

Prédiction à gros grain de la structure tridimensionnelle de l'ARN

Alexis Lamiable
D. Barth, A. Denise, F. Quessette, S. Vial, É. Westhof

Journées ARENA 2010

PRiSM (UVSQ) & LRI (Paris-Sud 11)

Plan

1 Introduction

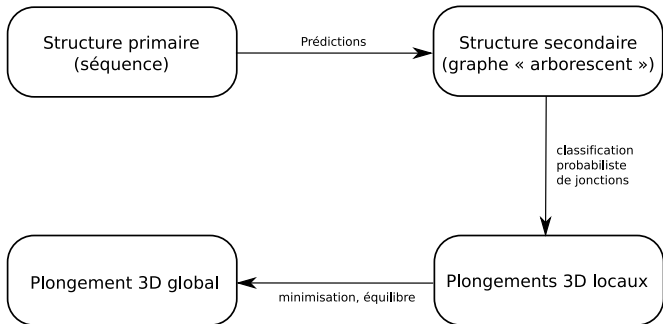
2 Formes locales

3 Forme globale

4 Conclusion

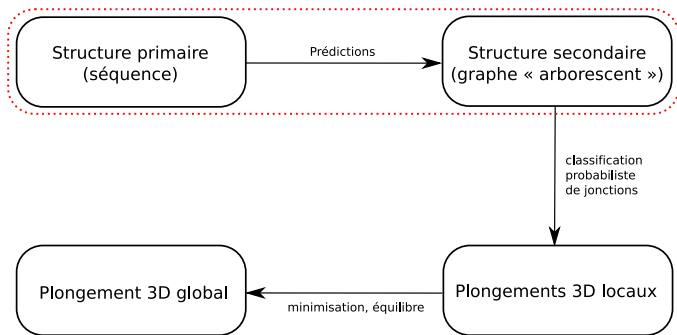
Projet PASAPAS

Prédiction de l'Architecture Spatiale d'un ARN à Partir de sa Séquence



Projet PASAPAS

Prédiction de l'Architecture Spatiale d'un ARN à Partir de sa Séquence



Introduction

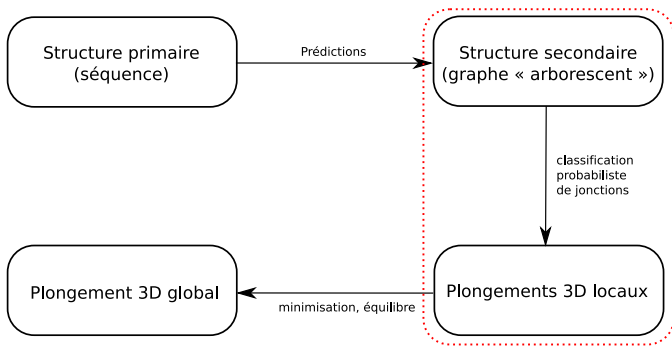
Formes locales

Forme globale

Conclusion

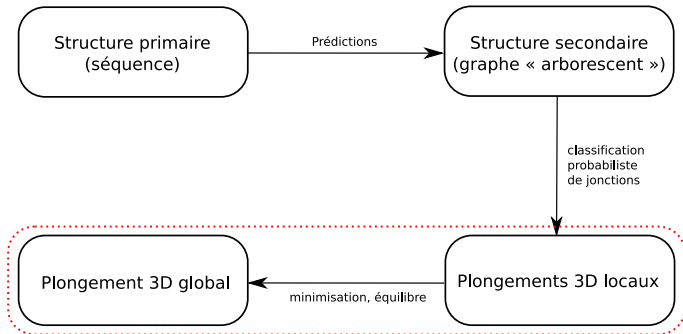
Projet PASAPAS

Prédiction de l'Architecture Spatiale d'un ARN à Partir de sa Séquence



Projet PASAPAS

Prédiction de l'Architecture Spatiale d'un ARN à Partir de sa Séquence



Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

Formes locales

Forme globale

Conclusion

Plan

1 Introduction

2 Formes locales

3 Forme globale

4 Conclusion

Jonctions triples

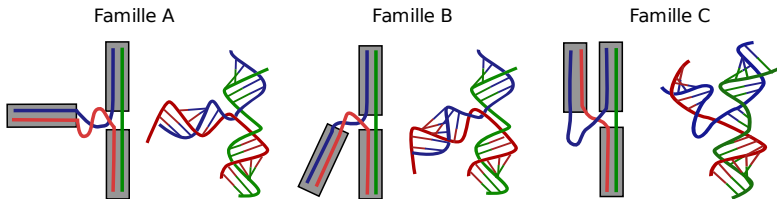
Lescoute and Westhof, *Topology of threeway junctions in folded RNAs*, RNA 12, 2006

Introduction

Formes locales

Forme globale

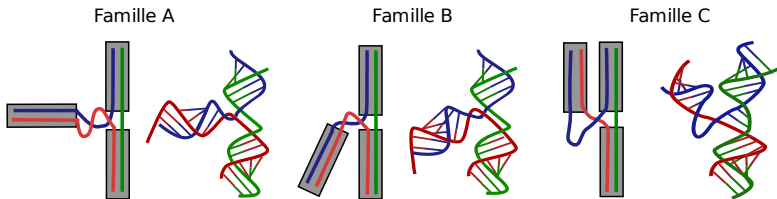
Conclusion



+ une famille "Y" sans stacking

Jonctions triples

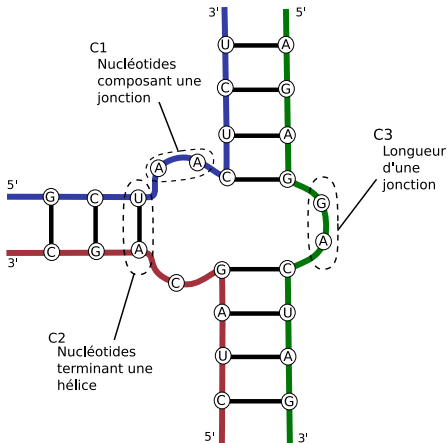
Lescoute and Westhof, *Topology of threeway junctions in folded RNAs*, RNA 12, 2006



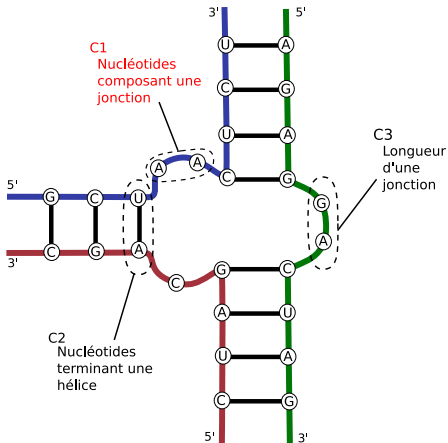
+ une famille "Y" sans stacking

Avec seulement la structure secondaire ?

