

# Outils informatiques du chercheur

## I Outils à disposition du chercheur

Que ce soit pour taper un simple texte, comparer ses résultats ou modéliser une molécule, l'ordinateur est devenu un outil indispensable au chercheur scientifique, particulièrement pour faire de la **bibliographie**.

### A. Logiciels

Le chercheur devra manipuler des **logiciels très généraux** :

- ➔ Traitements de textes (*Word*), pour écrire ses publications, qui doivent posséder un formatage spécifique
- ➔ Tableurs (*Excel*), afin d'effectuer automatiquement des tris ou des calculs sur une liste importante de données
- ➔ Éditeurs d'équation et grapheurs (*Deltagraph*), qui permettent d'écrire par ordinateur des formules mathématiques, ou de mettre ses résultats sous forme de graphiques

Plus spécifiquement, le chercheur sera amené à manipuler des **logiciels scientifiques** qui l'aideront dans son travail, voire qui sont indispensables pour manipuler ses résultats. Peu de logiciels existent pour la chimie organique.

- ➔ Dessin de molécules (*IsisDraw, ChemDraw*), pour dessiner des réactions en chimie organique ou des mailles cristallines en chimie inorganique
- ➔ Statistiques, souvent requis pour comparer plusieurs essais, méthodes, etc ...
- ➔ Gestion de documents (*Endnote*), pour classer et retrouver facilement des publications que l'on a cherchées
- ➔ Modélisation, pour visualiser une molécule dans l'espace, la conformation d'une protéine, l'arrangement d'un cristal ... Indispensable en chimie théorique
- ➔ Simulation de spectres, des logiciels pratiques qui permettent d'obtenir le spectre théorique d'une molécule. On peut ensuite le comparer au spectre réel.
- ➔ Aide à la synthèse, ces logiciels peuvent imaginer les étapes nécessaires à la synthèse d'une molécule ; cependant l'ordinateur déclare forfait plus facilement qu'un humain.

### B. Sources d'informations

Aujourd'hui l'importance des réseaux de communication (surtout Internet) offre une quantité gigantesque d'informations et de publications disponibles pour le chercheur.

Le problème qui se pose est de trouver les informations qui nous intéressent. Selon l'étude de **March**, on distingue différentes sources d'informations.

#### 1) Sources primaires : littérature originale

C'est l'origine de l'information, celle directement trouvée par un chercheur. On peut la trouver dans divers documents.

#### ▶ **Publications et revues scientifiques**

Ce sont des journaux où les chercheurs publient ce qu'ils ont trouvé. Ils constituent 80 % des sources primaires. Elles sont toujours en Anglais, et maintenant quasiment toujours sous forme électronique.

#### ► **Brevets**

Ce sont des inventions industriellement intéressantes, déposées au nom de l'inventeur. Ils constituent 20 % des sources primaires.

#### ► **Mémoires et thèses / rapports techniques**

Un **mémoire** (50 pages) ou une **thèse** (300 pages) est un rapport détaillé de la recherche menée dans une université. Ils présentent l'avantage de comporter ce qui n'a pas fonctionné.

Un **rapport technique** est aussi un rapport de recherche, mais réservé à l'entreprise où il a été produit.

#### ► **Conférences et congrès**

Pendant ces colloques, un chercheur (généralement le directeur de recherche) présente les avancées des recherches de son laboratoire. Ils ont l'avantage de pouvoir directement s'adresser au chercheur. Les conférences ont souvent lieu dans des endroits agréables, avec l'apéritif et la nourriture offerte ... C'est aussi un avantage ! ☺

On peut aussi faire ce que l'on appelle des **séances de poster**, où l'on peut voir des affiches décrivant les recherches en cours.

### **2) Sources secondaires : pointeurs bibliographiques**

Il n'est pas possible de lire toutes les publications (*3000 publications par jour* paraissent en moyenne). Heureusement les sources secondaires permettent de retrouver plus facilement des sources primaires.

#### ► **Recueil de résumés**

Aujourd'hui devenus obsolètes, c'étaient des journaux contenant des résumés faits par des lecteurs chargés de lire les nouvelles publications.

#### ► **Journaux spécialisés dans l'extraction de données**

Eux aussi maintenant peu utilisés, ils classaient les publications par mot-clé. À la manière d'une recherche informatique, on cherchait un mot-clé qui nous intéresse dans la liste, et on regardait les publications correspondantes. Avec l'informatique, c'est maintenant l'ordinateur qui fait ce travail.

#### ► **Abstracts et index**

De nos jours, les publications doivent posséder un bref résumé, fait par le chercheur lui-même. Ces journaux publient directement les résumés des publications, qui renvoient à la publication elle-même si elle nous intéresse.

#### ► **Revue**

Les revues sont faites par un chercheur, qui compile des publications (à peu près 300) à propos d'un même thème. Il donne l'état des connaissances sur le sujet, les recherches déjà effectuées et les pistes possibles. Ces revues sont très pratiques, car elles résument un sujet précis et donnent les sources auxquelles il s'attache.

Une **monographie** est une revue qui ne concerne que les publications du rédacteur.

### ► Livres et dictionnaires

Ils permettent de trouver rapidement une information précise. En chimie organique, on utilise par exemple des dictionnaires de noms de réactions, de molécules, etc ...

### ► Encyclopédies

Souvent bien faites et bien expliquées, on peut obtenir des informations générales à propos d'un sujet. Les meilleures sont *Encyclopedia Universalis* (la plus complète et détaillée) et *Wikipedia* (très vaste et en ligne). Elles comportent cependant peu d'informations techniques.

### **3) Sources tertiaires : guides**

Les guides permettent de retrouver des informations, après les avoir classées.

## II Littérature originale

### A. Publication

#### **1) Publication de l'article**

Les publications sont le résultat direct d'une recherche. Lorsque le chercheur pense avoir trouvé quelque chose d'intéressant, il envoie son article à l'éditeur. C'est le chercheur lui-même qui met en forme son article.

L'article est lu par des **referees**, des personnes spécialisées dans certains domaines qui valident la publication selon qu'elle présente un intérêt ou qu'elle a déjà été publiée. Ils demandent parfois une correction au chercheur. Aujourd'hui, grâce à l'informatique, l'article est très rapidement publié.

#### **2) Structure de la publication**

L'en-tête d'une publication comporte toujours un **titre** (2 lignes maximum). Il doit être pertinent et accrocheur. Ensuite la **liste des auteurs**, afin d'attribuer la publication aux auteurs respectifs ; elle comporte presque toujours une adresse e-mail, celle de l'auteur avec un astérisque. Enfin la **liste des universités** d'où les publications proviennent, pour le classement de Shanghai, qui donne une idée de la qualité de la publication.

Le classement de Shanghai est une liste internationale des universités, classées en fonction du nombre de publications. Les universités de Nancy sont classées 375/500 ☺.

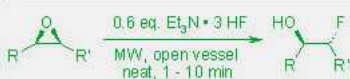
L'en-tête doit aussi posséder la **date de réception** de l'article, qui précise quand l'éditeur a reçu l'article. Si l'article a été révisé, on écrit aussi la **date de révision** et la **date d'acceptation**.

L'article doit ensuite comporter un **résumé (abstract)** de l'article, avec parfois (presque toujours en chimie organique) une image. Elle permet de créer des classements de résumés, qui permettent de trouver rapidement un résumé qui nous intéresse.

**Effective Fluorination Reaction with Et<sub>3</sub>N • 3 HF Under Microwave Irradiation**  
Tomotake Inagaki, Tsuyoshi Fukuhara, Shoji Hara\*

\*Division of Molecular Chemistry, Graduate School of Engineering, Hokkaido University, Sapporo 060-8628, Japan, Email: hara@org-mc.eng.hokudai.ac.jp

T. Inagaki, T. Fukuhara, S. Hara, *Synthesis*, **2003**, 1157-1159.  
DOI: 10.1055/s-2003-39162



**Abstract**  
Fluorinations of epoxides and alkyl mesylates can be effectively achieved by reaction with Et<sub>3</sub>N • 3 HF under microwave irradiation. The reactions were completed in a few minutes and the use of large excess of reagents could be avoided.

