Polarizabilities of Solvents from the Chemical Composition

Ramón Bosque and Joaquim Sales*

Departament de Química Inorgànica, Universitat de Barcelona, Martí i Franquès, 1, 08028-Barcelona, Spain

From the experimental polarizability values, α, of a large set of solvents containing 426 compounds with very different chemical characteristics, an additive model for the estimation of the polarizability is proposed. The derived average atomic polarizability of 10 elements, C, H, O, N, S, P, F, Cl, Br, and I, allows the calculation of the molecular polarizability of solvents from their chemical composition alone, without any other structural consideration. The average errors are 2.31% and 1.93% for the working set of 340 solvents and the prediction set of 86 solvents, respectively. Semiempirical quantum methods, such as, AM1, PM3, and MNDO, gave errors of about 35%. The density functional theory (DFT) calculations give better results than the semiempirical ones but poorer results than those obtained by the additive approach.

INTRODUCTION

Chemical transformations can be performed in the gaseous, liquid, or solid state. With good reason, however, the majority of such reactions are carried out in solution. Thus, it is very important to know the physicochemical properties of solvents.

The polarizability of a molecule, α, is one of the most significant electrical properties of solvents. It measures the ability of the electronic system of a molecule to be distorted by an electric field, and it is defined as the constant of proportionality between the strength of an applied electric field, ϵ, and the magnitude of the electric dipole moment the field induces

$$\mu_{\text{induced}} = \alpha\epsilon \tag{1}$$

Several studies have shown the relevant role of polarization response on properties in intermolecular weak interaction, and it has been argued that the primary electronic structure change that occurs in weakly interacting closed-shell species is that attributable to charge polarization. Many properties can be predicted by accounting for charge polarization. Some of these properties are as follows: boiling and melting points, vaporization and fusion enthalpies, Trouton constants, solubility parameters, polarity solvent scales, and also the structural properties of liquids.[1] Besides, polarizability parameters, such as the molar refraction and also the molar volume and parachor, are typical parameters in classic QSAR studies.[2] Excess molar refraction is also one descriptor of the well-known Abraham general solvation equation.[3] Very recently, polarizability has been used to describe molecular size and dispersion interactions with solvents in the prediction of logP.[4]

The polarizability of a molecule is governed by the strength with which the nuclear charges control the electrons and prevent their distortion by the applied field. If the molecule contains only a few electrons, the distribution of electrons is tightly controlled by the nuclear charge and the

polarizability is low. If the molecule contains large atoms with many electrons some distance from the nuclei, the degree of nuclear control is less, the electron distribution flabbier, and the polarizability greater. There is also a relation between the HOMO−LUMO energy separation in atoms and molecules. The electron distribution can be distorted readily if the LUMO energy lies close to the HOMO energy, so that the polarizability is then large. If the LUMO lies high above the HOMO, an applied field cannot significantly perturb the electron distribution and the polarizability is low. Molecules with small HOMO−LUMO gaps are typically large and with numerous electrons.

Experimentally, polarizability can be determined from the values of the molecules' refractive index and density, using the Lorentz-Lorentz equation that relates α to the molar refraction, R

$$R = \left(\frac{n_{\text{D}}^2 - 1}{n_{\text{D}}^2 + 2}\right)\frac{M}{\rho} = \frac{4}{3}\pi N_0 \alpha \tag{2}$$

where n_{D} is the refractive index at the sodium D-line, and the quotient between the molecular weight, M, and the density, ρ, is the molar volume, V, of the molecule. According to the Lorentz-Lorentz equation, the relationship between the polarizability, α, and the molar refraction, R, is α = 0.3964 R, where R is expressed in mL, and α in Å^3. There are other classic experimental methods that allow the determination of the polarizability from measures of magnitudes, such as from dipole moments or dielectric constants.

Quantum mechanical calculation of molecular polarizability may be carried out by solving the coupled perturbed Hartree−Fock (CPHF) equations with electric field perturbations.[5] Recently, Allinger et al.[6] have used a general formula based on MM3 force constants and bond lengths to compute bond polarizabilities and molecular polarizabilities by molecular mechanics.

A view that has prevailed for some time is that the polarizability of a molecule is simply the sum of the polarizabilities of its parts.[7] This is based on the finding that the molar refraction, which is proportional to the molecular

* Corresponding author phone: +34934021266; fax: +34934907725; e-mail: joaquim.sales@qi.ub.es.