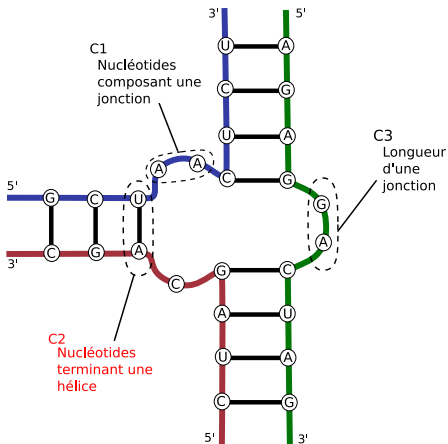
Critères utilisés



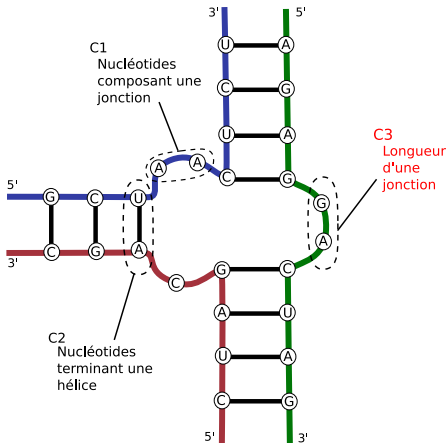
Critères utilisés

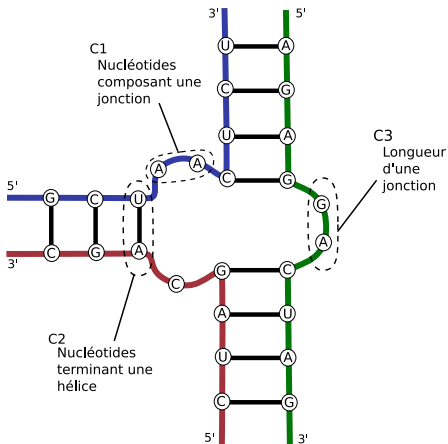


Critères utilisés



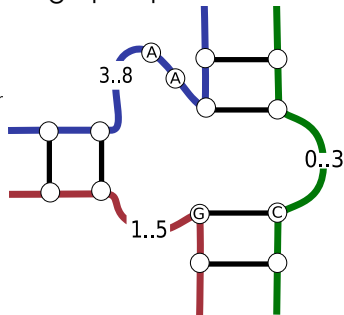
Critères utilisés





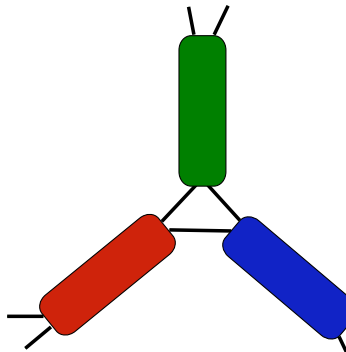
Critères utilisés

Comparaison avec un graphe patron



Prédiction en deux temps

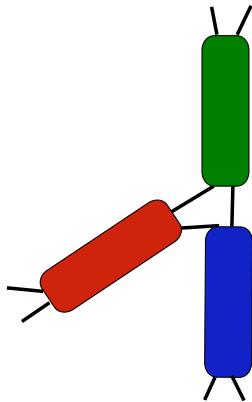
Comment prédit-on la famille ?



Prédiction en deux temps

Comment prédit-on la famille ?

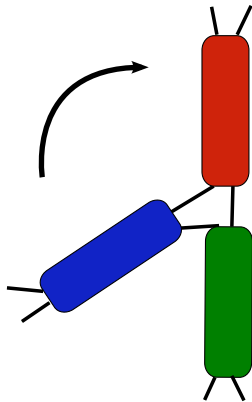
- On décide du stacking (3 possibilités)



Prédiction en deux temps

Comment prédit-on la famille ?

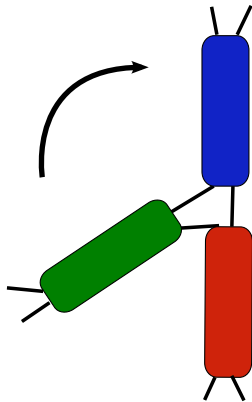
- On décide du stacking (3 possibilités)



Prédiction en deux temps

Comment prédit-on la famille ?

- On décide du stacking (3 possibilités)



Prédiction en deux temps

Comment prédit-on la famille ?

- On décide du stacking (3 possibilités)
- On décide de l'orientation de la troisième hélice (3 possibilités)

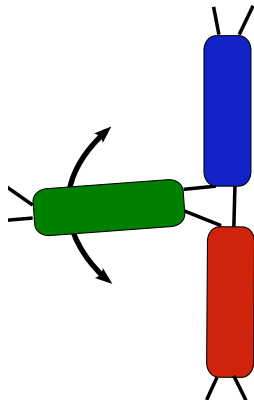


Tableau de scores

16S H25-25-26a

Stack.	Fam.	C1	C2	C3	Score
1.75	A	0.00	1.40	3.00	1.76
1.75	B	0.00	1.00	0.00	0.28
1.75	C	0.00	0.25	-3.00	0.10
2.33	A	2.50	1.40	-2.00	1.16
2.33	B	0.00	1.00	1.00	0.48
2.33	C	1.50	0.25	2.00	0.66
0.75	A	2.00	1.40	-1.00	0.49
0.75	B	0.00	1.00	5.00	0.33
0.75	C	0.50	0.25	1.00	0.16

Tableau de scores

16S H25-25-26a

Stack.	Fam.	C1	C2	C3	Score
1.75	A	0.00	1.40	3.00	1.76
1.75	B	0.00	1.00	0.00	0.28
1.75	C	0.00	0.25	-3.00	0.10
2.33	A	2.50	1.40	-2.00	1.16
2.33	B	0.00	1.00	1.00	0.48
2.33	C	1.50	0.25	2.00	0.66
0.75	A	2.00	1.40	-1.00	0.49
0.75	B	0.00	1.00	5.00	0.33
0.75	C	0.50	0.25	1.00	0.16

Introduction

Formes locales

Forme globale

Conclusion

30 jonctions étudiées par Lescoute et Westhof

- Première position : 70 %
- Dans les 3 premiers : 87 %

72 jonctions supplémentaires (FR3D)

