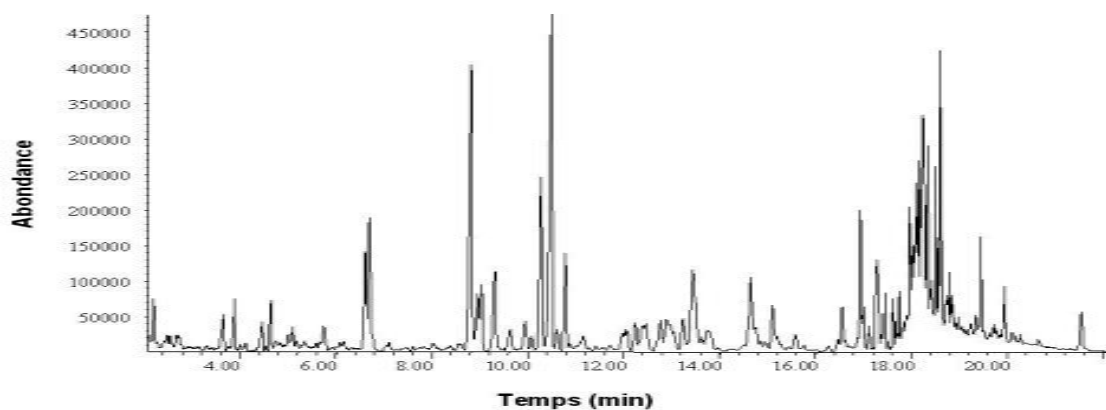
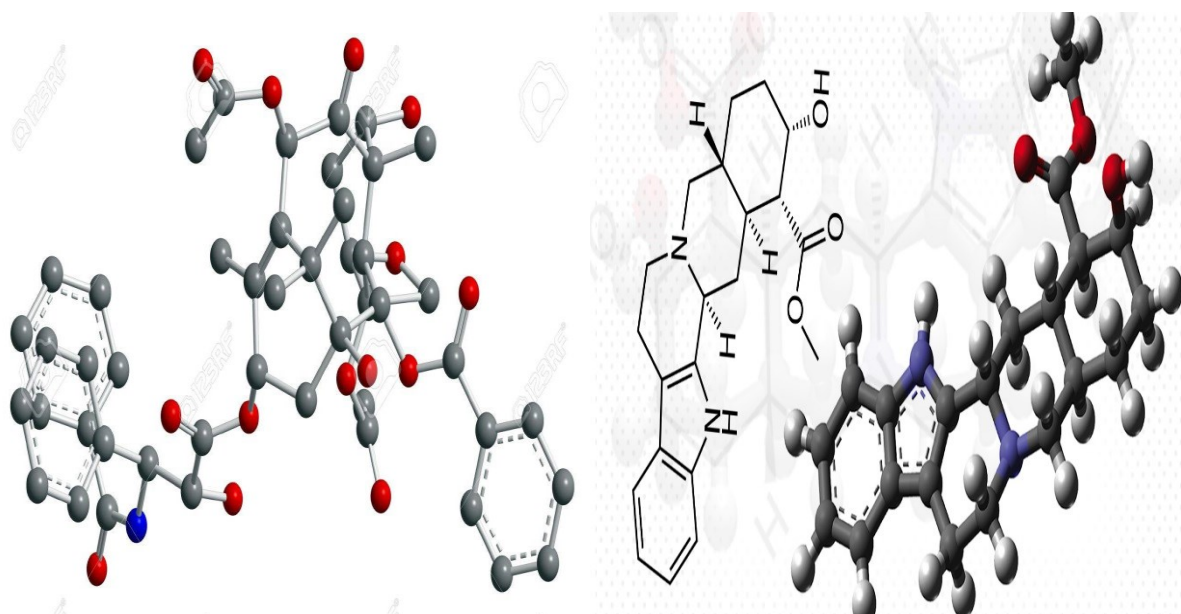


Polycopie de cours

Informatique pour la Chimie



Préparé par : Dr. BOUGHRARA Boudjema
Laboratoire de Physicochimie des Matériaux (LPCM)
Département de Chimie
Faculté des Sciences et de la Technologie
Université Chadli Bendjedid El-Tarf

Année universitaire 2025/2026

Le présent manuscrit a pour but de rassembler à l'usage des étudiants en chimie et en génie de procédés les données indispensables à la compréhension et à la pratique de développement, la création, l'organisation, le stockage, la diffusion, l'analyse, la visualisation et l'utilisation de l'information chimique, elle représente un domaine en évolution rapide

A l'heure actuelle les unités d'enseignements et de recherche dans le domaine des outils informatiques appliqués au domaine de la chimie

Ce cours s'adresse aux étudiantes troisièmes années licence chimie analytique, département de chimie, faculté des sciences et de la technologie Université Chadli Bendjedid El-Tarf désirant approfondir leurs connaissances dans les traitements statistiques et graphiques de données, Etude de banques de données chimiques indexées par structure.

Nous espérons qu'il est aussi clair, aussi précis et aussi simple.

Intitulé du Licence : chimie Analytique

Intitulé de l'UE : Unité d'Enseignement méthodologie

Intitulé de la matière : informatique pour la chimie

Crédits : 2 **Coefficient :** 1

Programme d'enseignement

Titres	Page
Introduction	07
1. Initiation aux outils informatiques appliqués au domaine de la chimie	09
• La chemoinformatique	09
2. Introduction aux systèmes d'exploitation type Unix/Linux	12
• La démarche expérimentale par l'ordinateur	12
• Unix est multi-utilisateurs	18
3. Traitement statistique et graphique de données	21
4. Méthodologie de la recherche d'informations en Chimie	27
5. Représentation de la structure 3D et applications locales	31
6. Initiation à la modélisation moléculaire	36
Référence bibliographique	40

Liste des figures

	Titres	Page
Figure.01	Les éléments de l'ordinateur	12
Figure.02	Représentation de données	14
Figure.03	Liaison entre les ressources matérielles d'un ordinateur	16
Figure.04 :	Quelques logiciels et leurs fonctions	18
Figure.05 :	Rapport entre les descripteurs de structure et la propriété	22
Figure.06 :	Méthodes pour apprendre des données	23
Figure.07 :	Quelques types des représentations 3D	32
Figure.08 :	Types de modélisation et leurs applications	37

Liste des tableaux

Titres	Page
Tableau 01 : Type de Capteur et leur Application	13
Tableau 02 : Quelques exemples de bases de données et leur domaine d'utilisation	24
Tableau 03 : Quelques questions préalables de la recherche d'information en chimie	27
Tableau 04 : Quelques logiciels et Base de données employés la représentation 3D	31
Tableau 05 : Essentiels types des représentations 3D	31

LISTE DES PHOTOS :

Titre	page
Photo 01 : Les Capteurs	12
Photo 02 : L'Interface	12
Photo 03 : Représentation en fil de fer	32
Photo 04 : Représentation (CPK)	32
Photo 05 : Représentation en rubans	32

ABBREVIATIONS

pH : Potentiel hydrogène

2D : Représentation les molécules en 2 dimensions

3D : Représentation les molécules en 3 dimensions

SE : Système d'exploitation

UNICS : Uniplexed Information and Computer Service.

CHIRBASE : banque de données pour la séparation chromatographique chirale

THERMODATA : Banques de données experts pour la thermochimie

SXD : Extraction liquide-liquide des métaux

Introduction :

La Chemoinformatique est un nouveau domaine apparu il y a quelques années. Il concerne le développement, la création, l'organisation, le stockage, la diffusion, l'analyse, la visualisation et l'utilisation de l'information chimique, elle représente un domaine en évolution rapide. Actuellement, les méthodes spécifiques de la chemoinformatique indispensables pour le développement de nouveaux composés, matériaux et processus.

Ainsi, les méthodes ont été développées pour la construction des bases de données sur les composés chimiques et leurs réactions, pour la prédiction des propriétés physiques, chimiques et biologiques des composés et des matériaux à l'échelle de l'atome et de la molécule, dans tous les secteurs de l'activité humaine, la conception des médicaments, pour élucider la structure, pour la prévision des réactions chimiques et de la conception de synthèse organique. La recherche et le développement sont essentiels en chemoinformatique d'une part pour accroître notre compréhension des phénomènes chimiques et d'autre part pour que l'industrie reste compétitive dans une économie mondiale.

La Chémoinformatique vise à exploiter les connaissances en chimie et chimie-physique pour aider à rationaliser la recherche et le développement de nouveaux composés et substances. Le concept n'est pas neuf et, au fond, constitue le corpus de ce que fait tout chimiste. Mais l'informatique et la fouille de données ont complètement bouleversé les choses, faisant émerger une discipline scientifique à part entière et de nouveaux métiers. Elle a principalement été développée par l'industrie pharmaceutique. Ces technologies sont un élément essentiel pour la compétitivité des laboratoires. Mais elles ont aujourd'hui des applications dans n'importe quel domaine de la chimie, depuis la chimie analytique (chimie des solutions, réactions cinétiques, etc...) à la chimie organique (synthèse de réactions, mécanismes, etc...). La chemoinformatique peut être appliquée pour le développement de produits phytosanitaires, de produits cosmétiques, agroalimentaires, de solvants, etc... En outre, elle a une importance particulière le développement des médicaments.