Table 1. Experimental and Calculated Polarizability (Å^3) of the Working Set

no.	solvent	exptl	calcd	diff	error[a] (%)	no.	solvent	exptl	calcd	diff	error[a] (%)
1	3-methyltetrahydropyran	11.71	12.05	0.34	2.92	72	2,3-butanediol	9.33	9.23	−0.10	1.12
2	chloromethyloxirane	8.19	8.45	0.26	3.11	73	2,4,6-collidine	15.48	15.40	−0.08	0.49
3	propylene carbonate	8.55	9.09	0.54	6.35	74	2,4-dimethyl-3-pentanone	13.7	13.92	0.22	1.57
4	1,1,1,3,3,3-hexafluoro-2-propanol	7.2	7.09	−0.11	1.54	75	2,4-lutidine	13.4	13.54	0.14	1.05
5	1,1,1-trichloroethane	10.44	10.34	−0.10	0.97	76	2,4-pentanedione	10.99	10.40	−0.59	5.37
6	1,1,1-trichlorotrifluoroethane	10.47	10.48	0.01	0.11	77	2,5-dibromotoluene	18.72	18.57	−0.15	0.80
7	1,1,2,2-tetrachloroethane	12.21	12.32	0.11	0.93	78	2,6-di-*tert*-butylpyridine	24.98	24.72	−0.26	1.03
8	1,1,2-trichloro-1,2,2-trifluoro-ethane	10.43	10.48	0.05	0.49	79	2,6-difluoropyridine	9.43	9.91	0.48	5.07
						80	2,6-dimethyl-4-heptanone	17.5	17.64	0.14	0.82
9	1,1,2-trichloroethane	10.33	10.34	0.01	0.09	81	2,6-lutidine	13.54	13.54	0.00	0.01
10	1,1,3,3-tetraethylurea	20.17	20.08	−0.09	0.46	82	2-(hydroxymethyl)furan	9.84	10.06	0.22	2.18
11	1,1,3,3-tetramethylurea	12.86	12.62	−0.24	1.85	83	2-(*n*-butoxy)ethanol	13.1	12.95	−0.15	1.12
12	1,1-dichlorethylene	8.36	8.01	−0.35	4.20	84	2-(*tert*-butyl)phenol	18.6	18.47	−0.13	0.69
13	1,1-dichloroethane	8.38	8.35	−0.03	0.31	85	2-ethoxyethanol	9.48	9.23	−0.25	2.68
14	1,10-dichlorodecane	23.04	23.26	0.22	0.97	86	2-amino-1-butanol	10.09	9.89	−0.20	2.02
15	1,2,3,4-tetrahydronaphthalene	17.05	17.57	0.52	3.05	87	2-aminoethanol	6.48	6.16	−0.32	4.96
16	1,2,3,4-tetramethylbenzene	17.81	17.91	0.10	0.59	88	2-bromopyridine	12.78	12.94	0.16	1.23
17	1,2,3-propanetriol	8.14	7.92	−0.22	2.71	89	2-butanol	8.77	8.67	−0.10	1.15
18	1,2,3-trichloropropane	12.09	12.20	0.11	0.93	90	2-butanone	8.25	8.32	0.07	0.90
19	1,2,4-trichlorobenzene	16.32	16.41	0.09	0.58	91	2-chloroaniline	14.05	13.66	−0.39	2.76
20	1,2-butanediol	9.32	9.23	−0.09	1.01	92	2-chloroethanol	7.03	6.93	−0.10	1.48
21	1,2-diaminoethane	7.21	6.82	−0.39	5.43	93	2-chloropyridine	11.6	11.80	0.20	1.71
22	1,2-dibromoethane	10.75	10.63	−0.12	1.11	94	2-cyanoethanol	6.92	7.33	0.41	5.96
23	1,2-dibromopropane	12.54	12.49	−0.05	0.36	95	2-ethylphenol	14.6	14.74	0.14	0.99
24	1,2-dichloro-1,1,2,3,3,3-hexa-fluoropane	10.52	10.50	−0.02	0.17	96	2-hexanone	11.95	12.05	0.10	0.85
						97	2-iodopropane	11.67	11.54	−0.13	1.14
25	1,2-dichlorobenzene	14.3	14.43	0.13	0.91	98	2-isopropylphenol	16.4	16.61	0.21	1.27
26	1,2-dichloroethane	8.43	8.35	−0.08	0.90	99	2-methoxyethanol	7.62	7.36	−0.26	3.38
27	1,2-difluorobenzene	10.42	10.55	0.13	1.29	100	2-methoxyphenol	13.6	13.44	−0.16	1.19
28	1,2-dimethoxybenzene	15.69	15.30	−0.39	2.48	101	2-methyl-1-propanol	8.81	8.67	−0.14	1.60
29	1,2-dimethoxyethane	9.6	9.23	−0.37	3.90	102	2-methyl-2-butanol	10.64	10.53	−0.11	1.01
30	1,2-ethanediol ethylen glycol	5.72	5.50	−0.22	3.88	103	2-methyl-2-nitropropane	10.3	10.10	−0.20	1.96
31	1,3,5-triflurobenzene	10.24	10.60	0.36	3.53	104	2-methylbutane	10.13	9.98	−0.15	1.53
32	1,3-butanediol	9.37	9.23	−0.14	1.54	105	2-nitropropane	8.59	8.23	−0.36	4.14
33	1,3-dibromobenzene	16.59	16.71	0.12	0.70	106	2-pentanol	10.57	10.53	−0.04	0.35
34	1,3-dichlorobenzene	14.38	14.43	0.05	0.35	107	2-pentanone	10.11	10.19	0.08	0.77
35	1,3-dichloropropane	10.17	10.22	0.05	0.47	108	2-picoline	11.57	11.68	0.11	0.92
36	1,3-dimethyl-2-oxo-hexahydro-pyrimidine	13.8	14.14	0.34	2.47	109	allyl alcohol	6.71	6.46	−0.25	3.72
						110	2-propenenitrile	6.24	6.43	0.19	3.05
37	1,3-dimethylimidazolidin-2-one	12.14	12.28	0.14	1.13	111	2-propyne-1-ol	5.98	6.12	0.14	2.26
38	1,3-dioxolane	6.79	7.02	0.23	3.35	112	3,3-dimethyl-2-butanone	11.85	12.05	0.20	1.70
39	1,3-propanediol	7.55	7.36	−0.19	2.49	113	3,4-lutidine	13.34	13.54	0.20	1.50
40	1,4-butanediol	9.35	9.23	−0.12	1.33	114	3-bromopyridine	12.76	12.94	0.18	1.38
41	1,4-cyclohexadiene	10.48	10.80	0.32	3.10	115	3-chloro-1,1,1-trifluoropropane	8.23	8.38	0.15	1.76
42	1,4-dichlorobutane	12.02	12.08	0.06	0.51	116	3-chlorophenol	13.6	13.00	−0.60	4.40
43	1,4-difluorobenzene	10.27	10.55	0.28	2.77	117	3-chloropropyne	7.52	7.54	0.02	0.31
44	1,4-dioxane	8.63	8.88	0.25	2.91	118	3-ethyl-3-pentanol	14.44	14.26	−0.18	1.25
45	1,5-pentanediol	11.2	11.09	−0.11	0.99	119	3-heptanone	13.7	13.92	0.22	1.57
46	1,8-cineole	18.1	19.16	1.06	5.87	120	3-methoxypropanenitrile	8.79	9.20	0.41	4.62
47	1-bromobutane	11.26	11.24	−0.03	0.22	121	3-methyl-1-butanol	10.61	10.53	−0.08	0.73
48	1-butanol	8.79	8.67	−0.12	1.38	122	3-methyl-2-butanone	10.02	10.19	0.17	1.67
49	1-chloronaphthalene	19.37	18.87	−0.50	2.61	123	3-methyl-2-oxazolidinone	9.29	9.75	0.46	4.99
50	1-decanol	19.83	19.85	0.02	0.11	124	3-pentanone	10.03	10.19	0.16	1.57
51	1-heptanol	14.3	14.26	−0.04	0.28	125	3-picoline	11.5	11.68	0.18	1.54
52	1-hexanol	12.46	12.40	−0.06	0.51	126	4-methylphenol	13.2	12.88	−0.32	2.42
53	1-hexene	11.79	11.49	−0.30	2.51	127	5-acetyl-5-methyl-1,3-dioxane	14.26	14.68	0.42	2.98
54	1-hexyne	11.09	11.15	0.06	0.54	128	5-isopropyl-2-methylphenol	18.6	18.47	−0.13	0.69
55	1-iodonaphthalene	22.49	22.17	−0.32	1.43	129	5-nonanone	17.44	17.64	0.20	1.16
56	1-iodopropane	11.5	11.54	0.04	0.32	130	*N*,*N*,*N*′,*N*′-tetraethylsulfamide	21.32	22.11	0.79	3.71
57	1-methyl-2-pyrrolidinone	10.66	11.06	0.40	3.75	131	*N*,*N*,*N*,*N*-tetramethylguanidine	13.85	13.28	−0.57	4.10
58	1-methylimidazole	9.24	9.51	0.27	2.94	132	*N*,*N*-diethylformamide	11.56	11.40	−0.16	1.34
59	1-methylpiperidine	12.62	12.71	0.09	0.73	133	*N*,*N*-dimethylacetamide	9.69	9.54	−0.15	1.54
60	1-mettylnaftalene	19.42	18.74	−0.68	3.48	134	*N*,*N*-dimethylaniline	16.29	15.40	−0.89	5.44
61	1-nitropropane	8.6	8.23	−0.37	4.25	135	*N*,*N*-dimethylbenzylamine	17.53	17.27	−0.26	1.49
62	1-octanol	16.14	16.12	−0.02	0.10	136	*N*,*N*-dimethylcyclohexylamine	16.09	16.44	0.35	2.17
63	1-pentanol	10.61	10.53	−0.08	0.73	137	*N*,*N*-dimethylformamide	7.93	7.68	−0.25	3.19
64	1-phenylethanol	14.62	14.74	0.12	0.85	138	*N*,*N*-dimethylhtioformamide	11.39	10.12	−1.27	11.18
65	1-propanol	6.96	6.81	−0.16	2.23	139	*N*-methylacetamide	7.85	7.68	−0.17	2.20
66	2,2,2-trichloroethanol	11	10.90	−0.10	0.94	140	*N*-methylaniline	14.21	13.54	−0.67	4.71
67	2,2,2-trifluoroethanol	5.2	5.08	−0.12	2.24	141	*N*-methylcyclohexylamine	14.02	14.58	0.56	3.96
68	2,2,3,3-tetrafluoro-1-propanol	7.08	6.99	−0.09	1.21	142	*N*-methylformamide	6.01	5.81	−0.20	3.27
69	2,2,4,4-tetramethyl-3-pentanone	17.37	17.64	0.27	1.57	143	*N*-methylformanilide	15.89	15.62	−0.27	1.72
70	2,2,5,5-tetramethyltetrahydrofuran	15.4	15.78	0.38	2.46	144	acetic acid	5.18	5.15	−0.03	0.51
71	2,3,4,5,6-pentafluorotoluene	12.24	12.56	0.32	2.62	145	acetic anhydride	8.93	9.09	0.16	1.83

1156 *J. Chem. Inf. Comput. Sci., Vol. 42, No. 5, 2002*

BOSQUE AND SALES

Table 1 (Continued)

