

A priori calculation of chemical shifts in ^{19}F NMR spectroscopy. 3. Data treatment for saturated molecules
 Calcul *a priori* des déplacements chimiques en RMN du ^{19}F . 3. Traitement informatique pour les molécules saturées

Gérard Bauduin*, Yves Piétrasanta

Laboratoire de Chimie Appliquée, URA CNRS D 11930, Ecole Nationale Supérieure de Chimie de Montpellier, 8, Rue Ecole Normale, 34053 Montpellier-Cedex 1 (France)

et Mohamed Belbachir

Laboratoire de Chimie des Polymères, Institut de Chimie, Université d'Oran Es Senia, B.P. 1524 El Menaouar, 31000 Oran (Algérie)

(Reçu le 8 mars 1992; accepté le 7 septembre 1992)

Abstract

This paper presents two programs which allow the ready empirical calculation of fluorine chemical shifts. The first program deals with the management of paramagnetic increment files and the second with the calculation itself.

Résumé

Dans cet article, nous présentons les deux programmes qui permettent de calculer empiriquement et très facilement les déplacements chimiques du fluor. Le premier programme concerne la gestion des fichiers d'incrément paramagnétiques et le second le calcul proprement dit.

Introduction

Nous avons montré dans deux précédentes publications [1, 2] que le déplacement chimique du noyau fluor dans les molécules saturées pouvait être calculé de façon empirique, avec une bonne précision, comme étant la somme d'un terme diamagnétique et d'un terme paramagnétique. Le terme diamagnétique est proportionnel à la charge nette de l'atome de fluor considéré. Il est très voisin de 2×10^{-6} . Le terme paramagnétique, en général le plus important, est calculé comme étant la somme d'incrément paramagnétiques caractéristiques de chaque substituant distant de moins de six liaisons du noyau considéré, ainsi que de leur position dans la chaîne et du degré de substitution du carbone qui les porte. En général, l'incrément paramagnétique du substituant X peut s'écrire de la façon suivante:

$$\begin{matrix} \text{(C ou E)} & \text{P}^{\text{z}} \\ \text{(m, d ou t)} & \text{P}^{\text{X}} \end{matrix}$$

*Auteur auquel la correspondance doit être adressée.

où n est le nombre de liaisons entre X et le noyau considéré. Les autres indices caractérisent le carbone qui porte X: (i) l'indice E signifie que X est lié à un atome de carbone en extrémité de chaîne et C, à un carbone de la chaîne; et (ii) les indices m, d et t sont employés selon que ce carbone est lié à un, deux ou trois substituants autres que le fluor ou un carbone de la chaîne. Ainsi, il ne peut pas exister d'incrément c_tP , ni ${}^E_tP^2$, ni ${}^c_dP^2$.

Le calcul du terme paramagnétique, qui prend en compte tous les substituants situés à cinq liaisons au plus du noyau F considéré, devient assez rapidement fastidieux. Par ailleurs, s'il existe de nombreux substituants différents, il faut avoir recours aux tableaux de valeurs des incréments pour chacun d'entre eux. Si l'on ajoute à cela les risques d'erreurs qui augmentent avec la taille de la molécule, on voit combien il est intéressant de disposer d'une méthode informatisée sûre pour effectuer ces calculs.

Naturellement, la première étape de l'informatisation consiste à organiser et construire les fichiers d'incréments paramagnétiques, qui pourront être consultés par le programme principal de calcul. Ces fichiers qui contiennent les valeurs des incréments pour chacun des substituants doivent être stockés sur disquette et rapidement accessibles. La programmation étant effectuée en basic sur microordinateur APPLE II, nous avons choisi d'utiliser la procédure des fichiers à accès direct.

Structure et organisation des fichiers d'incréments paramagnétiques

Nous fixons tout d'abord la longueur de l'enregistrement pour chaque incrément à cinq caractères, ce qui correspond à au moins trois chiffres significatifs, de façon tout à fait compatible avec la précision attendue.

La solution la plus simple serait alors de créer un fichier par substituant, mais au détriment des temps d'accès à la mémoire. Le problème majeur reste le positionnement des enregistrements dans le fichier en fonction des trois facteurs de définition de chaque incrément: nombre de liaisons n , position dans la chaîne et degré de substitution: n peut prendre les quatre valeurs entières 2, 3, 4 et 5. La position est définie de façon alternative: soit dans la chaîne, soit en extrémité. Enfin, le facteur de substitution présente trois niveaux: mono-, di- ou tri-substitué. Dans ces conditions, si on tient compte du fait que les incréments c_tP ne peuvent pas exister, il reste vingt combinaisons possibles. (Voir Tableau 1.)

