-86.7	79.6	18.51	9.2	1.3788	0.33	0.805	0.29	0.494	0.46	-1.84
42	98-86-2	Acetophenone	19.6	202	17.39	9.8	1.5342	0.31	1.0281	1.58	-1.553	0.44	2.28
43	75-09-2	Dichloromethane	-94.9	39.6	8.93	5.2	1.4242	0.31	1.33	1.25	-0.699	-0.76	-0.57
44	632-22-4	1,1,3,3-Tetramethylurea	-1.2	175.2	23.6	11.7	1.4493	0.32	0.969	0.19		1.50	0.66
45	680-31-9	Hexamethylphosphoramide	7.2	233	29.6	18.5	1.4588	0.32	1.024	0.28		2.76	1.38
46	108-94-1	Cyclohexanone	-32.1	155.65	16.1	10.3	1.451	0.28	0.9478	0.81	-1.236	0.32	0.47
47	110-86-1	Pyridine	-41.55	115.25	12.91	7.9	1.5102	0.3	0.982	0.65	1.094	0.64	0.22
48	79-20-9	Methyl acetate	-98.05	56.9	6.68	5.7	1.3614	**0.25**	0.933	0.18	0.486	-0.25	-2.07
49	108-10-1	4-Methyl-2-pentanone	-84.7	117.4	13.11	2.7	1.3958	0.27	0.7978	1.31	-0.871	-0.90	-1.19
50	75-34-3	1,1-Dichloroethane	-97	57.3	10	6.1	1.4164	0.27	1.176	1.79	-1.259	-1.03	-0.63
51	91-22-5	Quinoline	-14.85	237.1	8.95	7.3	1.6273	0.27	1.093	2.03		0.09	3.21
52	96-22-0	3-Pentanone	-39	102	17	9.4	1.3923	0.27	0.8138	0.91	-0.27	0.25	-0.86
53	67-66-3	Chloroform	-63.15	61.2	4.81	3.8	1.4459	0.26	1.48	1.97	-1.246	-1.14	0.52
54	112-49-2	Triglyme	45	216	7.5		1.4264	0.25	**0.99**	**1.15**		0.99	0.15
55	111-96-6	Diglyme	-64	162	**7.23**	6.6	1.4078	0.24	**0.945**	**-0.80**		0.23	-0.59
56	110-71-4	Glyme	-69	85	7.2	5.7	1.3796	0.23	0.8269	-0.21		-0.49	-1.53
57	95-50-1	1,2-Dichlorobenzene	-17	180.5	9.93	7.1	1.5515	0.23	1.305	3.38	-3.029	-1.18	2.90
58	141-78-6	Ethyl acetate	-83.55	77.1	6.02	6.1	1.3724	0.23	0.9	0.73	-0.041	-0.44	-1.59
59	462-06-6	Fluorobenzene	-42.2	84.7	5.42	4.9	1.4684	0.19	1.023	2.27	-1.797	-1.35	0.39
60	591-50-4	Iodobenzene	-31.35	188.3	4.49	4.7	1.62	0.17	1.831	3.25	-2.847	-1.43	4.32
61	108-90-7	Chlorobenzene	-45.6	131.7	5.62	5.4	1.5248	0.19	1.106	2.84	-2.45	-1.51	1.51
62	108-86-1	Bromobenzene	**-30.6**	155.9	5.4	5.2	1.559	0.18	1.495	2.99	-2.57	-1.37	2.88
63	109-99-9	Tetrahydrofuran	-108.4	66	7.58	5.8	1.4072	0.21	0.889	0.46	1.142	-0.24	-1.85
64	100-66-3	Methoxybenzene	-37.5	153.6	4.33	4.2	1.517	0.2	0.996	2.11	-1.917	-1.19	1.24
65	103-73-1	Ethoxybenzene	-29.5	169.8	4.22	4.5	1.5074	0.18	0.967	2.51	-2.332	-1.33	1.43
66	71-55-6	1,1,1-Trichloroethane	-30.4	74.1	7.25	5.7	1.438	0.17	1.339	2.49	-2.179	-1.35	0.89
67	123-91-1	1,4-Dioxane	11.8	101.3	2.21	1.5	1.4224	0.16	1.034	-0.27	1.07	0.09	-0.21
68	79-01-6	Trichloroethene	-86.4	87.2	3.42	2.7	1.4773	0.16	1.464	2.42	-1.995	-1.84	0.91
69	110-89-4	Piperidine	-10.5	106.7	5.8	4	1.4525	0.15	0.861	0.85		-0.85	0.16
70	101-84-8	Diphenyl ether	26.9	258.1	3.69	3.9	1.5763	0.14	1.075	4.21	-3.947	-1.81	3.85
71	60-29-7	Diethyl ether	-116.3	34.4	4.2	3.8	1.3524	0.12	0.714	0.89	0.076	-1.27	-2.71
72	71-43-2	Benzene	5.5	80.1	2.27	0	1.5011	0.11	0.8786	2.13	-1.559	-1.76	0.69
73	108-20-3	Di-isopropyl ether	-85.5	68.3	3.88	4.2	1.3681	0.1	0.7241	1.52	-0.198	-1.24	-1.89
74	108-88-3	Toluene	-95	110.6	2.38	1	1.4969	0.1	0.867	2.73	-1.747	-2.28	0.01

