

Cours M2 BIBS - Séance 1
Repliement de l'ARN *in silico*

Yann Ponty

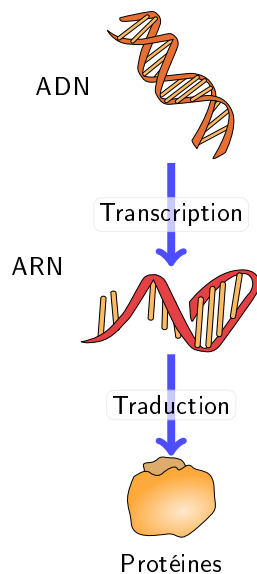
Bioinformatics Team
École Polytechnique/CNRS/INRIA AMIB – France

18 janvier 2010

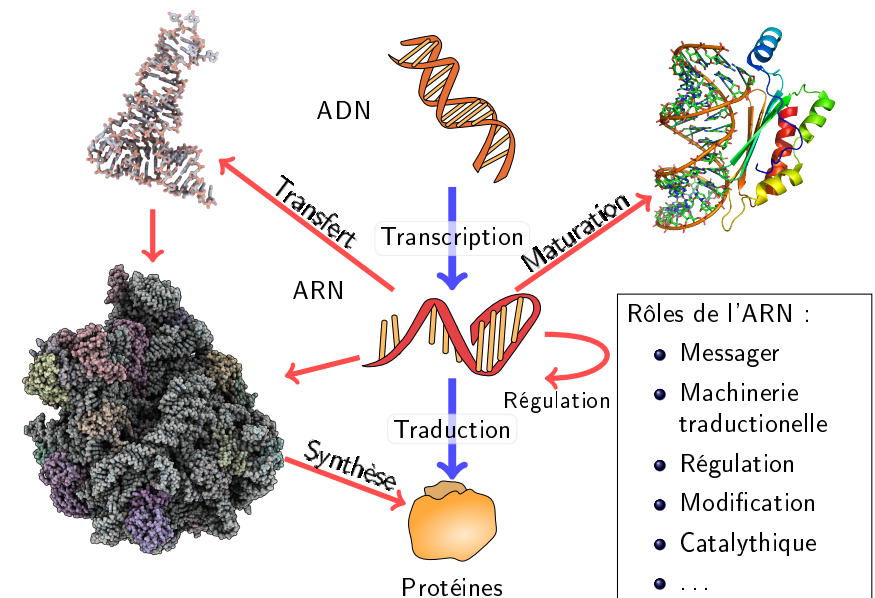
- 1 Introduction
 - Structure(s) de l'ARN
 - Simulation *in silico* du repliement
 - Performances et promesses

- 2 Principes algorithmique du repliement *ab initio*
 - Repliement *in silico*
 - Programmation dynamique
 - Implémentation

Dogme fondamental de la biologie moléculaire



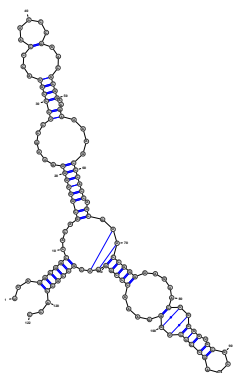
Dogme fondamental de la biologie moléculaire



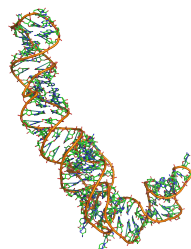
Structure de l'ARN

Trois¹ niveaux de représentation :

```
UUAGGGGGCCACAGC
GGUGGGGUUGCCUCC
CGUACCAUCCCGAA
CACGGAAGUAAGCC
CACCAGCGUUCGGG
GAGUACUGGAGUGCG
CGAGCCUCUGGGAAA
CCCGGUUCGCCGCA
CC
```



Structure primaire



Structure secondaire

Structure tertiaire

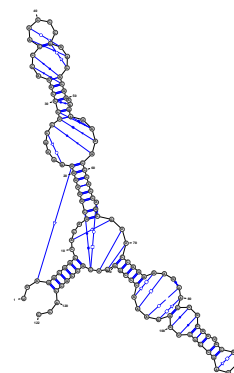
Source : 5s rRNA (PDBID : 1K73 :B)

