J. Phys. Chem. A **1999,** *103,* 11366−11377

Scaling Factors for the Prediction of the Frequencies of the Ring Modes in Benzene Derivatives

Mauricio Alcolea Palafox*

Departamento de Química-Física I (Espectroscopia), Facultad de Ciencias Químicas, Universidad Complutense, Madrid 28040 ES, Spain

Received: February 17, 1999; In Final Form: October 13, 1999

The performance of semiempirical, ab initio, and density functional methods in calculating and describing the vibrational frequencies of benzene and several derivatives was determined. Different levels were used. The normal modes of the ring were characterized by the magnitude and direction of the displacement vector. Two new procedures of scaling the frequencies were presented. Scaling factors were determined at different levels. A significant reduction in the predicted frequencies of the ring modes of several benzene derivatives was obtained over the one-factor overall scaling procedure.

Introduction

The computation of the vibrational spectrum of a polyatomic molecule of even modest size is lengthy. Despite the tremendous advances made both in theoretical methods and computer hardware, the most accurate of the quantum chemical methods are still too expensive and cumbersome to apply as routine research. Thus, one may be forced to work at a low level, and consequently, one must expect a considerable overestimation of the calculated vibrational frequencies. This overestimation can be remarkably reduced with the use of transferable empirical parameters for the force fields and calculated frequencies. Also, the difference between the computed and the experimental frequencies may be due to many different factors that are usually not even considered in the theory, such as anharmonicity, errors in the computed geometry, Fermi resonance, etc. Even solvent effects may lead to systematic differences between computed and observed frequencies. The introduction of scaling factors is also capable of accounting for all these various effects.

The values of the force constants and the frequencies are close among similar molecules and characteristic groups. Based on this assumption, two procedures are frequently used for an accurate scaling.

One possible approach involves the rescaling of the force constant matrix. A simple example is the proposal of Pulay and Meyer[1] to multiply all diagonal stretching force constants by 0.9 and to multiply all diagonal bending force constants by 0.8. This procedure will certainly improve the agreement between computed and experimental frequencies since the computed values are usually too high by 5−10%. The underlying assumption by Pulay is similar to ours, namely that it is possible to transfer scaling factors among similar molecules.

We believe that our introduction of scaling factors for the frequencies themselves is preferable to the more indirect approach of scaling the force constants.[2] First, the frequencies are physically observable and the force constants are not. Second, the definition of the force constant matrix is not unique because the number of matrix elements is usually much larger than the number of known frequencies. The third argument in favor of our procedure is the pragmatic observation that we could be primarily interested in a few major characteristic frequencies only and not in the whole spectrum.

The scaling of the frequencies is in general only carried out with the use of a unique scale factor, which is common for the same level of computation.[3] However, we find that it is not advisable to use a uniform correction factor for all vibrational modes and molecules and it is preferable instead to introduce different correction factors for different types of vibrational modes and/or related molecules. Thus we show in the present paper two new procedures for scaling, i.e., with a scaling equation and with specific scale factors for each mode determined in a simpler and similar molecule. Following these procedures may improve the accuracy of the computed frequencies by more than a 20% over the one-factor overall scaling procedure. In addition, our procedure leads to easier and more precise prediction for specific characteristic frequencies that are of special interest. As a first study, the present paper shows the work on the benzene molecule, from which are extracted scaling factors and scaling equations to be transferred and used in the interpretation of the ring modes in benzene derivatives. The performance of these procedures of scaling is analyzed in eight examples, the molecules of phenylsilane, aniline, benzoic acid, phenothiazine, *p*-aminobenzoic acid, *p*-methoxybenzoic acid, 1,4-dicyanobenzene, and 2,4-difluorobenzonitrile.

Computational Methods

Semiempirical methods were employed because they are potentially attractive for the computation of vibrational frequencies of large benzene derivatives, for their inherent low computational cost. Among these methods the MNDO, AM1, and PM3 were selected, which are included in the GAUSSIAN 94[4] and in the AMPAC 5.0[5] program packages. The most modern one, SAM1,[6] is included only in AMPAC 5.0. The keywords OPT and FREQ were utilized for optimization and frequency calculations, respectively, with the GAUSSIAN 94, while the PRECISE and FORCE keywords correspond to the AMPAC 5.0. In the optimization was used the TIGHT criterion of convergence with GAUSSIAN 94 and the GNORM=0.001 with AMPAC 5.0. The frequency calculations were carried out at the same level of the respective optimization process, and with their final results.

10.1021/jp990591j CCC: $18.00

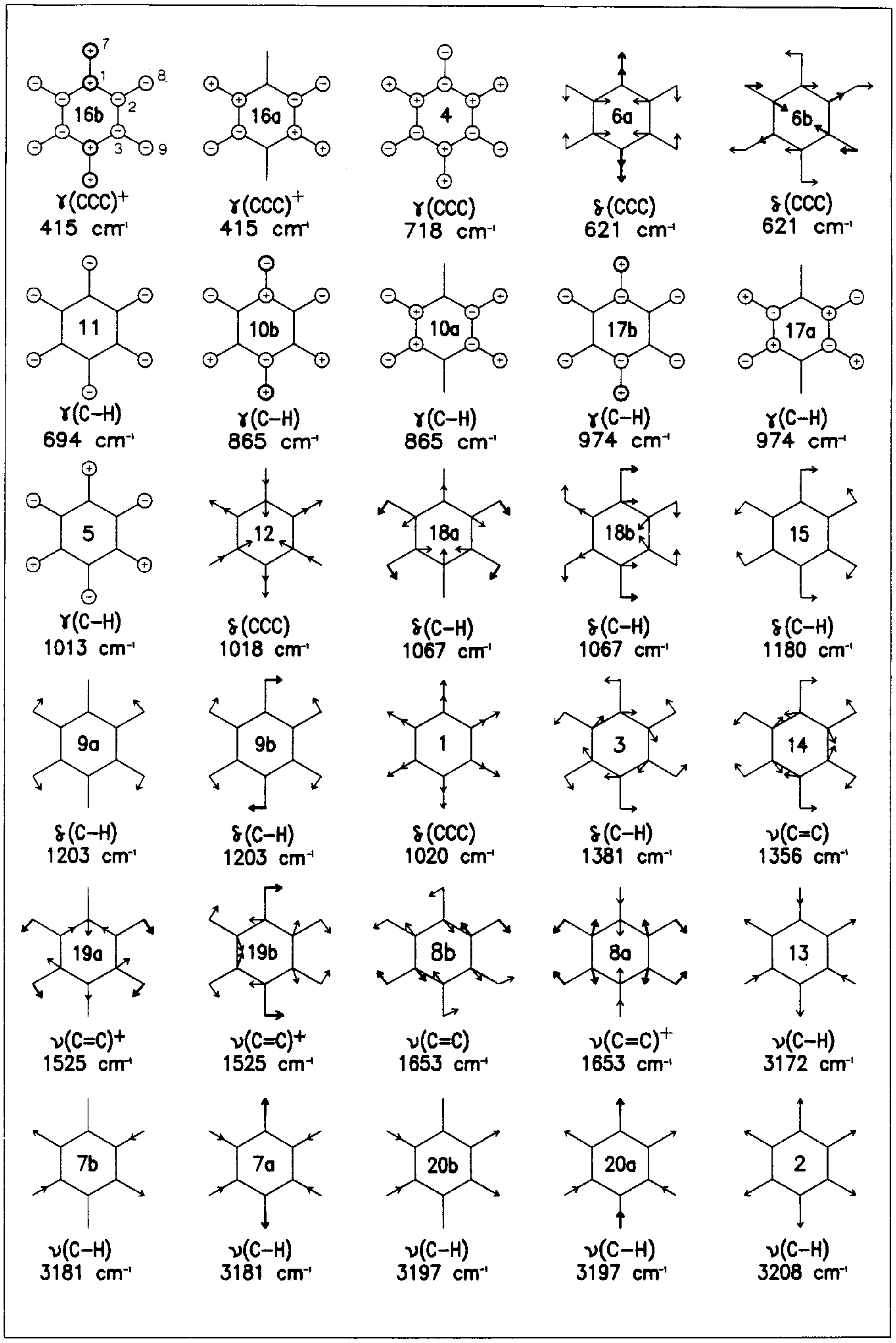

Figure 1. Characterization of the ring modes in benzene molecule at the B3LYP/6-31G** level with their calculated frequencies. $^+$The motions appear as (CCC + CH).

Ab initio calculations with wave function based HF and MP2, and density functional theory (DFT) methods were only performed with GAUSSIAN 94. Several basis sets were used, but the 6-31G** was established to initiate the comparisons. The results obtained with this basis set will be referred to by the short-hand notations, HF, MP2, B3LYP, etc.

In the DFT methods, the SLYP and SVWN with exchange zero[7] and the Becke exchange functional[8] (B) and Becke's three-parameter exchange functional[9] (B3) with correlation and exchange 1, were utilized. The B and B3 exchange functionals were used with the correlation functionals of Lee, Yang, and Parr[10] (LYP), P86,[11] and PW91.[12]

Results and Discussion

Characterization of the Ring Modes. The benzene ring modes characterized at the B3LYP/6-31G** level are plotted in Figure 1, while their calculated frequencies are shown in Table 1. The discrepancies observed with other methods are plotted in Figure 2. The motions are represented when the vector corresponding to the atomic displacement for each atom and computed frequency is ≥0.07 for the carbon atoms and ≥0.15 for the hydrogen atoms. Large displacements are indicated by the dark arrows and circles. In the literature,[13] the ring modes appear characterized by the direction of the displacement vector.

