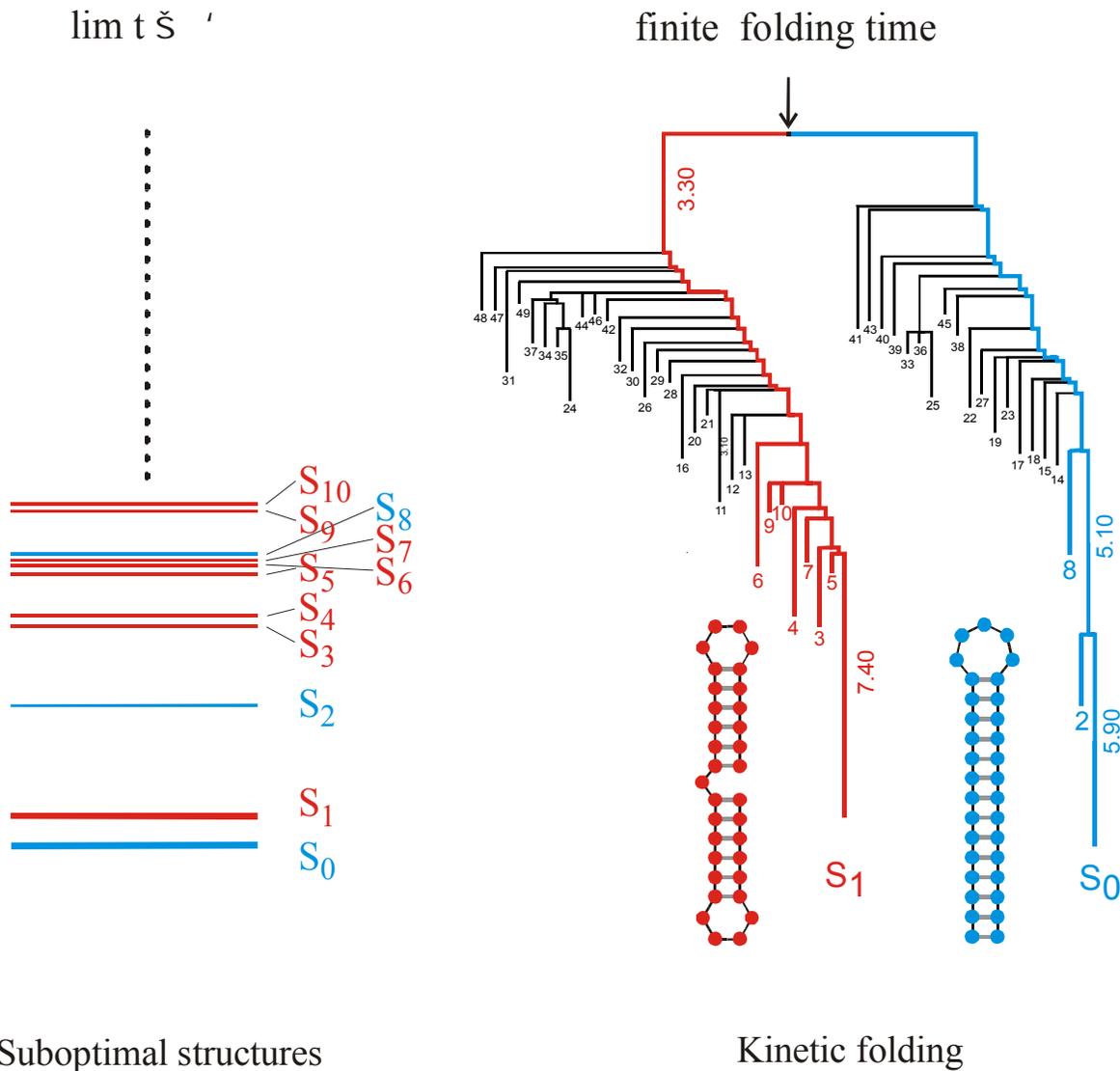






Web-Page for further information:

<http://www.tbi.univie.ac.at/~pks>



A typical energy landscape of a sequence with two (meta)stable conformations

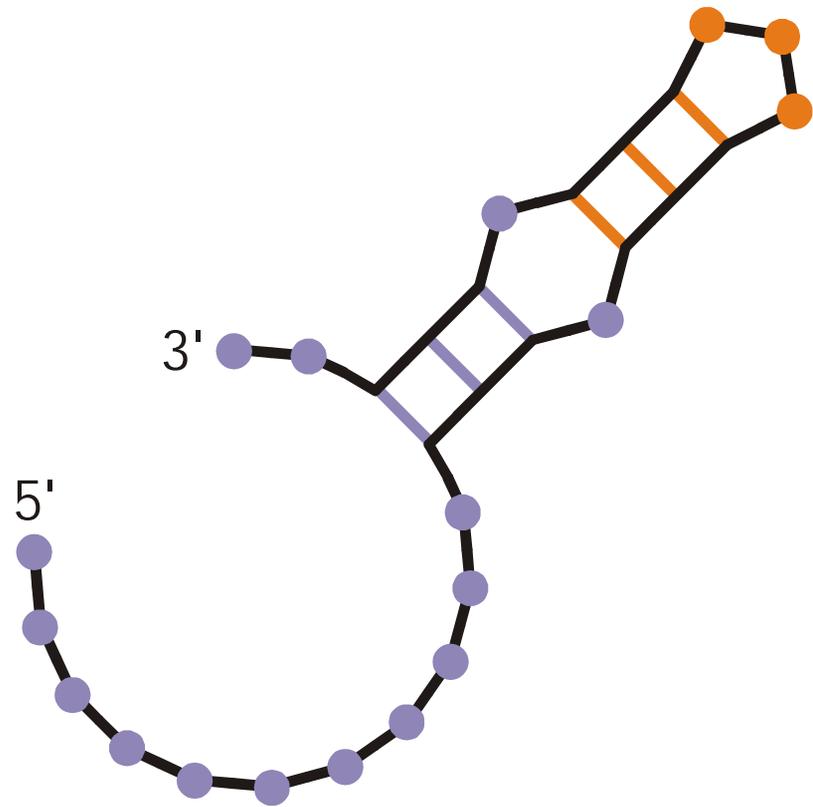
## Suboptimal RNA Secondary Structures

Michael Zuker. *On finding all suboptimal foldings of an RNA molecule*. Science **244** (1989), 48-52

Stefan Wuchty, Walter Fontana, Ivo L. Hofacker, Peter Schuster. *Complete suboptimal folding of RNA and the stability of secondary structures*. Biopolymers **49** (1999), 145-165

Total number of structures including all suboptimal conformations, stable and unstable (with  $\Delta G_0 > 0$ ):

#conformations = **1 416 661**

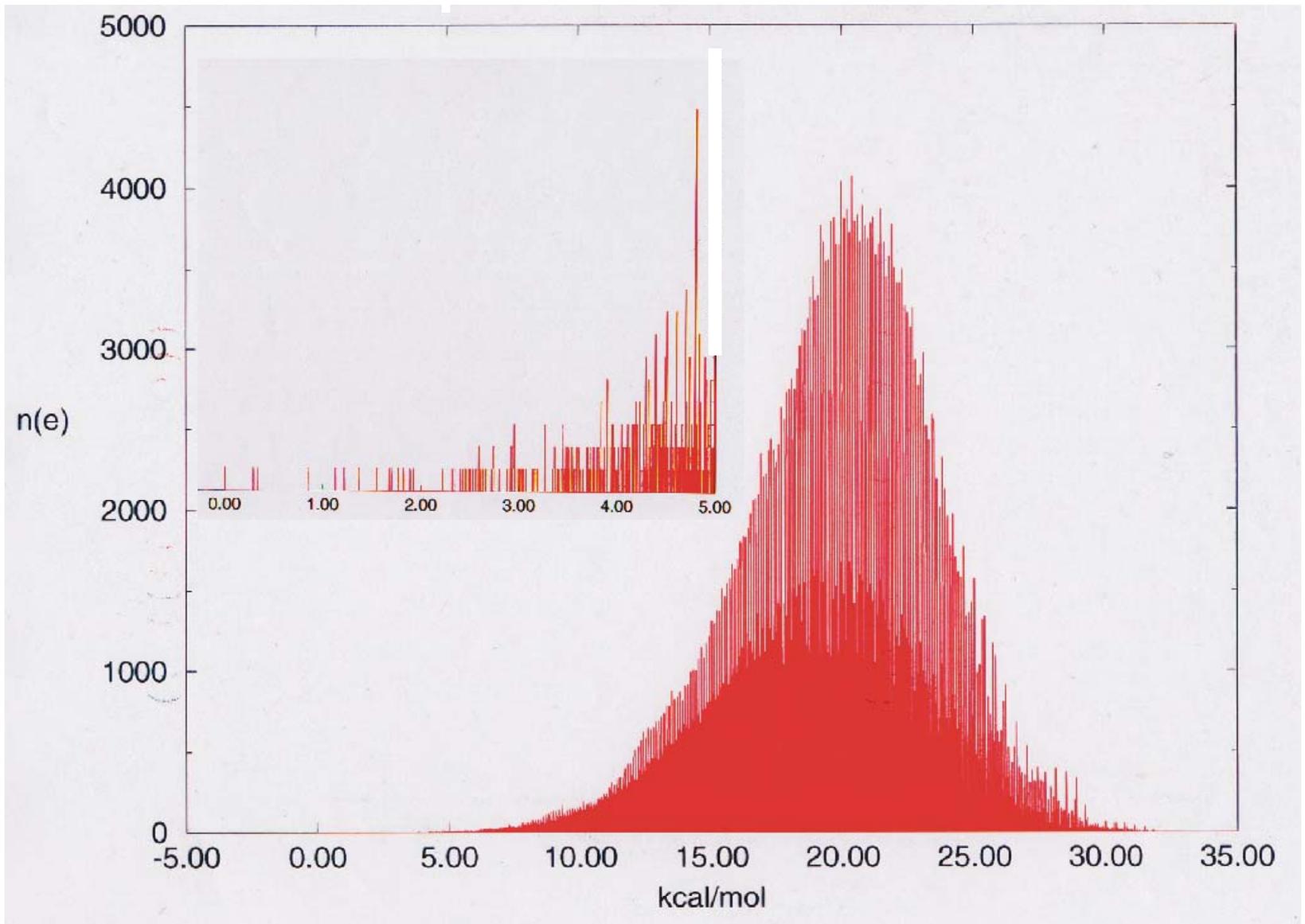


**Minimum free energy structure**

**AAAGGGCACAGGGUGAUUUCAAUAAUUUUA**

**Sequence**

Example of a small RNA molecule:  $n=30$



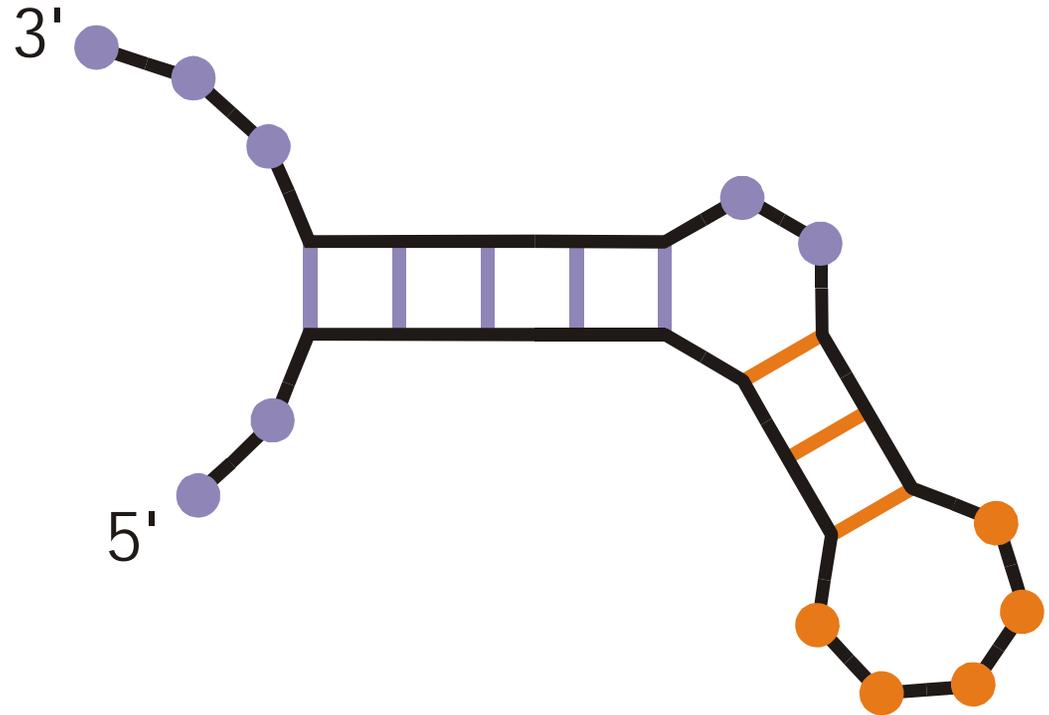
Density of states of suboptimal structures of the RNA molecule with the sequence:

**AAAGGGCACAGGGUGAUUUCAAUAAUUUUA**

## Partition Function of RNA Secondary Structures

John S. McCaskill. *The equilibrium function and base pair binding probabilities for RNA secondary structure*. Biopolymers **29** (1990), 1105-1119

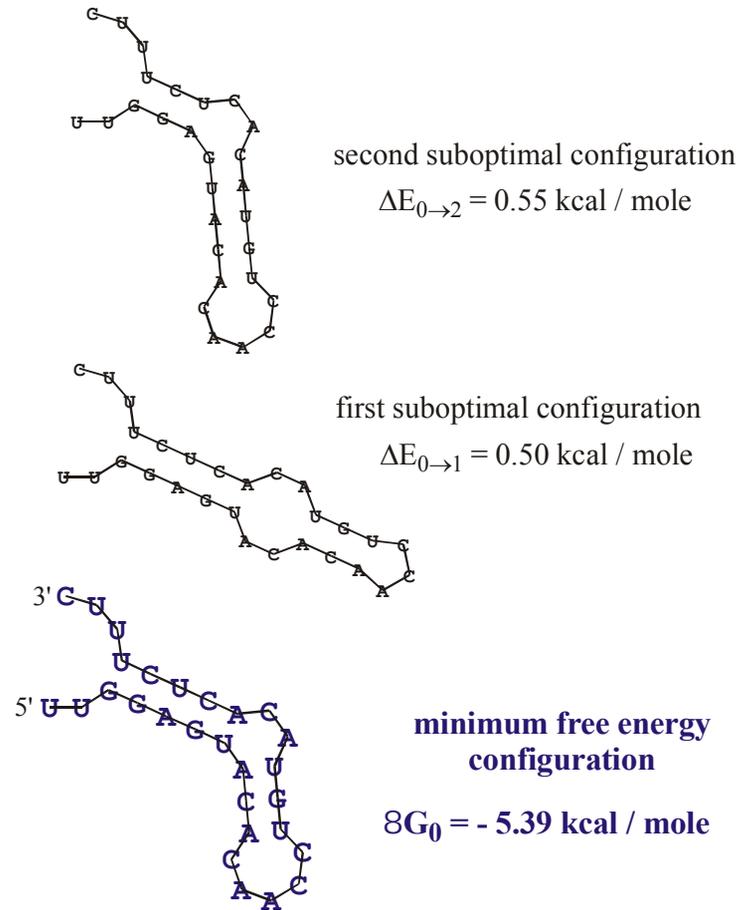
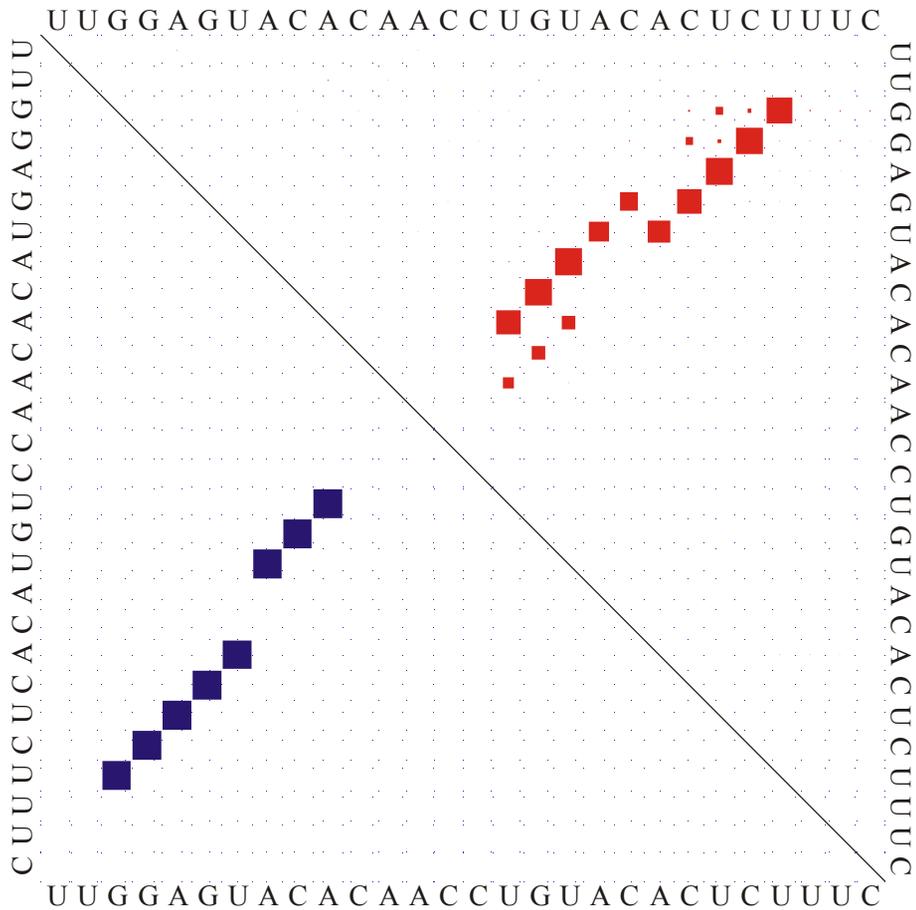
Ivo L. Hofacker, Walter Fontana, Peter F. Stadler, L. Sebastian Bonhoeffer, Manfred Tacker, Peter Schuster. *Fast folding and comparison of RNA secondary structures*. Monatshefte für Chemie **125** (1994), 167-188



**UUGGAGUACACAAC CUGUACACUCUUUC**

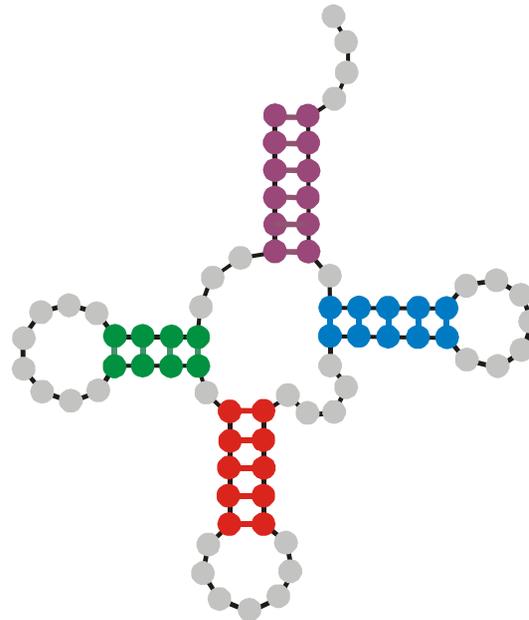
Example of a small RNA molecule with two low-lying suboptimal conformations which contribute substantially to the partition function

Example of a small RNA molecule:  $n=28$



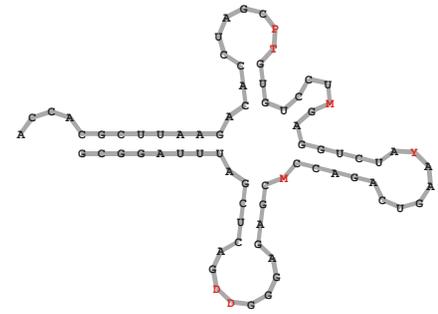
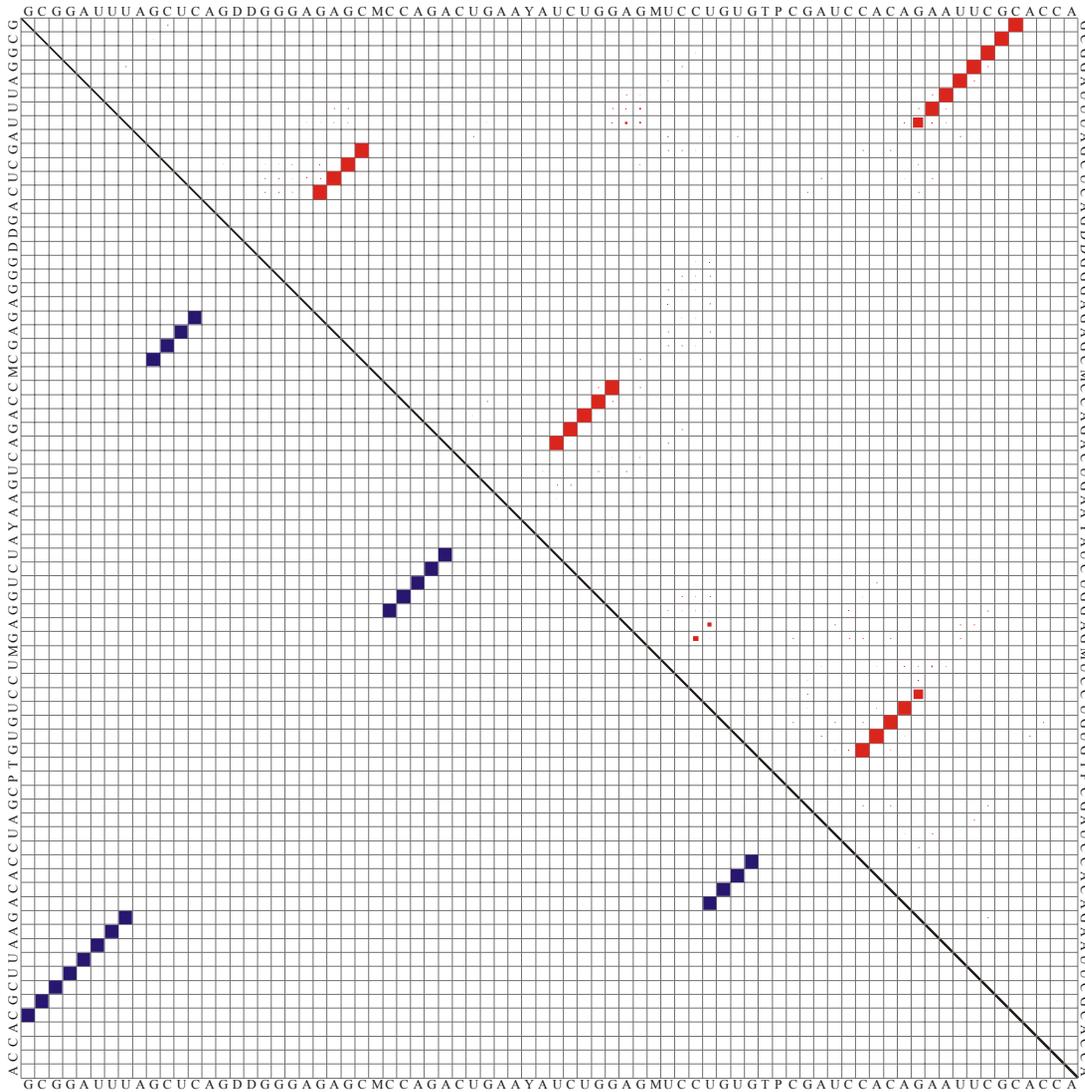
„Dot plot“ of the minimum free energy structure (**lower triangle**) and the partition function (**upper triangle**) of a small RNA molecule (n=28) with low energy suboptimal configurations

GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA  
GCGGAUUUAGCUCAGDDGGGAGAGCMCCAGACUGAAYAUCUGGAGMUCUGUGTPCGAUCCACAGAAUUCGCACCA



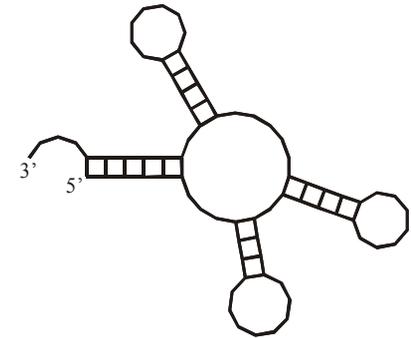
Phenylalanyl-tRNA as an example for the computation of the partition function





