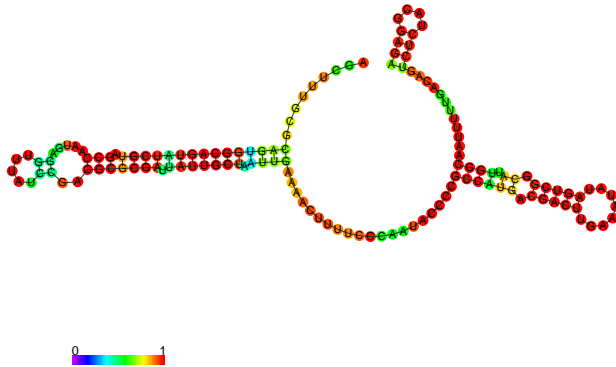


## Solution: RNA 3D modules discovery

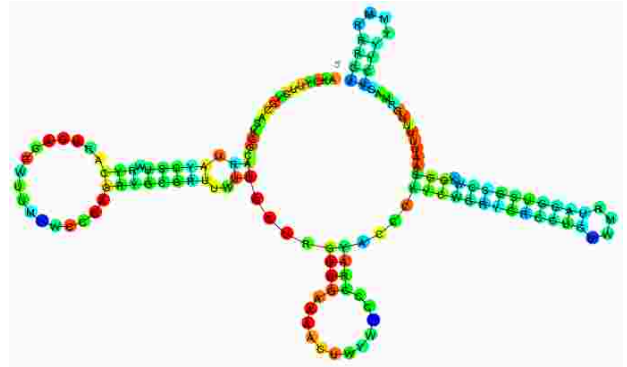
### 1. RNA secondary structure of U4

Three out of four stem-loops are predicted correctly from the human sequence by minimum-free energy folding.

RNAfold:



Rfam:



### 2. RNA 3D modules

1. RMDetect found 3 clusters of 3 different modules: Kink-turn, Tandem-GA, and C-loop (see file RF00015.stk.rmc\_cluster\_overview). Of the 170 sequences in the input (RF00015.stockholm.stk) 140 sequences have a predicted Kink-turn, 118 have a Tandem-GA and 16 have a C-loop.
2. C-loop is only predicted in few sequences and has a small score (describing how well the sequence fits to the bayesian network modelling the C-loop).
3. Tandem-GA has a large occurrence and a good score. However, the proportion of the RNA secondary structure ensemble (set of all possible folds for that sequence) compatible with the 3D module is low (BPP). In addition we see that the TGA cluster overlaps the KT cluster. When you look into the slides of the lecture you see that both TGA and KT consist of two trans oriented hoogsteen-sugar base pairs.
4. The only cluster that satisfies all our criteria pretty well is the cluster 1: Kink-turn. It is conserved in the majority of the sequences in the alignment, has a large score, and high base-pair probability (BPP). Its mutual information (MI) is lowest of all 3 clusters, however, there are still some compensatory base pair changes available in the canonical basepairs flanking the 3D module.

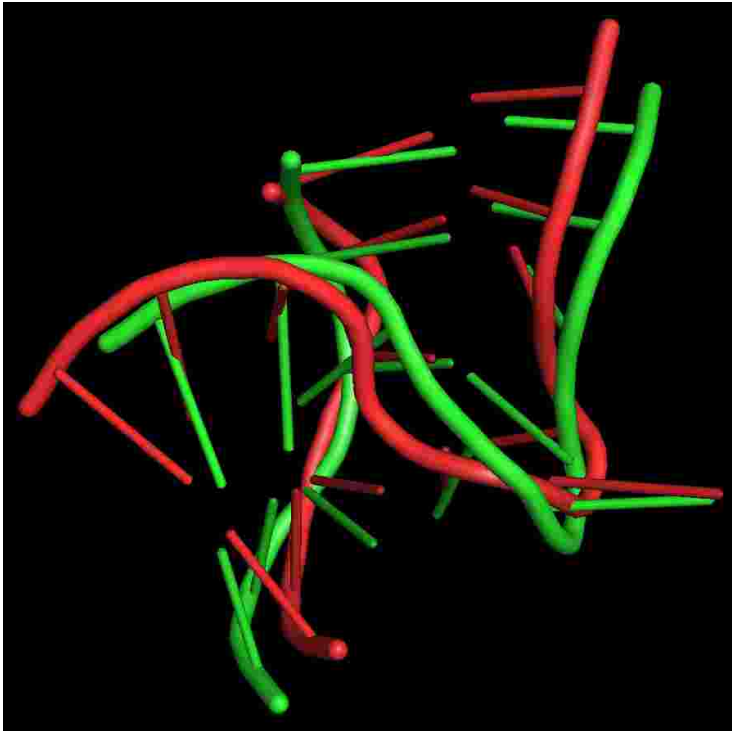
### 3. Integration of 3D modules in 2D structures

MFE structure:

```
AGCUUUGCGCAGUGGCAGUAUCGUAGCCAAUGAGGUUUUAUCCGAGGCGCGAUUAUUGCUAUUGAAAACUUUCCCAAUACCCGCAUGA
.....((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((
.....((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((((
```

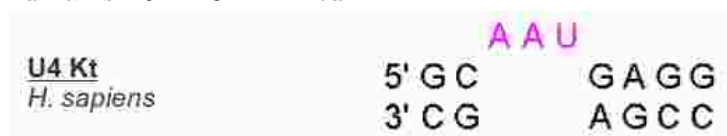
#### 4. Visualization and structure conservation of the kink-turn!

1. 1E7K is the crystal structure of the Spliceosomal 15.5Kd Protein Bound to a U4 snRNA Fragment in human.  
3CC2 is the crystal Structure of the Large Ribosomal Subunit at 2.4 Angstrom Resolution in Haloarcula Marismortui.
2. After aligning them the two modules are 2.914 RMS distant from each other.



3. Despite of their high structure similarity the sequence of the kink-turns in the two molecules is very different.

Human snRNA U4 kink-turn:



*H. marismortui* 23S large ribosomal subunit kink-turn:

