RNA structural ensembles from energy landscape explorations

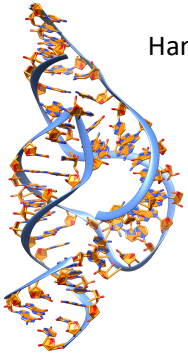
Dr Konstantin Röder

Randall Centre for Cell & Molecular Biophysics
School of Basic and Medical Biosciences
Faculty of Life Sciences and Medicine

KING'S College LONDON

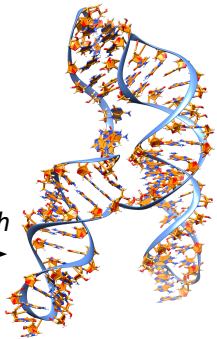
The importance of non-coding RNAs

Non-coding RNA **regulates** many **important processes**, e.g. in transcription or translation. Despite their centrality in regulation, ncRNAs are **underutilised** as **drug targets**.



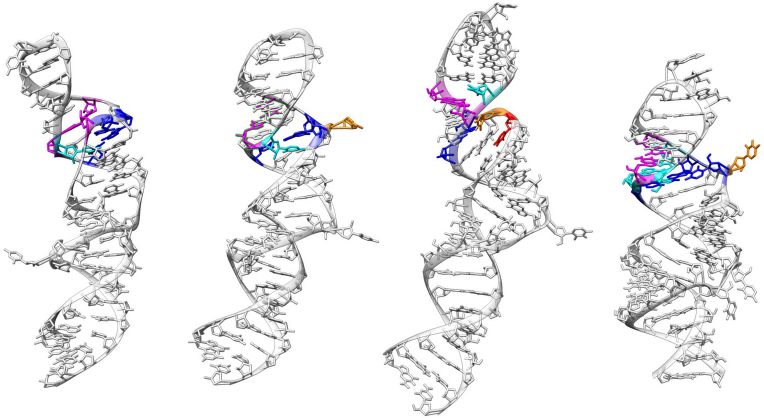
Hammerhead ribozyme
← *RNA cleavage*

RNA pseudoknot
from SARS CoV2
Regulatory switch →



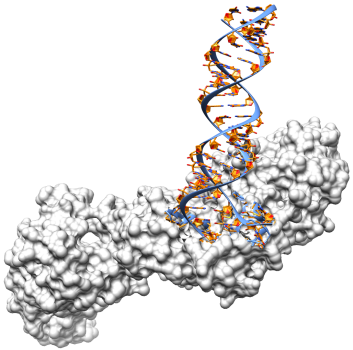
Structural polymorphism in RNAs

In contrast to proteins, RNAs exhibit **competing structural ensembles**. These may be characterisable in experiment, but their relative occupancy may be unobtainable.



External factors impacting RNA structural ensembles

- sequence changes
- epigenetic alterations
- environmental changes (pH, salt, ...)
- interaction partners
- ...



RNA-protein interaction in SARS-CoV2
Proofreading during RNA synthesis; critical for the viral life cycle

What insight can we gain from understanding RNA structures?

What insight can we gain from understanding RNA structures?

- The RNA 7SK HP1 hairpin interacts with an affector.

What insight can we gain from understanding RNA structures?

- The RNA 7SK HP1 hairpin interacts with an affector.
  Which structure binds affector and what us the mode of binding?

What insight can we gain from understanding RNA structures?

- The RNA 7SK HP1 hairpin interacts with an affector.
  Which structure binds affector and what us the mode of binding?

- This binding process is hijacked by HIV.

What insight can we gain from understanding RNA structures?

- The RNA 7SK HP1 hairpin interacts with an affector.
  Which structure binds affector and what us the mode of binding?

- This binding process is hijacked by HIV.
  Can we target this process in HIV, i.e. prevent the hijacking of transcription regulation?

What insight can we gain from understanding RNA structures?

- The RNA 7SK HP1 hairpin interacts with an affector.
  Which structure binds affector and what us the mode of binding?

- This binding process is hijacked by HIV.
  Can we target this process in HIV, i.e. prevent the hijacking of transcription regulation?

- m6A methylation in a stem-loop within a RNA transcript leads to viral activation in Kaposi's sarcoma-associated herpesvirus.

What insight can we gain from understanding RNA structures?

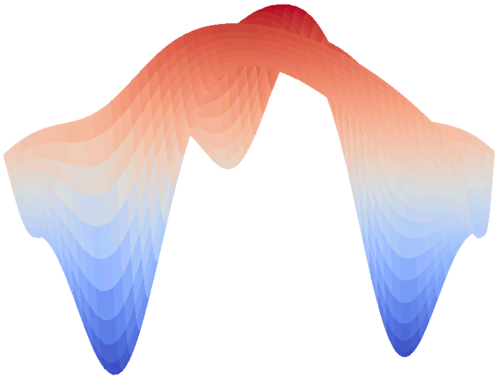
- The RNA 7SK HP1 hairpin interacts with an affector.
  Which structure binds affector and what us the mode of binding?

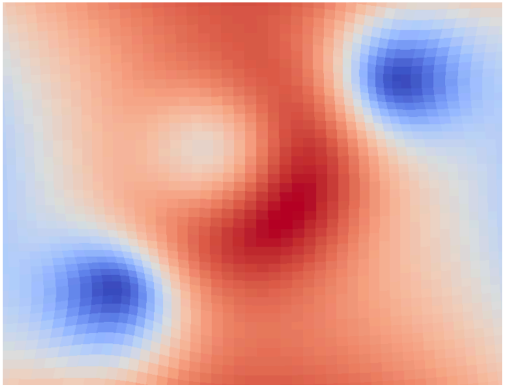
- This binding process is hijacked by HIV.
  Can we target this process in HIV, i.e. prevent the hijacking of transcription regulation?

- m6A methylation in a stem-loop within a RNA transcript leads to viral activation in Kaposi's sarcoma-associated herpesvirus.
  Why is single methyl group leading to viral activation and through which mechanism?