- **Initiation aux outils informatiques appliqués au domaine de la chimie**

I. Initiation aux outils informatiques appliqués au domaine de la chimie

Les premiers essais de l'utilisation de l'ordinateur dans l'enseignement des sciences ont concerné principalement son emploi comme outil de laboratoire, précisément dans l'acquisition et le traitement des données expérimentales. L'utilisation des simulations informatiques rencontrait par contre des difficultés, car on craignait de donner un prétexte aux enseignants pour ne pas faire des expériences [Durandea & Sarmant, 1996]. Aujourd'hui et grâce à l'évolution massive de l'informatique sur les plans matériel et logiciel, mais grâce aussi à une vision nouvelle de l'enseignement, un enseignement basé sur la mise en place des conflits cognitifs et des activités de recherche, la simulation informatique se situe au même niveau que l'expérimentation [Durey & Beaufile, 1998]. En effet, l'utilisation des simulations informatiques dans l'enseignement des sciences a fait l'objet de plusieurs recherches en sciences de l'éducation. White [1993] et Lewis & al [1993] ont proposé des démarches pédagogiques utilisant les simulations informatiques, permettant de développer des modèles mentaux chez les apprenants.

Gorsky & Fingold [1992] ont utilisé les simulations informatiques pour mettre en place des situations de conflits cognitifs permettant aux apprenants de raffiner leur conception. Utilisant les simulations informatiques, Simmons & Lunetta [1993] ont développé et évalué une stratégie de recherche dynamique permettant le développement des habilités de raisonnement chez des novices et des experts.

Dans une étude récente de Monaghan & Clement [1999] intitulée « Use of a computer simulation to develop mental simulations for understanding relative motion concepts », les auteurs montrent que l'activité de simulation informatique permet de créer un conflit cognitif et conduit les apprenants à construire une nouvelle représentation sur le concept « mouvement ». Il s'ensuit donc de ce panorama bibliographique que les simulations informatiques peuvent créer un contexte pédagogique permettant la mise en place des situations de recherche et de conflit cognitif.

La **chemoinformatique** également démontré chemoinformatique émergé pour la première fois lors du développement d'un médicament en 1998 par **FRANCK BROWN**. C'est une branche de la chimie et /ou de la physico-chimie qui consiste à l'application de l'informatique aux problèmes relatifs à la chimie. Elle utilise les lois de la chimie théorique exploitées dans des codes informatiques spécifiques afin de calculer les propriétés des structures

chimiques. Elle a pour objectif de fournir des outils et des méthodes pour le traitement des données issues des différents domaines de la chimie.

-Introduction aux systèmes d'exploitation type Unix/Linux.

II - L'expérimentation assistée par ordinateur

II.1. La démarche expérimentale par l'ordinateur

L'ordinateur peut jouer un rôle important dans l'apprentissage de la démarche expérimentale en sciences, et plus généralement, de la méthode scientifique. Il peut l'assurer en étant un triple outil pour l'étudiant : un instrument de mesure, de traitement et de communication. Si l'on décompose la méthode scientifique, on peut déjà voir le rôle que peut jouer l'ordinateur :

- **Observation**
- **Formulation d'Hypothèse**
- **Expérimentation**
- **Mesure et Modélisation**
- **Interprétation des Résultats**
- **Communication des résultats**

Branches où l'ordinateur a un rôle à jouer

Les compétences visées

- Gérer correctement un ordinateur et réaliser des installations simples
- Utiliser de façon performante les logiciels permettant de rédiger un rapport scientifique, et d'en faire un exposé ; d'effectuer du traitement de données simple ; de rechercher de façon efficace des informations sur internet.

Les Logiciels Word, statgraphics, Excel, Bureautique, Logiciel de traitement de texte, Logiciel de présentation, Logiciel statistique, Tableur...



Fig. 1. Les éléments de l'ordinateur

a. Les Capteurs

Ce sont les yeux de l'ordinateur. L'utilisateur doit savoir les choisir selon la mesure qu'il désire effectuer, ainsi que les connecter. Leur rôle est de transformer une grandeur physique en une grandeur électrique mesurable. Il en existe de nombreuses sortes, permettant ainsi de rendre l'ordinateur polyvalent.

Tableau 1 : Type de Capteur et leur Application

Domaine	Type de Capteur	Application
Chimie	pH-mètre, conductimétrie	Analyse de solutions
Thermodynamique	Thermorésistance, diode	Mesure de température
Electricité / Magnétisme	Ampèremètre, sonde à effet Hall	Etudes de circuits, de champs magnétiques

b. L'Interface

Il s'agit d'un boîtier électronique muni de nombreuses prises, destinées à connecter les sondes, et d'un fil qui le relie à l'ordinateur. Elle est caractérisée par ses tensions d'entrée (0: + 5V, -10V: + 10V,...), sa fréquence d'échantillonnage (40kHz), et sa précision (8 bits, 12 bits,...). Il peut être intéressant de parler de ces caractéristiques à l'élève, car elles vont conditionner les manipulations, aussi bien sur les branchements que sur l'acquisition. L'interface transforme les grandeurs électriques des capteurs en valeurs numériques interprétables par l'ordinateur.

c. Les Logiciels

Dans la lignée des capteurs et de l'interface, on trouve le logiciel d'acquisition. Complet, il permet de piloter l'interface, de lire la valeur qu'elle émet, de l'enregistrer, d'effectuer des calculs précis sur ces valeurs, d'afficher les courbes, et de procéder à la modélisation. La difficulté est de savoir manipuler le logiciel en exploitant toutes ses fonctions pour résoudre son problème. On utilise également un tableur (type Excel) pour la représentation de données saisies manuellement. Ses possibilités graphiques sont certaines, et il dispose également de fonctions de modélisation.

d. Communication :

- Il convient de mettre en place des solutions de communication des résultats.
- Les noms chimiques sont généralement utilisés à des fins de documentation et de communication.

- Pour faciliter la communication chimique via un support électronique, des formats linéaires ont été développés sur une période de temps.

Exemple ; Utilisation de ChemDraw

- Créez et ouvrez la molécule dans ChemDraw comme décrit ci-dessus
- Ensuite, utilisez « modifier » ➡ obtenir un modèle 3D
- La sortie sera la structure 3D de la molécule comme indiqué dans la Fig. 2

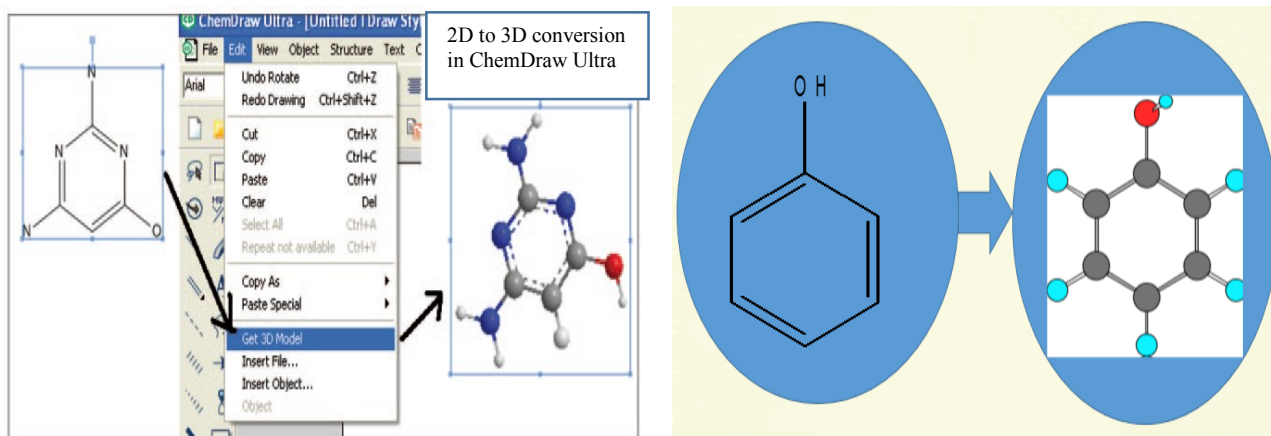


Fig 2 : représentation de données

- Démarrez ChemDraw
- Ouvrez le panneau d'outils chemdraw et dessinez la structure avec la barre d'outils suivante
- Après avoir dessiné la structure, faites un clic droit sur la structure, puis sélectionnez la molécule