Le corps de l'article contient l'information de la publication. Il est écrit en Anglais, et structuré. On précise les références ou les annotations, à l'aide d'un petit chiffre en exposant, rapportant à un petit texte en bas de page. Le texte peut contenir des figures et tableaux de résultats accompagnés d'une légende.

On écrit ensuite la **partie expérimentale**, qui décrit précisément et en détails le protocole expérimental utilisé. Elle est très importante, car elle permet de recréer et vérifier les résultats.

L'article finit par la partie **results**, qui donne les applications possibles, les interrogations et d'autres pistes possibles d'exploitation.

### **3) Accès à la publication**

L'accès à une publication est presque toujours payant. Les laboratoires sont souvent abonnés à des journaux scientifiques, disponibles en ligne.

Pour obtenir un article d'un journal dont on n'a pas l'abonnement, on peut l'acheter individuellement à la bibliothèque, ou l'obtenir à l'**INIST**, un organisme chargé d'acheter toutes les publications, pour les revendre individuellement en France.

## **B. Brevet**

### **1) Exclusivité**

Lorsqu'une invention permet une application industrielle intéressante (presque toujours économiquement ...), le chercheur peut déposer un **brevet**. Cela permet d'attribuer l'invention à l'inventeur. Elle lui donne l'exclusivité de l'exploiter.

Pour déposer un brevet, l'invention doit respecter certains critères : elle doit être nouvelle, inventive et posséder une application industrielle. Si un même brevet a déjà été validé dans un autre pays, le chercheur doit défendre son brevet. Cette procédure peut durer très longtemps. La défense d'un brevet a duré 18 mois, entre l'Europe et les États-Unis.

Le brevet n'est valable que dans les pays où on l'a validé, c'est-à-dire que l'on peut l'exploiter dans un autre pays.

### **2) Prix du brevet**

Pour déposer un brevet, le chercheur doit payer une somme d'argent conséquente. Il faut ensuite repayer le brevet chaque année pour le maintenir :

- \_ 7 500 € plus les frais de juristes en France
- \_ 20 000 € au niveau Européen
- \_ 20 000 € au niveau États-Unien ou mondial

Une fois en possession d'un brevet, on peut vendre des **licences** à des entreprises (environ 10 000 €), afin qu'elles puissent exploiter l'invention en versant une petite partie des bénéfices à l'inventeur. Il est possible de vendre directement le brevet si quelqu'un veut bien l'acheter.

Lorsqu'on arrête de payer le maintien du brevet, celui-ci passé dans le domaine public. Cela signifie que tout le monde peut l'exploiter ; cependant personne ne peut déposer une deuxième fois le même brevet. Un brevet devient automatiquement public après 20 ans.