1. Enfin, presque ...

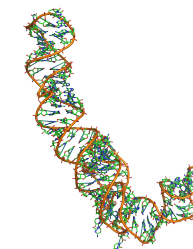
Structure de l'ARN

Trois¹ niveaux de représentation :

```
UUAGGGGGCCACAGC
GGUGGGGUUGCCUCC
CGUACCAUCCCGAA
CACGGAAGUAAGCC
CACCAGCGUUCGGG
GAGUACUGGAGUGCG
CGAGCCUCUGGGAAA
CCCGGUUCGCCGCA
CC
```



Structure primaire



Structure secondaire⁺

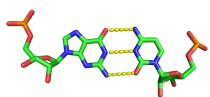
Structure tertiaire

Source : 5s rRNA (PDBID : 1K73 :B)

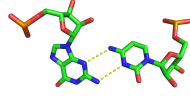
1. Enfin, presque ...

Ignorés par la structure secondaire

- Appariements non-canoniques :
Toute paire de base **autre que** {(A-U), (C-G), (G-U)}
Ou interagissant sur un bord non-standard (WC/WC-Cis) [LW01].

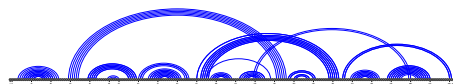


Paire CG canonique (WC/WC-Cis)



Paire CG non canonique (Sucre/WC-Trans)

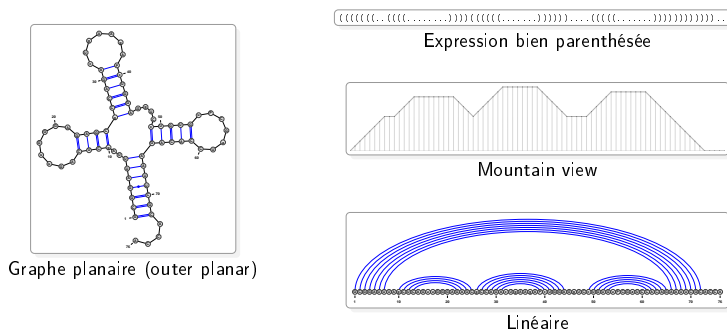
- Pseudonoëuds :



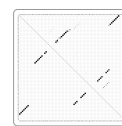
Structure pseudonoëud d'un Ribozyme du Groupe I (PDBID : 1Y0Q :A)

Plus expressif, mais repliement général *in silico* avec pseudonoëud :
⇒ NP-Complet [LP00] ...
... mais polynomial pour des classes restreintes [CDR+04].

Diversité de représentations pour une molécule omniprésente



Graphe planaire (outer planar)



Dot plot

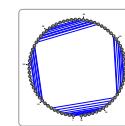


Diagramme de Feynman

Représentation différentes
mais
Structure combinatoire commune

Deux approches

Definition (Repliement ab initio)

Partant de la séquence, trouver la conformation minimisant une fonction d'énergie.

Avantages :

- Explication mécanique
- Complexité raisonnable $\mathcal{O}(n^3)/\mathcal{O}(n^2)$ temps/mémoire
- Exploration exhaustive

Limites :

- Pas de cinétique
- Pas d'info évolutive
- Performances limitées

Definition (Approche comparative)

Partant de plusieurs séquences homologues ou d'un alignement, trouver une conformation de score (énergie+alignement) élevé.

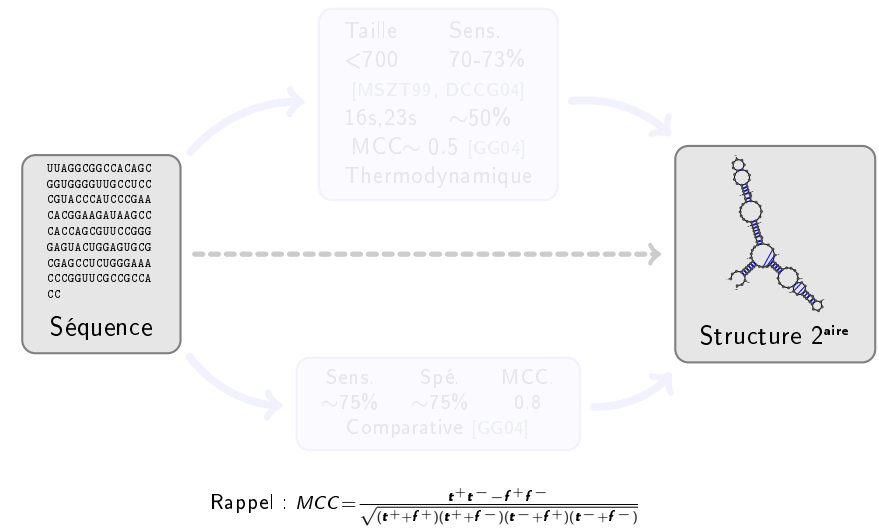
Avantages :

