

CHIMIE ORGANIQUE BIOLOGIQUE. — *Étude au moyen d'une analyse des correspondances multiples des propriétés insecticides et enzymatiques de N-méthyl carbamates de phényle.* Note (*) de **Jean-Christophe Doré, Jean-François Miquel, Georges Mrlina et Jean-Pierre Calmon**, présentée par Fernand Gallais.

Cette Note concerne l'étude des relations structure/activité de 40 molécules de type N-méthyl carbamate de phényle. Les propriétés insecticides d'une part et l'activité enzymatique vis-à-vis de l'acétylcholine estérase d'autre part ont été étudiées au moyen d'une analyse factorielle des correspondances multiples. Ce traitement mathématique a permis de préciser le mécanisme d'action biochimique de ces molécules et de définir les types de substitutions chimiques favorables à l'activité.

BIOORGANIC CHEMISTRY. — Study of the Insecticidal and Anticholinesterase Properties of Phenyl N-Methyl Carbamates by Correspondence Analysis.

This paper deals with the structure/activity relationships of forty phenyl N-methyl carbamates. The pesticide activity and the anticholinesterase activity were studied by means of correspondence analysis. This mathematical processing allowed the biochemical mode of action of these active molecules to be assessed and the most suitable chemical substitutions to be forecast.

Nous avons montré dans de précédents articles ([1], [2]) l'intérêt que présente l'utilisation des méthodes factorielles en général et de l'analyse des correspondances [3] en particulier, dans l'étude de problèmes complexes de structure/activité.

En effet, dans la mesure où une série de molécules montre une action, non plus sur un, mais sur plusieurs tests pharmacologiques [1] ou biochimiques [2], l'abord du problème devient délicat en raison de sa nature multidimensionnelle.

Les traitements factoriels apportent une solution intéressante dans la résolution de ces cas, dans la mesure où ils opèrent une restructuration des informations autour d'axes factoriels regroupant un pourcentage maximal de la variance du système. Ces axes factoriels, constituant en quelque sorte des super caractères théoriques, permettent une projection des données sous forme d'une représentation graphique plane réduisant ainsi, avec une perte minimale d'information, un système à n dimensions en un système à deux dimensions.

L'analyse factorielle des correspondances (AFC) par l'utilisation de la métrique du χ^2 permet en outre, contrairement à l'analyse en composantes principales (ACP) une présentation simultanée des deux champs (molécules et propriétés), ce qui facilite, moyennant certaines précautions, leur confrontation directe (phénomène barycentrique, représentativité des éléments des différents espaces, contributions relatives et absolues, conformité des profils marginaux...).

Rappelons que la « lecture » des cartes de proximité distributionnelle résultant du traitement factoriel s'effectue schématiquement, dans le cas qui nous préoccupe, de la manière suivante :

Deux ou plusieurs molécules, ou/et, deux ou plusieurs propriétés sont d'autant plus semblables qu'elles se trouvent proches topographiquement (et éloignées de l'origine des axes) et s'opposent d'autant plus que leurs coordonnées sont éloignées.

La présente Note se propose de montrer comment on peut analyser, simultanément, au moyen d'une disjonction des données en classes, d'une part des activités biologiques exprimées en DL 50 et d'autre part, des constantes de vitesse de réactions enzymatiques exprimées en diverses unités cinétiques.

Nous avons utilisé pour ce faire une variante de l'A.F.C., l'analyse factorielle des correspondances multiples (A.C.M.) [4]. En effet, cette méthodologie permet la prise en compte simultanée de données hétérogènes.

Les données expérimentales concernent l'étude de la toxicité *in vivo* et *in vitro* de 40 molécules insecticides, de type N-méthyl carbamate de phényle monosubstitué en *ortho*, *méta* ou *para* sur le noyau benzénique.