no.	solvent	exptl	calcd	diff	error[a] (%)	no.	solvent	exptl	calcd	diff	error[a] (%)
146	acetone	6.47	6.46	−0.01	0.15	220	ethyl lactate	11.34	11.30	−0.04	0.34
147	acetonitrile	4.44	4.91	0.47	10.62	221	ethyl propiolate	10.08	10.06	−0.03	0.25
148	acetophenone	14.45	14.40	−0.05	0.35	222	ethyl tricholoacetate	14.82	14.84	0.02	0.11
149	allylamine	7.57	7.12	−0.45	5.94	223	ethyl vinyl ether	8.74	8.32	−0.42	4.76
150	aniline	12.16	11.68	−0.48	3.97	224	ethylbenzene	14.24	14.19	−0.05	0.37
151	benzaldehyde	12.7	12.54	−0.16	1.29	225	ethylene carbonate	6.6	7.23	0.63	9.54
152	benzene	10.44	10.46	0.02	0.19	226	ethylene glycol dimethyl ether	9.58	9.23	−0.35	3.70
153	benzonitrile	12.57	12.85	0.28	2.24	227	ethylene sulfite	8	8.71	0.71	8.84
154	benzophenone	22.5	22.34	−0.16	0.72	228	ethyl propanoate	10.6	10.74	0.14	1.37
155	benzoyl chloride	14.7	14.52	−0.18	1.22	229	formamida	4.22	3.95	−0.27	6.40
156	benzyl alcohol	12.89	12.88	−0.01	0.07	230	formic acid	3.39	3.29	−0.10	2.96
157	biacetyl	8.3	8.54	0.24	2.84	231	furan	7.35	7.63	0.28	3.87
158	bis(methoxyethyl)ether	13.9	13.51	−0.39	2.80	232	γ-butyrolactone	7.94	8.54	0.60	7.51
159	bromobenzene	13.52	13.58	0.06	0.47	233	hexachloro-1,3-butadiene	19.75	19.33	−0.42	2.12
160	bromoform	11.87	11.89	0.02	0.17	234	hexachlorocyclopentadiene	20.67	20.85	0.18	0.87
161	bromotrichloromethane	11.67	11.60	−0.07	0.61	235	hexafluorobenzene	10.56	10.74	0.18	1.74
162	butanenitrile	8.12	8.64	0.52	6.39	236	hexamethylphosphoric acid triamide	19	18.75	−0.25	1.32
163	butyraldehyde	8.2	8.32	0.12	1.51						
164	carbon disulfide	8.51	7.82	−0.69	8.06	237	hexanenitrile	11.84	12.37	0.53	4.45
165	carbon tetrachloride	10.54	10.46	−0.08	0.76	238	iodoethane	9.66	9.67	0.01	0.13
166	chloroacetonitrile	6.4	6.90	0.50	7.76	239	iodomethane	7.68	7.81	0.13	1.68
167	cinnamaldehyde	17.5	15.92	−1.58	9.04	240	isoamyl acetate	14.48	14.47	−0.01	0.05
168	cis-1,2-dichloroethylene	8.08	8.01	−0.07	0.88	241	isooctane	15.63	15.57	−0.06	0.40
169	cycloheptane	12.84	13.36	0.52	4.04	242	isopropylbenzene	16.1	16.05	−0.05	0.30
170	cycloheptatriene	12.46	12.32	−0.14	1.10	243	m-xylene	14.33	14.19	−0.14	1.00
171	cyclohexane	11.04	11.49	0.45	4.12	244	mesitylene	16.25	16.05	−0.20	1.22
172	cyclohexanol	11.4	12.05	0.65	5.71	245	methanesulfonic acid	6.7	7.19	0.49	7.29
173	cyclohexanone	11.12	11.71	0.59	5.28	246	methyl oleate	36.37	36.49	0.12	0.34
174	cyclohexene	10.79	11.15	0.36	3.33	247	methyl acetate	7	7.02	0.02	0.25
175	cyclohexylamine	12.47	12.71	0.24	1.94	248	methyl acrylate	8.85	8.54	−0.31	3.55
176	cyclohexylbenzene	20.77	21.30	0.53	2.54	249	methyl benzoate	15.06	14.96	−0.10	0.69
177	cyclooctane	14.62	15.22	0.60	4.12	250	methyl butanoate	10.67	10.74	0.07	0.70
178	cyclooctatetraene	14.07	13.84	−0.23	1.62	251	methyl formate	5.21	5.15	−0.06	1.09
179	cyclopentane	9.23	9.63	0.40	4.34	252	methyl hexanoate	14.34	14.47	0.13	0.92
180	cyclopentanol	9.72	10.19	0.47	4.81	253	methyl methacrylate	10.6	10.40	−0.20	1.89
181	cyclopropyl methyl ketone	9.51	9.84	0.33	3.50	254	methyl octanoate	18.09	18.20	0.11	0.61
182	di-n-buthyl ether	16.31	16.12	−0.19	1.14	255	methyl pentanoate	12.53	12.61	0.08	0.63
183	di-n-buthyl phthalate	30.75	30.64	−0.11	0.37	256	methyl propionate	8.82	8.88	0.06	0.69
184	di-n-butylamine	16.9	16.78	−0.12	0.69	257	methyl salicylate	16	15.51	−0.49	3.04
185	di-ter-buthyl ether	16.27	16.12	−0.15	0.90	258	methyl trichloroacetate	12.87	12.97	0.10	0.79
186	di-tert-buthyl sulfide	19.31	18.56	−0.75	3.87	259	methyl trifluoroacetate	7.19	7.16	−0.03	0.43
187	dibromomethane	8.75	8.77	0.02	0.20	260	hexamethyl phosphoramide	16.05	15.67	−0.38	2.37
188	dichloroacetic acid	9.1	9.12	0.02	0.26	261	morpholine	9.4	9.54	0.14	1.50
189	dicyclopropyl ketone	12.52	13.23	0.71	5.63	262	morpholine-4-carbonitrile	11.22	11.93	0.71	6.35
190	diethanolamine	10.7	10.44	−0.26	2.40	263	n-butanoic acid	8.8	8.88	0.08	0.92
191	diethoxymethane	11.4	11.09	−0.31	2.72	264	n-buthyl acetate	12.57	12.61	0.04	0.31
192	diethyl carbonate	11.32	11.30	−0.02	0.16	265	n-buthyl methyl ether	10.8	10.53	−0.27	2.48
193	diethyl disulfide	14.59	14.11	−0.48	3.32	266	n-buthylamine	9.64	9.33	−0.31	3.23
194	diethyl malonate	15.11	15.24	0.13	0.87	267	n-decane	19.33	19.29	−0.04	0.18
195	diethyl sulfate	13	13.34	0.34	2.59	268	n-dodecane	23.03	23.02	−0.01	0.04
196	diethyl sulfide	11.38	11.11	−0.27	2.39	269	n-dodecanol	23.5	23.58	0.08	0.34
197	diethyl sulfite	12.71	12.78	0.07	0.55	270	n-heptane	13.81	13.70	−0.11	0.77
198	diethylamine	9.63	9.33	−0.30	3.13	271	n-hexadecane	30.42	30.48	0.06	0.19
199	di(ethylene glycol)	10.09	9.78	−0.31	3.04	272	n-octane	15.6	15.57	−0.03	0.21
200	di(ethyleneglycol) dimethyl ether	13.97	13.51	−0.46	3.29	273	n-pentane	10.11	9.98	−0.13	1.33
201	diiodomethane	12.96	13.10	0.14	1.06	274	n-pentanoic acid	10.6	10.74	0.14	1.37
202	diisopropyl ether	12.65	12.40	−0.25	2.01	275	n-propyl acetate	10.72	10.74	0.02	0.23
203	diisopropylamine	13.45	13.06	−0.39	2.93	276	n-propyl formate	8.93	8.88	−0.05	0.55
204	dimethoxymethane	7.68	7.36	−0.32	4.14	277	n-undecane	21.14	21.16	0.02	0.09
205	dimethyl carbonate	7.56	7.57	0.01	0.19	278	nitrobenzene	13	12.45	−0.55	4.26
206	dimethyl disulfide	10.84	10.38	−0.46	4.27	279	nitrocyclohexane	13.26	13.48	0.22	1.67
207	dimethyl phthalate	19.59	19.45	−0.14	0.70	280	nitroethane	6.76	6.37	−0.39	5.76
208	dimethyl sulfate	8.88	9.61	0.73	8.21	281	nitromethane	4.97	4.51	−0.46	9.32
209	dimethyl sulfide	7.63	7.38	−0.25	3.27	282	o-xylene	14.25	14.19	−0.06	0.44
210	dimethyl sulfite	8.94	9.05	0.11	1.25	283	p-chloroacetophenone	16.5	16.38	−0.12	0.70
211	dimethyl sulfoxide	8.03	7.94	−0.09	1.15	284	p-methoxybenzaldehyde	15.9	14.96	−0.94	5.93
212	diphenyl ether	21.02	20.82	−0.20	0.95	285	p-methylacetophenone	16.4	16.26	−0.14	0.83
213	diphenylmethane	21.9	22.13	0.23	1.04	286	pentachloroethane	14.17	14.31	0.14	0.98
214	ethanol	5.13	4.94	−0.19	3.68	287	pentafluorobenzene	10.43	10.70	0.27	2.56
215	ethyl acetate	8.87	8.88	0.01	0.12	288	pentafluoropyridine	10.18	10.05	−0.13	1.28
216	ethyl acetoacetate	12.82	12.82	0.00	0.01	289	pentanenitrile	9.99	10.50	0.51	5.13
217	ethyl benzoate	16.94	16.82	−0.12	0.71	290	perfluoro(methylcyclohexane)	13.71	14.02	0.31	2.27
218	ethyl chloroacatate	10.76	10.87	0.11	0.99	291	perfluoro-n-heptane	14.7	14.46	−0.24	1.63
219	ethyl formate	7.09	7.02	−0.07	1.03	292	perfluoro-n-hexane	12.71	12.50	−0.21	1.63