Nous créons donc un seul fichier, appelé PARA, contenant tous les incréments connus par enregistrement de cinq caractères. L'adresse 0 est réservée, les adresses 1 à 20 sont attribuées au premier substituant, les adresses 21 à 40 au deuxième et ainsi de suite. Pour préciser l'ordre d'attribution des blocs de 20 enregistrements dans le fichier PARA, nous créons un second fichier, appelé CATALOGUE où l'adresse 0 est réservée pour le nombre d'enregistrements dans le fichier, l'adresse 1 contenant le nom (aussi proche que possible de la formule chimique) du premier substituant,

TABLEAU 1

Localisation des incréments paramagnétiques en fonction des indices de définition

Incréments	Indice I	Indice J	Indice K	Adresse $U = I + J + K + 1$
${}^E_m P^2$	0	0	0	1
${}^E_d P^2$	0	0	8	9
${}^E_t P^2$	0	0	16	(17)
${}^C_m P^2$	0	4	0	5
${}^C_d P^2$	0	4	8	(13)
${}^C_t P^2$	—	—	—	—
${}^E_m P^3$	1	0	0	2
${}^E_d P^3$	1	0	8	10
${}^E_t P^3$	1	0	16	18
${}^C_m P^3$	1	4	0	6
${}^C_d P^3$	1	4	8	14
${}^C_t P^3$	—	—	—	—
${}^E_m P^4$	2	0	0	3
${}^E_d P^4$	2	0	8	11
${}^E_t P^4$	2	0	16	19
${}^C_m P^4$	2	4	0	7
${}^C_d P^4$	2	4	8	15
${}^C_t P^4$	—	—	—	—
${}^E_m P^5$	3	0	0	4
${}^E_d P^5$	3	0	8	12
${}^E_t P^5$	3	0	16	20
${}^C_m P^5$	3	4	0	8
${}^C_d P^5$	3	4	8	16
${}^C_t P^5$	—	—	—	—

et ainsi de suite. Pour des raisons tenant à l'utilisation ultérieure, la longueur de l'enregistrement dans le fichier CATALOGUE est limitée à quatre caractères. Nous affectons l'indice N au Nième substituant du fichier CATALOGUE.

Il reste à résoudre le problème de la localisation des 20 incréments dans les 20 adresses disponibles pour chaque substituant. Cette localisation doit dépendre de façon univoque des facteurs de définition des incréments paramagnétiques. Nous choisissons une procédure d'adressage 'pseudo-binaire' liée à trois indices I , J et K dépendant des facteurs de définition. L'indice I est égal au nombre de liaisons n diminué de 2, il varie ainsi entre 0 et 3 (00 et 11 en binaire). L'indice J vaut 0 si le substituant est sur un carbone en extrémité de chaîne et 4 (respectivement 000 et 100 en binaire) pour un substituant porté par un carbone de la chaîne. Pour un carbone monosubstitué par un substituant autre que le fluor ou un carbone de la chaîne, on pose $K=0$, pour un carbone disubstitué $K=8$ et pour un carbone

TABLEAU 2

Localisation des incréments paramagnétiques dans chaque bloc de 20 enregistrements attribué à chaque substituant

Incréments	Nombre de liaisons			
	2	3	4	5
E_mP	1	2	3	4
C_mP	5	6	7	8
E_dP	9	10	11	12
C_dP	(13)	14	15	16
E_tP	(17)	18	19	20

trisubstitué $K=16$ (respectivement 00000, 01000 et 10000 en binaire). A un facteur additif multiple de 20 près, l'adresse U de chaque incrément est donnée alors par la formule:

$$U=I+J+K+1$$

Le Tableau 1 montre comment fonctionne l'adressage ainsi conçu et le Tableau 2 présente l'occupation mémoire de 20 enregistrements pour un substituant. Il est à noter que, dans tous les cas, les adresses 13 et 17 qui correspondent aux incréments ${}^C_dP^2$ et ${}^E_tP^2$ sont vides. Mais nous avons préféré privilégier la facilité de gestion, même au détriment de l'occupation mémoire. Pour le substituant fluor, dont l'incrément ne dépend que de n [2], des valeurs identiques sont introduites dans le fichier à toutes les adresses correspondant à la même valeur de n .

