

UTILISATIONS PÉDAGOGIQUES DES VISUALISATIONS TRIDIMENSIONNELLES DE MOLÉCULES EN BIOLOGIE

J. Barrère, J-Y. Dupont, N. Salamé
INRP - Lycée Paul Louis Courier - TOURS

Mots-clés : données moléculaires, visualisation tridimensionnelle, biologie, structure, fonction

Résumé : Différentes techniques permettent de déterminer et modéliser la structure d'une molécule dans l'espace. Les coordonnées tridimensionnelles de centaines de molécules ainsi que différentes informations sur leurs structures sont disponibles dans les banques de données internationales. Les logiciels de visualisation et de manipulation interactive de ces molécules sont devenus disponibles sur micro-ordinateur avec des fonctionnalités assez riches pour envisager des utilisations pédagogiques. Les caractéristiques de l'un de ces logiciels (RASMOL) qui est facilement accessible sont présentées. Sur différents exemples, l'ADN, les globines, une enzyme, etc., des exploitations possibles sont détaillées qui initient à l'analyse des structures de ces molécules et à la mise en évidence de leurs fonctions biologiques en relation avec leur configuration spatiale.

DES TECHNIQUES DE SÉQUENÇAGE AUX MODÈLES MOLÉCULAIRES

Les techniques de séquençage permettent depuis une quarantaine d'années de déterminer les structures primaires qui entrent dans la composition de gènes et de protéines. Des milliers de séquences sont maintenant disponibles qui ont fait progresser considérablement les connaissances sur l'information génétique, sur les mécanismes à l'origine de la diversité des molécules biologiques et sur les conséquences des mutations qui peuvent se traduire par des modifications phénotypiques et être dans certains cas à l'origine de maladies.

Alors que l'hypothèse concernant l'importance de la forme des molécules avait été émise dès 1946, Pauling et Astbury en 1951 proposent les modèles théoriques de structure secondaire (l'hélice alpha et le feuillet bêta). C'est la période où le perfectionnement des techniques, notamment la diffraction aux rayons X, permet à Kendrew et Perutz de proposer la première maquette représentant la molécule de myoglobine. En 1953, Crick et Watson imaginent un modèle de l'ADN en double hélice. L'idée d'une relation entre structure et fonction d'une protéine commence à être précisée. Elle est maintenant au centre des préoccupations de beaucoup de molécularistes.

Les programmes de biologie pour les classes de lycée ont intégré ces évolutions et comportent de nombreuses références aux conformations spatiales en rapport avec les mécanismes biologiques. Ainsi, dans le programme de première scientifique, nous pouvons relever : "les propriétés de la molécule d'ADN sont liées à sa structure (...) c'est la séquence des acides aminés qui détermine la structure spatiale d'une protéine indispensable à son activité, la complémentarité entre une zone de la molécule enzymatique et une zone correspondante de la molécule de substrat est indispensable à la réalisation de la réaction",...

LES RESSOURCES UTILISABLES POUR L'ENSEIGNEMENT

On ne disposait jusqu'à ces toutes dernières années que de documents d'illustration et de maquettes pour faire appréhender par les élèves les concepts spatiaux dans le domaine moléculaire. Actuellement, les banques de données internationales rendent accessibles une grande partie des structures spatiales connues des molécules. C'est ainsi qu'en rapport avec l'enseignement actuel de la biologie, on peut noter l'existence dans Protein Data Bank (PDB, principale banque de référence) de molécules utilisables, par exemple, pour aborder les thèmes suivants : les acides nucléiques, porteurs d'une information génétique (avec des molécules d'ADN, d'ARNm), les protéines, leur structure, les rapports structure - fonction (avec l'insuline, l'hémoglobine oxydée, réduite, le trypsinogène, la chymotrypsine), les enzymes et la catalyse enzymatique (avec la chymotrypsine, l'acétylcholinestérase, la lactate déshydrogénase, la catalase, etc.), les rapports hormone-récepteur (avec l'enzyme de croissance), l'immunologie (avec les HLA et peptides du soi ou non soi, les récepteurs de Lymphocytes T4), les protéines et l'évolution (avec les globines, les sérines protéases), etc.