POLARIZABILITIES OF SOLVENTS

J. Chem. Inf. Comput. Sci., Vol. 42, No. 5, 2002 **1157**

Table 1 (Continued)

no.	solvent	exptl	calcd	diff	error[a] (%)	no.	solvent	exptl	calcd	diff	error[a] (%)
293	perfluoro-*n*-octane	17.55	16.42	−1.13	6.44	319	*tert*-butyl hydroperoxide	9.44	9.23	−0.21	2.27
294	perfluorodecalin	18.49	19.46	0.97	5.23	320	tetraethylene glycol	18.71	18.35	−0.36	1.91
295	phenol	11	11.02	0.02	0.15	321	tetrahydrofuran	7.97	8.32	0.35	4.44
296	phenylacethylene	13.91	13.50	−0.41	2.97	322	tetrahydrothiophene	10.41	10.76	0.35	3.40
297	phenylacetone	15.8	16.26	0.46	2.93	323	tetramethylene sulfoxide	10.77	11.32	0.55	5.11
298	phenylacetonitrile	13.96	14.71	0.75	5.41	324	tetrtahydropyran	9.85	10.19	0.34	3.43
299	piperidine	10.65	10.85	0.20	1.86	325	thiophene	9.72	10.07	0.35	3.64
300	propanenitrile	6.3	6.78	0.48	7.55	326	toluene	12.4	12.32	−0.08	0.62
301	propanoic acid	6.72	7.02	0.30	4.42	327	tri-*n*-buthyl phosphate	27.55	27.95	0.40	1.46
302	propargylamine	7.29	6.78	−0.51	7.06	328	tri-*n*-buthylamine	24.43	24.24	−0.19	0.78
303	propionaldehyde	6.4	6.46	0.06	0.94	329	tri-*n*-propyl phosphate	22.24	22.36	0.12	0.54
304	propiophenone	16.2	16.26	0.06	0.39	330	trichloroacetonitrile	10.48	10.87	0.39	3.69
305	pyridine	9.58	9.81	0.23	2.43	331	trichloroethylene	10.08	9.99	−0.09	0.85
306	pyrimidine	8.59	9.17	0.58	6.71	332	triethanolamine	15.1	14.73	−0.37	2.47
307	pyrrole	8.23	8.29	0.06	0.78	333	triethylamine	13.48	13.06	−0.42	3.14
308	pyrrolidin-2-one	8.66	9.20	0.54	6.19	334	trimethyl orthoacetate	11.91	11.65	−0.26	2.21
309	pyrrolidine	8.77	8.98	0.21	2.44	335	trimethyl orthoformate	10.09	9.78	−0.31	3.04
310	pyrrolidine-1-carbonitrile	11.09	11.38	0.29	2.57	336	trimethyl phosphate	11.07	11.18	0.11	0.98
311	quinoline	16.65	16.23	−0.42	2.50	337	trimethylacetonitrile	10	10.50	0.50	5.03
312	styrene	14.5	13.84	−0.66	4.54	338	trimethylene sulfide	8.64	8.90	0.26	3.01
313	sulfolane	10.73	11.88	1.15	10.70	339	vinyl acetate	8.87	8.54	−0.33	3.76
314	*tert*-butyl alcohol	8.82	8.67	−0.15	1.71	340	2-methylphenol	13	12.88	−0.12	0.92
315	*tert*-amyl methyl ether	12.48	12.40	−0.08	0.67				av	0.00	
316	*tert*-buthyl ethyl ether	12.62	12.40	−0.22	1.77				av absolute relative		
317	*tert*-buthyl methyl ether	10.65	10.53	−0.12	1.10				error 2.31%		
318	*tert*-buthylamine	9.69	9.33	−0.36	3.73						

[a] Absolute relative error: absolute value of 100[(calcd − exptl)/exp].

polarizability, is an additive property. This means that the various atoms and functional groups in a molecule can be assigned refraction values whose sum for the whole molecule is the molar refraction. Therefore, the value for a given group or atom is fairly constant for a variety of molecules. Other successful empirical approaches to the calculation of average molecular polarizabilities take into account the atomic hybridization. Thus, different sets of optimized atomic hybrid components and atomic hybrid polarizabilities have been proposed.[8]

The additivity hypothesis has been extended in the interpretation of anisotropy of polarizability; thus, polarizability tensors have been ascribed to various bonds and functional groups according to the hypothesis that component addition of the group tensors yields the molecular polarizability tensor.[9] Recently,[10] a model has been derived from ab initio calculations of polarizabilities of over 30 molecules containing up to four non-hydrogen (C, N, O, and F) atoms. This model assigns atomic polarizabilities of the three tensor components of each respective atomic polarizability. In all the studies reported, the atomic anisotropic and isotropic polarizabilities assigned to each kind of atom depend on the chemical and structural characteristics of the element and different values are assigned to sp^3, sp^2, sp, aromatic carbon atoms; amine, amide, nitrile nitrogen atoms; alcohol, ether, ester oxygen atoms, etc.

QSPR studies have been successfully applied to the correlation of many diverse physicochemical properties of chemical compounds. In the literature there are at least two papers on the estimation of molar refraction by QSPR methodology. Kier and Hall[11,12] have found a very good model for the molar refraction of alkanes and some derivatives using the four topological chi indexes. The refractive index of organic compounds[13,14] and polymers[15] have also been studied by the QSPR approach. Neural Network

methods have been used in the refractive index prediction of alkanes[16] as well as alkenes.[17]

The present paper studies a very large set of solvents with very different chemical and structural characteristics that contain up to nine non-hydrogen elements. We have found that the polarizability of these solvents can be easily estimated from the atomic composition of the molecule only. A set of average atomic polarizabilities is proposed for the elements C, H, O, S, N, P, F, Cl, Br, and I.

DATA AND COMPUTATIONAL METHODS

The set of solvents studied contains 426 compounds of very different chemical nature. These include the following: cyclic and acyclic nonaromatic hydrocarbons; aromatic hydrocarbons; halogenated and perfluorinated compounds; alcohols; phenols; ethers; esters; aldehydes; ketones; carboxylic acids; amines; nitriles; nitro-derivatives; amides; and sulfur- and phosphorus-containing solvents. The structures include C, H, O, N, S, P, and the four halogen elements.

The experimental values of polarizability, α, have been obtained from the refractive index and the density of the compounds using the Lorentz-Lorentz equation. The experimental refractive index values were measured at 20 or 25 °C and at the D-line of sodium (589 nm) wavelength. Even though the refractive index is affected by changes in temperature, the polarizability remains nearly constant with changes in temperature by virtue of the density factor, which is a function of temperature and thus offsets this effect. The compounds were taken from the lists of the solvents compiled by Marcus,[18] Abboud,[19] and Reichardt.[20] The data set ranged from 3.26 to 36.37 Å^3, with a mean value of 12.5 Å^3. This data set was split randomly into a 340 member working set (Table 1) and an external prediction set of 86 compounds (Table 2).

