RNA^{edit}
Interactive Visualisation of secondary RNA structures

Patrick AMAR[*] & friends…

[*]LRI (Laboratoire de Recherche en Informatique)

Université Paris-Sud, Orsay

Algorithmics of -1 frameshift RNA sequences

Michaël Bekaert[1], Laure Bidou[1], Alain Denise[2], Guillemette Duchateau-Nguyen[1], Céline Fabret[1] Jean-Paul Forest[2], Christine Froidevaux[2], Isabelle Hatin[1], Jean-Pierre Rousset[1], Michel Termier[1]

[1]IGM (Institut de Génétique et Microbiologie)
[2]LRI (Laboratoire de Recherche en Informatique)

Université Paris-Sud, Orsay

Flow of genetic information

DNA sequence

CATATGGATTACATGGTCTAAGAT

transcription

RNA sequence

CAU AUG GAU UAC AUG GUC UAA GAU

translation

Protein

Translation

5' ····· CAU AUG GAU UAC AUG GUC UAA GAU ····· 3'

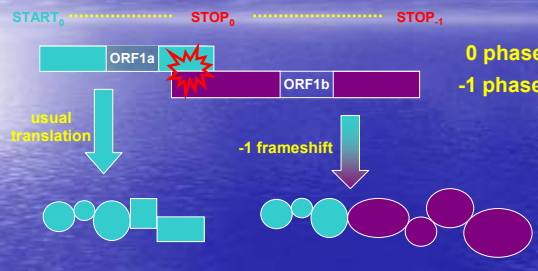
The synthesis goes on until a STOP codon is read

1 mRNA gives 1 protein

Experimental fact

- Some mRNAs encode two distinct proteins with the same beginning

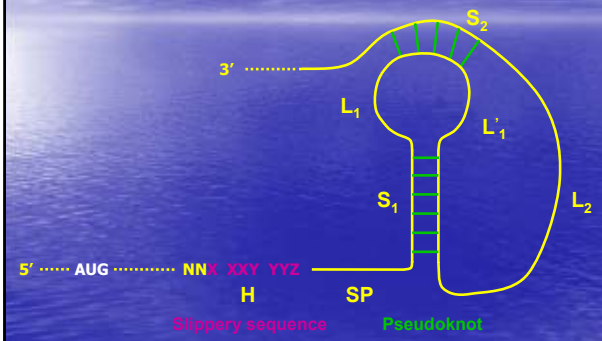
Programmed -1 frameshifting



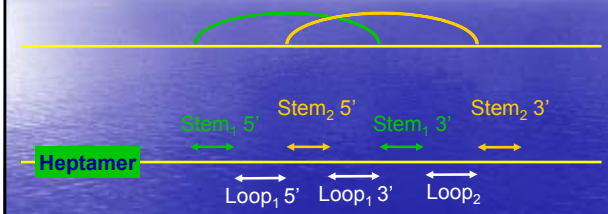
1 mRNA gives 2 distinct proteins with accurate ratio

