

Modélisation de la structure secondaire d'ARN



Les structures secondaires

- Molécule d'ARN : suite linéaire de nucléotides = **structure primaire**

- 3. Les bases A, C, G, U peuvent former des **appariements** (liaisons hydrogène) - 4 niveaux d'appariement :
 - paires "**Watson-Crick**" : A—U et G—C
 - + paires "**Wobble**" (plus faible énergie) : G—U
 - + paires de très faible énergie : G—A ou C—A
 - + autres paires (rares) : **toutes les paires sont a priori possibles**. [Leontis N., Westhof E. 2001. Geometric nomenclature and classification of RNA base pairs. RNA 7:499-512.]

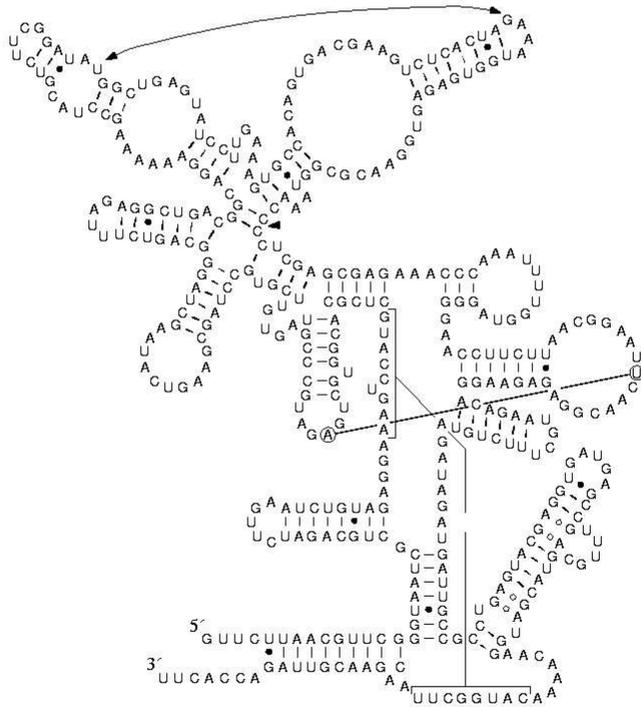


Vers une comparaison automatique

Ribonuclease P RNA *Bacillus subtilis* 168

Sequence : M13175, Reich, *et al.*, 1986 J. Biol. Chem. 261:7888
Structure : Harris, *et al.*, RNA (in press)

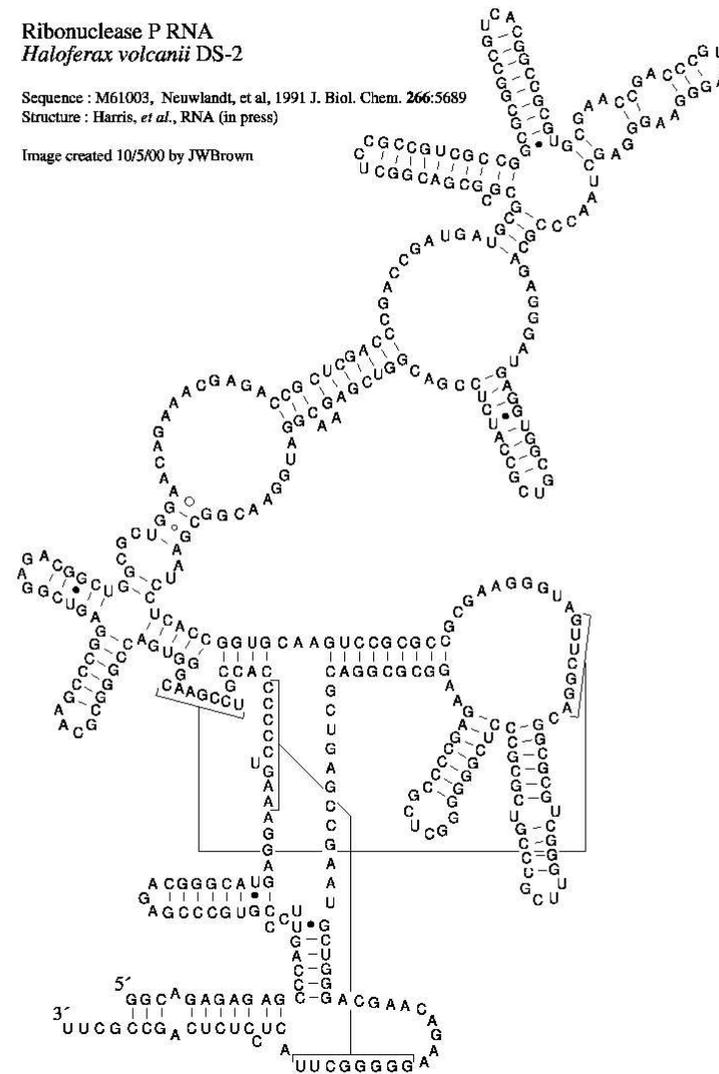
Image created 10/3/00 by JWBrown



Ribonuclease P RNA *Haloferax volcanii* DS-2

Sequence : M61003, Neuwandt, *et al.*, 1991 J. Biol. Chem. 266:5689
Structure : Harris, *et al.*, RNA (in press)