### III\_ Pointeurs bibliographiques

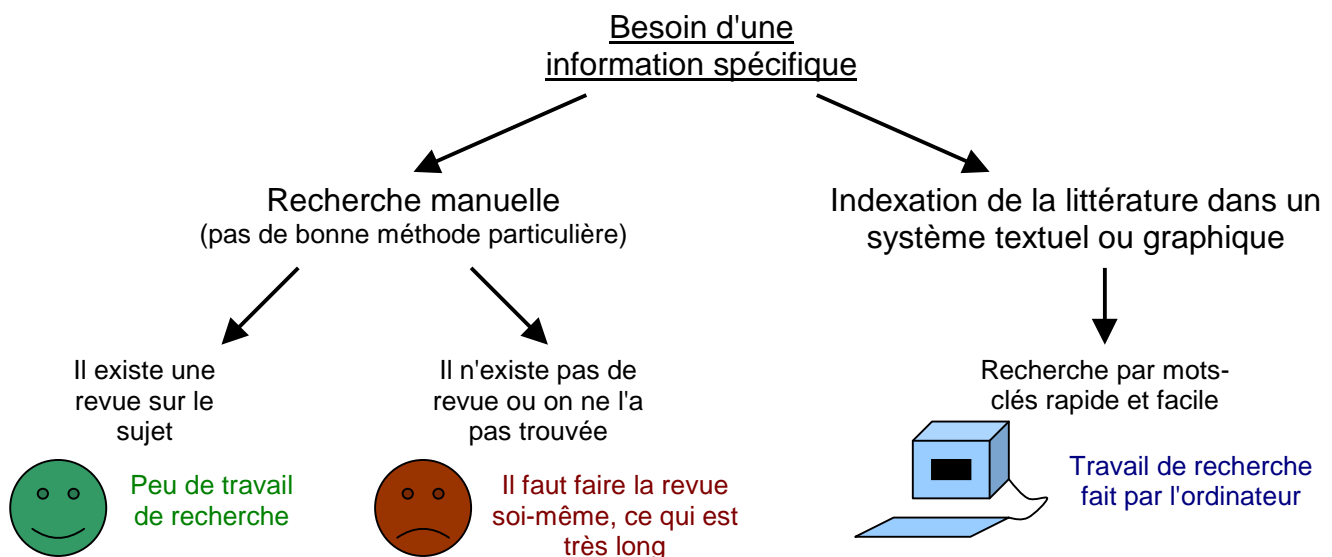
#### A. Chemical abstracts

Les *Chemical abstracts* sont des journaux qui ne publient que le résumé des publications, écrit par le rédacteur lui-même. Très pratiques, on les trouvait autrefois sous forme de gros livres, écrits en tout petit sur du papier très fin. Comprenez papier de merde, comme dans la Bible à la con, où on voit à travers le texte de derrière.

Aujourd'hui ces listes de résumés sont indexés dans des bases de données, où l'on peut effectuer une recherche informatique en ligne (ex : *PubMed*). Soit un milliard de fois plus pratique qu'avant.



#### B. Recherche d'informations



Grâce aux énormes progrès en informatique et en communication (Internet), la recherche informatique est une des meilleures (la meilleure ?) source d'information. La recherche peut être effectuée *on line*, sur des bases de données en ligne, ou *in house* à partir d'encyclopédies sur DVD.

Dans ce dernier cas, les encyclopédies électroniques sont souvent très bien faites et bien expliquées, cependant elles coûtent cher et deviennent vite obsolètes. On privilégiera la littérature en ligne pour une recherche d'informations.

#### C. Ressources en ligne

On peut trouver en ligne beaucoup de sources d'information, à la fois primaires et secondaires.

##### ► **Version électronique des journaux**

Aujourd'hui presque toute la littérature scientifique se trouve sur Internet. Cela permet une diffusion rapide des informations, et offre des possibilités de recherche informatique. On peut de plus obtenir des **compléments** (*supplementary materials*) sur les publications, tels que des tableaux de résultats, des spectres mesurés ...



C'est une excellente source bibliographique très complète, mais qui présente justement le problème d'être trop complète ... Il y a beaucoup de publications, dans beaucoup de journaux différents, il faut chercher les publications qui nous intéressent.

Il est possible de recevoir par courrier électronique la liste des publications sur le point d'être publiées. On peut alors les télécharger tout de suite (si on est abonné).

#### ► **Congrès internationaux électroniques**

Assez limitées, ces conférences permettent tout de même d'y participer sans avoir à se déplacer très loin. Le plus souvent cette conférence n'est que textuelle ...

#### ► **Sites généralistes, de laboratoires, d'universités**

Ces sites présentent souvent des ressources et/ou des liens vers des sites intéressants. On peut obtenir de la littérature scientifique, des cours, des programmes, des conférences, etc ...

On peut aussi participer à des **forums de discussion** scientifique. On peut alors poser des questions très spécifiques, y compris qu'est-ce qui ne pourrait pas marcher (non publié). Cependant n'importe qui peut y répondre, il faut donc apprécier la qualité des réponses.

#### ► **Fils RSS**

Ces petites fenêtres d'informations permettent de récupérer automatiquement et de faire défiler des nouvelles informations, sous forme de résumés, à partir d'un site précis. Il faut aussi s'abonner afin de récupérer l'information complète.