En définitive, si on désire localiser un incrément quelconque correspondant au substituant en position N dans le fichier CATALOGUE, son adresse sera donnée par:

$$U=I+J+K+1+(N-1)\times 20$$

dans le fichier PARA. Autrement dit, l'ordre des blocs successifs de 20 enregistrements du fichier PARA est lu directement comme l'ordre des enregistrements du fichier CATALOGUE.

Programme de gestion des fichiers

Dans les messages du programme l'expression 'terme paramagnétique' remplace les mots 'incrément paramagnétique' plus longs. Par ailleurs, le terme 'fichier' y désigne l'ensemble des 20 enregistrements correspondant aux incréments paramagnétiques d'un substituant. Nous le distinguerons en l'appelant fichier-substituant.

Le programme de gestion des fichiers comporte trois options: (i) création de fichiers-substituant; (ii) modification de fichiers-substituant; et (iii) listing d'un ou des fichiers-substituant.

Dans l'option création de fichier-substituant, le programme édite tout d'abord à l'écran le contenu du fichier CATALOGUE (liste des substituants existants) et demande à l'utilisateur s'il désire toujours créer un nouveau fichier. Dans l'affirmative, il faut alors introduire le nom choisi pour le substituant correspondant. La longueur de ce nom ne doit pas excéder quatre caractères, le programme en effectue le contrôle et, en cas de dépassement, demande un nouveau nom. En général, le nom est choisi aussi évocateur que possible, par exemple, F pour le fluor, CL pour le chlore, CO2H pour le groupement carboxyle, C2F5 pour le perfluoroéthyle, etc. Le nombre d'enregistrements du fichier CATALOGUE est alors augmenté d'une unité et stocké à l'adresse 0 du fichier, le nom du nouveau substituant est inscrit à la dernière adresse N du fichier.

Le programme demande alors à l'utilisateur la valeur de chaque incrément défini automatiquement en clair sur l'écran par le programme (Fig. 1). Il mémorise les valeurs des indices I , J et K correspondants et précise de donner la valeur 99999 pour un incrément indéterminé. Il contrôle également que la longueur de la valeur donnée n'excède pas cinq caractères. La valeur est enfin stockée dans le fichier PARA à l'adresse $U = I + J + K + 1 + (N - 1) \times 20$. Ainsi l'option création de fichier-substituant correspond en fait à une extension des deux fichiers CATALOGUE (+1 enregistrement) et PARA (+20 enregistrements).

Lorsqu'on souhaite modifier un fichier-substituant, le programme lit, comme dans le cas précédent, le fichier CATALOGUE et écrit la liste des

	1	5	10	15	20	25	30	35	40
1	VALEUR DU TERME PARAMAGNETIQUE POUR								
	L'ATOME OU LE GROUPEMENT <u>(1)</u>								
5	- SITUÉ A <u>(2)</u> LIAISONS DE L'ATOME								
	LIE A UN ATOME DE CARBONE								
10	<u>(3)</u>								
	ET <u>(4)</u> SUBSTITUÉ PAR DES								
	SUBSTITUANTS AUTRES QUE LE FLUOR :								
	<u>(5)</u>								
15	(POUR UNE VALEUR INDETERMINEE,								
	DONNEZ 99999)								
20									

Fig. 1. Schéma de l'écran pour l'entrée des termes (incréments) paramagnétiques. (1) Nom du substituant. (2) Nombre de liaisons entre le substituant et l'atome de fluor considéré. (3) 'en extrémité' de chaîne' ou 'dans la chaîne carbonée'. (4) 'mono', 'di' ou 'tri'. (5) position du curseur pour l'entrée de la donnée.

substituants à l'écran. Il demande ensuite le nom du substituant à modifier, dont il retient l'adresse N dans le fichier CATALOGUE.

L'incrément paramagnétique à modifier doit alors être précisé en fonction des trois facteurs de définition habituels. Les questions concernant ces facteurs apparaissent par des chiffres ou par O (oui) ou N (non). Le curseur se positionne à l'endroit où doit être donnée la réponse (Fig. 2). Le programme contrôle strictement les réponses données et redemande, en cas d'erreur, une nouvelle réponse. Ainsi:

1. Le nombre de liaisons entre le substituant et le fluor étudié ne peut être inférieur à 2 ou supérieur à 5.

2. Le carbone porteur du substituant ne peut être à la fois dans la chaîne carbonée et en extrémité de chaîne. En cas d'erreur, le curseur remonte à la première question (position 2 de la Fig. 2).