Table 1. Continued

Entry	CAS[b]	Solvent	Descriptors									PPs	
			1	2	3	4	5	6	7	8	9	t_1	t_2
75	142-96-1	Dibutyl ether	-95.2	140.3	3.08	3.9	1.3992	0.07	0.7689	3.21	-1.941	-2.10	-0.64
76	121-44-8	Triethylamine	-114.7	88.9	2.42	2.9	1.401	0.04	0.7275	1.45	0.224	-1.44	-1.83
77	108-67-8	1,3,5-Trimethylbenzene	-44.7	164.7	2.28	0	1.4994	0.07	0.865	3.42	-3.356	-2.68	1.22
78	75-15-0	Carbon disulfide	-111.6	46.2	2.64	0	1.6225	0.07	1.263	1.94	-1.552	-2.43	1.09
79	56-23-5	Carbon tetrachloride	-22.8	76.6	2.23	0	1.4602	0.05	1.59	2.83	-2.294	-2.22	1.63
80	127-18-4	Tetrachloroethene	-22.4	121.2	2.3	0	1.5057	**0.04**	1.623	3.4	-2.895	-2.42	2.56
81	110-82-7	Cyclohexane	6.7	80.7	2.02	0	1.4262	0.01	0.778	3.44	-3.069	-2.76	0.34
82	110-54-3	Hexane	-95.3	68.7	1.88	0	1.3749	0.01	0.66	3.98	-3.359	-3.53	-1.26
83	112-60-7	Tetraethylene glycol	-6.2	327.3	19.7	10.8	1.4577	0.66	1.1285	**-2.23**		3.74	1.29
84	71-41-0	1-Pentanol	-78.2	138	13.9	5.7	1.41	**0.59**	0.8144	1.56	-0.631	0.15	-0.92
85	80-73-9	1,3-Dimethylethyleneurea	8.2	225.5	37.6	13.6	1.4707	0.36	**1.052**	**-0.02**		2.49	1.36
86	7226-23-5	1,3-Dimethylpropyleneurea	-20	230	36.12	14.1	1.4881	0.35		**0.36**		2.18	1.54
87	107-87-9	2-Pentanone	-76.9	102.3	15.38	9	1.3908	0.32	0.8089	0.91	-0.328	0.10	-1.27
88	110-91-8	Morpholine	-4.8	128.9	7.42	5.2	1.4542	0.31	1.0005	-0.86	1.06	0.95	-0.04
89	563-80-4	3-Methyl-2-butanone	-92	94.2	15.87	9.2	1.388	0.32	0.8051	0.56	-0.164	0.15	-1.57
90	75-97-8	3,3-Dimethyl-2-butanone	-49.8	106.3	13.1	9.3	1.3952	0.26	0.8012	**1.07**	-0.706	-0.05	-0.84
91	565-80-0	2,4-Dimethyl-3-pentanone	-69	125.25	17.2	9.1	1.3999	0.25	0.8108	1.49	-1.299	-0.34	-0.69
92	108-83-8	2,6-dimethyl-4-heptanone	-46	168.2	9.91	8.9	1.4122	0.23	0.8053	2.96	-1.731	-0.75	0.18
93	112-36-7	Diethyleneglycol diethyl ether	-44.3	188.9	5.7		1.4115	0.21	0.9063	**0.26**	0.076	0.34	-0.16
94	105-58-8	Diethyl carbonate	-43	126.8	2.82	3	1.3837	**0.19**	0.9752	1.21	-0.798	-0.84	-0.49