## D. Recherche en Chimie Organique, avec SciFinder

### 1) Organisation des bases de données

L'ordinateur doit être capable de retrouver une certaine information dans une base de donnée. En Chimie Organique, les publications sont retrouvées à partir de deux bases de données :

#### ➡ **Base de données de molécules**

Cette base de données doit permettre de retrouver des informations à propos d'une molécule, par son nom ou sa structure.

Ainsi on doit faire comprendre au programme comment les atomes s'associent, ainsi que lui apprendre le comportement des groupements importants.

Cette base de données partage aussi un index qui relie les différentes molécules à leur rôle dans les réactions (produit, réactif, solvant, catalyseur ... ).

#### ➡ **Base de données de réactions**

Cette base de données trie les publications par rapport au type des réactions étudiées. On apprend donc au programme les principales réactions en chimie organique, à partir des différentes molécules de l'**index partagé**.

On lui apprend aussi à reconnaître les **centres réactionnels**, c'est-à-dire qu'il doit savoir sur quel groupement la réaction va s'effectuer.

Cette base de données possède un **index additionnel** à propos des conditions opératoires à utiliser (température, conditions physiques ...).

### 2) Généralités sur SciFinder



**SciFinder Scholar** est un programme qui recherche des informations en ligne. Il faut un abonnement afin de pouvoir l'utiliser. Il est interdit de télécharger toutes les publications pour faire sa propre base de données, ni de télécharger un livre complet si celui-ci est en vente.

Dans l'UHP, seules 5 connexions simultanées sont autorisées dans toute l'université.

Il se connecte aux bases de données **Chemical Abstracts** ou **CAS**, qui regroupe toutes les publications en rapport avec la chimie aux États-Unis, et **Medline** qui référence des publications en rapport avec la médecine.

On dispose ainsi d'un contenu très diverse (chimie, physique, ingénierie, biologie, médecine, etc ...) et très complet (60 ans d'informations, soit 20 millions de publications et 3 millions de brevets).

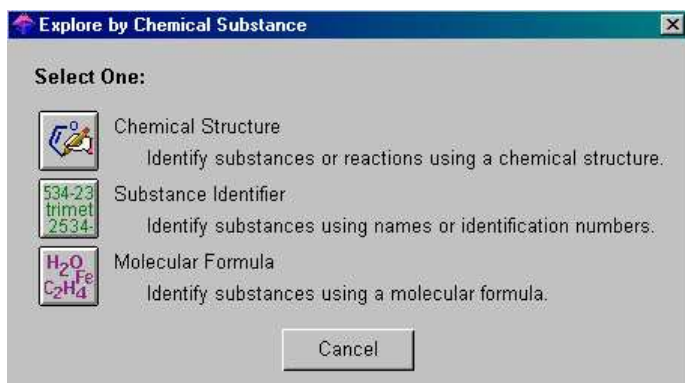
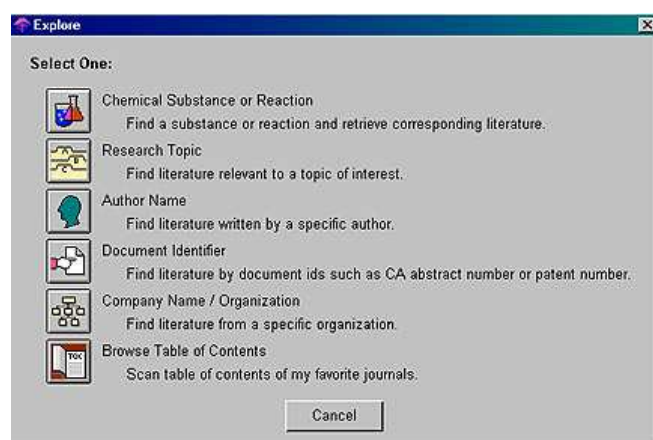
### 3) Principe d'utilisation

Dans les **préférences**, il est possible de limier ses recherches à certaines bases de données : bases de données de publications (Chemical Abstracts et Medline), base de données de molécules, base de données de réactions.

On peut aussi spécifier le mode d'affichage des publications, soit simplement le titre, soit avec la référence, soit tout le résumé.

En faisant une **nouvelle recherche**, plusieurs méthodes de recherche sont possibles.

