

SYNTHESE ASSISTEE PAR ORDINATEUR DE LA PHOSPHACARNEGINE-I

ETABLISSEMENT DU PLAN DE SYNTHESE AVEC L'AIDE DE PASCOP

C. LAURENCO, L. VILLIEN et G. KAUFMANN

Laboratoire de Modèles Informatiques Appliqués à la Synthèse, E.R.A. 671 du CNRS, Université Louis Pasteur, 4, rue Blaise Pascal, 67008 Strasbourg Cedex, France

(Received in France 5 May 1983)

Résumé—Par analyse retrosynthétique de la structure cible 1, l'équipe chimiste-PASCOP a développé un plan de synthèse pour cette nouvelle phosphine.

Abstract—Using a retrosynthetic analysis of the target structure 1, the chemist-PASCOP team has developed a synthetic plan for this new phosphine.

Notre laboratoire a pour vocation de créer des outils informatiques destinés aux chimistes organiciens. Au cours de ces dernières années il a réalisé différents systèmes dont PASCOP,¹ un programme de synthèse assistée par ordinateur. À notre connaissance l'expérimentation de chemins de synthèse multi-tapes et inédits proposés par de tels programmes n'a fait l'objet d'aucune publication². Il nous a paru intéressant d'entreprendre la synthèse de molécules organophosphorées originales, avec l'aide de PASCOP, pour montrer les différents modes d'utilisation et les limites de ce système et finalement essayer de dégager quelques règles générales devant permettre d'exploiter au mieux la technique de la synthèse assistée par ordinateur.

Nous avons choisi d'étudier la synthèse de la phosphine cyclique 1 qui appartient à une famille de composés de synthèse possédant le squelette tétrahydroisophosphinoline^{3,4} et dont l'intérêt est l'analogie de structure avec des composés azotés souvent biologiquement actifs. Ainsi 1 est l'isostère phosphoré de la carnégine, un alcaloïde isolé sous forme racémique d'un cactus américain.⁵ Pour simplifier et en prenant exemple de l'appellation donnée aux analogues phosphorés connus de quelques composés azotés comme l'atropine,⁶ la lilolidine ou la julolidine,⁷ nous nommerons 1 la phosphacarnégine plutôt que par son nom systématique: tétrahydro-1,2,3,4 diméthoxy-6,7 diméthyl-1,2 isophosphinoline.

Cet article décrit la première partie de notre étude de la synthèse de la phosphacarnégine, c'est-à-dire

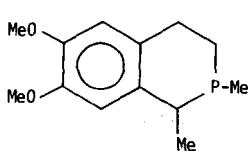
l'établissement du plan de synthèse avec l'aide de PASCOP.

Afin de faciliter la compréhension de nos résultats, il est nécessaire de rappeler succinctement les grandes lignes du fonctionnement de PASCOP.

Comme la plupart des autres systèmes de synthèse assistée par ordinateur connus,⁸ PASCOP exploite le principe de l'analyse retrosynthétique et, pour une molécule cible soumise par le chimiste, construit un graphe appelé arbre de synthèse qui réunit les précurseurs potentiels de cette cible. Dans l'arbre de synthèse le passage de la cible à un précurseur ou d'un précurseur à un autre de rang plus élevé s'effectue par l'application d'une opération topologique appelée transformation. Celle-ci décrit en fait, dans un langage spécialisé, une ou plusieurs réactions chimiques prises dans le sens retrosynthétique. Environ 700 transformations sont contenues dans la banque de données de PASCOP.²

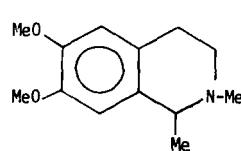
Pour soumettre la phosphacarnégine à PASCOP nous devons dessiner la formule semi-développée de celle-ci sur l'écran de visualisation du système. A partir de ce graphe PASCOP analyse les caractéristiques topologiques et chimiques de la molécule, il reconnaît notamment la nature et la taille des cycles, les groupes fonctionnels et la configuration des stéréocentres. Après cette perception le programme nous interroge sur le mode d'utilisation souhaité. Si nous n'avons pas d'idée préconçue sur la synthèse de notre molécule cible, nous pouvons demander une recherche systématique des précurseurs. Dans ce cas PASCOP explore le fichier des transformations sélectionne celles qui sont applicables à la molécule cible et construit les précurseurs correspondants. Ces précurseurs sont évalués pour que soient éliminées les structures redondantes ou chimiquement instables puis ceux qui restent sont visualisés sur l'écran sous la forme d'un premier niveau de l'arbre de synthèse. Ils peuvent ensuite être affichés individuellement ou par groupe à la demande de l'utilisateur.

Selon cette recherche systématique, il serait possible d'analyser chaque précurseur obtenu au premier niveau puis chacun de ceux obtenus au deuxième et ainsi de suite jusqu'à développer un arbre de synthèse comportant tous les chemins susceptibles d'être trou-



1

Phosphacarnégine



Carnégine

Fig. 1.

vés par le programme. Cette éventualité doit cependant être écartée car elle conduirait très rapidement à l'explosion combinatoire de l'arbre de synthèse par accumulation des précurseurs. Par exemple, si l'analyse d'une molécule cible comportant plusieurs groupes fonctionnels fournit un premier niveau de 40 précurseurs et que chacun de ceux-ci fournit à son tour 40 précurseurs et ainsi de suite, le programme aura développé après 5 niveaux un arbre comportant 40⁵ soit plus de 100 millions de structures.⁸ Il est donc demandé au chimiste de choisir à chaque niveau le ou les précurseurs à partir desquels il souhaite voir poursuivre le développement de l'arbre de synthèse. De plus, l'encombrement des niveaux ayant surtout pour origine l'application anarchique des transformations introduisant ou échangeant des groupes fonctionnels, PASCOP n'a directement accès qu'à des transformations dont l'application conduit à une simplification du problème pose, c'est-à-dire à une fragmentation du squelette de la cible, et à quelques transformations échangeant des groupes fonctionnels. La sélection et l'application de toutes les autres transformations s'effectue uniquement dans les autres modes d'utilisation de PASCOP, sous le contrôle de l'utilisateur ou selon des procédures spécialisées qui seront indiquées plus loin.