3. Un carbone de la chaîne ne peut pas être trisubstitué par des substituants autres que le fluor ou un carbone de la chaîne. En cas d'erreur, le curseur remonte à la question concernant la position du carbone (position 2 de la Fig. 2).

4. Le nombre de substituants autres que le fluor ou un carbone de la chaîne doit être compris entre 1 et 3.

Au fur et à mesure des réponses, le programme affecte les valeurs correspondantes aux indices I , J et K et va chercher à l'adresse $U=I+J+K+1+(N-1)\times 20$ l'ancienne valeur de l'incrément qu'il affiche à l'écran.

Si cette valeur est 99999, il précise qu'elle est indéterminée. Il demande alors immédiatement (Fig. 2) la nouvelle valeur qu'il inscrit à l'adresse U du fichier PARA après avoir vérifié que sa longueur n'excédait pas cinq

	1	5	10	15	20	25	30	35	40
1	LE	TERME	PARAMAGNETIQUE	QUE	VOUS	VOULEZ			
	CORRIGER	CORRESPOND	A	UN	_(a)_	SITUE			
5	A	(1)	LIAISONS	DE	L'ATOME	DE	FLUOR	ETUDIE	
	SUR	UN	CARBONE	:					
	-	DANS	LA	CHAINE	CARBONEE	O	OU	N?	- (2)
	-	EN	EXTREMITE	DE	CHAINE	O	OU	N?	- (3)
10	-	SUBSTITUE	PAR	(4)	SUBSTITUANTS	AUTRES			
	QUE	LE	FLUOR	.					
	D'	ACCORD	?	(O	OU	N)	-	(5)	
15		ANCIENNE	VALEUR	:	_(b)_	(INDETERMINE)			
		NOUVELLE	VALEUR	:	(6)				
20									

Fig. 2. Schéma de l'écran pour la modification des termes (incréments) paramagnétiques. (a) Nom du substituant. (b) Ancienne valeur de l'incrément en fichier. (1), (2), (3), (4), (5) et (6): positions successives du curseur.

TABLEAU 3

Lecture du fichier-substituant du chlore

<i>Groupe ment CL</i>			
Nombre de liaisons entre F et CL	CL sur un carbone	Substituants autres que le fluor sur ce carbone	Terme paramagnétique (en ppm)
2	de l'extrémité	1	- 3.00
2	de l'extrémité	2	7.000
2	de l'extrémité	3	ind
2	de la chaîne	1	25.70
2	de la chaîne	2	ind
3	de l'extrémité	1	11.00
3	de l'extrémité	2	17.00
3	de l'extrémité	3	13.00
3	de la chaîne	1	9.100
3	de la chaîne	2	11.30
4	de l'extrémité	1	.2000
4	de l'extrémité	2	3.800
4	de l'extrémité	3	2.400
4	de la chaîne	1	3.100
4	de la chaîne	2	3.100
5	de l'extrémité	1	-2.30
5	de l'extrémité	2	3.700
5	de l'extrémité	3	.5000
5	de la chaîne	1	6.900
5	de la chaîne	2	- 3.50

caractères. Il est alors possible de corriger un autre incrément pour le même substituant ou de revenir en début de programme.

La lecture des fichiers-substituant est réalisé sur imprimante. Le programme donne d'abord le choix entre la lecture de l'ensemble des fichiers-substituant ou un seul d'entre eux. Nous donnons l'exemple d'une lecture de fichier dans le Tableau 3. On note que les deux incréments ${}^E_P^2$ et ${}^C_P^2$ qui n'existent pas sont indiqués comme indéterminés.

Une fois les fichiers des incréments paramagnétiques établis, ils peuvent être utilisés pour le calcul des déplacements chimiques.

Programme de calcul des déplacements chimiques du fluor

Nous souhaitons présenter au chimiste un logiciel qui soit aussi proche que possible de ses connaissances et de ses habitudes et qui ne fasse appel à aucune connaissance informatique. Naturellement, le travail de programmation s'en trouve accru.

Le programme indique tout d'abord qu'il ne s'applique qu'à des molécules saturées, contenant exclusivement les substituants lus dans le fichier CATALOGUE. Vient ensuite le module de définition de la molécule qui peut être décomposé en trois éléments correspondant chacun à une page écran.