Exploring energy landscapes

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

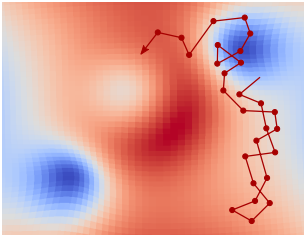


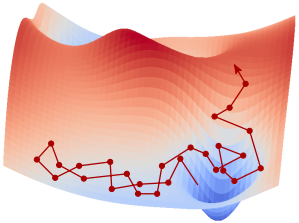


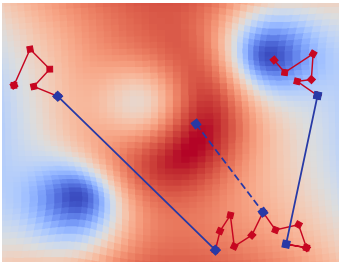
Classical explorations - MD and MC simulations

Molecular dynamics

- discrete time propagation using Newton's laws
- rare events remain rare
- need enhanced sampling methods
- challenge to simulate long time scales



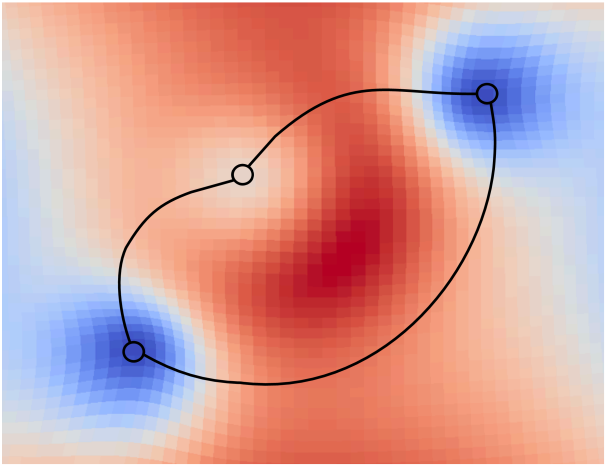




Monte Carlo sampling

- create "random" new configurations
- accept/reject based on their energy
- moves difficult to design

Changing the point of view: Minima and transition states



Stable configurations are **local minima** and transitions between them are characterised by Hessian-index 1 **saddle points** (*aka* transition states).

These stationary points **determine** structural, thermodynamic and kinetic **properties**.

Principles of the computational potential energy landscape framework

1. **Coarse-grain** the energy landscape as a set of minima and transition states.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

Principles of the computational potential energy landscape framework

1. **Coarse-grain** the energy landscape as a set of minima and transition states.
2. As configurations of interest are defined geometrically, finding structures is based **geometry optimisation**.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

Principles of the computational potential energy landscape framework

1. **Coarse-grain** the energy landscape as a set of minima and transition states.
2. As configurations of interest are defined geometrically, finding structures is based **geometry optimisation**.
3. Geometry optimisation is **independent** of energy barriers and hence **time scales**.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

Principles of the computational potential energy landscape framework

1. **Coarse-grain** the energy landscape as a set of minima and transition states.
2. As configurations of interest are defined geometrically, finding structures is based **geometry optimisation**.
3. Geometry optimisation is **independent** of energy barriers and hence **time scales**.
4. **All information** required to calculate kinetics, thermodynamics and mechanisms is **obtained** this way.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

Principles of the computational potential energy landscape framework

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

1. **Coarse-grain** the energy landscape as a set of minima and transition states.
2. As configurations of interest are defined geometrically, finding structures is based **geometry optimisation**.
3. Geometry optimisation is **independent** of energy barriers and hence **time scales**.
4. **All information** required to calculate kinetics, thermodynamics and mechanisms is **obtained** this way.
5. Transitions are represented as **discrete paths** containing a series of local minima connected by transition states.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

Principles of the computational potential energy landscape framework

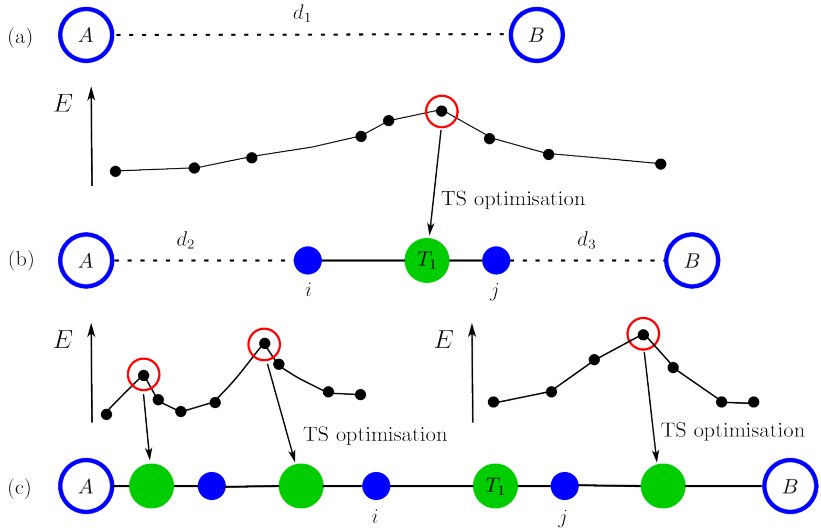
1. **Coarse-grain** the energy landscape as a set of minima and transition states.
2. As configurations of interest are defined geometrically, finding structures is based **geometry optimisation**.
3. Geometry optimisation is **independent** of energy barriers and hence **time scales**.
4. **All information** required to calculate kinetics, thermodynamics and mechanisms is **obtained** this way.
5. Transitions are represented as **discrete paths** containing a series of local minima connected by transition states.
6. **No projections** into lower dimensions or the use of reaction coordinates are required.