- Tout compris : 51 %, 68 %
- Sans les Y : 65 % et 87 %

La suite

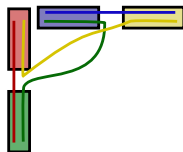
- Caractériser la "famille" Y
- 2-way : Kink-turns et C-loops
- 4-way :



Famille H



Famille X



Famille T

+ "hors famille"

Plan

1 Introduction

2 Formes locales

3 Forme globale

4 Conclusion

Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

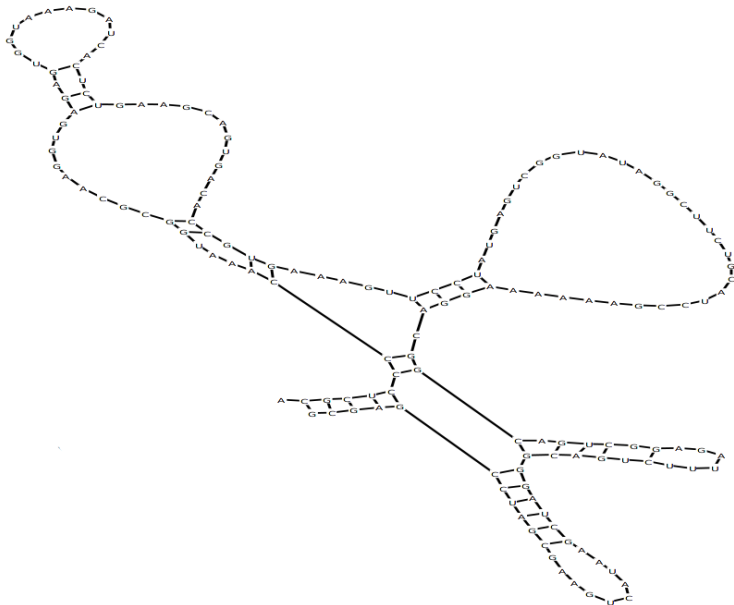
Introduction

Formes locales

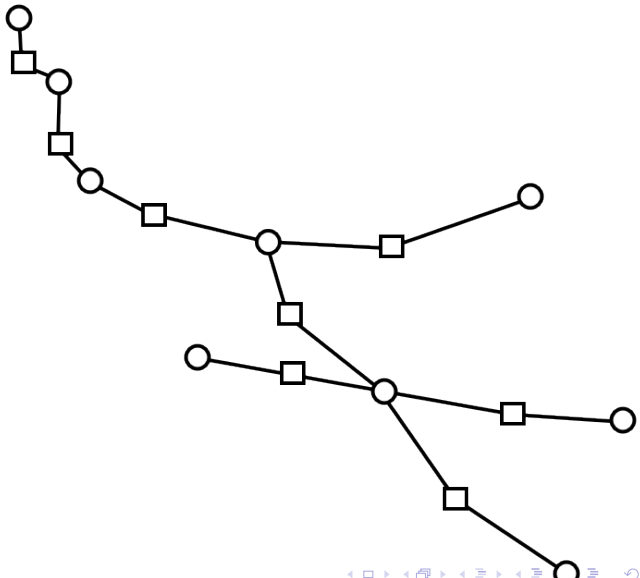
Forme globale

Conclusion

Extraction du squelette



Extraction du squelette



Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

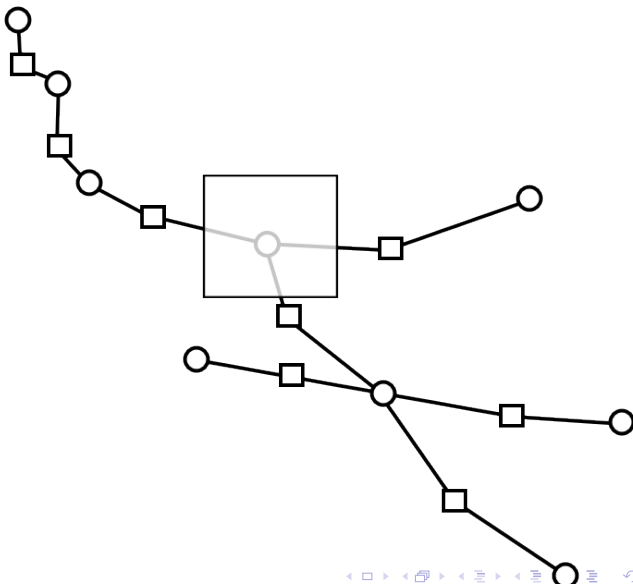
Introduction

Formes locales

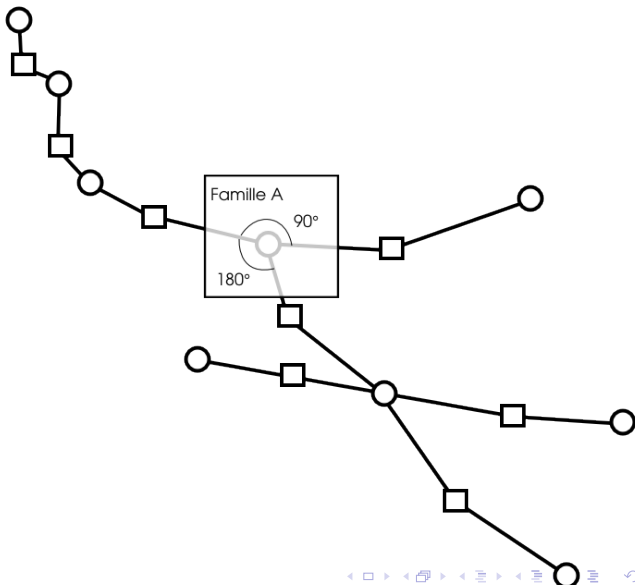
Forme globale

Conclusion

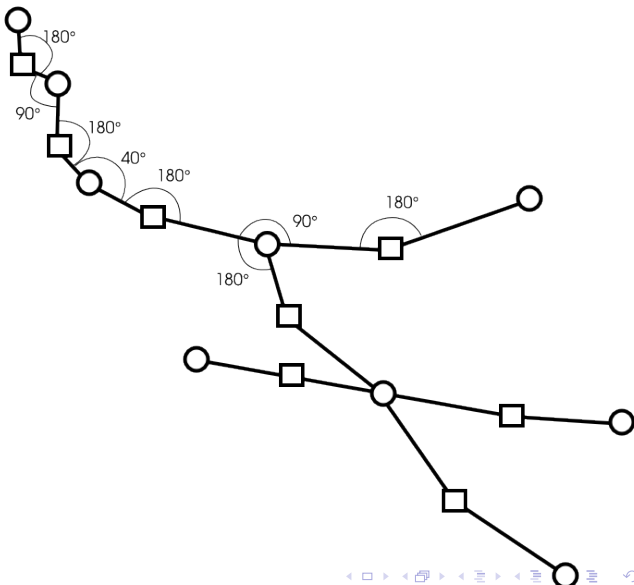
Classification



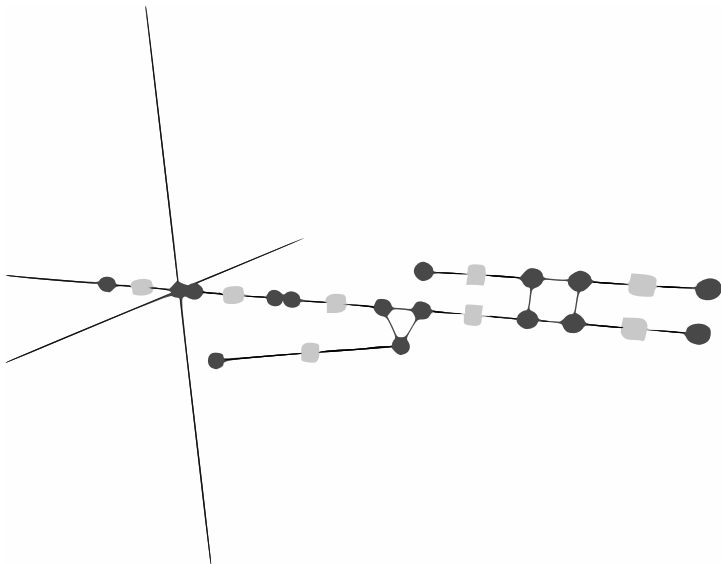
Classification



Classification



Positionnement initial



Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

Formes locales

Forme globale

Conclusion

“Vraie” forme



Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

Formes locales

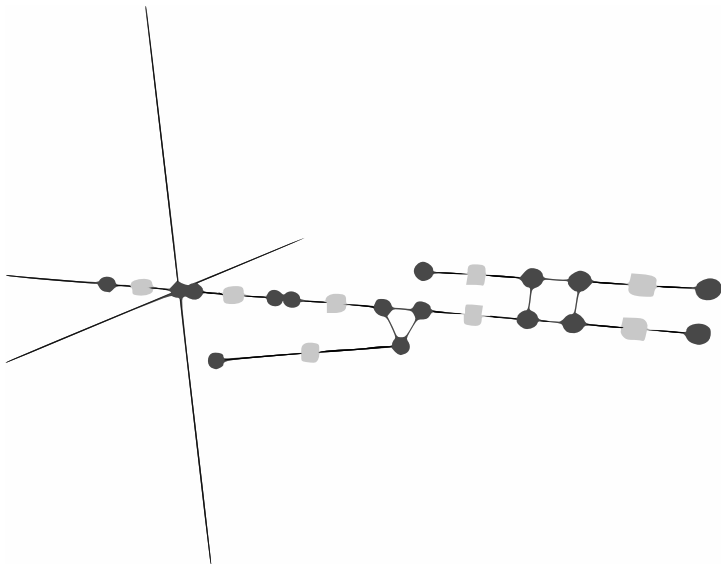
Forme globale

Conclusion

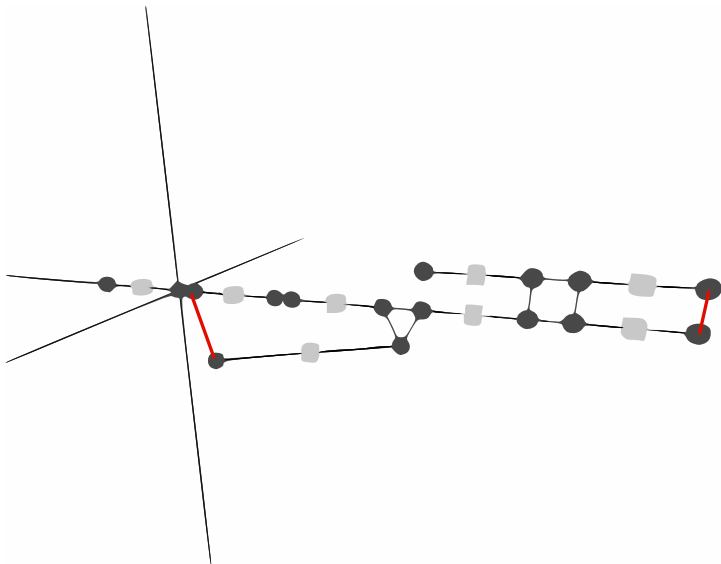
“Vraie” forme



Positionnement initial



Positionnement initial



Tertiaires → repliement

Respect de contraintes :

- Longueur et torsion des hélices
- Longueur des liaisons tertiaires
- Angles entre hélices induits par les jonctions

Tertiaires → repliement

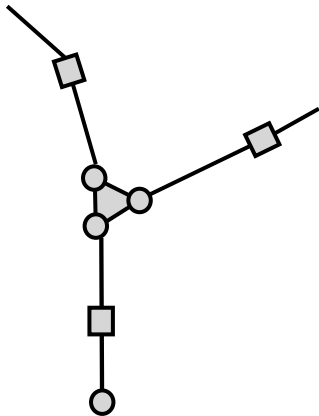
Respect de contraintes :

- Longueur et torsion des hélices
- Longueur des liaisons tertiaires
- Angles entre hélices induits par les jonctions

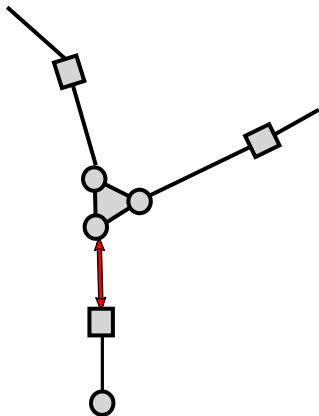
Algorithme de dessin de graphe

- Arêtes représentées par des ressorts
- Déplacement successifs
- Minimisation des forces exercées par les ressorts