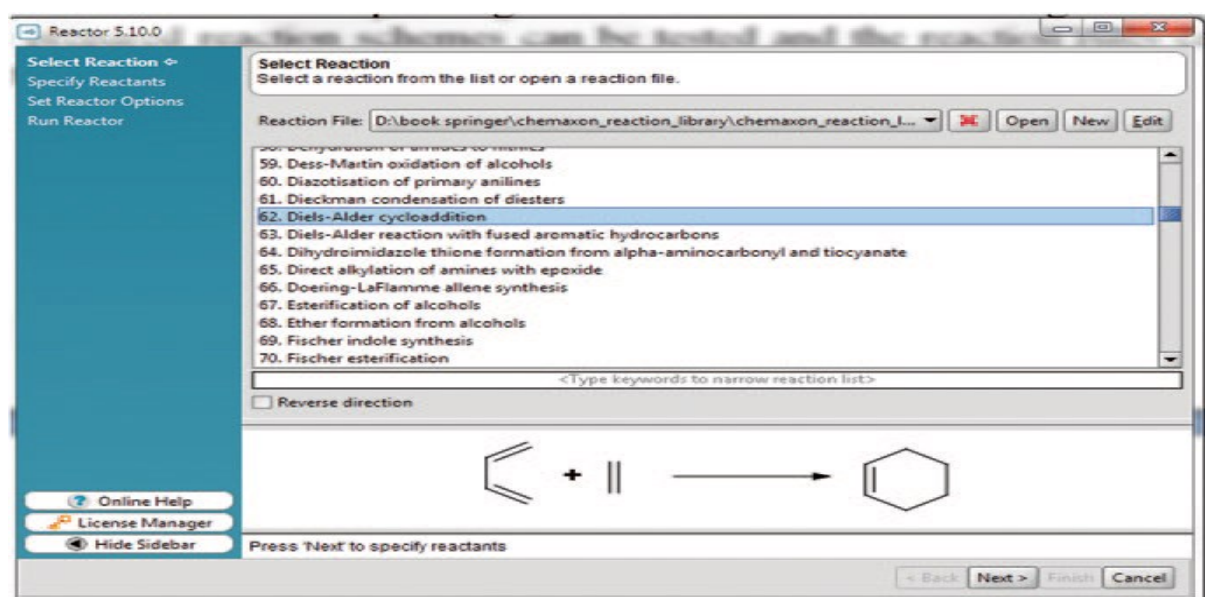
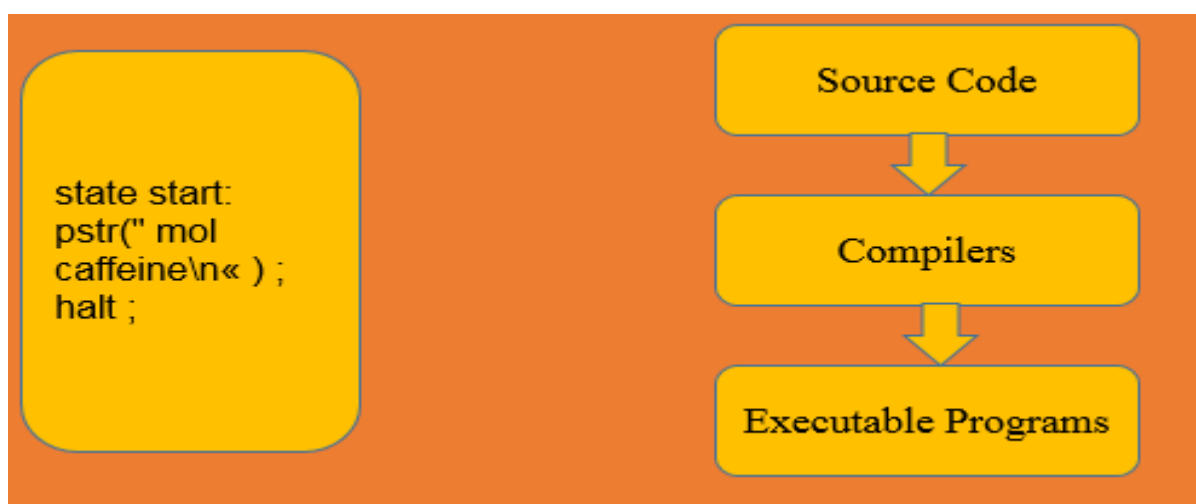
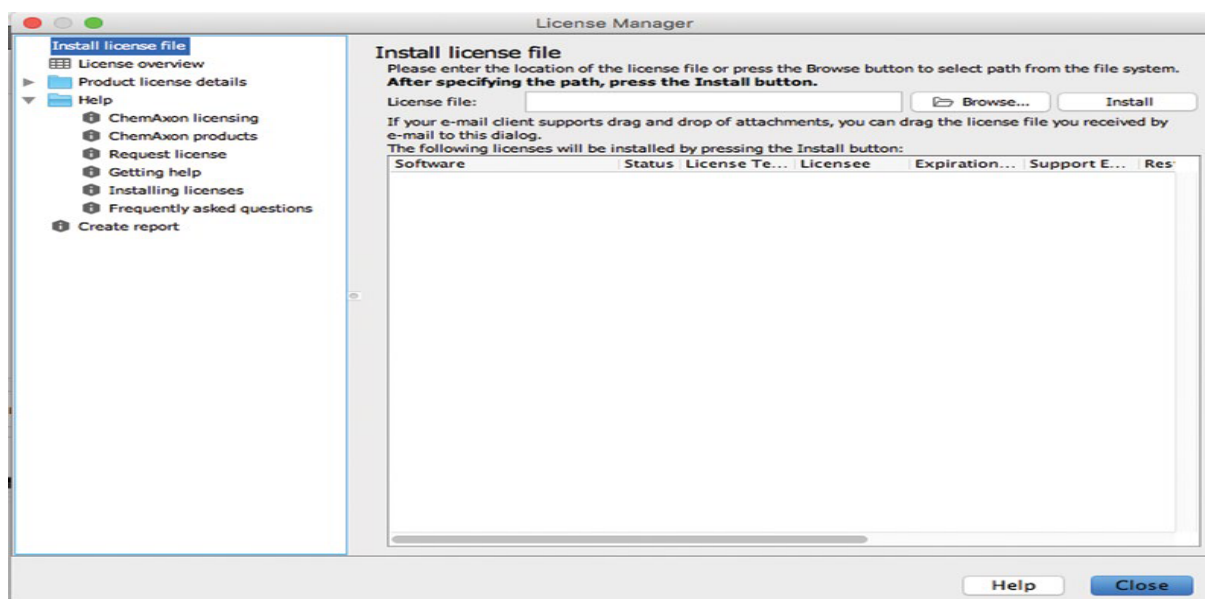
Copier comme ➡ SMILE ➡ collez-le où vous voulez

Obtenir une structure à partir de SMILES à l'aide de ChemDraw

- Copiez les SMILES du fichier source
- Cliquez avec le bouton droit de la souris sur la fenêtre de l'éditeur ChemDraw
- Cliquez sur Coller les sourires

II.2. Introduction aux systèmes d'exploitation type Unix/Linux

Un système d'exploitation (SE) est un ensemble de programmes responsables de la liaison entre les ressources matérielles d'un ordinateur et les applications informatiques de l'utilisateur (traitement de textes, vidéo des expériences (chimiques, physiques, biologiques). Il fournit aux programmes applicatifs des points d'entrées génériques pour les périphériques. Parmi les premiers systèmes à temps partagé, nous citons: **CTSS, MULTICS, UNIX, MINIX, LINUX.**