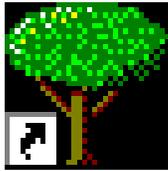
first suboptimal configuration

$$\Delta E_{0 \rightarrow 1} = 0.94 \text{ kcal / mole}$$



tRNA<sup>phe</sup>

with modified bases



RNASTudio.Ink

**GGCGCGCCCGGCGCC**

**GUAUCGAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA**

**UGGUUACGCGUUGGGGUAACGAAGAUUCCGAGAGGAGUUUAGUGACUAGAGG**

Suboptimal structures, partitionfunction, „dot“- and „mountain“-plots

## **Kinetic Folding of RNA Secondary Structures**

Christoph Flamm, Walter Fontana, Ivo L. Hofacker, Peter Schuster. *RNA folding kinetics at elementary step resolution*. RNA **6**:325-338, 2000

Christoph Flamm, Ivo L. Hofacker, Sebastian Maurer-Stroh, Peter F. Stadler, Martin Zehl. *Design of multistable RNA molecules*. RNA **7**:325-338, 2001

## The Folding Algorithm

A sequence  $\mathbf{I}$  specifies an energy ordered set of compatible structures  $\mathbf{S}(\mathbf{I})$ :

$$\mathbf{S}(\mathbf{I}) = \{\mathbf{S}_0, \mathbf{S}_1, \dots, \mathbf{S}_m, \mathbf{O}\}$$

A trajectory  $T_k(\mathbf{I})$  is a time ordered series of structures in  $\mathbf{S}(\mathbf{I})$ . A folding trajectory is defined by starting with the open chain  $\mathbf{O}$  and ending with the global minimum free energy structure  $\mathbf{S}_0$  or a metastable structure  $\mathbf{S}_k$  which represents a local energy minimum:

$$T_0(\mathbf{I}) = \{\mathbf{O}, \mathbf{S}(1), \dots, \mathbf{S}(t-1), \mathbf{S}(t), \\ \mathbf{S}(t+1), \dots, \mathbf{S}_0\}$$

$$T_k(\mathbf{I}) = \{\mathbf{O}, \mathbf{S}(1), \dots, \mathbf{S}(t-1), \mathbf{S}(t), \\ \mathbf{S}(t+1), \dots, \mathbf{S}_k\}$$

Transition probabilities  $P_{ij}(t) = \text{Prob}\{\mathbf{S}_i \rightarrow \mathbf{S}_j\}$  are defined by

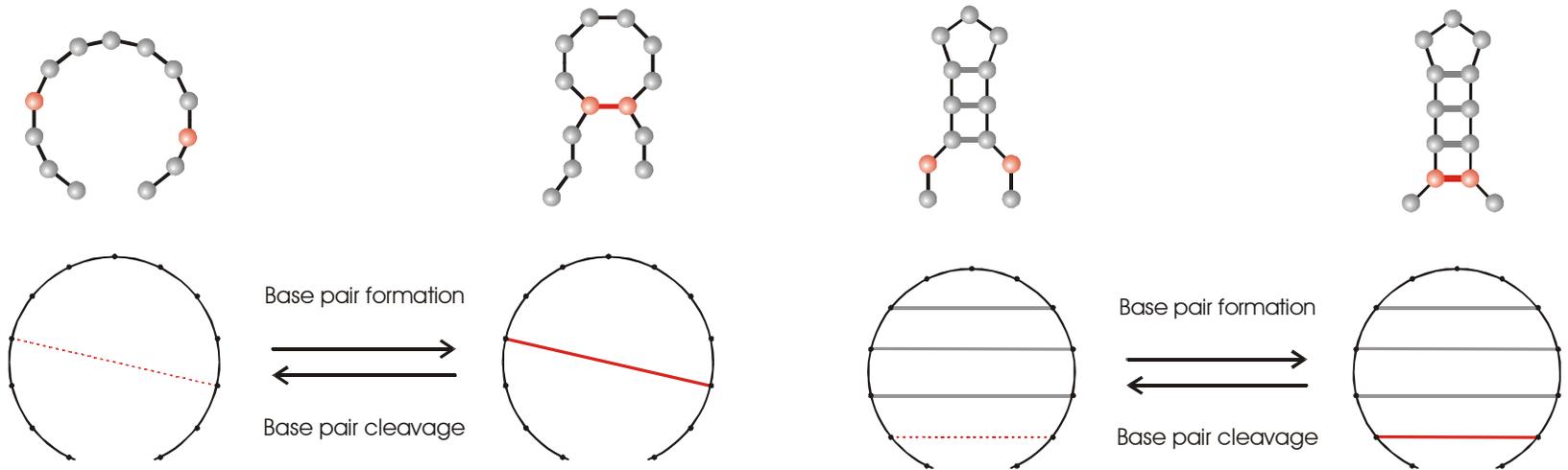
$$P_{ij}(t) = P_i(t) k_{ij} = P_i(t) \exp(-\Delta G_{ij}/2RT) / \Sigma_i$$

$$P_{ji}(t) = P_j(t) k_{ji} = P_j(t) \exp(-\Delta G_{ji}/2RT) / \Sigma_j$$

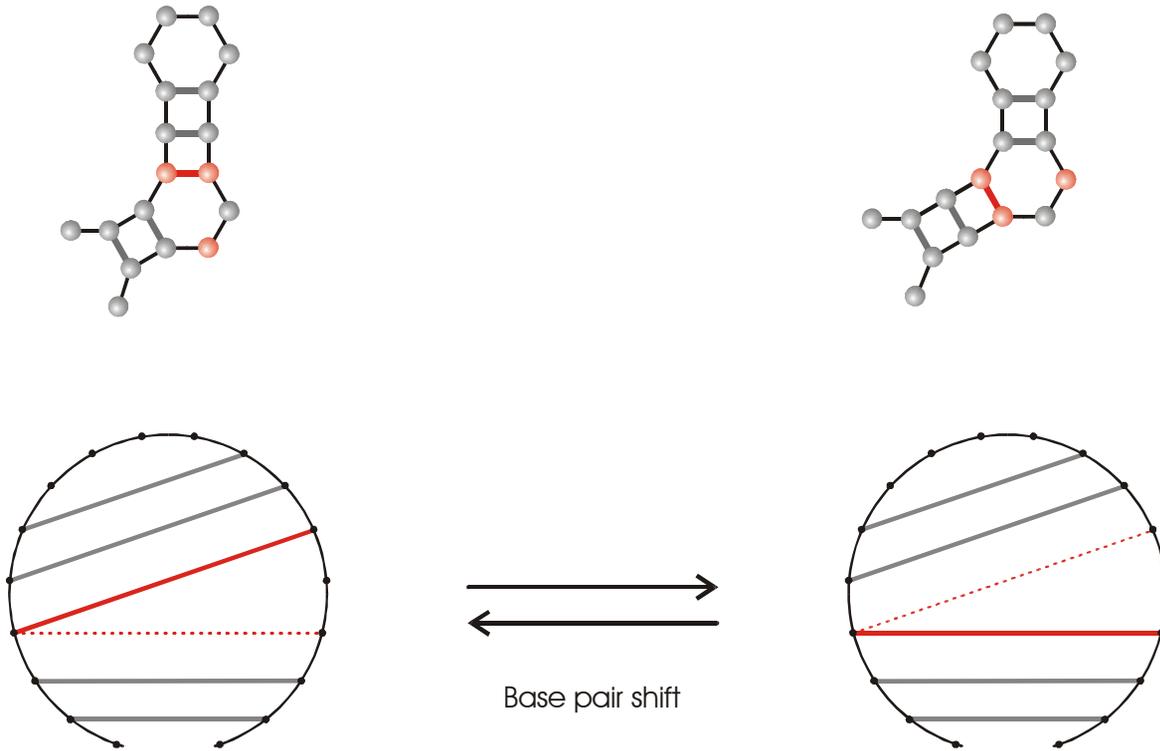
$$\Sigma_k = \sum_{k=1, k \neq i}^{m+2} \exp(-\Delta G_{ki}/2RT)$$

The symmetric rule for transition rate parameters is due to Kawasaki (K. Kawasaki, *Diffusion constants near the critical point for time dependent Ising models*. Phys.Rev. **145**:224-230, 1966).

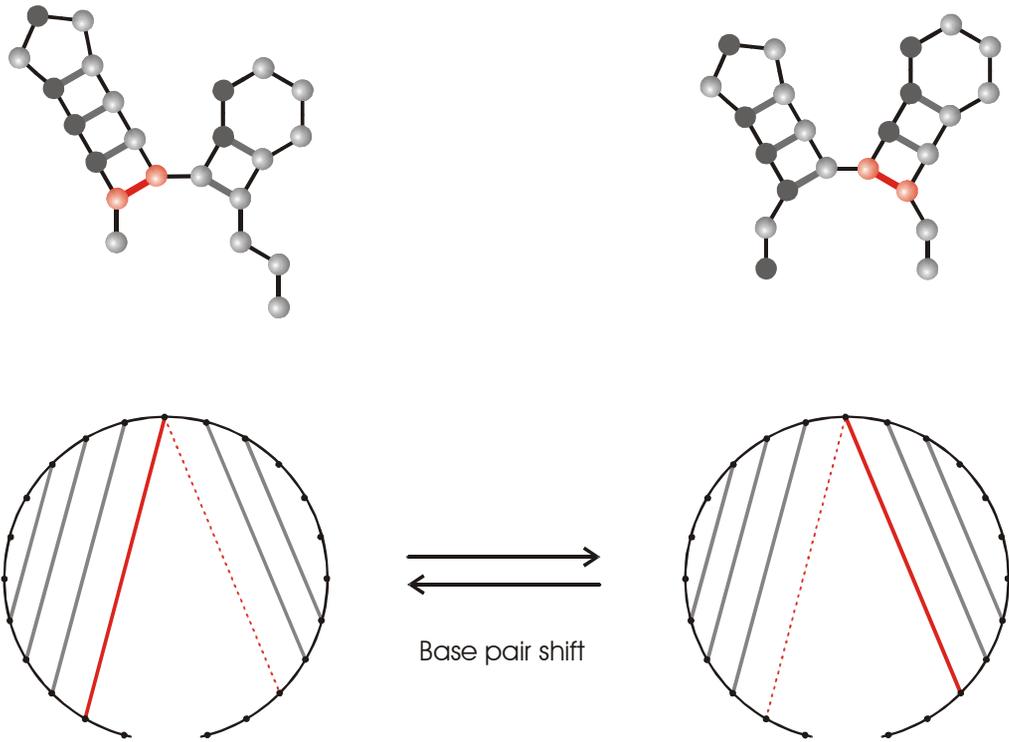
Formulation of kinetic RNA folding as a stochastic process



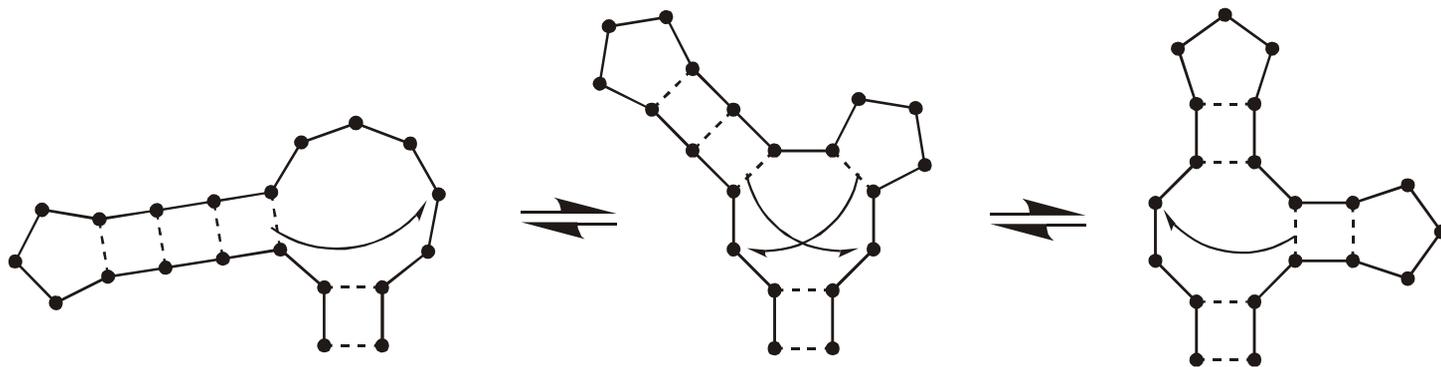
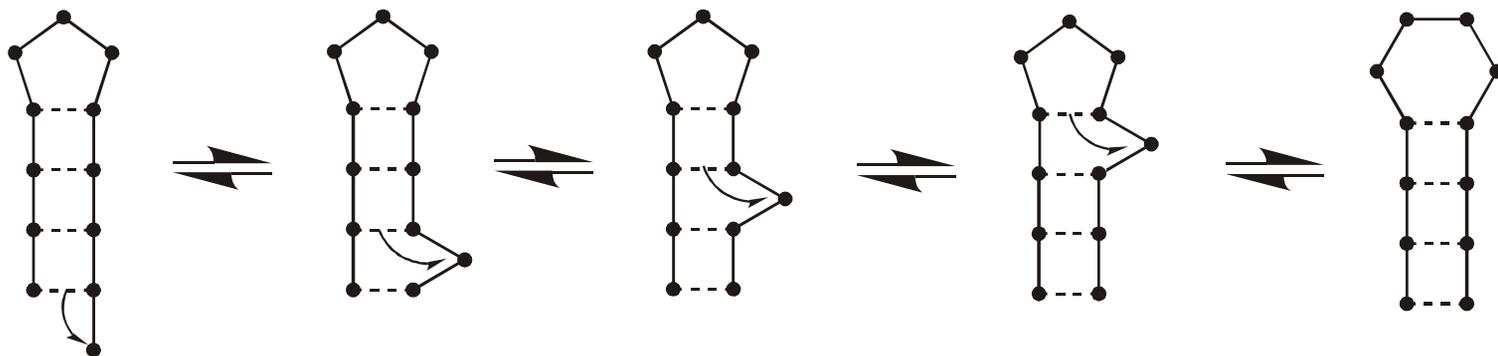
Base pair formation and base pair cleavage moves for nucleation and elongation of stacks