Les ressorts utilisés

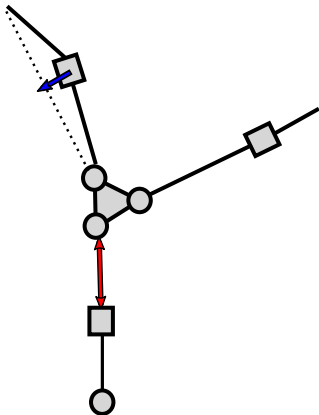


Les ressorts utilisés



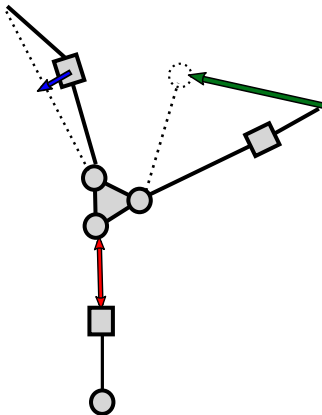
- Longueur des hélices

Les ressorts utilisés



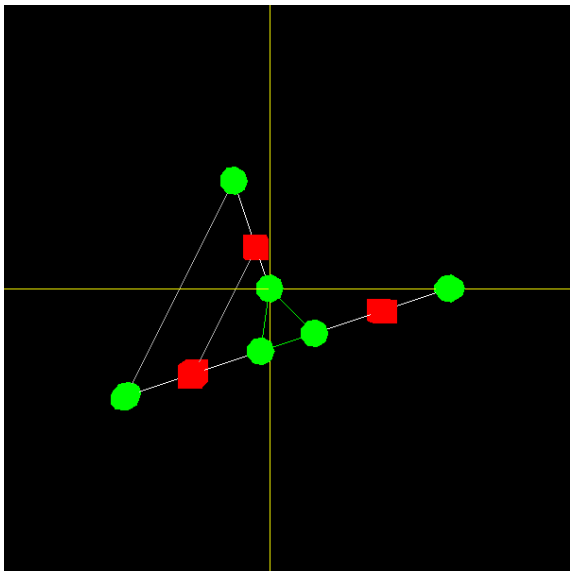
- Longueur des hélices
- Torsion des hélices

Les ressorts utilisés

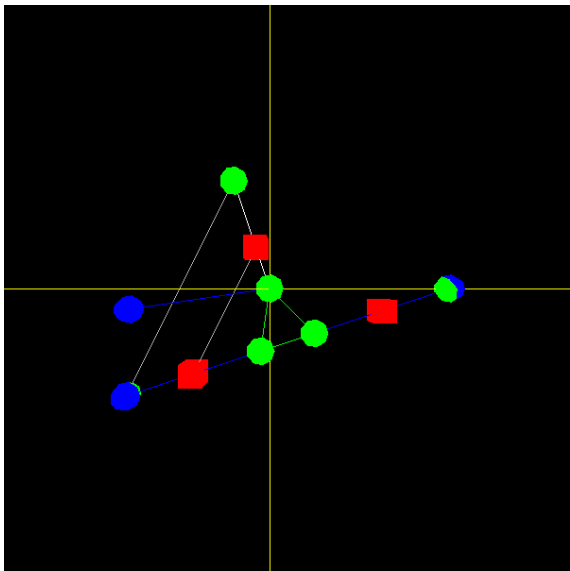


- Longueur des hélices
- Torsion des hélices
- Angles des jonctions

Repliement



Repliement



Repliement

Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

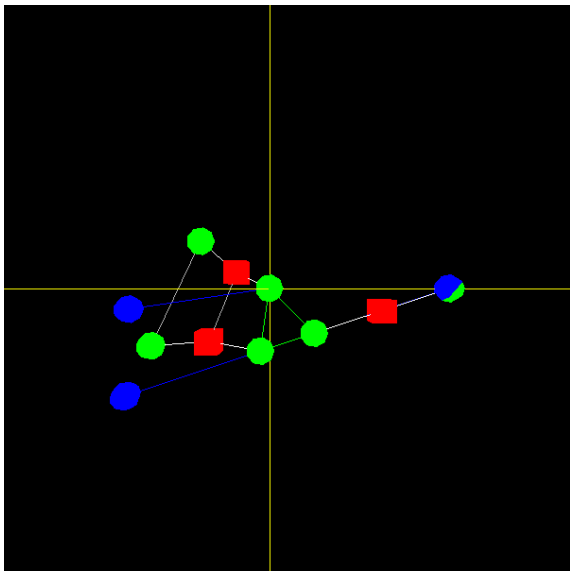
Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