Table 2. Experimental and Calculated Polarizability (Å³) of the Prediction Set

no.	structure	exptl	calcd	diff	error[a] (%)	no.	structure	exptl	calcd	diff	error[a] (%)
1	methyloxirane	6.25	6.46	0.21	3.36	45	cyclopentanone	9.25	9.84	0.59	6.41
2	(trifluoromethyl)benzene	12.25	12.47	0.22	1.76	46	di(2-chloroethyl)ether	12.6	12.64	0.04	0.31
3	1,2-dibromobenzene	16.43	16.71	0.28	1.68	47	di-*n*-buthyl sulfide	18.74	18.56	−0.18	0.94
4	1,2-propanediol	7.55	7.36	−0.19	2.49	48	di-*n*-propyl ether	12.67	12.40	−0.27	2.16
5	1,3-difluorobenzene	10.33	10.55	0.22	2.17	49	diallylamine	12.84	12.37	−0.47	3.69
6	1,5-cyclooctadiene	14.16	14.53	0.37	2.63	50	dibenzyl ether	24.5	24.55	0.05	0.19
7	1-amino-2-propanol	8.36	8.02	−0.34	4.04	51	dichloromethane	6.52	6.49	−0.03	0.46
8	1-bromopropane	9.42	9.37	−0.05	0.52	52	diethyl ether	8.98	8.67	−0.31	3.47
9	1-chlorobutane	10.12	10.10	−0.02	0.23	53	di(ethylene glycol) diethyl ether	17.67	17.24	−0.43	2.44
10	1-chloropropane	8.32	8.23	−0.09	1.05	54	diisopropyl sulfide	15.19	14.84	−0.35	2.33
11	1-iodobutane	13.33	13.40	0.07	0.53	55	dodecanenitrile	22.68	23.55	0.87	3.83
12	1-methylpyrrole	10.24	10.16	−0.08	0.80	56	ethyl phenyl ether	15.02	14.74	−0.28	1.83
13	1-nonyne	16.62	16.74	0.12	0.73	57	ethyl salicylate	17.8	17.38	−0.42	2.37
14	2,2-dichloropropane	10.32	10.22	−0.10	0.99	58	fluorobenzene	10.33	10.51	0.18	1.71
15	2,4-dimethyl-3-pentanol	14.21	14.26	0.05	0.35	59	fluorotrichloromethane	8.61	8.52	−0.09	1.01
16	2,4-dimethylphenol	14.8	14.74	−0.06	0.38	60	glyceryl triacetate	19.31	19.74	0.43	2.22
17	2,5-dimethyltetrahydrofuran	11.4	12.05	0.65	5.71	61	hexachloropropene	17.84	17.81	−0.03	0.15
18	2-cyanopyridine	11.7	12.20	0.50	4.31	62	iodobenzene	15.58	15.75	0.17	1.08
19	2-fluoropyridine	10.61	9.86	−0.75	7.06	63	isopropyl nitrate	9.61	8.79	−0.82	8.52
20	2-methoxy-1,3-dioxolane	9.36	9.44	0.08	0.83	64	methanol	3.26	3.08	−0.18	5.60
21	2-methyltetrahydrofuran	9.95	10.19	0.24	2.39	65	methyl decanoate	21.73	21.93	0.20	0.91
22	2-propanol	6.98	6.81	−0.18	2.51	66	methyl dodecanoate	25.4	25.65	0.25	1.00
23	3-methylphenol	13	12.88	−0.12	0.92	67	methyl linoleate	36.23	36.15	−0.08	0.23
24	3-methylsulfolane	12.65	13.74	1.09	8.63	68	methyl phenyl ether	13.11	12.88	−0.23	1.75
25	3-pentanol	10.62	10.53	−0.09	0.82	69	methyl phenyl sulfane	15.63	15.32	−0.31	1.98
26	3-phenyl-1-propanol	16.44	16.61	0.17	1.02	70	methylcyclohexane	12.96	13.36	0.40	3.07
27	4-heptanone	13.67	13.92	0.25	1.79	71	*n*-heptanoic acid	14.3	14.47	0.17	1.20
28	4-methyl-2-pentanone	11.98	12.05	0.07	0.60	72	*n*-hexane	11.94	11.84	−0.10	0.84
29	4-picoline	11.52	11.68	0.16	1.36	73	*n*-hexanoic acid	12.5	12.61	0.11	0.87
30	*N,N*-diethylacetamide	13.28	13.27	−0.01	0.09	74	*n*-nonane	17.45	17.43	−0.02	0.11
31	*N,N*-diethylcyanamide	11.71	11.72	0.01	0.09	75	*n*-pentadecane	28.55	28.61	0.06	0.22
32	*N,N*-diisopropylcyanamide	15.31	15.45	0.14	0.90	76	*p*-xylene	14.35	14.19	−0.16	1.13
33	*N,N*-dimethylcyanamide	7.94	7.99	0.05	0.66	77	pentamethylehe-sulfide	12.3	12.63	0.33	2.66
34	*N,N*-dimethylpropinoamide	11.46	11.40	−0.06	0.48	78	piperidine-1-carbonitrile	12.83	13.24	0.41	3.19
35	*N*-(*tert*-butyl)benzylamine	21.11	21.00	−0.11	0.54	79	tetrachloroethylene	12.07	11.98	−0.09	0.75
36	*N*-benzylmethylamine	15.52	15.40	−0.12	0.74	80	thiobis(2-ethanol)	12.5	12.22	−0.28	2.22
37	benzoyl bromide	15.7	15.66	−0.04	0.26	81	*trans*-1,2-dichloroethylene	8.22	8.01	−0.21	2.56
38	benzylamine	13.6	13.54	−0.06	0.44	82	triethyl phosphate	16.57	16.77	0.20	1.20
39	bis(2-chloroethyl)ether	12.63	12.64	0.01	0.07	83	tri(ethylene) glycol	14.36	14.07	−0.29	2.04
40	bromoethane	7.6	7.51	−0.09	1.22	84	tri(ethylene) glycol dimethyl ether	18.34	17.80	−0.54	2.97
41	chlorobenzene	12.4	12.44	0.04	0.36	85	trifluoroacetic acid	5.4	5.30	−0.10	1.94
42	chlorocyclohexane	13.03	13.48	0.45	3.45	86	undecanenitrile	20.94	21.69	0.75	3.56
43	chloroform	8.53	8.48	−0.05	0.64				av	0.03	
44	*cis*-decalin	17.4	18.60	1.20	6.92				av absolute relative error 1.93%		

[a] Absolute relative error: absolute value of 100[(calcd − exptl)/exp].

The additivity model was developed by fitting the experimental values of α with the number of atoms of each element present in the molecule. The goodness of the correlation is tested by the coefficient regression (R^2), the F-test, the standard deviation (SD), the relative standard deviation (RSD), that is the SD divided by the mean experimental polarizability, and by the average absolute relative error. The stability of the correlations was tested against the cross-validated coefficient, R^2_{cv}. The R^2_{cv} describes the stability of a regression model obtained by focusing on the sensitivity of the model to the elimination of any single data point.

The quantum calculations of the polarizability by the AM1, PM3, and MNDO semiempirical methods were carried out with the MOPAC 6.0 program.[21] DFT calculations were performed with the Gaussian 98 program[22] using the B3LYP functional[23] and the 6-31G basis set[24,25] with polarization and diffuse functions[26] on all atoms. In all cases a full geometric optimization was carried out, without imposing any symmetry constraint.

RESULTS AND DISCUSSION

To test the suitability of an atomic additive approach to evaluate the polarizability of the solvents, we selected the subsets of compounds that have the same molecular weight. There are two types of isomers, the structural type and the constitutional type. The structural isomers are the compounds with the same functional groups but with the substituents in different positions, i.e., *n*-octane and isooctane; *o*-, *m*-, and *p*-xylene; etc. The constitutional isomers have different functional groups, i.e., acetone and methyloxirane; 1-hexanol and di-*n*-propyl ether; cyclohexane and 1-hexene, etc. Among the solvents studied are 77 groups of isomers containing 190 different compounds. Table 3 contains some of these groups of isomers, showing the very close polarizability values within each group of isomers. The average of the differences

POLARIZABILITIES OF SOLVENTS

J. Chem. Inf. Comput. Sci., Vol. 42, No. 5, 2002 **1159**

Table 3. Experimental Polarizabilities ($Å^3$) of Some Groups of Isomer Compounds

structural isomers				constitutional isomers			
solvent	exptl	diff[a]	rel diff (%)[b]	solvent	exptl	diff[a]	rel diff (%)[b]
n-octane	15.60			propargylamine	7.25		
isooctane	15.63	0.03	0.19	propanenitrile	7.54	0.29	3.86
ethylbenzene	14.24			methyloxirane	6.25		
o-xylene	14.25	0.01	0.05	acetone	6.47	0.22	3.33
m-xylene	14.33	0.09	0.57	allyl alcohol	6.71	0.46	6.82
p-xylene	14.35	0.11	0.73	propanoic acid	6.72		
1-iodopropane	11.50			1,3-dioxolane	6.79	0.07	0.96
2-iodopropane	11.67	0.17	1.50	methyl acetate	7.00	0.28	3.96
1,4-difluorobenzene	10.27			ethyl formate	7.09	0.37	5.15
1,3-difluorobenzene	10.33	0.06	0.62	1,2−propanediol	7.55		
1,2-difluorobenzene	10.42	0.15	1.50	2-methoxyethanol	7.62	0.07	0.95
2-butanol	8.77			dimethoxymethane	7.68	0.13	0.70
1-butanol	8.79	0.02	0.19	cyclopentanol	9.72		
2-methyl-1-propanol	8.81	0.04	0.40	tetrahydropyran	9.85	0.13	1.29
tert-butyl alcohol	8.82	0.05	0.52	2-methyltetrahydrofuran	9.95	0.23	2.22
tert-amyl methyl ether	12.48			3-pentanone	10.03	0.31	3.05
tert-buthyl ethyl ether	12.62	0.14	1.16	hexanenitrile	11.69		
diisopropyl ether	12.65	0.17	1.35	diallylamine	12.91	1.22	9.42
di-n-propyl ether	12.67	0.19	1.54	benzyl alcohol	12.89		
ethyl propionate	8.82			2-methylphenol	13.01	0.12	0.92
ethyl acetate	8.87	0.05	0.48	methyl phenyl ether	13.11	0.22	1.66
n-propyl formate	8.93	0.11	1.15				

[a] Difference to the isomer with the lowest value of polarizability in the group. [b] Relative difference of each isomer expressed as %.