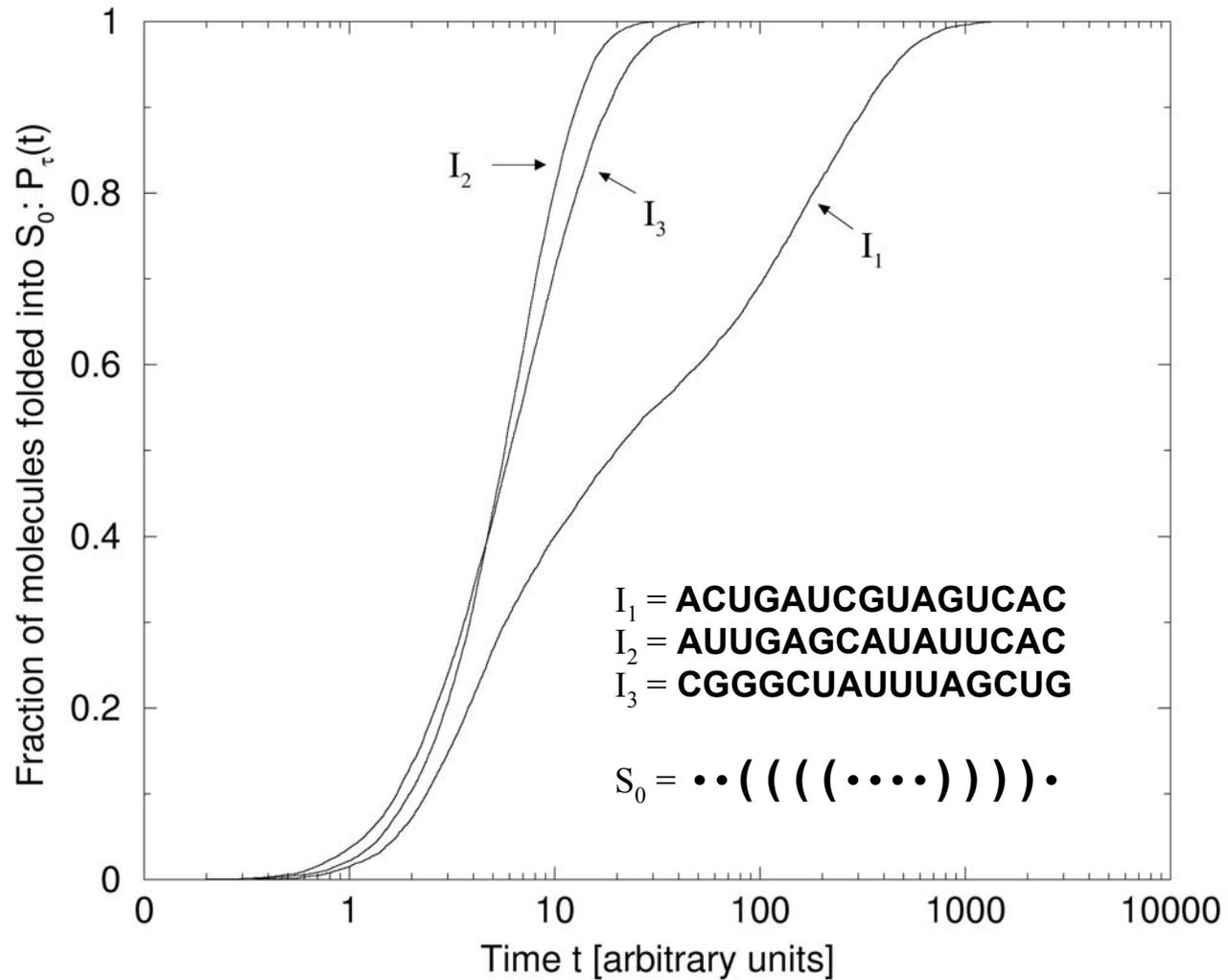
Base pair shift move of class 1: Shift inside internal loops or bulges



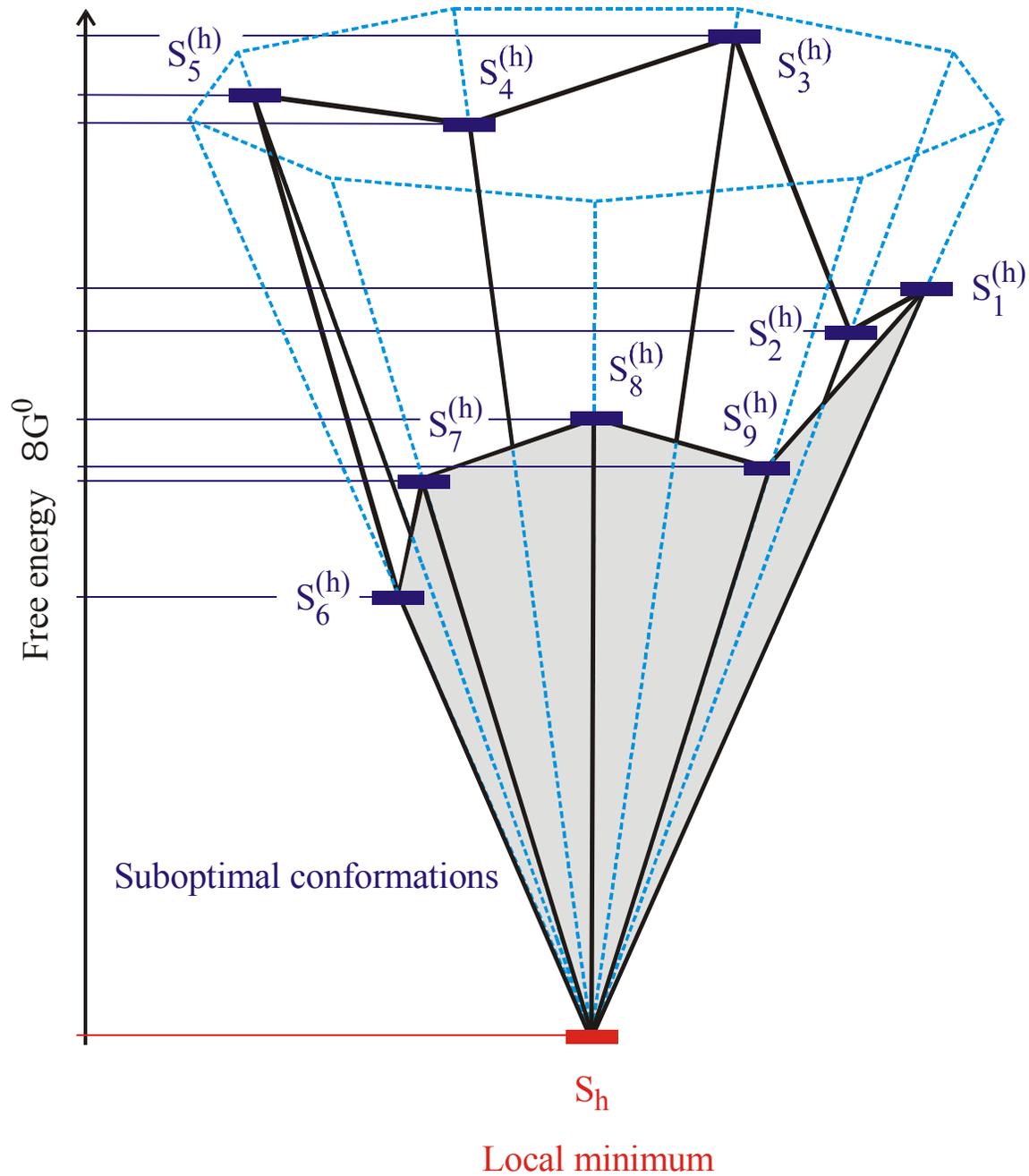
Base pair shift move of class 2: Shift involving free ends



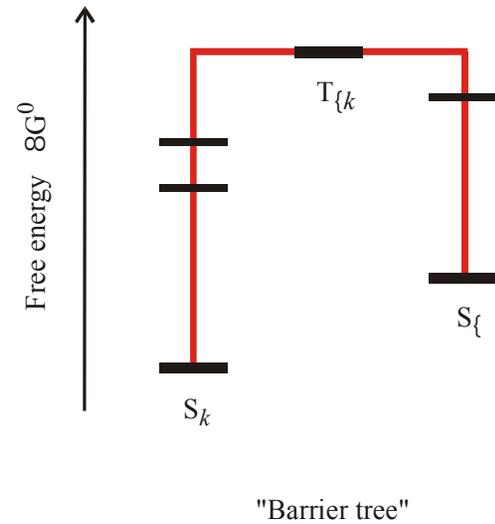
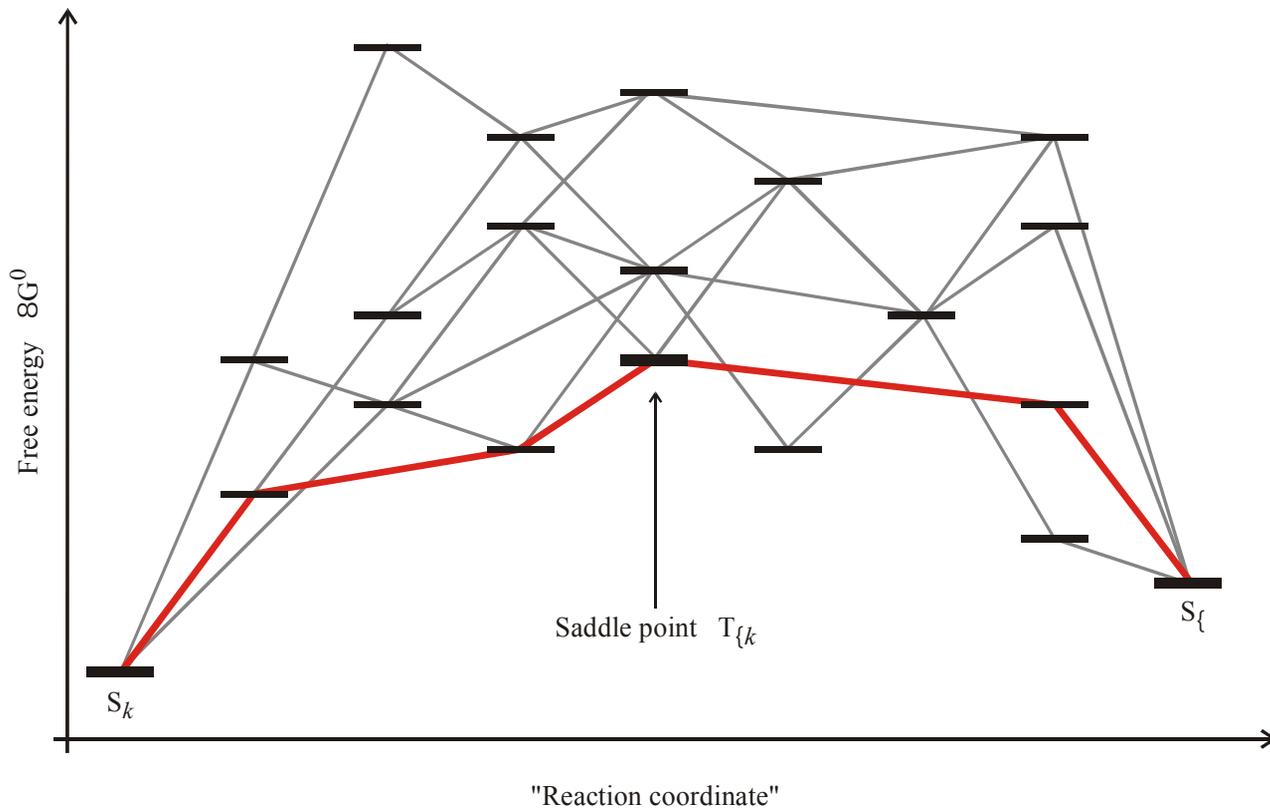
Examples of rearrangements through consecutive shift moves