Par la recherche systématique des précurseurs de la phosphacarnégine, PASCOP a construit un premier niveau de l'arbre de synthèse comportant 18 précurseurs ou couples de précurseurs (Fig. 2). Sans faire une étude détaillée de ce premier niveau nous remarquons que la plupart des précurseurs sont des dérivés du phosphore trivalent. Pour limiter le nombre d'étapes nécessitant la manipulation de tels composés très oxydables nous avons choisi de ne retenir à ce niveau que le dérivé tétracoordonné 3 devant conduire à la phosphacarnégine par réduction du groupe phosphoryle avec HSiCl3, PhSiH3 etc.

Un deuxième mode d'utilisation de PASCOP fait plus particulièrement appel à l'imagination et à l'initiative du chimiste. Dans ce mode, dit de stratégie interactive, l'utilisateur a la possibilité d'exprimer ses idées, d'explorer des stratégies de synthèse et de confronter ses connaissances avec celles de la machine. Pour cela il dispose d'un module de stratégie par l'intermédiaire duquel il peut fixer au programme, à l'aide d'un langage spécialisé,^{2d} des objectifs à atteindre prioritairement tels que la cassure rétrosynthétique d'une liaison particulière, l'échange ou l'introduction d'un groupe fonctionnel, etc.

Ainsi, si nous pensons que la formation du cycle phosphoré est une étape clé de la synthèse de la phosphacarnégine et de son oxyde 3 nous demanderons à PASCOP la rupture rétrosynthétique de ce cycle à l'aide de la directive: BREAK RING 3, 2, 14, 13, 12, 11. La numérotation utilisée ici et dans les figures suit l'ordre d'entrée des atomes en machine et n'a donc pas de rapport avec celle de la nomenclature habituelle. En exécutant cette directive PASCOP développe à partir de 3 un deuxième niveau de l'arbre de synthèse comportant 17 précurseurs (Fig. 3). Nous constatons que dans chaque structure engendrée l'objectif a été effectivement atteint et que, à l'exception de la liaison aromatique 2-3, chaque liaison du cycle a été rompue par l'application d'une ou plusieurs transformations. La plupart des solutions proposées à ce niveau sont acceptables, ce-

pendant nous remarquons que dans les composés 29, 30 et 31 la partie carbonée est plus simple que dans les autres précurseurs où le noyau aromatique reste tétrasubstitué et ceci peut être un critère pour effectuer une première sélection des structures à partir desquelles le développement de l'arbre mérite d'être poursuivi. Un examen plus attentif nous permet d'écarter le précurseur 30 dont la cyclisation selon une réaction de Friedel et Crafts devrait conduire à un composé comportant un cycle à 7 chaînons plutôt qu'à la molécule 3. En effet, lorsqu'une oléfine est conjuguée avec un groupe électro-attracteur, l'attaque électrophile de celle-ci sur un noyau aromatique, en présence d'un catalyseur, a lieu par l'atome de carbone situé en β du groupe c'est-à-dire par celui sur lequel le carbocation est le mieux stabilisé. L'oxyde de phosphacarnégine 3 devrait donc être obtenu à partir des composés 29 et 31 par formation de la liaison 2-14. Si nous admettons que la formation de cette liaison est d'autant plus facilitée que l'atome 14 est plus électrophile alors nous pouvons tester une stratégie en deux étapes consistant à demander d'abord la fonctionnalisation de l'atome 14 de la cible 3 puis la rupture rétrosynthétique de la liaison 2-14. Pour cela nous soumettons à PASCOP le programme de stratégie suivant:

```
1) MAKE GROUP AT 14
NEXSTRAT
2) BREAK BOND 2-14
END.
```

L'exécution de ce programme par PASCOP engendre 12 précurseurs répartis sur deux niveaux (Fig. 4). Au premier de ces niveaux la fonctionnalisation de l'atome 14 est effectivement réalisée dans chaque structure alors que la cassure de la liaison 2-14 l'est au second niveau. Les deux séquences rétrosynthétiques ($3 \Rightarrow 41 \Rightarrow 47$) et ($3 \Rightarrow 43 \Rightarrow 49$), correspondant respectivement à une cyclodéhydratation de 47 suivie de l'hydrogénéation de l'oléfine 41 et à une acylation intramoléculaire de 49 suivie de la réduction de la cétone 43 nous paraissent les plus satisfaisantes.

