



HAL
open science

Analyse des systèmes de visualisation et d'interaction 3D pour la biologie moléculaire

Mouna Essabbah, Samir Otmane, Guillaume Bouyer, Joan Hérisson, Malik
Malle

► To cite this version:

Mouna Essabbah, Samir Otmane, Guillaume Bouyer, Joan Hérisson, Malik Malle. Analyse des systèmes de visualisation et d'interaction 3D pour la biologie moléculaire. *Revue des Sciences et Technologies de l'Information - Série TSI: Technique et Science Informatiques*, 2012, 31 (2), pp.187–214. 10.3166/tsi.31.187-214 . hal-00633364

HAL Id: hal-00633364

<https://hal.science/hal-00633364v1>

Submitted on 30 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse des systèmes de visualisation et d'interaction 3D pour la biologie moléculaire

Mouna Essabbah* — **Samir Otmane*** — **Guillaume Bouyer*** — **Joan Hérisson**** — **Malik Mallem***

* *Laboratoire IBISC (CNRS - EA 4526)*
40 rue du Pelvoux, 91020 - Evry Courcouronnes, France
mouna.essabbah@ibisc.fr

** *iSSB, EA4527*
Genopole® Campus 1 - Bât. Genavenir 6, 5 rue Henri Desbruères - 91030 Evry
Joan.Herisson@issb.genopole.fr

RÉSUMÉ. En biologie, des études montrent que derrière la réalité fonctionnelle des molécules se cache une architecture spatiale. Ainsi, pour comprendre la corrélation entre l'aspect architectural et fonctionnel de la molécule il est nécessaire de comprendre les raisons et les mécanismes mis en place. Toutefois, la plupart des recherches menées demandent un ensemble d'expérimentations qui s'avèrent coûteuses en temps et en moyens financiers. Les simulations numériques ont montré leur intérêt dans ce domaine. Ainsi, un grand nombre de systèmes d'analyse moléculaire ont vu le jour. Certains sont dédiés uniquement à la visualisation 3D et d'autres à la manipulation des molécules, bénéficiant des techniques d'interaction 3D et des dispositifs de la réalité virtuelle.

ABSTRACT. In biology, studies have proven that behind the functional reality molecules there is a spatial architecture. Thus, to understand the correlation between the functional and architectural aspect of the molecule it is necessary to understand the reasons and mechanisms implemented. However, most of the research carried out required series of experiments that are costly in time and financial resources. Rapidly, computer simulations have shown their interest in this domain. Therefore, a large number of molecular analysis systems have emerged. Some of them are dedicated to 3D visualization while others allow the molecules' manipulation, benefiting from 3D interaction techniques and Virtual Reality devices.

MOTS-CLÉS : Modélisation moléculaire, analyse moléculaire, réalité virtuelle.

KEYWORDS: Molecular modeling, molecular analysis, Virtual Reality.

1. Introduction

Afin de mieux comprendre les phénomènes biologiques et de les interpréter, il est indispensable d'avoir un minimum de connaissances sur leurs comportements *in vivo*¹. Toutefois, il est très difficile, voire impossible, d'étudier les molécules *in vitro*² par des moyens directs tels que l'observation microscopique. Cela l'est encore moins *in vivo*, sans extraire les molécules de leurs environnements.

Il faut donc trouver une solution qui s'affranchit le plus possible des techniques expérimentales *in vivo* et *in vitro*. Pour cela, il est possible de mettre en place des moyens de modélisation approximative, qui permettraient d'avoir la meilleure adéquation avec les résultats expérimentaux. L'expérimentation *in silico*³ est un moyen d'effectuer des simulations numériques afin de modéliser les molécules. Le plus souvent, le processus de modélisation comprend deux étapes distinctes : l'acquisition de données expérimentales et la génération du modèle.

Les premiers modèles qui ont émergé étaient des modèles matériels, tels que ceux assemblés à partir de boules et bâtonnets. Ils ont été utilisés pendant plus d'un siècle pour représenter les règles de la Chimie d'une manière très simple (eg. la double hélice de l'ADN en fil de fer de Watson et Crick (Watson *et al.*, 1953)). À présent la recherche s'intéresse plus aux modèles informatiques et aux simulations conduisant à la conception d'algorithmes d'analyse de molécules. La modélisation moléculaire est devenue une discipline scientifique indépendante adoptant un ensemble de techniques pour modéliser et imiter le comportement des molécules (modèles mathématiques, modèles 2D, modèles 3D, etc.).

En particulier, la modélisation 3D moléculaire consiste à reproduire numériquement la structure tridimensionnelle des molécules. Un grand nombre de systèmes de modélisation 3D, de plus en plus précis, ont vu le jour, offrant des outils performants pour l'analyse des structures voire de la dynamique moléculaire⁴. Parmi ces différents systèmes de modélisation 3D, certains sont dédiés uniquement à la visualisation 3D des molécules, alors que d'autres sont plus axés sur la manière d'interagir avec les modèles 3D afin de pouvoir les modifier et notamment à travers les techniques d'interaction et les dispositifs de réalité virtuelle. Ceci place l'utilisateur au centre des expérimentations virtuelles, on parle d'analyse *in virtuo*.

Cet article présente un aperçu des systèmes d'analyse moléculaire. Il est divisé en six sections : nous commencerons par présenter quelques définitions. Nous introduirons ensuite le principe de l'analyse en biologie moléculaire ainsi que la solution *in virtuo* de cette analyse. Nous présenterons alors les trois composantes de l'analyse *in virtuo*, en commençant par la modélisation 3D et les différentes approches que nous avons recensés. Par la suite, nous décrirons quelques systèmes dédiés que nous diviserons selon les deux composantes visualisation et interaction 3D. Nous présenterons les ré-

1. Au sein du vivant

2. Test en tube, en dehors de l'organisme vivant

3. Test effectués au moyen d'outils informatiques

4. La dynamique moléculaire consiste à étudier l'évolution de l'ensemble des atomes d'un système moléculaire au cours du temps, grâce à des simulations numériques

sultats de l'étude de ces systèmes sous forme de discussion ce qui nous amènera à proposer une nouvelle approche d'analyse moléculaire.

2. Quelques définitions

Il est nécessaire de donner quelques brèves définitions de certaines notions qui seront utilisées, notamment la définition de la BM, de la réalité virtuelle, ainsi que de l'interaction 3D.

2.1. La biologie moléculaire

La Biologie Moléculaire (BM) est un sous-domaine de la Biologie. Il s'agit de l'étude des mécanismes employés dans l'organisation et le fonctionnement des cellules au niveau moléculaire. Elle est souvent utilisée pour modéliser ou "mimer" le comportement des macromolécules biologiques. Son débouché industriel majeur est la mise au point de nouveaux médicaments. Cette sous-discipline de la Biologie traite de l'ADN (support de l'information génétique) et de la synthèse des protéines (ouvriers et briques de la cellule) chez les organismes vivants. Plus particulièrement, la BM s'apparente souvent aux techniques utilisées pour la manipulation de l'ADN et de l'ARN ainsi qu'à l'étude de leurs structures, synthèses et altérations (ie. mutations). Elle étudie les mécanismes biologiques, en fonction des structures des constituants cellulaires et les interactions existantes entre ceux-ci.

2.2. La réalité virtuelle

La RV définie par une majorité d'acteurs du domaine est "un domaine scientifique et technique exploitant l'informatique et des interfaces comportementales en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en interaction en temps réel entre elles et avec un ou des utilisateurs en immersion pseudo naturelle par l'intermédiaire de canaux sensori-moteurs." (Fuchs *et al.*, 2006b). La multimodalité permet d'appréhender la RV par nos sens grâce aux différents canaux sensori-moteurs (l'audio, la vidéo et l'haptique). La RV, contrairement à la simulation qui représente des phénomènes physiques réels, permet d'aller au delà du réel (étude de la structure d'une molécule ou de la radioactivité, la collaboration à distance, etc.). L'immersion pseudo naturelle doit apporter une intégration, la plus efficace possible, de l'utilisateur dans le monde virtuel. Cette sensation est une notion en partie subjective qui dépend de l'application et du dispositif utilisé (interfaces, logiciels, etc.)(Fuchs *et al.*, 2006a).

2.3. L'interaction 3D

L'interaction 3D est un aspect essentiel de la RV car elle représente le pouvoir d'action de l'utilisateur dans le monde virtuel. Elle s'effectue par l'intermédiaire de périphériques appelés Interfaces Homme-Machine, qui peuvent être sensorielles ou motrices, complétées par des outils logiciels appelés techniques d'interactions.

Selon Bowman, il existe trois types d'actions naturelles définies comme la navigation (se déplacer dans l'espace virtuel), la sélection (désigner un objet) et la manipulation des objets (agir dans cet espace). Enfin, un quatrième type de tâche, plus spécifique à l'informatique, est le contrôle d'applications (communication avec autrui ou avec le système) (Bowman, 1999). Un récent bilan de l'état actuel du domaine de l'interaction 3D en RV a été établi dans (Ouramdane *et al.*, 2009).

3. Problématique : L'analyse en biologie moléculaire

L'analyse en BM consiste à observer des systèmes biologiques et à examiner des résultats expérimentaux dans le but de comprendre des phénomènes du Vivant. Les outils d'analyse en BM permettent, pour la plupart, la combinaison de la visualisation des modèles moléculaires et de l'interaction limitée au changement de point de vue du modèle, ou de son mode de représentation. Cette phase d'analyse peut être assimilée à une phase d'observation du modèle moléculaire. Dans le cas particulier des modèles 3D, l'analyse moléculaire est apparentée à la tâche de navigation autour d'une molécule virtuelle (ou dans un environnement virtuel).

L'infographie moléculaire est née dans les années 60 pour visualiser les structures résolues à cette époque (Levinthal, 1966). La modélisation moléculaire graphique permet la synthèse et ensuite la visualisation de structures moléculaires ainsi que leurs propriétés. Elle offre la possibilité d'afficher des structures dans une grande variété de styles. Il existe différents modes de représentation graphique des structures des macromolécules biologiques. Ils permettent dans chaque cas de mettre en valeur une information demandée. Par exemple, un mode de représentation filaire⁵ va permettre d'observer le squelette de la macromolécule, tandis qu'une représentation surfacique informe sur l'aspect extérieur de celle-ci.