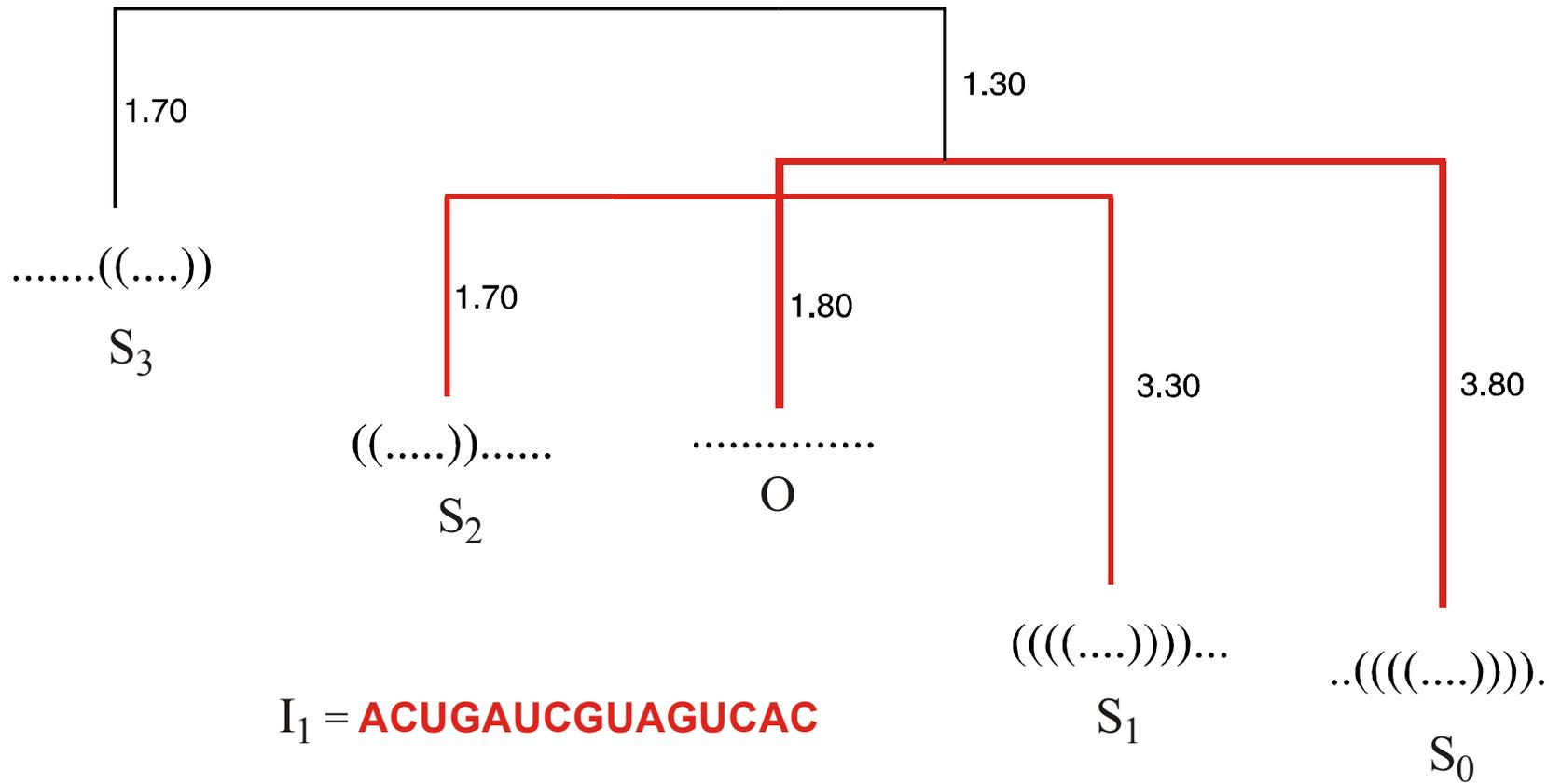
Mean folding curves for three small RNA molecules with different folding behavior



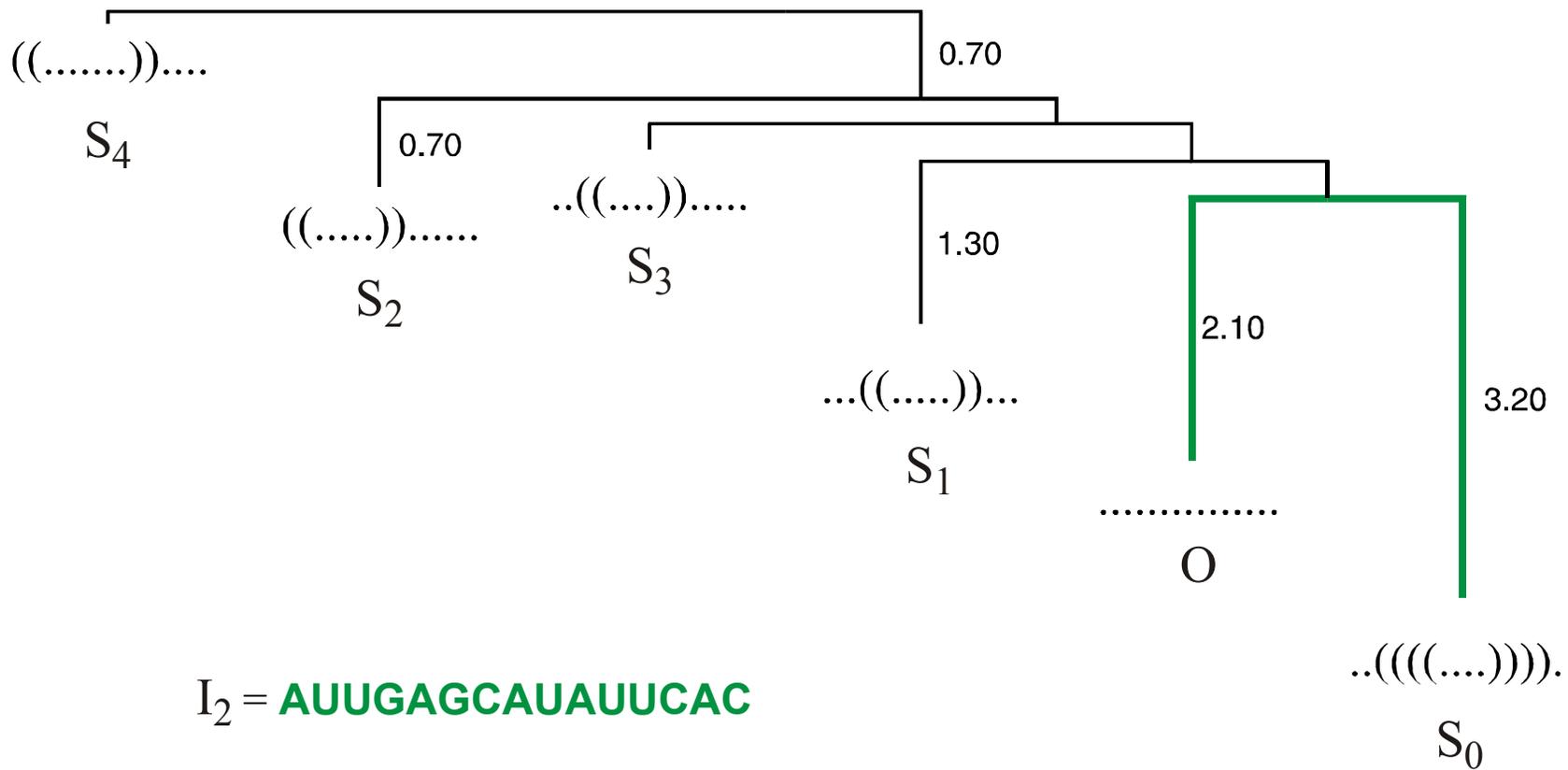
Search for local minima in conformation space



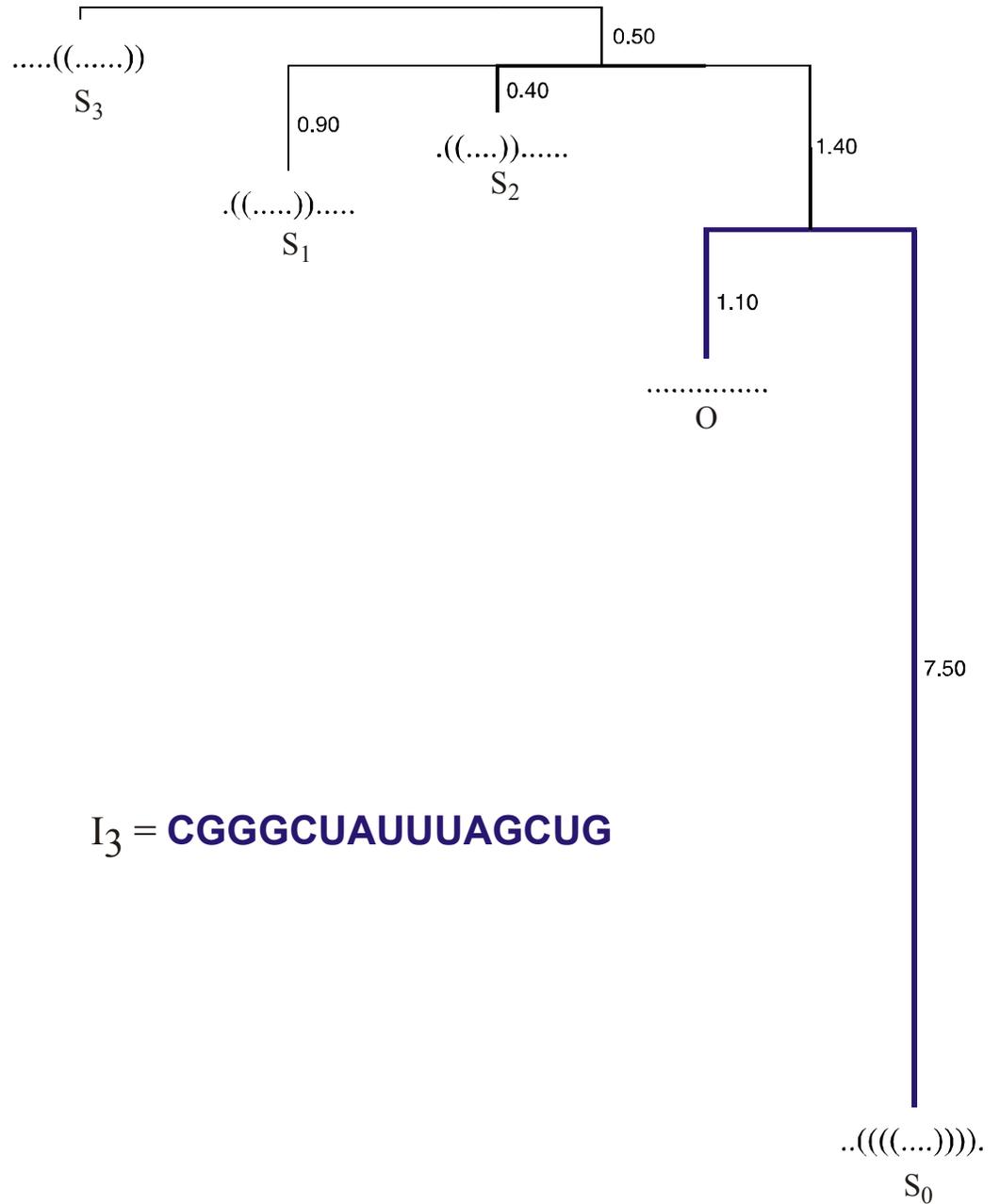
Definition of a 'barrier tree'