Les informations sur les molécules que l'on trouve dans les fichiers de données sont pour chaque molécule, sa séquence primaire, la désignation de chaque atome, le groupe auquel il appartient (acide aminé ou acide nucléique ou autre), ses coordonnées selon Ox, Oy, Oz et, suivant les cas, des indications concernant les structures (nombre de chaînes, nombre et longueur des hélices, les feuillettes, les coudes, les ponts dissulfure, etc.) et éventuellement les groupes qui participent au site actif de la molécule.

Pour traiter ces données sur micro-ordinateur il existe maintenant plusieurs logiciels accessibles parmi lesquels on s'est intéressé particulièrement au logiciel RASMOL dont on détaille ici quelques traitements.

Le logiciel RASMOL, fonctionnant sous Windows, est exploitable à partir d'un ensemble de trois fenêtres : une fenêtre donnant l'accès à des macro-commandes (Menus), une fenêtre d'affichage de l'image, une fenêtre dans laquelle, à la fois, on récupère des informations (séquences, nature, position, appartenance d'un atome désigné dans la fenêtre d'affichage, etc.) et on signifie des instructions en utilisant un langage simple de commande. Il ouvre des fichiers de molécules assez volumineuses (plus de 10 000 atomes).

Il comporte en premier lieu des fonctions d'affichage des informations (nom de la molécule, sa séquence par chaîne, ses structures) qui sont lues dans le fichier PDB. Il permet ensuite différents modes de visualisation du modèle moléculaire : sphères, bâtonnets, sphères et bâtonnets, squelette carboné, ruban, enveloppe matérialisée par la surface de Van der Waal's. Les liaisons sont figurées différemment suivant leur nature : liaisons covalentes, liaisons "hydrogène" et ponts dissulfure. Par défaut, les atomes et leurs liaisons sont colorés de manière conventionnelle, mais on peut en modifier les couleurs. Il offre également la possibilité de manipuler de manière interactive la molécule visualisée à travers des rotations suivant les trois axes, des translations et des changements d'échelle.

Tous les éléments composant une molécule sont accessibles en cliquant directement à l'écran. On peut ainsi les identifier, obtenir des indications sur les distances ou les angles entre atomes. Pour faciliter la lecture des volumes on fera varier l'intensité des contrastes, on inclura la molécule dans un parallélépipède ou on demandera l'affichage d'un système d'axes dont l'origine est au centre de gravité de la molécule. L'exploration du volume peut aussi être faite en réalisant des coupes sériées.

On a enfin la possibilité de sélectionner un sous-ensemble quelconque d'éléments en fonction des propriétés physico-chimiques connues de ses constituants. Des "scripts" peuvent enchaîner une série de traitements et des exportations des images obtenues sont prévues dans différents formats.

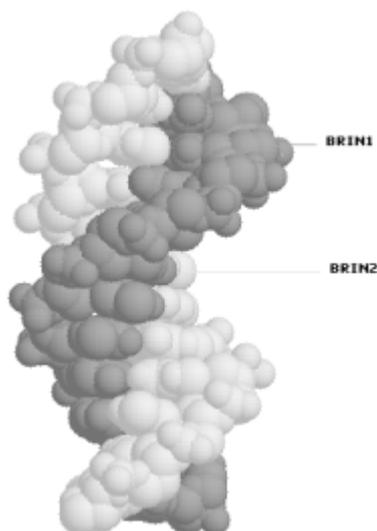
APPLICATIONS PÉDAGOGIQUES

Plutôt que de s'en tenir à l'illustration pour faire appréhender l'importance de la structure spatiale d'une molécule, notre démarche consiste à mettre l'élève en activité et lui permettre d'explorer et de découvrir.