For more details, see: Joseph *et al.*, *Chem. Commun.* 2017, Röder *et al.*, *Adv. Theory Simul.* 2019 and Röder and Wales, *Front. Mol. Biosci.* 2022

For and instructive guide, see: Röder and Pasquali, *in:* RNA Scaffolds (Methods in Molecular Biology series), 2021

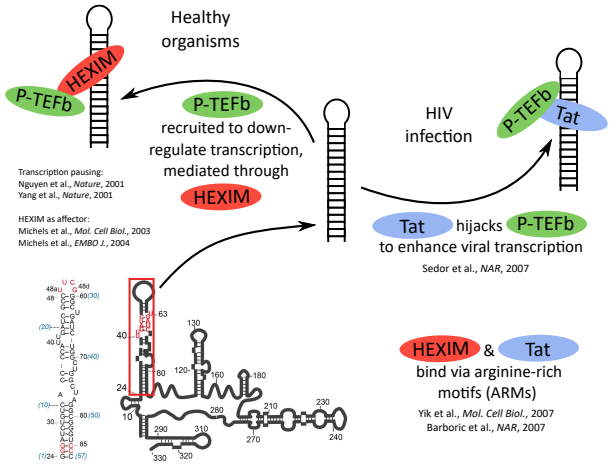
Opening the bonnet - what we actually do

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



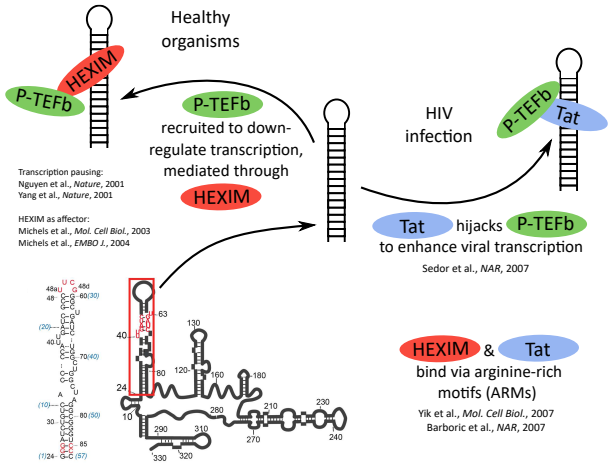
The regulatory role of 7SK RNA HP1

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



The regulatory role of 7SK RNA HP1

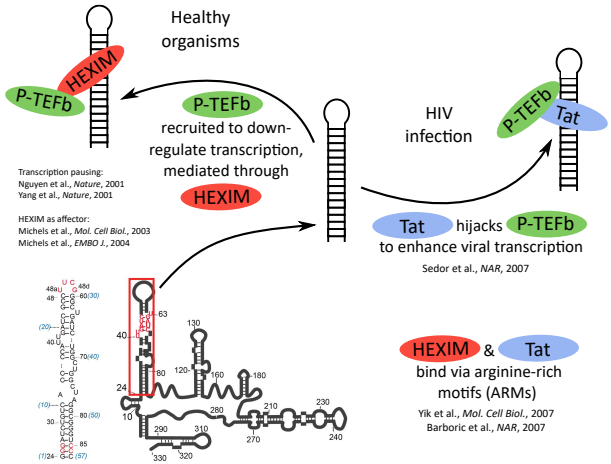
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



- HP1 hairpin contains binding motif

The regulatory role of 7SK RNA HP1

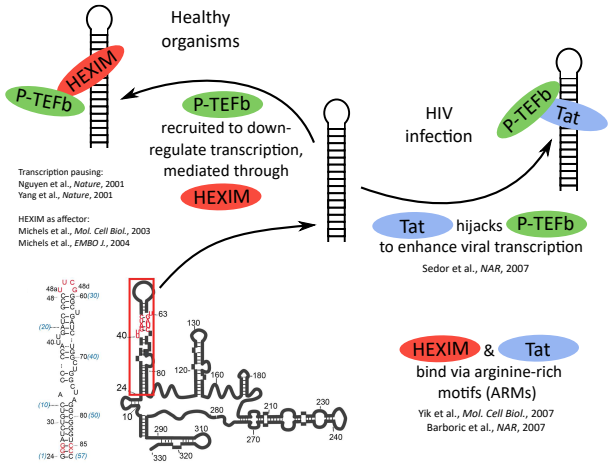
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



- HP1 hairpin contains binding motif
- use of stable, shortened model

The regulatory role of 7SK RNA HP1

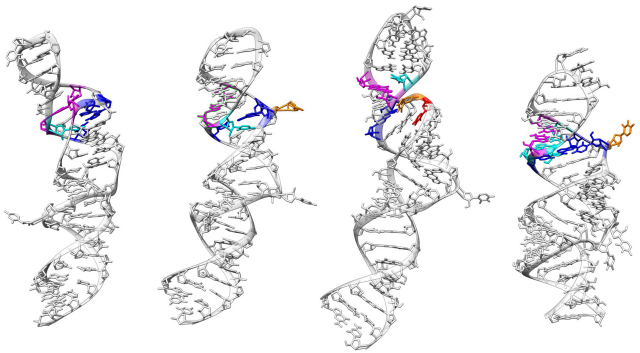
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



- HP1 hairpin contains binding motif
- use of stable, shortened model
- key motif: U(U)GAUC

Experimental structure - a puzzle with missing pieces

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

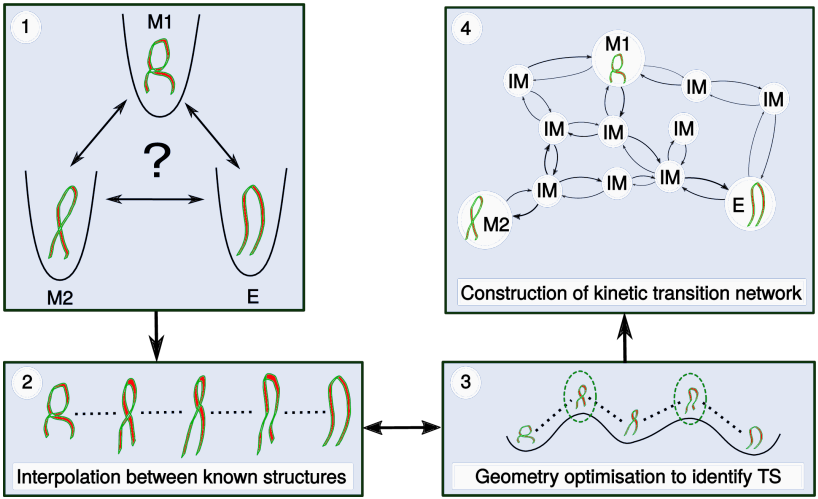


Two co-existing structures from crystallography and two structures from NMR

1,2 - Martinez-Zapien et al., *NAR*, 2017; 3 - Bourbigot et al., *RNA*, 2016; 4 - Pham et al., *Nat. Commun.*, 2018