Table 4. Intercept and Coefficients of the Fit

	coefficient	SD	t-test
intercept	3.18E−01	6.05E−02	5.3
number of Br atoms	3.29E+00	6.29E−02	52.4
number of C atoms	1.51E+00	1.05E−02	143.7
number of Cl atoms	2.16E+00	2.35E−02	92.0
number of F atoms	2.22E−01	1.10E−02	20.2
number of H atoms	1.74E−01	5.73E−03	30.4
number of I atoms	5.45E+00	1.19E−01	45.9
number of N atoms	1.03E+00	3.65E−02	28.3
number of O atoms	5.71E−01	1.98E−02	28.9
number of P atoms	2.48E+00	1.65E−01	15.0
number of S atoms	2.99E+00	6.62E−02	45.2

of the isomers of all groups is 0.17, which represents a relative difference of 1.63%. The relative difference is lower for the set of the structural isomers (0.72%) than for the constitutional isomers (2.91%). The largest relative difference, 9.42%, is observed between hexanitrile and diallylamine. In both types of isomers the differences are low enough to indicate that the polarizability of the solvents can be calculated from their chemical composition.

We used multilinear regression to fit the polarizability of the 340 solvents of the working set to the number of atoms of each element present in the molecule. Table 4 shows the results obtained. In all cases, the standard deviation of each coefficient is 1 or 2 orders of magnitudes lower than the value of the respective coefficient, as can be seen in the high values of the t-test. Thus, indication is provided of the robustness of the coefficients for each element. It is also worth noticing that the intercept has a value of 0.3, which is 10 times lower than the lowest value of the experimental polarizability and forty times lower than the mean experimental polarizability values. These figures show that the intercept has very low significance in the estimation of the polarizability of the compounds studied. The correlation obtained is very good, $R^2 = 0.9943$; F = 5784; $R^2_{cv} = 0.9938$, with a standard deviation (SD) of 0.34 that represents a relative standard deviation (RSD) of 2.76%. The average

absolute relative error is 2.31%. The residuals are randomly positive and negative, and their mean value is zero (Table 1). Figure 1 shows the relationship between calculated and experimental polarizability values. The regression equation has values for the slope and the intercept that are not significantly different from unity and zero, respectively. This again demonstrates the excellent prediction ability of the derived coefficients.

With these parameters, the 86 solvents of the prediction set also gives a very good correlation, $R^2 = 0.9989$; F = 20338, with a SD of 0.33 and a RSD of 2.69%, and the average absolute relative error is 1.93%, see Figure 2.

These results show clearly that it is possible to calculate the polarizability of the solvents from only the chemical composition of the compound. The coefficients of Table 4, can be taken as the average atomic polarizability of the 10 elements involved. We use the expression "average atomic polarizability" because the value corresponds to the global polarizability of each element, not divided in its three tensor components. Moreover, these values are also average polarizabilities because only one value is proposed for each element, independently of its chemical and structural characteristics. Hence, in this model all the carbon atoms, sp^3, sp^2, sp or aromatic, have the same atomic polarizability.

The values of these average atomic polarizabilities decrease in the following order: I, Br, S, P, Cl, C, N, O, F, and H. Thus, the elements with large covalent radius and low electronegativity have the highest average atomic polarizabilities, while a high electronegativity and a small covalent radius result in a low polarizability. These results are in agreement with the general and qualitative explanation of polarizability.

The literature shows several sets of values of atomic, structural, and group contributions to molar refraction. Some of these are rather old, such as those proposed by Einselhor[27] and Vogel,[28] derived from experimental determinations of the molar refraction of several organic compounds. In these
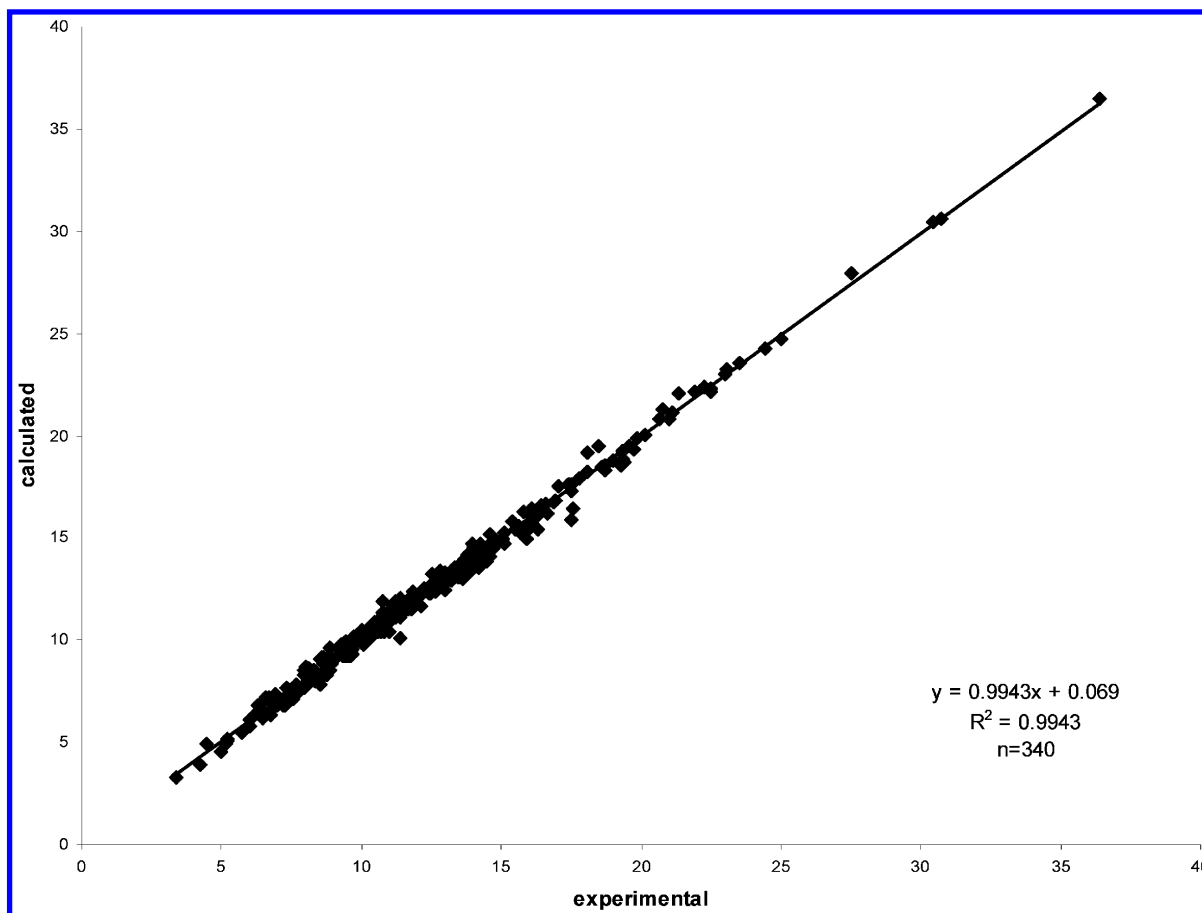
Figure 1. Plot of calculated vs experimental polarizability of the working set with the corresponding regression equation.
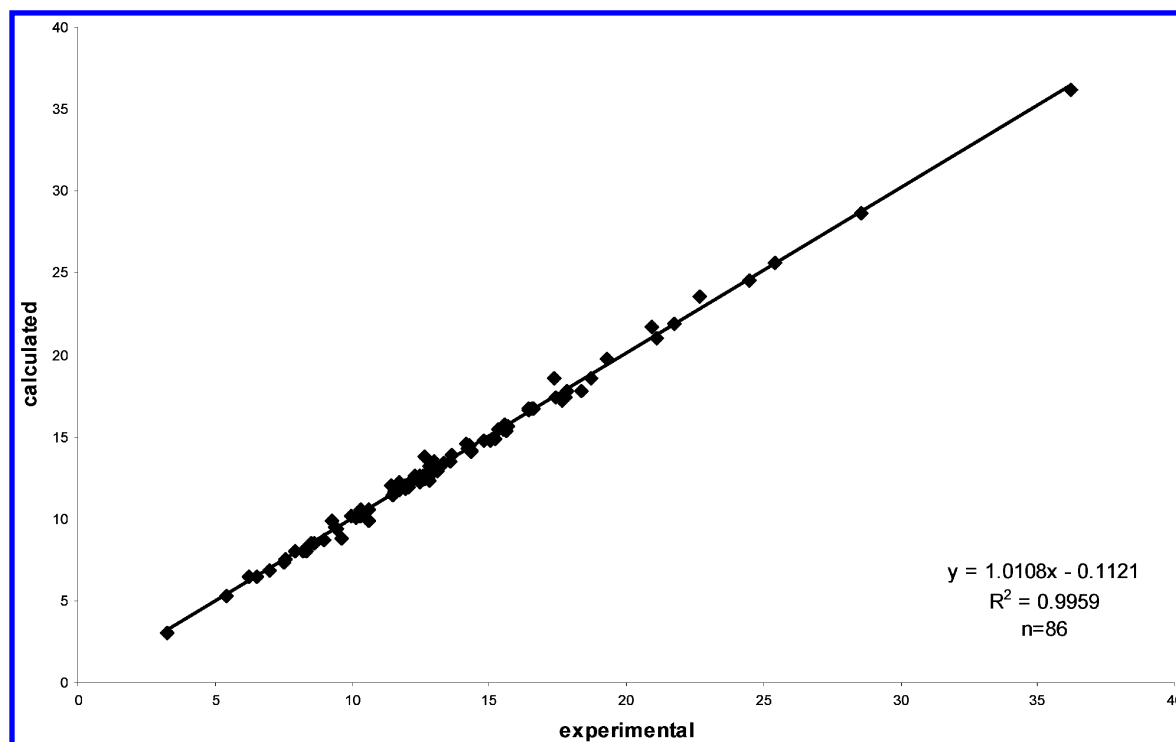