- Meilleures performances
- Affinement permanent

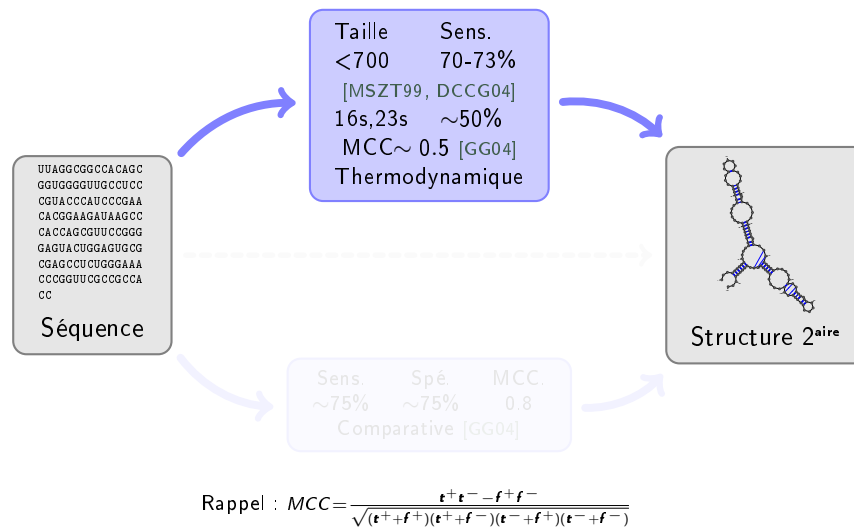
Limites :

- Complexité élevée
- Exploration non-exhaustive

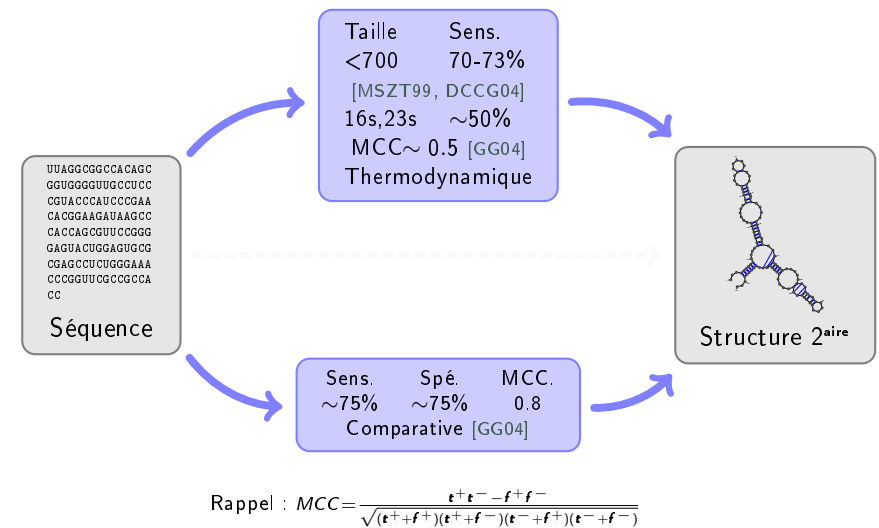
Performances



Performances



Performances



Futur (proche) : Vers une prédiction 3D

But : De la séquence à des modèles tri-dimensionnels!!!

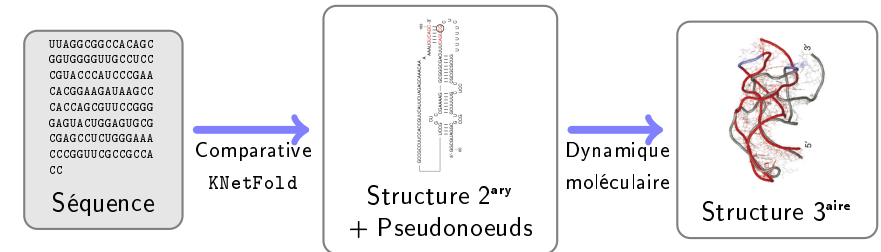
- Modèles comparatifs + Dynamique moléculaires : RNA2D3D [SYKB07]
- Pipeline MC-Fold/MC-sym [PM08]



Futur (proche) : Vers une prédiction 3D

But : De la séquence à des modèles tri-dimensionnels!!!

