

# The effect of modified nucleotides on RNA structure

---

Yuliia Varenyk

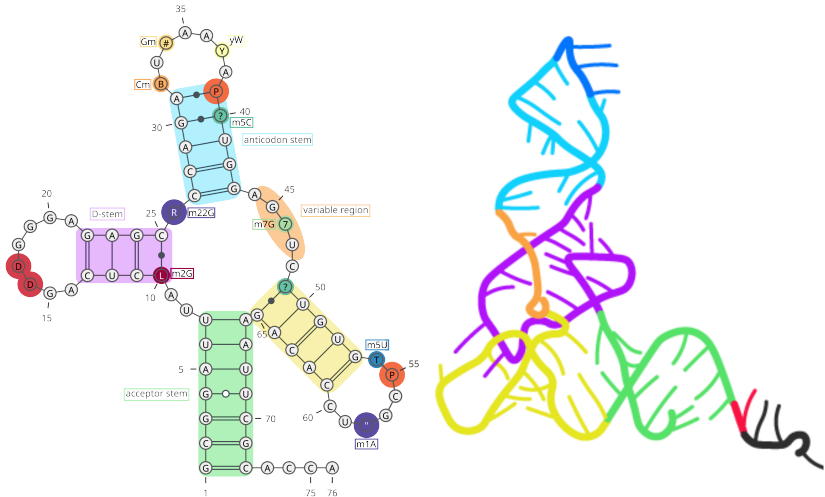
TBI Vienna

3th Winterseminar Bled

February 13, 2024



# Modifications in tRNA sequences



## *Saccharomyces cerevisiae* cytosolic tRNA<sup>Phe</sup> (76 nt, 14 nt modified)

Along the lines of Lorenz, C., Lünse, C.E., & Mörl, M. (2017). tRNA Modifications: Impact on Structure and Thermal Adaptation. *Biomolecules*, FIG1 [4]

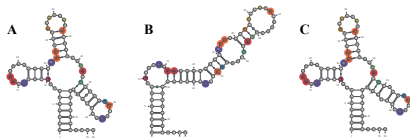
# RNA modifications & secondary structure prediction

Traditional methods lack the means for handling modifications in RNA structure prediction.

- Implementations are limited to the standard RNA nucleotide alphabet AUGC, no complete NN energy parameter set for other nucleotides is available.
- *ViennaRNA Package*<sup>[3]</sup> now provides a mechanism to adjust free energy parameters\* and to improve the prediction accuracy<sup>[6]</sup>.

\*energy corrections for modified bases with parameters for the six modifications: **inosine** (I), **pseudouridine** ( $\Psi$ ), **N6-methyl-adenosine** (m6A), **dihydrouridine** (D), **7-deaza-adenosine** (7DA), and **purine** (P) are already available.

Example: *Bos taurus* tRNA-Phe (tRNAdb ID: tdbR00000096)<sup>[1]</sup>



(A) Reference structure, (B) MFE prediction for unmodified sequence, and (C) MFE prediction using *RNAfold* with modified base support.

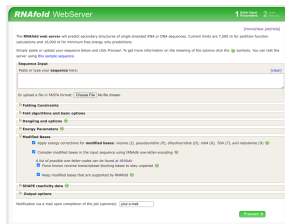
[6] Varenkyk, Y., Spicher, T., Hofacker, I.L., Lorenz, R., "Modified RNAs and predictions with the ViennaRNA Package", Bioinformatics, Nov. 2023

# Modification support in ViennaRNA Package 2.6

## ViennaRNA Package

- *RNAfold* - predicting global structures, equilibrium probabilities.
- *RNAsubopt* - prediction of suboptimal structures.
- *RNALfold* - local structure predictions.
- *RNAplfold* - accessibility predictions.
- *RNAcofold* - predictions for two interacting RNAs, where one or both may contain modifications.

