

CHROM. 6813

IDENTIFICATION DE COMPOSÉS ALIPHATIQUES SATURÉS PAR CHROMATOGRAPHIE EN PHASE GAZEUSE

I. COMPOSÉS À UNE DIRECTION DE DÉVELOPPEMENT

MAURICE CHASTRETTE

Laboratoire de Chimie Organique Physique, Faculté des Sciences, 43 Bd du 11 Novembre 1918, 69100 Villeurbanne (France)

GÉRARD LENFANT

Laboratoire de Chimie Organique, E.N.S.C.L., Cité Scientifique, BP 40, 59650 Villeneuve D'Ascq (France)

ANNIE REMY

Laboratoire de Calcul, U.E.R. d'I.E.E.A., Université des Sciences et Techniques, Cité Scientifique, BP 36, 59650 Villeneuve D'Ascq (France)

et

MOÏSE COHEN-MAKABEH

Laboratoire de Chimie Organique Physique, Faculté des Sciences, 43 Bd du 11 Novembre 1918, 69100 Villeurbanne. (France)

(Reçu le 14 mars 1973)

SUMMARY

Identification of aliphatic saturated compounds by means of gas chromatography. I. Compounds developed in one direction

Retention-structure relations are used to identify the carbon chain R of saturated aliphatic compounds obeying the general formula RX (X being a monovalent chemical function). From a limited number of results, a program, written in Algol 60, enables us to propose one or more structures for a compound whose family and retention time are known. The organigram of this program is presented.

INTRODUCTION

L'identification de composés chimiques analysés à l'aide d'un appareil de chromatographie en phase gazeuse est un problème fréquemment rencontré. Les méthodes incrémentales¹⁻¹¹, parfois utilisées pour le résoudre, permettent d'obtenir rapidement une idée de la structure de composés dont on connaît la fonction chimique mais demeurent assez imprécises.

Il a été montré précédemment que l'emploi du système topologique DARC mis au point par J. E. Dubois et coll.¹²⁻¹⁵ permet de relier de manière très satisfaisante, par des relations de topologie-information, les grandeurs de rétention (temps de rétention relatifs, indices de Kováts) et les structures de solutés appartenant à plusieurs familles: cétones¹⁶⁻¹⁸, esters¹⁸⁻²⁰, aldéhydes²¹, nitriles²², thiols²² et éther-oxydes²³.

Nous montrons dans cet article comment les relations de topologie-information établies précédemment peuvent être utilisées, avec l'aide d'un ordinateur, pour l'identification de composés dont on connaît la fonction chimique.

PRINCIPE DE LA MÉTHODE

L'identification d'un composé appartenant à une famille chimique procède selon deux étapes.

La première étape est l'établissement de la relation de topologie-information dont nous rappellerons seulement le principe, les détails ayant été exposés antérieurement¹⁶⁻²⁰. On choisit une population de définition, comprenant un petit nombre de composés de la famille étudiée dont on mesure les grandeurs de rétention, dans des conditions chromatographiques précises, à plusieurs températures. On exprime la structure de ces composés dans le langage du système DARC et on la relie à leurs grandeurs de rétention par une relation de topologie-information dont on détermine les paramètres par la méthode des moindres carrés. Puis en introduisant la relation entre ces paramètres et la température on obtient pour tout le domaine de températures étudié une seule relation rétention-structure générale.

La deuxième étape est l'utilisation de la relation rétention-structure générale qui permet de calculer pour une température quelconque du domaine étudié les grandeurs de rétention de tout composé inclus dans l'empreinte^{15,17} de la population de définition. Une technique simple d'identification consisterait à calculer, pour chacun des composés de la famille inclus dans l'empreinte, ses grandeurs de rétention, et à les comparer aux grandeurs observées, en constituant de ce fait un catalogue. Pour éviter d'avoir à dresser un tel catalogue, nous avons mis au point un programme de calcul qui permet d'obtenir directement les structures des composés dont les grandeurs de rétention calculées sont proches des valeurs expérimentales. Pour cela on introduit dans la mémoire de l'ordinateur les paramètres de la relation de topologie-information, le nom de la famille à laquelle appartient le composé, la température et la grandeur de rétention observée. Le traitement fournit une liste de composés dont la structure est compatible avec la grandeur de rétention observée, aux erreurs d'expérience près.