Sans décrire ici les séquences pédagogiques et leur progression, différents niveaux de cette découverte seront présentés en relation avec les fonctionnalités du logiciel. Un premier niveau consiste à appréhender la structure de la molécule, en analyser l'organisation, en comprendre la cohésion. Un second niveau aura pour objectif de relier la structure à la fonction. Un troisième niveau visera à découvrir que la fonction d'une molécule n'est pas nécessairement liée à une structure figée.

Appréhender la structure d'une molécule

Dans de nombreux cas, s'il s'agit d'une molécule de grandes dimensions dont l'allure globale aura déjà été vue par d'autres techniques : pour l'ADN, des observations au microscope électronique révèlent sa silhouette filamenteuse. Pour une protéine, par exemple, l'hémoglobine ou une immunoglobuline, c'est sa silhouette globulaire qui est connue. La première visualisation moléculaire en 3D doit permettre de retrouver cette allure. Dans le cas de l'ADN, la visualisation en sphères fait apparaître l'allure globale de la molécule : en la manipulant dans l'espace, on repère facilement sa caractéristique filamenteuse. Une rotation selon son plus grand axe permet également de retenir sa structure plus ou moins torsadée, son agencement spiralé (document 1).



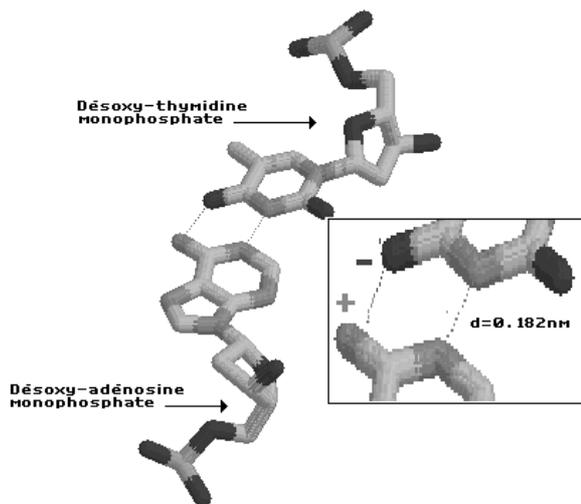
Document 1 :

la visualisation de la molécule d'ADN en sphères fait apparaître sa caractéristique filamenteuse et sa structure torsadée.

De la même manière, une vue apicale montrera sa section globalement circulaire et la position centrale, rayonnante d'unités moléculaires (riches en azote). En connaissant les couleurs conventionnelles de désignation des atomes, on identifie ainsi les principaux éléments chimiques présents (P,C,O,N,H).

La visualisation en bâtonnets, plus “transparente”, permet de voir comment les différents éléments chimiques repérés précédemment sont agencés : c’est alors qu’on remarque certains motifs qui se succèdent tout au long de la molécule et se répètent. Une identification de chacun d’eux est réalisable : on fait un “zoom” sur un secteur de la molécule et on identifie, en faisant appel à une table de référence comportant les formules développées, des constituants possibles de cette molécule (bases azotées, ribose, déoxyribose, acide phosphorique). On repérera surtout qu’ils sont agencés de façon à constituer deux brins enlacés dans l’espace (double hélice). On découvrira aussi que certains groupes (les bases azotées) appartenant à chaque brin s’affrontent sans liaisons chimiques apparentes.

Il s’agit alors de comprendre la cohésion de l’ensemble. L’exploration du modèle visualisé en bâtonnets a déjà révélé que ce sont les bases azotées qui occupent l’espace central. On constate que ces bases, à un même niveau, sont coplanaires, qu’elles ne sont pas quelconques (il y a des paires de bases complémentaires). L’absence de liaisons “apparentes” montre qu’elles ne sont pas unies par des liaisons chimiques “fortes” covalentes. On peut rechercher alors des liaisons “faibles”, électrostatiques, qui s’établissent entre groupes chimiques pourvu qu’ils soient de charges opposées et à une distance convenable (0.15 à 0.22 nm). On sélectionne à partir de l’image globale une paire de nucléotides ; avec l’option Géométrie du logiciel, en cliquant sur deux atomes opposés de chaque base (par exemple O de la guanine et le NH₂ de la thymine en vis-à-vis), on obtient la distance (ex : 0.18 nm). On montre ainsi que les conditions sont réunies pour l’établissement de liaisons faibles (document 2).