Dans le troisième mode d'utilisation de PASCOP, mode dit de stratégie automatique, le développement de l'arbre de synthèse est confié à des sous-programmes, appelés automates, spécialisés dans l'application de transformations d'un grand intérêt synthétique ou dans la recherche de précurseurs pour des molécules comportant un motif structural précis. Lorsqu'il est activé par le chimiste, l'automate vérifie la présence dans la molécule cible de la sous-structure qu'il doit transformer et analyse son environnement, puis au vu des résultats de cette perception procède à toute modification de la cible (protection, déprotection, échange, introduction d'un groupe fonctionnel etc.) qu'il juge nécessaire pour atteindre l'objectif pour lequel il a été programmé. En appliquant des techniques d'intelligence artificielle et sans faire intervenir l'utilisateur, un automate est capable de fournir un arbre de synthèse émondé comportant plusieurs niveaux. Dans le domaine de la synthèse des composés organophosphores PASCOP dispose de l'automate Airbus qui a pour objectif la rupture rétrosynthétique de la liaison phosphore-carbone

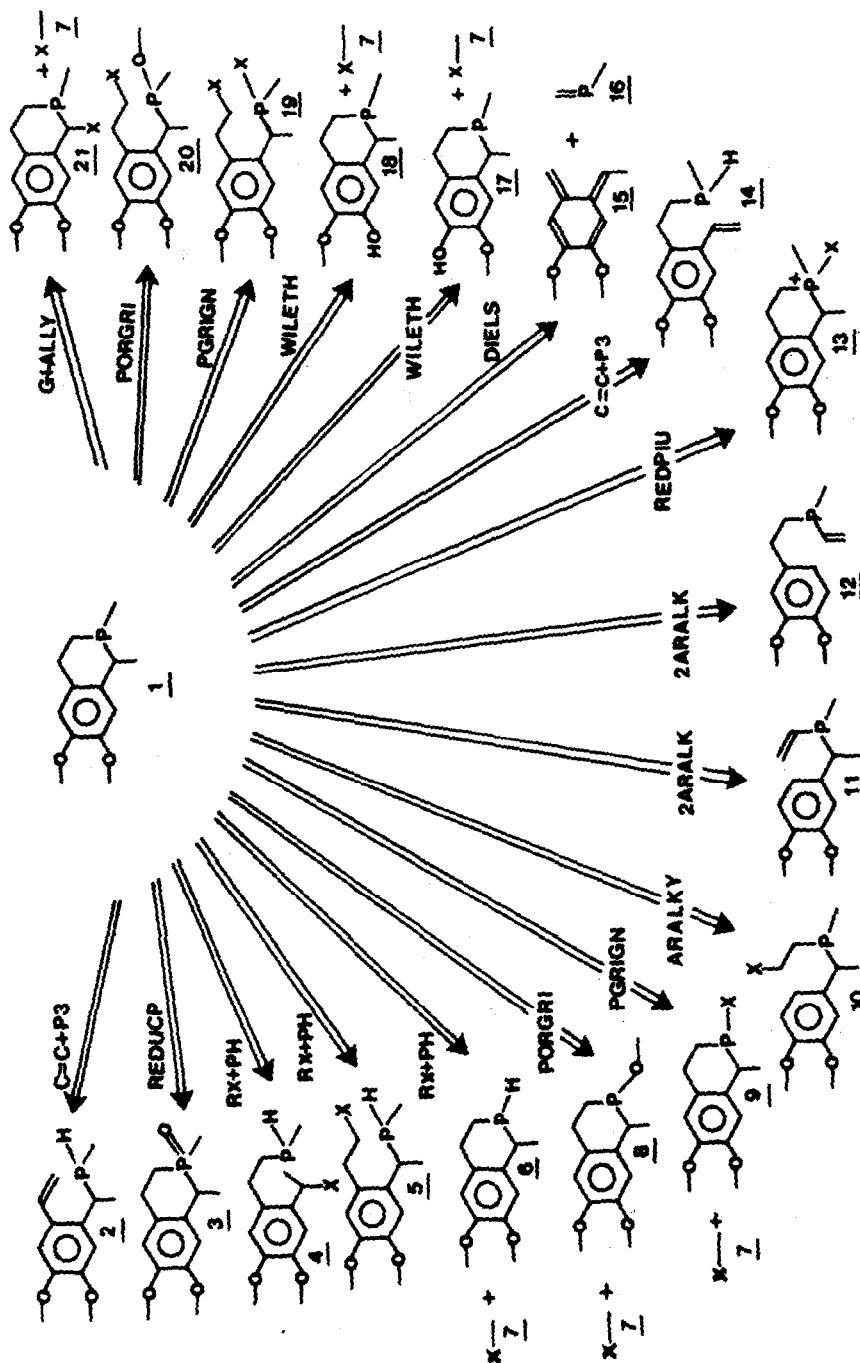


Fig. 2. Recherche systématique des précurseurs de la phosphacarnégine 1. Les transformations sont désignées par des noms mnémonomiques, la liste des réactions chimiques correspondantes est donnée en annexe.