Nous présentons ici les résultats obtenus dans le cas des familles de composés ne comportant qu'une seule direction de développement (composés de formule générale RX , X étant une fonction chimique quelconque). Le cas beaucoup plus complexe de composés ayant deux directions de développement sera traité ultérieurement²⁴.

DÉTERMINATION DES POSITIONS OCCUPÉES

Le problème est de déterminer les positions occupées dans le topomodèle¹³ d'un composé non identifié dont on connaît d'une part l'appartenance à la population

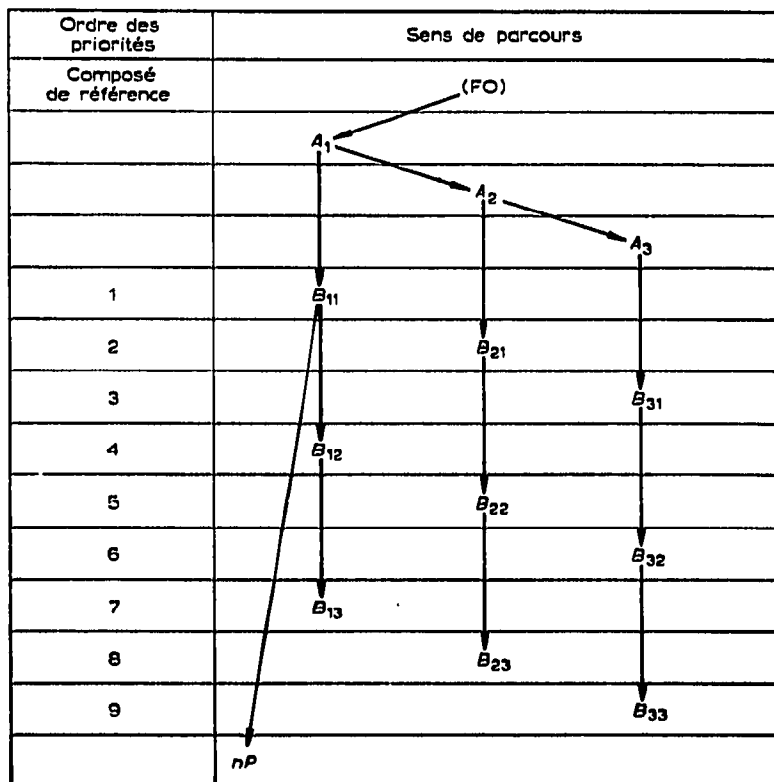


Fig. 1. Interdépendance des positions dans le topomodèle.

considérée, d'autre part le temps de rétention relatif au composé le plus simple de la famille.

La résolution mathématique de ce problème est liée au respect des priorités entre les positions (Fig. 1). La solution consiste à envisager successivement trois branches de développement I_1 , définies selon la Fig. 2. Chacune de ces branches doit être parcourue dans le sens dicté par la règle de priorité des positions et matérialisé par des flèches sur cette même figure.

La valeur du logarithme du temps de rétention relatif d'un composé RX inconnu déterminée expérimentalement avec une erreur de 0.006 unité logarithmique environ sera appelée YD . (Il est possible de faire varier l'erreur expérimentale admise.)

