Evaluation of Relevant Molecular Descriptors for the Solvation of Paramagnetic Relaxation Reagents

Hans Grahn and Ulf Edlund

NMR Research Group, Department of Organic Chemistry, Umeå University, S-901 87 Umeå, Sweden

Grahn, H. and Edlund, U., 1987. Evaluation of Relevant Molecular Descriptors for the Solvation of Paramagnetic Relaxation Reagents. – Acta Chem. Scand. B 41: 24–28.

In order to quantify the importance of the physical/chemical properties of the ligand in the second solvation sphere of a neutral paramagnetic relaxation reagent, $Cr(acac)_3$, a multivariate approach was used to relate ^{13}C electron-nuclear relaxation data and induced chemical shifts of a variety of monosubstituted benzenes to different physical descriptors. When compounds capable of hydrogen bonding were excluded, two main effects could be recognized in the analysis. The first effect, which is similar for both relaxation and shift data, accounts for most of the variance in relaxation and shift data, and was best described by the dipole moment. For the second effect, which separated the two NMR parameters, electronic factors were important for the description of the induced shifts, while variables like log P were shown to have relevance in the description of the T_1^e process. The predictive potential of the model may have relevance in the design of catalytic reactions.

The application of NMR spectroscopy to the study of preferential solvation of electrically neutral paramagnetic species has produced valuable knowledge about the different interacting mechanisms that control this complexation.[1] Parameters such as the chemical shift and the relaxation times (T_1, T_2), are very sensitive to changes in the second coordination sphere of the metal chelate. Several reports in this field have been concerned with ligand substitution processes and homogeneous catalysis, where information about bonding in the second coordination sphere is related to the catalytic activity.[2] Considerable attention has also recently been directed towards paramagnetic compounds used as spin relaxants to circumvent unfavourable nuclear Overhauser effects and to decrease long relaxation times.[3]

In previous papers, we reported preferential solvation-orientation effects of several aromatic systems toward a paramagnetic relaxation reagent (PARR), tris(acetylacetonato)chromium (III), $[Cr(acac)_3]$.[4] For solutes unable to form hydrogen bonds towards $Cr(acac)_3$, the interaction mechanism has been claimed to be of a dipole-dipole or dipole-induced dipole type. This assumption was based on the induced chemical shifts and the electron-nuclear spin-lattice relaxation times of the aromatic outer-sphere ligands. The model was qualitatively acceptable for a variety of systems having functionalized groups, such as tropone, nitrobenzene and *N,N*-dimethylaniline.

To be able to further increase our understanding of outer-sphere complexation towards neutral paramagnetic complexes, we have approached the problem by a statistical analysis in an attempt to quantify the importance of physical properties of the ligands in the solvation sphere.[5] Such an analysis might tell whether the observations made in the T_1^e and the induced shift measurements could be described by suitable physical descriptors for the substrates. In the analysis, we have excluded compounds that could be expected to form hydrogen bonds toward the PARR. A variety of descriptors with expected relevance to the motional and electronic structure of the solute were chosen, such as molecular dipole moment, dielectric constant, density, molecular weight and van der Waals volume.

Thus, we have extended our previous studies

Table 1.

Compound	Substituent[a]	Compound	Substituent[a]
1	– F	*8*	– CHO
2	– Cl	*9*	– $N(CH_3)_2$
3	– Br	*10*	– OCH_3
4	– I	*11*	– H
5	– CH_3	*12*	– NO_2
6	– CH_2CH_3	*13*	– $COCH_3$
7	– CH_2Cl		

[a]Monosubstituted benzenes.

by using a principal components data analysis (PCA) method to examine the systematic variation in T_1^e and paramagnetically induced shift data. One would expect that factors governing the rotational/translational motion of the substrate would be of greater significance for the T_1^e process, while electronic factors, possibilities for orbital overlap between the inner and outer sphere ligand heteroatoms, etc. would be more important for the induced shifts. In the present study, by using tabulated data for the substrates, we have determined to what extent the variation in relaxation/induced shift data can be accounted for by these descriptors.

Experimental

All NMR measurements were made on a Bruker WM 250 NMR instrument at 22±0.5 °C. The ^{13}C relaxation times were measured by the usual inversion recovery method, using at least 15 random variable delays. The diamagnetic T_1 values were obtained by the fast IRFT sequence. The Cr $(acac)_3$ concentration was 0.5 M in all solutions to ensure a sufficient induced paramagnetic shift; the substrate concentration was 0.5 M, using CCl_4 as solvent. All materials were commercially available and purified according to standard procedures. Cyclohexane (0.1 M) was used as an internal standard. Samples (1.5 ml) were kept in 8 mm tubes, fitted into a 10 mm outer tube containing the lock substance (acetone-d_6). The experimental error was 10–15 % for T_1^e and 0.01 ppm for the induced shift measurements. Shift assignments were taken from Ref. 7c.