95	75-35-4	1,1-Dichloroethene	-122	31.6	4.82	4.3	1.4247	0.19	1.218	1.86	-1.666	-1.77	-0.83
96	1634-04-4	*tert*-Butyl methyl ether	-108.6	55.2	4.5	4.1	1.369	**0.12**	0.7405	0.94		-1.81	-2.09
97	109-89-7	Diethylamine	-49.8	55.55	3.78	4	1.3846	0.15	0.7056	0.58		-1.31	-1.53
98	111-43-3	Dipropyl ether	-123.2	90.1	3.39	4.4	1.3805	0.1	0.736	2.03	-1.388	-1.83	-1.72
99	106-42-3	1,4-Dimethylbenzene	13.3	138.4	2.27	0	1.4958	0.07	0.8611	3.15	-2.806	-2.24	1.44
100	102-82-9	Tributylamine	-70	214		2.6	1.4291	0.04	0.7771	4.56	-3.116	-2.84	0.72
101	493-01-6	*cis*-Decalin	-43	195.8	2.2	0	1.481	0.02	0.8965	5.08	-5.192	-3.64	1.89
102	142-82-5	*n*-Heptane	-90.6	98.4	1.92	0	1.3876	0.01	0.6838	4.57	-4.046	-3.75	-0.67
103	109-66-0	*n*-Pentane	-129.7	36.1	1.84	0	1.3575	0.01	0.6262	3.39	-3.129	-3.59	-2.14
104	142-68-7	Tetrahydropyran	**-49.2**	**88**	**5.66**	**5.8**	**1.4211**	**0.17**	**0.8814**	**0.82**	**-0.031**	-0.44	-0.77
105	693-65-2	Dipentyl ether	**-69.43**	**186.75**	**2.8**		**1.412**		**0.7833**	**4.17**		-2.23	0.26
106	544-01-4	Di-isopentyl ether		**172**	**2.82**		**1.408**		**0.775**	**3.8**		-2.12	0.05
107	93-18-5	2-Ethoxynaphthalene	**37.5**	**280**			**1.5932**		**1.064**	**3.896**		0.51	3.95
108	994-05-8	tert-Amyl methyl ether	**-80**	**86.3**			**1.3885**	**0.11**	**0.7703**	**1.678**		-2.01	-1.22
109	637-92-3	tert-Butyl ethyl ether	**-94**	**70**			**1.376**	**0.12**	**0.7404**	**1.678**		-1.94	-1.67
110	91-16-7	1,2-Dimethoxybenzene	**22.5**	**206.5**				**0.24**	**1.084**	**1.955**		0.54	2.41
111	103-50-4	Dibenzyl ether	**3.6**	**296**	**3.82**		**1.5406**	**0.14**	**1.0014**	**3.948**		-0.81	3.27

Table 1. Continued

Entry	CAS[b]	Solvent	Descriptors									PPs	
			1	2	3	4	5	6	7	8	9	t_1	t_2
112	109-87-5	Dimethoxymethane	**-105**	**43**	**2.64**		**1.3541**	**0.16**		**-0.265**		-1.26	-2.58
113	462-95-3	Diethoxymethane	**-66.5**	**88**	**2.53**			**0.1**		**0.797**		-1.39	-0.82