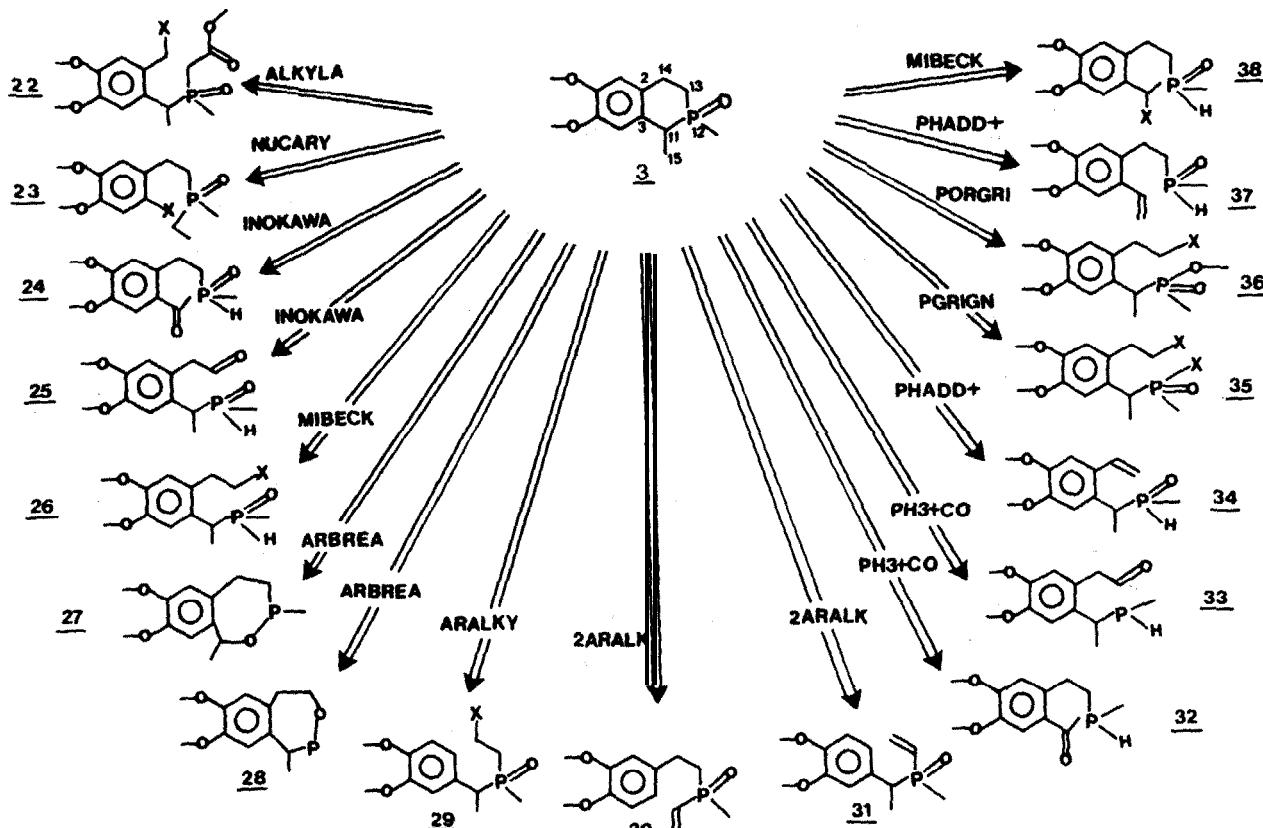


Fig. 3. Application de la directive: "BREAK RING 3, 2, 14, 13, 12, 11". Voir remarque Fig. 2.