En parcourant la branche $I_1 = 1$, par exemple, on cherchera un topomodèle correspondant à un composé RX dont le logarithme du temps de rétention relatif théorique Y , calculé au moyen de la relation rétention-structure générale, soit tel que :

$$|YD - Y| \leq 0.006 \quad (1)$$

Les trois branches seront parcourues successivement de façon à trouver toutes les solutions en tenant compte de la possibilité d'existence, à partir de la position

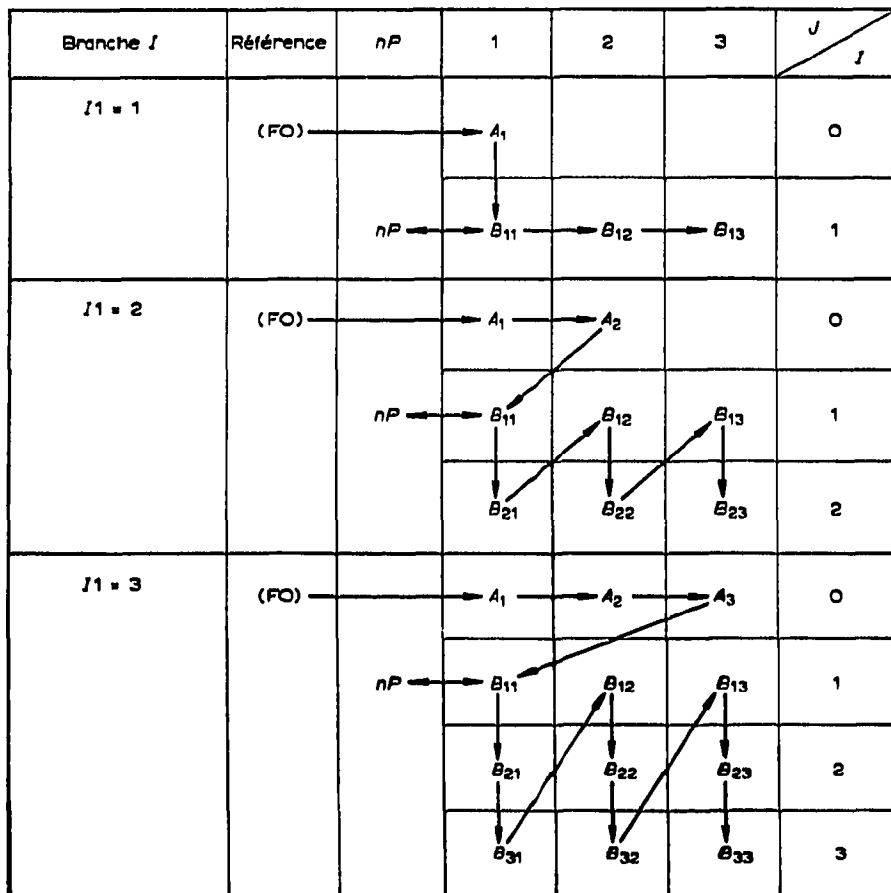


Fig. 2. Principe de la détermination des positions occupées dans le topomodèle.

B_{11} , de n positions p de la chaîne linéaire extérieure à l'environnement E_B^1 (bibl. 13). De cette façon, il est possible de déterminer tous les topomodèles solutions, c'est-à-dire toutes les structures des composés dont le logarithme du temps de rétention relatif Y obéit à la relation 1, à la température envisagée.

ORGANIGRAMME DU PROGRAMME

Le programme permettant l'identification de la chaîne carbonée d'un composé aliphatique saturé monofonctionnel est écrit en Algol 60 pour un ordinateur Bull-Honeywell M40. L'organigramme du programme comprend trois parties: l'organigramme principal; l'organigramme de la procédure "Branche"; l'organigramme de la procédure "P".