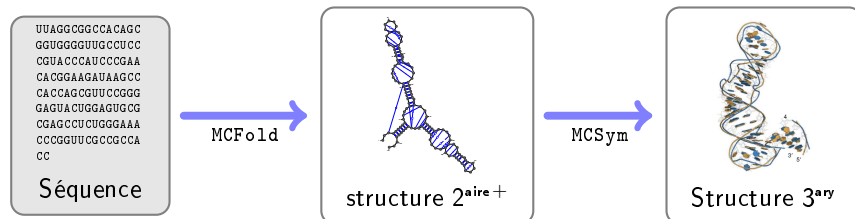
- Modèles comparatifs + Dynamique moléculaires : RNA2D3D [SYKB07]
- Pipeline MC-Fold/MC-sym [PM08]



Futur (proche) : Vers une prédiction 3D

But : De la séquence à des modèles tri-dimensionnels!!!

- Modèles comparatifs + Dynamique moléculaires : RNA2D3D [SYKB07]
- Pipeline MC-Fold/MC-sym [PM08]



Résumé

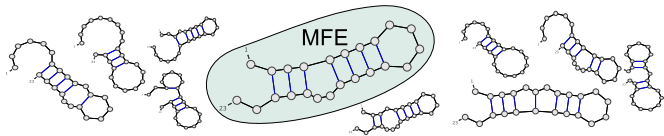
- 1 Introduction
 - Structure(s) de l'ARN
 - Simulation in silico du repliement
 - Performances et promesses
- 2 Principes algorithmique du repliement ab initio
 - Repliement in silico
 - Programmation dynamique
 - Implémentation

Formalisation du repliement

Repliement = Déterminer la struct. sec. native d'un ARN de seq. ω .

Principe thermodynamique

Structure native = Conformation d'énergie libre minimale (MFE).



- Conformations : Ensemble S_ω des structures secondaires compatibles avec la structure primaire ω (contrainte d'appariements).
- Fonction d'énergie E associant une valeur numérique molaire $E_{\omega,S}$ à tout couple séquence/conformation (ω, S) .
- Structure native : Conformation *fonctionnelle* de la molécule.
Remarques :
 - Pas nécessairement unique (Cinétique ou structures bi-stables)
 - Présence de pseudo-noeuds : Quelle est la structure native ?

Principe général

Programmation dynamique = Technique générale pour les problèmes d'optimisation.

Condition : Solution optimale pour P peut être reconstruite à partir de solutions pour des sous-problèmes strictes de P .

Bioinformatique :

- Espaces de solutions *discrets* (alignements, repliements)
- + Fonctions objectifs *additives* (score, énergie)
- ⇒ Schémas de programmation dynamique

Exemple : Alignement local (Smith/Waterman)

$$\begin{aligned}
 W(i, m+1) &= 0 \\
 W(n+1, j) &= 0 \\
 W(i, j) &= \max \left\{ \begin{array}{l} W(i+1, j+1) + m_{i,j} \\ W(i+1, j) + p_i \\ W(i, j+1) + p_d \end{array} \right.
 \end{aligned}$$

Modèle de Nussinov/Jacobson

Modèle de Nussinov/Jacobson (NJ)

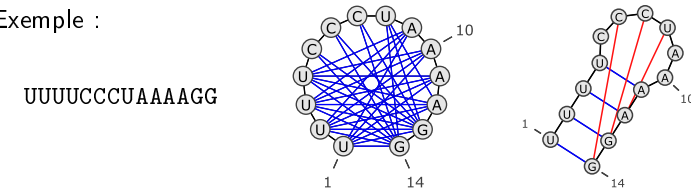
Plus proche voisins simple :

- Seuls les appariements contribuent à l'énergie
- Uniquement liaisons Watson/Crick (A/U, C/G) et Wobble (G/U)

$$\Rightarrow E_{\omega,S} = -\#Paires(S)$$

Repliement dans NJ \Leftrightarrow Maximisation du nombre de paires de bases.