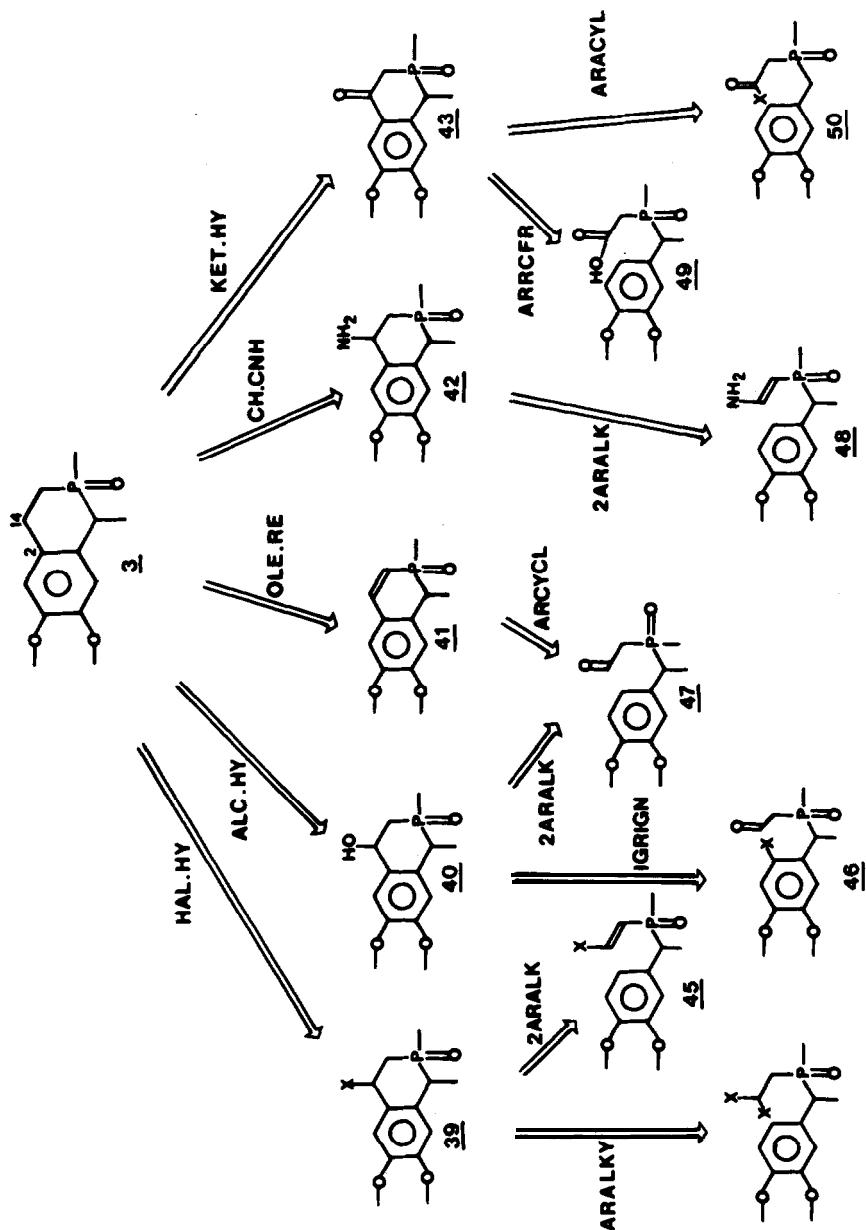


Fig. 4. Application des directives: "MAKE GROUP AT 14-BREAK BOND 2-14". Voir remarque Fig. 2.

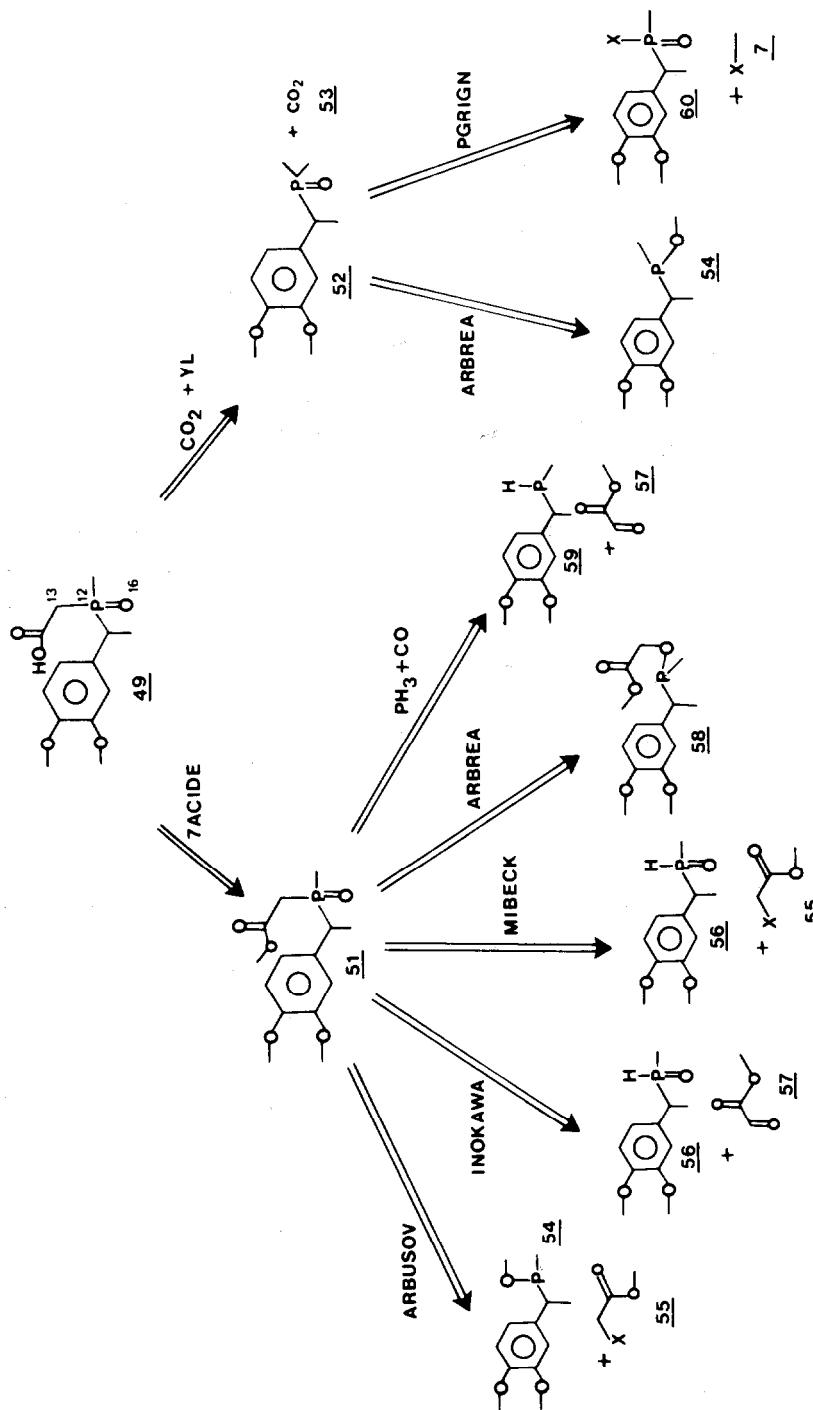


Fig. 5. Application de l'instruction: "PERFORM ARBUSOV AT 16, 12, 13". Voir remarque Fig. 2.

