

## Le repliement tridimensionnel des biomolécules : un problème de satisfaction de contraintes

La connaissance de la conformation tridimensionnelle d'une protéine ou d'un acide nucléique est un pas indispensable vers la compréhension de son mécanisme d'action. Aussi, la détermination d'une nouvelle structure est toujours un événement attendu connaissant des retombées aussi bien dans les études fondamentales de la régulation et de l'évolution moléculaire que dans les travaux de manipulation des macromolécules à des fins médicales ou agronomiques. Ces structures tridimensionnelles ne peuvent être déterminées avec certitude que par des techniques physiques comme la cristallographie aux rayons X ou la résonance magnétique nucléaire. Pourtant, le nombre de structures élucidées de cette manière est encore très limité et les chercheurs ont le plus souvent recours à des études par modifications chimiques accompagnées d'une modélisation manuelle ou informatique. La construction de modèles moléculaires rigoureux à partir de ce type de données expérimentales est actuellement l'objet de recherches très actives. En l'absence de contraintes structurales fortes, l'espace conformationnel d'une biomolécule prend rapidement des dimensions gigantesques interdisant une exploration systématique par les méthodes numériques classiques de type dynamique moléculaire. Un groupe de biochimistes et d'informaticiens de l'université de Montréal vient de publier, dans la revue *Science*, le fruit d'une collaboration de plusieurs années sur ces problèmes de prédiction. Le nouveau logiciel, nommé MC-SYM (*Macromolecular structures by SYMbolic programming*), est la première application de la programmation symbolique et des techniques de satisfaction de contraintes, deux produits de l'intelligence artificielle, à un problème habituellement traité par des méthodes numériques. Dans cette nouvelle approche, la molécule à étudier est considérée

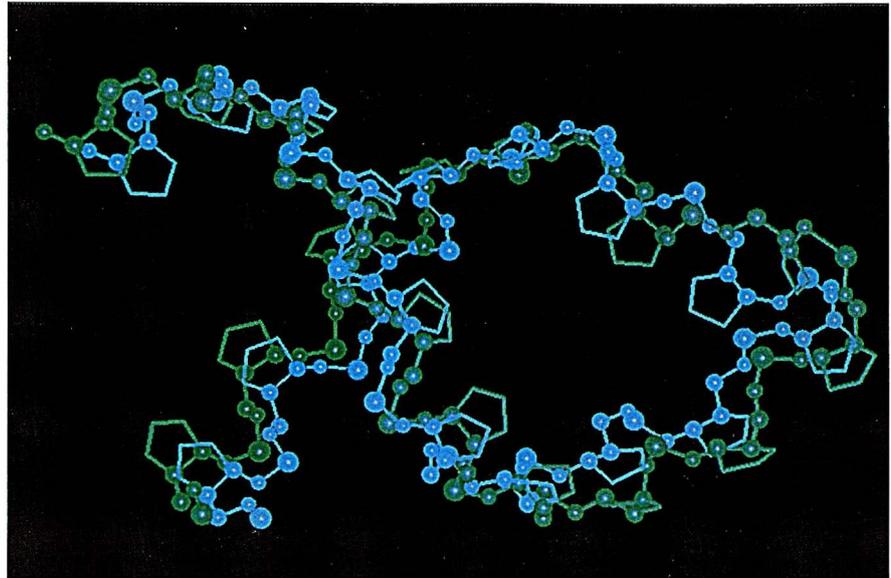


Figure 1. **Superposition des squelettes ribose-phosphate d'une boucle de l'anticodon de l'ARNt engendrée par le logiciel MC-SYM (vert) et de la structure (bleu).**

comme une liste de résidus (nucléotides ou acides aminés) dont chacun peut adopter un nombre prédéterminé de conformations, dicté principalement par son appartenance ou non à une structure secondaire. Le biopolymère est construit en ajoutant les résidus un à un et en vérifiant à chaque étape qu'un ensemble de contraintes spatiales fixé par le biochimiste est respecté. Ces contraintes sont, par exemple, des fermetures de boucles ou des liaisons hydrogène. De cette façon, une grande partie des structures potentielles est éliminée en cours de recherche et il n'est pas rare qu'un problème décrivant plus de  $10^{10}$  conformations soit rapidement réduit à quelques dizaines de solutions, pouvant toutes être évaluées énergétiquement. Ainsi, des molécules de structure encore indéterminée se trouvent à présent à portée d'une modélisation rigoureuse. L'efficacité de cette technique est démontrée par la comparaison des structures d'ARN

prédites aux structures déterminées par cristallographie : la précision des structures prédites est de l'ordre de 2 à 3 Å pour l'ensemble des atomes, ce qui est suffisant pour une étude fonctionnelle détaillée.

D. G.

1. Major F, Turcotte M, Gautheret D, Lapalme G, Fillion E, Cedergren R. The combination of symbolic and numerical computation for three-dimensional modeling of RNA. *Science* 1991 ; 253 : 1255-60.