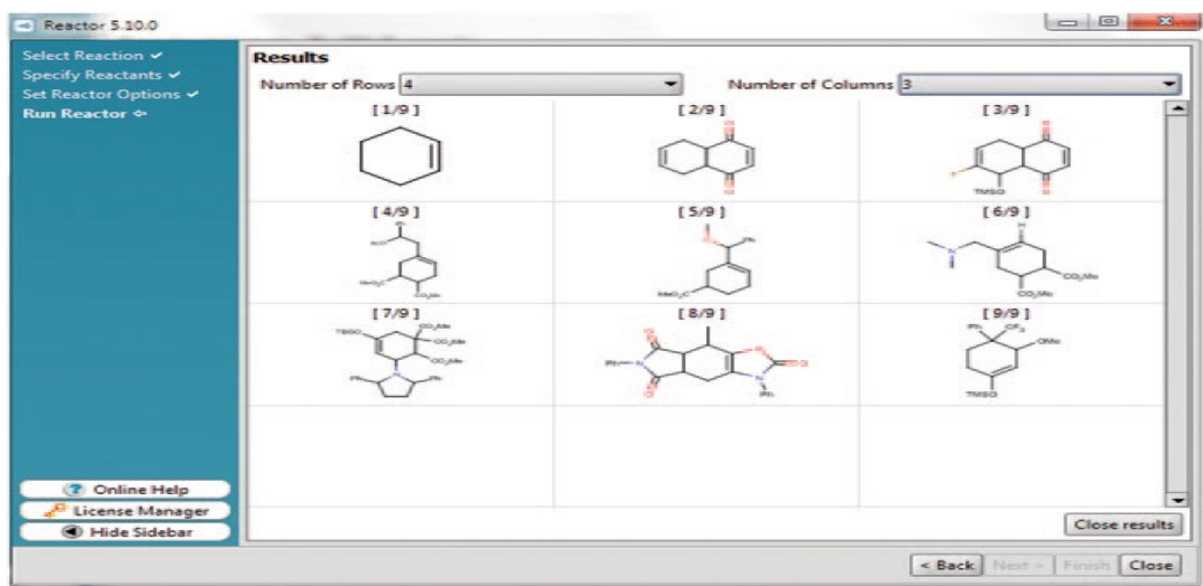
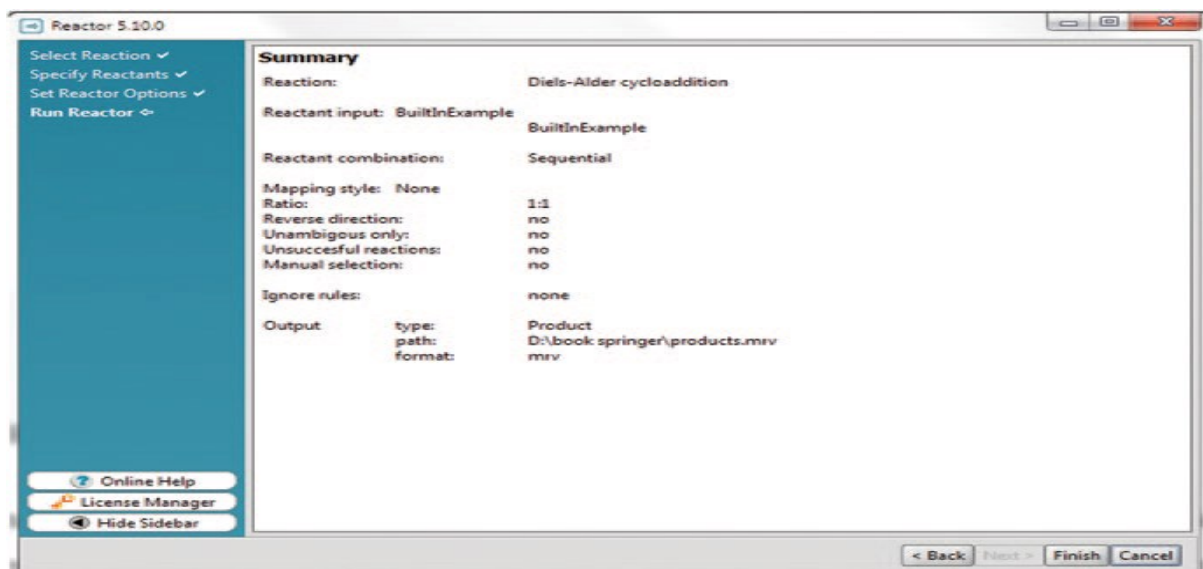
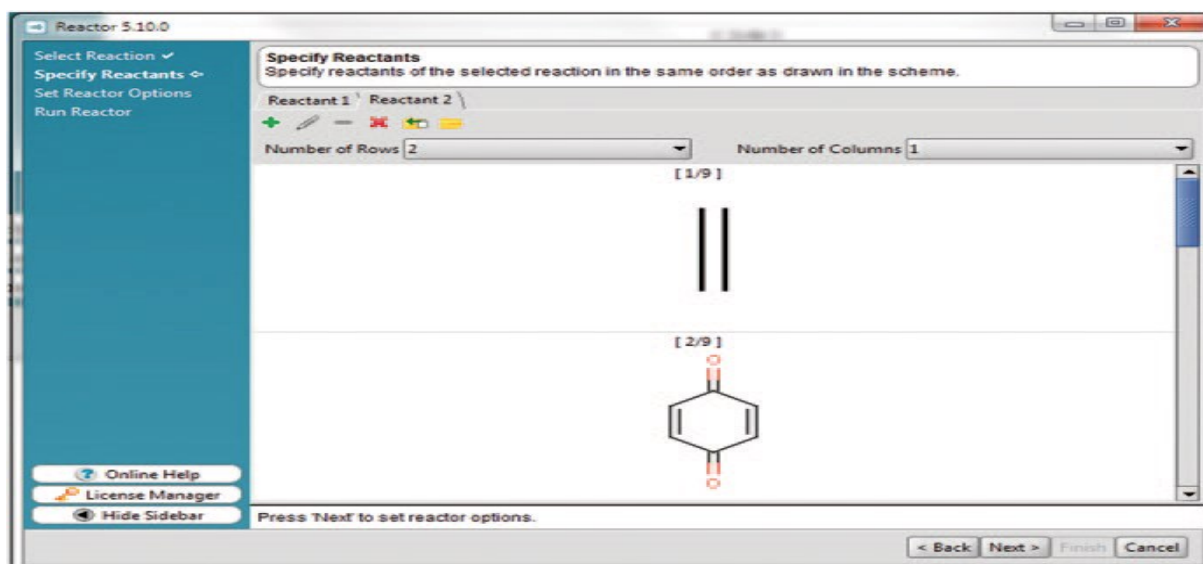


Fig 3 : Liaison entre les ressources matérielles d'un ordinateur

L'histoire du système d'exploitation commence en 1969, et ceci avec le développement d'une version simplifiée du système **MULTICS** par **Ken Thompson**. **Brian Kernighan** appela cette version **UNICS (Uniplexed Information and Computer Service)**.

- ❖ La possibilité d'utiliser l'ordinateur par une interface graphique ou plutôt une interface homme-machine.

1 - Lancer des programmes

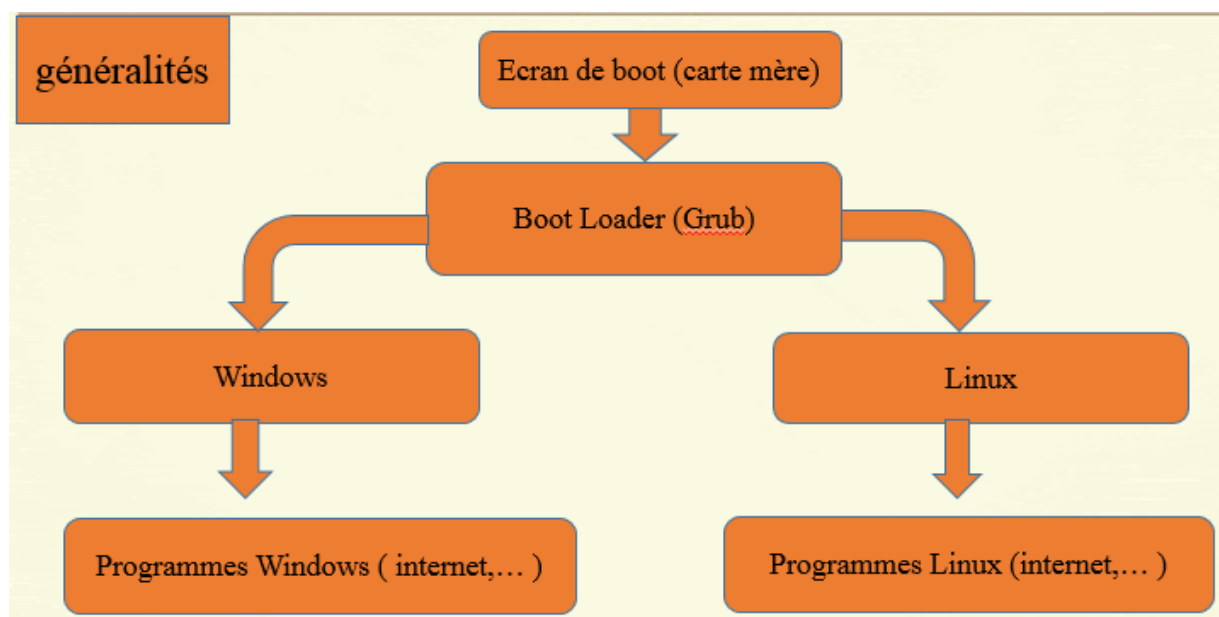
2 - copier/déplacer/ des fichiers

- ❖ Permettre aux programmes de fonctionner quel que soit le matériel.

Jouer à un jeu vidéo quelle que soit la carte vidéo et sa performance, avec plus ou moins d'options. Pour plus des détails :

L'OS (Operating System) gère

- ❖ La mémoire : il la partage entre tous les programmes
- ❖ Les périphériques : écran, imprimante, disque dur, réseau. Il s'assure que les programmes puissent les utiliser de façon standard.
- ❖ Le processeur : il le partage entre tous les programmes pour qu'ils aient l'air de fonctionner parallèlement
- ❖ Les utilisateurs : gérer les droits d'accès aux fichiers, comme au matériel
- ❖ La standardisation des programmes : offre des interfaces de programmation simplifiées et standardisées.