Results and discussion

A set of 13 monosubstituted aromatic compounds was chosen (Table 1) according to previous recommendations.[6] The induced chemical shifts were obtained as the difference in ppm with and without PARR, using a given molar ratio Cr $(acac)_3$/solute in CCl_4. The electron-nuclear relaxation times, T_1^e, were calculated as:

$$1/T_1^e = 1/T_1^{obs} - 1/T_1^{dia}. \quad (1)$$

All four positions in the aromatic substances were used in the analysis. Three matrices of experimental data were formed: (I) T_1^e (s), (II) induced shifts (ppm), and (III) physical descriptors. The last-named descriptors were taken from Ref. 7. The statistical approach used in this study employs the methodology of principal components[8] and the related two-block partial least-squares (PLS) method.[9] The relevance of using

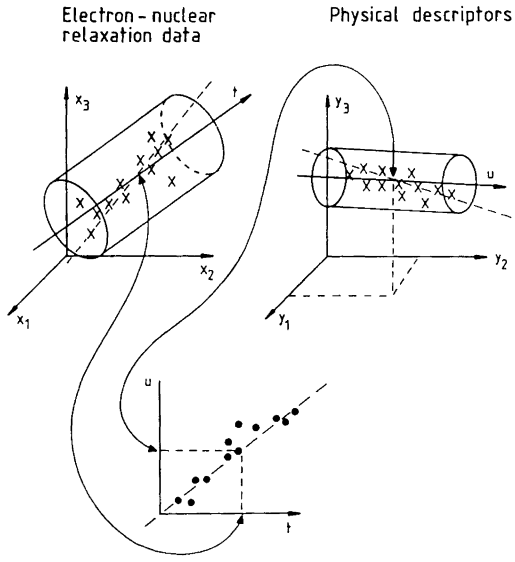

Fig. 1. A geometrical interpretation of the PLS model. The **X** block is formed by the electron-nuclear relaxation data and the **Y** block is represented by the physical descriptors. In each space, the data are modelled as lines. The projections of the object points drawn on these models have the coordinates **T** and **U** for the **X** and **Y** parts of the data, respectively. Each column in **U** is related to the corresponding column in **T** by a linear model (bottom).

these methods for interpreting NMR data has already been recognized.[6,9] The algorithms will be only briefly described here. Details can be found elsewhere.[5]

Principal components analysis. Data x_{ik} with K variables ($k = 1,2,\ldots,K$) for compounds ($i = 1,2,\ldots,n$) were collected in the matrix **X**. This matrix **X** is modelled by the t-dimensional model as shown in eqn. (2)

$$x_{ik} = \bar{x}_k + \sum_{a=1}^{A} t_{ia}\, p_{ak} + e_{ik} \qquad (2)$$

where $\bar{x}_k$ is the mean value of variable k, p_{ak} is the loading of variable k in dimension a, t_{ia} the component score value of compound i in dimension a and e_{ik} represents the residuals. In matrix form this is:

$$\mathbf{X} = 1 \cdot x + \mathbf{TP} + \mathbf{E}. \qquad (3)$$

Cross-validation is used to determine the appropriate number of terms, A (columns in **T** and **P**), i.e. the predictive ability of the model is maximized.[10]

PLS analysis decomposes two connected block matrices into score matrices times loading matrices, similarly to the PC model. The PLS model is computed so as to improve correlations between the two-block score vectors of the same component (Fig. 1). This corresponds to the approximation of the data by linear geometrical structures. The model constructs a tolerance interval in x-space, where each object has a maximum allowed residual standard deviation. In x-space this is illustrated by a cylinder, the length of which is determined by the range of the class object points. This gives a tool to relate, for example, NMR relaxation data to physical descriptors that are relevant in the case in question. This approach also has a predictive potential, as is shown later.

The initial statistical (PCA) analysis was performed in order to account for the systematic variation in each of the two measured NMR data sets, I and II, T_1^e and induced shifts, respectively. Set I was scaled to unit variance (global scaling), to ensure equal weight of each variable in the modelling. In set II, global scaling would have introduced additional noise due to the increased importance of the shift values with small initial variance. Hence, no scaling of shift data was undertaken. The first analysis was applied on the relaxation data set (I). A two component model

Table 2.

Physical descriptor	T_1^e/s		Induced shift/ppm	
	Residual variance after A=1	Residual variance after A=2	Residual variance after A=1	Residual variance after A=2
Π^a	0.96	0.33	1.01	0.84
Dipole moment	0.25	0.24	0.31	0.01
ε^a	0.91	0.97	0.94	0.44
Bp^{760}	0.62	0.64	0.61	0.52
d^{20}	0.74	0.06	0.64	0.50
n_D^{20}	0.67	0.25	0.59	0.62
M_w	0.83	0.06	0.73	0.69
$\lg P(H)^b$	0.96	0.36	1.01	1.12
v_W	1.08	1.11	1.08	0.99
σ_m	0.20	0.22	0.23	0.23
σ_p	0.28	0.24	0.33	0.35
C-1[c]	0.98	0.58	0.91	0.19
C-2[c]	0.94	0.91	0.98	0.23
C-3[c]	1.02	0.68	1.00	1.09
C-4[c]	0.44	0.30	0.49	0.49

[a]See Ref. 7d. [b]See Ref. 7b. [c] ^{13}C chemical shift.

described 89 % of the systematic variation in T_1^e, the major first component accounting for 46 %.

The PC analysis of the induced shift data (II) resulted in a somewhat weaker model. Of the shift variance, 37 % could be explained by the first component and 60 % by the two significant components together. Thus, the first step in the data analysis (PC), showed that the data sets contained systematic information comprising from 60 % to 89 %.

Partial least-squares data analysis. The next step was to investigate by PLS whether relevant information from the physical descriptor matrix (III) could "explain" the relaxation (I) and induced shift data (II). This analysis showed that 51 % of the variance in relaxation data (I) could be described by the first component; 68 % was described by the two significant components together. The induced shift set (II) was modelled to 54 %, also by two significant components, where the first component was the dominant one and described 45 % of the variance. Table 2 shows the residual variance of each variable for the two components describing the T_1^e data and the shift matrix. The residual variance accounts for the unmodelled, unsystematic data.

These results show that the separation of the first and second component can be interpreted as two separate effects, both being important for the interaction with the PARR. For the NMR data, the first component is related mainly to C-4 while component 2 relates to C-1 and C-3 (T_1^e) and C-1 and C-2 (induced shifts). The first dominant component in both PLS models uses the same physical descriptors to explain shift and relaxation data. Descriptors such as the dipole moment, σ_m and σ_p are most important in the first component. Hence, variables describing electronic structure and size (steric information) seem to be of relevance for the first component.

The second component contains less information, but it is interesting to note that the importance of the different physical descriptors (decrease in residual variance) is not the same for the shift and relaxation matrices. Descriptors such as log *P*, density, n_D , M_w and log *P*(H) are of greatest explanatory value for T_1^e. Correspondingly, the dipole moment, dielectric constant, ε, and the *ipso* and *ortho* chemical shifts form the second component in the induced shift matrix. It is thus possible by a multivariate statistical analysis to identify one common effect that describes both the relaxation and the induced shift data. The second components model different effects for the two data sets, thus separating the two NMR chemical shift and relaxation parameters.

Table 3.

Compound	T_1^e/s	T_1^e(exp.-pred.)/s	T_1^e/s	T_1^e(exp.-pred.)/s
	meta		*para*	
1	0.94	−0.13	0.97	−0.07
2	0.68	−0.33	0.61	−0.38
3	0.71	−0.03	0.95	0.19
4	0.77	0.04	0.67	−0.08
5	1.71	−0.25	1.53	−0.29
6	2.48	0.44	2.11	0.22
7	0.66	−0.10	0.54	−0.22
8	0.51	−0.16	0.59	0.22
9	2.29	−0.12	2.14	−0.02
10	1.36	−0.09	1.38	−0.01
11	1.98	0.21	1.98	0.33
12	0.31	0.02	0.27	−0.01
13	0.65	0.15	0.59	0.09

Five descriptors are especially significant in the derivation of the PLS model using induced shift data. All these variables (dipole moment, σ_m, σ_p, and ^{13}C chemical shifts) describe the electronic distribution in the substrate. The T_1^e PLS model is based mostly on descriptors that can be related to the molecular motion, such as the density, the molecular weight and the partition coefficient.