Document 2 :

les 2 brins de la molécule d'ADN sont solidarisés grâce à l'établissement de liaisons faibles entre nucléotides complémentaires.

On peut ainsi appréhender la structure de cette molécule en variant les modes de représentation, en l’orientant dans l’espace, en sélectionnant et identifiant des motifs, en faisant des mesures.

Relier les structures avec les fonctions , l’exemple des globines

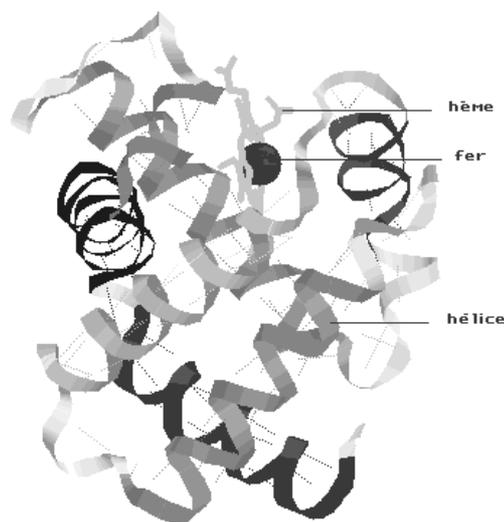
Il ne s’agit pas, à ce niveau de la démarche, de découvrir la structure d’une protéine en général, mais de mettre en relation la structure d’une famille de molécules avec leur fonction, ici la fixation et le transport du dioxygène. Déjà l’élève sait qu’une protéine est constituée par l’enchaînement d’un certain nombre d’acides aminés, que la chaîne est repliée dans l’espace, que ce repliement prend souvent l’aspect de motifs communs à l’ensemble des protéines (enroulement en hélice, repliement en feuillets).

La visualisation d’une molécule de bêta globine en bâtonnets montre la complexité de sa structure. La première question qui pourra se poser est de savoir si cette molécule comporte

une ou plusieurs chaînes, ce que l'on peut mettre en évidence en demandant l'affichage de la séquence.

Pour identifier les caractéristiques de la disposition spatiale de cette chaîne, il est nécessaire de visualiser une image simplifiée : en squelette carboné ou en ruban¹. On identifie ainsi des motifs de structure, essentiellement des portions en hélices réunies par des segments de liaisons. Une coloration sélective de ces différents domaines est possible. On comprend ainsi parfaitement l'origine de l'allure globuleuse de la molécule.

Pour comprendre comment cette molécule assure la fonction de fixation du O₂, l'une des démarches consiste à rechercher et à localiser une molécule de O₂. On peut sélectionner les atomes d'oxygène, les colorer de façon à les repérer aisément. En pratiquant des changements d'échelle et en manipulant la molécule dans l'espace, on découvre un seul assemblage de deux atomes d'oxygène liés à un atome de fer qui appartient à une unité appelée "hème" (HEM101:FE). En choisissant une représentation particulière pour mettre ce dernier en évidence, on constate qu'il est logé dans une "poche" ménagée au cœur de la molécule et délimitée par les hélices (document 3).