Le premier élément comporte deux questions concernant le nombre d'atomes de carbone de la chaîne N_C et le numéro N_F dans la chaîne du carbone portant le fluor étudié. Ces deux entrées sont contrôlées: N_C doit être supérieur à 1 et N_F doit être inférieur ou égal à N_C , sinon la question correspondante est reformulée.

Les deux éléments suivants concernent la définition de la structure de la molécule. Le premier élément concerne la gauche du fluor étudié et le second, la droite. Le programme écrit sur l'écran, dans les deux cas (Figs. 3 et 4), la structure éclatée de la partie correspondante et le curseur se positionne à tour de rôle à la place des substituants, dans l'ordre de haut en bas en partant de la gauche. En bas de l'écran sont rappelés les substituants disponibles du fichier CATALOGUE. Naturellement, seuls sont représentés dans la structure les carbones porteurs de substituants à moins de six liaisons du fluor étudié. Si le premier substituant (pour la partie gauche) ou le dernier (pour la partie de droite) est la suite de la chaîne carbonée, le programme peut le détecter en comparant N_C et N_F et écrit automatiquement R à la place du premier ou du dernier substituant respectivement. Au fur et à mesure de l'introduction des substituants, le programme détermine pour chacun d'eux les indices I , J et K en effectuant des tests sur les substituants. Concrètement, la partie gauche de la molécule apparaît d'abord à l'écran, puis lorsque tous les substituants ont été introduits, l'écran s'efface et fait apparaître la partie droite. Dans les deux cas, l'atome de fluor étudié apparaît en mode inverse sur l'écran. Le programme demande ensuite confirmation de la structure écrite et, en cas de réponse négative, revient au début du module de définition de la molécule. Remarquons que c'est la représentation de la structure sur l'écran qui nous a contraint à limiter à 4 caractères la longueur des noms des substituants.

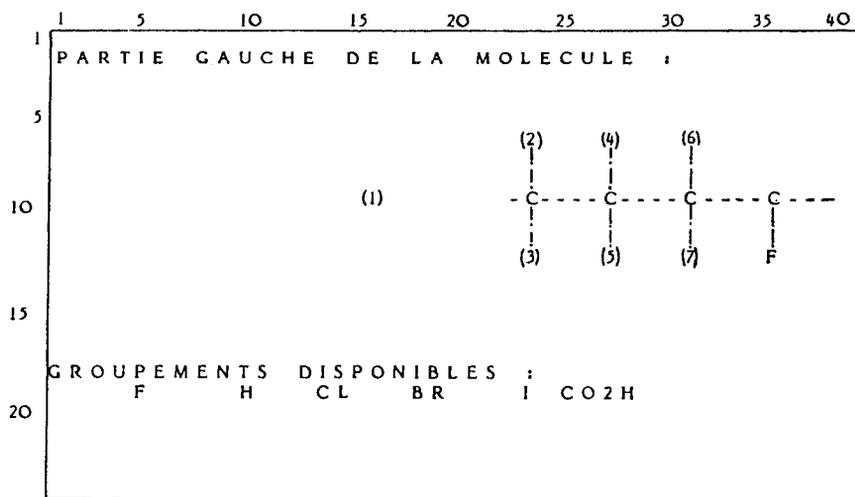


Fig. 3. Schéma de l'écran pour l'entrée des substituants sur la partie gauche de la molécule. (1) à (7): positions successives du curseur sur l'écran

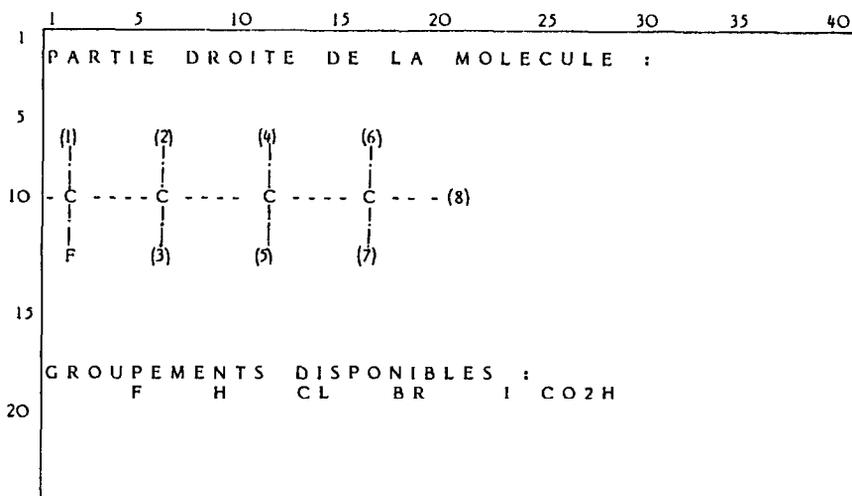


Fig. 4. Schéma de l'écran pour l'entrée des substituents de la partie droite de la molécule. (1) à (8): positions successives du curseur