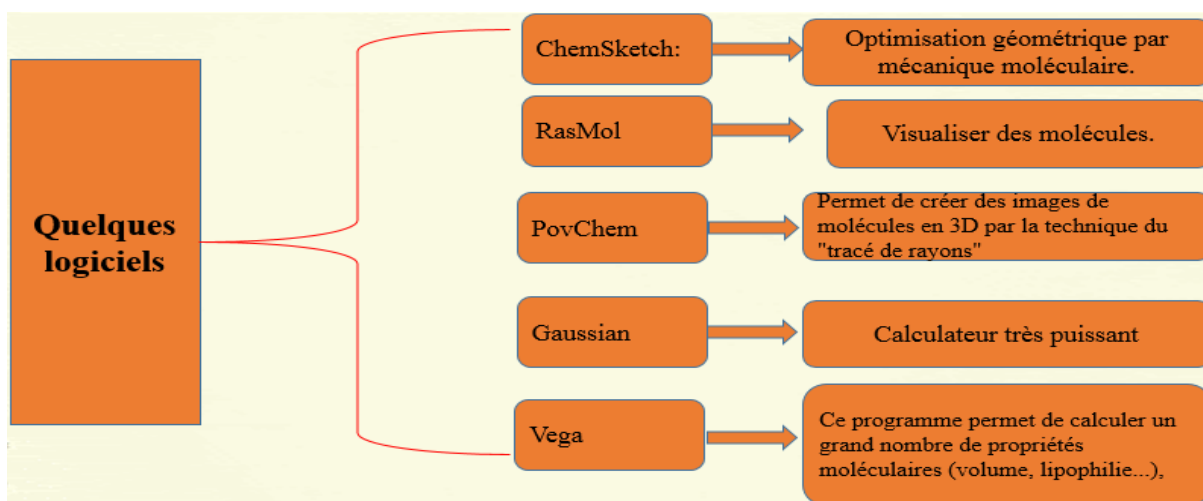


Fig 4 : quelques logiciels et leurs fonctions

Le système Unix est un système d'exploitation multi-utilisateur et multi-tâche

✓ **Unix est multi-utilisateurs**

Multi-User

Plusieurs utilisateurs sous Unix. Chacun dispose de l'ensemble des ressources du système. Comme tout système multi-utilisateur, Unix comporte des mécanismes d'identification et de protection permettant d'éviter toute interférence entre utilisateurs.

Deux types de Users

1- Users normaux : compte avec

- ✓ Login
- ✓ Password
- ✓ Espace de travail protégé
- ✓ Mail

2 - Super-User root gère tout le système

Unix est multi-tâche

Unix est multi-tâche car plusieurs programmes peuvent être en cours d'exécution en même temps sur une même machine.

Un processus est une tâche en train de s'exécuter. On appelle processus, l'image de l'état du processeur et de la mémoire au cours de l'exécution du programme. En fait, à chaque instant, le processeur ne traite qu'au plus un seul des programmes lancés. La gestion des processus est effectuée par le système.

On essayera d'énumérer les principales étapes de l'analyse des données qualitatives ainsi que les différentes techniques de traitement associées à ces données, notamment les traitements sémantiques et statistiques.

-Traitement statistique et graphique de données.

III. Traitement statistique et graphique de données.

Beaucoup de **connaissances chimiques ont été dérivées des données**. La chimie doit offrir une gamme riche des données sur les propriétés physiques, chimiques, et biologiques, par exemple, données binaires pour la classification, vraies données pour la modélisation, et données spectrales ayant une densité élevée de l'information. Ces données doivent être introduites dans une forme favorable à l'échange d'information facile et analyse de données.

III. 1. Les sources de données et les bases de données

L'énorme quantité de données en chimie a mené au développement des bases de données pour stocker et disséminer les données en forme électronique. Par exemple, des bases de données ont été développées pour la littérature chimique, composés chimiques, structures 3D, réactions, et spectres. L'Internet est de plus en plus employé pour distribuer des données et l'information en chimie.

III. 2. Méthodes pour calculer des données physiques et chimiques

Une variété de données physiques et chimiques des composés peut directement être calculée par une gamme des méthodes. Les premiers sont les calculs mécaniques de quantum de divers degrés de sophistication. Cependant, des méthodes simples telles que des arrangements d'additivité peuvent également être employées pour estimer une variété de données avec l'exactitude raisonnable.

IV. Calcul des descripteurs de structure

Dans la plupart des cas, cependant, physique, chimique, ou les propriétés biologiques ne peuvent pas être directement calculés à partir de la structure d'un composé. Dans cette situation, une approche indirecte doit être adoptée : la structure du composé est représentée par des descripteurs de structure par la suite un rapport entre les descripteurs de structure et la propriété est établie en analysant une série de paires de descripteurs de structure et les propriétés associées par les méthodes d'étude inductives

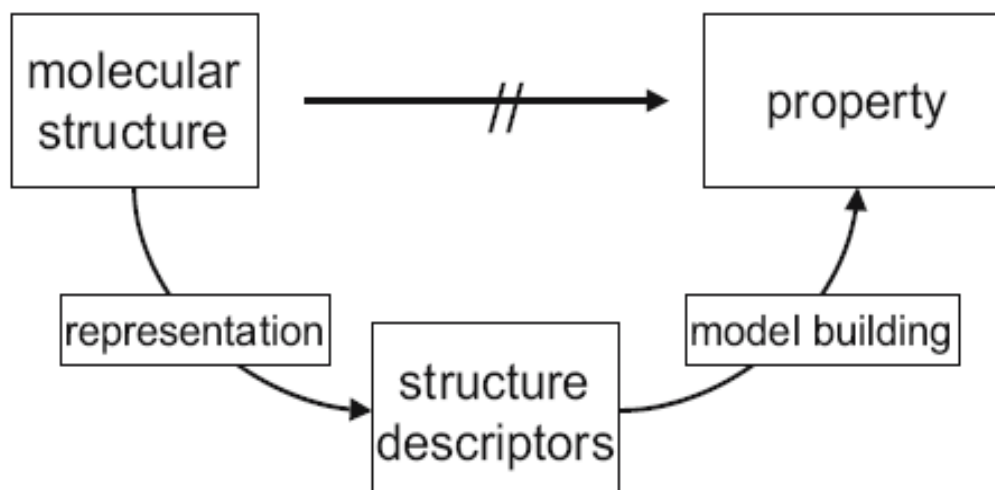


Fig : Établissement des relations de structure-propriété/activité.

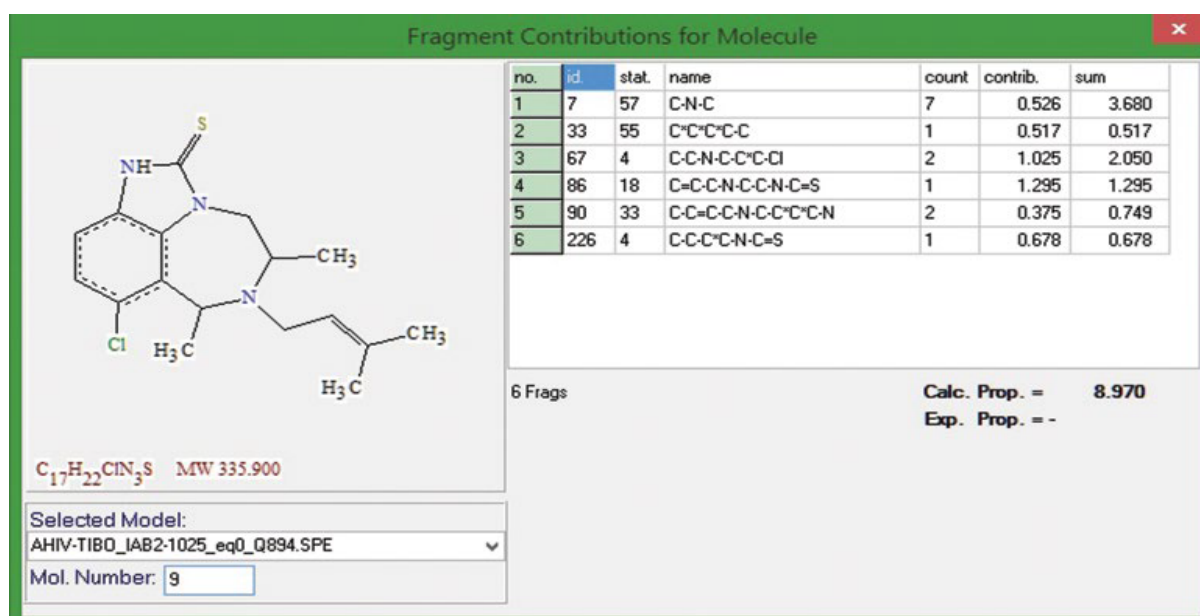


Fig 5 : Rapport entre les descripteurs de structure et la propriété

Une variété de descripteurs de structure ont été développées 1D, 2D, ou information de la structure 3D ou les propriétés de la surface moléculaires.