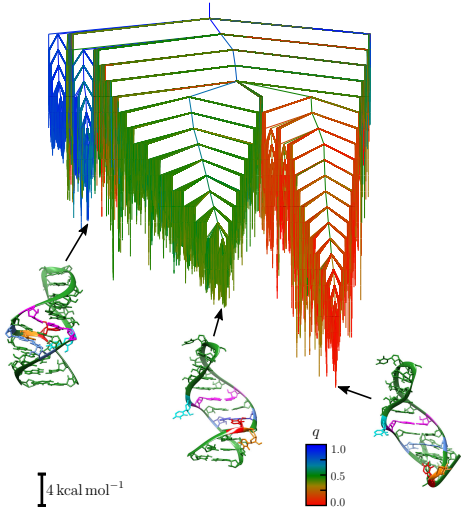
How to explore the energy landscape

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

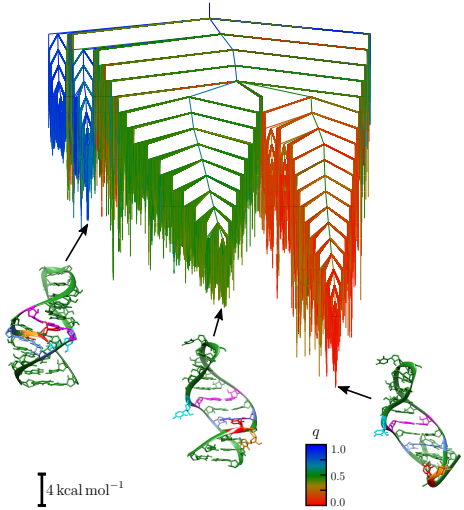


The energy landscape for the native sequence

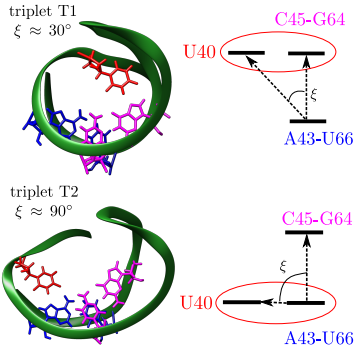
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



The energy landscape for the native sequence

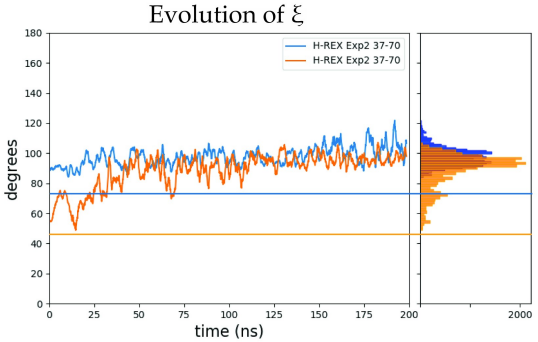


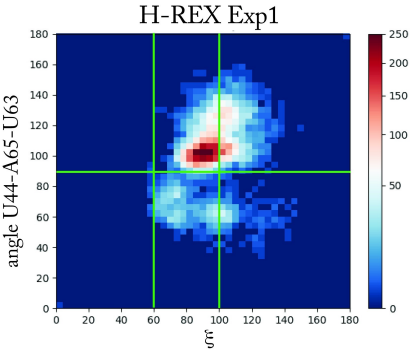
The configurations can be identified by an order parameter based on the angle ξ.



H-REX MD simulations in agreement

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead





Binding of 7SK RNA HP1 and ARM

RNA structures from ELs

Dr Konstantin Röder

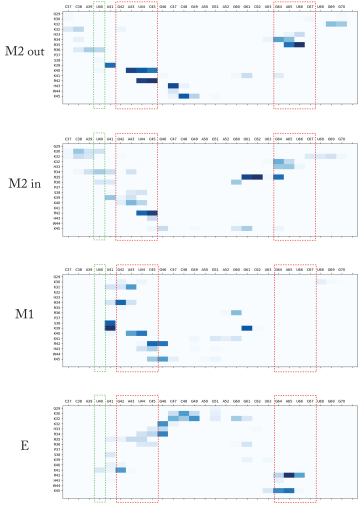
RNA polymorphism

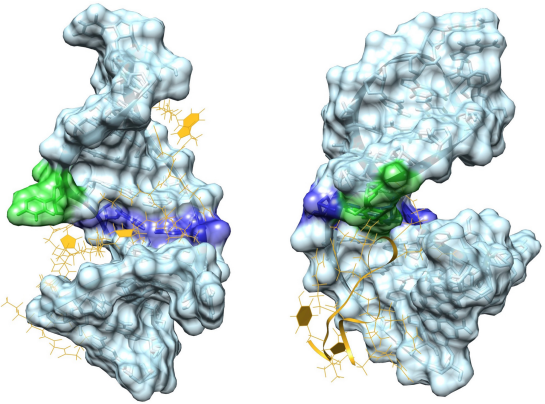
Energy landscape explorations

7SK RNA HP1 hairpin

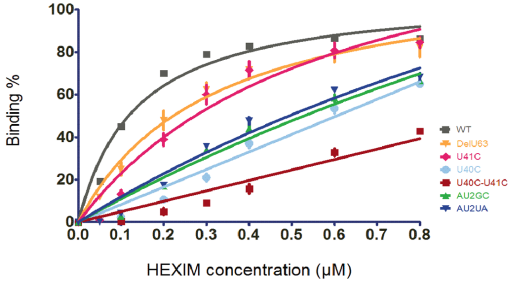
Further case studies

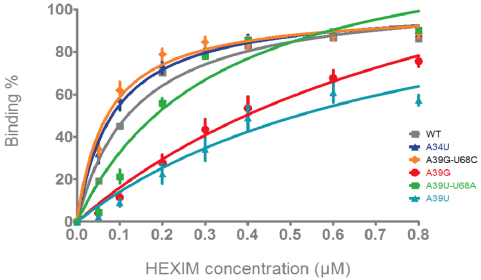
Pseudoknots and a look ahead





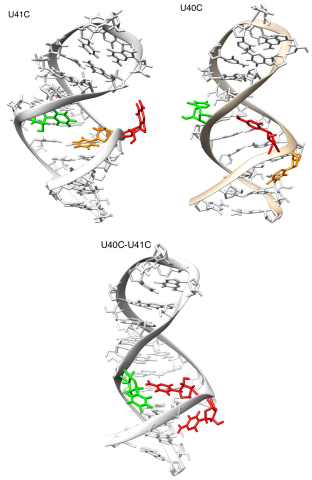
Mutations and their impact on peptide binding