En outre, elle offre la capacité de déplacer les structures de manière interactive en trois dimensions. La partie graphique de la modélisation du système peut également fournir une interface pour les outils de chimie quantique, permettant de définir et d'exécuter des calculs et enfin de les analyser lorsqu'ils seront terminés.

Ensuite, avec l'apparition de la génomique⁶, la BM ne traite plus une vision statique et locale des molécules mais plutôt une vision dynamique et globale (Dardel *et al.*, 2002). Parallèlement, la RV s'est développée particulièrement dans des applications militaires. Vers les années 90 la RV a gagné progressivement les domaines scienti-

5. marque les liaisons entre les atomes de la molécule par des cylindres de faible diamètre (segments)

6. La génomique est l'étude du génome, la disposition, la séquence et la fonction de l'ensemble des gènes

figes ((Brooks *et al.*, 1990)(Leech *et al.*, 1996)(Humphrey *et al.*, 1996)(Levine *et al.*, 1997)(Loftin *et al.*, 1998)(Anderson *et al.*, 1999)). De plus, les récents progrès des technologies de visualisation et la puissance croissante des ordinateurs ont permis d'envisager l'utilisation de la RV afin améliorer la perception des résultats de simulation. Cependant, la modélisation et la visualisation de structures spatiales moléculaires sont limitées par certaines contraintes (taille des molécules, espace de recherche, etc.). La difficulté est donc de proposer des moyens et des outils informatiques qui pourraient aider à l'analyse moléculaire. Nous pouvons recenser deux problématiques :

- D'une part, la modélisation moléculaire s'est développée rapidement bénéficiant de l'évolution du matériel et du logiciel informatique. Ainsi, les exigences des biologistes ont grandi peu à peu, et ont tendance à rajouter de plus en plus de contraintes à leurs systèmes de modélisation en essayant de simuler le plus fidèlement possible la structure spatiale et le comportement des molécules. La modélisation moléculaire 3D se doit d'être de plus en plus précise, mais surtout avec un niveau de réalisme ou de crédibilité plus élevé.

- D'autre part, les scientifiques réclament des systèmes qui permettent une analyse plus approfondie des modèles moléculaires 3D proposés. Le rôle de l'aide à l'analyse prend alors toute son ampleur. L'évolution des systèmes d'analyse moléculaire doit accompagner celle de la modélisation 3D. Ces systèmes devraient permettre à l'utilisateur un certain nombre de manipulations sur le modèle moléculaire construit, offrant ainsi une analyse avancée tout en respectant les contraintes (physiques, biologiques, etc.) du modèle 3D.

4. Analyse *in virtuo* en BM

Face à la complexité des systèmes biologiques, l'analyse en BM devient une tâche difficile pour les biologistes, qui ont besoin de nouveaux outils d'analyse.

L'analyse *in virtuo*⁷ représente un nouveau moyen permettant de remédier aux limites des techniques d'analyse "classique". Elle associe les avantages des simulations numériques à ceux de la RV en plaçant le biologiste au centre des expérimentations virtuelles.

L'analyse *in virtuo* (figure 1) se compose alors de trois phases fondamentales : modélisation 3D, visualisation et interaction 3D. D'abord, la phase de modélisation permet de construire l'architecture spatiale de la molécule ou du complexe moléculaire à analyser. Ensuite, l'étape suivante est la visualisation du résultat établi. La modélisation et la visualisation constituent l'étape d'observation, proposée par la plupart des systèmes de visualisation moléculaire (ie. *viewers*). Au cours de cette étape, l'utilisateur peut visualiser le modèle 3D moléculaire sous différents points de vue. De plus, ces deux premières phases de l'analyse peuvent être enrichies par des

7. L'expression *in virtuo* (dans le virtuel) est un néologisme construit ici par analogie avec les locutions adverbiales d'étymologie latine *in vivo* (dans le vivant) et *in vitro* (dans le verre). Une expérimentation *in virtuo* est ainsi une expérimentation conduite dans un univers virtuel de modèles numériques auquel l'homme participe (Tisseau, 2001).

outils de mesure (eg. calcul de volume, distance entre points d'intérêt, etc.).

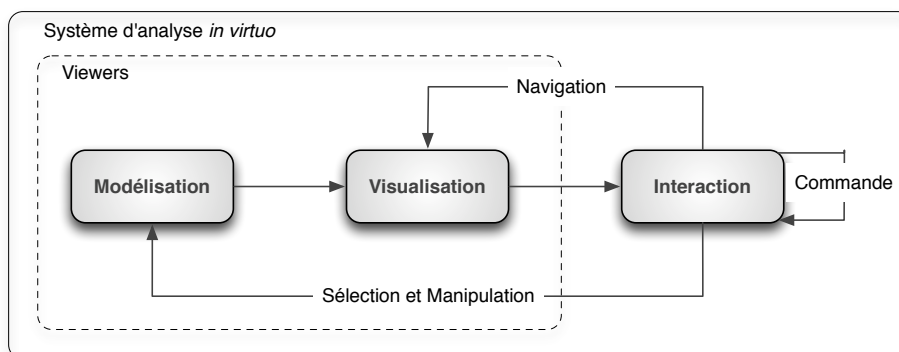


Figure 1. Composantes d'un système d'analyse in virtuo

Enfin, la phase d'interaction (Homme-Molécule) permet à l'utilisateur d'animer le modèle 3D afin de le corriger, de l'améliorer voir simplement de tester certaines hypothèses. Cette phase peut être assurée à travers une interface et des techniques d'interaction 3D adaptées.

Les systèmes d'analyse moléculaire (modélisation 3D moléculaire) que nous allons traiter désignent des outils logiciels pour des simulations moléculaires associés au dispositif matériel adopté par les auteurs.

4.1. Modélisation 3D

La modélisation 3D est un processus qui consiste à générer un modèle numérique (ie. représentation abstraite) de la géométrie d'un objet. Le modèle est souvent représenté par un nuage de points, d'une triangulation ou d'une fonction implicite. Le but du modèle 3D construit est de donner une approximation de l'objet modélisé afin d'en étudier l'aspect, l'organisation ou encore le comportement.

En BM, la modélisation 3D permet de reproduire les molécules réelles par des données numériques, principalement les coordonnées spatiales des composantes des molécules.

La modélisation 3D représente un outil visuel de l'analyse moléculaire. Elle permet au scientifique de mieux comprendre certains phénomènes biologiques, principalement à l'échelle moléculaire (ie. structure et réactivité des molécules). La connaissance des structures moléculaires aide à la compréhension des transformations physiques, chimiques ou biologiques, et leur rapport avec l'activité de la molécule. Les recherches se sont intensifiées autour de la modélisation 3D des protéines (eg. pour le docking de protéines) et du génome à différents niveaux d'intégration.

Dans ce qui suit, nous présentons les principales approches de modélisation 3D moléculaire, et de l'ADN en particulier. Nous avons recensé trois approches de modélisation 3D issues de données : expérimentales brutes, expérimentales extrapolées et interprétées (figure 2). Le principal objectif de cette classification est de distinguer les trois types de données sur lesquelles sont basées les approches de modélisation.

La plupart des techniques expérimentales sont complétées par des outils de modélisation pour obtenir les coordonnées 3D des atomes. Toutefois, nous distinguons la modélisation à partir de données expérimentales associées à une molécule donnée, et un ensemble de données expérimentales servant de base statistique à la modélisation de plusieurs molécules. Nous appelons le premier cas "données expérimentales brutes" et le deuxième "données expérimentales extrapolées".

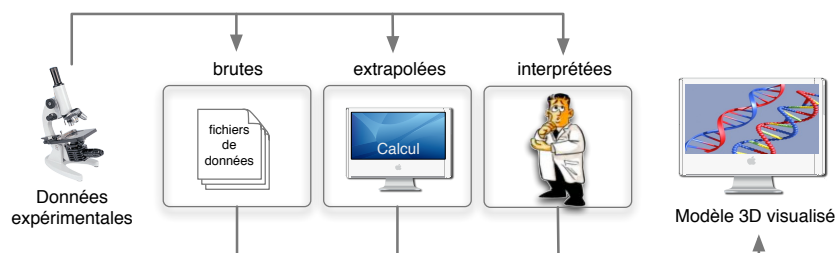


Figure 2. Les principales approches de modélisation 3D issues de données : expérimentales brutes, expérimentales interpolées et interprétées

4.1.1. Approche issue de données expérimentales brutes

Cette approche de modélisation 3D se base sur des données expérimentales brutes effectuées sur la molécule à modéliser. Cela consiste à reproduire la structure 3D de la molécule à partir de données 3D réelles. Les données expérimentales brutes concernent particulièrement les protéines. Il existe principalement deux méthodes d'observation afin d'établir la cartographie 3D de la molécule : la résonance magnétique nucléaire (RMN) pour les molécules de petites tailles, et la cristallographie par rayons X (propriétés structurales par diffraction X) particulièrement efficace pour les grosses molécules.

Concernant la molécule d'ADN, ces données sont souvent utilisées pour la modélisation de petits fragments (ie. modélisation locale). Les techniques actuelles d'études expérimentales de la structure 3D de l'ADN (cristallographie, microscopie cryo-electron, microscopie AFM, etc.) présentent différentes limites : taille de la molécule, déformation des molécules, etc. Elles sont également coûteuses en temps et en moyens financiers.

Le résultat obtenu par la modélisation issue de données expérimentales brutes, qu'elle soit par cristallographie ou par RMN, est le plus souvent stocké dans un fichier standardisé (PDB : Protein Data Bank). Ce type de fichier renferme, entre autre, le

nuage de points 3D qui constituent le modèle 3D de la molécule⁸. Ces points peuvent être ensuite visualisés à travers différents outils informatiques (ie. systèmes de visualisation moléculaire) tels que gOpenMol, MoleKel, ViewMol, MolMol, PyMol (figure 3), RasMol (figure 3), etc. (voir section 4.2).