Figure 2. Plot of calculated vs experimental polarizability of the prediction set with the corresponding regression equation.

methods, different values are assigned to an element according to its chemical and structural characteristics, i.e., the nitrogen atoms have different values for primary, secondary, or tertiary amine, nitriles, etc. Moreover, the structural units

such as double and triple bonds, rings of different number of carbon atoms, etc. also make their own contributions to the molar refraction of the compound. Recently, atomic contributions to the molecular polarizability tensor and the

POLARIZABILITIES OF SOLVENTS

J. Chem. Inf. Comput. Sci., Vol. 42, No. 5, 2002 **1161**

Table 5. Average Atomic Polarizabilities (Å^3) of the Elements Derived by Different Methods

element	this paper	ref 27	ref 28	ref 9	ref 10
C	1.51	0.96	1.03	0.74	1.86
N	1.03	1.30	1.24	0.52	1.75
O	0.57	0.71	0.65	0.45	0.50
F	0.22	0.38	0.32	0.32	0.03
H	0.17	0.44	0.41	0.15	~0
Cl	2.16	2.37	2.32	1.91	
Br	3.29	3.45	3.46	2.88	
I	5.45	5.51	5.53	4.69	
S	2.99	3.09	3.10		
P	2.48				

global atomic polarizability have been proposed,[10,11] derived from theoretical calculations of the polarizability of sets of organic compounds containing only 30 and 40 molecules, respectively. In these models, only the contributions of the atoms are considered, but each element has different atomic polarizabilities according to its chemical and structural characteristics. To compare with our results, we have calculated the average value of the atomic polarizabilities assigned to each type of atom of one element. For example, from the values assigned to O(hydroxyl), O(ether), and O(carbonyl) atoms, the average atomic polarizability of oxygen has been obtained. Table 5 shows these average values and the values proposed in this paper. It is easy to see that the series of values calculated by the different methods are consistent, supporting the idea that the polarizability of the solvents can be predicted with fair precision from the chemical composition of the molecule.

To test the suitability of semiempirical quantum calculations in the estimation of the polarizability, we used AM1, PM3, and MNDO methods. The alpha values obtained from these methods for the 426 solvents gave very high average absolute relative errors, 34.7%, 39.1%, and 36.4%, respectively. In all instances the calculated values are lower than the experimental ones, although good cor-

relations with the experimental values are obtained for all of them.

AM1:
$$\alpha_{calc} = 0.70\alpha_{exp} - 0.54; \text{R}^2 = 0.9308; \text{F} = 5707; \text{SD} = 1.23$$

PM3:
$$\alpha_{calc} = 0.64\alpha_{exp} - 0.33; \text{R}^2 = 0.9378; \text{F} = 6319; \text{SD} = 1.16$$

MNDO:
$$\alpha_{calc} = 0.70\alpha_{exp} - 0.69; \text{R}^2 = 0.9279; \text{F} = 5457; \text{SD} = 1.26$$

The high errors obtained in the semiempirical calculations of the polarizability could be attributed to the fact that no polarizability data have been used for deriving the semiempirical atomic parameters. On the other hand, the tendency in the accuracy of the different methods is similar to those arising in the calculation of other molecular properties, such as dipole moments.[29]

Density functional theory methods were also used in the estimation of the polarizability; only 22 solvents were studied, due to the time needed to perform calculations. Table 6 shows the results obtained with the 22 solvents studied. As expected, the DFT method gives lower average absolute relative error, 12.88%, than the other quantum methods used, because the former gives a better description of the wave function. There is a good correlation with the experimental values

DFT:
$$\alpha_{calc} = 1.01\alpha_{exp} - 1.21; \text{R}^2 = 0.9790; \text{F} = 934; \text{SD} = 0.74$$

In all the systems studied the calculated polarizability is lower than the experimental. In this respect, it should be noted that in this kind of calculation, both DFT and semiempirical, the intermolecular interactions are not taken into account. Moreover, the calculated α values are those corresponding to 0 K.

Table 6. Experimental and Calculated Values by Different Methods of a Subset of 22 Solvents

solvent	exptl	additive	error[a]	DFT	error	AM1	error[a]	PM3	error[a]	MNDO	error[a]
1,1,1,-trichloroethane	10.44	10.34	0.96	9.18	12.07	5.31	49.14	5.94	43.10	5.24	49.81
2,6-di-*tert*-butylpyridine	24.98	24.72	1.04	23.85	4.52	16.79	32.79	15.17	39.27	16.56	33.71
acetic acid	5.18	5.15	0.58	3.96	23.55	3.26	37.07	3.04	41.31	3.1	40.15
acetone	6.47	6.46	0.15	5.93	8.35	3.99	38.33	3.63	43.89	3.86	40.34
acetonitrile	4.44	4.91	10.59	3.11	29.95	2.8	36.94	2.76	37.84	2.67	39.86
allyl alcohol	6.71	6.46	3.73	6.39	4.77	4.28	36.21	3.87	42.32	4.13	38.45
benzaldhyde	12.7	12.54	1.26	10.51	17.24	9.16	27.87	8.64	31.97	9.09	28.43
1-butanol	8.79	8.67	1.37	8.15	7.28	5.43	38.23	4.78	45.62	5.21	40.73
carbon disulfide	8.51	7.82	8.11	6.21	27.03	5.92	30.43	6.53	23.27	5.43	36.19
cyclohexane	11.04	11.49	4.08	10.26	7.07	7.11	35.60	6.19	43.93	6.88	37.68
diethyl ether	8.98	8.67	3.45	8.24	8.24	5.68	36.75	4.97	44.65	5.41	39.76
dimethyl sulfide	7.63	7.38	3.28	6.77	11.27	4.3	43.64	3.96	48.10	3.94	48.36
diphenyl ether	21.02	20.82	0.95	21.3	1.33	16.37	22.12	15.9	24.36	17.21	18.13
ethanol	5.13	4.94	3.70	3.97	22.61	3.01	41.33	2.68	47.76	2.87	44.05
hexafluorobenzene	10.56	10.74	1.70	10.25	2.94	9.39	11.08	8.13	23.01	8.91	15.63
methyl acetate	7	7.02	0.29	6.46	7.71	4.66	33.43	4.24	39.43	4.47	36.14
N,N-dimethylformamide	7.93	7.68	3.15	6.34	20.05	5.26	33.67	4.94	37.70	5.02	36.70
nitrobenzene	13	12.45	4.23	10.63	18.23	9.63	25.92	8.87	31.77	9.72	25.23
perfluoro-*n*-hexane	12.71	12.5	1.65	12.27	3.46	10.56	16.92	7.48	41.15	9.47	25.49
pyridine	9.58	9.81	2.40	7.55	21.19	6.75	29.54	6.45	32.67	6.69	30.17
tetrahydrofuran	7.97	8.32	4.39	7.39	7.28	5.11	35.88	4.5	43.54	4.92	38.27
trimethyl phosphate	11.07	11.18	0.99	9.16	17.25	6.73	39.21	7.74	30.08	8.38	24.30
av absolute relative error			2.82		12.88		33.28		38.03		34.89

[a] Absolute relative error: absolute value of [(calcd − exptl)/exp].

Table 7. One-Descriptor Correlations for All the Solvents

descriptor	R^2	F	s	RSD (%)	error[a] (%)
HOMO−LUMO gap	0.0588	26.5	4.54	36.35	35.93
molecular weight	0.4606	362.1	3.44	27.52	18.82
number of atoms	0.7277	1133	2.43	19.46	17.45
number of bonds	0.7486	1262	2.35	18.78	16.56
molecular volume	0.8970	3693	1.50	12.03	9.03
set-5[b]	0.9272	1069	1.27	10.15	7.04
set-4[b]	0.9149	1131	1.37	10.97	7.44
set-3[b]	0.9105	1430	1.40	11.24	7.7

[a] Average absolute relative error. [b] See text.