Les activités biologiques sont évaluées par les paramètres suivants :

- la mesure de la DL 50 sur mouche, déterminée en ug/g par Metcalf et coll. [5];
- la mesure de la DL 50 sur mouche ($\mu\text{g/g}$) avec adjonction d'un synergiste, le butoxyde de pipéronyle [5];
- l'étude de l'inhibition 50 sur broyat de cerveaux de mouche, exprimée en concentration molaire [5];
- la mesure de la constante de vitesse K_d relative à la cinétique d'inhibition réalisée à partir d'acétylcholine estérase d'érythrocyte bovin, E.C.3.1.1.7, déterminée par Nishioka et coll. [6].

Le traitement factoriel a été réalisé au moyen d'un microordinateur à partir du tableau logique constitué de 40 lignes (I), les 40 molécules et de 20 colonnes (J), les quatre propriétés précédemment décrites, éclatées chacune en cinq classes d'importance comparable. Le contenu de chaque élément $K_{i,j}$ de la matrice étant (0) ou (1) suivant la situation de la I-ième molécule pour la J-ième propriété.

Le choix d'un éclatement de la population totale en cinq classes a été dicté par le souci de disposer d'une échelle de niveaux suffisante tout en ne réduisant pas trop l'effectif de chacun d'entre eux.

Le résultat des calculs vectoriels permet de constater que les trois premières valeurs propres de la matrice carrée symétrique intermédiaire regroupent un peu moins de 50% de la variance du système, ce qui rend l'interprétation significative, tout en permettant l'élimination d'une part importante du « bruit de fond » contenu dans les données expérimentales.

Le premier axe factoriel fournit une bonne représentation du phénomène étudié dans la mesure où la projection des cinq niveaux d'activité pour les deux tests toxicologiques comme pour les deux tests enzymatiques se distribue dans l'ordre du 1^{er} niveau (bonne activité) au 5^e niveau (très faible activité).

La représentation plane constituée par la projection des deux champs sur les deux premiers facteurs qui rassemblent environ un tiers de l'inertie totale, apporte à différents niveaux des informations sur la structure du problème étudié :

- (1) en terme de comparaison des différents tests les uns par rapport aux autres;
- (2) en terme de typologie des différents produits actifs étudiés (P.A.);
- (3) en terme d'étude des relations entre structure chimique et activité biologique.

EXPLICATION DE LA PLANCHE

Fig. 1. – Activité biologique de N-méthyl carbamate de phényle.

Fig. 1. – *Biological activity of phenyl N-methylcarbamates.*

Fig. 2. – Les différents halogènes en (*ortho*, *meta*, *para*).

Fig. 2. – *Effect of the position of the halogenated substituents.*

Fig. 3. – Effet de chaîne et encombrement *ortho-meta*.