dans des **molécules-cible possédant** la sous-structure **P(O)-C.**^{2e} Arbusov est donc adapté à la recherche des **précurseurs** de molécules telles que les oxydes de phosphines tertiaires dissymétriques et fonctionnelles 29, 31, 47 et 49 engendrés précédemment (Fig. 3 et 4). Ces composés possèdent trois fois la **sous-structure P(O)-C** et une **stratégie simple consiste à** fixer comme **objectif au système la cassure** de la liaison reliant au phosphore la **partie la plus fonctionnaliste** de la molécule. Ainsi pour développer l'arbre de synthèse à partir du compost 49 nous activons l'automate par la directive: PERFORM ARBUSOV AT 16, 12, 13. Celui-ci considre que la présence de la fonction acide carboxylique l'**empêche d'atteindre son objectif** et pour surmonter cet obstacle il **échange ou élimine cette fonction** (Fig. 5). L'elimination fournit le précurseur 52 qui offre deux

possibilités de rupture de la liaison 12-13. Cette voie n'est cependant pas acceptable car la présence du proton **ArCH-P** rend peu probables la formation et la carbonatation **selectives d'un** carbanion sur l'un des deux méthyles liés au phosphore dans le compose 52. L'**échange** de la fonction acide de 49 conduit à former 51 dont la liaison 12-13 est rompue par cinq transformations. L'automate Arbusov peut encore être utilisé pour développer l'arbre de **synthèse** à partir de l'oxyde de phosphine secondaire 56 qui **comporte** deux fois la sous-structure **P(O)-C**. Par exécution de la directive PERFORM ARBUSOV AT 1 I, 12, 16 le **système** engendre deux niveaux supplémentaires renouissant 20 **précurseurs** ou couples de précurseurs.

A partir de chaque structure des 7 niveaux de l'arbre de synthèse déjà obtenus il est possible de

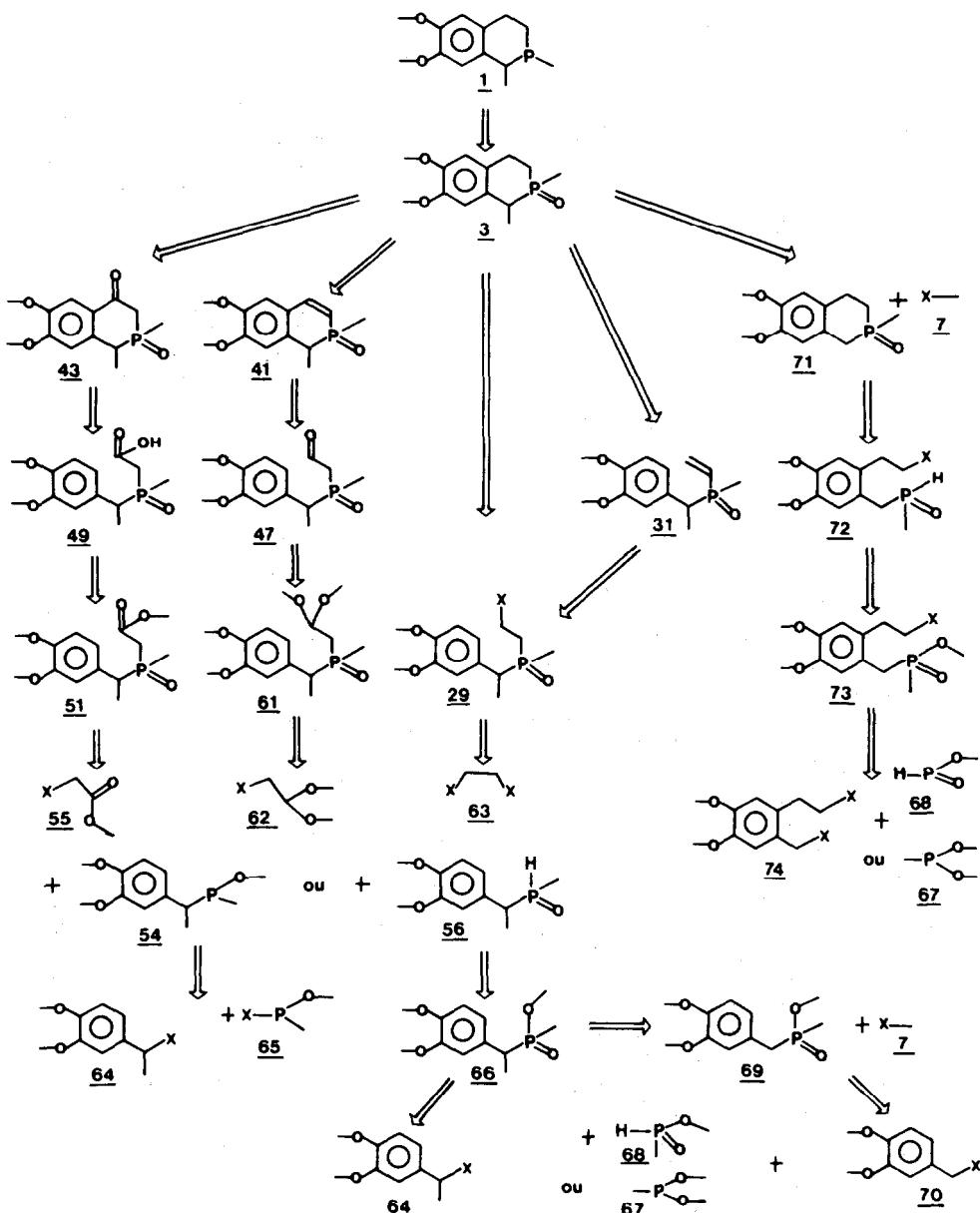


Fig. 6. Plan pour la synthèse de la phosphacarnégine établi avec l'aide de PASCOP.

developper de nouvelles ramifications. Ainsi après avoir explore notre **stratégie basée** sur la rupture du cycle phosphori: de 3, nous pouvons en tester une autre, à partir de la **même cible** 3, dans laquelle l'**élimination** du **stéréocentre** situé en a du phosphore est un **objectif** prioritaire. Pour cela, par exemple, nous proposons à PASCOP le programme de strattgie suivant:

```

1) BREAK BOND 1 1-15
NEXSTRAT
2) BREAK RING 3, 2, 14, 13, 12, 11
END.