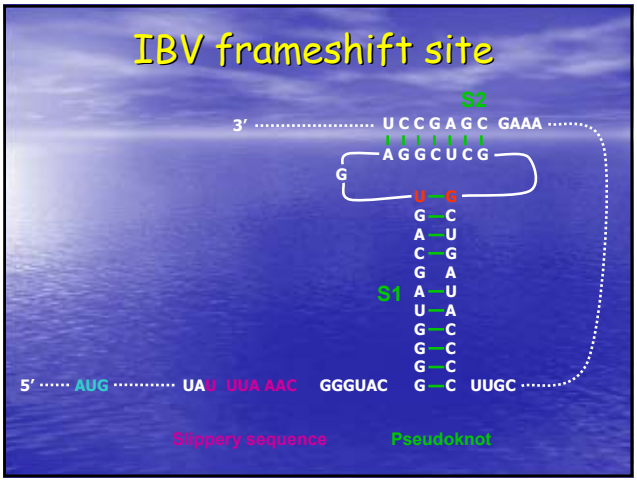
Typical -1 frameshift site

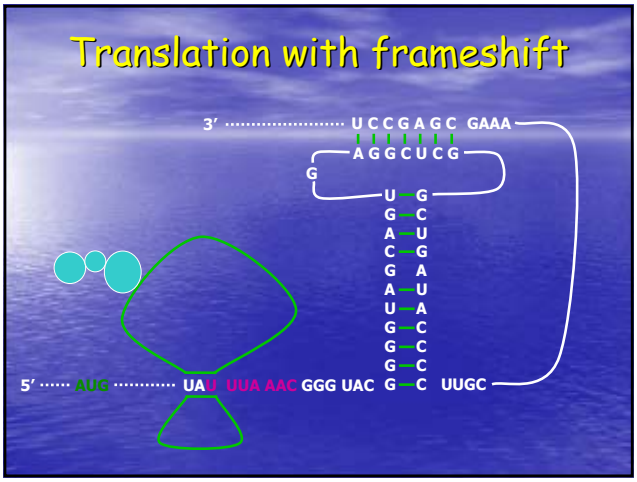
[Brierley, 1989]

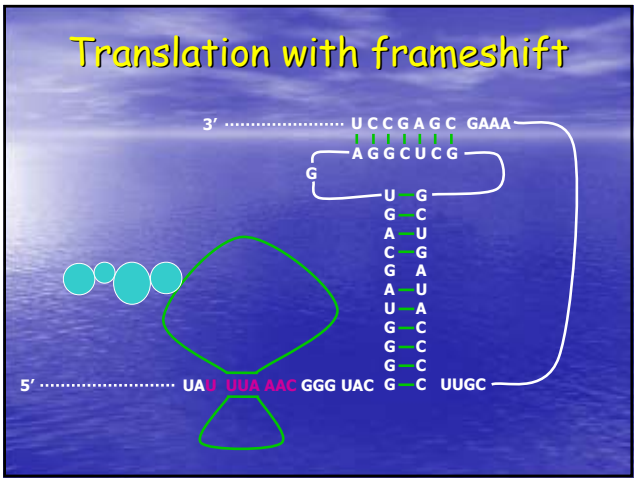


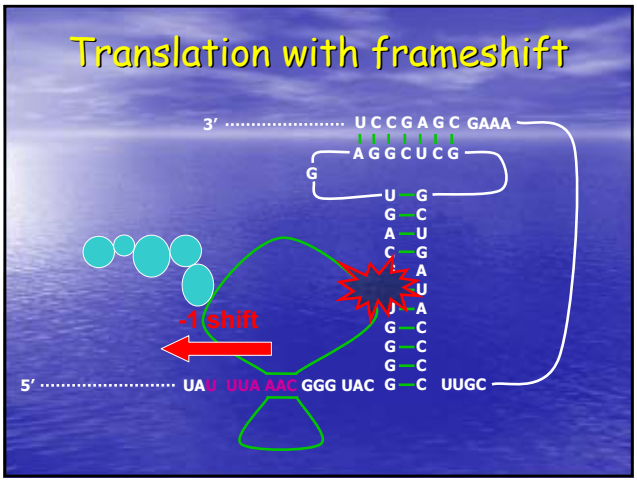
Find secondary structure (pseudoknot)