TABLE 1: Calculated with the 6-31G Basis Set and Experimental Frequencies for the Normal Ring Modes of Benzene Molecule**

ring mode			semiempirical				ab initio		density functional methods							infrared		
Wilson no.	Herzberg no.	symmetry	MNDO	AM1	PM3	SAM1	HF	MP2	SVWN	SLYP	BP86	BLYP	B3P86	B3LYP	B3PW91	liq[a]	est[b]	gas[c]
1	2	a_{1g}	1197	1276	1230	1270	1083	1028	1025	1024	994	986	1029	1020	1026	993	1008	993.1
2	1	a_{1g}	3423	3205	3085	2908	3373	3283	3143	3105	3128	3123	3224	3208	3218	3073	3208	3073.9
3	3	a_{2g}	1372	1328	1225	1202	1501	1396	1312	1290	1329	1342	1374	1381	1375	1350	1390	1350
4	8	b_{2g}	598	618	619	600	778	500	708	704	695	696	720	718	719	707	718	707
5	7	b_{2g}	1075	1012	1024	852	1135	919	977	965	968	974	1014	1013	1015	990	1011	990
6	18	e_{2g}	613	648	620	615	665	618	601	595	600	605	618	621	618	606	613	608.1
7	15	e_{2g}	3407	3188	3060	2871	3343	3258	3120	3082	3102	3096	3198	3181	3192	3056	3191	3056.7
8	16	e_{2g}	1719	1768	1785	1753	1795	1675	1644	1641	1601	1589	1667	1653	1664	1599	1639	1601.0
9	17	e_{2g}	1252	1222	1147	1151	1288	1224	1161	1144	1167	1171	1202	1203	1203	1178	1192	1177.8
10	11	e_{1g}	917	891	849	914	961	854	833	823	830	835	864	865	866	846	866	847.1
11	4	a_{2u}	771	744	712	771	764	689	665	657	666	674	691	694	694	673	686	674.0
12	6	b_{1u}	957	1028	976	998	1096	1020	981	970	984	991	1014	1018	1013	1010	1024	1010
13	5	b_{1u}	3402	3184	3055	2862	3332	3248	3111	3073	3092	3087	3188	3172	3182	3057	3172	3057
14	9	b_{2u}	1277	1368	1303	1365	1350	1467	1423	1431	1354	1328	1383	1356	1378	1309	1318	1309.4
15	10	b_{2u}	1164	1179	1153	1115	1192	1201	1135	1115	1146	1154	1177	1180	1178	1146	1167	1149.7
16	20	e_{2u}	368	371	356	382	453	395	399	396	397	401	412	415	412	398	407	398
17	19	e_{2u}	1037	989	978	892	1102	916	935	920	930	933	976	974	978	967	989	967
18	14	e_{1u}	1147	1146	1068	1088	1139	1082	1050	1041	1037	1034	1072	1067	1071	1037	1058	1038.3
19	13	e_{1u}	1555	1579	1547	1546	1646	1541	1481	1469	1474	1480	1525	1525	1524	1482	1512	1484.0
20	12	e_{1u}	3415	3195	3073	2889	3361	3274	3134	3096	3117	3111	3213	3197	3208	3064	3191	3064.4

[a] In the liquid state, ref 13. [b] Estimated harmonic frequencies, ref 14. [c] In the gas phase, ref 15.

In the present discussion, they are also identified by the value of this displacement vector. For this purpose, Table 2 shows these values of the displacement vector by AM1, HF, MP2, and B3LYP methods. For symmetry considerations of the ring structure, the value of the displacement is only given in the carbon and hydrogen atoms numbered in the first ring of Figure 1. The identification and description of the different modes by the various methods used are established with respect to Figure 1, which is selected as a reference system because their scaled results are the nearest to the experimental data and are as follows:

Modes with the Same Characterization by All the Theoretical Methods. Modes 1, 2, 3, 11, 12, and 13 show similar values of the displacement vector by all the theoretical methods, Table 2. In mode 1, however, this fact is not associated with a similar calculated frequency, the methods differing greatly, until ca. 300 cm^{-1}. In benzene derivatives this mode 1 appears notably deformed.

In modes 2, 11, and 13 the displacements of all the C are very short (0.03−0.04) thus they are not plotted in Figure 1, although they appear in the literature.[13] In modes 2 and 13 a slight difference between the values of the displacement of the H atoms appear with BLYP.

Modes with Differences in the Characterization by the Theoretical Methods. (a) Tangential Vibrations. In the characterization of mode 8, by our theoretical methods, remarkable differences with its definition in the classical literature appear.[13] Thus, the hydrogen atoms move in opposite direction to the carbon atoms, when they appear to move in the same direction. Moreover, in 8a all the atoms have large displacements, when in its definition two pairs of C−H atoms have no displacement. These features give rise to small differences between the a and b modes, which are mainly in the magnitude of the displacement, with a C−H pair shorter (ca. twice) than the other two pairs of C−H atoms in mode 8a, and larger (ca. 50%) in 8b, Table 2. Differences also appear with the different methods, some representative examples of the discrepancies with Figure 1 are shown in Figure 2.

The C and the H of the CH bonds of mode 14 are described by all the semiempirical methods with the same direction of displacement vector, Figure 2. However by HF, MP2, and DFT methods they appear in opposite directions (like mode 15) and are in contrast to that reported in the literature;[13] i.e., the directions of modes 14 and 15 are interchanged. Thus by MP2 and DFT methods the only difference between modes 14 and 15 is in the value of the displacement, which is more C−C in mode 14 (between 0.29 and 0.39 for the C, and 0.14 and 0.29 for the H) and more C−H in mode 15 (0.01−0.03 for the C and 0.41 for the H). In HF, with very short (0.10) displacement of the C in mode 14, and slightly long (0.14) in mode 15, in contrast to the characterization of mode 14 as ν(C═C) and mode 15 as δ(C−H), Figure 2, mode 14 corresponds to mode 15 and vice versa. By MNDO, PM3, and SAM1 the displacements of the H atoms are so large, 0.40−0.36, and the C atoms so short, 0.07−0.17, that mode 14 is represented as a δ(C−H) mode instead of ν(C═C).

In mode 19 the direction of the displacement of the H is in the opposite direction to the C while in mode 18 the motion of the H is in the same direction as the C, both in contrast to that reported in the literature,[13] in which they are in the same direction (mode 19) or in the opposite (mode 18). Thus mode 19 resembles mode 18 and vice versa. However, in the value of the displacement they can be clearly identified, with greater values in the C (ca. 4 times with the semiempirical methods) for mode 19, a C═C mode, than for mode 18, a C−H mode. The main difference between the a and b forms is in the value and direction of the displacement of the H, e.g., with two displacements large (ca. 0.50) and one small (ca. 0.15) in 18a, and with one displacement large (ca. 0.56) and two medium (ca. 0.30) in 18b. Small differences are also noted by the different methods.

Forms 9a and 9b differ only in the magnitude of the displacement of the first pair of H, which is in general very small or insignificant in a and very large in b. However, in MP2 the values are coincident and thus a and b forms are indistinguishable.

(b) Radial Vibrations. Modes 6a and 6b represent bending of the C and H atoms that appear more deformed than defined in the literature,[13] in exactly a straight line toward the C−H bond. Depending on the degree of deformation, the characterization by the different methods varies slightly. Figure 2 shows an example with MP2 and SVWN. Our characterization of mode

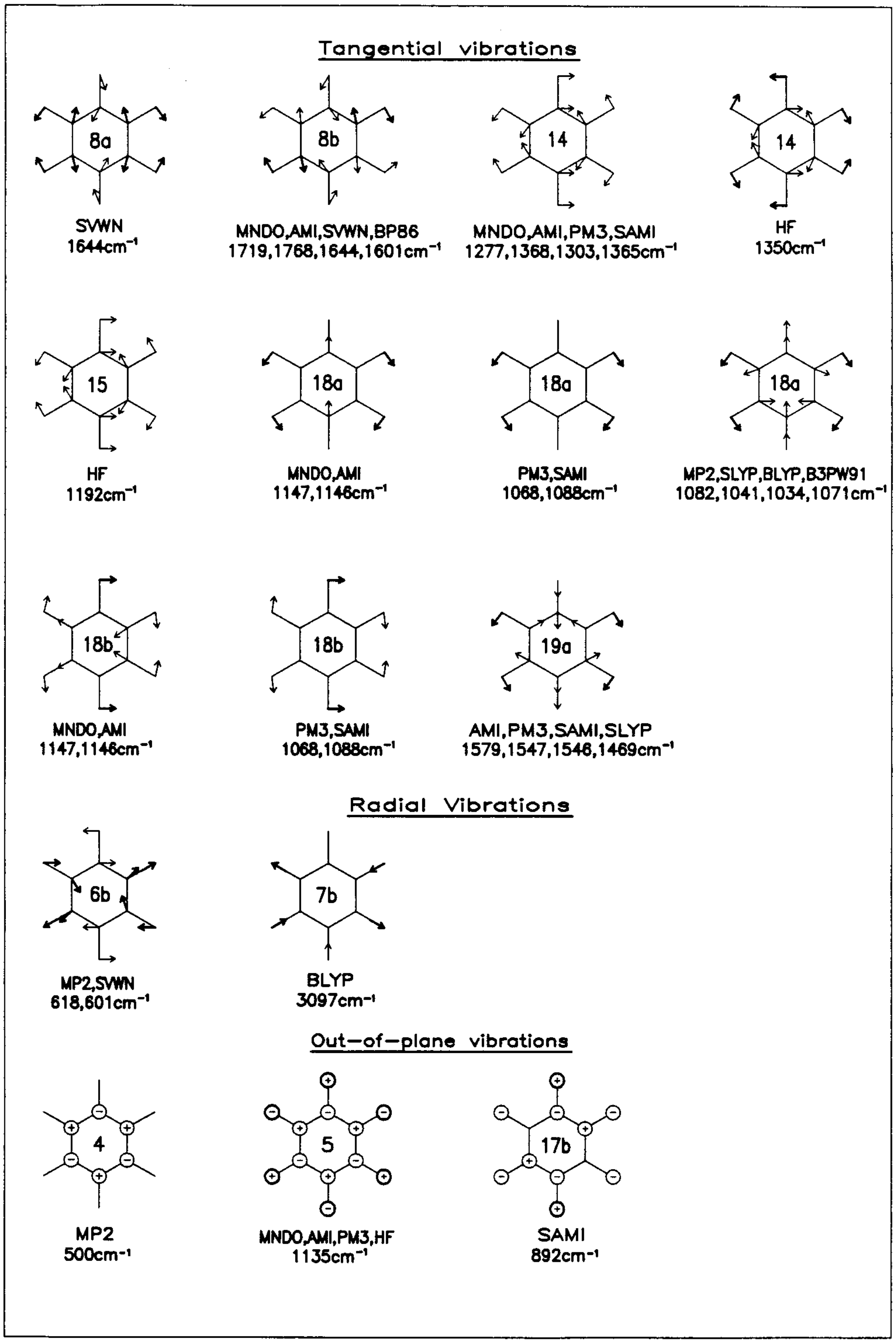

Figure 2. Specific characterization of several ring modes by the theoretical methods, with their respective calculated frequencies.

6a with MP2 differs in the direction of the displacement vector of the H atoms from that reported in Figure 1 of ref 14, together with other discrepancies in some of the C atoms of mode 18a.

In mode 7 the displacements of the C are very short (between 0.01 and 0.05) with all the methods, and they are not shown in Figure 1. Slight differences appear with SLYP and SVWN; however with PM3, MP2, and the other DFT methods, the three displacements are different, the differences decreasing between a and b forms, especially with B3PW91.