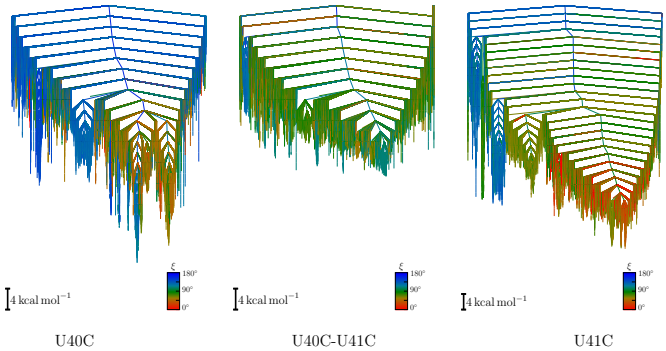




Mutations and their impact on peptide binding

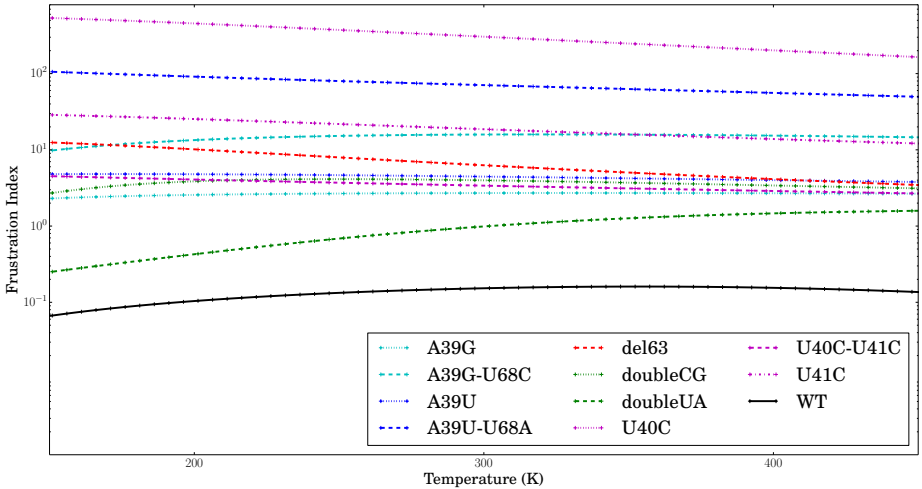
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead





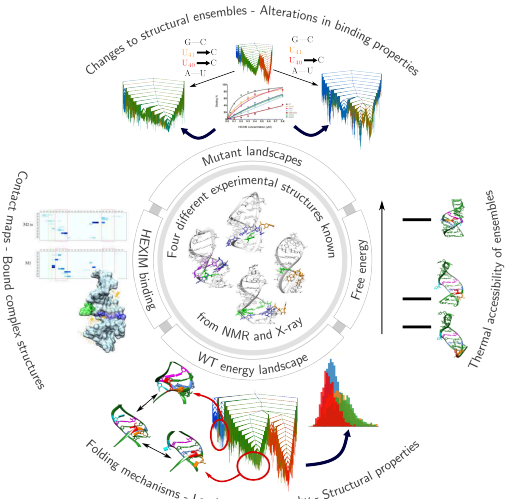
Side note: Evolutionary signatures in the energy landscapes

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



7SK RNA work - a summary

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



Röder et al., *Nucleic Acids Res.*, 2020

KSHV ORF50 transcript - viral reactivation

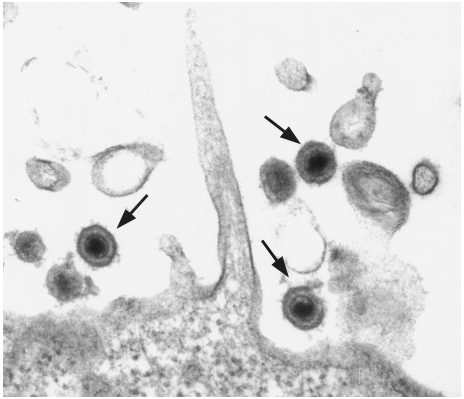


Image: Ablashi et al., *Clin. Microbiol. Rev.*, 2002

KSHV is a human **oncovirus**, and its transcription is relying on a viral protein encoded by **ORF-50**.

KSHV ORF50 transcript - viral reactivation

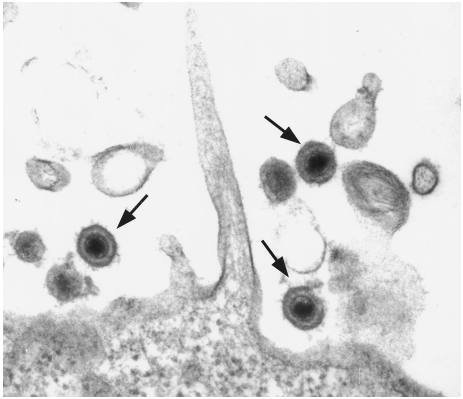


Image: Ablashi et al., *Clin. Microbiol. Rev.*, 2002

KSHV is a human **oncovirus**, and its transcription is relying on a viral protein encoded by **ORF-50**.

KSHV manipulates the host cell N6-methyl adenosine (m6A) RNA modification pathway.

KSHV ORF50 transcript - viral reactivation

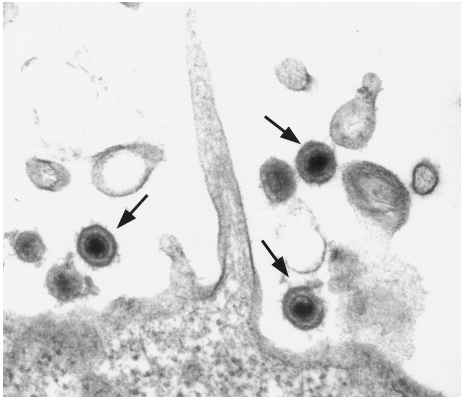


Image: Ablashi et al., *Clin. Microbiol. Rev.*, 2002

KSHV is a human **oncovirus**, and its transcription is relying on a viral protein encoded by **ORF-50**.