L'organigramme principal (Fig. 3)

Après avoir indiqué le nom de la famille chimique concernée et celui de la phase stationnaire utilisée, on introduit en mémoire toute une série de données. Tout

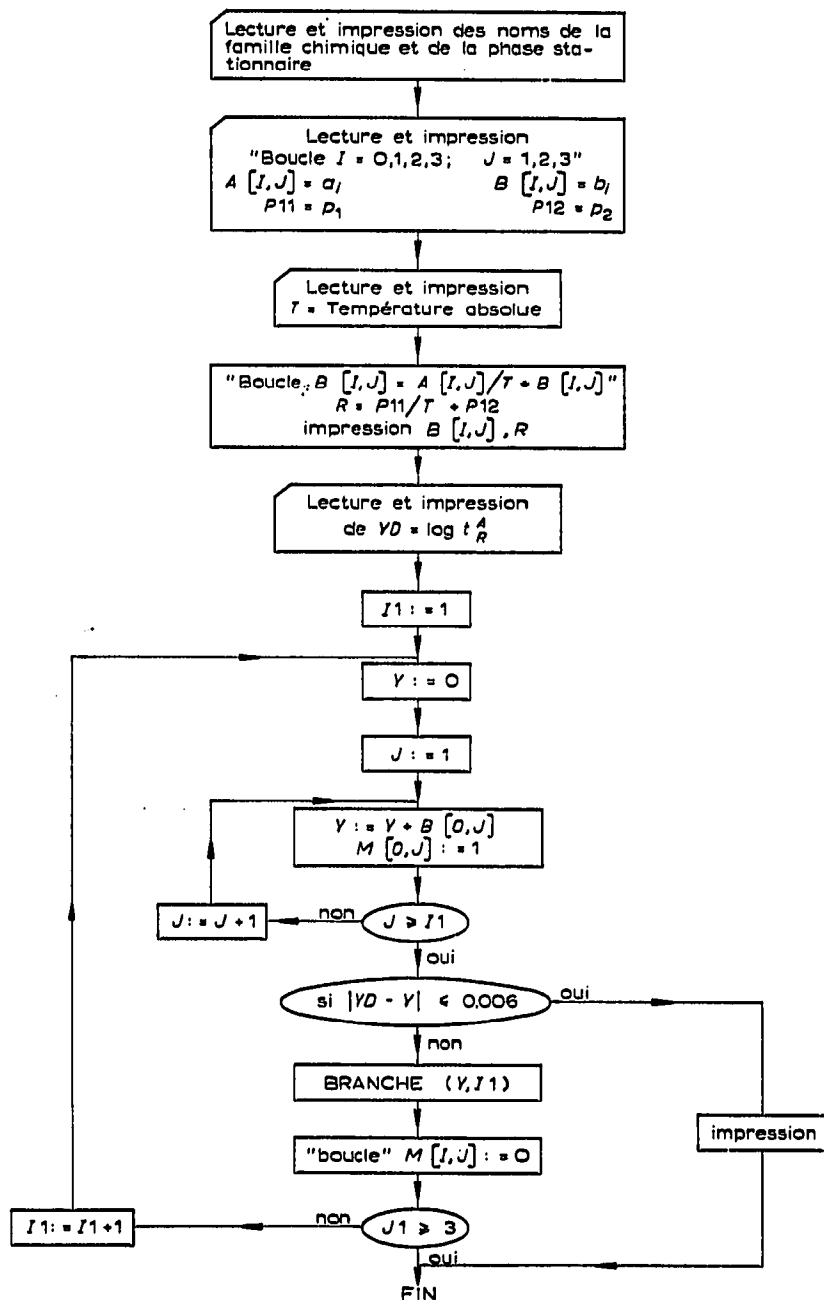


Fig. 3. Organigramme principal du programme.