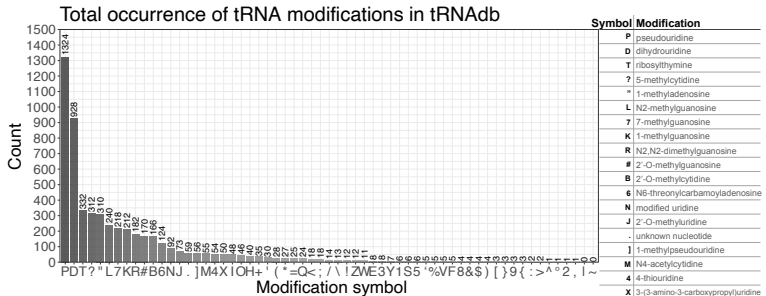
**ViennaRNA WebService** - RNAfold WebService  
(<http://rna.tbi.univie.ac.at>) integrates an input sequence pre-processing step to handle *tRNAdb* one-letter-codes.



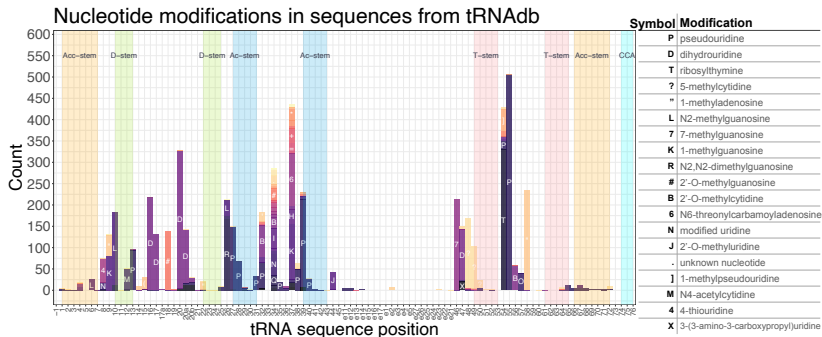
The screenshot shows the 'RNAfold WebService' interface. The 'Request form' section includes a text input for 'Paste or type your sequence here:' with a 'Submit' button. Below this, there are several expandable sections: 'Adding Constraints', 'Add algorithms and base options', 'Energy Parameters', and 'Modified Bases'. The 'Modified Bases' section is expanded, showing options for 'Apply energy corrections for modified bases' (checked), 'Consider modified bases in the input sequence using tRNAdb one-letter encoding' (checked), 'List of possible one-letter codes can be found at tRNAdb', 'Base bases reverse transcription blocking bases in the sequence' (checked), and 'Skip modified bases that are supported by tRNAdb' (checked). At the bottom, there is an 'Output options' section and a 'Notification via e-mail upon completion of the job (optional):' field with a 'Send e-mail' button.

# Databases of RNA modifications

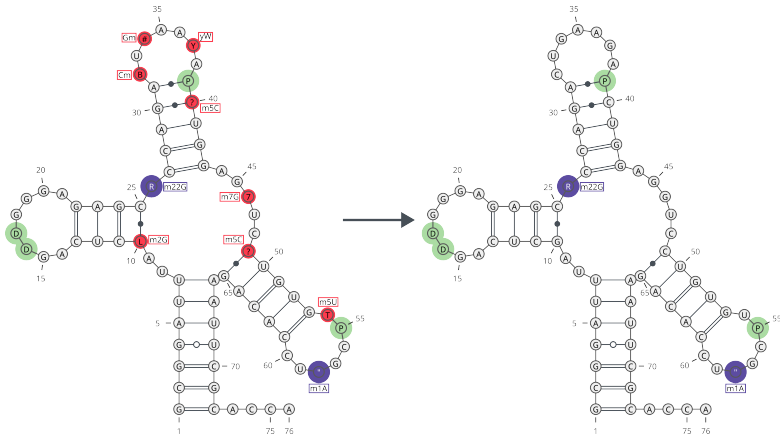
- MODOMICS [Dunin-Horkawicz et al., 2006, Boccaletto et al., 2018]<sup>[5]</sup>
- Transfer RNA database (tRNAdb) [Jühling et al., 2009]<sup>[1]</sup>
- RNA Modification Database (*RNAMDB*) [Cantara et al., 2011]
- Small Subunit rRNA Modification Database (*SSUmods*) [McCloskey and Rozenki, 2005]