In mode 20 the displacements of all the C are very short (ca. 0.04) with all the methods, thus they are not plotted in Figure 1, although they appear in the literature.[13] In 20a one of the displacements of the H is great (ca. 0.57) while the other two H have almost half that value (ca. 0.29) by AM1, HF, SLYP, and SVWN. With the other methods the three displacements are different.

(c) Out-of-Plane Vibrations. Modes 4 and 5 are badly characterized by MP2, which leads to a very great error in their calculated and scaled frequencies. For example, in mode 4 the displacement of the H is very short (0.07) at MP2, but it is very long (0.35−0.39) with the other methods; and in the displacement of the C it is very high (0.40), while it is relatively low (0.11) (AM1, PM3) or medium (0.18) in the rest of the methods.

TABLE 2: Values of the Displacement Vector Obtained in the Wavenumbers of Several Ring Modes by AM1, HF, MP2, and B3LYP Methods and with the 6-31G Basis Set**[a]

Wilson no.	AM1						HF						MP2						B3LYP					
	carbons			hydrogens			carbons			hydrogens			carbons			hydrogens			carbons			hydrogens		
	1	2	3	7	8	9	1	2	3	7	8	9	1	2	3	7	8	9	1	2	3	7	8	9
1	0.28	0.28	0.28	0.30	0.30	0.30	0.28	0.28	0.28	0.30	0.30	0.30	0.27	0.28	0.28	0.30	0.30	0.30	0.28	0.28	0.28	0.30	0.30	0.30
2	0.04	0.04	0.04	0.40	0.40	0.41	0.04	0.04	0.04	0.41	0.40	0.40	0.04	0.04	0.04	0.41	0.40	0.40	0.04	0.04	0.04	0.40	0.40	0.41
3	0.06	0.06	0.06	0.40	0.40	0.40	0.06	0.06	0.06	0.40	0.40	0.40	0.06	0.06	0.06	0.40	0.40	0.40	0.06	0.06	0.06	0.40	0.40	0.40
4	0.12	0.12	0.12	0.39	0.39	0.39	0.18	0.18	0.18	0.36	0.36	0.36	0.40	0.40	0.40	0.07	0.07	0.07	0.21	0.21	0.21	0.35	0.35	0.35
5	0.11	0.11	0.11	0.39	0.39	0.39	0.07	0.07	0.07	0.40	0.40	0.40	0.01	0.01	0.01	0.41	0.41	0.41	0.06	0.06	0.06	0.40	0.40	0.41
6a	0.38	0.23	0.23	0.36	0.24	0.24	0.36	0.22	0.22	0.36	0.27	0.27	0.35	0.18	0.27	0.35	0.24	0.30	0.36	0.21	0.23	0.36	0.26	0.27
6b	0.15	0.34	0.34	0.19	0.33	0.33	0.16	0.32	0.32	0.22	0.34	0.34	0.17	0.28	0.35	0.24	0.30	0.35	0.15	0.31	0.32	0.22	0.32	0.34
7a	0.04	0.02	0.02	0.57	0.29	0.29	0.05	0.02	0.02	0.58	0.28	0.29	0.04	0.02	0.03	0.56	0.19	0.37	0.05	0.01	0.04	0.57	0.20	0.37
7b	0	0.04	0.04	0	0.50	0.50	0	0.04	0.04	0	0.50	0.50	0.01	0.04	0.05	0.10	0.44	0.54	0.01	0.04	0.04	0.09	0.45	0.54
8a	0.19	0.44	0.44	0.15	0.15	0.15	0.15	0.29	0.29	0.17	0.38	0.38	0.15	0.29	0.31	0.17	0.36	0.37	0.15	0.30	0.30	0.17	0.37	0.37
8b	0.30	0.50	0.30	0.15	0.15	0.15	0.21	0.32	0.21	0.26	0.42	0.26	0.20	0.33	0.22	0.24	0.41	0.27	0.33	0.22	0.22	0.41	0.25	0.26
9a	0.01	0.02	0.02	0.02	0.50	0.50	0	0.06	0.06	0	0.50	0.50	0.02	0.04	0.06	0.14	0.40	0.56	0	0.05	0.06	0.02	0.48	0.51
9b	0.02	0.02	0.02	0.58	0.29	0.29	0.07	0.03	0.03	0.57	0.29	0.29	0.06	0.02	0.04	0.56	0.15	0.40	0.06	0.03	0.03	0.57	0.26	0.31
10a	0	0.07	0.07	0	0.49	0.49	0.01	0.07	0.08	0.03	0.48	0.51	0	0.07	0.07	0	0.49	0.49	0	0.07	0.07	0	0.49	0.50
10b	0.09	0.04	0.04	0.57	0.29	0.29	0.09	0.04	0.05	0.57	0.26	0.31	0.09	0.04	0.04	0.57	0.29	0.29	0.09	0.04	0.04	0.57	0.29	0.29
11	0.03	0.03	0.03	0.41	0.41	0.41	0.03	0.03	0.03	0.41	0.41	0.41	0.03	0.03	0.03	0.41	0.41	0.41	0.03	0.03	0.03	0.41	0.41	0.41
12	0.30	0.30	0.30	0.28	0.28	0.28	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
13	0.04	0.04	0.03	0.40	0.40	0.40	0.03	0.04	0.04	0.41	0.40	0.40	0.03	0.04	0.04	0.41	0.40	0.40	0.04	0.04	0.03	0.40	0.40	0.41
14	0.24	0.24	0.24	0.33	0.33	0.33	0.10	0.10	0.10	0.39	0.39	0.39	0.34	0.34	0.34	0.22	0.22	0.22	0.29	0.29	0.29	0.29	0.29	0.29
15	0.04	0.04	0.05	0.40	0.40	0.41	0.13	0.14	0.14	0.39	0.38	0.38	0.02	0.02	0.02	0.40	0.40	0.41	0.03	0.03	0.04	0.41	0.40	0.40
16a	0	0.17	0.17	0	0.47	0.47	0	0.21	0.21	0	0.45	0.46	0	0.23	0.23	0	0.44	0.44	0	0.21	0.21	0	0.45	0.45
16b	0.20	0.10	0.10	0.54	0.27	0.27	0.24	0.12	0.12	0.53	0.26	0.26	0.26	0.13	0.13	0.51	0.26	0.26	0.24	0.12	0.12	0.52	0.26	0.26
17a	0	0.11	0.11	0	0.49	0.49	0	0.09	0.09	0	0.49	0.49	0	0.08	0.08	0	0.49	0.49	0	0.09	0.09	0.01	0.49	0.49
17b	0.13	0.06	0.06	0.56	0.28	0.28	0.10	0.05	0.05	0.57	0.28	0.28	0.09	0.05	0.05	0.57	0.28	0.28	0.10	0.05	0.05	0.57	0.28	0.29
18a	0.07	0.05	0.05	0.07	0.50	0.50	0.12	0.09	0.09	0.13	0.46	0.49	0.12	0.09	0.10	0.15	0.50	0.50	0.12	0.09	0.09	0.13	0.46	0.49
18b	0.03	0.07	0.07	0.57	0.29	0.29	0.08	0.11	0.11	0.54	0.27	0.32	0.08	0.11	0.11	0.54	0.24	0.35	0.08	0.11	0.11	0.54	0.27	0.32
19a	0.18	0.22	0.22	0.18	0.41	0.41	0.09	0.14	0.14	0.12	0.47	0.47	0.09	0.13	0.14	0.13	0.44	0.50	0.09	0.14	0.14	0.12	0.47	0.47
19b	0.24	0.20	0.20	0.46	0.28	0.28	0.15	0.11	0.11	0.54	0.28	0.28	0.15	0.10	0.11	0.53	0.24	0.34	0.15	0.11	0.11	0.54	0.28	0.30
20a	0.05	0.02	0.02	0.58	0.29	0.29	0.05	0.02	0.02	0.58	0.29	0.29	0.05	0.02	0.03	0.57	0.22	0.35	0.05	0.02	0.03	0.57	0.24	0.33
20b	0	0.04	0.04	0	0.50	0.50	0	0.04	0.04	0.01	0.50	0.50	0.01	0.04	0.05	0.08	0.45	0.53	0	0.04	0.04	0.05	0.48	0.52

[a] The notation refers to the three first carbon and hydrogen atoms indicated in the first ring of Figure 1.

In modes 10, 16, and 17 the forms a and b differ mainly in the value of the displacement of the first H, being very short or null in a, and very long in b. An exception in mode 17 is SAM1 with different displacements of the three C and H atoms, which leads to a poor calculated frequency.

Frequency Calculations. Table 1 collects the calculated frequencies of the benzene ring modes by our theoretical methods and only with the 6-31G** basis set. An exploration of the basis set was carried out in specific cases, and the results are shown in further tables of the present manuscript. The first and second columns of Table 1 list, respectively, the Wilson and Herzberg notation of these ring modes. The symmetry of the vibration appears in the third column. Several studies have appeared on the frequencies of benzene.[13,15−17] The last three columns show the experimental frequencies of benzene reported in the liquid state[13] and in the gas phase[15] and the theoretical estimation for the harmonic frequencies,[14] $\omega_{\text{est.}}$, obtained after using the anharmonic correction in the experimentally observed bands.

In the C−H stretching region, modes 2, 7, 13, and 20 appear strongly overestimated with all the methods, except by SAM1. The closest values to the experimental ones correspond to PM3. It is a consequence of the discrepancies obtained in the bond lengths, which have the effect on the predicted vibrational stretching frequencies. In the medium and low region we observe a strong change in the calculated frequency of modes 4, 5, 1, and 3 with the method used and an insignificant variation in the modes 16, 6, and 11.

The difference between the calculated and experimental frequencies, $\Delta(\omega^{\text{cal}} - \nu^{\text{exp}})$, are listed in Table 3. The largest absolute errors are printed in bold. The bottom row shows the root-mean-square error, rms, obtained with the different methods. The average errors obtained are in agreement with those reported for 61 molecules.[18]

In the semiempirical methods it is noted that in general AM1 overestimates the frequencies while SAM1 underestimates the values. Although it is the newest semiempirical method, the SAM1 shows larger errors than AM1 and PM3. PM3 gives the best calculated frequency in the C−H region (even better than the ab initio and DFT methods) with a very low error, ca. 6 cm^{-1}, although it fails dramatically in the prediction of modes 1, 3, 4, and 8. MNDO is the oldest of the four semiempirical methods used, and consequently the calculated frequencies are the poorest.