(a) Values different from those in ref. 1 are in red, newly inserted data in blue. Descriptors are those reported in ref. 1: 1, melting point (°C); 2, boiling point (°C); 3, dielectric constant; 4, dipole moment $\times 10^{30}$ (C m), 5, refractive index; 6, E_T, the normalized Reichardt–Dimroth parameter (kcal/mol); 7, density $\times 10^3$ (kg/m^3); 8, lipophilicity as measured by logP where P is the equilibrium constant of the distribution of the solvent between 1-octanol and water at 25°C; 9, water solubility (log mol/L).
(b) Chemical Abstracts Registry Number

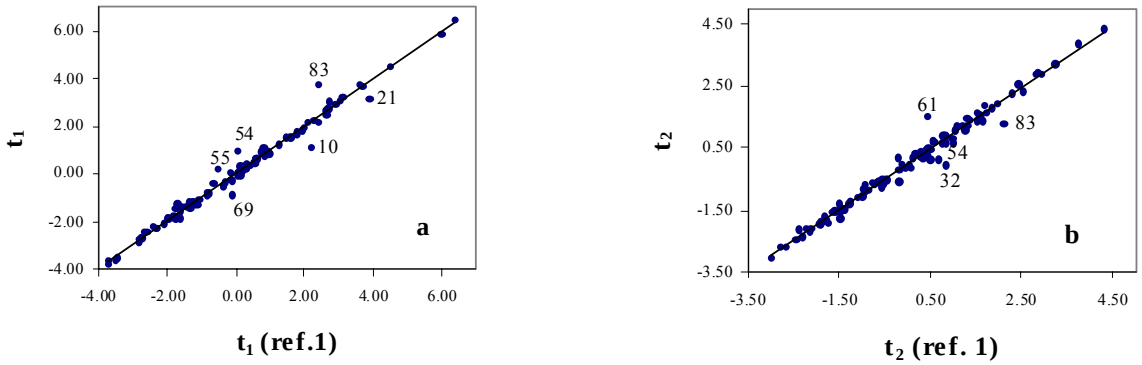

Figure 1. Correlation between PPs for solvents 1–103 derived in the present work and those in ref. 1.

Figure 1 shows that, as expected, the PPs for solvents 1–103 derived from the 113 solvents model, with a few exceptions (differences above 0.5 are discussed below), closely resemble those of the 103 solvents model reported in ref. 1. Differences in both t_1 and t_2 found for triglyme (54) and tetraethylene glycol (83) can be ascribed to the addition of new values for 54 and correction of logP for 83. Significant differences in t_1 for ethanol (10) and in t_2 for N,N-dimethylacetamide (32) are due to corrections of descriptor values, while those in t_1 for 4-methyl-1,3-dioxol-2-one (21) and diglyme (55) are the result of the insertion of new values. Differences in t_2 for chlorobenzene (61) and in t_1 for piperidine (69) are probably due to printing errors in Table 15A.1 of ref. 1, as the plot in the same book and PPs in ref. 12 are consistent with our values.

Figure 2, the PCA "p_1–p_2 loadings plot" (see equation 1 in the Experimental Section), elucidates the descriptors information content, and provides guidance for interpreting the physico-chemical meaning of solvents PPs. The first PC, exhibiting high p_1 values for

descriptors such as dielectric constant (3), dipole moment (4), E_T (6), and water-solubility (9), and the lowest p_1 value for logP (8), can be related to the solvent polarity.

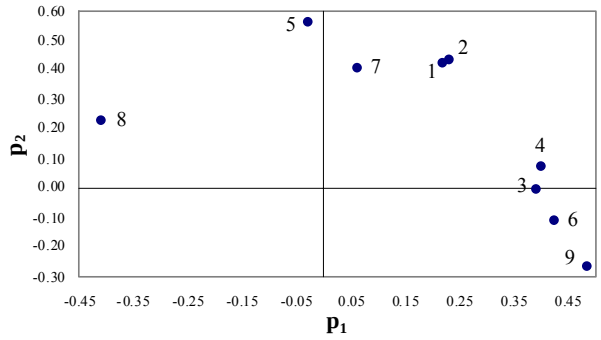

Figure 2. Loadings plot for the 113 solvents model.