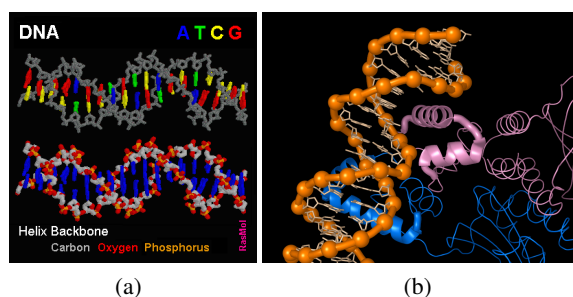


Figure 3. Modélisation de complexes moléculaires grâce à la cristallographie à travers (a) RasMol (b) PyMol

Cependant, cette approche génère une image 3D d'une molécule à un instant donné impliquant des modèles moléculaires statiques. Ainsi, l'approche issue de données expérimentales brutes est associée à la visualisation et à l'observation des molécules, ce qui rend la manipulation ou la modification de la structure difficile voir impossible. De ce fait, en adoptant cette approche, l'utilisateur est dans un processus d'examen du modèle 3D et demeure un simple observateur avec une marge de manoeuvre limitée.

4.1.2. Approche issue des données extrapolées (ou approche automatique)

Cette approche consiste à générer le modèle 3D d'une molécule par des algorithmes de calcul à partir de données expérimentales extrapolées. Ces données sont expérimentales mais qui ont été post-traitées (ie. calculées) suivant certaines contraintes biologiques. Elles tendent à prédire la structure 3D des molécules. Contrairement aux données expérimentales brutes, celles extrapolées sont fondées sur le calcul des coordonnées 3D de la molécule. Pour la prédiction des structures 3D des protéines, ces données sont utilisées par trois grands types de méthodes : la modélisation par homologie (Jones *et al.*, 1986)(Claessens *et al.*, 1989), la reconnaissance de repliements, les méthodes de prédiction *ab initio*⁹/*de novo*¹⁰ (minimum

8. Un fichier PDB est une archive de structures tridimensionnelles de macromolécules biologiques déterminées expérimentalement. Les données contenues dans un fichier PDB comprennent le nom des molécules, les coordonnées atomiques, la bibliographie, les informations structurales, les facteurs de la structure cristallographique et les données expérimentales de RMN.

9. depuis les premiers principes

10. nouvellement synthétisé

global d'énergie libre dont la méthode Rosetta (Simons *et al.*, 1999)). En revanche, concernant l'ADN, ces données de modélisation font l'objet de beaucoup de critiques car pour les obtenir, la molécule est extraite de son milieu, de plus elle est dénaturée par les techniques d'acquisition. Elles consistent à fournir une représentation spatiale de la molécule à partir de sa séquence textuelle (suite des molécules A, C, T et G) et ce grâce à des statistiques établies par les biologistes sur des petits fragments d'ADN extraits de leurs environnements (ce sont les modèles de conformation spatiale d'ADN nu). Les données extrapolées sont considérablement plus faciles à traiter, moins coûteuses et plus rapides à mettre en place qu'une méthode expérimentale *in vitro*, sachant que l'ADN à grande échelle est quasiment impossible à observer *in vitro* et encore moins *in vivo*. Une visualisation globale de la molécule d'ADN s'est basée sur ce type de données (Hérisson, 2004).

Toutefois, les algorithmes sont souvent coûteux en temps de calcul, et plus particulièrement quand il s'agit de dynamique moléculaire. Dans le passé, les systèmes de modélisation moléculaire 3D les plus intéressants ont été soit trop grands soit trop complexes à modéliser en temps réel. À présent, de nombreux programmes de simulation 3D moléculaire fonctionnent de façon efficace sur des ordinateurs parallèles. Par ailleurs, l'approche automatique, comme l'approche issue de données expérimentales brutes, génère le plus souvent des modèles moléculaires figés (ie. non modifiables) et la marge de manoeuvre de l'utilisateur reste très limitée.

4.1.3. Approche issue de données interprétées (ou approche manuelle)

L'approche manuelle consiste à construire la totalité du modèle 3D moléculaire en se basant uniquement sur les connaissances et les interprétations de l'utilisateur (expert). Cette approche se base sur une démarche itérative qui consiste à construire le modèle localement et à l'enrichir au fur et à mesure que l'on évalue et valide l'étape de modélisation précédente. Dans la phase de construction, l'approche permet à l'utilisateur d'être un acteur libre d'agir sur le modèle et de le modifier.

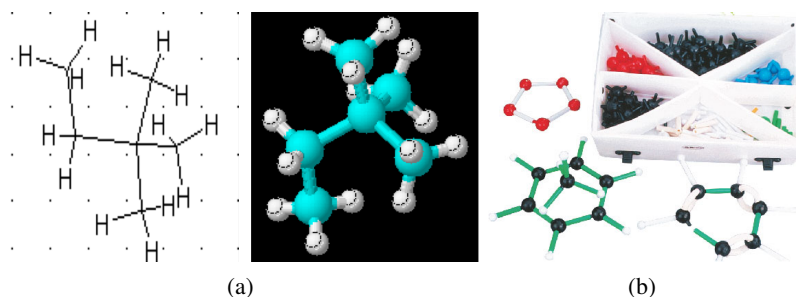


Figure 4. Modélisation manuelle pour les modèles numériques (Chemsketch) (a) et les modèles matériels (Kit de construction de molécules) (b)

Les modèles matériels de molécules (décrits dans l'introduction) ont été la première application de cette approche. Dans ce cas le modèle 3D est construit en assemblant un kit de modèles moléculaires (Fieser, 1963) en fonction de l'expertise de l'utilisateur (figure 4). Le kit est souvent constitué d'un ensemble de boules (atomes) et de bâtonnets (liaisons).

D'autres outils informatiques ont été développés grâce à cette approche. En fonction du squelette de la molécule (que l'on dessine à travers un mode d'édition de structure) et en attachant des atomes entre eux, il est possible d'analyser le modèle 3D en sortie (eg. le logiciel ACDC/Chemsketch¹¹). En plus d'offrir une liberté totale à l'utilisateur, les systèmes basés sur cette approche sont particulièrement intéressants dans l'enseignement. En effet, la création du modèle 3D est conditionnée par la connaissance de l'utilisateur à travers une interactivité Homme-Molécule continue. L'utilisateur est le seul acteur dans cette approche.

Le degré de liberté de l'utilisateur étant alors très important, le risque d'erreur dans la modélisation devient élevé. Les contraintes physiques et biologiques imposées à la modélisation 3D doivent être respectées. Cela représente une tâche délicate car c'est à l'utilisateur qu'elle revient exclusivement. Dans ce cas, un système d'assistance aiderait l'utilisateur à accomplir sa modélisation en respectant certaines contraintes, voir en le déchargeant de la vérification d'autres.

4.2. Visualisation 3D

Les modèles moléculaires 3D générés ou construits par les approches précédentes sont destinés à être analysés, c'est-à-dire visualisés et manipulés. Toutefois, la notion d'interaction décrites dans les systèmes de visualisation 3D diffère de celle d'interaction 3D. Dans ce cas, il s'agit de l'interaction nécessaire à tout système pour établir une communication entre lui et l'utilisateur. Cette interaction "classique" peut être étendue à la manipulation des modèles 3D, mais cela reste limité par le dispositif matériel inadapté et l'absence de techniques particulières.

Les systèmes que nous appellerons "non immersifs" utilisent un écran monoscopique de taille standard et des périphériques de commande classiques pour poste de travail (eg. clavier et souris). L'immersion visuelle n'est donc pas recherchée, et les interactions avec les modèles 3D ne se veulent pas naturelles.

Les outils de visualisation moléculaire "classiques" permettent d'afficher sur un écran d'ordinateur des molécules virtuelles et de les manipuler en translation, en rotation et en facteur d'échelle (ie. agrandissement et rétrécissement de l'objet). La manipulation des molécules et des vues est faite à l'aide d'une souris. Le clavier peut être utilisé, parfois, pour commander d'autres paramètres. Différents modes de représentation graphique sont possibles (modèles surfaciques, filaires, nuage de points, etc.). Les systèmes de visualisation 3D moléculaire, plus communément appelés les *viewers* (Essabbah *et al.*, 2009b), n'offrent qu'une exploration visuelle des molécules. Ils se basent sur une approche automatique, soit issue de données expérimentales brutes, soit

11. <http://www.acdlabs.com/chemsketch/>

de données expérimentales extrapolées. Ils proposent uniquement une observation du modèle 3D, sans pouvoir interagir avec.

Ces *viewers* sont majoritairement dédiés à l'analyse des protéines, moins à celle de l'ADN. Ils se distinguent par le type de données observées (section 4.1) qu'ils utilisent (modélisation par homologie, dynamique moléculaire, cristallographie, etc.). Les *viewers* ont quasiment tous la possibilité d'interroger une banque de données pour récupérer les fichiers PDB qui contiennent les informations structurales de la molécule. Ainsi, l'utilisateur a également le choix de rechercher des molécules qui ont des caractéristiques spécifiques.

Leur principal avantage est la facilité d'utilisation. Ils proposent tous des rendus, de différentes qualités.

La liste de ces logiciels est longue, nous n'en citons que quelques uns. Parmi ces systèmes il y a WPDB (Windows Protein Data Bank) (Shindyalov *et al.*, 1995), Spock (Structural Properties Observation and Calculation Kit) (christopher, 1998), MolMol (Molecule analysis and Molecule display) (Smith, 2000), PyMOL (DeLano, 2002) et DS ViewerPro (Discovery Studio ViewerPro) (Accelrys, 2003). Certains permettent d'analyser plusieurs protéines en même temps et proposent des calculs complémentaires, comme Swiss-Pdb Viewer¹² (Guex *et al.*, 1997) qui est un logiciel distribué par la Swiss Protein DataBank. D'autres ont pour points forts les rendus en haute résolution et l'affichage en temps réel comme QuteMol (Tarini *et al.*, 2006). Il permet, entre autres, d'animer les molécules.

En plus des fonctionnalités de QuteMol, RASMOL/OpenRasMol (RASter display of MOlecules) (Irisa *et al.*, 2007) est capable de gérer un très grand nombre de formats de modèles moléculaires. Les auteurs y ont ajouté certaines techniques de RV, en particulier l'affichage stéréoscopique sur grand écran et l'utilisation d'une souris 3D pour la sélection des atomes.