d'abord, comme chaque paramètre structural I_i est fonction de la température absolue T selon :

$$I_i = a_i/T + b_i \quad (2)$$

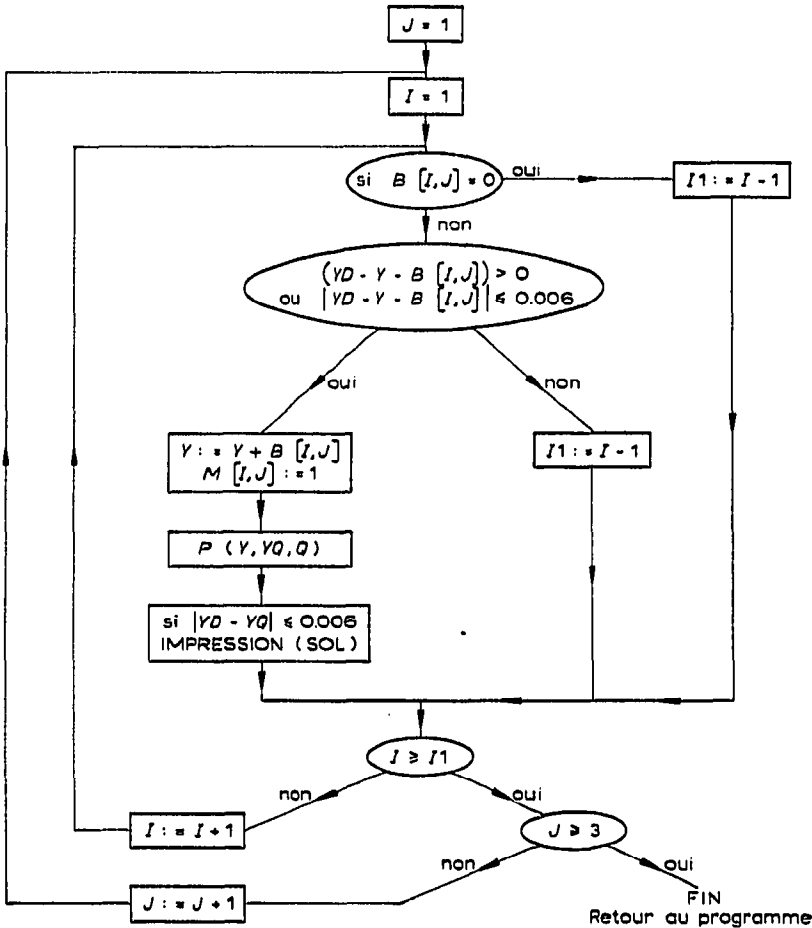


Fig. 4. Organigramme de la procédure Branche (Y, I_1).

les valeurs des coefficients thermiques a_i et b_i sont inscrites dans les tableaux A et B . En outre, le paramètre structural p jouant un rôle particulier, on le dissocie des paramètres I_i et les coefficients $P11$ et $P12$ tels que $p = P11/T + P12$ sont introduits en mémoire.

Puis on affiche la température absolue T à laquelle a été faite la mesure YD . Il est alors possible de procéder au calcul, à cette température, de la valeur de chacun des paramètres structuraux I_i et p notés respectivement $B [I,J]$ (dans le but de réserver moins de place en mémoire) et R .

Enfin on introduit la valeur YD du logarithme du temps de rétention relatif du composé à identifier.

La séparation des branches explicitées plus haut se fait au moyen d'une boucle sur I_1 , à chaque valeur de I_1 on repart à zéro soit $Y = 0$.

On ajoute alors, suivant la valeur de I_1 , les paramètres structuraux I_{A_i} , notés $B [O,J]$, J variant de 1 à I_1 et simultanément les positions solutions sont notées 1 dans le tableau $M [O,J]$. La solution est testée selon le critère $|YD - Y| \leq 0.006$; si on a une solution correcte on imprime et on va à la fin; si non on fait intervenir la procé-

sure "Branche ($Y, I1$)", qui introduit les paramètres structuraux $I_{B_{IJ}}$ notés $B [I,J]$. À la fin de cette procédure, la branche $I1$ est totalement parcourue et les résultats sont imprimés. On réinitialise le tableau M puis on passe à la branche suivante $I1 + 1$ si $I1 < 3$.

Organigramme de la procédure "Branche" (Fig. 4)

Sur la Fig. 2, on remarque qu'on parcourt à partir de B_{11} un nombre de lignes égal à $I1$, donc I doit varier de 1 à $I1$ et la règle des priorités nous impose la boucle sur I , intérieure.