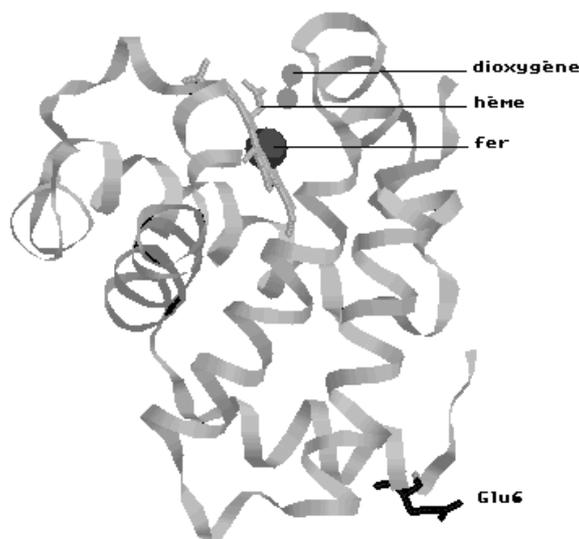
Document 3 :

la visualisation en ruban de la globine alpha humaine fait apparaître la présence de 8 hélices stabilisées par de nombreuses liaisons hydrogène.

La relation entre la disposition structurale et la fonction de la molécule n'est pas évidente si elle n'est constatée que sur une seule molécule. Compte-tenu de la richesse des banques de données, il est possible de réaliser la même recherche sur des molécules différentes (myoglobine, alpha et bêta globines, globines fœtale ou appartenant à différentes espèces). La découverte des mêmes structures, des hélices en nombre et disposition comparables, ménageant la même poche pour le hème, amène à fonder l'idée qu'une telle disposition structurale de la molécule est bien en rapport avec la fonction (document 4).

¹ Dans la représentation d'une molécule, un premier niveau de simplification consiste à débarrasser les acides aminés de leurs chaînes latérales. On ne conserve que les C-alpha et les liaisons peptidiques. On sait que les atomes impliqués dans la liaison peptidique entre deux C-alpha sont dans un même plan. On pourra même supprimer la représentation des atomes en ne conservant que le plan propre à chaque liaison. C'est ainsi que la chaîne sera figurée par un "ruban" dont la position dans l'espace épouse précisément le plan des liaisons peptidiques successives.

Un deuxième niveau de simplification ne retient que les C-alpha. La représentation, plus anguleuse, correspond à la liaison entre C-alpha successifs (liaison purement arbitraire, les C-alpha étant trop éloignés pour qu'une quelconque liaison chimique puisse s'établir entre eux) : c'est la représentation en squelette carboné alpha (en "fil de fer"). Cette visualisation ne fait apparaître par défaut que les structures protéiques. Les autres groupes (cofacteurs, ions eau, etc.) ne seront affichés que s'ils sont sélectionnés séparément.



Document 4 :

la globine bêta humaine présente une structure spatiale comparable à celle de la globine alpha.

L'analyse pourra être poussée plus loin en se posant la question du maintien de l'hème dans cette "poche". On peut poser l'hypothèse que ce maintien est réalisé par certains des acides aminés les plus voisins. Afin d'en établir l'inventaire on exploite une fonctionnalité qui permet, autour d'un atome ou d'un groupe désigné, de sélectionner automatiquement tous les atomes situés dans un rayon de dimension déterminée. On identifiera les acides aminés auxquels ils appartiennent en cliquant à l'écran. La comparaison des inventaires établis sur différentes molécules de globines révèle la présence systématique de quatre acides aminés (deux histidines, une phénylalanine, une valine).

Ceux-ci occupent dans la séquence de la molécule des positions éloignées (42,58,62,87) mais très comparables d'une molécule à l'autre. On conforte alors l'idée que la structure tridimensionnelle de la molécule a pour résultat de rapprocher dans l'espace des acides aminés très distants dans la séquence et que cette structure a un rapport étroit avec la fonction.

Montrer qu'une protéine est une structure dynamique, exemple de la carboxypeptidase

Il ne s'agit pas de déterminer la structure de la carboxypeptidase mais de comprendre les interactions enzyme-substrat. La visualisation en sphères montre que l'enzyme a une forme globuleuse. L'affichage de la séquence permet de découvrir l'existence d'une chaîne polypeptidique de 307 acides aminés.