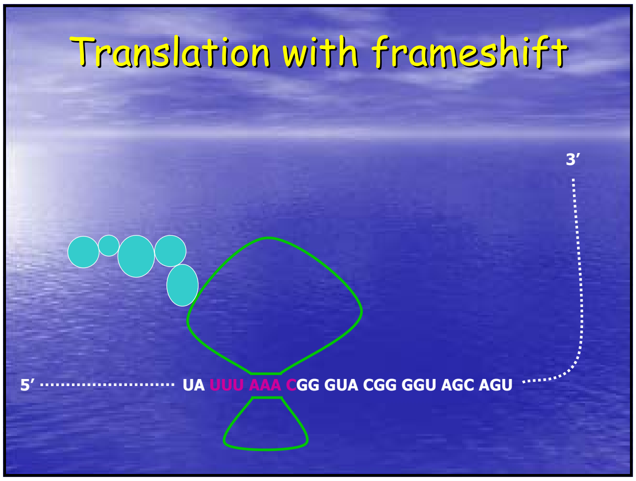


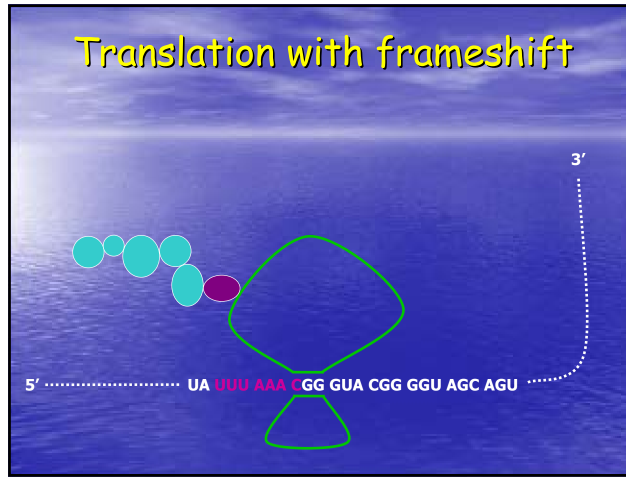

IBV frameshift site
S2
3′ UCCGAGC GAAA
AGGCUCG
G
U G
G C
A U
C G
G A
S1 A U
U A
G C
G C
G C
5′ AUG UAU UUA AAC GGGUAC G C UUGC
Slippery sequence
Pseudoknot

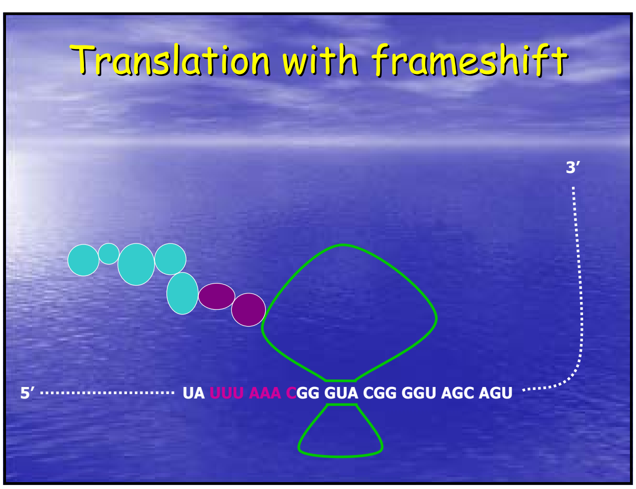

Translation with frameshift
3′ UCCGAGC GAAA
AGGCUCG
G
U G
G C
A U
C G
G A
A U
U A
G C
G C
G C
5′ AUG UAU UUA AAC GGG UAC G C UUGC

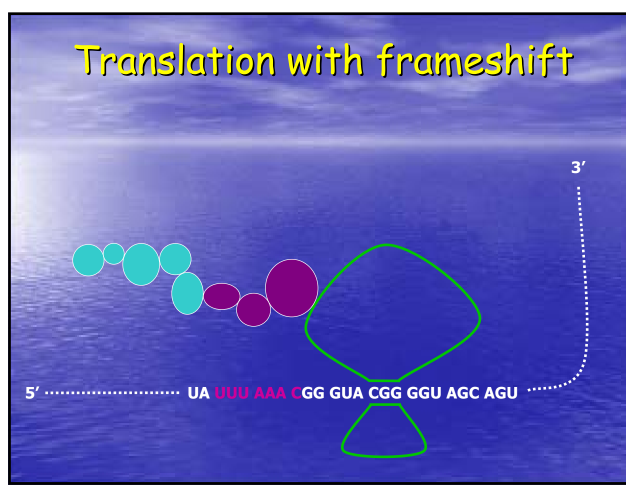

Translation with frameshift
3′ UCCGAGC GAAA
AGGCUCG
G
U G
G C
A U
C G
G A
A U
U A
G C
G C
G C
5′ UAU UUA AAC GGG UAC G C UUGC