V. Méthodes d'analyse de données

Une variété de méthodes pour apprendre des données par des méthodes d'étude inductives sont employées dans la chimie, par exemple, statistiques, méthodes d'identification de modèle, réseaux neurones artificiels, et algorithmes génétiques. Ces méthodes peuvent être classifiées dans des méthodes d'apprentissage supervisé et non supervisé et sont employés pour la classification ou modélisation quantitatif.

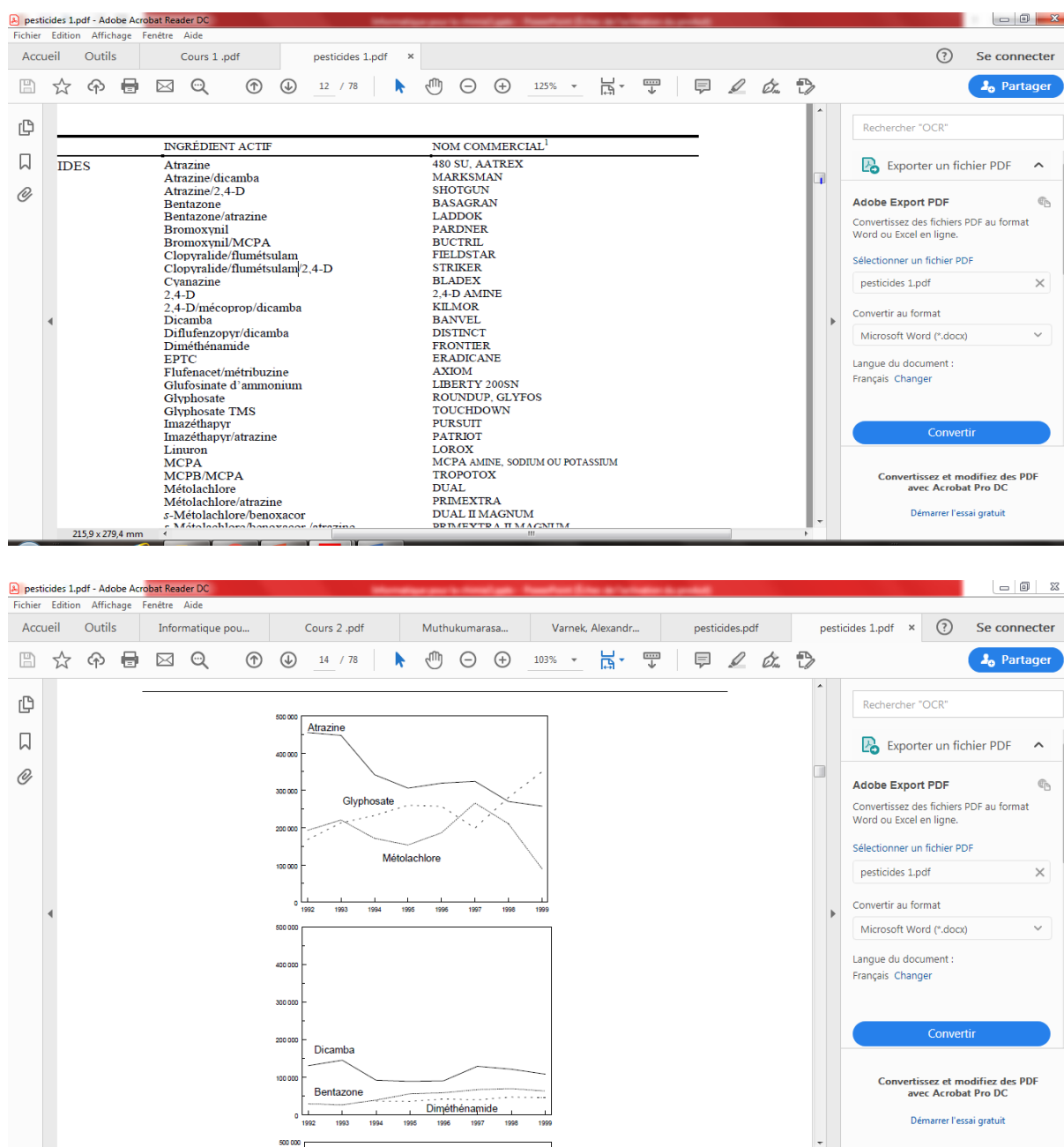


Fig6 : méthodes pour apprendre des données

V.1. Construction de bases de données de molécules et de réactions

Une base de données chimique est un stock d'informations organisé et structuré de manière à pouvoir être facilement manipulé. La plupart des bases de données chimiques stockent des informations sur des molécules stables.

Les grandes bases de données chimiques devraient être capables d'assurer le stockage et la recherche d'informations sur des millions de molécules sur des téraoctets de mémoire physique (terabytes of physical memory).

Exemples de banques de données :

- ✓ **CHIRBASE** : banque de données pour la séparation chromatographique chirale
- ✓ **THERMODATA** : Banques de données experts pour la thermochimie
- ✓ **SXD** : Extraction liquide-liquide des métaux

V.2. Etude de banques de données chimiques indexées par structure

Cette étude est très importante dans les domaines de bio-informatique et de recherche en chimie médicinale et pharmaceutique. La nomenclature de cette étude est composée en deux parties :

➤ Banques de données chimiques

Ces bases contenant des informations sur les composés chimiques, particulièrement leurs structures (2D et 3D), propriétés, spectres, activités biologiques, etc.

➤ Indexées par structure

Ce terme indique l'organisation des données d'une façon qui facilite la recherche basant sur la similarité ou l'identité exacte des structures et sous-structures.

Cette étude vise plusieurs objectifs tels que l'identification des composés similaires à un composé, la recherche des analogues de structure d'un composé, l'analyse des relations structure-activité (SAR), obtenir les propriétés prédites ou expérimentales d'une structure.

V.2.1. Principe

Le principe de l'indexation par structure est basé sur la représentation de la molécule par SMILES ou par des formats structuraux tels que Molfile, SDF ect. La deuxième étape consiste à effectuer une des trois recherches, soit la recherche de la structure afin de prédire les molécules identiques (Exact match), soit, la détection d'un motif au sein d'une molécule de poids moléculaire plus important (Substructure search) ou la recherche des molécules qui se ressemblent à une structure déterminée (Similarity search). La troisième étape de ce principe concerne la visualisation des structures obtenus utilisant différents logiciels tels que ChemDraw, MarvinSketch...

V.2.2. Domaine d'utilisation

Quelques exemples de bases de données sont illustrés dans le tableau suivant :

Tableau 2 : Quelques exemples de bases de données et leur domaine d'utilisation.

Base de données	Domaine d'utilisation	Spécification
PubChem	Chimie, Bioinformatique	Plus que 100 millions de composés
ChemSpider	Chimie, Bioinformatique	Intègre plusieurs bases chimiques

ChEMBL	Biologie, Chimie médicinale et Pharmaceutique	Activité des Biomolécules
Reaxys	Chimie	Riche en réactions et structures chimiques
Protein Data Bank (PDB)	Chimie médicinale et Pharmaceutique	Données de structures 3D protéines- ligands
DrugBank	Chimie médicinale et Pharmaceutique	Données sur médicaments et molécules candidates

-Méthodologie de la recherche d'informations en Chimie

VI. Méthodologie de la recherche d'information en chimie

Il s'agit de l'ensemble des étapes permettant de chercher, identifier et trouver des documents relatifs à un sujet par l'élaboration d'une stratégie de recherche. Dans un contexte de surabondance d'informations et de diversité de ses supports, les processus de recherche documentaire et de validation de l'information requièrent la mise en application d'une méthodologie efficace.

VI.1. Une ou des méthodes pour la recherche documentaire ?

Il existe différentes méthodes permettant de définir une bonne stratégie de recherche en chimie. La finalité d'une bonne méthodologie de recherche est de faciliter la production d'un travail universitaire alliant richesse documentaire, celles-ci reposent néanmoins toujours sur quelques principes clés, que nous allons voir successivement. et rigueur scientifique. Elle va s'articuler autour de 5 étapes successives

- 1 - Préparer sa recherche (domaine de chimie analytique, chimie organique, électrochimie, chimie de l'eau...
- 2 - Sélectionner les sources d'information (article scientifique, résultats des expériences ...).
- 3 - Chercher et localiser les documents.
- 4 - Évaluer la qualité et la pertinence des sources.
- 5 - Mettre en place une veille documentaire.