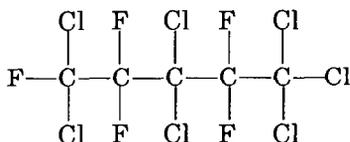
Si la structure est confirmée, le programme entre dans le module de calcul. Il cherche d'abord l'adresse N de chaque substituant dans le fichier CATALOGUE. Si un substituant n'existe pas, un message d'erreur apparaît. Les indices, I , J , K et N permettent de chercher l'incrément paramagnétique de chaque substituant dans le fichier PARA. Si un incrément est indéterminé, un message apparaît sur l'écran.

Dans le cas contraire, le module d'édition sur imprimante est exécuté. C'est d'abord la formule éclatée globale, regroupant les parties gauche et droite de la molécule qui est présentée. Le fluor étudié est distingué (Tableau 4). Il apparaît ensuite un tableau où tous les substituents, présentés dans l'ordre de haut en bas en passant de la gauche, sont caractérisés par leurs trois facteurs de définition (sauf le fluor). La première colonne indique le nom du substituant, la deuxième, le nombre de liaisons entre le fluor étudié et le substituant, la troisième, la place du carbone portant le substituant (dans la chaîne ou en extrémité) et la quatrième, le degré de substitution de ce carbone par des substituents autres que le fluor ou un carbone de la chaîne. Le programme effectue alors un retour à l'écran et demande à l'utilisateur s'il connaît la charge nette du fluor étudié. En cas de réponse positive, la charge nette doit être entrée et le terme diamagnétique est calculé. Sinon le terme diamagnétique est pris égal à 2. Le programme retourne alors à l'imprimante où il imprime éventuellement la charge nette, puis les termes diamagnétique et paramagnétique, ainsi que le déplacement chimique calculé. A titre d'exemple, le Tableau 4 donne les résultats imprimés pour l'atome de fluor souligné de la molécule $\text{CFCl}_2\text{-CF}_2\text{-CCl}_2\text{-CF}_2\text{-CCl}_3$.

TABLEAU 4

Exemple de calcul de déplacement chimique du fluor

Formule de la molécule



Caractérisation des substituants

Groupement	Nombre de liaisons	Place dans la chaîne	Substitution
F	5	extrémité	-
CL	5	extrémité	di
CL	5	extrémité	di
F	4	dans	-
F	4	dans	-
CL	3	dans	di
CL	3	dans	di
F	2	dans	-
CL	3	extrémité	tri
CL	3	extrémité	tri
CL	3	extrémité	tri

Les atomes sont dans l'ordre de gauche à droite et de haut en bas.

Pour le fluor encadré:

Charge nette totale: -1.32

Terme diamagnétique: 2.112

Terme paramagnétique: 95.7

Je calcule, pour ce fluor, un déplacement chimique de:

97.81

(le déplacement chimique est donné en ppm par rapport à CFCl_3)

Conclusion

Les programmes que nous avons mis au point permettent au chimiste, même ignorant du modèle que nous avons proposé, de calculer le déplacement chimique d'un atome de fluor dans n'importe quelle molécule saturée comportant les substituants pour lesquels les incréments paramagnétiques ont été préalablement déterminés. Nous avons pris également le maximum de précautions pour qu'un chimiste, même peu au fait de l'utilisation d'un clavier puisse utiliser facilement le programme sans risque d'erreur. De plus, l'utilisateur connaissant la méthode de calcul trouve dans l'impression des résultats les éléments nécessaires pour vérifier le bon fonctionnement du programme.

En résumé, ce programme, même s'il peut sans nul doute être amélioré, nous paraît posséder les qualités nécessaires de souplesse et de sécurité, qui le rendent accessible au plus large public.

References

- 1 A. Battais, G. Bauduin, B. Boutevin et Y. Pietrasanta, *J. Fluorine Chem.*, 31 (1986) 197.
- 2 G. Bauduin, M. Belbachir, A. Benzaza et Y. Pietrasanta, *J. Fluorine Chem.*, 52 (1991) 277.