Le test $(YD - Y - B [I,J]) > 0$ ou $|YD - Y - B [I,J]| \leq 0.006$ signifie que l'on regarde si la valeur obtenue en ajoutant $B [I,J]$ à Y dépasse YD ou encore se trouve dans la zone d'erreur; si oui, on ajoute à Y cette valeur et on inscrit 1 dans le tableau M à l'emplacement correspondant. Ensuite on regarde s'il est possible d'ajouter un multiple de p selon la procédure "P" et on teste la solution. Si l'élément $B [I,J]$ n'intervient pas, on s'oriente définitivement vers la branche $I1 = I - 1$ (Fig. 2).

Organigramme de la procédure "P" (Fig. 5)

La procédure $P(Y, YQ, Q)$ consiste essentiellement en la recherche du multiple Q , tel que:

$$(YD - Y - Q * R) \leq 0.006$$

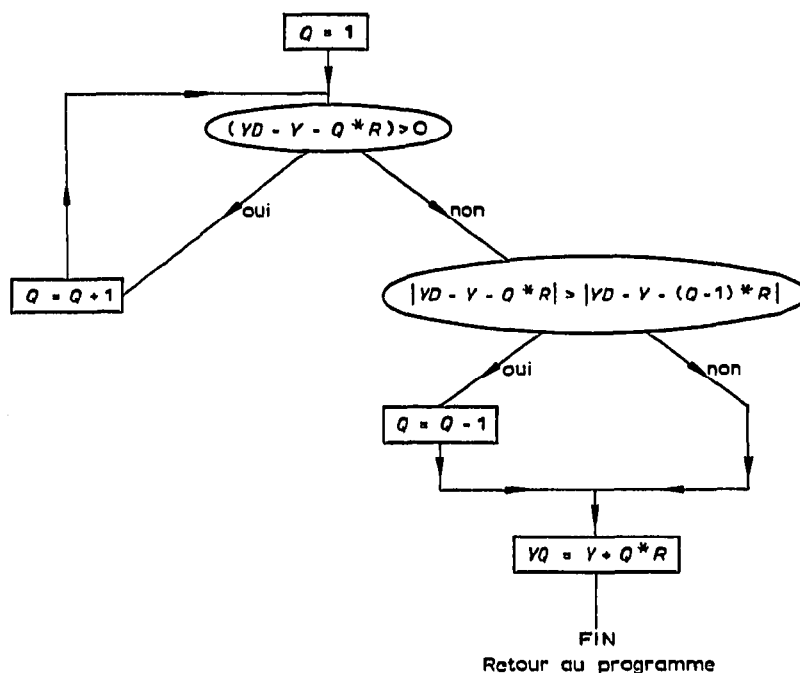


Fig. 5. Organigramme de la procédure $P(Y, YQ, Q)$.

TEST DU PROGRAMME

Ce programme est utilisé pour déterminer la structure de la chaîne carbonée d'un composé non identifié appartenant à la famille des méthyl-cétones saturées.

Précédemment, une relation rétention-structure a été proposée pour un ensemble de quarante cétones aliphatiques saturées^{17,18}. Dans le cas des méthyl-cétones elle peut s'écrire:

$$Y = \log t_R^A = A_1 I_{A_1} + A_2 I_{A_2} + (A_3 + B_{12}) I_{A_3} + B_{11} I_{B_{11}} + B_{13} I_{B_{13}} + B_{21} I_{B_{21}} + B_{31} I_{B_{31}} + p I_p$$

Chacun des paramètres I_i de structure dépend de la température selon la relation 2; les valeurs des coefficients thermiques a_i et b_i sont portées dans le Tableau I.