In ab initio methods the vibrational frequencies are usually calculated using the simple harmonic oscillator model. Therefore, they are typically greater than the fundamentals observed experimentally.[19] In general, the calculated ab initio frequencies are overestimated, at the Hartree−Fock level by about 10−20% and at the MP2 level by about 5−10%, which agrees with our calculations. This overestimation in the frequencies also depends on the type of vibrational mode and on the frequency range, varying between 1 and 12%, Table 3. Thus for modes that appear at high frequencies, the difference between the harmonic oscillator prediction and the exact or Morse potential like behavior is about 10%. The highest differences correspond to the ν(C−H) modes: 2, 7, 13, and 20. However, at a very low frequency, below a few hundred wavenumbers, this difference is off by a large amount, especially in mode 4 and by MP2. In HF and MP2, the rms error with the 6-31G** basis set is higher than that obtained with the modest AM1 and PM3 semiempirical methods. The inclusion of two rows of "diffuse" functions to the HF/6-31G** level leads to significant changes, ca. 10 cm^{-1}, only in the modes 1, 5, 8, and 19. They are also higher than

TABLE 3: Values of the Absolute Error, $\Delta(\omega^{cal} - \nu^{exp})$ in cm^{-1}, Obtained at Several Levels in the Calculated Frequencies of the Ring Modes of Benzene

Wilson no.	semiempirical				ab initio			density functional methods								
	MNDO	AM1	PM3	SAM1	HF/ 6-31G**	MP2/ 6-31G**	MP2/ TZ2P+f	BP86/ 6-31G**	BP86/ 6-311G (2d,p)	BLYP/ 6-31G**	B3P86/ 6-31G*	B3P86/ 6-31G**	B3LYP/ 6-31G*	B3LYP/ 6-31G**	B3LYP/ 6-311+G (2d,p)	B3PW91/ 6-31G**
1	204	**283**	237	**277**	90	35	25	1	−5	−7	37	36	28	27	19	33
2	**349**	131	11	−166	**299**	209	168	54	37	49	152	150	140	134	116	144
3	22	−22	−125	−148	151	46	24	−21	−12	−8	32	24	37	31	40	25
4	−109	−89	−88	−107	71	−207	−23	−12	−7	−11	12	13	11	11	6	12
5	85	22	34	−138	145	−71	6	−22	−29	−16	24	24	20	23	13	25
6	5	40	12	7	57	10	0	−8	−4	−3	11	10	14	13	17	10
7	**350**	131	3	−186	**286**	201	160	45	29	39	143	141	130	124	108	135
8	118	167	184	152	194	74	36	0	−15	−12	71	66	56	52	33	63
9	74	44	−31	−27	110	46	17	−11	−14	−7	31	24	30	25	21	25
10	70	44	2	67	114	7	18	−17	−20	−12	19	17	17	18	12	19
11	97	70	38	97	90	15	9	−8	−14	0	19	17	19	20	9	20
12	−53	18	−34	−12	86	10	29	−26	−13	−19	6	4	10	8	21	3
13	**345**	127	−2	−195	**275**	191	161	35	19	30	133	131	121	115	99	125
14	−32	59	−6	56	41	158	152	45	17	19	76	74	49	47	23	69
15	14	29	3	−35	42	51	28	−4	−7	4	35	27	35	30	26	28
16	−30	−27	42	−16	55	−3	6	−1	−1	3	15	14	17	17	13	14
17	70	22	11	−75	135	−51	12	−37	−40	−34	6	9	2	7	10	11
18	109	108	30	50	101	44	36	1	−6	−4	38	34	31	29	21	33
19	71	95	63	62	162	57	31	−10	−14	−4	49	41	47	41	34	40
20	**351**	131	9	−175	**297**	210	174	53	37	47	152	149	139	133	116	144
rms[a]	174	105	79	125	163	114	84	27	20	22	62	73	70	65	53	68

[a] The root-mean-square errors for the wavenumbers, defined as, $\Sigma\Delta(\omega^{cal} - \nu^{exp})^2/n)^{1/2}$, where the sum is over all the modes of benzene, n.

with MP2. In MP2 with the increase of the basis set, the rms error is only slightly reduced, e.g. with the very large basis[14] of TZ2P+f the rms error is still great, 84 cm^{-1}. MP2 fails dramatically in the calculation of modes 2, 7, 13, 14, and 20.

Only with density functional methods are the frequencies close to the experimental values. The BLYP is the best for this purpose. The 6-31G* or 6-31G** basis sets are the most appropriate, due to the small improvement obtained with its increase. In B3LYP the addition of a row of "diffuse" functions to the 6-311G(2d,p) level only gives rise to a small variation of 10 cm^{-1} in mode 4 and 17. However, the results obtained with DFT methods do not adequately reproduce all the experimental pattern of frequencies with enough accuracy. The use of scale factors solves this problem, improving the predicted frequency.

Scaling the Frequencies. The calculated frequencies of benzene were directly scaled by two procedures: using an overall scale factor for the calculated level and using a scaling equation determined for the benzene ring modes at each level.

Single Overall Scale Factor for the Calculated Level. This section describes the scaling of all the computed frequencies of the benzene molecule at a specific level of theory, with a unique scale factor (or correction factor), which is common for all the molecules with the same level of calculation. These scale factors differ for the different levels of theory, and it is the procedure generally used in the literature to scale the calculated frequencies.

As scale factors we used for benzene those reported[3,20] from a data set containing 122 molecules and 1066 frequencies. In them, Scott et al.[3] recommended the use of two scale factors, one for the high and medium spectra, and the other for the low-frequency modes, except with AM1 and PM3, in which is only reported a single scale factor, and in HF in which the two scale factors are reasonably close to one another so that the use of a single scale factor does not introduce significant error. The values of these scale factors are included in Table 4, together with the absolute errors obtained with this procedure, between the scaled frequencies and the corresponding experimental ones in the gas phase.[15] The largest absolute errors are printed in bold. Scale factors have not been found in the literature for the semiempiricals MNDO and SAM1 and for the SLYP method, and thus they are not included in Table 4. Neither have we found scale factors for the 6-31G**, 6-311G**, 6-311G(2d,p), and 6-311+G(2d,p) basis sets in DFT methods; the scale factors at the 6-31G* level are thus used for our 6-31G** results.

Comparing Tables 3 and 4, in general, with this procedure of scaling one obtains a remarkable reduction of the error. With B3LYP and B3P86 the drop is 3 times, while with AM1, MP2, and SVWN it is almost half. A special reduction is observed with HF, from the highest error (rms of 163 cm^{-1}) in the calculated frequencies of Table 3, to a rms error of 38 cm^{-1} and much better than MP2. The exceptions are the BLYP method (which slightly increases the error, although it continues low) and the PM3 method.

The lowest error is obtained with B3LYP. However, its value is yet slightly too high, especially in modes 12 and 17, to be used routinely. An additional reduction of the error can be reached using a scaling equation, as described below.

With a Scaling Equation Determined for the Benzene Ring Modes. A linear relation is observed between the calculated and the experimental frequencies. For example, in Figure 3 is plotted the result at the MP2 level. Therefore, by linear regression we obtain the equations listed in Table 5. A good correlation coefficient, *r*, is found in all cases, especially in DFT methods with values close to 1. The poorest corresponds to SAM1, which is the only one with a pendent higher than the value 1. The absolute errors obtained with this procedure are also collected in Table 4.

It is noted that the scaled frequency of mode 12 is underestimated with all the methods, except with MP2. Its value is sensitive to the basis set, and its error is not reduced drastically until two rows of "d" functions are added, e.g., 6-311G(2d,p).

Also noted is the continued great error of the sophisticated MP2 procedure, as great as the simplest semiempirical methods. To investigate the cause, in an analysis of the different modes with MP2, it is observed that the frequency of mode 4 was calculated in Table 1 to be far too low (by 200−300 cm^{-1}), being the poorest determined by MP2. Its error is not significatively reduced with the use of a scale factor or a scaling equation. With the increase of the basis set the error is

TABLE 4: Values of the Absolute Error, Δ($\nu^{scaled} - \nu^{exp}$) in cm^{-1}, Obtained in the Scaled Wavenumbers of the Benzene Ring Modes at Several Levels Using a Single Overall Scale Factor and a Scaling Equation

	overall scale factor										scaling equation													
Wilson no.	AM1 ×0.9532	PM3 ×0.9761	HF/ 6-31G** ×0.8992	MP2/ 6-31G** ×0.9427	BP86/ 6-31G** ×0.9914	BLYP/ 6-31G** ×0.9945	B3P86/ 6-31G** ×0.9558	B3LYP/ 6-31G* ×0.9614	B3LYP/ 6-31G** ×0.9614	B3PW91/ 6-31G** ×0.9573	AM1	PM3	SAM1	HF/ 6-31G**	MP2/ 6-31G**	MP2/ TZ2P+f	BP86/ 6-31G**	BLYP/ 6-31G**	B3P86/ 6-31G**	B3LYP/ 6-31G**	B3LYP/ 6-311+G (2d,p)	B3PW91/ 6-31G**
1	**223**	**208**	−19	−30	−8	−12	−9	−11	−12	−11	**221**	**223**	**293**	−9	25	1	9	−1	9	2	−0.6	7
2	−19	−63	−41	2	27	32	7	16	10	7	−22	−21	−36	8	−7	2	9	11	8	9	8.8	8
3	−84	−154	0	−42	−32	−15	−37	−17	−22	−34	−86	−138	−136	17	2	−23	−21	−9	−21	−10	4.9	−19
4[a]	−118	−103	0	**−196**	24	32	7	12	12	7	−120	−95	−137	−3	**−169**	−25	3	2	2	0	−1.6	1
5	−25	10	31	−129	−30	−21	−21	−19	−16	−18	−27	23	**−150**	41	−71	−16	−13	−9	−2	−1	−5.8	−1
6[a]	10	3	−4	24	23	35	5	15	14	6	8	5	−22	−7	37	2	10	12	5	7	12.5	4
7	−18	−70	−51	−4	18	22	0	7	1	−1	−21	−29	−59	−3	−13	−5	1	1	1	1	2.2	0
8	84	141	13	−32	−14	−21	−8	−8	−12	−8	82	**165**	**202**	35	5	−27	−7	−18	6	−2	−11.4	5
9	−13	−58	−20	−31	−21	−13	−29	−17	−21	−26	−15	−44	−19	−7	18	−18	−7	−4	−12	−8	−5.6	−10
10	2	−18	17	−47	−24	−17	−21	−16	−15	−18	0	−7	59	25	12	4	−5	−2	−1	0	−1.5	1
11[a]	35	21	20	31	26	42	12	20	21	15	33	30	79	17	36	7	8	13	8	10	3.1	10
12	−30	−57	−24	−54	−34	−24	−41	−29	−31	−40	−32	−45	−14	−14	0	4	−18	−13	−22	−16	0.2	−23
13	−22	−75	−61	−14	8	13	−10	−2	−7	−11	−25	−34	−68	−13	−22	−4	−9	−7	−9	−8	−7.1	−9
14	−5	−37	−95	66	33	12	13	−3	−5	10	−7	−20	79	**−81**	**108**	**100**	44	18	29	7	−9.7	25
15	−26	−25	−78	−25	−14	−2	−25	−11	−16	−22	−28	−10	−29	−66	25	−6	0	7	−7	−2	0.3	−6
16[a]	−44	−51	14	6	19	28	11	18	18	11	−46	−46	−61	8	44	22	22	22	20	20	17.3	18
17	−24	−12	24	−109	−45	−39	−34	−35	−31	−31	−26	0	−85	34	−51	−9	−27	−26	−15	−15	−8.2	−13
18	54	4	−14	−24	−10	−10	−13	−10	−12	−13	52	18	54	−3	29	9	6	2	5	2	−0.7	4
19	21	26	−4	−40	−23	−12	−26	−12	−18	−25	19	46	98	15	0	−25	−14	−8	−12	−7	−6.0	−9
20	−19	−64	−42	4	26	30	7	15	10	7	−21	−23	−46	7	−5	8	8	9	8	9	8.9	9
rms	67	80	38	65	25	24	20	16.5	16.9	19	67	77	109	29	53	26	16	12	13	8.9	7.4	12