```

En réponse le **système** applique la transformation ALKYLA (cfannexe) pour donner le **précurseur 71** (Fig. 6) à partir duquel 15 prcursors sont engendrés par application de la plupart des transformations employées **précédemment** pour rompre le cycle phosphoré de 3 (Fig. 3).

La Fig. 6 représente le plan de synthtse que nous avons établi pour la phosphacarnégine en suivant différentes stratégies, en combinant les 3 modes d'utilisation de PASCOP et en silectionnant, dans l'arbre ainsi développé les séquences nous paraissant les plus acceptables. Il est difficile d'évaluer directement les mérites respectifs des différents chemins contenus dans ce plan, aussi nous bornerons-nous à comparer ces voies avec celles suivies précédemment pour préparer des molécules ayant quelque analogie avec la phosphacamtgine. Nous constatons que l'étape cl6 31 → 3 s'apparente à la formation du squelette ttrahydroisophosphinoline obtenue par Dilbeck et al.¹⁰ en cyclisant un sel de benzyl diphenylvinylphosphonium dans l'acide polyphosphorique à température élevée. De même, l'étape 49 → 43 est analogue à la formation d'isophosphinolones décrite par Henning¹¹ puis par De Graaf et al.¹² Par contre 2 méthodes de préparation des ttrahydroisophosphinolines n'ont pas d'équivalence dans notre plan. L'une a été décrite par Märkl et Baier¹³ et suit une stratégie opposée à celles que nous avons explorées puisqu'elle vise à créer le cycle aromatique. Elle ne convient pas à la synthtse de la phosphacarnégine en raison de la substitution du noyau aromatique imposée par l'emploi d'un sel de triarylpyrilium comme produit de départ. L'autre méthode a permis à Mann et al.¹⁴ de synthétiser la première ttrahydroisophosphinoline. Elle crée le cycle phosphoré par formation de la liaison entre le phosphore et le carbone benzylique dans des conditions expérimentales où les éthers de méthyle et d'aryle sont clivés. Enfin il semble que des réactions analogues à celles des étapes 47 → 41 et 72 → 71 n'aient jamais été utilisées pour construire la structure ttrahydroisophosphinoline, la réaction impliquée dans l'étape 72 → 71 est cependant connue pour avoir conduit à divers hétérocycles phosphorés.^{15,16}

Dans cette première partie de notre travail nous avons montré comment on peut utiliser PASCOP dans l'établissement d'un plan de synthèse pour une molécule telle que la phosphacamtgine. En explorant quelques stratégies simples nous avons développé différents chemins parmi lesquels certains sont analogues à des voies déjà suivies pour la synthèse de dérivés de la ttrahydroisophosphinoline tandis que d'autres sont plus originaux. La validité d'un plan de

synthèse ne pouvant être affirmée sans vérification expérimentale, la seconde partie de notre travail a été consacrée à l'expérimentation des différents chemins constituant le plan Ctabli pour la phosphacarnégine.¹⁷

ANNEXE

Transformations citées et réactions chimiques correspondantes

2ARALK	Alkylation de Friedel et Crafts avec une oléfine
7ACIDE	Hydrolyse des esters d'acides carboxyliques
ALC.HY	Reduction d'alcools benzyliques ou aldehyques
ALKYLA	Alkylation en alpha d'un groupe électro-attracteur
ARACYL	Acylation de Friedel et Crafts avec un chlore d'acide
ARALKYL	Alkylation de Friedel et Crafts avec un halogénure d'alkyle
ARBREA	Rearrangement d'Arbusov
ARBUSOV	Reaction d'Arbusov
ARCYCL	Cyclodehydratation d'aldehydes et de cétones
ARRCFR	Acylation de Friedel et Crafts avec un acide carboxylique
C = C + P3	Addition d'une phosphine primaire ou secondaire sur une oléfine
CHCNH	Reduction des amines primaires
co2 + YL	Préparation d'acides carboxyliques par carbonatation d'ylures ou d'organométalliques
DIELS	Reaction de Diels-Alder
G + ALLYL	Couplage d'un organométallique avec un halogénure d'alkyle ou de benzyle
HAL.HY	Hydrogenolyse des esters, tosylates, halogénures etc
IG RIGN	Addition intramoléculaire d'un organométallique sur une &one
INOKAWA	Addition d'un composé possédant un motif P(O)H sur une tosylhydrazone puis reduction de l'additut
KET.HY	Reduction des c&ones (réactions de Clemmensen, Wolff-Kishner etc)
MIBECK	Reaction de Michaelis-Becker
NUCARY	Condensation d'un organométallique avec un halogénure d'aryle
OLE. RE	Reduction des oléfines
PG RIGN	Reaction d'un organométallique avec un composé à liaison phosphore-halogène
PHADD +	Addition d'un composé à liaison phosphore-halogène sur une oléfine
PH3 + CO	Reaction d'une phosphine primaire ou secondaire sur une c&one ou un aldehyde
PORGRI	Reaction d'un organométallique sur un composé à liaison phosphore-alkoxy
REDUCP	Reduction des oxydes de phosphines
REDPIU	Reduction des sels de phosphonium
RX+PH	Reaction d'une phosphine primaire ou secondaire avec un halogénure
WILETH	Synthèse des éthers de Williamson.