KSHV manipulates the host cell N6-methyl adenosine (m6A) RNA modification pathway.

m6A methylation of the ORF50 transcript enhances virus replication.

(Baquero-Perez et al., *eLife* 2019)

KSHV ORF50 transcript - viral reactivation

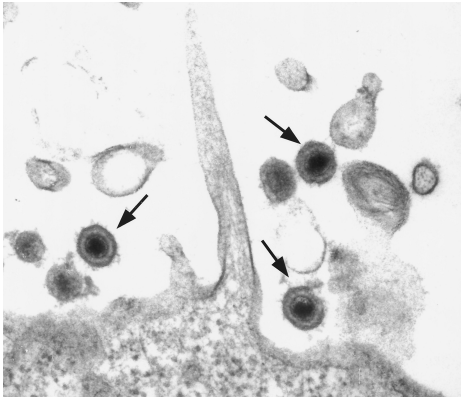


Image: Ablashi et al., *Clin. Microbiol. Rev.*, 2002

KSHV is a human **oncovirus**, and its transcription is relying on a viral protein encoded by **ORF-50**.

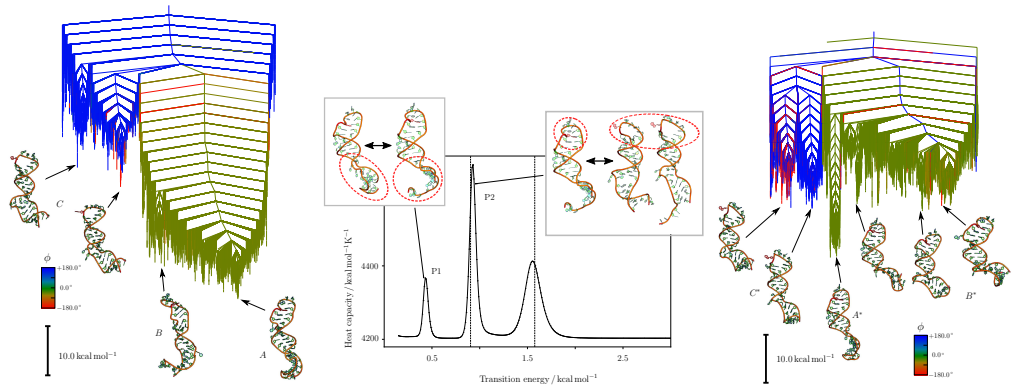
KSHV manipulates the host cell N6-methyl adenosine (m6A) RNA modification pathway.

m6A methylation of the ORF50 transcript enhances virus replication.

(Baquero-Perez et al., *eLife* 2019)

m6A modified site within the ORF50 transcript is a 43mer RNA stem-loop.

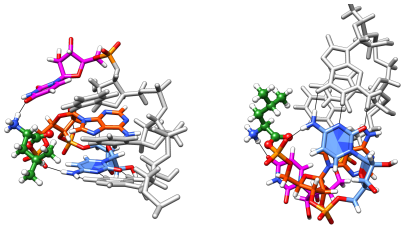
KSHV ORF50 transcript - m6A methylation changes everything

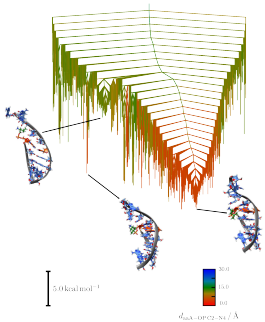


Röder et al., *PLoS Comput. Bio.*, 2022

Structure determines rates in tRNA-amino acyl transfer reactions

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



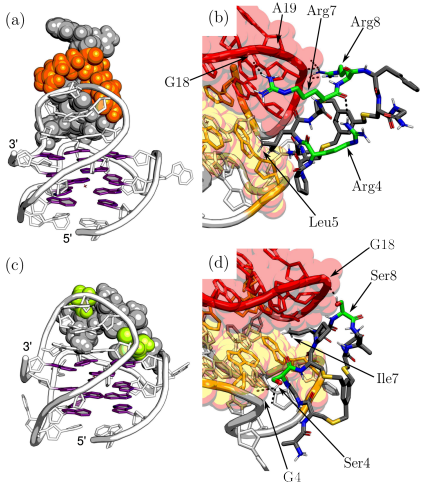


Amino acyl transfer in tRNA overhangs is dependent on the **tRNA stem** sequence.

Collaboration with Meng Su and John Sutherland, LMB, under review.

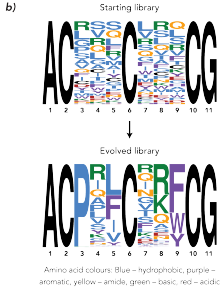
Rationalising bicyclic peptide design

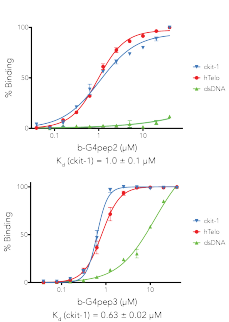
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



Bicyclic peptides mimic **protein-like binding** with G4s, allowing DNA **recognition by structure.**

Liu et al., *JACS*, 2020





(K)not complex ?!

SARS-CoV2 frameshifting RNA element employs **shifting to backtrack and pause** during translation. Graph-based computation and chemical probing show **multiple alternative structures**. Schlick et al., *JACS*, 2021

(K)not complex ?!

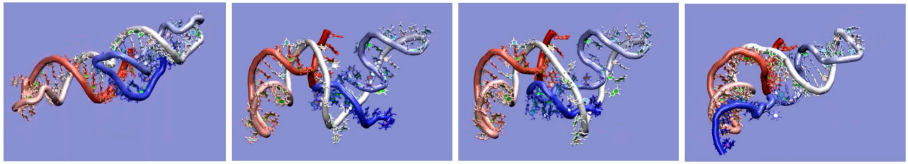
SARS-CoV2 frameshifting RNA element employs **shifting to backtrack and pause** during translation. Graph-based computation and chemical probing show **multiple alternative structures**. Schlick et al., *JACS*, 2021

What is the transition mechanism between the structures?