Les logiciels spécialisés dans l'analyse et la reconstruction des structures tridimensionnelles de la molécule d'ADN sont moins nombreux. 3DNA (3-Dimensional Nucleic Acid structures) (Lu *et al.*, 2003) peut manipuler les structures simples, triplexes, quadruplexes et d'autres motifs trouvés dans la structure de l'ADN et de l'ARN (figure 5) à partir d'un fichier de format PDB. Le programme se sert des données extrapolées sous la forme d'une armature de référence pour la description de la géométrie de paires de bases d'acides nucléiques¹³. Cela permet de calculer des paramètres locaux de conformation (ie. coordonnées 3D) et de reconstruire la structure 3D correspondante. Des outils sont fournis pour localiser les paires de bases et les régions hélicoïdales dans une structure.

Enfin, certains *viewers* permettent, en plus de la lecture des fichiers standards PDB et l'affichage des structures correspondantes, l'animation et l'analyse de la trajectoire de simulation de la dynamique moléculaire. VMD (Visual Molecular

12. <http://spdbv.vital-it.ch/>

13. Une paire de base, en abrégé pb, désigne l'appariement de deux bases azotées (Adénine (A), Thymin (T), Guanine (G) et la Cytosine (C)) situées sur deux brins complémentaires d'ADN ou ARN.

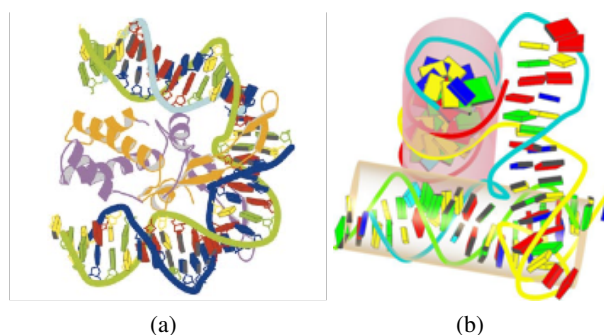


Figure 5. (a) 3DNA modélise en 3D localement l'ADN, (b) certaines parties de la double hélice relativement droites peuvent être représentées par des cylindres (Lu et al., 2003)

Dynamics)(Humphrey *et al.*, 1996)(Caddigan *et al.*, 2003) est, sans doute, l'un des premiers outils conçus pour l'analyse de larges systèmes moléculaires (protéines, acides nucléiques, lipides, etc.) (figure 6). Sa particularité est qu'il peut être utilisé sur un ordinateur distant. Il supporte également des périphériques de commandes 3D, de retour haptique et une large variété de dispositifs visuels. YASARA (Yet Another Scientific Artificial Reality Application)¹⁴ permet également l'étude de la dynamique moléculaire en plus de la modélisation et la visualisation. Il peut être utilisé dans différentes applications scientifiques. Les solutions, pour la réalité virtuelle, proposées par YASARA peuvent aller du simple écran stéréoscopique (pour une vision 3D sans lunettes) au dispositif de modélisation 3D stéréoscopique à grande échelle. Il fournit nativement les logiciels pour utiliser différents périphériques d'interaction 3D. Dans la même catégorie on trouve également gOpenMol (Bergman *et al.*, 1997) qui permet en plus l'affichage des orbitales moléculaires, des densités d'électrons, et des potentiels électrostatiques.

Très similaire à gOpenMol, Qmol (Gans *et al.*, 2001) est basé sur le logiciel molview et inspiré par l'application Xmol, et CHIMERA (Pettersen *et al.*, 2004) permet la visualisation des résultats d'amarrage moléculaire (docking) et des structures complexes (architectures supramoléculaires) (figure 6).

Molekel (Varetto, 2000) est un logiciel de modélisation moléculaire permettant de visualiser différents types de surfaces (orbitales, Van del Waals, etc.) et de les animer. Il propose un contrôle complet sur la génération des structures moléculaires (par la résolution, boîtes englobantes, etc.).

BallViewer utilise la méthode de Ackland et Jones (Ackland *et al.*, 2006) pour définir la structure moléculaire. Cette méthode est basée sur la distribution des angles de liaison entre un atome et ses plus proches voisins, et offre une précision similaire à celle

14. <http://www.yasara.org>

d'autres méthodes plus courantes (ANC, l'analyse de Voronoï), mais nécessite moins de calculs. Ballview offre différents modes de visualisation stéréoscopique.

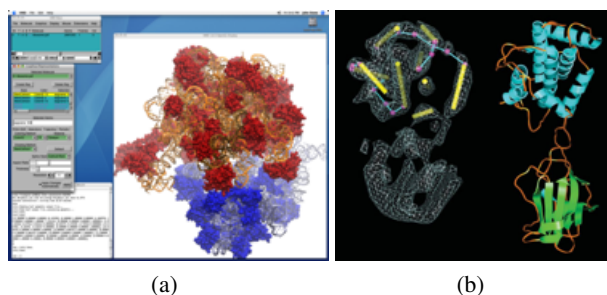


Figure 6. *Systèmes d'analyse moléculaire pour la modélisation et l'observation de dynamique moléculaire (a) VMD (b) CHIMERA*

4.3. Interaction 3D

Il existe une deuxième catégorie de systèmes d'analyse en BM, celle qui intègre la troisième phase d'interaction 3D (Essabbah *et al.*, 2009b). Nous appellerons ces systèmes "immersifs" ou "semi-immersifs", ils offrent différents degrés d'immersion, qui dépendent essentiellement du matériel utilisé : dispositifs d'affichages et périphériques d'interactions. L'immersion visuelle varie selon la taille des dispositifs d'affichage (eg. CAVE, grand écran et stéréoscopie¹⁵, workbench, etc.) et l'occlusion du monde réel qu'ils proposent. Les interactions tendent à être naturelles et multisensorielles par l'utilisation de technologies de RV .

La manipulation des modèles moléculaires 3D au clavier et à la souris (périphériques 2D), s'avère parfois complexe. Une interaction immersive et plus naturelle avec ces modèles virtuels suscite un intérêt certain chez les bio-informaticiens. En parallèle, de nouveaux dispositifs de RV sont apparus, tels que le trackball, la souris 3D, le flystick, ou encore les périphériques haptiques, et se sont avérés plus pratiques pour effectuer des tâches de navigation, de sélection et de manipulation (Ai *et al.*, 1998). En intégrant ces dispositifs, les systèmes d'analyse moléculaire ont permis aux utilisateurs, au delà de la simple observation et du changement du point de vue, de manipuler les molécules virtuelles plus naturellement et d'une façon intuitive. Ces dispositifs peuvent également apporter une multimodalité qui améliore la perception dans ces systèmes.

Notons au passage que la révolution au niveau de la visualisation et la manipulation

15. La stéréoscopie est une technique qui existe depuis les débuts de la photographie. Elle a pour principe d'envoyer deux images à chaque oeil de l'observateur afin de créer la sensation de relief dans la vision. Elle peut être passive ou active.

immersive des molécules n'implique pas forcément une nouveauté au niveau de la modélisation 3D. En effet, la plupart des systèmes que nous allons évoquer recueillent les modèles 3D moléculaires en sortie de *viewers*, similaires à ceux décrits dans la section précédente (VMD, YASARA, PyMol, etc.).

Le premier système moléculaire interactif de visualisation a été élaboré en 1965 par Cyrus Levinthal (Levinthal, 1966). Ce système a permis à un utilisateur d'orienter un modèle moléculaire tridimensionnel avec une boule en plastique (trackball) qui commande l'orientation et la vitesse du modèle manipulé (figure 7). En plaçant un pointeur lumineux sur l'écran de l'ordinateur, il était possible de choisir un objet parmi un menu ou d'agrandir les parties importantes d'une molécule.



Figure 7. *Le système de visualisation moléculaire interactif de Levinthal (Levinthal, 1966)*

Depuis, beaucoup de systèmes immersifs, de plus en plus précis, ont été développés, ainsi la BM a continué à envahir les plateformes de RV. Toutefois, et par soucis d'accessibilité vis-à-vis des utilisateurs, ces systèmes peuvent également être utilisés sur des stations de travail classiques (sur des écrans d'ordinateurs stéréoscopiques ou pas). Par exemple, VRMol (Haase *et al.*, 1996) est un système qui a été conçu pour étudier les propriétés d'une molécule dans un environnement immersif. Le dispositif mis en place est composé d'une souris 3D, d'un gant de données pour manipuler la molécule et d'un grand écran pour la projection stéréoscopique. Un menu virtuel propose plusieurs choix pour la visualisation. Celui-ci est appelé par un geste de la main. La sélection est réalisée par pointage. Dans le même registre, RealMol (Real-time Molecular dynamics simulation) (Ai *et al.*, 1998) a été développé pour la modélisation moléculaire interactive et la simulation de la dynamique moléculaire (figure 8). La seule différence avec VRMol, est que RealMol propose une immersion totale grâce à un CAVE (Cave Automatic Virtual Environment) ou un visiocasque.

ADN-Viewer (Gherbi *et al.*, 2002) (figure 9) est un logiciel qui permet la modélisation 3D et l'exploration visuelle de la molécule d'ADN. Il est basé sur les données extrapolées (ie. prédites) donnant ainsi un modèle 3D qui prévoit une

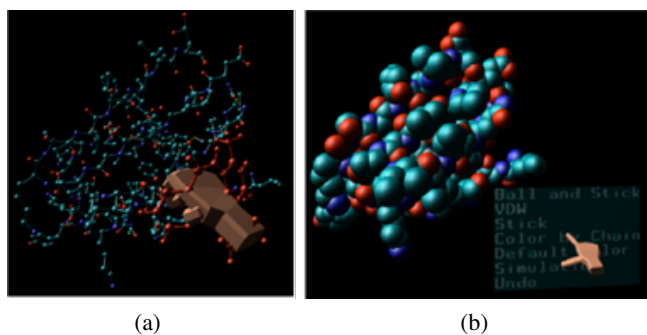


Figure 8. (a) Des molécules peuvent être saisies avec un gant de données et être déplacées dans une position désirée. (b) le menu 3D peut être activé et un choix peut être fait via le gant de données (Ai et al., 1998)

configuration spatiale de n'importe quelle séquence d'ADN nu. Cet outil offre, entre autres, plusieurs représentations de la molécule (représentation génomique ou génique). ADN-Viewer a été testé avec un affichage stéréoscopique et une interaction immersive. L'affichage stéréoscopique est assuré par deux grands écrans verticaux disposés à angle droit. L'interaction immersive est réalisée en utilisant un Wand¹⁶ (Hérisson, 2004). Pour désigner ou sélectionner aisément les portions de l'ADN tout en naviguant dans la scène, un rayon virtuel a été représenté.