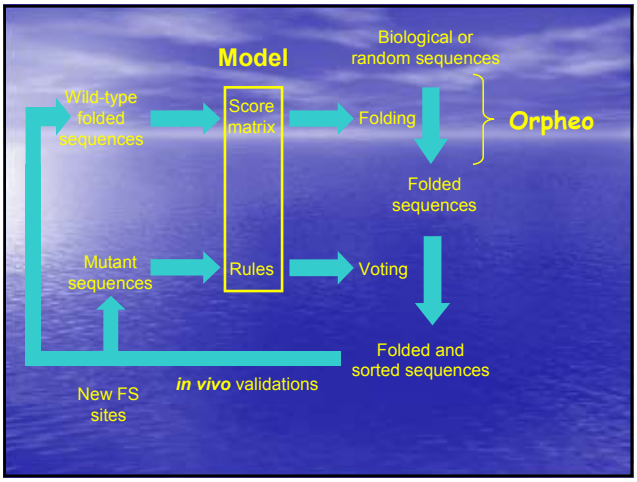

Translation with frameshift
3′ UCCGAGC GAAA
AGGCUCG
G
U G
G C
A U
C G
A U
U A
G C
G C
G C
-1 shift
5′ UAU UUA AAC GGG UAC G C UUGC

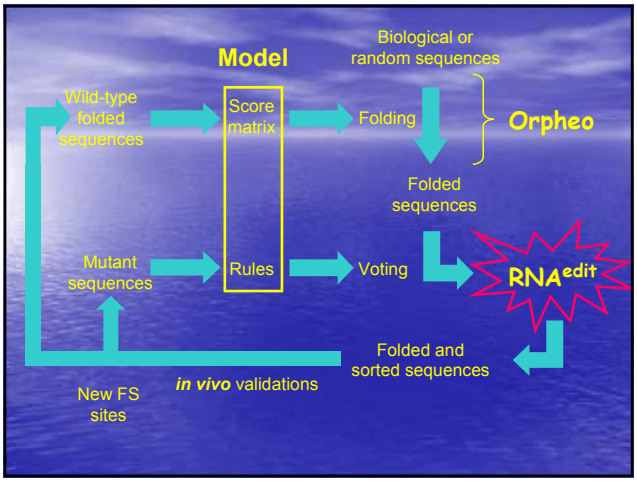

Translation with frameshift
3′
5′ UA UUU AAA CGG GUA CGG GGU AGC AGU


Translation with frameshift
3′
5′ UA UUU AAA CGG GUA CGG GGU AGC AGU


Translation with frameshift
3′
5′ UA UUU AAA CGG GUA CGG GGU AGC AGU


Translation with frameshift
3′
5′ UA UUU AAA CGG GUA CGG GGU AGC AGU





RNAedit

Multiple views

- Secondary structure displayed in 3D
- Travel through the entire structure (pilot view)
- Focus on a specific area:
  - Geometric zoom (cubic area)
  - Structural zoom (pseudoknot, stem-loop, etc.)

RNAedit

Subparts selection & edition

- Single nucleotide
  - Rename (mutation)
- Base pairs (bonds)
  - Break or create bonds (force a new folding)

RNAedit

- Input
  - CT format, bracket format
  - Specific format (Orpheo)
- Output
  - Base sequence (mutation)
  - bonds information (manually forced or break)

RNAedit

- Feedback from a control program
  - Committed changes submitted to programs to:
    - Calculate the free energy of the structure
    - Calculate a folding based on the new sequence
  - Number of G-C pairs in each stem
- Tightly coupled (linked) to folding algorithms
  - Orpheo, Vienna RNA, etc.

RNAedit

Implementation

- Portability
  - Language: C++
  - Graphic engine: *OpenGL*
  - GUI: *GLUI* (based on *GLUT*)
- Works on main platforms
  - Unix / Linux
  - MacOS X
  - Windows