(K)not complex ?!

SARS-CoV2 frameshifting RNA element employs **shifting to backtrack and pause** during translation. Graph-based computation and chemical probing show **multiple alternative structures**. Schlick et al., *JACS*, 2021

What is the transition mechanism between the structures?



QCI - the magic ingredient

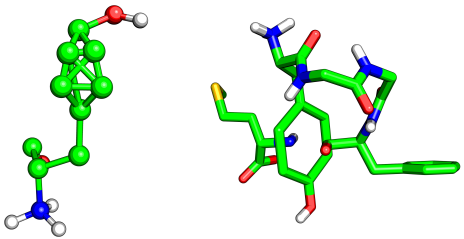
The Problem

- NEB and DNEB rely on **discrete** images.
- Even if neighbouring images are physical (correct cis-trans isomers, chirality, ...), the continuous path between the might be **unphysical**.
- In particular, **bond crossings** may occur. (NB: This affects all methods with discrete steps!)

QCI - the magic ingredient

The Problem

- NEB and DNEB rely on **discrete** images.
- Even if neighbouring images are physical (correct cis-trans isomers, chirality, ...), the continuous path between the might be **unphysical**.
- In particular, **bond crossings** may occur. (NB: This affects all methods with discrete steps!)



QCI - the magic ingredient

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

Such unphysical structures **can be detected** with relative ease, and often (not always!) they leave high-energy signatures. Specifically designed **sampling algorithms** also help.

```
              E+        Ets - E+           Ets       Ets - E-           E-          S       D      gamma   ~N
  -10013.5448812  3.8330       -10009.7118548  3.8330       -10013.5448812   3.907   3.0731537.088   3.000
 chirality_check> Atom   996          inverted.
 chirality_check> Originally S chirality.
 chirality_check> Atom   996          inverted.
 chirality_check> Originally S chirality.
tryconnect> Chirality inversion detected in at least one of the carbon centres, rejecting
```

```
       E+        Ets - E+           Ets       Ets - E-           E-          S       D      gamma   ~N
  -9986.9012644  21.360        -9965.5408247  13.747        -9979.2879197  86.082  65.325 172.965  26.659
 cis_trans_check> Peptide bond    29    28    30    31 changed to cis.
 cis_trans_check> Originally trans.
 cis_trans_check> Peptide bond    29    28    30    31 changed to cis.
 cis_trans_check> Originally trans.
tryconnect> Cis-trans isomerisation of a peptide bond detected (wrt. the original structure), rejecting
tryconnect> Transition state with energy      -9965.540825     ignored, cis-trans isomerisation detected in one or more peptide bonds.
```

```
 tryconnect> 400-iteration DNEB run for minima 1_S and 2_F using 40 images  (DNEB attempt #1)  ...
 lbfgs> Cold fusion diagnosed - step discarded, energy, limit=    -2054448299.        -1000000.000
 lbfgs> Final DNEB force constant      10.00000000
 Time to go through NEB:   0.553106999999997
 Double-ended search iterations= 0 RMS= ******************** Dev= 0.00% S= 0.00 time= 0.55
```

QCI - the magic ingredient

The Solution - Quasi-continuous interpolation (QCI)

Auxiliary potential for interpolation

- Constraint term: chemical bonds and geometric constraints (planarity, cis-trans ...)
- Repulsion term: Coulombic repulsion for non-constraint atom at close proximity
- Quasi-continuous term: repulsion based on local minima in the distances between bonds in adjacent images

Wales and Carr, *JCTC*, 2012 and Röder and Wales, *JCTC*, 2018

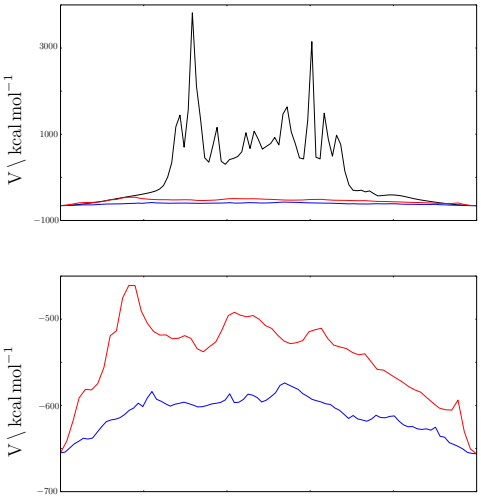
QCI - the magic ingredient

The Solution - Quasi-continuous interpolation (QCI)

Auxiliary potential for interpolation

- Constraint term: chemical bonds and geometric constraints (planarity, cis-trans ...)
- Repulsion term: Coulombic repulsion for non-constraint atom at close proximity
- Quasi-continuous term: repulsion based on local minima in the distances between bonds in adjacent images

Wales and Carr, *JCTC*, 2012 and Röder and Wales, *JCTC*, 2018



Simplifying things

Use of multi-scale simulations for further acceleration

Require a physical coarse-grained scheme with enough information to obtain good all-atom structures.

Simplifying things

Use of multi-scale simulations for further acceleration

Require a physical coarse-grained scheme with enough information to obtain good all-atom structures.

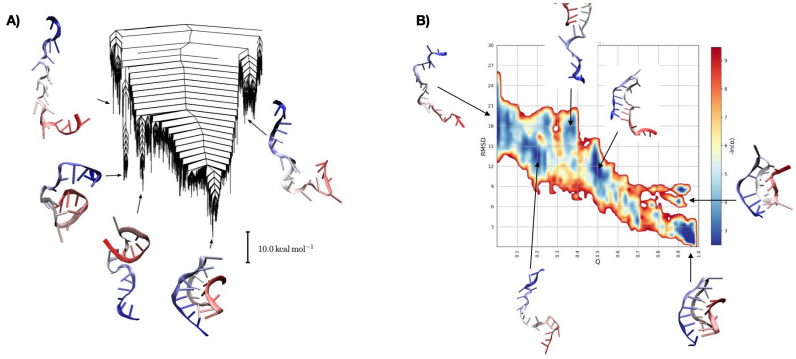
Use of HiRE-RNA (High-resolution CG Energy Model for RNA)