[a] With the scale factors corresponding to the low wavenumbers modes, ref 3: ×0.9089 (HF/6-31G**), ×1.0214 (MP2/6-31G**), ×1.0512 (BP86), ×1.0620 (BLYP), ×0.9923 (B3P86), ×1.0013 (B3LYP), ×0.9930 (B3PW91).

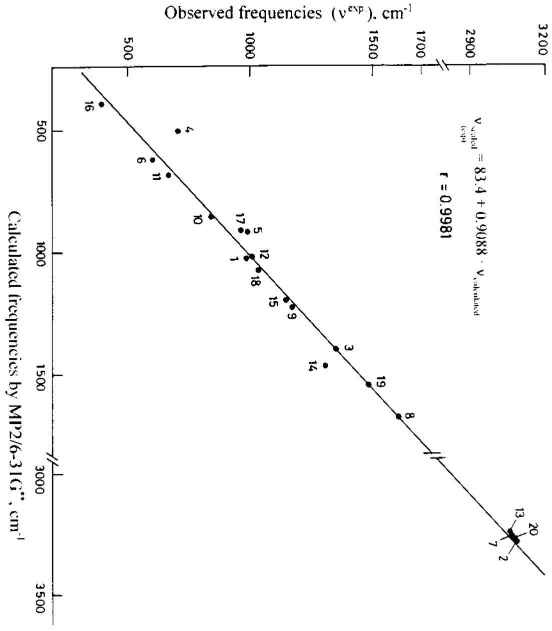

Figure 3. Calculated frequencies by MP2/6-31G** versus experimental ones in benzene molecule.

TABLE 5: Scaling Equations $\nu^{scaled} = a + b\omega^{calculated}$ for the Benzene Ring Modes

methods	*a*	*b*	correlation coeff, *r*
semiempirical			
MNDO	89.4	0.8729	0.9977
AM1	−1.6	0.9529	0.9977
PM3	−0.7	0.9898	0.9960
SAM1	−71.5	1.0692	0.9920
SCF			
HF/4-21G*	−3.8	0.9128	0.9993
HF/6-31G*	−4.0	0.9103	0.9994
HF/6-31G**	−8.6	0.9162	0.9994
HF/6-31++G**	−6.2	0.9153	0.9994
post-SCF			
MP2/6-31G**	83.4	0.9088	0.9981
MP2/6-311G**	97.3	0.9156	0.9972
MP2/TZ2P+f	41.4	0.9360	0.9995
density functional			
SVWN/6-31G**	30.3	0.9692	0.9994
SLYP/6-31G**	27.8	0.9811	0.9991
BP86/6-31G**	32.7	0.9752	0.9998
BP86/6-311G(2d,p)	28.8	0.9819	0.9999
BLYP/6-31G	31.6	0.9530	0.9996
BLYP/6-31G**	27.2	0.9791	0.9999
B3P86/6-31G*	25.0	0.9473	0.9999
B3P86/6-31G**	27.2	0.9476	0.9999
B3LYP/6-31G*	23.3	0.9519	0.9999
B3LYP/6-31G**	22.1	0.9543	0.9999
B3LYP/6-311G**	17.8	0.9614	0.9999
B3LYP/6-311G(2d,p)	18.6	0.9616	0.9999
B3LYP/6-311+G(2d,p)	21.2	0.9597	1.0000
B3PW91/6-31G**	24.8	0.9501	0.9999

incremented, except when “f” basis functions are included in a very large basis with triple-ζ plus double polarization (TZ2P+f), which remarkably reduces the error. This fact has been interpreted[14] as a lack of balance of the basis set which is almost solved with “f” basis functions. However, even with this basis set, the frequency of mode 4 has not converged completely.

Mode 14 is also a failure of MP2. It is poorly calculated, and the effect of scaling or addition of “f” functions to the basis set is slight, although it is also poorly scaled with other methods especially with SAM1, SVWN, SLYP, HF, and MP2. This mode 14 shows in Table 4 the greatest error of 100 cm^{-1} with TZ2P+f. An examination of the mode itself reveals the problem. In this mode the displacement of the C atoms is toward or away one from one another, in a tendency to break the C_6 ring into three C_2 molecules. Furthermore, the H atoms displacement is

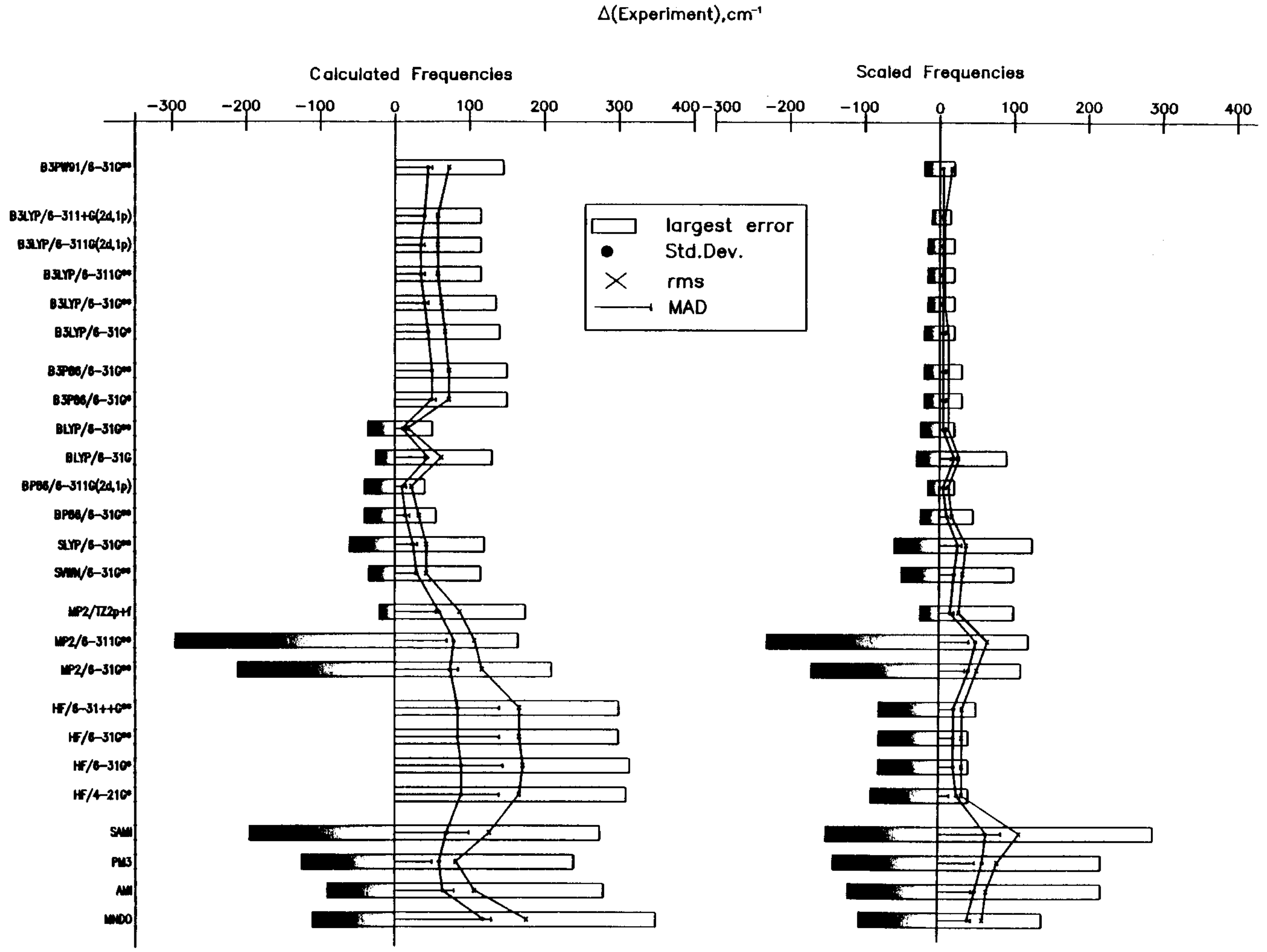

Figure 4. Error obtained by the different methods and basis sets in the calculated frequencies and in the scaled values with the scaling equations.

less than the C atom displacement. This contrasts with the SCF description of the mode, where the usual larger displacement of the H atoms occurs, Table 2. Thus we conclude that the MP2 method itself is certainly deficient for special vibrations, as modes 4 and 14. This failure in the prediction of the frequencies has been observed in other molecules.[21]

B3LYP procedure at the 6-311+G(2d,p) gives the best result, but the simpler 6-31G* is almost as successful; i.e., the rms error is only reduced 0.3 cm^{-1} at 6-311G(2d,p) and 1.6 cm^{-1} at 6-311+G(2d,p), because the error in mode 16 is still great. In general, DFT procedures with B-bases and B3-bases yield very low errors that are of comparable reliability. The use of the scaling equations reduces the errors to ca. half of those found with an overall scale factor, showing that the errors in the calculated frequencies with DFT methods are systematic and partially associated with the kind of molecules studied, and therefore, they can be reduced by employing specific scale factors or scaling equations.

HF, SVWN, and SLYP are of comparable reliability but have larger errors than B3LYP. The semiempirical methods (especially SAM1) and MP2 are the poorest procedures, and although with the use of a scaling equation the error is remarkably reduced, it is still too large to be recommended.