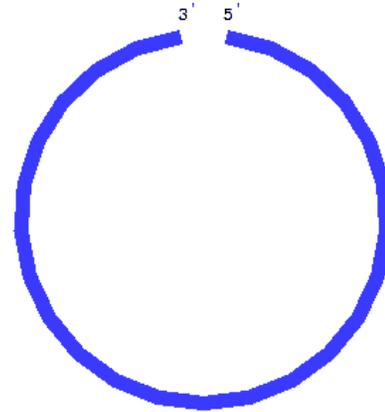
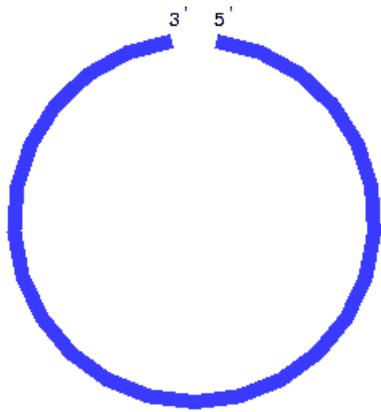
Example of an unefficiently folding small RNA molecule with  $n = 15$



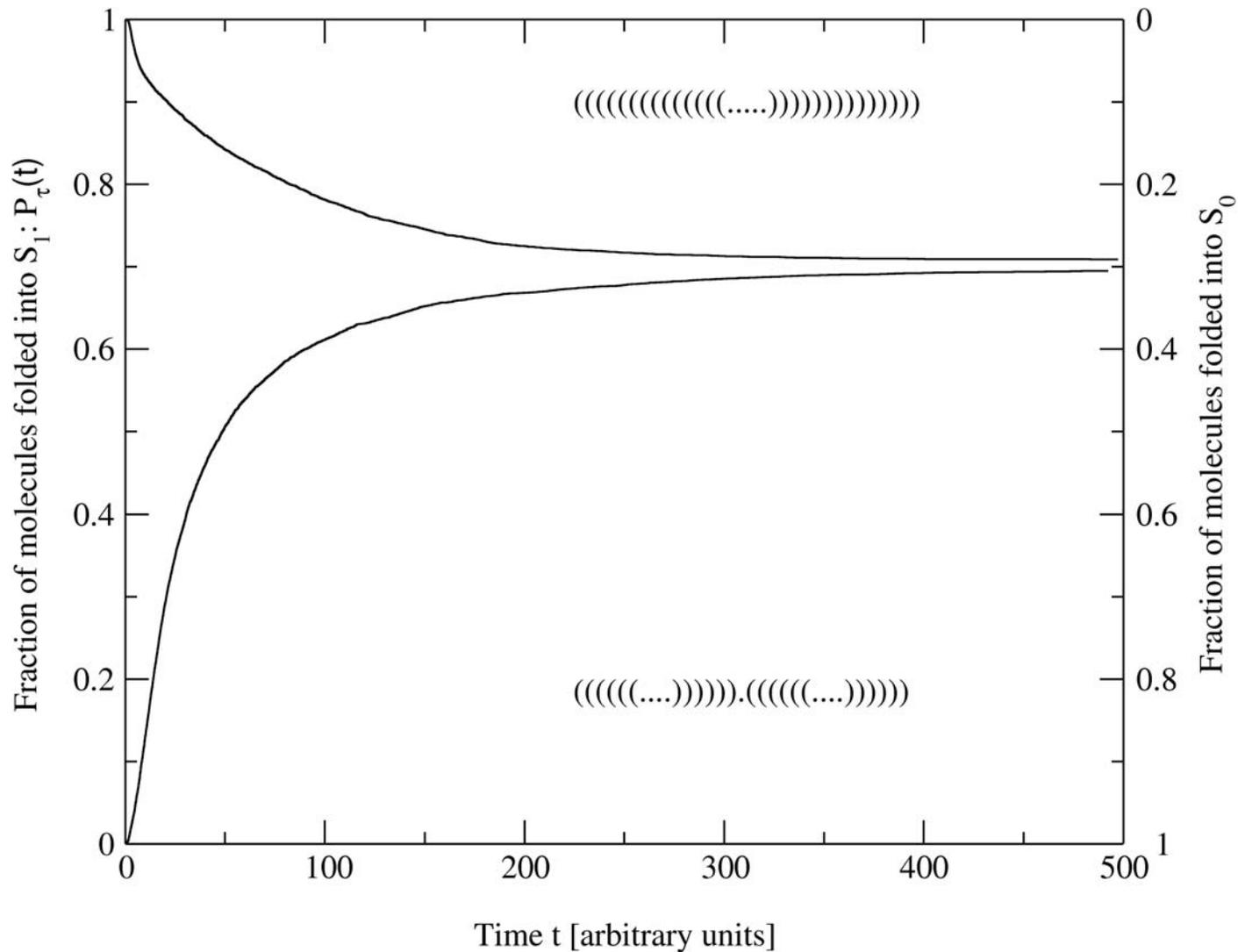
Example of an easily folding small RNA molecule with  $n = 15$



Example of an easily folding  
and especially stable small  
RNA molecule with  $n = 15$



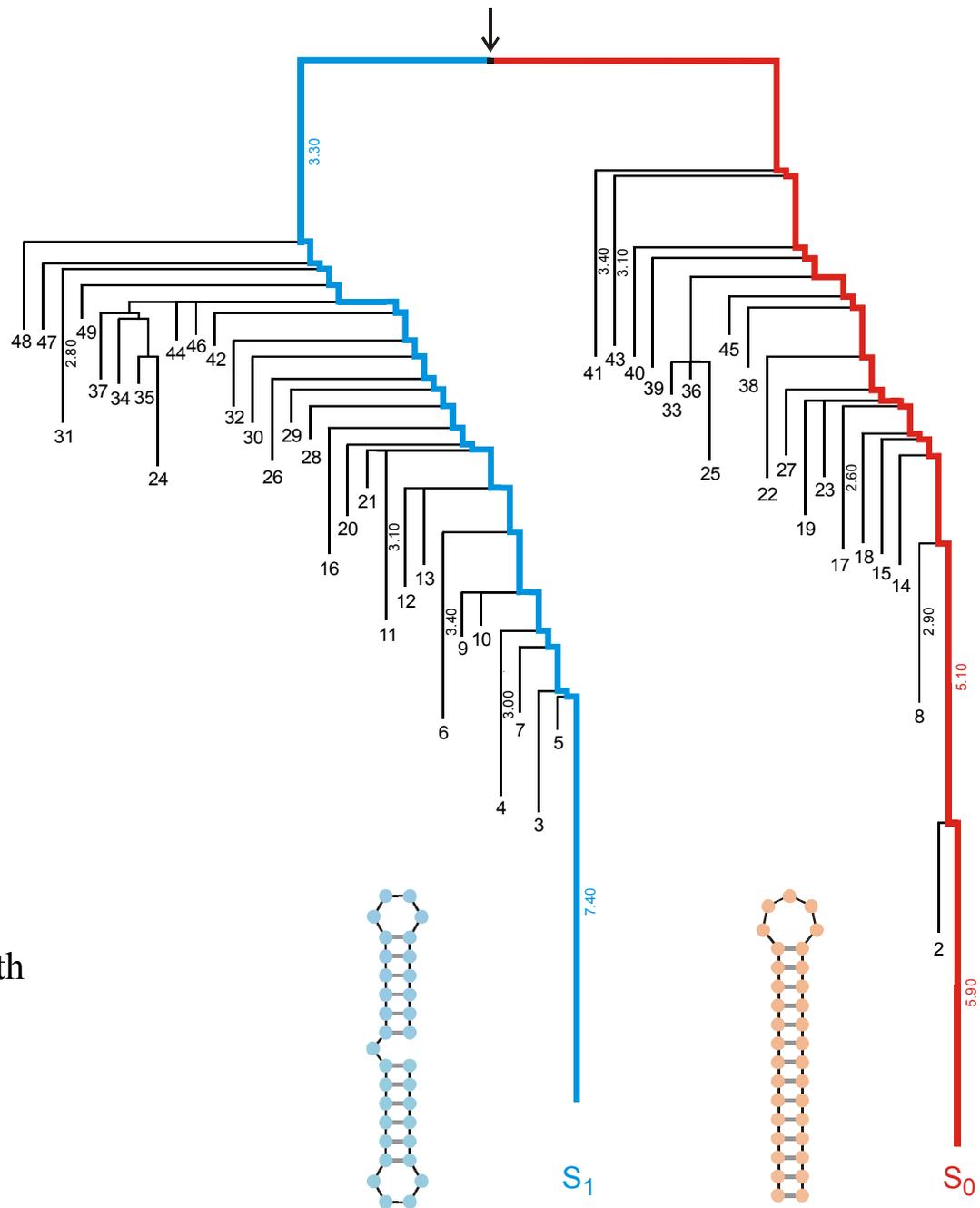
Examples of two folding trajectories leading to different local minima

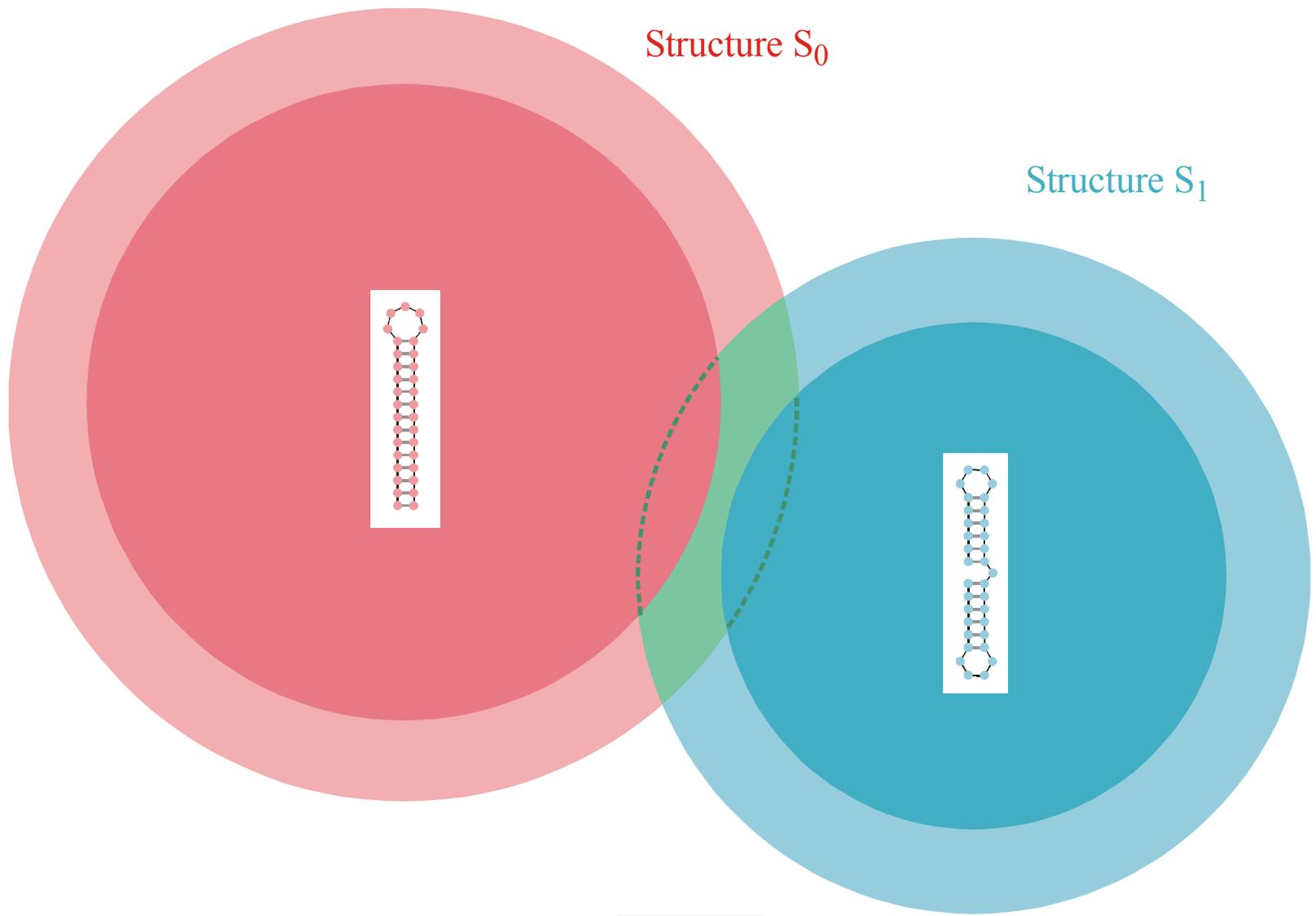


Folding dynamics of the sequence **GGCCCUUUGGGGCGAGACCCUAAAAAGGGUC**



Barrier tree of a sequence with two conformations





**Intersection** of two compatible sets:  $C_0 \cap C_1$

The intersection of two compatible sets is always non empty:  $C_0 \cap C_1 \neq \emptyset$



S0092-8240(96)00089-4

## GENERIC PROPERTIES OF COMBINATORIAL MAPS: NEUTRAL NETWORKS OF RNA SECONDARY STRUCTURES<sup>1</sup>

■ CHRISTIAN REIDYS\*, †, PETER F. STADLER\*, ‡ and PETER SCHUSTER\*, ‡, §, ¶

\*Santa Fe Institute,  
 Santa Fe, NM 87501, U.S.A.