Exemple :



Variante : Pondérer les paires selon leur nombre de liaisons hydrogènes
 $\Delta G(G \equiv C) = -3$ $\Delta G(A = U) = -2$ $\Delta G(G - U) = -1$

Détails algorithmiques

Un schéma fait intervenir des *classes* de sous-problèmes dont on sait calculer le score du *champion*.

Étant donné un schéma, deux étapes :

- Calcul matrices : Sauvegarde des meilleurs scores sur classes de sous-problèmes (Ordre inverse de celui induit par les dépendances).
- Remontée : Reconstitue le parcours ayant mené au meilleur score. (Parcours = Instance)

Complexité du calcul dépend alors :

- Taille de l'espace des sous-problèmes
- Nombre de sous-problèmes considérés ($\#$ Termes décomposition)

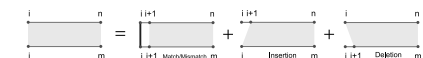
Exemple S/W :

$i : 1 \rightarrow n+1 \Rightarrow \Theta(n)$

$j : 1 \rightarrow m+1 \Rightarrow \Theta(m)$

Trois opération pour chaque sous-calcul

⇒ $\Theta(m.n)$ temps/mémoire



Propriétés des schémas

Propriétés requise d'un schéma :

- Validité : \forall sous-problème, la valeur obtenue doit être celle de la fonction objectif.

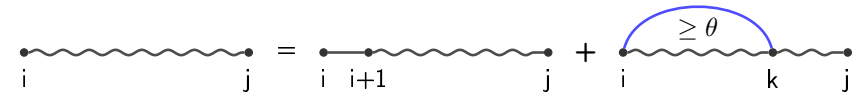
Preuve souvent assez technique.

Propriétés souhaitables d'un schéma :

- Complétude : Espace des solutions engendré par la décomposition. Des astuces algorithmiques peuvent *couper des branches*...
- Non-ambiguïté : Chaque solution est *engendrée* au plus une fois.

⇒ Possibilité d'énumérer l'espace des solutions.

Décomposition de Nussinov/Jacobson



$$N_{i,t} = 0, \quad \forall t \in [i, i + \theta]$$

$$N_{i,j} = \min \begin{cases} N_{i+1,j} & i \text{ non apparié} \\ \min_{k=i+\theta+1}^j E_{i,j} + N_{i+1,k-1} + N_{k+1,j} & i \text{ apparié à } k \end{cases}$$

Nussinov/Jacobson

| | C | G | G | A | U | A | C | U | U | C | U | U | A | G | A | C | G | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|
| C | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 9 | 9 | 11 | 14 | 14 |
| G | | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 7 | 9 | 11 | 11 | 11 |
| G | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 6 | 8 | 10 | 10 | 10 |
| A | | | | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 4 | 4 | 5 | 7 | 7 | 8 | 10 |
| U | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 4 | 5 | 7 | 7 | 8 | 10 |
| A | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 5 | 5 | 5 | 8 | 8 |
| C | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 5 | 5 | 8 | 8 |
| U | | | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 6 | 7 |
| U | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 5 | 7 |
| C | | | | | | | | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 |
| U | | | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 3 |
| U | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| A | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| G | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 |
| A | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 |
| C | | | | | | | | | | | | | | | | 0 | 0 | 0 |
| G | | | | | | | | | | | | | | | | | 0 | 0 |
| A | | | | | | | | | | | | | | | | | | 0 |

Nussinov/Jacobson

| | C | G | G | A | U | A | C | U | U | C | U | U | A | G | A | C | G | A |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|
| C | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 9 | 9 | 11 | 14 | 14 |
| G | | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 7 | 9 | 11 | 11 | 11 |
| G | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 6 | 8 | 10 | 10 | 10 |
| A | | | | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 4 | 4 | 5 | 7 | 7 | 8 | 10 |
| U | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 4 | 5 | 7 | 7 | 8 | 10 |
| A | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 5 | 5 | 5 | 8 | 8 |
| C | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 5 | 5 | 8 | 8 |
| U | | | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 6 | 7 |
| U | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 5 | 7 |
| C | | | | | | | | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 |
| U | | | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 3 |
| U | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| A | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 |
| G | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 |
| A | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 |
| C | | | | | | | | | | | | | | | | 0 | 0 | 0 |
| G | | | | | | | | | | | | | | | | | 0 | 0 |
| A | | | | | | | | | | | | | | | | | | 0 |