Fig. 3. – *Chain effect and ortho-meta steric hindrance.*

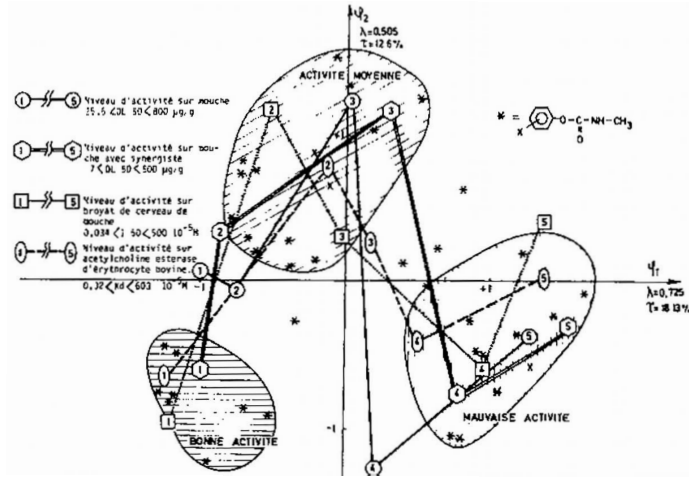


Fig. 1

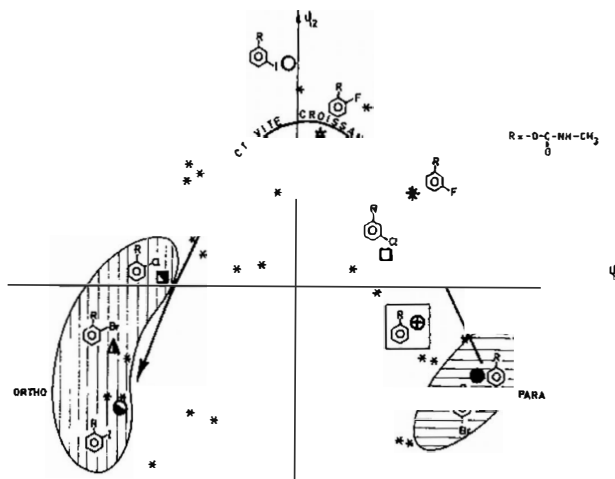


Fig. 2

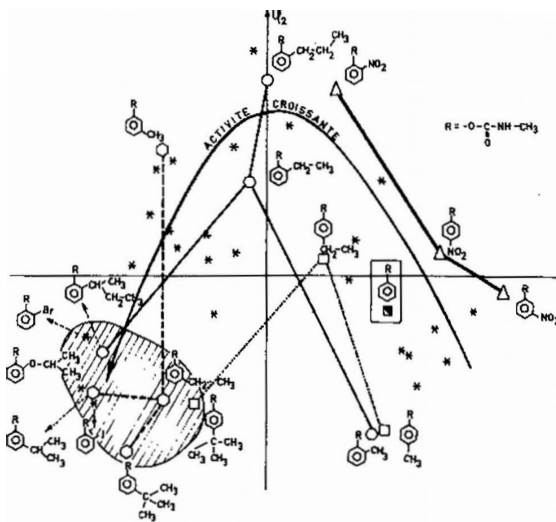


Fig. 3

Au niveau de la représentation des différents essais biologiques, le graphe (fig. n° 1) confirme le bien fondé du mécanisme d'action proposé (l'inhibition de l'acétylcholine estérase) [7]. En effet, l'évolution des niveaux de réponse, qu'il s'agisse d'action biologique ou enzymatique suit globalement la même trajectoire. La réponse des différentes molécules sur mouche entière, sur broyat de cerveaux ou sur AcChE pure obéit donc aux mêmes « lois ».

L'effet du synergiste (butoxyde de pipéronyle) est, quant à lui, bien mis en évidence dans la mesure où il empêche l'inactivation prématurée des molécules les plus actives dans le cas d'une expérimentation sur l'insecte entier (cf. niveau n° 1 de la DL 50 sur mouche avec et sans synergiste). Il agirait soit comme un autre substrat, soit comme un inhibiteur au niveau des différentes oxydases [8].

En ce qui concerne la typologie obtenue pour la population de molécules étudiées au travers de la réponse biologique analysée, on peut constater :

— un partitionnement sommaire des P.A. le long du premier axe factoriel. Les molécules les plus actives ayant des coordonnées fortement négatives par rapport à cet axe et les moins actives des coordonnées fortement positives ;

— en regardant simultanément les deux champs, on peut définir trois rassemblements caractéristiques de molécules (groupes 1, 2 et 3) qui correspondent (grosso modo) au niveau 1, (2 et 3), (4 et 5) des quatre tests pris en compte.

Mais, c'est au niveau de l'étude détaillée des relations entre structure chimique et activité biologique que l'AFC se révèle un bon instrument d'aide à l'interprétation de résultats expérimentaux complexes.

En effet, la représentation graphique des différentes molécules dérivées de l'archétype de base sous forme d'une carte de proximité distributionnelle, permet une simulation directe par confrontation immédiate des différentes hypothèses susceptibles d'expliquer l'activité biologique.

Pour illustrer ces possibilités, nous avons mis en évidence, sur la figure 2, les différents composés halogénés étudiés (F, Cl, Br, I) dans leurs différentes positions (*ortho-méta-para*). Cette opération permet la visualisation directe de l'effet comparé des différents types d'encombrement stérique (volume atomique et positions) sur l'activité biologique globale (correspondant à une sorte de moyenne pondérée de la réponse sur les quatre tests). Ce résultat rejoint les observations faites par Metcalf [7].