†Los Alamos National Laboratory,  
 Los Alamos, NM 87545, U.S.A.

‡Institut für Theoretische Chemie der Universität Wien,  
 A-1090 Wien, Austria

§Institut für Molekulare Biotechnologie,  
 D-07708 Jena, Germany

(E.mail: pks@tbi.univie.ac.at)

Random graph theory is used to model and analyse the relationships between sequences and secondary structures of RNA molecules, which are understood as mappings from sequence space into shape space. These maps are non-invertible since there are always many orders of magnitude more sequences than structures. Sequences folding into identical structures form *neutral networks*. A neutral network is embedded in the set of sequences that are *compatible* with the given structure. Networks are modeled as graphs and constructed by random choice of vertices from the space of compatible sequences. The theory characterizes neutral networks by the mean fraction of neutral neighbors ( $\lambda$ ). The networks are connected and percolate sequence space if the fraction of neutral nearest neighbors exceeds a threshold value ( $\lambda > \lambda^*$ ). Below threshold ( $\lambda < \lambda^*$ ), the networks are partitioned into a largest “giant” component and several smaller components. Structures are classified as “common” or “rare” according to the sizes of their pre-images, i.e. according to the fractions of sequences folding into them. The neutral networks of any pair of two different common structures almost touch each other, and, as expressed by the conjecture of *shape space covering* sequences folding into almost all common structures, can be found in a small ball of an arbitrary location in sequence space. The results from random graph theory are compared to data obtained by folding large samples of RNA sequences. Differences are explained in terms of specific features of RNA molecular structures. © 1997 Society for Mathematical Biology

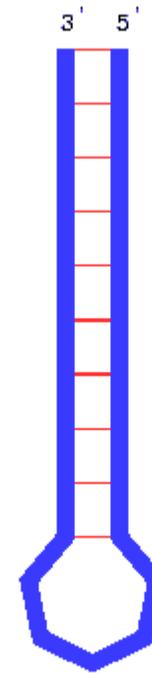
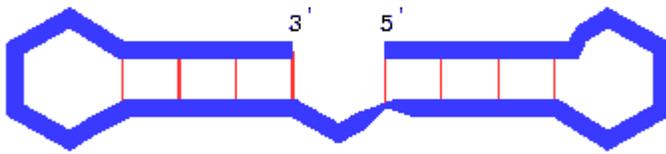
**THEOREM 5. INTERSECTION-THEOREM.** *Let  $s$  and  $s'$  be arbitrary secondary structures and  $C[s], C[s']$  their corresponding compatible sequences. Then,*

$$C[s] \cap C[s'] \neq \emptyset.$$

*Proof.* Suppose that the alphabet admits only the complementary base pair  $[XY]$  and we ask for a sequence  $x$  compatible to both  $s$  and  $s'$ . Then  $f(s, s') \cong D_m$  operates on the set of all positions  $\{x_1, \dots, x_n\}$ . Since we have the operation of a dihedral group, the orbits are either cycles or chains and the cycles have even order. A constraint for the sequence compatible to both structures appears only in the cycles where the choice of bases is not independent. It remains to be shown that there is a valid choice of bases for each cycle, which is obvious since these have even order. Therefore, it suffices to choose an alternating sequence of the pairing partners  $X$  and  $Y$ . Thus, there are at least two different choices for the first base in the orbit. ■

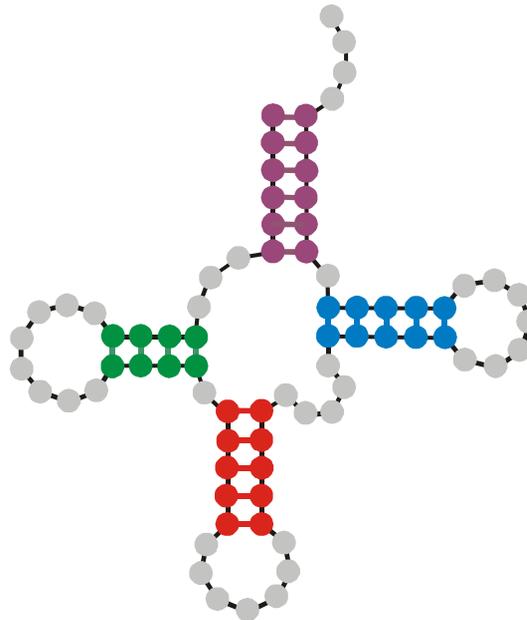
*Remark.* A generalization of the statement of theorem 5 to three different structures is false.

Reference for the definition of the intersection and the proof of the **intersection theorem**

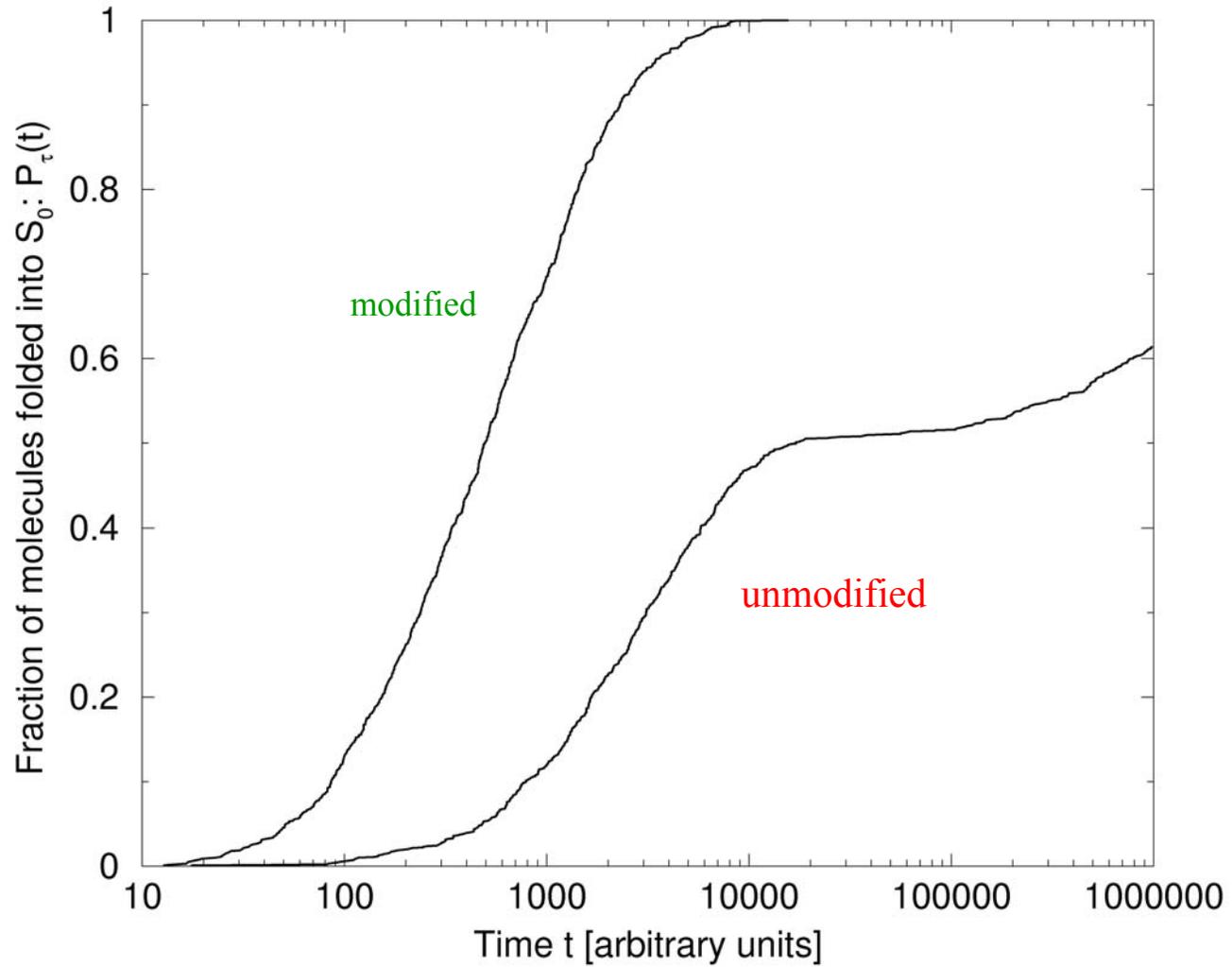


Kinetics RNA refolding between a long living metastable conformation  
and the minimum free energy structure

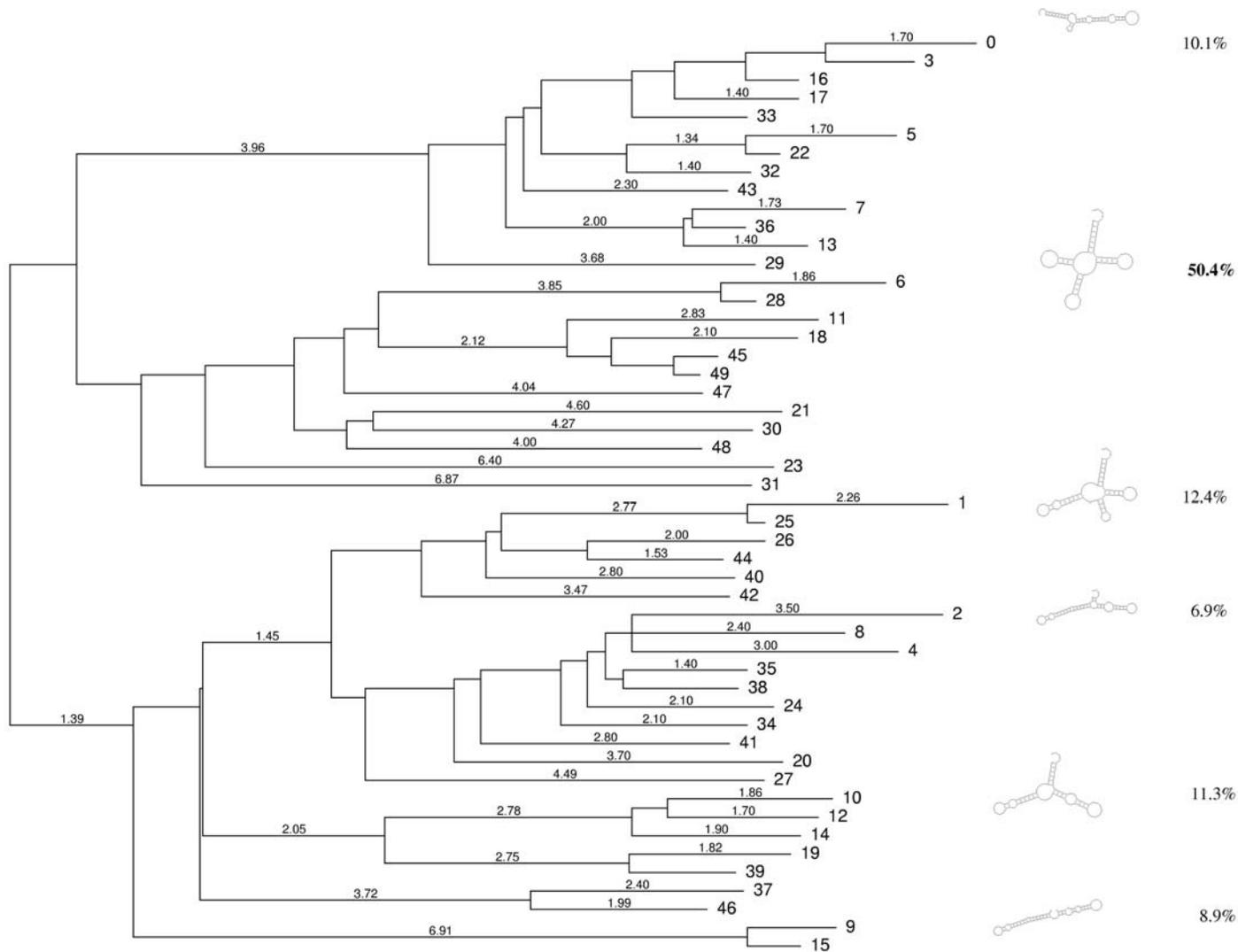
GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUC CUGUGUUCGAUCCACAGAAUUCGCACCA  
GCGGAUUUAGCUCAGDDGGGAGAGCMCCAGACUGAAYAUCUGGAGMUC CUGUGTPCGAUCCACAGAAUUCGCACCA