Table 2. Principal properties (t_1, t_2 and t_3) for the 24 ethers model

Entry	Solvents	PPs		
		t_1	t_2	t_3
54	Triglyme	1.20	-2.82	0.33
55	Diglyme	0.39	-2.62	0.34
56	Glyme	-0.73	-2.32	0.45
63	Tetrahydrofuran	-0.83	-2.05	0.90
64	Methoxybenzene	1.54	-0.09	-0.16
65	Ethoxybenzene	1.50	0.30	-0.03
67	1,4-Dioxane	0.62	-0.17	-2.42
70	Diphenyl ether	3.43	1.67	-0.02
71	Diethyl ether	-2.64	0.11	0.08
73	Di-isopropyl ether	-2.01	0.70	0.02
75	Dibutyl ether	-1.28	1.97	0.54
93	Diethyleneglycol diethyl ether	0.52	-1.30	0.14
96	*Tert*-Butyl methyl ether	-2.18	0.01	0.18
98	Dipropyl ether	-2.07	1.03	0.39
104	Tetrahydropyran	-0.29	-0.81	-0.06
105	Dipentyl ether	-0.05	2.36	0.70
106	Diisopentyl ether	-0.16	2.17	0.54
107	2-Ethoxynaphthalene	4.15	1.16	0.06
108	Tert-Amyl methyl ether	-1.45	0.75	-0.16
109	Tert-Butyl ethyl ether	-1.83	0.57	0.08
110	1,2 Dimethoxybenzene	3.06	-0.90	-0.29
111	Dibenzyl ether	2.89	1.63	0.46
112	Dimethoxymethane	-2.39	-0.24	-1.04
113	Diethoxymethane	-1.64	0.92	-0.92

Accordingly, very high t_1 values are exhibited in Table 1 by amides (9, 5, and 2) and water (1), and very low t_1 values by hydrocarbons (82, 101–103). Figure 2 also shows high p_2 values for m.p. (1), b.p. (2), refractive index (5), and density (7) descriptors, properties affected by the molecular weight within each class of chemically different solvents. Accordingly, t_2 values in Table 1 increase on increasing the molecular weight of hydrocarbons (*e.g.*, 103, 82, 102), of primary alcohols (*e.g.*, 4, 10, 14, 15) and of ethers (*e.g.*, 71, 98, 75, 105).

The PPs in Table 1 can be conveniently adopted as solvent descriptors in multivariate optimization of reactions in which a wide range of solvents may be used. However, severe limitations in solvent selection may occur owing to chemical reasons, *e.g.*, the Grignard reaction can be carried out only in ethers. In this case, where a small portion of the experimental space can be investigated, it appears appropriate to derive PPs from a specific "class" model including only chemically similar solvents. Therefore PCA was carried out on a data matrix including 24 objects (ethers 54–56, 63–65, 67, 70, 71, 73, 75, 93, 96, 98, 104–113) and seven variables (descriptors 1–3, 5–8). Exclusion of descriptors 4 and 9 from the analysis is dictated by the lack of the above data for many solvents (see Table 1). PCA provided a 3-PC model explaining 95.1% of variance (56.5 first PC, 31.6% second PC and 7.0% third PC). Ether PPs (t_1, t_2 and t_3) derived from this model are reported in Table 2 and plotted in Figure 3.