Dans le même ordre d'idée, on peut procéder à la superposition de paramètres caractéristiques des propriétés moléculaires (physicochimiques par exemple), avec les 40 points représentatifs des 40 dérivés, l'utilisation d'un paramètre électronique exprimé suivant trois niveaux par exemple permet d'examiner si la propriété biologique est sous la dépendance direct. de ce facteur, en fonction de la plus ou moins grande dispersion des trois classes sur le graphe.

Cependant, c'est par la confrontation directe des caractéristiques des molécules ayant une proximité ou une opposition « topographique » que cette approche permet une saisie précise du pourquoi d'un effet biologique complexe. Ainsi, nous avons fait figurer sur le graphe n° 3 un certain nombre de molécules significatives pour lesquelles on peut :

1. Suivre l'évolution de l'activité pour différentes séries homologues :
 - du 2-méthyle ou 2-sec-butyle en passant par le 2-éthyle et le 2 *n*-propyle;
 - ou bien du 3-méthyle au 3-isopropyle ou 3-tertiobutyle en passant par le 3-éthyle.

2. Comparer pour une zone d'activité biologique donnée (la réponse insecticide maximale par exemple) les dénominateurs communs structuraux qui lient les différentes molécules à cette zone.

Dans le cas présent, on peut constater la très bonne corrélation entre le maximum d'activité et le fort encombrement *ortho* ou *méta* avec des substituants plutôt électro-donneurs. Les groupements électro-attracteurs type NO₂, bien que présentant un caractère carbamoylant élevé, s'avèrent en effet moins bien placés. Pour les N-méthylcarbamates insecticides, le maximum d'activité reste un compromis entre la faculté de fixation sur le site anionique récepteur de l'AcChE et la capacité de carbamoyer l'hydroxyle libre de la sérine appartenant au site actif de cette enzyme [9].

Sans entrer dans les détails, signalons que ce résultat est conforme aux observations faites sur la structure du site d'action biochimique [7].

Précisons, pour conclure, que cette courte Note n'a pas la prétention de développer en profondeur le sujet de recherche qui lui a servi de support, mais de montrer comment les traitements factoriels dont la mise en œuvre est dorénavant relativement simple en raison de la banalisation des techniques informatiques peut apporter une aide efficace dans la résolution de problèmes de structure/activité faisant appel à une pluralité de tests.

(*) Remise le 4 octobre 1982, acceptée après révision le 20 juin 1983.

[1] J. C. DORÉ et J. F. MIQUEL, *Comptes rendus*, 293, Série II, 1981, p. 1061.

[2] J. C. DORÉ, J. GILBERT, A. CRASTES DE PAULET, F. MICHEL et J. F. MIQUEL, *Comptes rendus*, 294, série III, 1982, p. 731.

[3] J. P. BENZECRI et coll., *L'analyse des données (II), l'analyse des correspondances*, Dunod, Paris, 1980.

[4] L. LEBART, J. P. FENELON, *Statistique et informatique appliquées*, Dunod, Paris, 1975.

[5] R. L. METCALF, *Bull. O.M.S.*, 44, 1971, p. 43.

[6] T. NISHIOKA, T. FUJITA, K. KAMOSHITA et M. NAKAJIMA, *Pest. Bioch. Phys.*, 7, 1977, p. 107.

[7] R. L. METCALF et T. R. FUKUTO, *J. Agr. Food chem.*, 13, 1965, p. 220.

[8] J. E. CASIDA, *J. Agr. Food chem.*, 18, 1970, p. 753.

[9] G. MRLINA et J. P. CALMON, *J. Agr. Food chem.*, 28, 1980, p. 675.

J.-C. D. et J.-F. M. : *Équipe de Recherche sur l'Innovation Pharmaceutique, C.E.R.C.O.A.-C.N.R.S.,
2, rue Henry-Dunant, 94320 Thiais;*

G. M. et J.-P. C. : *Laboratoire de Chimie organique Biologique et Physicochimie du Sol,
E.N.S.A.P.-I.N.P., 145, avenue de Muret, 31076 Toulouse Cedex.*