An analysis of the Tables 3 and 4 can permit us to examine for the benzene molecule the relative accuracies in the frequencies of a variety of different model chemistries. The results are shown in graphical form in Figure 4, which plots the mean absolute deviation (MAD), the root-mean-square (rms) error, the standard deviation (StdDev), and the greatest positive and negative deviations from experiment for each model chemistry. The errors plotted correspond to the calculated and scaled frequencies (with the scaling equations) at the different methods and levels under study. From these results the following is observed:

In the semiempirical methods the errors obtained in the calculated frequencies are in accordance with those reported[18] in general for the ring modes, the best being PM3 and the worst MNDO. However, in the scaled frequencies the best is MNDO and the worst SAM1. PM3 is certainly no better than AM1 and MNDO. These results indicate that in PM3 and SAM1 the error in the calculated frequencies is largely inherent to the method and it is not appreciably reduced by the simple use of an overall scale factor or a scaling equation. However, by AM1 and especially MNDO, it is partially systematic, and therefore it can be remarkably reduced, in general, with the use of a scaling equation, except in modes 1, 3, 4, and 8. It is because the other part of the error is associated with the mode itself, with its calculated frequency, and this error is systematic for the same mode in related molecules and can be only eliminated by using specific scale factors for each vibration, as described in the next section.

Although the semiempirical methods have better MAD's than the Hartree−Fock-based methods, indicating that their parametrizations have accounted for some of the effects of electron correlation, their errors are less systematic than the HF, and thus they are less reduced with the use of a single overall scale

TABLE 6: Calculated Specific Scale Factors, ν^{exp}/ω^{cal}, for Each Normal Mode of the Benzene Molecule and at Some of the Levels Considered

Wilson no.	semiempirical			ab initio			density functional methods						
	AM1	PM3	SAM1	HF/ 6-31G*	HF/ 6-31G**	MP2/ 6-31G**	BP86/ 6-31G**	BLYP/ 6-31G**	B3P86/ 6-31G**	B3LYP/ 6-31G*	B3LYP/ 6-31G**	B3LYP/ 6-311G**	B3PW91/ 6-31G**
1	0.7783	0.8074	0.7819	0.9161	0.9169	0.9660	0.9991	1.0072	0.9651	0.9726	0.9736	0.9707	0.9679
2	0.9588	0.9964	1.0571	0.9068	0.9113	0.9363	0.9827	0.9843	0.9535	0.9564	0.9582	0.9636	0.9552
3	1.0166	1.1020	1.1231	0.8952	0.8994	0.9670	1.0158	1.0060	0.9825	0.9733	0.9775	0.9768	0.9818
4	1.1440	1.1422	1.1783	0.9099	0.9088	1.4140	1.0173	1.0158	0.9819	0.9846	0.9846	0.9779	0.9833
5	0.9783	0.9668	1.1620	0.8707	0.8715	1.0772	1.0227	1.0164	0.9764	0.9802	0.9773	0.9734	0.9754
6	0.9352	0.9809	0.9888	0.9145	0.9145	0.9841	1.0135	1.0052	0.9841	0.9777	0.9792	0.9746	0.9840
7	0.9589	0.9989	1.0647	0.9097	0.9147	0.9382	0.9854	0.9873	0.9558	0.9591	0.9609	0.9661	0.9576
8	0.9049	0.8969	0.9132	0.8909	0.8919	0.9558	1.0000	1.0075	0.9604	0.9662	0.9780	0.9780	0.9621
9	0.9640	1.0268	1.0232	0.9102	0.9144	0.9623	1.0092	1.0058	0.9798	0.9750	0.9790	0.9839	0.9790
10	0.9495	0.9978	0.9268	0.8806	0.8814	0.9920	1.0206	1.0145	0.9805	0.9804	0.9793	0.9816	0.9782
11	0.9046	0.9466	0.8741	0.8810	0.8821	0.9782	1.0120	1.0000	0.9753	0.9725	0.9712	0.9782	0.9711
12	0.9825	1.0349	1.0120	0.9207	0.9216	0.9902	1.0264	1.0192	0.9960	0.9902	0.9922	0.9970	0.9970
13	0.9602	1.0007	1.0681	0.9128	0.9178	0.9412	0.9887	0.9903	0.9589	0.9619	0.9638	0.9692	0.9607
14	0.9568	1.0049	0.9592	0.9685	0.9699	0.8925	0.9671	0.9860	0.9468	0.9642	0.9656	0.9808	0.9502
15	0.9720	0.9971	1.0311	0.9605	0.9637	0.9573	1.0032	0.9963	0.9768	0.9702	0.9743	0.9785	0.9760
16	1.0727	1.1179	1.0419	0.8786	0.8786	1.0076	1.0025	0.9925	0.9660	0.9590	0.9590	0.9637	0.9660
17	0.9778	0.9887	1.0841	0.8791	0.8775	1.0556	1.0398	1.0364	0.9908	0.9979	0.9929	0.9847	0.9888
18	0.9049	0.9722	0.9543	0.9092	0.9116	0.9596	1.0012	1.0041	0.9685	0.9712	0.9730	0.9795	0.9694
19	0.9386	0.9592	0.9599	0.8983	0.9016	0.9630	1.0068	1.0027	0.9731	0.9693	0.9731	0.9808	0.9737
20	0.9590	0.9972	1.0607	0.9072	0.9117	0.9360	0.9831	0.9850	0.9537	0.9567	0.9585	0.9639	0.9552

factor or a scaling equation. In HF the reduction is drastic and the used basis has little effect. However, in several modes the error is again very great (e.g. modes 14 and 15), and thus to reduce the error, specific scale factors should be used.

MP2 does not appear to offer a significant improvement in performing the calculated frequencies over HF, and in the scaled frequencies it shows an almost double error. In modes 4, 5, 14, 16, and 17 the error is especially great. For these reasons and for the excessive time and memory computer consuming, it is preferable to use a method other than MP2.

DFT methods are the most cost-effective procedures found for predicting the vibrational frequencies, and in general they can be used efficiently. Differences are observed between the results obtained with these DFT methods. Thus the scaled frequencies with SLYP and SVWN have the greatest error of the DFT methods because the error is quite great in modes 3 and 14, which is not reduced with scaling. However, the error is generally systematic in B3P86 and the B3LYP procedures and the use of a scale factor or a scaling equation lead to accurate frequencies. The B-based DFT procedures, while not performing quite as well as the corresponding B3-based procedures, have the attraction of standard frequency scale factors close to unity, meaning that they can often be used without scaling. The LYP functional is superior in precision to the P86 and PW91 functionals, and also superior to the CAP functional.[14] Thus combining the most accurate exchange with the correlation functional leads to B3LYP, which gives the lowest errors in benzene and is the recommended method.

It is noted that the frequencies of modes 16 and 6 appear in general overestimated by ab initio and DFT methods, and by using an overall scale factor or a scaling equation, and this error is not reduced even by increasing the basis set. This fact can tentatively suggest that there is a slight inexactitude in the frequency of the observed bands. To solve this problem we look to the estimated harmonic frequencies reported by Handy et al.[14] (Table 1) and by Goodman et al.[15] They differ remarkably in the ν(C−H) modes 2, 7, 13, and 20 with the gas-phase values.[15] These estimated frequencies, although used to obtain the semiempirical scaled valence force field of benzene[17] with very low error, fail in comparison with our scaled frequencies, except for the estimations by Handy et al.[14] in modes 6 and 16. Thus, replacing the ν_{obs}[15] of modes 6 and 16 at 608.13 and 398 cm^{-1}, respectively, by the ω_{est}[14] at 613 and 407 cm^{-1} reduces the absolute error in the scaled frequencies, e.g., to 7.6 and 8.3 cm^{-1} at B3LYP/6-311G(2d,p), modes 6 and 16 respectively, instead of 12.5 and 17.3 cm^{-1}.

Calculation of Specific Scale Factor for Each Mode. The procedure is based on the assumption that the ratios between experimental and computed frequencies are fairly constant for each type of characteristic frequency (in our case for each ring mode in benzene). It is then possible to derive for a known experimental spectrum a correction factor for each characteristic frequency by taking the average of the ratios between the experimental and computed frequencies, $\lambda = \nu^{exp}/\omega^{cal}$ (or $\lambda' = \omega^{cal}/\nu^{exp}$) and to use them for predicting or assigning unknown spectra. The scale factors determined for the benzene ring modes are listed in Table 6, which can be used for the ring modes of benzene derivatives. The introduction of different scale factors for distinct types of vibrational modes, instead of using a uniform correction factor for all the modes, removes the error that is systematic and associated with the same mode in related molecules and, therefore, remarkably improves the accuracy of the methods in predicting the frequencies, especially the semiempiricals.

In general, the majority of the published works with ab initio methods use a single overall correction factor for the frequencies, with no consideration for the different modes. Sometimes a simplification is used with only two or three scale factors for the modes, e.g., 0.9 for stretches and bends and 1.0 for torsion. However, a better accuracy is obtained if a specific scale factor is used for each mode and level of calculation, although it requires slightly more effort. This procedure of scaling is relatively new and, to our knowledge, the only scale factors reported with this procedure are for tertiary amines[2] and for toluene molecule.[22] However, the scale factors reported for toluene are not as detailed as the calculations carried out by us on the benzene molecule and listed in Table 6.

With semiempirical methods this procedure of scaling was tested to the ring modes of several benzene derivatives with biological and pharmacological applications, and errors lower than 2% in the majority of the cases[23] were obtained. With ab initio and DFT methods the errors were remarkably reduced.