Image created 10/5/00 by JWBrown



Intérêts pour le biologiste

- structure \leftrightarrow fonction :
meilleure compréhension des **fonctions des molécules d'ARN**
- 4. secondaire \rightarrow tertiaire :
prédiction avec une meilleure efficacité **du repliement dans l'espace** d'une molécule
- 6. si séquences \neq mais SS identiques :
identification des motifs structuraux conservés (lors des processus de mutation par exemple)
- 8. recherche de **structures communes** à un ensemble de structures macromoléculaires
- 9. construction **d'arbres phylogénétiques**

Approche pour la comparaison de structures secondaires

Représentation arborescente **des**
structures



comparaison d'arborescences

Structures secondaires

→ Arborescences ordonnées étiquetées

Arborescences dont les sommets sont étiquetés :

1. A, U, G, C
2. A-U, U-A, G-C, C-G, G-U, U-G, G-A, A-G, ...
3. M, I, B₅, B₃, H, S
4. ...

selon la granularité du modèle arborescent choisi.



Arborescence fine

codage

modèle 1

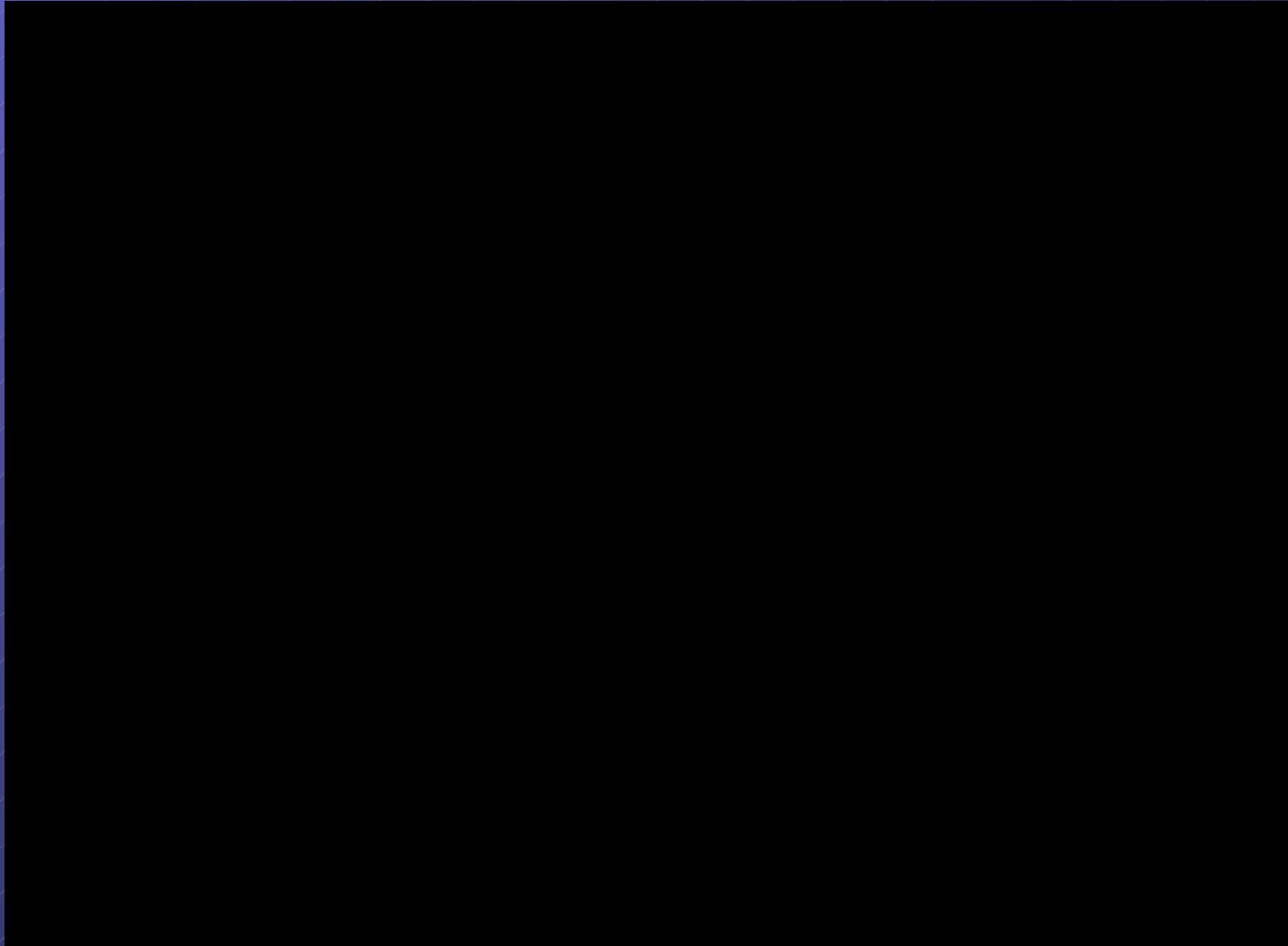


Arborescence fine

codage

modèle 2





Arborescence fine

codage

modèle 5



Arborescence grossière

modèle 6

- **B3'**, **B5'** : bulges
(renflements)
- **S** : stem (hélice)
- **I** : internal loop
(boucle interne)
- **M** : multibranch loop
(boucle multibranchée)
- **H** : hairpin
(épingle à cheveux)



Arborescence grossière

modèle 7



Utilisation de toutes les représentations

→ comparaison multi-échelles