- potential: Pasquali and Derreumaux, *JPCB*, 2010 and Cragnolini et al., *JCTC*, 2015
- 6/7 bead per nucleotide model
- physically motivated with focus on non-canonical interactions

A test case - PK1 pseudoknot

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead

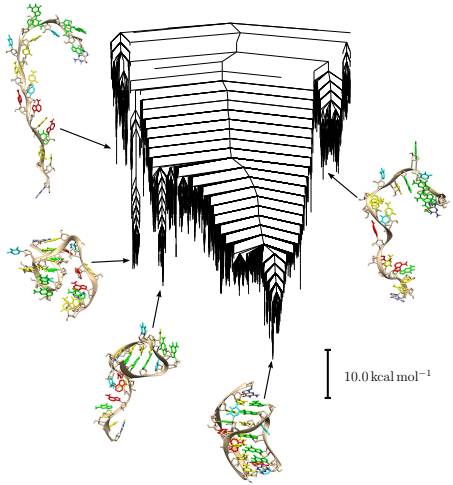
Pseudoknot1 (PK1) from the thermophilic bacterium *A.aeolicus* - predicted to be **smallest pseudoknot** (21 nucleotides) and NMR studies available.
All-atom ff with different methods: Correct folding/unfolding of PK1



Röder et al., *QRB Discovery*, 2022

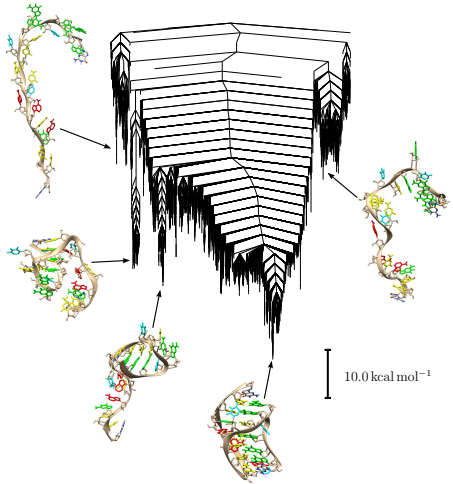
A test case - pK1 pseudoknot

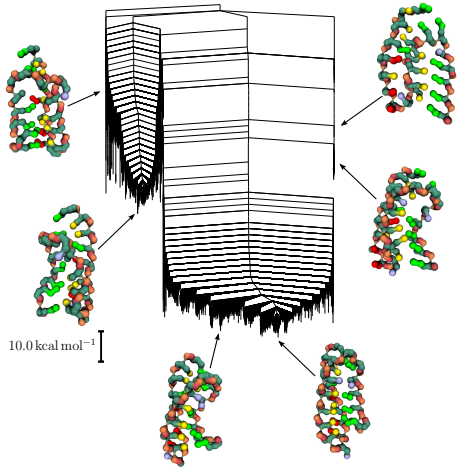
RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead



A test case - pK1 pseudoknot

RNA structures from ELs
Dr Konstantin Röder
RNA polymorphism
Energy landscape explorations
7SK RNA HP1 hairpin
Further case studies
Pseudoknots and a look ahead





A test case - pK1 pseudoknot

- potential too sticky (not surprising: parameters based on folded states)
- wrong structures
- global minimum is correct

A test case - pK1 pseudoknot

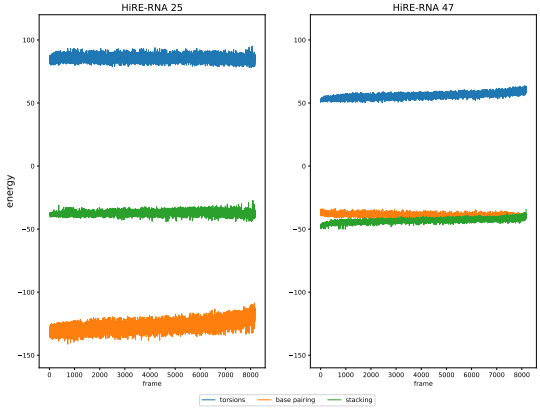
- potential too sticky (not surprising: parameters based on folded states)
- wrong structures
- global minimum is correct

- introduce **more parameters** to capture physics better
- **reoptimise parameters** based on folded and partially folded structure
- **balance** individual energy **contributions**

A test case - pK1 pseudoknot

- potential too sticky (not surprising: parameters based on folded states)
- wrong structures
- global minimum is correct

- introduce **more parameters** to capture physics better
- **reoptimise parameters** based on folded and partially folded structure
- **balance** individual energy **contributions**



HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions
3. new parameter optimisation including partially-unfolded and excited structures

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions
3. new parameter optimisation including partially-unfolded and excited structures
4. all input can be created from pdb file or sequence

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions
3. new parameter optimisation including partially-unfolded and excited structures
4. all input can be created from pdb file or sequence

5. available with additional simulation tools:

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions
3. new parameter optimisation including partially-unfolded and excited structures
4. all input can be created from pdb file or sequence

5. available with additional simulation tools:
   - basin-hopping global optimisation (GMIN)

HiRE - a new look (available soon)

1. potential completely rewritten and available as stand-alone library
2. improved terms for base stacking and pairing and implementation of phosphate - nucleobase interactions
3. new parameter optimisation including partially-unfolded and excited structures
4. all input can be created from pdb file or sequence

5. available with additional simulation tools:
   - basin-hopping global optimisation (GMIN)
   - MD engine (including temperature and Hamiltonian REX)

RNA structures from ELs

Dr Konstantin Röder

RNA polymorphism

Energy landscape explorations

7SK RNA HP1 hairpin

Further case studies

Pseudoknots and a look ahead

Acknowledgements

Collaborations

Prof. Samuela Paquali (Université de Paris Cité, France)
Prof. David J. Wales (University of Cambridge, UK)
Dr Amy M. Barker and Prof. Adrian Whitehouose (University of Leeds, UK)
Dr Meng Su and Prof. John Sutherland (LMB, Cambridge, UK)
Dr Kim C. Liu (University of Cambridge and LMB, UK)

Funding

King's College London
Henslow Research Fellowship, Cambridge Philosophical Society
Yusuf Hamied Departmental Fund, University of Cambridge
Visiting Fellowship, Univeristé de Paris Cité