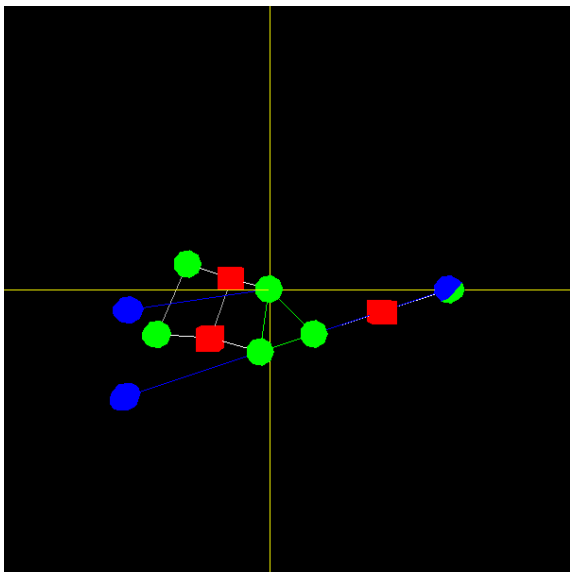
Formes locales

Forme globale

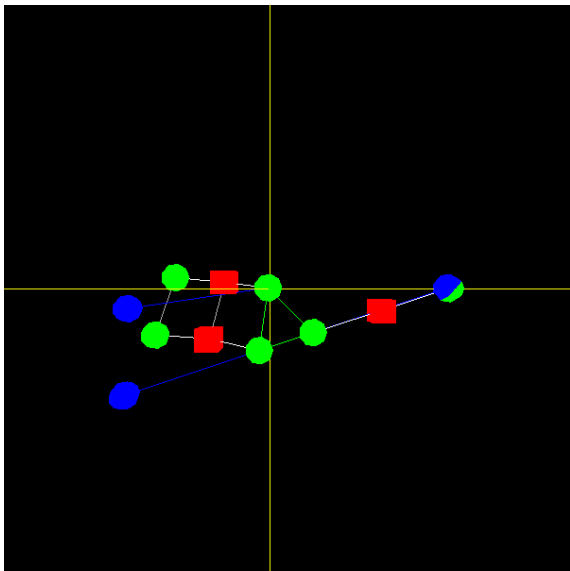
Conclusion



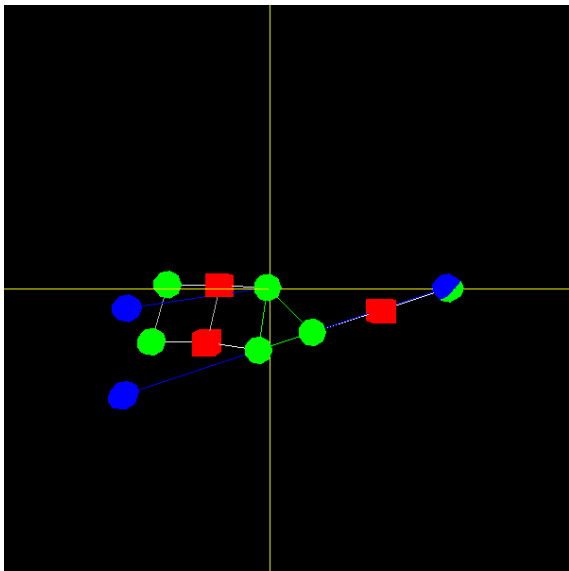
Repliement



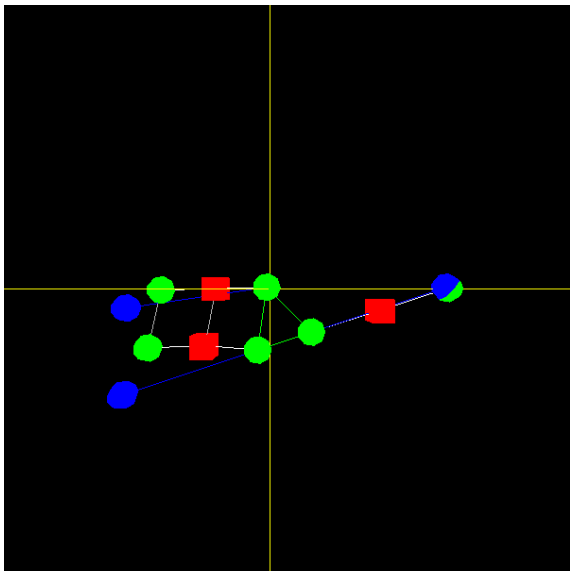
Repliement



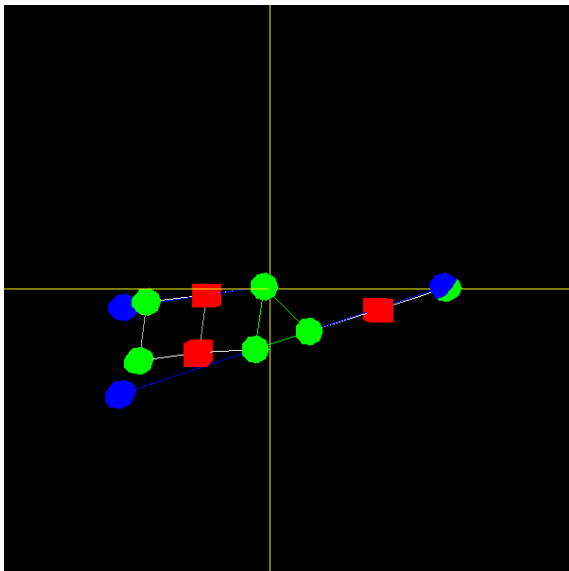
Repliement



Repliement



Repliement



Repliement

Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

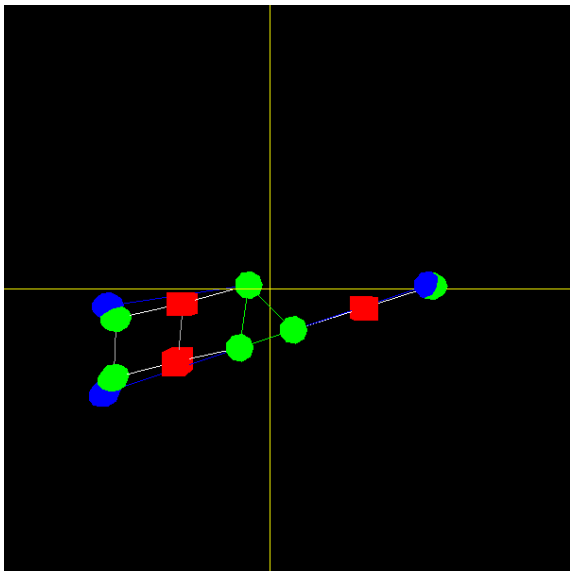
Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

Formes locales

Forme globale

Conclusion



Découverte des liaisons tertiaires

- tertiaire \rightarrow repliement
- repliement \rightarrow équilibre entre plusieurs forces
- somme de ces forces \rightarrow coût pour un sommet
- tertiaire = gain + coût

\Rightarrow Trouver les tertiaires qui minimisent le coût des sommets

- Fonction de cout :
somme des logs des couts des sommets
- Recherche exhaustive :
minimum global pour de petits exemples
- Algorithme évolutionniste
 - converge sur l'optimum global
 - ou bien hésite sur des solutions proches en terme de forme 3D.
 - à valider sur de vrais exemples

Limites de l'optimisation globale

- Ne tient pas compte de la dynamique de repliement
- Permet de pénaliser un sommet pour favoriser le plus grand nombre

⇒ on cherche une optimisation où chaque sommet est égoïste

Théorie des jeux

- Plusieurs joueurs
- Plusieurs stratégies pour chaque joueur
- Une fonction de gain pour chaque joueur
- Le gain dépend du choix de tous
- Optimisation locale, égoïste

Exemples : dilemme du prisonnier, pierre-papier-ciseaux

Équilibre de Nash

Un équilibre de Nash :

- Un choix de stratégie pour chaque joueur
- Si personne ne change sauf le joueur i , le joueur i n'a pas intérêt à changer

Un équilibre mixte :

- Un choix *probabilisé* de stratégies pour chaque joueur
- Ex : pierre-papier-ciseaux

Exemple : guerre des sexes

Un couple décide de sa sortie du soir

- la femme voudrait voir un match de football
- l'homme voudrait aller à l'opéra
- ils préfèrent être ensemble

F – H	Opéra	Match
Opéra	(2, 3)	(0, 0)
Match	(1, 1)	(3, 2)

Exemple : guerre des sexes

Un couple décide de sa sortie du soir

- la femme voudrait voir un match de football
- l'homme voudrait aller à l'opéra
- ils préfèrent être ensemble

F – H	Opéra	Match
Opéra	(2, 3)	(0, 0)
Match	(1, 1)	(3, 2)

Découverte d'un équilibre de Nash

- Un équilibre de Nash (mixte) existe toujours
- Trouver un équilibre de Nash est PPAD-complet
- PPAD-complet est probablement difficile (Papadimitriou)