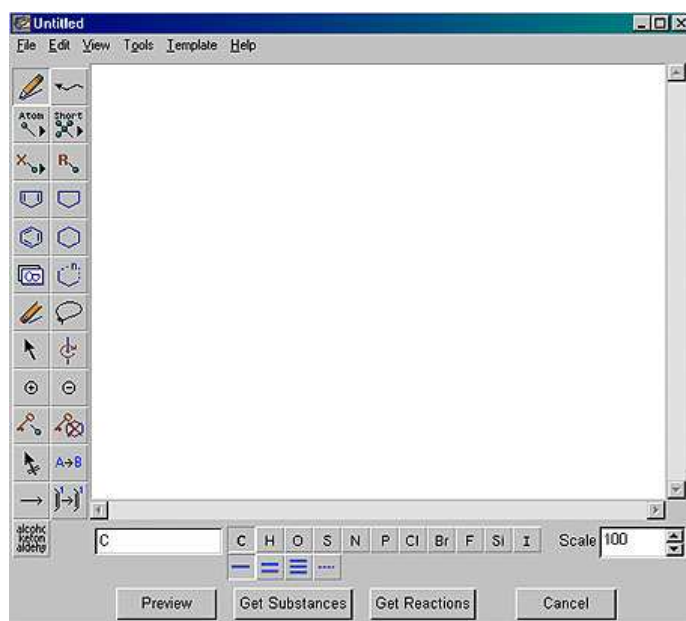
#### ► Recherche à partir d'une substance



Cette option permet de retrouver des publications relatives à une certaine molécule, ou un certain groupement.

Il est possible d'utiliser une fenêtre graphique qui nous permettra de tracer la molécule.

Cette fenêtre présente plusieurs options intéressantes : le **lissage** permet d'améliorer automatiquement la forme de la molécule ; on peut indiquer manuellement les centres réactionnels, les liaisons mises en jeu et numéroter les atomes afin d'aider l'ordinateur à mieux comprendre la réaction.

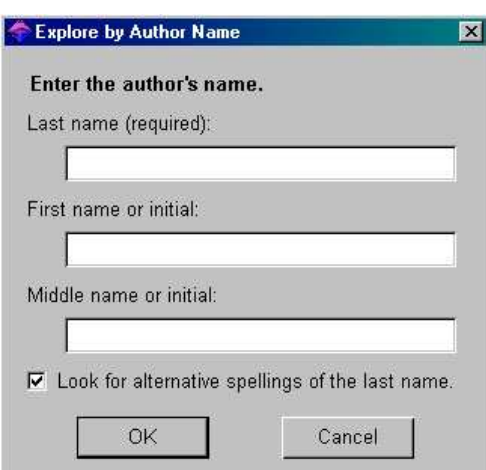
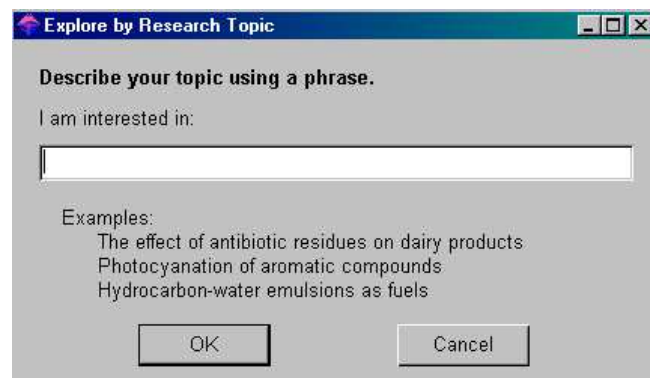


Si le programme trouve trop de publications relatives à cette molécule, il propose différentes molécules de structure proche, afin de limiter les recherches. C'est une méthode de recherche très utilisée en chimie organique.

La molécule peut aussi être trouvée par sa dénomination : **CAS number** (numéro unique assigné à chaque molécule), nom trivial, nom officiel, nom technique (ex : E323) ...

### ► Recherche par sujet

Cette méthode de recherche permet de rechercher des publications à propos d'un sujet général. Cette méthode n'est utile que pour se documenter ou pour trouver des sujets de recherche. SciFinder permet de rechercher en entrant des mots-clés, ou en tapant une phrase.