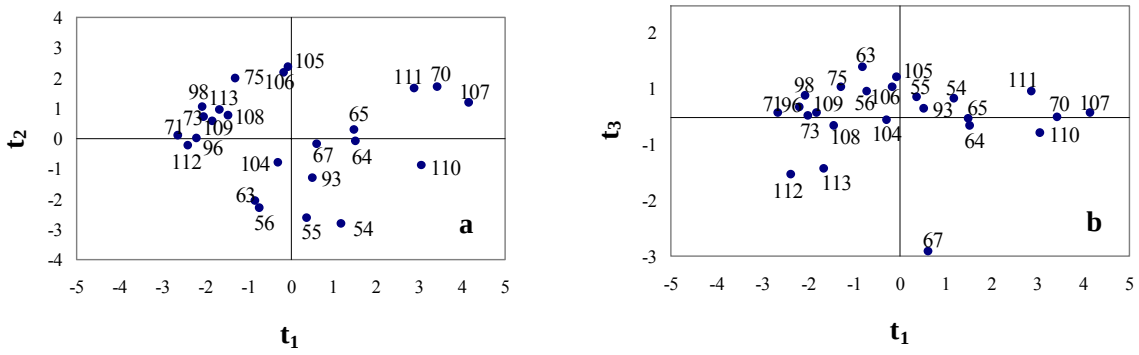

Figure 3. Scores Plots t_1–t_2 (a) and t_1–t_3 (b) for the 24-ethers model.

The first PP (t_1 in Table 2), exhibiting an excellent correlation (R^2= 0.93) with the second PP derived from the general model (t_2 in Table 1), can be related to the ether molecular weight, *i.e.*, with the lowest value for diethyl ether (71) and very high values for 2-ethoxynaphthalene (107) and diphenyl ether (70). This correlation is not surprising, as by restricting the model to a class of chemically similar solvents, it is expected that molecular weight becomes the first systematic variation evidenced by PCA. The interpretation of the 2nd and 3rd PPs is not straightforward. The loadings plot p_1–p_2 shows (Figure 4a) grouping of descriptors 1, 2, 5 and 7 and a clear differentiation of descriptors 3 and 6 (low p_2) from 8 (high p_2) resembling the trend already observed in the general model (Figure 2), where variable 8 was discriminated from 3 and 6 by the first component. However, the correlation between t_2 for 24 ethers in the ethers model (Table

2) and t_1 in the overall model (Table 1) is very poor (R^2 = 0.53). This can be reasonably explained by considering that the first component in the general model accounted for, "interclass solvent polarity", which can be roughly represented by the dielectric constant, a bulk property measuring non-specific solvation effects opposite to those of logP as, in general, highly polar solvents are not very lipophilic. Accordingly, high t_2 values are exhibited in Figure 3a by symmetrical ethers with a C_4- or C_5- chain (75,105,106) and low t_2 values by glymes (54–56).

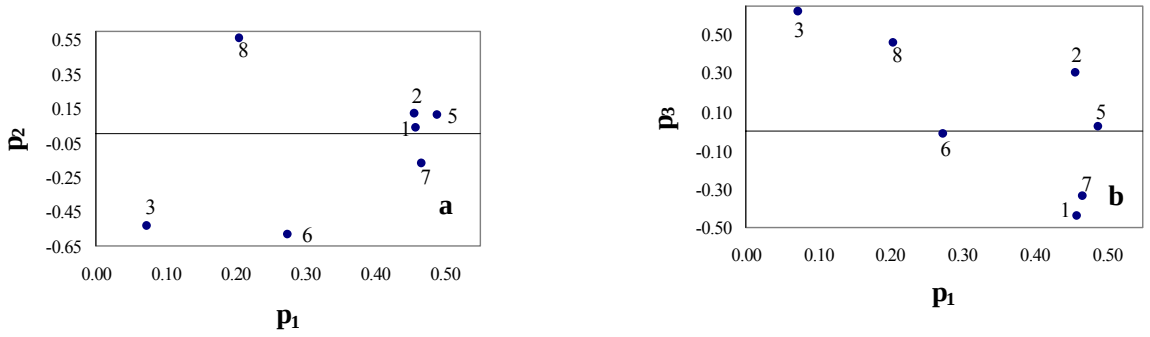

Figure 4. Loadings plot p_1–p_2 (a) and p_1–p_3 (b) for the 24 ethers models.