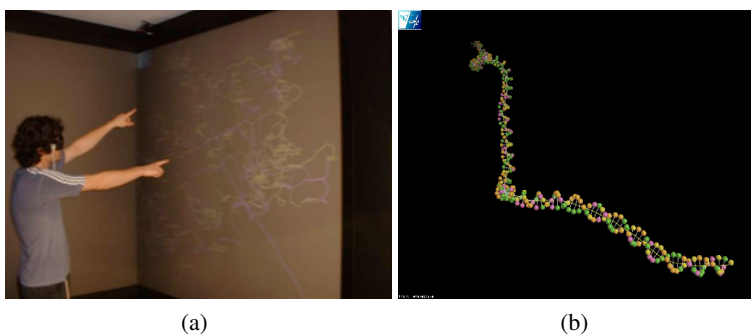


Figure 9. (a) Interaction immersive et (b) représentation génique d'ADN-Viewer

16. un périphérique 3D muni d'un capteur de position et d'orientation 3D, de 3 boutons et d'une boule à deux degrés de liberté

Les représentations visuelles stéréoscopiques peuvent être enrichies par d'autres modalités graphiques (vecteur 3D, ressort, curseur, sphère, etc.). La figure 10 représente le fonctionnement de MolDRIVE (Koutek *et al.*, 2002) (Molecular Dynamics in Real-time Interactive Virtual Environments) sur une table stéréoscopique immersive ainsi qu'un utilisateur effectuant une manipulation en temps réel d'une dynamique atomique. La manipulation s'effectue grâce à un marqueur à six degrés de liberté.

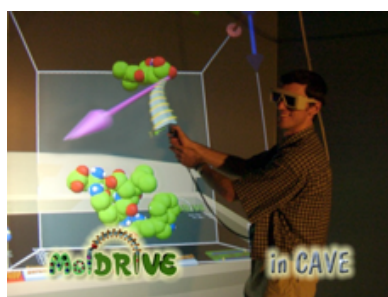


Figure 10. Exemple d'utilisation de MolDRIVE pour l'analyse de dynamique atomique

Jusqu'à présent nous avons présenté des systèmes moléculaires immersifs qui jouissent essentiellement des avantages de la modalité visuelle, et cela à travers une exploration immersive (souvent à échelle humaine) des modèles 3D moléculaires.

La modalité haptique, quant à elle, est de plus en plus appliquée à la Biologie. Le premier système permettant la visualisation et l'interaction haptique sur des molécules était GROPE-III, développé par Brooks (Brooks *et al.*, 1990) (figure 11). Ce système utilise un bras manipulateur ARM (Argonne-III Remote Manipulator) à six degrés de liberté et un casque stéréoscopique pour la visualisation. Le système GROPE-III est complété par un retour d'effort et une meilleure précision dans la manipulation par le système Nanomanipulator (Taylor *et al.*, 1993). La figure 11 montre l'interface utilisateur de ce système. L'utilisateur peut contrôler ses actions sur le STM (Scanning Tunneling Microscope) en s'aidant du retour d'effort du bras ARM. En outre, il peut sentir les contours de la surface et spécifier l'intensité de vibration en appuyant du doigt sur la manette. La visualisation des molécules virtuelles s'effectue à l'aide d'un visiocasque.

Depuis, d'autres outils de visualisation interactive ont exploité la modalité haptique. CAMD (Lai-Yuen *et al.*, 2006) et HAMStER (Haptic Application for Molecular Structure and Energy Refinement) (Wollacott *et al.*, 2007) sont des systèmes de conception moléculaire assistée par ordinateur pour le docking (amarrage) de protéines. Un écran d'ordinateur est utilisé pour la visualisation. Pour la manipulation des molécules, CAMD utilise un dispositif haptique à cinq degrés de liberté (électromécanique)(figure 12), tandis que HAMStER dispose d'une interface haptique PHANTOM (figure 12) à six degrés de liberté.

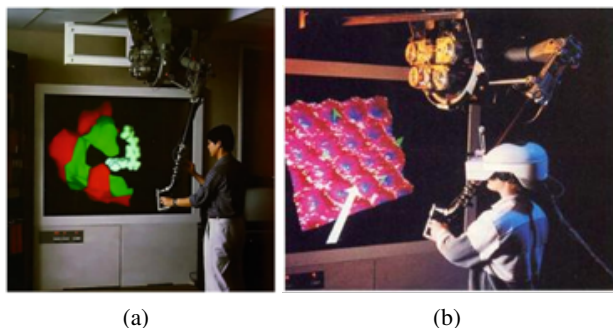


Figure 11. (a) Le système GROPE-III (b) et l'interface utilisateur pour le système Nanomanipulator

Dans le projet FVNano, dédié à la simulation de dynamique moléculaire en temps réel et en mode interactif, la modalité haptique représente la force de liaison à un site donné (Delalande *et al.*, 2010). MDDriver (Férey *et al.*, 2008)(Delalande *et al.*, 2009), une librairie générique a été développée afin de faciliter le développement de ces simulations interactives. Un environnement de réalité virtuelle est dédié à la simulation immersive et interactive. Des essais ont été établies avec différents modes de visualisation (écrans CRT stéréo, écran LCD haute résolution et un grand écran LCD) et d'interaction (souris 3D, Wiimotes, dispositifs haptiques, etc.).



Figure 12. (a) Interface utilisateur avec le dispositif haptique de CAMD (b) Le dispositif haptique PHANTOM pour l'application HAMStER

D'autres systèmes proposent diverses interactions multimodales sur les trois canaux sensori-moteurs, à savoir le visuel, l'audio et l'haptique. Dans ce sens, un système pour l'enseignement des interactions moléculaires a été développé (Sato *et al.*, 2008). La modalité haptique est assurée par un SPIDAR-G (Sato, 2002). C'est

une interface haptique à six degrés de liberté, dans laquelle une boule est attachée à huit câbles reliés à huit moteurs et encodeurs. Les rendus haptique et audio sont utilisés pour communiquer les collisions entre les molécules.

Le projet CoRSAIRE (Combinaison de Rendus Sensori-moteurs pour l'Analyse Immersive de Résultats) (Férey *et al.*, 2009) porte sur l'étude du docking de protéines (figure 13). L'immersion visuelle est assurée par un CAVE à stéréoscopie active. Les retours sonores 3D sont diffusés par un casque ou par un ensemble de huit haut-parleurs répartis dans le dispositif immersif. Les manipulations des protéines sont assurées par un bras haptique *Virtuose*, à six degrés de liberté, et une souris 3D sans fil. L'expert cherche à emboîter des protéines et le retour haptique informe sur les collisions et les forces d'attraction et de répulsion, représentées par un score.

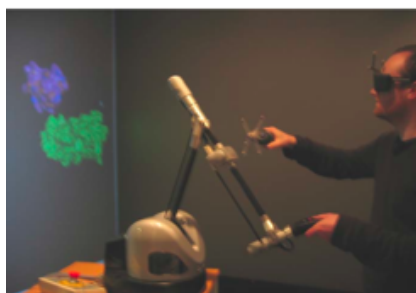


Figure 13. Le système interactif pour le docking de protéines (projet CoRSAIRE)

IGD (Interactive Global Docking) (Heyd *et al.*, 2009) a les mêmes objectifs que CoRSAIRE. Une visualisation stéréoscopique, combinée à un suivi de tête et un retour haptique, génère un environnement virtuel idéal pour une véritable interaction permettant à l'utilisateur de résoudre les problèmes de docking les plus délicats.

Les techniques de Réalité Augmentée¹⁷ sont aussi utilisées avec des modèles auto-fabriqués tangibles. Le système de Réalité Augmentée permet aux représentations 3D virtuelles (produites par un *viewer* moléculaire) d'être superposées sur un modèle moléculaire réel (obtenu par impression 3D¹⁸) (Gillet *et al.*, 2005). Les utilisateurs peuvent facilement changer cette superposition d'information, commutant entre différentes représentations de la molécule, l'affichage des propriétés moléculaires telles que l'électrostatique, ou des informations sur la dynamique (figure 14). Une fois l'objectif principal atteint, à travers les modalités mises en oeuvre, l'étape suivante est souvent de faire évoluer le système vers un outil collaboratif. Certains systèmes ont

17. La Réalité Augmentée est la superposition d'images virtuelles 3D ou 2D sur des images réelles. Elle vise à accroître notre perception du monde réel en y incrustant des objets virtuels.

18. L'impression 3D permet de produire un objet réel à partir d'un fichier. Il s'agit d'une technique de prototypage rapide qui dépose ou solidifie de la matière couche par couche afin d'obtenir la pièce désirée.

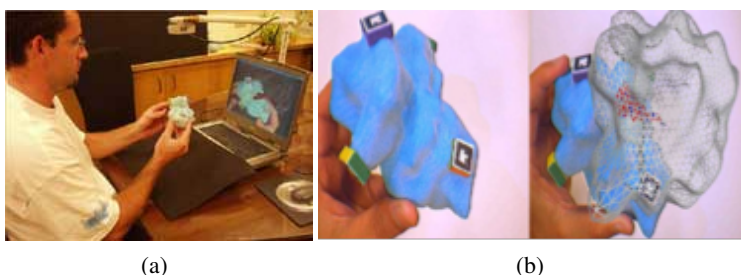


Figure 14. (a) Manipulations moléculaire à travers une interface tangible (b) exemple d'interface tangible avec et sans Réalité Augmentée (Gillet et al., 2005)

su bénéficier des avantages de la multimodalité pour proposer un outil collaboratif. NAVRNA (Bailly *et al.*, 2006) est un système multi-surface collaboratif pour l'analyse de l'ARN. Il permet de visualiser en même temps la structure tridimensionnelle projeté sur un mur blanc et la structure secondaire également projetée sur une table. Les deux représentations sont fortement liées. L'interaction est assurée par une paire de jeton (manipulés par l'utilisateur) et un système de vision par ordinateur.