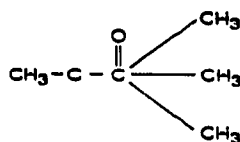
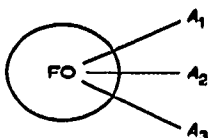
TABLEAU I

VARIATION DES PARAMÈTRES I_i DE SUBSTITUTION DES MÉTHYL-CÉTONES EN FONCTION DE LA TEMPÉRATURE¹⁸

No.	Paramètres I_i	a_i	b_i	Critère d'Exner ²⁵
1	I_{A_1}	156	-0.134	0.009
2	I_{A_2}	108	-0.100	0.029
3	$I_{A_3} = I_{B_{12}}$	73	-0.049	0.031
4	$I_{B_{11}}$	172	-0.194	0.010
5	$I_{B_{13}}$	62	-0.045	0.062
6	$I_{B_{21}}$	131	-0.112	0.030
7	$I_{B_{31}}$	235	-0.269	0.028
8	I_p	188	-0.221	0.003

Supposons que l'on veuille identifier, par exemple, la méthyl-cétone dont le logarithme du temps de rétention relatif à l'acétone, à 160°, sur SE-30, est $\log t_R^A = 0.499$.

L'identification menée selon le schéma précédemment exposé conduit à une seule solution pour la matrice de la Fig. 2, c'est-à-dire à un seul topomodèle dans lequel les positions occupées sont A_1 , A_2 et A_3 . On en déduit donc que la méthyl-cétone testée est la diméthyl-3,3 butanone-2



À cette solution topologique correspond une valeur calculée du logarithme du temps de rétention relatif (Tableau II) qui est $\log t_R^A = Y = 0.495$.

D'autres exemples d'identification de structure sont mentionnés dans le Ta-

TABLEAU II

IDENTIFICATION DE LA STRUCTURE DE LA CHAÎNE CARBONÉE DE MÉTHYL-CÉTONES ALIPHATIQUES SATURÉES

Information sur le composé <i>R-CO-CH₃</i>		Solutions proposées		
<i>log t_R^A mesuré</i>	<i>T (°K)</i>	<i>Positions occupées dans le topomodèle</i>	<i>Nom du composé R-CO-CH₃</i>	<i>log t_R^A calculé</i>
0.499	433	<i>A₁, A₂, A₃</i>	Diméthyl-3,3 butanone-2	0.495
0.647	433	<i>A₁, B₁₁, p = 1</i>	Hexanone-2	0.643
0.535	413	<i>A₁, B₁₁, B₁₂, B₁₃</i>	Diméthyl-4,4 pentanone-2	0.647
0.699	413	<i>A₁, A₂, A₃</i>	Diméthyl-3,3 butanone-2	0.533
		<i>A₁, B₁₁, p = 1</i>	Hexanone-2	0.700
		<i>A₁, B₁₁, B₁₂, B₁₃</i>	Diméthyl-4,4 pentanone-2	0.699

bleau II. On remarque que plusieurs structures peuvent être proposées pour une même valeur du temps de rétention relatif, ce qui est d'ailleurs en accord avec l'expérience.

Il est possible d'identifier les seize méthyl-cétones de la population expérimentale à l'aide d'une relation de topo-information établie à partir de la mesure du temps de rétention de dix méthyl-cétones bien choisies. L'utilisation de cette relation permet d'étendre l'identification à un nombre bien plus élevé de composés dont l'environnement présente la même empreinte à l'intérieur de l' E_B^1 , cet environnement pouvant en outre se développer linéairement en dehors de l' E_B^1 .

À fin de nous assurer de la validité du programme dans sa totalité, nous l'avons testé pour l'ensemble des données relatives à la population concernée, avec des résultats satisfaisants. L'extension à d'autres familles chimiques ne pose aucun problème particulier si l'on possède les données nécessaires.

CONCLUSION

Un programme d'identification écrit en Algol 60 a été mis au point dans le cas relativement simple des composés à une direction de développement RX . Cette méthode présente le grand avantage du traitement direct sur ordinateur sans avoir à consulter un catalogue. En outre, le temps de calcul est extrêmement faible; ce qui se traduit par un prix de revient particulièrement réduit.