The 3rd PC is required to explain descriptor 3, whose information content, rather than being similar to that of descriptor 6 (see Figure 4b) is closer to that of descriptor 8, and to differentiate descriptor 2 (boiling point) from descriptor 1 (melting point). In fact, by considering a more homogeneous class, specific solvation effects such as hydrogen bonding and dipole–dipole interactions may be evidenced. Accordingly in Fig. 3b the 3rd PC is required to differentiate solvents with very low t_3 values such as dioxane (67), dimethoxymethane (112) and diethoxyethane (113), having both low ε and logP values (*i.e.*, high water-affinity in spite of their low "bulk" polarity).

Conclusions

Solvent PPs (t_1 and t_2) were extended to 113 solvents and specific PPs for ethers (t_1, t_2 and t_3), suitable for solvent optimization in the Grignard reaction, also derived. The score in the ether model t_1, related to the molecular weight, exhibits an excellent correlation with t_2 in the overall model, while t_2 and t_3 account for specific solvent effects.

Experimental Section

General Procedures. The data set used for PCA[13] was a table (matrix) in which 113 solvents were characterized by nine physico-chemical properties. The variables have been autoscaled by multiplying the variables by appropriate weights (the reciprocal of the variable standard deviation) to give them unit variance (*i.e*., the same importance). PCA was carried out by using the SIMCA software package[14] on a data matrix containing x_{ik} elements (113 x 9 for the overall model and 24 x 7 for the ethers model, respectively), where the index *k* is used for the physico-chemical properties (variables) and index *i* for the solvents (objects). Autoscaled matrix elements were then fitted into a model given by Equation (1), where the number A of significant cross terms (components), and the parameters p_{ak} and t_{ia} are calculated by minimizing the residuals, e_{ik}, after subtracting x_k (the mean value of the i^{th} experimental quantities x_k).

$$x_{ik} = \bar{x}_k + \sum_{a=1}^{a=A} t_{ia} p_{ak} + e_{ik} \qquad (1)$$

Parameters $\bar{x}_k$ and p_{ak} (the loadings) depend only on the physico-chemical properties (variables), and the t_{ia} (scores) only on the solvents.
The deviations from the model are expressed by the residuals, e_{ik}. The number of significant components (A) was determined using the cross-validation technique.[15]

Acknowledgments

We thank the University of Catania for financial support of this work

References

1. Carlson, R. *Design and Optimization in Organic Synthesis;* Elsevier: Amsterdam, 1992.
2. Skagerberg, M.; Bonelli, D.; Clementi, S.; Cruciani, G.; Ebert, C. *Quant. Struct.-Act. Relat*. **1989**, *8*, 32.
3. Hellberg, S.; Sjöström, M.; Wold, S. *Acta Chem. Scand*. **1986**, *B40*, 135.
4. Hellberg, S.; Sjöström, M.; Wold, S. *J. Med. Chem.* **1987**, *30*, 1127.
5. Skagerberg, M.; Sjöström, M.; Wold, S. J. *Chemometrics* **1990**, *4*, 241.
6. Caruso, L.; Musumarra, G.; Katritzky, A. R. *Quant. Struct.–Act. Rel*. **1993**, *12*, 146.
7. Clementi, S.; Cruciani, G.; P. Fifi, P.; Riganelli, D.; Valigi, R.; Musumarra, G. *Quant. Struct.–Act. Rel.* **1996**, *15*, 108.
8. Prochazka, M. P.; Carlson, R. *Acta Chem. Scand*. **1989**, *43*, 651.

9. Prochazka, M. P.; Carlson, R. *Acta Chem. Scand*. **1990**, *44*, 614.
10. Fischer, E.; Hess, V. F. *Ber. Dtsch. Chem. Ges*. **1884**, *17,* 559.
11. Robinson, B. *The Fischer Indole Synthesis;* Wiley: Chichester, 1982.
12. Carlson, R.; Lundsted, T.; Albano, C. *Acta Chem. Scand*. **1985***, B39*, 79.
13. Wold, S. *Intell. Lab. Syst*. **1987**, *2*, 37.
14. SIMCA-P 8.0, *Umetrics AB*, Umea, Sweden, 1999.
15. Wold, S. *Technometrics* **1978**, *20*, 397.