Once the atomic polarizabilities had been determined, a second study was undertaken in order to ascertain if a quantitative relationship could be established between the experimental polarizability and some of the properties to which they have been qualitatively related. These properties include the HOMO−LUMO energy gap, the molecular weight, the number of atoms and bonds, and the molecular volume. Table 7 shows the obtained results for the complete set of solvents studied. The best results are found with the molecular volume. The values of the molecular volumes were calculated[30] by applying a 3D grid of cubes in the parallelepiped box with the dimensions X_{max}, Y_{max}, and Z_{max} containing the molecule and by summation of the volume of the cubes overlapped with atomic spheres. The factorized molecular volume is calculated as the ratio of the molecular volume and the volume defined by X_{max}, Y_{max}, and Z_{max}. The average absolute relative error found is of 9.03%. The other properties, molecular weight, number of atoms, and number of bonds, give poorer correlations. It is worth noting that the HOMO−LUMO energy gap, calculated at AM1 level, is not correlated with the polarizability. The last rows of Table 7 give the parameters of the correlation of experimental polarizability with the five previous descriptors (set-5), without the HOMO−LUMO energy gap (set-4), and with only the last three descriptors (set-3). These results show that the addition of number of atoms and number of bonds to the molecular volume slightly improves the correlation in relation to the one obtained with the molecular volume only.

CONCLUSIONS

The results obtained clearly demonstrate the validity of an additive model in the polarizability calculation of the solvents. The proposed model allows the calculation of the polarizability from the chemical composition alone, that is, from the number of atoms present of each element. Thus, an "average atomic polarizability" is assigned to each element independently of its chemical and structural characteristics. Contrary to the classic additive methods, it is not necessary to consider the existence of structural units such as number of double and triple bonds, number and size of rings present, etc.

The values of the average atomic polarizability reflect the characteristics of the elements that affect the polarizability of the molecule. Thus, the elements with high average atomic polarizabilities are those with large covalent radius and low electronegativity. The average atomic polarizability values are consistent with the average of the atomic polarizabilities derived from other experimental determinations or theoretical calculations.

The polarizability of the solvents, expressed in Å^3, can be calculated from the following equation

$$\alpha_{calc} = 1.51\#C + 0.17\#H + 0.57\#O + 1.05\#N + 2.99\#S + 2.48\#P + 0.22\#F + 2.16\#Cl + 3.29\#Br + 5.45\#I + 0.32$$

The intercept has a value of 0.32 that is forty times lower than the average value of the experimental polarizability, showing that it does not have any significance in the calculation of α. The average absolute relative errors of the working set and the prediction set, 2.31% and 1.92% respectively, are within the experimental errors. The regression equations of α_{calc} vs α_{exp} of these two sets of solvents support the goodness of the model.

The correlations of the α_{exp} versus several molecular physicochemical properties qualitatively related to the molecular polarizability (molecular volume, molecular weight, number of bonds and atoms, HOMO−LUMO energy gap) show that only the molecular volume, calculated from a geometric algorithm, has a significant R^2 of 0.89, while the average absolute relative error is about 9%. The HOMO−LUMO energy gap, which is usually considered to be related to the polarizability, does not have any quantitative correlation with this magnitude. The use of quantum methods in the calculation of the polarizability produces worse results that our additive model. The obtained results with the 426 solvents by the semiempirical methods, AM1, PM3, and MNDO, give very high average absolute relative errors, about 35−40%; although the correlation with the experimental values are good, $R^2 \approx 0.93$. In all the cases examined, the calculated values are significantly lower than the experimental ones. The DFT approach has here been applied only to a short set of 22 solvents due to the great amount of time needed to carry out this method. DFT gives better results than the other quantum methods but poorer results than those obtained from the additive model.

ACKNOWLEDGMENT

REFERENCES AND NOTES

(1) Marcus, Y. *Ion Solvation*; Wiley: 1985.
(2) Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches*; VCH: 1993.
(3) Abraham, M. H. Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biomedical Processes. *Chem. Soc. Rev.* **1993**, 73−83.
(4) Xing, L.; Glen, R. C. Novel Methods for the Prediction of logP, pK_a, and logD. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 796−805.
(5) Yamaguchi, Y.; Osamura, Y.; Goddard, J. D.; Schaefer, H. F., III. *A New Dimension to Quantum Chemsitry, Analytical Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; Oxford University Press: Oxford, New York, 1994.
(6) Ma, B.; Lii, J. H.; Allinger, N. L. Molecular Polarizabilities and Induced Dipole Moments in Molecular Mechanics. *J. Comput. Chem.* **2000**, *21*, 813−825.
(7) Partington, J. R. *An Advanced Treatise on Physical Chemistry*; Longmans: London, 1953; Vol 4, p 42 ff.
(8) Miller, K. J. Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, *112*, 8533−8542.
(9) Applequist, J.; Carl, J. R.; Fung, K. K. An Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atomic Polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952−2960.
(10) Stout, J. M.; Dykstra, C. E. Static Dipole Polarizabilities of Organic Molecules. Ab Initio Calculations and a Predictive Model. *J. Am. Chem. Soc.* **1995**, *117*, 5127−5132.

POLARIZABILITIES OF SOLVENTS

J. Chem. Inf. Comput. Sci., Vol. 42, No. 5, 2002 **1163**

(11) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press: Wiley: Letchworth, England, 1986.

(12) Hall, L. B.; Aaserud, D. Structure−Activity Models for Molar Refraction of Alkylsilanes Based on Molecular Connectivity. *Quant. Struct.-Act. Relat.* **1989**, *8*, 296−304.

(13) Katritzky, A. R.; Sild, S.; Karelson, M. General Quantitative Structure−Property Relationship Treatment of the Refractive Index of Organic Compounds. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 840−844.

(14) Cocchi, M.; Benedetti, P. G.; Seeber, R.; Tassi, L.; Ulrici, A. Development of Quantitative Structure−Property Relationships Using Calculated Descriptors for the Prediction of the Physicochemical Properties (n_D, ρ, bp, η) of a Series of Organic Solvents. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1190−1203.

(15) Katritzky, A. R.; Sild, S.; Karelson, M. Correlation Prediction of the Refractive Indices by QSPR. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1171−1176.

(16) Gakh, A. A.; Gakh, E. G.; Sumpter, B. G.; Noid, D. W. Neural Network-Graph Theory Approach to the Prediction of the Physical Properties of Organic Compunds. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 832−839.

(17) Liu, S.; Zhang, R.; Liu, M.; Hu, Z. Neural-Network Topological Indices Approach to the Prediction of Properties of Alkene. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1146−1151.

(18) Marcus, Y. *The Properties of Solvents*; Wiley: 1998; p 95 ff.

(19) Abboud, J.-L. M.; Notario, R. Critical compilation of scales of solvent parameters. Part I. Pure, non-hydrogen bond donor solvents. *Pure Appl. Chem.* **1999**, *71*, 645−718.

(20) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; VCH: Weinheim, 1990.

(21) Stewart, J. J. P. *MOPAC 6.0*; QCPE No. 455, 1989.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewsi, V. G.; Montgomery, J. A., jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Menucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98, revision A.11*; Gaussian, Inc.: Pittsburgh, PA, 2001.

(23) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange *J. Chem. Phys.* **1993**, *98*, 5648−5652. Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti correlation formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785−789. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(24) Hariharan, P. C.; Pople, J. A. Influence of polarization functions on MO hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213−222.

(25) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second row elements. *J. Chem. Phys.* **1982**, *77*, 3654−3665.

(26) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleer, P.v. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements. *J. Comput. Chem.* **1983**, *4*, 294−301.

(27) Eisenlohr, F. A New Calculation of Atomic Refractions. *Z. Physik. Chem.* **1911**, *75*, 585−607.

(28) Vogel, A. I. Physical Properties and Chemical Constitution, XXIII. Miscellaneous Compounds, Investigation of the so-called Coordinate or Dative Link in Esters of Oxyacids and in Nitroparaffins by Molecular Refraction Determinations. Atomic, Structure and Refractivities. *J. Chem. Soc.* **1948**, 1833−1854.

(29) Boyd, D. B. Aspects of Molecular Modelling. *Rev. Comput. Chem.* **1990**, *1*, 321−354.

(30) Katriztky, A. R.; Lovanov, V. S.; Karelson, M. *CODESSA Reference Manual, Version 2.0*; University of Florida: 1996.