AMMP-Vis (Another Molecular Modeling Program - Vis) (Chastine *et al.*, 2005) est

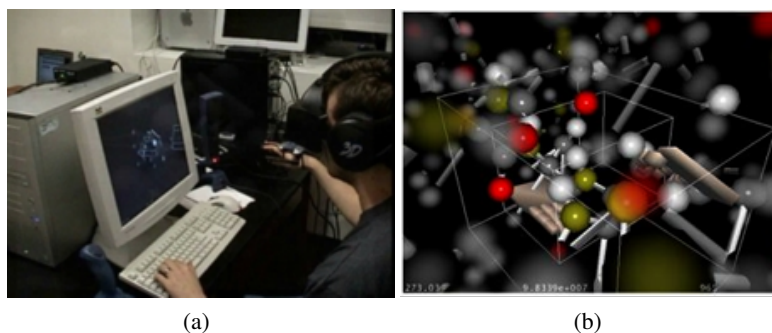


Figure 15. (a) L'interaction 3D du système AMMP-Vis (b) Exemple de collaboration entre deux utilisateurs (Chastine et al., 2005)

un environnement virtuel collaboratif pour la modélisation 3D moléculaire. Il permet aux chercheurs biologistes de saisir et déplacer des modèles moléculaires par l'intermédiaire des gants de données¹⁹ et de naviguer dans l'environnement virtuel en utilisant un joystick (figure 15). AMMP-Vis permet aussi de recevoir et de visualiser en temps réel le résultat d'un simulateur de dynamique moléculaire, de partager des

19. gants munis de capteurs pouvant détecter les mouvements des doigts

vues adaptées aux besoins des utilisateurs et de fournir un support pour les recherches collaboratives locales et distantes. Pour faciliter la sensibilité à la présence, chaque utilisateur est représenté par une main virtuelle qui est visible dans l'environnement. Les positions et les orientations de la main sont mises à jour d'une manière continue et transmises aux autres utilisateurs (figure 15). Un signal audio permet aux utilisateurs du système de savoir lorsqu'un nouvel utilisateur est présent.

5. Discussion

Après un tour d'horizon des systèmes d'analyse moléculaire, nous proposons une synthèse de ces systèmes en fonction de leur type d'immersion dans un premier temps, et en fonction des modalités employés et des tâches de l'interaction 3D, dans un deuxième temps.

5.1. Classification en fonction du type d'immersion

La plupart des systèmes moléculaires mentionnés (non-immersifs, semi-immersifs, immersifs) sont portables, ce qui signifie que ces applications sont indépendantes du dispositif matériel employé et peuvent être utilisées dans différentes plateformes. Toutefois, pour les besoins de cette étude nous n'avons tenu compte que des dispositifs utilisés par les auteurs pour pouvoir ainsi exploiter leurs résultats.

De manière générale, le nombre de systèmes non immersifs est plus important que les systèmes immersifs ou semi-immersifs. Cela peut s'expliquer par le fait que la synthèse d'images a vu le jour avant la RV qui, par ailleurs, exige des dispositifs matériels souvent coûteux et volumineux. De plus, la plupart des interfaces de visualisation (*viewers*) sont libres et simples d'utilisation puisqu'elles se basent sur des interfaces homme/machine 2D usuelles (clavier, souris, menu 2D, etc.). Ces interfaces représentent un outil idéal pour modéliser rapidement et analyser une structure moléculaire autant pour illustrer des documents que pour vérifier une hypothèse.

Toutefois, ces *viewers* demeurent limités d'une part par l'aspect automatique de la génération des modèles 3D, ce qui produit des modèles moléculaires fixes (ie. géométrie non modifiables). De plus, la structure de données de ces modèles (adoptée pour une application particulière) n'est souvent pas assez générique pour être réadaptée pour d'autres applications, mais l'approche de modélisation quant à elle peut être reprise.

D'autre part, ces systèmes non immersifs restreignent considérablement l'interaction Homme-Molécules. La visualisation des systèmes complexes est difficile pour le scientifique qui doit user de toute son expertise et toute son attention pour essayer de "démêler" les modèles 3D. L'exploration visuelle de ces molécules reste très limitée sur un écran d'ordinateur. Cela est particulièrement laborieux pour l'observation globale de la molécule.

C'est précisément sur ce verrou qu'interviennent les systèmes immersifs et semi-immersifs. L'immersion sensorielle apportée par ces systèmes offrent une nouvelle vision et une interaction moins contraignante et plus intuitive. Toutefois, ces disposi-

tifs sont nouveaux pour des utilisateurs tels que les biologistes, et requièrent donc un certain temps d'adaptation. Ceci nous pousse à nous intéresser à la façon dont on peut les aider à les utiliser.

5.2. Classification en fonction des tâches de l'interaction 3D et de la multimodalité

Cette classification concerne en particulier les systèmes immersifs et semi-immersifs traités dans la section 4.3, à l'exception de certains logiciels comme VMD, YASARA ou encore BallViewer qui intègrent des techniques de RV et sont fréquemment utilisés comme base ou outils de synthèse du modèle 3D avec lequel on va interagir (eg. CoRSAIRe utilise PyMol pour le rendu visuel). Ces systèmes d'analyse moléculaire utilisent les différentes techniques de RV qui correspondent aux tâches de l'interaction 3D (N : navigation, S : sélection, M : manipulation et CA : contrôle d'application) ainsi que les différentes modalités (A : audio, V : vidéo et H : haptique). Le tableau 5.2 présente les différentes techniques d'interaction 3D, modalités et dispositifs matériels employés dans les systèmes d'analyse moléculaire étudiés.

Tous ces systèmes utilisent la modalité visuelle, alors que les deux autres modalités complémentaires, n'ont pas toujours été exploitées. La visualisation 3D stéréoscopique apporte une notion de relief et de profondeur à la scène qui permet une observation moins ambiguë d'un environnement complexe et dense telle qu'une macromolécule biologique. Ainsi, les biologistes bénéficient à la fois d'une vue globale du modèle 3D (parfois à échelle humaine), mais aussi d'une visualisation locale où l'immersion visuelle est utile dans la distinction des détails du modèle virtuel. La modalité visuelle peut être monoscopique ou stéréoscopique. Mais la plus grande différence réside dans le matériel d'affichage utilisée (écran d'ordinateur ou grand écran à échelle humaine). En revanche, la modalité audio est très peu utilisée (AMMP-Vis, CoRSAIRe, Intermolecular learning system). La modalité haptique a souvent été employée pour rendre compte des collisions entre les molécules (comme dans GROPE-III, CAMD, HAMStER, Intermolecular learning system, CoRSAIRe et FVNano).

En ce qui concerne la classification en terme d'interaction 3D, nous constatons d'abord que la tâche de navigation est la moins employée dans les systèmes d'analyse moléculaire immersifs étudiés. La navigation, au sens classique de recherche de chemin et de déplacement n'est pas une activité centrale de ces applications. Cela n'empêche pas de fréquentes manipulations en position et rotation des modèles moléculaires, afin de varier le point de vue (global, local, etc.).

Par ailleurs, la sélection d'un objet virtuel précède souvent une manipulation. Cette tâche de l'interaction 3D est donc très utilisée bien qu'elle soit insignifiante dans le cas où l'objet est unique et/ou déjà sélectionné par défaut.

Les tâches de manipulation sont les interactions 3D principales des systèmes d'analyse moléculaire, car elles permettent de modifier relativement naturellement la structure modélisée, de la déplacer ou d'y apporter de la dynamique. Le passage d'un mode d'interaction à un autre se fait à travers le contrôle d'application, pour la plupart grâce à un menu 3D comme pour VRMol ou RealMol (section 4.3).

Systèmes moléculaires	Tâches de l'interaction 3D						Modalités			Dispositif utilisé
	N	S	M	CA	A	V	H			
Levinthal (Levinthal, 1966)		+	+	+Menu		+Mono		Ecran, pointeur lumineux, trackball		
GROPE-III (Brooks <i>et al.</i> , 1990)		+	+	+		+Stéréo	+	Grand écran, bras haptique		
VRMol (Haase <i>et al.</i> , 1996)	+	+MVS	+MVS	+ MV		+Stéréo		Grand écran, gants de données, souris 3D		
RealMol (Ai <i>et al.</i> , 1998)		+	+	+ MV		+Stéréo		Grand écran, gants de données		
MolDRIVE (Koutek <i>et al.</i> , 2002)		+	+	+		+Stéréo		Grand écran, marqueurs et tracking		
ADN-Viewer (Hérisson <i>et al.</i> , 2004)	+	+RLV				+Stéréo		Grand écran, joystick		
AMMP-Vis (Chastine <i>et al.</i> , 2005)	+	+MVS	+MVS	+	+	+Stéréo		Ecran, gants de données, joystick		
CAMD (Lai-Yuen <i>et al.</i> , 2006)		+	+			+Mono	+	Ecran, dispositif haptique		
NAVRNA (Bailey <i>et al.</i> , 2006)	+	+	+			+Mono		2 projections, 2 jetons		
HAMSTER (Wollacott <i>et al.</i> , 2007)		+	+			+Mono	+	Ecran, PHANToM		
RasMol (Irisa <i>et al.</i> , 2007)		+	+			+Stéréo		Grand écran, Isotrack		
ILS (Sato <i>et al.</i> , 2008)			+			+Mono	+	Ecran, SPIDAR		
CoRSAIRE (Férey <i>et al.</i> , 2009)			+	+Menu	+	+Stéréo	+	Grand écran, souris 3D, Virtuouse, hauts-parleurs		
FVNano (Delalande <i>et al.</i> , 2009)	+	+	+			+Stéréo	+	Grand écran, PHANToM, spaceball		
IGD (Heyd <i>et al.</i> , 2009)	+		+			+Stéréo	+	Grand écran, PHANToM, tracking de tête		

Les différents systèmes de modélisation 3D moléculaire avec leurs techniques d'interaction 3D et dispositifs associés
MVS : Main Virtuelle Simple, MV : Menu Virtuel, RLV : Rayon Laser Virtuel.