Découverte d'un équilibre de Nash

Introduction

Formes locales

Forme globale

Conclusion

- Un équilibre de Nash (mixte) existe toujours
- Trouver un équilibre de Nash est PPAD-complet
- PPAD-complet est probablement difficile (Papadimitriou)

Algorithme de Sastry :

- Vecteurs de probabilité de choix de stratégie
- À chaque étape, choix de stratégies
- Puis mise à jour du vecteur de probabilité

Preuves de convergence

Notre jeu

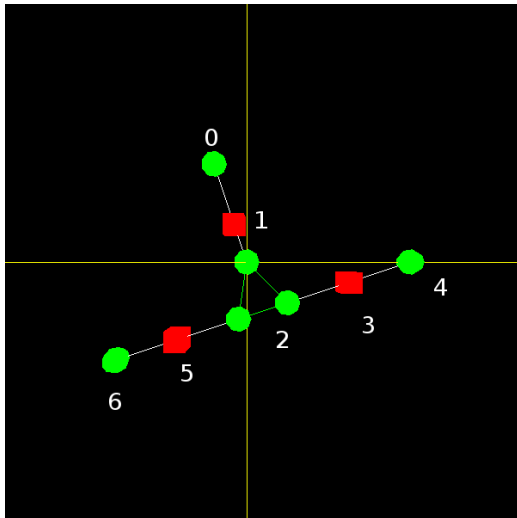
- Un joueur : un sommet (jonction ou hélice)
- Une stratégie : une tertiaire avec un autre joueur, ou aucune tertiaire
- Le coût : le coût de repliement - le gain des tertiaires

Esquisse d'algorithme

P_i : vecteur de probabilité des stratégies du joueur i

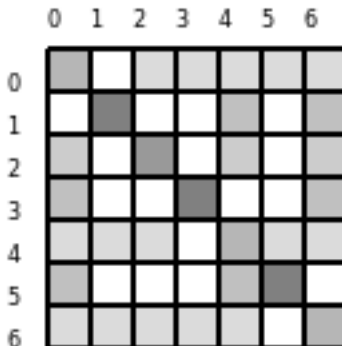
- 1: **while** not stable **do**
- 2: **for all** player i **do**
- 3: choisir une stratégie au hasard selon P_i
- 4: **end for**
- 5: replier le graphe
- 6: **for all** player i **do**
- 7: calculer les gains/coûts pour i
- 8: mettre à jour P_i
- 9: **end for**
- 10: **end while**

Sur une jonction 3-way simple :



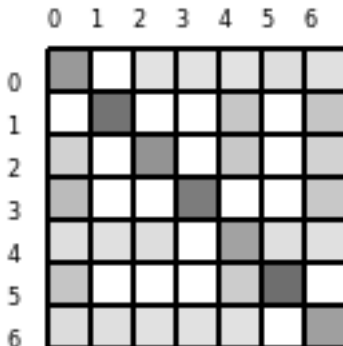
Résultats

Pour une famille A :

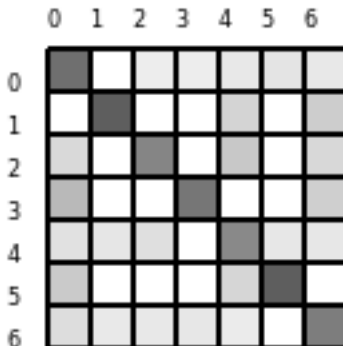


Résultats

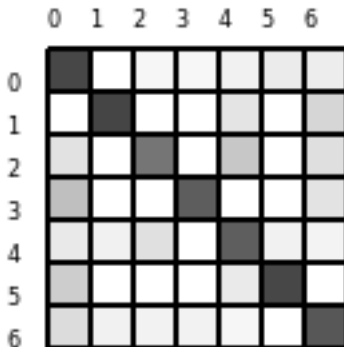
Pour une famille A :



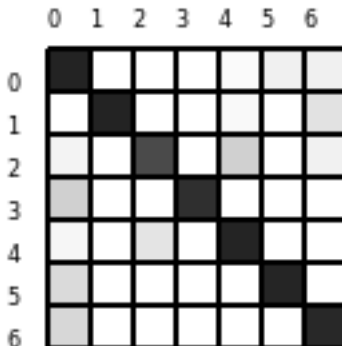
Pour une famille A :



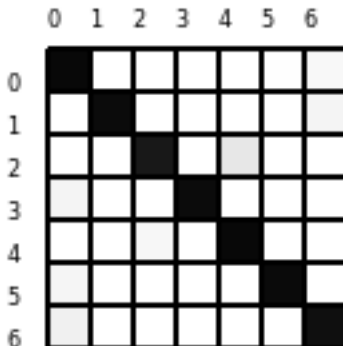
Pour une famille A :



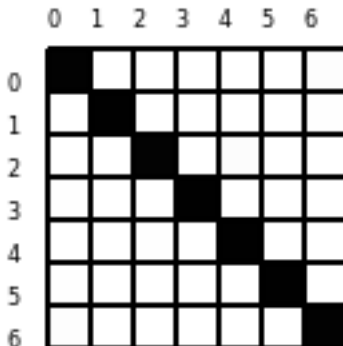
Pour une famille A :



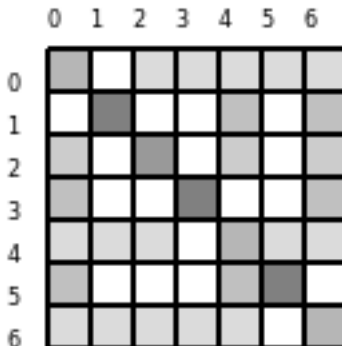
Pour une famille A :



Pour une famille A :

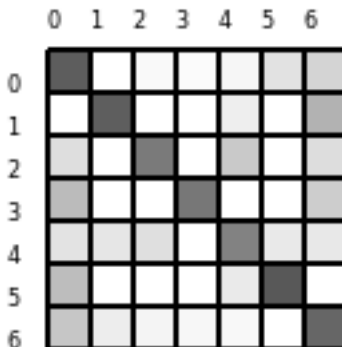


Pour une famille C :