TABLE 7: Vibrational Wavenumbers at the B3LYP/6-31G** Level and Experimental Ones, in cm^{-1}, and the Error Obtained in the Scaled Values with the Three Procedures of Scaling

Wilson no.	aniline					benzoic acid					phenylsilane					*p*-aminobenzoic acid				
	frequencies		absolute error[a]			frequencies		absolute error[a]			frequencies		absolute error[a]			frequencies		absolute error[a]		
	cal	exp[d]	(1)	(2)	(3)	cal	exp[f]	(1)	(2)	(3)	cal	exp[j]	(1)	(2)	(3)	cal	exp[f]	(1)	(2)	(3)
1	835	812	−9	7	1	1127		−8	9	8	704[g]	690	15	4	−5	853	839	−19	−3	−9
2	3207	3084	−1	−2	−11	3225		−25	−10	−9	3204	3080	0	0	−10	3222	3060[p]	38	37	27
3	1370[b]	1308				1347		17	9	3	1363	1331	−21	−8	1	1342	1322[p]	−32	−19	−10
4	705	690	16	5	4	705	686	−7	3	20	770[h]	752	−12	5	6	712	698	15	4	3
5	978	968	−28	−13	−12	1012	998	−11	−1	16	1006	985	−18	−3	−2	983	940	5	20	21
6a	536	526	11	8	−1	637		−35	−22	−14	388[h]	388	1	4	−8	352[k]	342	10	16	3
6b	635	619	17	9	3	631	615	−21	−5	−5	633	619	15	7	1	652	636	17	8	2
7a	3184	3056	5	4	4	778		−9	8	−2	3182	3067	−8	−8	−9	1394	1352	−12	0	−13
7b	3168	3025	21	20	19	3195		−22	−7	9	3169	3033	14	13	12	3177				
8a	1677[c]	1603				1661	1604	8	21	14	1648	1596	−12	−1	16	1678[l]	1622	−9	1	19
8b	1643	1590	−10	0	17	1639	1587	30	32	12	1626	1567	−4	7	23	1623	1586[p]	−26	−15	1
9a	1206	1180	−21	−7	1	1195	1184	7	8	−12	1218	1190	−19	−6	2	1191	1172	−27	−13	−6
9b	1184	1152	−14	0	7	1189		−22	−7	9	1190	1165	−21	−7	0	509[k]				
10a	827	820	−45	−9	−10	869	856	−19	−4	−6	870	847	−11	5	5	827	815	20	−4	−5
10b	224	233[e]	−9	3	−14	817		−16	−2	−4	153[i]	159	−6	9	−9	774[m]	770	5	−9	−12
11	763	745	20	5	−4	729	710	−20	−9	−2	717	699	19	7	−3	280	258	22	31	14
12	1010	996	−25	−10	6	1019	1002	−20	−8	−2	1015	998	−22	−7	9	740[n]				
13	3167	3074	−29	−30	−22	3183		14	14	5	3167	3023	22	21	29	1336	1290	−6	7	−2
14	1367	1324	−10	3	−4	1367	1306	30	29	20	1320	1298	−29	−16	−23	1380[k]	1341	−14	−2	−8
15	384	388	−4	0	−14	498					201	203	−2	11	−7	384[k]	385	−1	4	−11
16a	418	415[e]	4	6	−14	416	387				404	386	19	22	1	422	427[p]	−4	−2	−22
16b	507	500	8	6	−14	437	431				437	420	18	19	−1	518[o]	497	22	19	0
17a	953	957	−41	−26	−11	990	974				979	920	−21	−6	8	968	962	−31	−16	−1
17b	879	874	−29	−13	−1	959					935	1030	−17	−2	−4	847	841	−27	−11	0
18a	1055	1028	−14	1	−1	1052	1030				1054	1064	−2	13	11	1027	1009	−22	−7	−10
18b	1142	1090	8	22	21	1101	1075				1105	1487	−16	−5	2	1156	1128	−17	−3	−3
19a	1546	1503	−17	−6	1	1537	1498				1530	1432	−18	−6	−1	1563	1520	−17	−6	1
19b	1513	1470	−15	−4	2	1493	1455				1471	1123	−35	−21		1484	1440	−13	−2	4
20a	1315	1278	−12	1	−16	3205	3067				1132[g]	3055	14	13	5	3178	3060	−5	−5	−14
20b	3189	3046	20	19	11	3221	3067				3192					3218				

Wilson no.	*p*-methoxybenzoic acid					1,4-dicyanobenzene					2,4-difluorobenzonitrile					phenothiazine				
	frequencies		absolute error[q]			frequencies		absolute error[q]			frequencies		absolute error[q]			frequencies		absolute error[q]		
	cal	exp[u]	(1)	(2)	(3)	cal	exp[x]	(1)	(2)	(3)	cal	exp[y]	(1)	(2)	(3)	cal	exp[aa]	(1)	(2)	(3)
1	1070	1025	4	18	17	825	794	−1	15	9	760	743	−12	4	−3					
2	3232	3075	32	31	22	3229	3088	16	16	6	3244	3108	11	10	0	3222	3095.2[bb]	3	−5	−13
3	1332	1304	−23	−11	−2	1338	1328	−42	−29	−20	1281	1276	−44	−31	−24	1284	1248	−14	−2	2
4	711	696	16	5	4	740	725	16	3	4						739	727.3	13	0	1
5	991	958	−5	10	11	985	971	−24	−9	−8	851					971	972.4	−38	−24	−20
6a	604	612	−7	−14	−21	379[v]	372	7	12	−1	518	510	8	6	−3	685	672.7	13	2	−3
6b	647	632	16	8	2	667	662	6	−3	−9	408	401	8	10	−1	588	577.0	12	6	−2
7a	835[r]	821	−18	−2	−19	1229	1193	−11	1	−12	1323	1290	−18	−5	−19					
7b	3218					3215	3078	13	12	11	993	975[z]	−20	−5	−21	3188	2991.0	74	67	67
8a	1667	1602	1	11	28	1663	1619	−20	−10	7	1672	1620	−13	−2	−1	1632	1568.8	0	8	8
8b	1625	1575	−13	−2	14	1598	1609	−73	−62	−46	1635	1588	−16	−6	−5	1636	1579.5[bb]	−6	2	2
9a	1191	1178	−33	−19	−12	1203	1177	−20	−7	1	322	325	−3	4	−10	1200	1164.1	−10	2	6
9b	511	503	9	7	−3	571	566	6	1	−7	445	434	12	13	2					
10a	831					857	820	4	20	19										
10b	775	772	−27	−10	−13	532[w]	524	9	6	−3	232	220	12	23	7					
11	306					379[w]	361	18	23	7	831	818	−19	−3	−11	762	729.0[bb]	34	20	12
12	726[s]					648[v]	636	13	4	7	713	700	14	3	7	1065	1035.1	−11	2	20
13	1305[t]	1263	−8	4	−5	1229	1200	−18	−5	−15	1187	1150	−9	5	−6					
14	1372	1322	−3	9	3	1330	1294	−15	−3	−10	1363	1310	0	13	6	1336	1289.6	−6	5	−2
15	427	398	30	32	11	537[v]	530	8	5	−7	141									
16a	512					408	392	17	19	−1	632	615	18	10	−9					
16b	967	963	−33	−18	−3	585	562	24	18	−1	454									
17a	864	843	−12	4	15	984	977	−31	−16	0	102					931	927.6[bb]	−33	−18	1
17b	1027	1004	−17	−2	−5	862	846	−17	−1	10	962					863	850.3[bb]	−20	−5	11
18a	1140	1105	−9	5	4	1039	1028	−29	−14	−17	1216	1180	−11	3	3	1154	1121.7[bb]	−13	0	−1
18b	1560	1515	−15	−4	3	1142	1123	−25	−11	−12	1127	1100	−17	−2	−3	1072	1052.4	−21	−8	−11
19a	1466	1426	−17	−5	1	1545	1504	−19	−8	−1	1547	1500	−13	−2	5	1495	1447.8	−11	−2	1
19b	3209					1446	1401	−11	1	6	1482	1430	−5	6	12	1510	1458	−6	3	6
20a	3224					3214	3053	37	36	28	3236	3070	41	40	32	3211	3063.5	24	17	9
20b						3227	3099	3	3	−6	3221	3048	49	48	39	3197	3054.1[bb]	20	13	5

[a] $\Delta(\nu^{\mathrm{scaled}} - \nu^{\mathrm{exp}})$, where ν^{scaled} is (1) with an overall scale factor, (2) with a scaling equation, (3) with specific scale factors for each mode. [b] Strongly coupled with $\Gamma(\mathrm{NH_2})$. [c] Strongly coupled with $\beta_s(\mathrm{NH_2})$. [d] Infrared in vapor or solution phase, ref 24. [e] Raman in liquid phase, ref 24. [f] Infrared, ref 25. [g] $+\nu(\mathrm{C{-}Si})$. [h] $+\delta(\mathrm{Si{-}H})$. [i] $+\gamma(\mathrm{Si{-}H})$. [j] Infrared and Raman, ref 28. [k] $+\Gamma(\mathrm{NH_2}) + \Gamma(\mathrm{COO})$. [l] $+\beta_s(\mathrm{NH_2})$. [m] $+\gamma(\mathrm{COOH})$. [n] $+\delta(\mathrm{COO})$. [o] $+\gamma(\mathrm{NH_2})$. [p] Raman. [q] $\Delta(\nu^{\mathrm{scaled}} - \nu^{\mathrm{exp}})$, where ν^{scaled} is (1) with an overall scale factor, (2) with a scaling equation, (3) with specific scale factors for each mode. [r] $\nu(\mathrm{C4{-}O})$, $\nu(\mathrm{C{-}C7})$. [s] $+\delta(\mathrm{COOH})$. [t] $\nu(\mathrm{C4{-}O})$. [u] Infrared and Raman, refs 23c and 25. [v] $+\delta(\mathrm{C{\equiv}N})$. [w] $+\gamma(\mathrm{C{\equiv}N})$. [h] Infrared, ref 29. [y] Infrared in Nujol, ref 26. [z] Infrared in KBr, ref 26. [aa] Raman in solid state, ref 27. [bb] Infrared in KBr, ref 27.

TABLE 8: Error Obtained in the Calculated and Scaled Wavenumbers of Several Benzene Derivatives at the B3LYP/6-31G Level**

	calculated wavenumbers			scaled wavenumbers with an overall factor			scaled wavenumbers with the scaling equations			scaled wavenumbers with specific scale factors		
molecules	rms	MAD	std dev	rms	MAD	std dev	rms	MAD	std dev	rms	MAD	std dev
benzene	62	45.2	42.4	17	15.2	7.4	8.8	6.8	5.5			
phenylsilane	60.1	41.9	42.4	17.0	14.9	8.2	10.7	8.8	6.0	10.5	7.6	7.3
aniline	60	42.3	42.3	19	15.6	10.4	12.4	9.3	8.2	11.0	10.5	9.0
benzoic acid	55.9	40.7	38.4	19.7	17.1	8.5	13.9	12.1	9.2	10.7	11.6	9.5
p-aminobenzoic acid	47.8	32.5	35.1	19.3	16.8	9.6	13.7	10.2	9.2	11.3	8.5	7.4
p-methoxybenzoic acid	46.3	33.8	31.7	18.6	15.8	9.8	13.4	10.5	8.4	12.5	9.9	7.7
1,4-dicyanobenzene	56.5	36.5	43.2	23.1	18.4	14.0	17.7	12.4	12.6	13.4	9.7	9.3
2,4-difluorobenzonitrile	64.1	42.7	47.9	20.2	16.2	12.1	16.5	11.0	12.3	14.2	10.0	10.1
Phenothiazine	75.0	53.3	52.8	24.2	18.2	15.9	17.6	10.0	14.5	17.1	9.7	14.1

For example, Tables 7 and 8 show this improvement at the B3LYP/6-31G** level in the values of eight molecules, e.g. aniline, benzoic acid, phenylsilane, *p*-aminobenzoic acid, *p*-methoxybenzoic acid, 1,4-dicyanobenzene, 2,4-difluorobenzonitrile, and phenothiazine. As reference, in Table 8 the results obtained in a benzene molecule appear. Aniline, benzoic acid, and phenylsilane are monosubstituted benzene derivatives, and with them is expected a great transferability of the scale factors and scaling equations from benzene. Thus the errors should be very low except in modes strongly coupled with the substituents. *p*-Aminobenzoic acid, *p*-methoxybenzoic acid, and 1,4-dicyanobenzene are disubstituted benzene derivatives, and, analogously to the monosubstituted derivatives, with them a great transferability of the scale factors and scaling equations from benzene is expected. It leads to errors similar to the monosubstituted derivatives or slightly worse, depending on the substituent, as in 1,4-dicyanobenzene.