5.3. Apports des techniques de la RV pour l'aide à l'analyse en BM

Par le biais des systèmes d'analyse moléculaire que nous avons étudiés, nous avons pu identifier un certain nombre d'intérêts de l'utilisation de la RV en BM. Elle permet des expérimentations et des simulations en grandeur nature et dans un environnement proche de la réalité. C'est un support de simulation qui permet une observation et une exploration aisée de n'importe quel type de données grâce à une immersion et une navigation naturelle. Nous disposons pour cela de nombreuses métaphores de manipulation intuitive et réaliste. Ces environnements peuvent être modifiés aisément à l'aide de nombreux logiciels de modélisation et même de façon dynamique en temps réel. Ainsi, en plus d'une interaction intuitive, l'utilisateur peut bénéficier d'une visualisation stéréoscopique immersive, lui permettant de gérer tous les objets de la scène à la fois d'une manière globale mais aussi localement s'il s'agit de zones d'intérêt particulier, grâce à une notion de relief et de profondeur. Le rendu visuel est d'autant plus important que la taille des écrans utilisés est grande. Ce type de visualisation a permis de démontrer que la configuration de l'ADN n'était prédisposée aux besoins biologiques (transcription, co-régulation, etc.) (Matte-Tailliez *et al.*, 2006). De plus, en BM, les scènes atteignent très vite une masse importante de données, il est donc intéressant de bénéficier de la multimodalité ainsi que d'une gestion habile des objets virtuels. La multimodalité apporte :

- Une répartition de l'effort de perception et d'interprétation de l'information selon différents types de rendus (audio, visuel ou encore haptique)
- Une baisse de la charge cognitive (travail mental) de l'utilisateur au profit de la tâche principale

En effet, les environnements virtuels multimodaux réussissent mieux que les technologies sensorielles uniques dans la création d'une sensation de présence (Hecht *et al.*, 2006).

La modalité haptique permet d'atteindre des performances élevées en augmentant la compréhension des systèmes moléculaires étudiés. Elle permet à l'utilisateur de ressentir les forces internes d'un système moléculaire de manière transparente et stable (par contrôle de position et de force) (Boloignon *et al.*, 2009). Elle est également indispensable à une exploration efficace du potentiel électrostatique des complexes (Delalande *et al.*, 2010).

Par ailleurs, la RV peut aider à la prise de décision en montrant, par des simulations proches de la réalité, la faisabilité d'un projet en mettant en évidence les exploitations possibles. Elle facilite aussi l'identification des éventuels points sensibles. Enfin, la RV représente un support de communication et de collaboration. Une plate-forme de RV (souvent à échelle humaine) permet de réunir plusieurs acteurs d'un même projet autour d'une représentation compréhensible par tous.

5.4. Approche hybride

En dehors des avantages cités ci-dessus, nous pensons qu'un des apports de la RV pour l'analyse en BM est de permettre aux utilisateurs experts de combler certaines lacunes des approches automatiques.

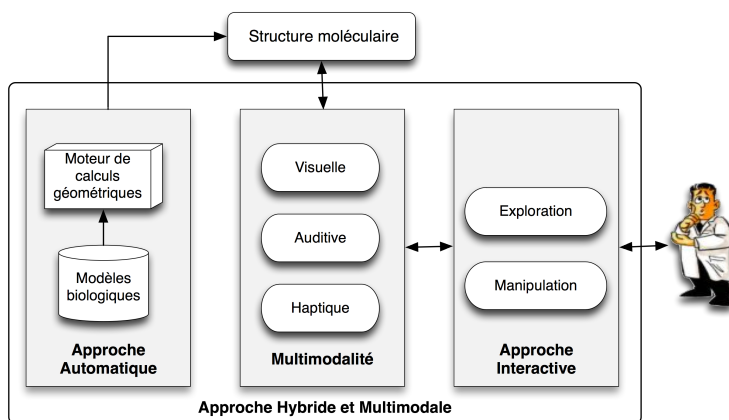


Figure 16. Une approche hybride et multimodale pour l'assistance à l'analyse moléculaire (Essabbah et al., 2009a)

Cette approche précédemment employée pour le docking de protéines (Férey *et al.*, 2009) peut être adoptée pour la modélisation de la fibre de chromatine (Essabbah *et al.*, 2008). En effet, certains algorithmes de simulation présentent des limitations en terme de temps de calcul, de ressources nécessaires, ou encore de nombre ou de validité des solutions obtenues. De plus, certaines données biologiques, connues par les experts, sont difficilement intégrables dans la modélisation. Une des solutions serait de permettre à l'utilisateur d'intervenir plus tôt dans le processus de résolution, en lui permettant d'exprimer ses connaissances directement, par l'interaction naturelle et multimodale avec les modèles moléculaires 3D (par exemple pour rapprocher spatialement deux gènes éloignés dans la séquence à cause de leur forte interaction). Plus précisément, l'utilisateur est intégré au centre du protocole développé qui devient semi-automatique. Ainsi, il serait possible de réduire le temps de calcul et les risques de modèles 3D peu probables. Cette approche que l'on peut qualifier d'hybride entre automatique et interactive vise à combiner les atouts des techniques de RV, de la multimodalité, des capacités de calcul des algorithmes et des connaissances des experts, au cours de certaines phases de l'étude.

L'utilisateur alterne l'utilisation de l'approche automatique et l'approche interactive de la modélisation, en fonction de son besoin (illustré dans la figure 16). La première étape est associée à l'approche automatique de la procédure de modélisation, elle va générer une première structure 3D, avec le moteur géométrique. Puis, à travers l'envi-

ronnement immersif, le biologiste sera en mesure d'améliorer le modèle 3D dans les configurations qui lui semblent intéressantes par le biais de retours visuels, auditifs et haptiques. Pendant cette phase, le système exploite l'image mentale de l'expert qui est issue de son expérience. Cette première étape permettra de réduire rapidement l'ensemble de recherche de conformations possibles, fondée sur la capacité de l'expert à analyser des modèles 3D et sa connaissance de la molécule et des modèles biologiques. Enfin, l'expert peut choisir de relancer la génération automatique du modèle 3D, localement ou globalement, mais en prenant en compte les changements précédemment effectués.

6. Conclusion

Ce travail constitue un aperçu des recherches menées autour des systèmes d'analyse 3D moléculaire. Dans un premier temps, nous avons donné les définitions de la BM, de la RV et de l'interaction 3D. Ensuite nous avons exposé les problématiques de la BM au niveau de l'exploration des structures moléculaires ainsi que les besoins qui en ont découlé : l'exploration 3D des modèles moléculaires et l'interaction Homme-Molécules.

Nous avons présenté une nouvelle approche d'expérimentation : l'analyse *in virtuo*. Cette analyse se compose de trois phases : modélisation 3D, visualisation et interaction 3D. Après la description des différentes approches de la modélisation 3D moléculaire nous avons proposé une revue des systèmes d'analyse moléculaire existants. Les applications existantes dans ce domaine tendent à se concentrer sur la visualisation. L'utilisation de la souris et du clavier reste dominante mais certaines applications commencent à intégrer des interactions plus naturelles et intuitives par l'utilisation des dispositifs d'interaction 3D tels que les gants de données, le joystick, la souris 3D, le bras haptique, etc.

D'autre part, les techniques d'interaction 3D ont suscité un grand intérêt pour faciliter la manipulation dans la scène virtuelle et pour guider l'opérateur à optimiser la solution. Nous avons proposé une classification des systèmes d'analyse moléculaire, immersifs, en terme de tâches d'interaction 3D et des différentes modalités employées. Ceci nous a amené à deux conclusions, la première est que toutes les tâches d'interaction 3D sont employées, bien qu'il y ait une utilisation moindre pour la navigation. La deuxième conclusion est que la multimodalité peut être exploitée pour compléter l'Interaction Homme-Molécule : par exemple les rendus haptiques pour ressentir les collisions ou les interactions énergétiques, ou bien les sons pour diminuer la charge cognitive.

Toutefois, nous avons noté que la majorité des systèmes d'analyse moléculaire dissocient la part de la modélisation 3D des molécules de la part de l'interaction Homme-Molécules. Dans un cas idéal, le scientifique pourrait, au moment des expérimentations biologiques, exploiter immédiatement ses résultats pour une modélisation 3D numérique et semi-automatique. L'approche semi-automatique préconisée repose sur une phase de modélisation issue des données expérimentales extrapolées (ie. approche automatique), et sur une phase basée sur des données interprétées par le scientifique

(ie. approche manuelle). L'intervention du scientifique devrait se faire d'une manière naturelle et intuitive à travers les dispositifs de RV et les techniques d'interaction 3D. Nous pensons également qu'il est intéressant de développer un système d'assistance à l'analyse *in virtuo* afin de permettre à l'utilisateur de respecter les contraintes biophysiques régissant les macromolécules.