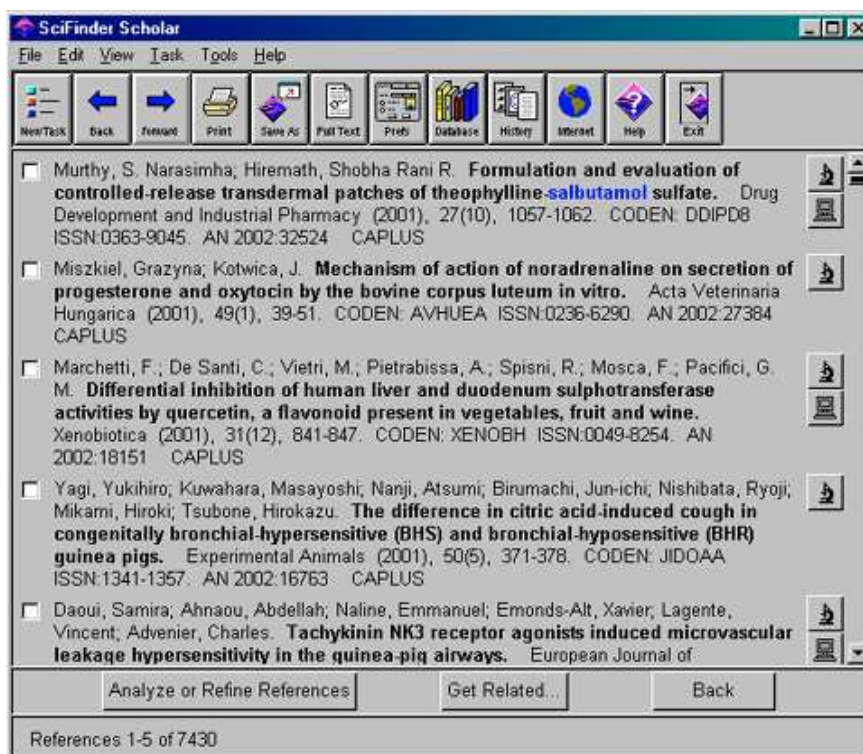


### ► Recherche par auteur

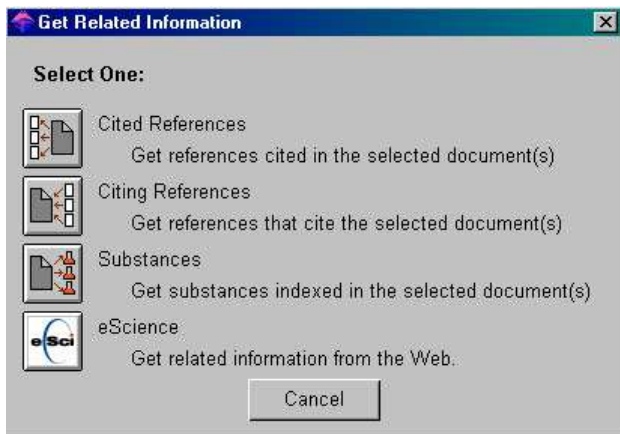
À partir du nom d'un auteur, on peut retrouver ses différentes publications. On peut aussi ajouter son prénom ou ses initiales.

Le logiciel peut tenter une recherche approchée, où il essaiera de trouver des auteurs ayant un nom qui ressemble à celui entré, au cas où on ne serait pas sûr de l'orthographe.

Les résultats sont alors affichés sous forme d'une liste de titres ou de résumés, accompagnés de leur référence afin de les consulter. Il est parfois possible de télécharger directement la publication, si on est abonné au journal correspondant.







Il est aussi possible d'obtenir très facilement des informations en rapport avec la publication grâce à SciFinder, qui permet de retrouver toutes les publications citées dans l'article, citant l'article, ou les molécules relatives à la publication.

Enfin on peut affiner chaque recherche en appliquant successivement plusieurs filtres de recherche (nouveau mot-clé, molécule, auteur, etc ... ).