REFERENCES

- ^aCommunication préliminaire: C. Laurenço et G. Kaufmann, *Tetrahedron Letters* 2243 (1980).
- ^bL'organisation et le fonctionnement de PASCOP (Programme d'Aide à la Synthèse des Composés Organiques et Organophosphorés) b nt été décrits précédemment: ^aF. Choplin, R. Marc, G. Kaufmann et W. T. Wipke, *J. Chem. Inf. Comput. Sci.* **18**, 110 (1978); ^bF. Choplin, C. Laurenço, R. Marc, G. Kaufmann et W. T. Wipke, Now. *J. Chim.*

2,285 (1978); ^cF. Choplin, P. Bonnet, M. H. Zimmer et G. Kaufmann, *Ibid.* 3, 223 (1979); ^dP. Bonnet, J. C. Demiaume, M. H. Zimmer, F. Choplin et G. Kaufmann, *R.A.I.R.O.* 13, 287 (1979); 'M. H. Zimmer, F. Choplin, P. Bonnet et G. Kaufmann, *J. Chem. Inf. Comput. Sci.* 19, 235 (1979). Pour une description simplifiée du fonctionnement de PASCOP voir: G. Kaufmann et C. Laurenço, *La Recherche* 12, 743 (1981).

^eL. D. Quin, *The Heterocyclic Chemistry of Phosphorus*. Wiley-Interscience, New York (1981).

^fK. D. Berlin et D. M. Hellwege, *Topics in Phosphorus Chemistry* (Edited by M. Grayson et E. I. Griffith), Vol. 6, p. 1. Interscience, New York (1969); ^gS. D. Venkataramu, G. D. MacDonell, W. R. Pudrum, M. El-Deek et K. D. Berlin, *Chem. Rev.* 77, 121 (1977).

^haG. Heyl, *Arch. Int. Pharmacodyn.* 239, 451 (1901); ^h*Ibid.* 266, 668 (1928). Les propriétés pharmacologiques de la carnégine ont été étudiées par E. Santi-Soncin et M. Furlanut, *Fitoterapia* 43, 21 (1972).

ⁱO. Awerbouch et Y. Kashmann, *Tetrahedron* 31, 33 (1975).

^jC. H. Chen, K. E. Brighty et F. M. Michaels, *J. Org. Chem.* 46, 361 (1981).

^ka*Computer-Assisted Organic Synthesis*. (Edited by W. T. Wipke et W. J. Howe), A.C.S. Symposium Series, Vol. 61. American Chemical Society (1977); ^bM. Bersohn et A. Escaks, *Chem. Rev.* 76, 269 (1976); ^cP. Gund, *Ann. Reports Medicin. Chem.* (Edited by F. H. Clarke), Vol. 12. v. 288, Academic Press., New York (1977); ^dM. Chanoh' et R. Barone, *Encyclopedia Universalis*, supplément, p. 1369, Paris (1979-80).

^lR. Koncos et B. S. Friedman, *Friedel-Crafts and Related*

Reactions (Edited by G. A. Olah), Vol. 2, p. 289. Interscience, New York (1964).

^mG. A. Dilbeck, D. L. Morris et K. D. Berlin, *J. Org. Chem.* 40, 1150 (1975).

ⁿH. G. Henning, *Z. Chem.* 5, 417 (1965).

^oaH. G. de Graaf, J. Dubbeldam, H. Vermeer et F. Bickelhaupt, *Tetrahedron Letters* 2397 (1973); ^bH. G. de Graaf et F. Bickelhaupt, *Tetrahedron* 31, 1057 (1975).

^pG. Märkl et H. Baier, *Tetrahedron Letters* 4439 (1972).

^qaF. G. Holliman et F. G. Mann, *J. Chem. Soc.* 1634 (1947); ^bM. A. Beeby and F. G. Mann, *Ibid.* 411 (1951).

^raR. B. Wetzel et G. L. Kenyon, *J. Am. Chem. Soc.* 94, 1774 (1972); ^b*Idem. J. Org. Chem.* 39, 1531 (1974).

^sT. H. Chan et K. T. Nwe, *Phosphorus* 3, 225 (1974).

^tC. Laurenço, J. Villien et G. Kaufmann, *Tetrahedron* 40, 2731 (1984).

^uUne synthèse d'indazoles, en une étape, trouvée avec l'aide d'un programme de synthèse assistée par ordinateur et testée expérimentalement a été décrite par R. Barone, P. Camps et J. Elguero, *Ann. Quim.* 75, 736 (1979). Par ailleurs, deux chemins de synthèse proposés par de tels systèmes, l'un pour l'alcool de patchouli: E. J. Corey et W. T. Wipke, *Science* 166, 178 (1969) et l'autre pour le thiazole: R. Barone, M. Chanon et J. Metzger, *Tetrahedron Letters* 2761 (1974) ont été expérimentés ultérieurement par d'autres équipes, respectivement: F. Naf et G. Ohloff, *Helv. Chim. Acta* 57, 1868 (1974) et P. Dubs et R. Stuessi, *Synthesis* 696 (1976), mais il semble que le travail de ces dernières soit indpendant et ne résulte pas d'une volonté de tester les propositions d'un logiciel informatique.