VI.2. Objectifs de la recherche la recherche d'information en chimie

Les objectifs de cette recherche peuvent être réalisés en répondant aux quelques questions préalables, qui sont réunies dans le tableau suivant :

Tableau 3 : Quelques questions préalables de la recherche d'information en chimie

Question	Réponse
Quelle est la nature du travail à produire ?	Exposé, mémoire, article, rapport d'étude, etc.
Quel est le niveau attendu de l'information ?	- Information de base : un ou deux ouvrages et une encyclopédie suffiront. - Information plus pointue ou plus développée : Se diriger vers des articles de périodiques, conférences, rapports, thèses, normes...
Quelle est la nature de l'information recherchée ?	Développements sur un sujet, données statistiques, schémas techniques, etc.
Quel est le degré d'actualité de l'information recherchée ?	Des documents d'archives aux derniers résultats connus de la recherche...

VI.3. Préparer une recherche

Il s'agit de questionner le sujet dans toutes ses dimensions, de le formuler en une phrase courte, de sélectionner les concepts importants et de chercher des synonymes. Cette étape doit permettre de poser la problématique, de cerner les besoins documentaires et de sélectionner les concepts/mots clés nécessaires à l'interrogation des sources documentaires.

VI.3.1. Définir le sujet

a. Choix du sujet

Vos connaissances préalables et vos intérêts personnels sur le sujet ainsi que sa pertinence par rapport à l'enseignement auquel il se rattache sont des éléments fondamentaux qui doivent guider votre choix du sujet

b. Première approche documentaire

A ce stade, la documentation utilisée, la question qui se pose est quels types de documents ?

Ace stade, la documentation utilisée concernera essentiellement des dictionnaires et encyclopédies (support papier ou en ligne), des manuels ou des articles de synthèse.

c. Cerner le sujet

Au cours de cette étape, on peut appliquer la méthode QQQOCP. Il s'agit d'un moyen mnémotechnique permettant de retenir un ensemble de questions simples qui vont être utilisées pour cerner, préciser et approfondir un sujet :

- * **Qui ?** = Quels sont les chercheurs, les personnes impliquées ?
- * **Quoi ?** = Quels sont les aspects qui m'intéressent ?
- * **Quand ?** = Quelle est la période concernée ?
- * **Où ?** = Le sujet est-il circonscrit à un domaine précis ?
- * **Comment ?** = Quelles approches ou points de vue faut-il considérer ? (Environnementaux, chimie de l'environnement, socioéconomique, etc.)
- * **Pourquoi ?** = Quelle est l'importance du sujet dans le contexte actuel

d. Formuler le sujet

Afin de formuler un sujet avec succès, certaines instructions doivent être suivies

✓ Une phrase courte

Le sujet doit être exprimé en une phrase courte, si possible sous forme de question et à l'aide de termes significatifs. Cet énoncé de recherche doit être le plus précis possible

✓ Des concepts clés

Chaque terme de l'énoncé est important et va correspondre à des concepts/mots clés qui vont servir à élaborer les équations de recherche.

e. Restreindre ou élargir le sujet

A l'issue de cette étape, deux types de difficultés peuvent se présenter :

1- Difficulté n° 1

Le sujet est trop général, trop vaste, risques majeurs, surabondance de la documentation, traiter le sujet de manière superficielle...

Que faire ?

Restreindre le sujet en se concentrant sur un aspect de la question, en délimitant une zone géographique ou une période donnée. Ce choix ne sera pas sanctionné s'il est argumenté.

2 - Difficulté n° 2 Le sujet est :

- Trop précis, trop pointu.
- Très ardu à traiter
- Difficulté à repérer la documentation
- Nécessite d'interroger de nombreuses sources

Que faire ?

Mettre le sujet en perspective. Replacer le sujet dans un contexte plus large vous permettra d'aborder des aspects de la question auxquels vous n'auriez pas pensé dans un premier temps et ainsi d'élargir sa portée.

VI.3.2. Sélectionner les sources d'information

Après avoir analysé et délimité le sujet, il faut choisir les meilleures sources d'information pour effectuer la recherche. Cette démarche comprend deux dimensions :

1 - Le type de documents que l'on recherche : monographies, articles de revues, thèses, etc.

2 - Le type de ressources à interroger : catalogues de bibliothèque, bases de données, moteur de recherche du Web, portails spécialisés, etc.

Le type de documentaire Il dépend du niveau et de la nature de l'information recherchée

VI.3.3. Evaluer la qualité et la pertinence des sources

Les principaux critères de fiabilité des sources sont comme suit :

- ✓ L'auteur, l'éditeur de la ressource
- ✓ La date de publication du document
- ✓ Le domaine de la ressource (adresse URL)
- ✓ La notoriété, l'indice de popularité du site
- ✓ Le contenu de l'information (structuration, argumentation, sources, etc.)

- **Représentation de la structure 3D et applications locales**

VII. Représentation de la structure 3D et applications locales

VII. 1. Représentation de la structure 3D

Cette représentation désigne la façon d'illustration spatiale d'une molécule, une protéine ou un solide afin d'illustrer l'arrangement des atomes dans l'espace. Les logiciels et les bases de données communément employés pour avoir cette représentation sont regroupés dans le tableau suivant :

Tableau 4 : Quelques logiciels et Base de données employés la représentation 3D

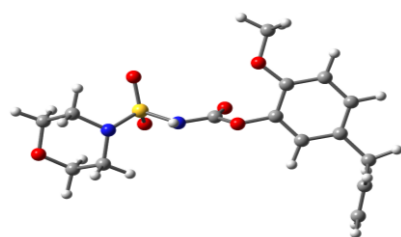
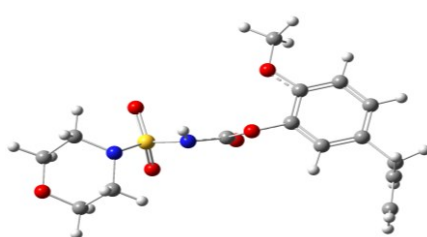
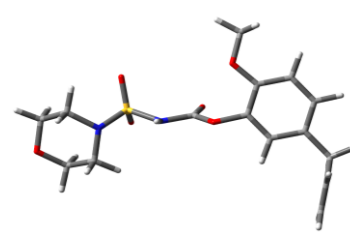
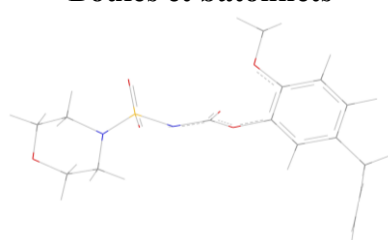
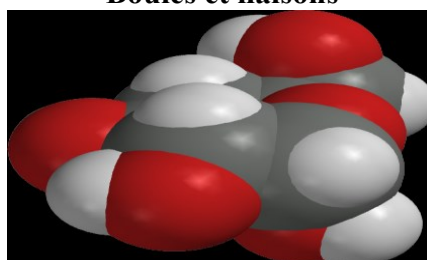
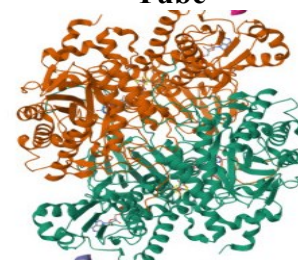
	Logiciels ou BDs	Niveau	
Logiciels	Chem3D	Débutant	Molécule
	Avogadro	Débutant	
	GaussView	Avancé	
	PyMOL	Intermédiaire	
	Chimera	Intermédiaire	
Bases de données	Protein Data Bank (PDB)	/	Protéines/ADN
	PubChem		Petites molécules

Il existe plusieurs types de représentations 3D, les plus importants sont cités dans le tableau 5 et la figure suivante.

Tableau 5 : Essentiels types des représentations 3D

Représentations 3D	Description
Bâtonnets (Stick model)	Les atomes sont reliés par des segments. Elle prend en considération les liaisons et la connectivité
Boules et bâtonnets (Ball and stick)	Les atomes sont représentés par des sphères (en fonction de leur taille) et les liaisons par des tiges.
Boules et liaisons (Ball and bond)	Les atomes sont représentés par des sphères (en fonction de leur taille) et les liaisons par des liaisons.
Tube (Cartoon/Backbone)	Dans cette représentation chaque atome n'est pas montré individuellement et la chaîne principale (backbone) est représentée sous forme de tube lisse et continu.
Surface de Van der Waals (CPK)	Les atomes sont représentés par des sphères à (en fonction de leur rayon de van der Waals). Elle assure la visualisation du volume et de surface d'une molécule.
Représentation en rubans (Ribbon)	Destinée aux protéines/ADN. Elle montre les structures secondaires

Représentation en fil de fer (wireframe)	Chaque liaison symbolisée par une ligne. Elle est simple, adaptée aux grandes molécules
---	---

**Boules et bâtonnets****Boules et liaisons****Tube****Représentation en fil de fer****Représentation (CPK)****Représentation en rubans****Fig 7 : Quelques types des représentations 3D**

VII. 2. Applications locales de la représentation 3D

Il s'agit des logiciels qui donnent la possibilité de dessiner une représentation 3D de la structure d'une molécule. Voici quelques informations sur les logiciels utilisés qui sont déjà cités dans le tableau 4.