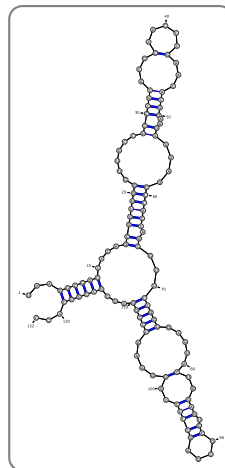
| | C | G | G | A | U | A | C | U | U | C | U | U | A | G | A | C | G | A | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|--|
| | (| (| (| . | . | . |) | . | (| (| . | . | . |) |) |) |) | . | |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 9 | 9 | 11 | 14 | 14 | |
| G | | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 7 | 9 | 11 | 11 | 11 | |
| G | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 6 | 8 | 10 | 10 | 10 | |
| A | | | | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 4 | 4 | 4 | 5 | 7 | 7 | 8 | 10 | |
| U | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 4 | 5 | 7 | 7 | 8 | 10 | 10 | |
| A | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 5 | 5 | 5 | 8 | 8 | 8 | |
| C | | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 5 | 5 | 8 | 8 | 8 | |
| U | | | | | | | | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 6 | 7 | 7 | |
| U | | | | | | | | | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 5 | 7 | 7 | |
| C | | | | | | | | | | 0 | 0 | 0 | 0 | 3 | 3 | 5 | 5 | 5 | |
| U | | | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 3 | |
| U | | | | | | | | | | | | 0 | 0 | 0 | 0 | 1 | 2 | 2 | |
| A | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| G | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | |
| A | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | |
| C | | | | | | | | | | | | | | | | 0 | 0 | 0 | |
| G | | | | | | | | | | | | | | | | | 0 | 0 | |
| A | | | | | | | | | | | | | | | | | | 0 | |

| | C | G | G | A | U | A | C | U | U | C | U | U | A | G | A | C | G | A | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|----|----|--|
| | (| (| (| . | . | . |) | . | (| (| . | . | . |) |) |) | . | | |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 9 | 9 | 11 | 14 | 14 | |
| G | | 0 | 0 | 0 | 0 | 0 | 3 | 4 | 4 | 6 | 6 | 6 | 6 | 7 | 9 | 11 | 11 | 11 | |
| G | | | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 6 | 8 | 10 | 10 | 10 | |
| A | | | | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 4 | 4 | 4 | 5 | 7 | 7 | 8 | 10 | |
| U | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 4 | 5 | 7 | 7 | 8 | 10 | 10 | |
| A | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 5 | 5 | 5 | 8 | 8 | 8 | |
| C | | | | | | | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 5 | 5 | 8 | 8 | 8 | |
| U | | | | | | | | 0 | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 6 | 7 | 7 | |
| U | | | | | | | | | 0 | 0 | 0 | 2 | 3 | 5 | 5 | 5 | 7 | 7 | |
| C | | | | | | | | | | 0 | 0 | 0 | 0 | 3 | 3 | 5 | 5 | 5 | |
| U | | | | | | | | | | | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 3 | |
| U | | | | | | | | | | | | 0 | 0 | 0 | 0 | 1 | 2 | 2 | |
| A | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| G | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | |
| A | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | |
| C | | | | | | | | | | | | | | | | 0 | 0 | 0 | |
| G | | | | | | | | | | | | | | | | | 0 | 0 | |
| A | | | | | | | | | | | | | | | | | | 0 | |

Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

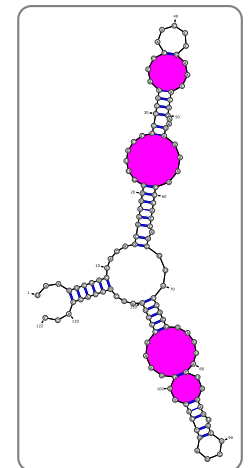
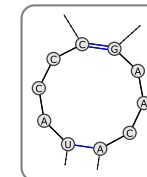
Déterminées expérimentalement
+ Interpolation pour grandes boucles.

Meilleure résultats grâce à la prise en compte de l'empilement.

Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

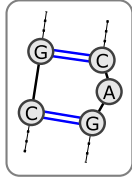
Déterminées expérimentalement
+ Interpolation pour grandes boucles.

Meilleure résultats grâce à la prise en compte de l'empilement.

Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

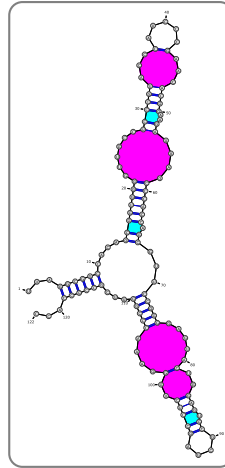
- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

Déterminées expérimentalement
+ Interpolation pour grandes boucles.

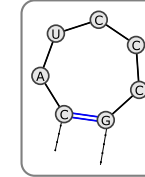
Meilleure résultats grâce à la prise en compte de l'empilement.



Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

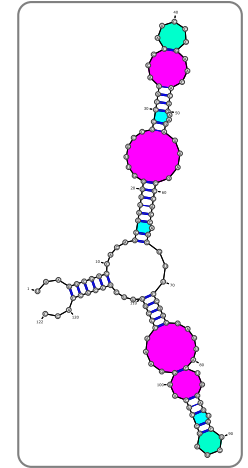
- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

Déterminées expérimentalement
+ Interpolation pour grandes boucles.

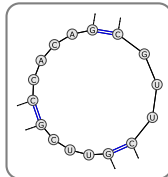
Meilleure résultats grâce à la prise en compte de l'empilement.



Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

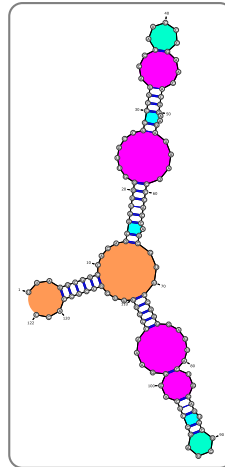
- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

Déterminées expérimentalement
+ Interpolation pour grandes boucles.

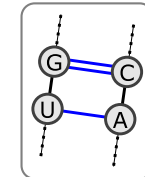
Meilleure résultats grâce à la prise en compte de l'empilement.



Modèle de Turner

Basée sur décomposition **non-ambiguë** en **boucles** de la structure 2^{aire} :

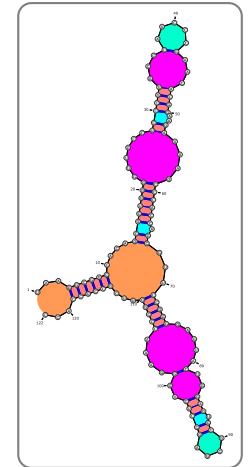
- Boucles internes
- Renflements
- Boucles terminales
- Boucles multiples
- Empilements



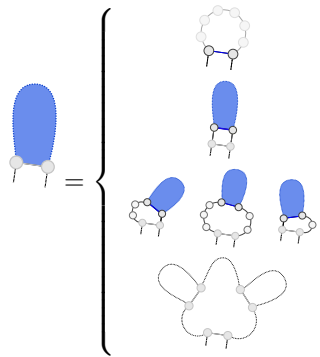
Énergies libres ΔG des boucles dépendent des bases, asymétrie, bases *libres* (dangle) ...

Déterminées expérimentalement
+ Interpolation pour grandes boucles.

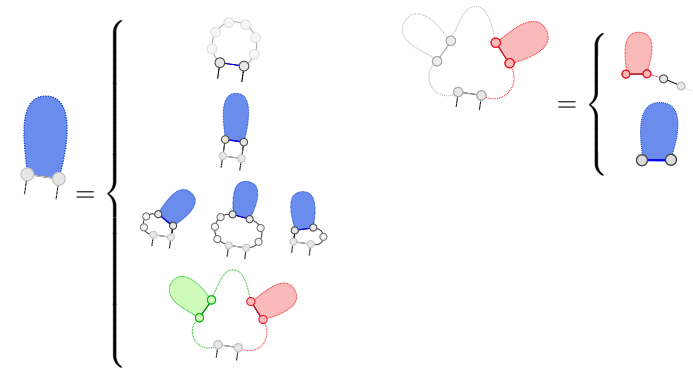
Meilleure résultats grâce à la prise en compte de l'empilement.