Afin de repérer le site catalytique, on travaille sur l'image du complexe enzyme-substrat : le marquage coloré du substrat (en rouge par exemple) facilite le repérage et l'analyse du site catalytique. On découvre ainsi que le substrat se niche dans une poche de l'enzyme au fond d'un profond sillon. Une complémentarité de forme apparaît entre le substrat et certains acides aminés qui tapissent la "poche" catalytique. On admet que les éléments chimiques qui participent à la catalyse se situent à proximité du substrat. Il est possible de colorer tous les éléments situés dans un rayon de 0.2nm autour du substrat. Ce marquage permet le repérage et l'identification d'un certain nombre d'éléments chimiques : un atome de zinc, six acides aminés (HIS69, GLU72, HIS196, ARG145, TYR248 et GLN270).

La comparaison de deux images, l'enzyme seule d'une part et le complexe enzyme-substrat d'autre part, permet de préciser le mode d'interaction entre enzyme et substrat. L'analyse comparative des deux images est facilitée par le repérage et le marquage coloré des acides aminés du site. On observe des réarrangements de certains acides aminés lors de la fixation du substrat. Cette observation peut être précisée en effectuant les mesures des distances séparant

les éléments chimiques déplacés au cours de la formation du complexe. On montre ainsi que la TYR248 subit un déplacement de près de 1.2nm par rapport à l'atome de Zn. La conséquence de ce déplacement est la fermeture du site catalytique, point de départ de l'hydrolyse du substrat. On découvre ainsi que la structure d'une enzyme n'est pas rigide, elle subit des modifications conformationnelles indispensables à son activité.

BILAN, RÉFLEXIONS PÉDAGOGIQUES ET PERSPECTIVES

L'imagerie moléculaire est nouvelle, abondante dans de nombreux ouvrages et manuels. Il est donc indispensable de familiariser les utilisateurs (professeurs et élèves) avec ces représentations et les traitements dont elles résultent, afin de développer leurs capacités de lecture et d'interprétation de tels documents. L'expérience dans les classes montre que des activités pratiques sur ces données sont à la portée des élèves : elles leur permettent sur du matériel biologique non manipulable (les molécules) de mener des explorations et d'atteindre des objectifs méthodologiques. Sur des problématiques posées, et grâce à la visualisation, les élèves observent avec des angles de vue et des grossissements différents, des marquages, des colorations (comme avec le microscope), ils retiennent les représentations pertinentes, ils recueillent des informations par la mesure ou l'application de traitements. La richesse des données et la facilité d'accès favorise le travail sur des exemples diversifiés et contribue à l'établissement du caractère général de tel ou tel concept que l'on veut fonder.

Ces activités débouchent sur des repérages, comparaisons et schématisations qui peuvent améliorer les représentations mentales que les élèves ont des phénomènes étudiés.

BIBLIOGRAPHIE

DAUNE [1995]. *Biophysique moléculaire*. Inter Éditions.

FERMI PERUTZ [1984]. The crystal structure of human deoxyhaemoglobin. *Journal of molecular biology*, V175.

La recherche [1993]. Le sang, N° spécial 254.

La recherche [1991]. Le repliement des protéines. N° 230.

LEHNINGER. *Biochimie*. Flammarion médecine.

LUBERT STRYER. *Biochimie*. Médecine Science, Flammarion.

MC CALL [1985]. The crystal structure of d(GGGCCC)ADN. *Journal of molecular biology*, V183.

PHILLIPS [1980]. Structure and refinement of oxymyoglobin. *Journal of molecular biology*, V142.

Pour la science [1993]. Voir les molécules biologiques. N° 183.

Pour la science [1993]. Les protéines, modèles moléculaires. N° 194.

SINGER [1992]. *Gènes et génome*. Vigot.

Logiciel

RASMOL : Roger Sayle, Biomolecular structure, Glaxo Research and Development, Greenford, Middlesex