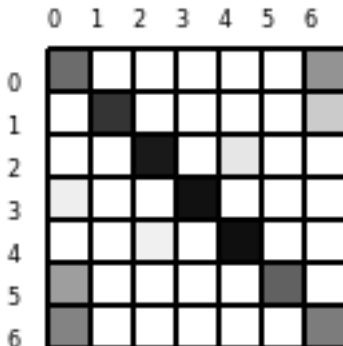


Résultats

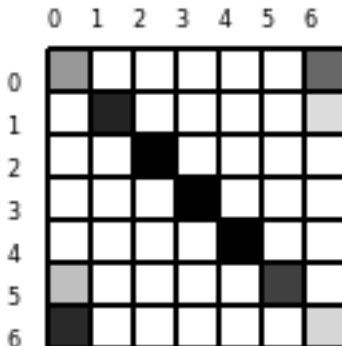
Pour une famille C :



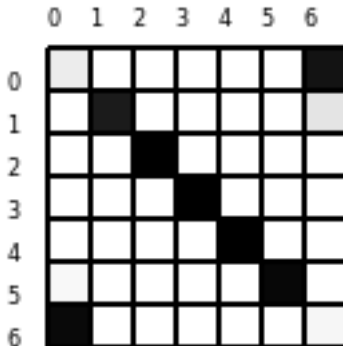
Pour une famille C :



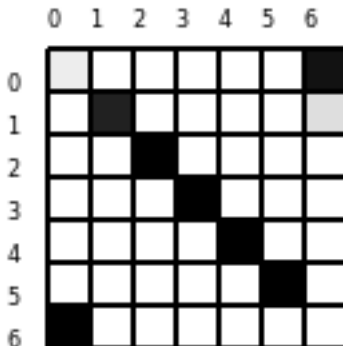
Pour une famille C :



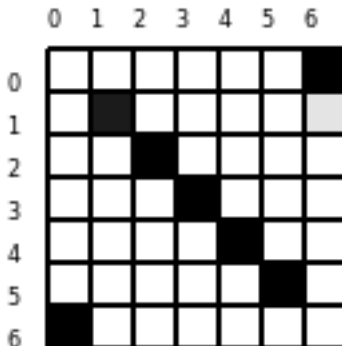
Pour une famille C :



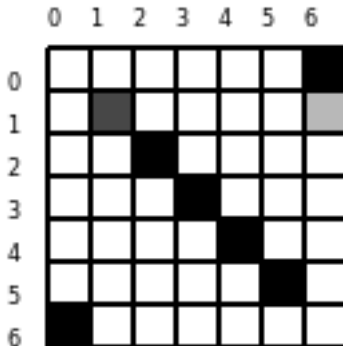
Pour une famille C :



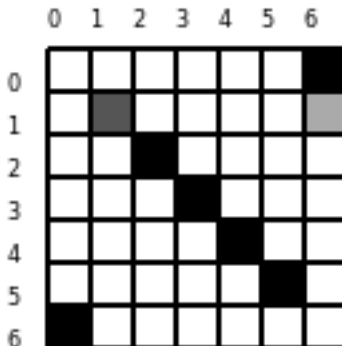
Pour une famille C :



Pour une famille C :



Pour une famille C :



Résultats

Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

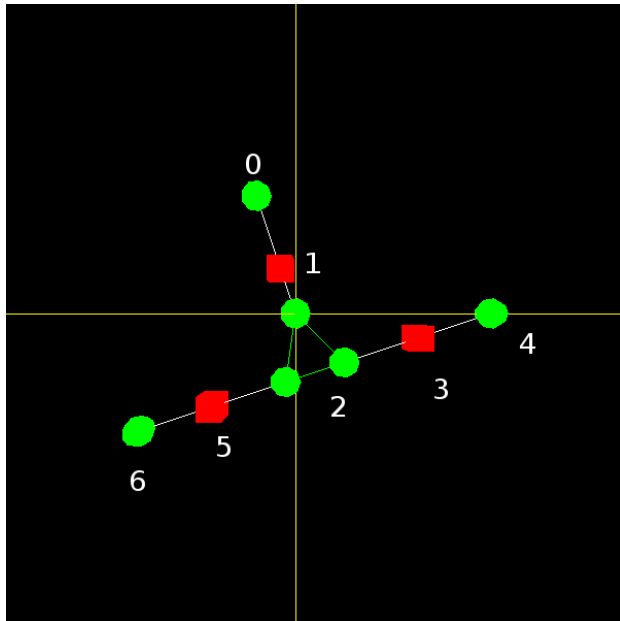
Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

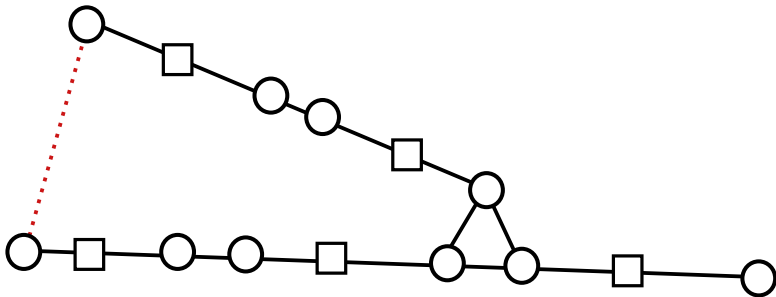
Formes locales

Forme globale

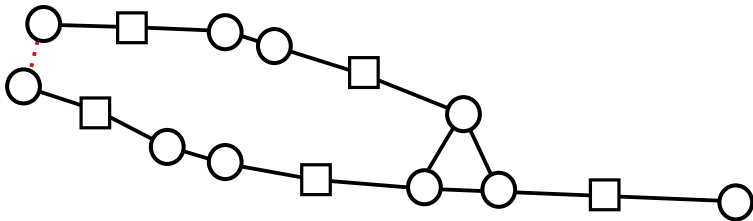
Conclusion



Longue distance



Longue distance



Plan

1 Introduction

2 Formes locales

3 Forme globale

4 Conclusion

Conclusion

- Prédiction à gros grain

Conclusion

- Prédiction à gros grain
- Prédiction en plusieurs étapes
local → global

Conclusion

- Prédiction à gros grain
- Prédiction en plusieurs étapes
local → global
- Peu de données pour valider

Conclusion

- Prédiction à gros grain
- Prédiction en plusieurs étapes
local \rightarrow global
- Peu de données pour valider
- Une certaine souplesse vis-à-vis des erreurs

Sur la classification :

- Identifier mieux les jonctions hors-famille
- Validation sur plus de structures
- Jonctions à 5, 6 branches ?

Sur la forme globale :

- Améliorer fonction de coût
- Exemples traitables et intéressants
- Valider la qualité de la prédiction

Prédiction à
gros grain de la
structure tridi-
mensionnelle
de l'ARN

Alexis Lamiable
D. Barth, A.
Denise, F.
Quessette, S.
Vial, É. Westhof

Introduction

Formes locales

Forme globale

Conclusion

Merci