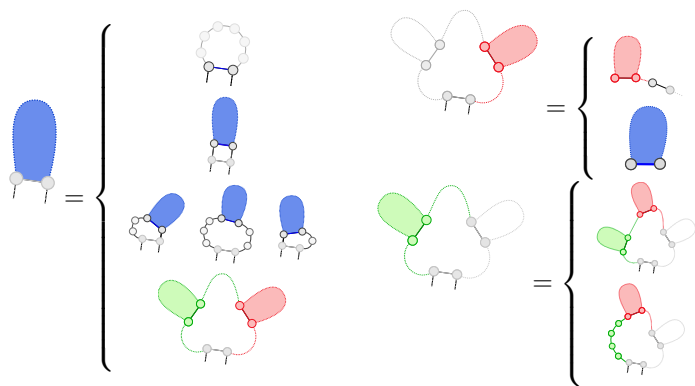
MFE DP equations



MFE DP equations

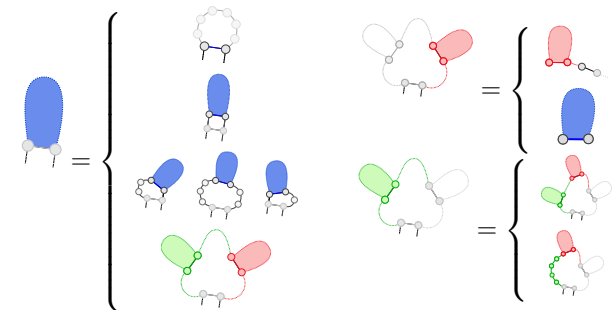


MFE DP equations



MFold/Unafold

- $E_H(i, j)$: Energie d'une boucle terminale *fermée* par une paire (i, j)
- $E_{BI}(i, j)$: Energie d'un renflement ou boucle interne *fermée* par une paire (i, j)
- $E_S(i, j)$: Energie d'un empilement $(i, j)/(i + 1, j - 1)$
- a, c, b : Pénalité pour boucle multiple, hélice et non-appariées dans multiboucle.





- $E_H(i, j)$: Energie d'une boucle terminale *fermée* par une paire (i, j)
- $E_{BI}(i, j)$: Energie d'un renflement ou boucle interne *fermée* par une paire (i, j)
- $E_S(i, j)$: Energie d'un empiement $(i, j)/(i + 1, j - 1)$
- a, c, b : Pénalité pour boucle multiple, hélice et non-appariées dans multiboucle.


$$\mathcal{M}'(i, j) = \text{Min} \left\{ \begin{array}{l} E_H(i, j) \\ E_S(i, j) + \mathcal{M}'(i + 1, j - 1) \\ \text{Min}(E_{BI}(i, i', j', j) + \mathcal{M}'(i', j')) \\ a + c + \text{Min}(\mathcal{M}(i + 1, k - 1) + \mathcal{M}^1(k, j - 1)) \end{array} \right\}$$


$$\mathcal{M}(i, j) = \text{Min} \{ \text{Min}(\mathcal{M}(i, k - 1), b(k - 1)) + \mathcal{M}^1(k, j) \}$$


$$\mathcal{M}^1(i, j) = \text{Min} \{ b + \mathcal{M}^1(i, j - 1), c + \mathcal{M}'(i, j) \}$$


 B. A. Shapiro, Y. G. Yingling, W. Kasprzak, and E. Bindewald.
Bridging the gap in rna structure prediction.
Curr Opin Struct Biol, 17(2) :157-165, Apr 2007.


 A. Condon, B. Davy, B. Rastegari, S. Zhao, and F. Tarrant.
Classifying RNA pseudoknotted structures.
Theoretical Computer Science, 320(1) :35-50, 2004.


 K. Doshi, J. J. Cannone, C. Cobaugh, and R. R. Gutell.
Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for rna secondary structure prediction.
BMC Bioinformatics, 5(1) :105, 2004.

 P. Gardner and R. Giegerich.
A comprehensive comparison of comparative rna structure prediction approaches.
BMC Bioinformatics, 5(1) :140, 2004.

 R. B. Lyngsø and C. N. S. Pedersen.
RNA pseudoknot prediction in energy-based models.
Journal of Computational Biology, 7(3-4) :409-427, 2000.

 N. Leontis and E. Westhof.
Geometric nomenclature and classification of RNA base pairs.
RNA, 7 :499-512, 2001.

 D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner.
Expanded sequence dependence of thermodynamic parameters improves prediction of rna secondary structure.
Journal of Molecular Biology, 288(5) :911-940, May 1999.

 M. Parisien and F. Major.
The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data.
Nature, 452(7183) :51-55, 2008.

Exercice : Implémentation en python de Nussinov-Jacobson