7. Bibliographie

- Accelrys, « DS ViewerPro 5.0 », Available : <http://www.accelrys.com/products/>, 2003.
- Ackland G., Jones A., « Applications of local crystal structure measures in experiment and simulation », *Physical Review B*, vol. 73, n 5, p. 54104, 2006.
- Ai Z., Frohlich T., « Molecular dynamics simulation in virtual environments », *Computer Graphics Forum*, vol. 17, n 3, p. 267-273, Jan, 1998.
- Anderson A., Weng Z., « VRDD : applying virtual reality visualization to protein docking and design », *Journal of Molecular Graphics and Modelling*, vol. 17, n 3-4, p. 180-186, Jan, 1999.
- Bailly G., Auber D., Nigay L., « From Visualization to Manipulation of RNA Secondary and Tertiary Structures », *Information Visualization, 2006. IV 2006. Tenth International Conference on*, p. 107-116, 2006.
- Bergman D., Laaksonen L., Laaksonen A., « Visualization of solvation structures in liquid mixtures », *Journal of Molecular Graphics and Modelling*, vol. 15, n 5, p. 301-306, Jan, 1997.
- Bolopion A., Cagneau B., Redon S., Régnier S., « Haptic feedback for molecular simulation », *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, p. 237-242, 2009.
- Bowman D., « Interaction techniques for common tasks in immersive virtual environments : Design, evaluation and application », *Thèse de doctorat de l'Institut de Technologie de Georgie*, Jan, 1999.
- Brooks F., Ouh-Young M., Batter J., « Project GROPEHaptic displays for scientific visualization », *ACM SIGGraph Computer Graphics*, vol. 24, n 4, p. 177-185, Jan, 1990.
- Caddigan E., Cohen J., Gullingsrud J., Stone J., « VMD User's Guide », *Urbana*, 2003.
- Chastine J., Brooks J., Zhu Y., Owen G., Harrison R., Weber I., « AMMP-Vis : a collaborative virtual environment for molecular modeling », *Proceedings of the ACM symposium on Virtual reality software and technology*, p. 8-15, 2005.
- christopher J. A., « SPOCK : The Structural Properties Observation and Calculation Kit (Program Manual) », *The Center for Macromolecular Design*, 1998.
- Claessens M., Cutsen E., Lasters I., Wodak S., « Modelling the polypeptide backbone with 'spare parts' from known protein structures », *Prot. Eng.*, vol. 4, p. 335, 1989.
- Dardel F., Képès F., « Bioinformatique : Génomique et post-génomique », *Ecole Polytechnique*, Jan, 2002.
- Delalande O., Férey N., Grasseau G., Baaden M., « Complex molecular assemblies at hand via interactive simulations », *J. Comput. Chem.*, vol. 30, n 15, p. 2375-2387, 2009.

- Delalande O., Ferey N., Laurent B., Gueroult M., Hartmann B., Baaden M., « Multi-resolution approach for interactively locating functionally linked ion binding sites by steering small molecules into electrostatic potential maps using a haptic device », *Pac. Symp. Biocomput.* 205-215, 2010.
- DeLano W., « Pymol : An open-source molecular graphics tool », *CCP4 Newsletter On Protein Crystallography*, Jan, 2002.
- Essabbah M., Herisson J., Otmane S., Mallem M., « Towards a biophysical 3D model of the DNA », *1th IEEE International Workshops on Image Processing Theory, Tools and Applications* p. 1-6, Jan, 2008.
- Essabbah M., Otmane S., Hérisson J., Mallem M., « A New Approach to Design an Interactive System for Molecular Analysis », *Human-Computer Interaction, Part IV, HCII 2009, LNCS 5613* p. 713-722, Jan, 2009a.
- Essabbah M., Otmane S., Mallem M., « 3D Molecular Interactive Modeling », *Human-Computer Systems Interaction, Advances in Soft Computing, Springer Berlin / Heidelberg*, vol. 60, p. 493-504, 2009b.
- Férey N., Delalande O., Grasseau G., Baaden M., « A VR framework for interacting with molecular simulations », *ACM symposium on Virtual reality software and technology* p. 91-94, 2008.
- Férey N., Nelson J., Martin C., Picinali L., Bouyer G., Tek A., Bourdot P., Burkhardt J., Katz B., Ammi M., Etchebest C., Autin L., « Multisensory VR interaction for protein-docking in the CoRSAIRE project », *Virtual Reality*, vol. 13, n 4, p. 257-271, 2009.
- Fieser L., « Chemistry in three dimensions », *Selbstverl.*, Jan, 1963.
- Fuchs P., Moreau G., Berthoz A., Coquillart S., « Le traité de la réalité virtuelle : Volume 2, Interfaçage, immersion et interaction en environnement virtuel », *Presses des Mines Paris Tech*, Jan, 2006a.
- Fuchs P., Moreau G., Berthoz A., Vercher J., Auvray M., « Le traité de la réalité virtuelle : Volume 1, L'homme et l'environnement virtuel », *Presses des Mines Paris Tech*, Jan, 2006b.
- Gans J., Shalloway D., « Qmol : a program for molecular visualization on Windows-based PCs », *Journal of Molecular Graphics and Modelling*, vol. 19, n 6, p. 557-559, Jan, 2001.
- Gherbi R., Herisson J., « Representation and processing of complex DNA spatial architecture and its annotated content », *Proceedings of the International Pacific Symposium on Biocomputing* p. 151-162, 2002.
- Gillet A., Sanner M., Stoffler D., Olson A., « Tangible interfaces for structural molecular biology », *Structure*, vol. 13, n 3, p. 483-491, 2005.
- Guex N., Peitsch M., « SWISS-MODEL and the Swiss-Pdb Viewer : an environment for comparative protein modeling », *Electrophoresis*, vol. 18, n 15, p. 2714-2723, Jan, 1997.
- Haase H., Strassner J., Dai F., « VR techniques for the investigation of molecule data », *Computers & Graphics*, vol. 29, n 2, p. 207-217, Jan, 1996.
- Hecht D., Reiner M., Halevy G., « Multimodal virtual environments : Response times, attention, and presence », *PRESENCE : Teleoperators and Virtual Environments*, vol. 15, n 5, p. 515-523, Jan, 2006.
- Hérisson J., « Représentation spatiale et exploration virtuelle des génomes : Une approche globale pour l'analyse des éléments architecturaux des séquences », *Thèse de l'Université de Paris XI*, Apr, 2004.

- Hérisson J., Gros P., Férey N., Magneau O., Gherbi R., « DNA in Virtuo visualization and exploration of 3D genomic structures », *Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africap*. 35-40, Jan, 2004.
- Heyd J., Birmanns S., « Immersive structural biology : a new approach to hybrid modeling of macromolecular assemblies », *Virtual Reality*, vol. 13, n 4, p. 245-255, Dec, 2009.
- Humphrey W., Dalke A., Schulten K., « VMD : visual molecular dynamics », *Journal of Molecular Graphics*, vol. 14, n 1, p. 33-38, Jan, 1996.
- Irisa M., Gondo S., Fujishima Y., Kakizaki K., « Stereoscopic viewing system for proteins using OpenRasmol : a tool for displaying a filament of proteins », *BIOPHYSICS*, vol. 3, p. 57-61, Jan, 2007.
- Jones T., Thirup S., « Using know substructures in protein model building and crystallography », *EMBO J*, vol. 5, p. 819-822, 1986.
- Koutek M., van Hees J., Post F., Bakker A., « Virtual spring manipulators for particle steering in molecular dynamics on the responsive workbench », *Proceedings of the workshop on Virtual environments 2002*, Jan, 2002.
- Lai-Yuen S., Lee Y., « Interactive computer-aided design for molecular docking and assembly », *Computer-Aided Design and Applications*, vol. 3, n 6, p. 701-709, Jan, 2006.
- Leech J., Prins J., Hermans J., « SMD : Visual steering of molecular dynamics for protein design », *IEEE Computational Science & Engineering*, vol. 3, n 4, p. 38-45, Jan, 1996.
- Levine D., Facello M., Hallstrom P., Reeder G., « Stalk : An interactive system for virtual molecular docking », *IEEE Computational Science & Engineering*, vol. 4, n 2, p. 55-65, Jan, 1997.
- Levinthal C., « Molecular model-building by computer », *Scientific American*, Jan, 1966.
- Loftin R., Pettitt B., Su S., Chuter C., « PaulingWorld : an immersive environment for collaborative exploration of molecular structures and interactions », *Proceedings of the 17th Nordic Internet Conference Nordunet*, Jan, 1998.
- Lu X., Olson W., « 3DNA : a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures », *Nucleic Acids Research*, vol. 31, n 17, p. 5108-5121, Jan, 2003.
- Matte-Tailliez O., Herisson J., Férey N., Magneau O., Gros P., Képès F., Gherbi R., « Yeast naked DNA spatial organization predisposes to transcriptional regulation », *Lecture Notes in Computer Science*, vol. 3980, p. 222, 2006.
- Ouramdane N., Otmame S., Mallem M., « Interaction 3D en Réalité Virtuelle-Etat de l'art », *Technique et Science Informatiques (TSI)*, vol. 28, n 8, p. 1017-1049, Jan, 2009.
- Petterson E., Goddard T., Huang C., Couch G., « UCSF Chimera-a visualization system for exploratory research and analysis », *J. Comput. Chem.*, vol. 25, p. 1605-1612, Jan, 2004.
- Sato M., « Development of string-based force display : SPIDAR », *In Proc. The Eighth International Conference on Virtual Systems and Multi Media VSMM*p. 1034-1039, Jan, 2002.
- Sato M., Liu X., Murayama J., Akahane K., Isshiki M., « A Haptic Virtual Environment for Molecular Chemistry Education », *Lecture Notes in Computer Science*, vol. 5080, p. 28-39, Jan, 2008.
- Shindyalov I., Bourne P., « WPDB-PC Windows-based interrogation of macromolecular structure », *Journal of Applied Crystallography*, vol. 28, p. 847-852, Jan, 1995.

- Simons K., Bonneau R., Ruczinski I., Baker D., « Ab initio protein structure prediction of CASP III targets using ROSETTA », *Proteins : Structure, Function, and Bioinformatics*, vol. 37, n 3, p. 171-176, Jan, 1999.
- Smith J., « MOLMOL : A free biomolecular graphics/analysis package », *Genome Biology*, Jan, 2000.
- Tarini M., Cignoni P., Montani C., « Ambient occlusion and edge cueing to enhance real time molecular visualization », *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, n 6, p. 1237-1244, Jan, 2006.
- Taylor R., Robinett W., Chi V., Jr F. B., « The nanomanipulator : a virtual-reality interface for a scanning tunneling microscope », *Proceedings of the 20th annual conference on Computer graphics and interactive technique*, p. 127-134, Jan, 1993.
- Tisseau J., « Réalité virtuelle : autonomie in virtuo », *Habilitation à Diriger des Recherches*, Jan, 2001.
- Varetto U., « Molekel », *Swiss National Supercomputing Centre : Manno, Switzerland.*, 2000.
- Watson J., Crick F., « Molecular structure of nucleic acids », *Nature*, n 171, p. 737-738, 1953.
- Wollacott A., Merz K., « Haptic applications for molecular structure manipulation », *Journal of Molecular Graphics and Modelling*, vol. 25, n 6, p. 801-805, Jan, 2007.