L'extension de ce traitement à des familles chimiques plus complexes est en cours. En particulier, l'identification de composés aliphatiques saturés présentant deux directions de développement fera l'objet d'une prochaine publication²⁴.

RÉSUMÉ

Des relations de topologie-information entre la structure de composés chimiques et leur temps de rétention sont utilisées pour identifier la chaîne carbonée R de composés aliphatiques saturés monofonctionnels de formule générale RX . À partir de résultats expérimentaux en nombre limité, un programme écrit en Algol 60 permet de proposer une ou plusieurs structures pour un composé dont on connaît la famille chimique et le temps de rétention. L'organigramme de ce programme est présenté.

BIBLIOGRAPHIE

- 1 G. Schomburg, *Z. Anal. Chem.*, 200 (1963) 360; *J. Chromatogr.*, 14 (1964) 157; *J. Chromatogr.*, 23 (1966) 1, 18; *Anal. Chim. Acta*, 38 (1967) 45; *Separ. Sci.*, 1 (1966) 339.
- 2 P. G. Robinson et A. L. Odell, *J. Chromatogr.*, 57 (1971) 1.
- 3 J. Takács, C. Szita et G. Tarján, *J. Chromatogr.*, 56 (1971) 1.
- 4 J. Takács, Z. Talas, I. Bernath, G. Czako et A. Fisher, *J. Chromatogr.*, 67 (1972) 203.
- 5 G. Castello et G. D'Amato, *J. Chromatogr.*, 58 (1971) 127.
- 6 L. E. Cook et F. M. Raushel, *J. Chromatogr.*, 65 (1972) 556.
- 7 T. Feher et L. Bodrogi, *J. Chromatogr.*, 71 (1972) 17.
- 8 T. L. Kwa, O. Korver et C. Boelhouwer, *J. Chromatogr.*, 30 (1967) 17.
- 9 A. J. Solo et S. W. Pelletier, *Anal. Chem.*, 35 (1963) 1584.
- 10 J. Jonas, J. Janak et M. Kratochvil, *J. Gas Chromatogr.*, 4 (1966) 332.
- 11 W. Ebing, *Chromatographia*, 2 (1969) 442.
- 12 J. E. Dubois, D. Laurent et H. Viellard, *C.R. Acad. Sci., Sér. C*, 264 (1967) 1019.
- 13 J. E. Dubois et H. Viellard, *Bull. Soc. Chim. Fr.*, (1968) 900, 905 et 913.
- 14 J. E. Dubois et D. Laurent, *C.R. Acad. Sci., Sér. A*, 268 (1969) 405.
- 15 J. E. Dubois et D. Laurent, *Bull. Soc. Chim. Fr.*, (1969) 2449.
- 16 M. Chastrette, G. Lenfant et J. E. Dubois, *C.R. Acad. Sci., Sér. C*, 265 (1967) 602.
- 17 G. Lenfant, M. Chastrette et J. E. Dubois, *J. Chromatogr. Sci.*, 9 (1971) 220.
- 18 M. Chastrette et G. Lenfant, *J. Chromatogr.*, 77 (1973) 255.
- 19 M. Chastrette et G. Lenfant, *C.R. Acad. Sci., Sér. C*, 271 (1970) 79.
- 20 M. Chastrette et G. Lenfant, *J. Chromatogr.*, 68 (1972) 19.
- 21 M. Chastrette, P. Couillault, G. Casteignau et M. Mazet, *Bull. Soc. Chim. Fr.*, (1973) 605.
- 22 M. Chastrette et G. Lenfant, à paraître.
- 23 J. Cuvelier, *Thèse de Doctorat de Spécialité, Nice, 1972.*
- 24 M. Chastrette, G. Lenfant et A. Remy, à paraître.
- 25 O. Exner, *Collect. Czech. Chem. Commun.*, 131 (1966) 3222.