1 - Avogadro

Ce logiciel peut être fonctionné avec Windows et Linux. Il peut dessiner les molécules, optimiser leur géométrie et les exporter en 3D, il est très simple pour la construction et la visualisation rapide des molécules.

2 - Chem3D

Il est inclus dans le ChemOffice, peut être fonctionné seulement avec Windows, ce logiciel est largement utilisé en chimie organique. Il permet de visualiser les molécules ainsi que calculer leur géométrie.

3 - GaussView

Ce logiciel est utilisé avec Gaussian, il peut être fonctionné avec Windows et Linux. Il est utilisé pour la modélisation moléculaire avancée notamment pour la préparation et l'analyse des calculs quantiques.

4 - PyMOL

Ce programme peut être fonctionné avec Windows et Linux. Il est destiné pour les protéines/ADN ainsi que les petites molécules. La représentation des molécules est sous forme de rubans, boules et bâtonnets.

5 - Chimera

Ce programme peut être fonctionné avec Windows et Linux. Il est très puissant pour les structures de protéines et de complexes biomoléculaires.

- **Initiation à la modélisation moléculaire.**

VIII. Initiation à la modélisation moléculaire

La modélisation moléculaire est le point de rencontre de la chimie, de la biologie et de l'informatique. C'est l'ensemble des méthodes informatiques employées afin de visualiser l'arrangement des molécules en 3D. Elle est employée en Chimie organique dans le but d'optimisation des structures des molécules et prédire leur réactivité, en Biochimie et Biologie structurale (Bioinformatique) afin d'étudier les protéines, l'ADN et les complexes ligand-récepteur, ainsi qu'en Pharmacologie (Chimie médicinale) pour le Drug design et le Docking moléculaire. Ces méthodes ont l'objectif également de simuler les comportements des structures moléculaires tels que les mouvements et les interactions.

VIII.1. Types de modélisation

VIII.1.1. Mécanique moléculaire (MM)

Ces méthodes utilisées pour une approche rapide comme la pré-optimisation, elles reposent sur la mécanique classique. Ces techniques ont pour objectifs l'optimisation rapide des structures, l'étude de la conformation (torsions, flexibilité) et l'étude du mouvement des molécules dans le temps (la dynamique moléculaire), parmi ces méthodes on peut citer MMFF94, AMBER, CHARMM et OPLS.

VIII.1.2. Méthodes quantiques

Ces méthodes utilisées sont fondées sur la mécanique quantique, Ces calculs nécessitent beaucoup de temps, Ils visent à calculer les orbitales moléculaires (HOMO/LUMO), prédire les énergies de transition, spectres théoriques (IR, UV et RMN), la réactivité chimique et les mécanismes réactionnels. Les plus fréquents méthodes quantiques employés sont les méthodes Ab initio et Hartree qui sont très précises mais coûteuses et la méthode Density Functional Theory (DFT) qui combine la précision et le coût. Parmi les logiciels qui utilise ces méthodes on peut citer Gaussian, Orca, etc.

VIII.1.3. Docking moléculaire

Cette approche serve à prédire les interactions entre une petite molécule qui est généralement le ligand (médicament) et une macromolécule qui est la cible biologique (protéine ou ADN). Elle a pour objectif l'identification des composés biologiquement actifs.

VIII.1.4. Dynamique moléculaire (MD)

C'est une simulation de l'évolution d'un système moléculaire dans le temps (nanosecondes → microsecondes). Elle consiste à étudier la stabilité des protéines et les interactions ligand-récepteur.

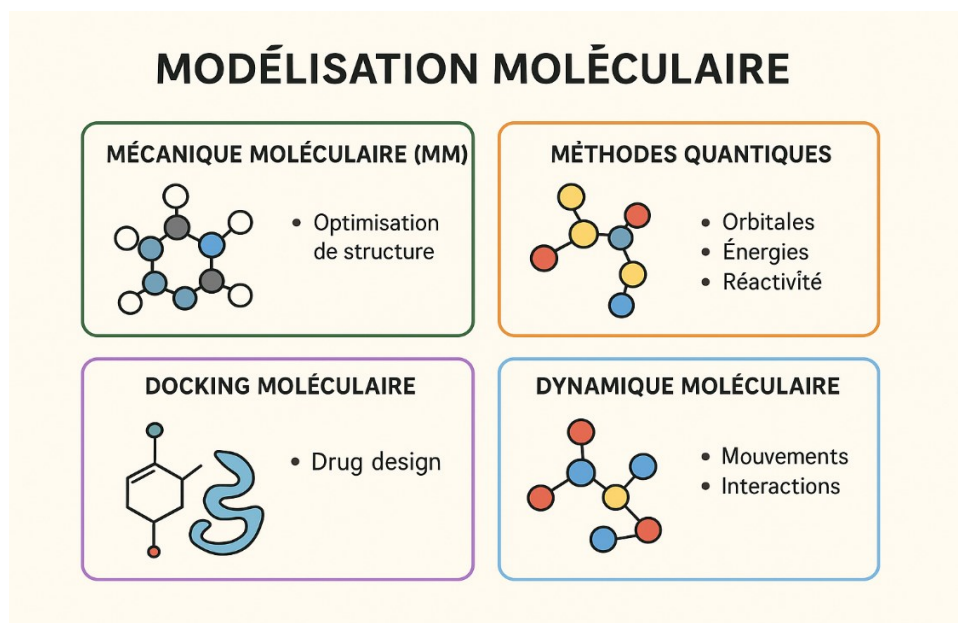


Fig 7 : Types de modélisation et leurs applications.

VIII.2. Les étapes de la modélisation moléculaire

VIII.2.1. Construction de la molécule

À partir de la structure plane (2D) de la molécule, les atomes de cette dernière doivent être dessinés et ses liaisons définies à partir d'un logiciel.

VIII.2.2. Optimisation de la géométrie

Il s'agit de l'arrangement automatique des longueurs de liaisons et angles afin de se rapprocher à la structure réelle de la molécule. Les méthodes couramment employés dans cette étape sont les méthodes de la mécanique moléculaire (AMBER, MM⁺) et celles de la mécanique quantiques (DFT, Hartree-Fock)

VIII.2.3. Visualisation 3D

Au cours de cette visualisation, différentes représentations sont disponibles telles que les boules et bâtonnets, rubans pour protéines etc. (Voir tableau. 5).

VIII.2.4. Analyse et simulations

Cette étape est destinée pour deux objectifs, le calcul des énergies, les surfaces de densité électronique, les interactions ligand-récepteur, ainsi que la prédiction de l'évolution dans le temps à l'aide des simulations de dynamique moléculaire.

VIII.3. Exemple pratique (avec Gaussian)

1 – Construction de la molécule

Dessiner la molécule d'éthanol dans Chemdraw

2 – Optimisation de la géométrie

A) Optimisation par la mécanique classique

- Ouvrir le Chem 3D
- Optimiser la molécule par la méthode MM⁺

B) Optimisation par la mécanique quantique

- Ouvrir le logiciel Gaussian.
- Changer la vue de la molécule en Ball and Stick, Ball and Bond Type ou Tube.
- Optimiser la molécule par la méthode DFT.

3 - Analyse

Discuter les résultats trouver par calcul DFT (charges, orbitales, énergie).

REFERENCE BIBLIOGRAPHIQUE

1. Benard VIDAL Chimie quantique de l'atome à la théorie de Huckel. Masson. 1993.
2. Claude Leforestier. Introduction à la chimie quantique cours et exercices corrigés. DUNOD 2005.
3. P. Chaquin. Pratique de la chimie theorique. LCT-UPMC
4. Jean-Louis Rivail. Elements de chimie quantique à l'usage des chimists. CNRS editions. 1999.
5. David C. Young Cytoclonal Pharmaceuticals Inc. Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems Copyright. John Wiley & Sons, Inc. 2001
6. Bernard Silvi et Najib Fourati Molecular Physics: An International Journal at the Interface Between Chemistry and Physics 1984, 52: 415-430.
7. James J. P. Stewart J Mol Model. 2007, 13:1173–1213
8. T. Tung Nguyen-Dang Chimie quantique, Université Laval, Quebec A-2005
9. Koch, W.; Holthausen, M.C. in A Chemist's guide to Density Functional Theory, Oxford University Press, New-York, 1989.
10. Jensen, F. Introduction to computational chemistry, Ed. J. Wiley & Sons Ltd. 1999