Kinetid folding of phenylalanyl-tRNA



Folding dynamics of tRNA<sup>phe</sup> with and without modified nucleotides



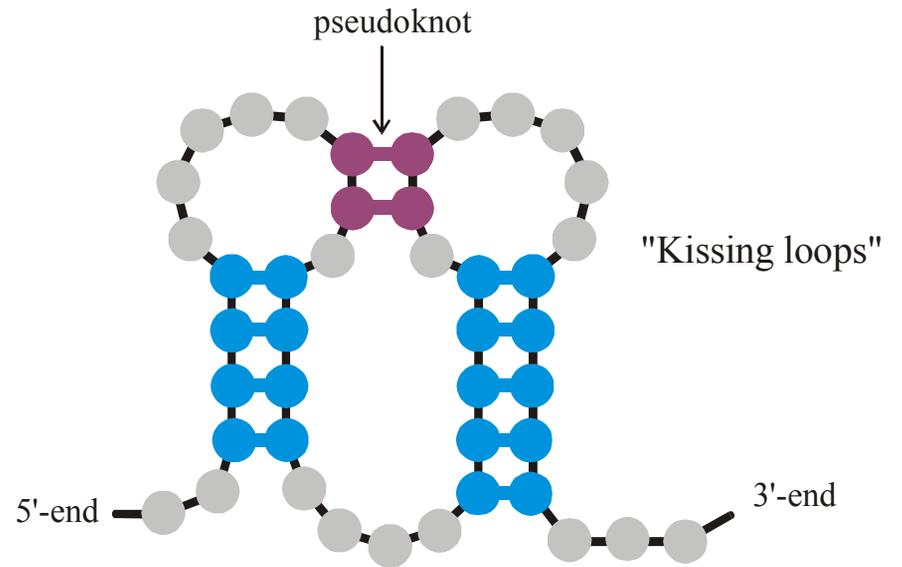
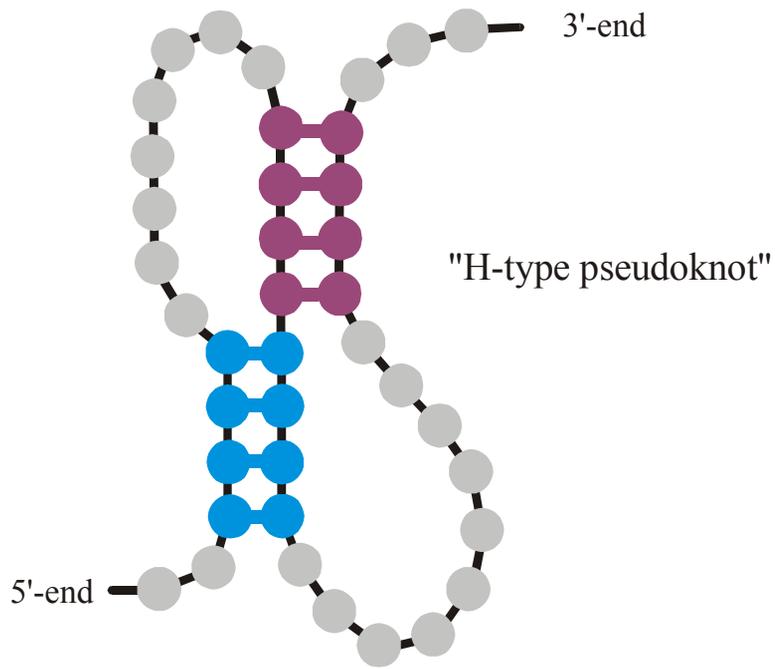
Barrier tree of tRNA<sup>phe</sup> without modified nucleotides

## **Tertiary elements in RNA structure**

1. Different classes of pseudoknots
2. Different classes of non-Watson-Crick base pairs
3. Base triplets, G-quartets, A-platforms, etc.
4. End-on-end stacking of double helices
5. Divalent metal ion complexes,  $Mg^{2+}$ , etc.
6. Other interactions involving phosphate, 2'-OH, etc.

## Tertiary elements in RNA structure

- 1. Different classes of pseudoknots**
2. Different classes of non-Watson-Crick base pairs
3. Base triplets, G-quartets, A-platforms, etc.
4. End-on-end stacking of double helices
5. Divalent metal ion complexes,  $Mg^{2+}$ , etc.
6. Other interactions involving phosphate, 2'-OH, etc.

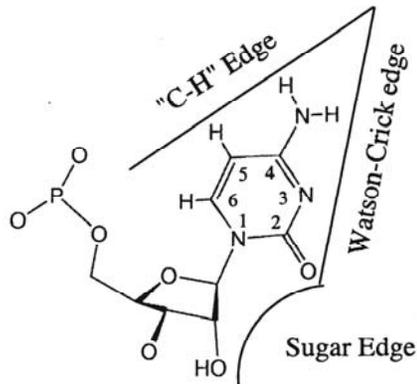
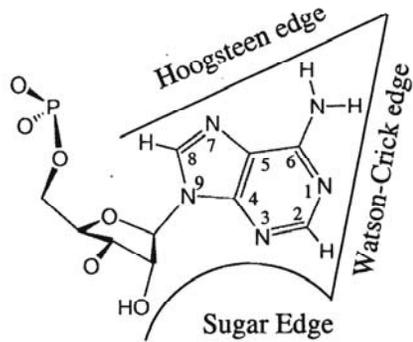


Two classes of pseudoknots in RNA structures

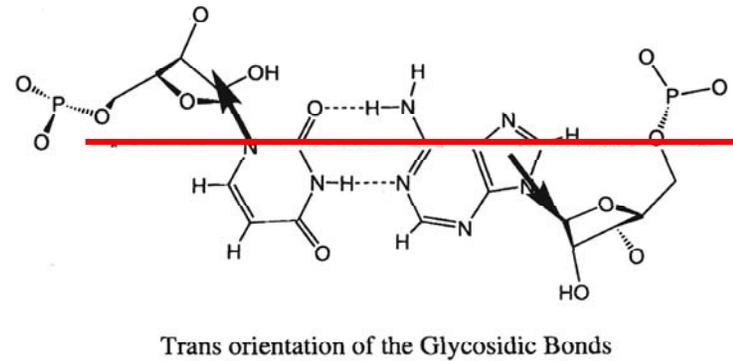
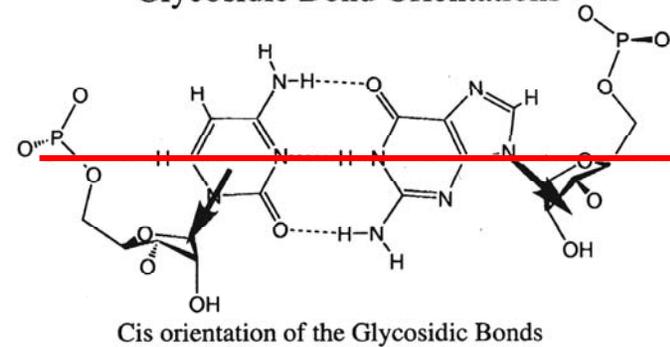
## Tertiary elements in RNA structure

1. Different classes of pseudoknots
- 2. Different classes of non-Watson-Crick base pairs**
3. Base triplets, G-quartets, A-platforms, etc.
4. End-on-end stacking of double helices
5. Divalent metal ion complexes,  $Mg^{2+}$ , etc.
6. Other interactions involving phosphate, 2'-OH, etc.

## Interacting Edges



## Glycosidic Bond Orientations



## Twelve families of base pairs

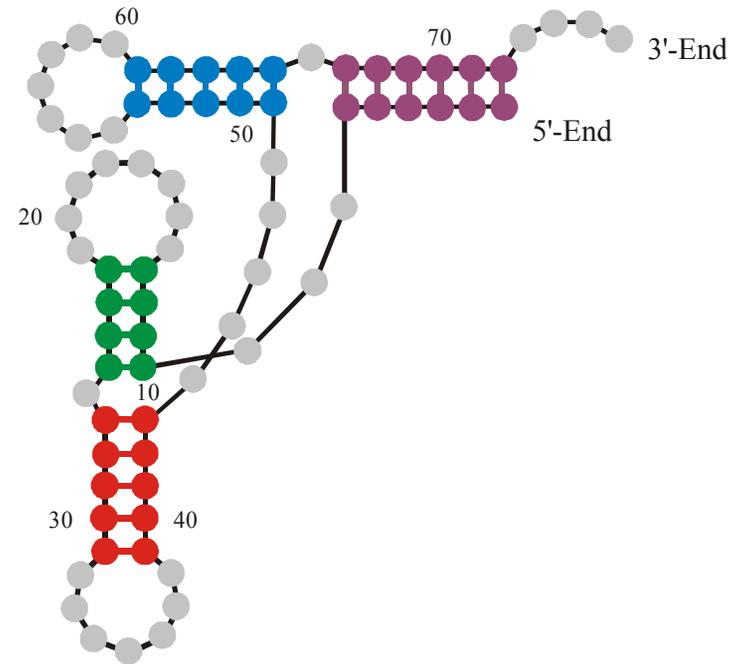
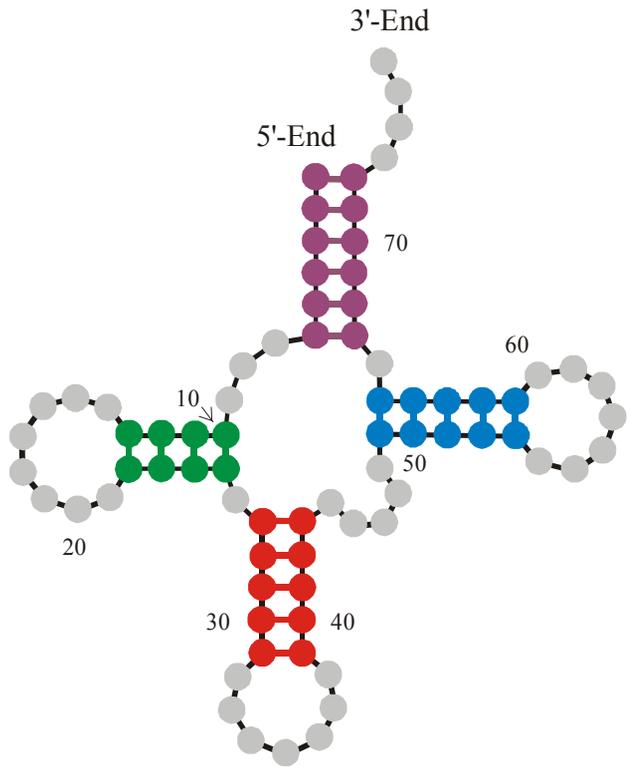
Watson-Crick / Hogsteen / Sugar  
edge

Cis / Trans  
orientation

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

## Tertiary elements in RNA structure

1. Different classes of pseudoknots
2. Different classes of non-Watson-Crick base pairs
3. Base triplets, G-quartets, A-platforms, etc.
- 4. End-on-end stacking of double helices**
5. Divalent metal ion complexes,  $Mg^{2+}$ , etc.
6. Other interactions involving phosphate, 2'-OH, etc.



End-on-end stacking of double helical regions yields the L-shape of tRNA<sup>Phe</sup>

## Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)

Projects No. 09942, 10578, 11065, 13093  
13887, and 14898

Jubiläumsfonds der Österreichischen Nationalbank

Project No. Nat-7813

European Commission: Project No. EU-980189

Siemens AG, Austria

The Santa Fe Institute and the Universität Wien

The software for producing RNA movies was developed by  
Robert Giegerich and coworkers at the Universität Bielefeld



Universität Wien

# Coworkers



Universität Wien

**Walter Fontana**, Santa Fe Institute, NM

**Christian Reidys, Christian Forst**, Los Alamos National Laboratory, NM

**Peter Stadler, Bärbel Stadler**, Universität Leipzig, GE

**Ivo L.Hofacker, Christoph Flamm**, Universität Wien, AT

**Andreas Wernitznig, Michael Kospach**, Universität Wien, AT

**Ulrike Langhammer, Ulrike Mückstein, Stefanie Widder**

**Jan Cupal, Kurt Grünberger, Andreas Svrček-Seiler, Stefan Wuchty**

**Ulrike Göbel**, Institut für Molekulare Biotechnologie, Jena, GE

**Walter Grüner, Stefan Kopp, Jaqueline Weber**

Web-Page for further information:

<http://www.tbi.univie.ac.at/~pks>