PLS prediction. To further illustrate the usefulness of this multivariate approach, it was decided to test how well the electron-nuclear relaxation times could be predicted from a combination of the induced shift set II and the set of physical descriptors III. These latter sets formed the **Y** block as in Fig. 1. By using two significant components, a total of 73 % of the variance in relaxation data could be explained.

In Table 3, the predictive and the measured values of T_1^e are tabulated for the *meta* and *para* carbon positions of in all 13 monosubstituted benzenes. The results show a good correlation between the measured and predicted values. The difference in T_1^e (exp.–pred.) is in most cases less than or close to the overall experimental error

(10–15 %). The given example also illustrates future possibilities for choosing suitable substrate systems for homogenous catalysis. By relating NMR parameters that reflect a desired type of interaction with the chelate to molecular properties of the substrate, it would be possible to predict the catalytic activity of a given chelate in reaction with substrates.

Conclusions

In solutions of monosubstituted aromatics containing paramagnetic relaxation reagents, an electronic factor described by the dipole moment of the substrate and Hammett sigma constants accounts for most of the variance of the electron-nuclear relaxation times and the induced chemical shifts. A second effect, different for the two NMR parameters, could also be identified. Electronic descriptors constitute the second component for the shift matrix, while variables related to motion and solubility are more significant for the relaxation data. The results support our previous interpretation that the preferred solvation is of a dipole–dipole type. A partial least-squares method was used to predict the electron-relaxation times from the induced chemical shifts and various physical descriptors of the aromatic substrates.

Acknowledgements. The authors are indebted to Dr. Svante Wold, Umeå University, Umeå, Sweden, for valuable discussions. Grants from the Swedish Natural Science Research Council are gratefully acknowledged.

References

1. (a) Langford, C. H. and Stengle, T. R. In: La Mar, G. N., Horrocks, W. D., Jr. and Holm, R. H., Eds., *NMR of Paramagnetic Molecules,* Academic Press, New York 1973, p. 372; (b) Levy, G. C., Edlund, U. and Holloway, C. E. *J. Magn. Reson. 24* (1976) 375; (c) Kitaigorodski, A. N., Nekipelov, V. M. and Zamarev, K. I. *Zh. Struct. Khim. 19* (1978) 796.
2. (a) Eaton, D. R. *Can. J. Chem. 47* (1969) 2645; (b) Frankel, L. S., Langford, C. H. and Stengle, T. R. *J. Phys. Chem. 74* (1970) 1376.
3. Lyerla, J. R. and Levy, G. C. In: Levy, G. C., Ed., *Topics in Carbon-13 NMR Spectroscopy,* Wiley, New York 1974, Vol. 1, p. 121.
4. (a) Grahn, H., Edlund, U. and Levy, G. C. *J. Magn. Reson 56* (1984) 61; (b) Holak, T. A., Aksnes, D. W., Grahn, H. and Edlund, U. *Magn. Reson. Chem. 24* (1986) 575.
5. Wold, S., Albano, C., Dunn, W. J., Edlund, U., Esbensen, K., Geladi, P., Hellberg, S., Johansson, E., Lindberg, W. and Sjöström, M. *Chemometrics Mathematics and Statistics in Chemistry* (NATO ASI Series C No. 138), D. Reidel, Dordrecht, The Netherlands 1984, p. 17.
6. Johnels, D., Edlund, U., Grahn, H., Hellberg, S., Sjöström, M., Wold, S., Clementi, S. and Dunn, W. J. *J. Chem. Soc., Perkin Trans. 2* (1983) 863 and references therein.
7. (a) *Handbook of Chemistry,* CRC Publishing Co., Cleveland, Ohio 1971; (b) Seydel, J. K. and Schaper, K. J. *Chemische Struktur und Biologische Aktivität von Wirkstoffen,* Verlag Chemie, Weinheim and New York 1979; (c) Ewing, D. F. *Org. Magn. Reson. 12* (1979) 499; (d) Dunn, W. J., Johansson, E. and Wold, S. *Quant. Struct.-Act. Relat. 2* (1983) 156.
8. Albano, C., Dunn, W. J., Edlund, U., Johansson, E., Nordén, B., Sjöström, M. and Wold, S. *Anal. Chim. Acta 103* (1978) 429.
9. Hellberg, S., Wold, S. and Dunn, W. J. In: Höskuldsson, A. and Espensen, K., Eds., *Nordic Symposium in Applied Statistics and Data Processing,* NEUCC, Det regionale edb-center, Danmarks Tekniske Højskole, Lyngby, Denmark 1982.
10. Geisser, S. *J. Am. Statist. Assoc. 70* (1975) 320.

Received August 15, 1986.