2,4-Difluorobenzonitrile is a trisubstituted benzene derivative with strong electronegativity substituents that affect the description of the ring modes, and thus a worse transferability of the scale factors is expected. The phenothiazine molecule is two benzene rings fused in ortho and meta positions through a nitrogen and a sulfur atom. This structure strongly deforms the bonds and angles of the benzene ring and thus the characterization of their normal modes. Therefore, in this molecule a worse error than in the others is expected.

The 1st column of Table 7 gives the ring modes according to Wilson notation. The 2nd, 3rd, 12th, and 17th columns list the calculated frequencies at only the B3LYP/6-31G** level in the molecules under testing. The 3rd, 8th, 13th, and 18th columns collect the experimental frequencies reported.[24−29] The rest of the columns show the absolute error obtained in the scaled frequencies using the three procedures of scaling mentioned in the present manuscript, i.e., with an overall scale factor, with a scale equation, and with specific scale factors for each mode. The rms errors obtained are listed in Table 8. The use of the scaling equations represent always an improvement of the frequencies than the use of a single overall scale factor, although the use of specific scale factors for each mode leads to the best scaled frequencies. As expected and discussed before, the rms error in phenothiazine is larger than in 2,4-difluorobenzonitrile and these are also larger than in monosubstituted derivatives.

In an analysis of these molecules at other levels and methods, we note the following: The reduction of the error is especially drastic in semiempirical, HF, MP2, SVWN, and SLYP methods, and thus if these methods are used, it is recommended to scale the frequencies with specific scale factors for each mode. The reduction in HF is so high that the scaled frequencies are of comparable reliability (or even better) than those obtained with B and B3 bases. The errors in semiempirical calculations are far less systematic than in HF and DFT methods and should not be used except for large benzene derivatives.

Finally, it is concluded that the procedure with a scaling equation represents a remarkable improvement of the scaled frequencies over the single overall factor procedure, being at the B3LYP/6-31G** level and in monosubstituted benzene about 30% in the rms error. The procedure with specific scale factors represents again an improvement over the scaling equation, about 10−15% in the rms error and at B3LYP, and should be the procedure recommended for scaling if a high accuracy is required in the scaled frequencies.

Summary and Conclusions

The ring modes were identified for the different methods by the value of the displacement vectors, in addition to its direction reported in the literature. Depending of the method used, several modes drastically change its nature, e.g., modes 14 and 15.

From the estimated harmonic frequencies we selected as ω_{best} that value reported by Goodman et al.,[15] except in modes 6 and 16 where we chose the values of Handy et al.[14]

The accuracy of the methods in calculating the frequency of the benzene ring modes was determined. For the C−H stretching region, the best calculated frequencies were those with PM3, while for the entire spectrum the BLYP method with the 6-31G* basis set or higher gave the best frequencies. There was no demonstrated advantage with SAM1 over the other semiempirical methods, which fail in the predicted frequencies. The calculations at the HF level failed to give the observed experimental pattern, although the inclusion of electron correlation slightly improved the values. DFT methods have shown a more reliable prediction for the calculated frequencies than with the expensive HF and MP2 methods.

To improve the calculated frequencies, three procedures are proposed. The scaling equations procedure gives rise to a more noticeable improvement in the predicted frequencies, than when a single overall scale factor is used.

For the computed frequencies of the semiempirical methods it is not recommended to use a single overall scale factor or a scaling equation, instead specific scale factors for each mode should be used. The density functional frequencies appeared to be more reliable predictions than the MP2 frequencies, and they were obtained at significantly less cost. The LYP correlation functional was slightly better than P86 and PW91. The best was B3LYP. The basis set had little influence.

For benzene derivatives, the procedure selected for scaling depends on the size of the organic molecule and the accuracy required for the predicted frequencies. For this purpose a complete set of specific scale factors for each mode and scaling equations were determined and recommended to be used for the different levels and methods.

If the benzene derivative has less than 20−30 atoms, DFT methods and 6-31G* basis sets can be used for calculating frequencies. If the accuracy required is not very high (the errors in the predicted frequencies could be between 0 and 4%), then the use of the scaling equation is the simplest and easiest procedure, and among the DFT methods, the most cost-effective is the B3LYP/6-31G*. If the accuracy required is high, then with B3LYP large basis sets and specific scale factors for each mode should be used .

If the benzene derivative is larger than 30 atoms, semiempirical methods and DFT methods, with a small basis set, can be used for calculating frequencies. However, the cost-effective ratio with DFT methods is very high versus that with semiempirical methods, and therefore their use is not recommended at all. In contrast, the AM1 and PM3 semiempirical methods, when used with a specific scale factor for each mode, lead to acceptable scaled frequencies.

References and Notes

(1) (a) Meyer, W.; Pulay, P. *Theor. Chim. Acta* **1974**, *32*, 253. (b) Pulay, P.; Meyer, W. *Mol. Phys.* **1974**, *27*, 473. (c) Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037. (d) Pulay, P.; Zhou, X.; Fogarasi, G. In *Recent Experimental and Computational Advances in Molecular Spectroscopy*; Fausto, R., Ed.; Kluwer Academic Publishers: Netherlands, 1993; p 99.

(2) (a) Hameka, H. F.; Famini, G. R.; Jensen, J. O.; Newhouse, E. I. *Gov. Rep. Announce Index* **1990**, *90*, 13. (b) Hameka, H. F.; Famini, G. R.; Jensen, J. O.; Jensen, J. L. *Gov. Rep. Announce Index* **1991**, *91*, 15.

(3) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

(4) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian, Inc.: Pittsburgh, PA, 1995.

(5) Holder, A. J. *AMPAC 5.0*; Semichem Inc.: 7128 Summit, Shawnee KS 66216, 1994.

(6) Dewar, M. J. S.; Jie, C.; Yu, J. *Tetrahedron* **1993**, *49*, 23, 5003.

(7) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 7, 5612.

(8) Becke, A. D. *Phys. Rev.* **1988**, *A38*, 3098.

(9) Becke, A. D. *J. Chem. Phys.* **1992**, *97*, 9173; **1993**, *98*, 5648.

(10) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.

(11) Perdew, J. P. *Phys. Rev.* **1986**, *B33*, 8822; *B34*, 7406 (E).

(12) (a) Perdew, J. P.; Wang, Y. *Phys. Rev.* **1992**, *B45*, 13244. (b) Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev.* **1992**, *B46*, 6671.

(13) (a) Wilson, E. B. *Phys. Rev.* **1934**, *45*, 706. (b) Varsanyi, G. *Assignment for vibrational spectra of seven hundred benzene derivatives*; Adam Hilger: London, 1974.

(14) Handy, N. C.; Maslen, P. E.; Amos, R. D.; Andrews, J. S.; Murray, C. W.; Laming, G. J. *Chem. Phys. Lett.* **1992**, *197*, 4−5, 506.

(15) Goodman, L.; Ozkabak, A. G.; Thakur, S. N. *J. Phys. Chem.* **1991**, *95*, 23, 9044.

(16) (a) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093. (b) Thakur, S. N.; Goodman, L.; Ozkabak, A. G. *J. Chem. Phys.* **1986**, *84*, 12, 6642.

(17) Baranović, G.; Schrader, B. *THEOCHEM* **1994**, *306*, 165.

(18) Coolidge, M. B.; Marlin, J. E.; Stewart, J. P. *J. Comput. Chem.* **1991**, *12*, 8, 948.

(19) Clark, T. In *Recent Experimental and Computational Advances in Molecular Spectroscopy*; Fausto, R., Ed.; Kluwer Academic Publishers: Netherlands, 1993; p 99.

(20) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc, Pittsburgh, PA, 1996.

(21) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Replogle, E. *Chem. Phys. Lett.* **1989**, *158*, 207.

(22) Hameka, H. F.; Jensen, J. O. *THEOCHEM* **1996**, *362*, 325.

(23) (a) Alcolea Palafox, M. *J. Mol. Struct*. **1988**, *175*, 81. (b) Alcolea Palafox, M. *THEOCHEM* **1991**, *236*, 161. (c) Alcolea Palafox, M.; Gil, M.; Núñez, J. L. *Appl. Spectrsoc.* **1993**, *48*, 1, 27; *Vib. Spectrosc*. **1993**, *6*, 95. (d) Alcolea Palafox, M. *Spectrosc. Lett.* **1994**, *27*, 5, 613; *Vib. Spectrosc*. **1994**, *6*, 149. (e) Alcolea Palafox, M.; Melendez, F. J. *THEOCHEM* **1999**, *459*, 239.

(24) Evans, J. C. *Spectrochim. Acta* **1960**, *16*, 428.

(25) (a) Sanchez de la Blanca, E.; Núñez, J. L.; Martinez, P. *An. Quim. Ser. A* **1986**, *82*, 3, 480. (b) Sanchez de la Blanca, E. Ph.D. Thesis, Universidad Complutense, Madrid 1986.

(26) Rastogi, V. K.; Alcolea Palafox, M.; Singh, S.; Singhal, S. K. *Asian J. Phys*. **1998**, *7*, 2, 229.

(27) Alcolea Palafox, M.; Gil, M.; Núñez, J. L.; Tardajos, G. *J. Phys. Chem.*, submitted for publication.

(28) (a) Alcolea Palafox, M.; Boggs, J. E. *THEOCHEM* **1993**, *284*, 23. (b) Durig, J. R.; Hellams, K. L.; Mulligan, J. H. *Spectrochim. Acta* **1972**, *28A*, 1039.

(29) Higgins, J.; Zhou, X.; Liu, R. *Spectrochim. Acta* **1997**, *53A*, 721.