# Location of modified bases in tRNA sequences (tRNAdb)



# Treatment of modified bases with ViennaRNA Package



# Evaluation of RNA secondary structure prediction

- Matthews Correlation Coefficient

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

- Positive Predictive Value

$$PPV = \frac{TP}{TP+FP}$$

- Sensitivity

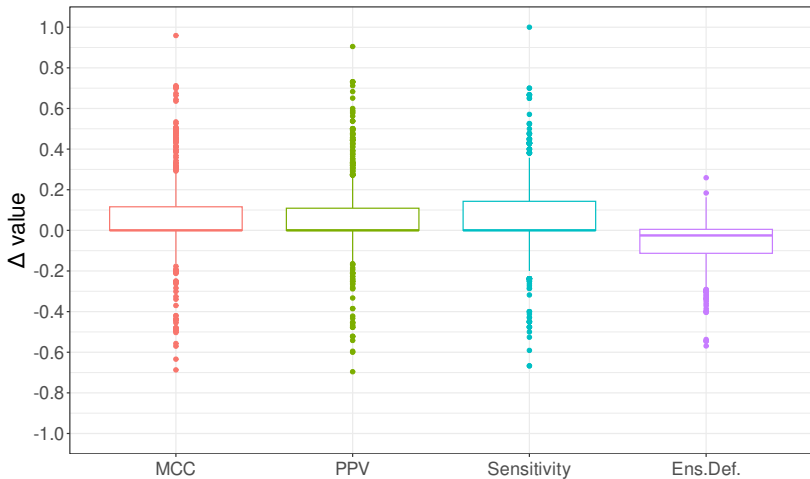
$$Sensitivity = \frac{TP}{TP+FN}$$

- Ensemble defect - the average number of incorrectly paired nucleotides at equilibrium evaluated over the ensemble of secondary structures.



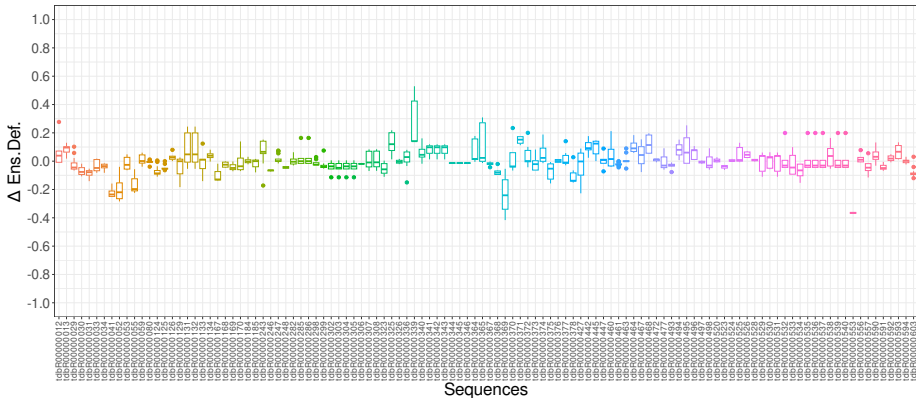
# Secondary structure prediction upon modification constraints

tRNAdb dataset - change in prediction quality upon modification constraints



# Secondary structure prediction upon modification constraints - methylation at pos.9 (m1A, m1G)

trRNAbd dataset - sequences with methylation (m1G or m1A) at pos.9



- Positional constraints analysis - determine which locations of modifications lead to improvement in structure prediction.
- Further updates of *ViennaRNA Package* - add modification support for services like *energy evaluation, alignment programs, consensus structure prediction*.

# Thank you!

Ivo Hofacker

Ronny Lorenz

Thomas Spicher

TBI group

RNA Deco

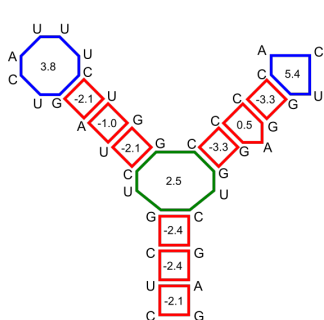
Winterseminar Bled

*tbi*



universität  
wien

# RNA secondary structure prediction - Nearest Neighbor energy model



$$E(\text{RNA}) = \begin{array}{c} \begin{array}{c} \text{U} \quad \text{A} \\ \text{C} \quad \text{G} \end{array} \begin{array}{|c|} \hline -2.1 \\ \hline \end{array} + \begin{array}{c} \text{C} \quad \text{G} \\ \text{U} \quad \text{A} \end{array} \begin{array}{|c|} \hline -2.4 \\ \hline \end{array} + \begin{array}{c} \text{G} \quad \text{C} \\ \text{C} \quad \text{G} \end{array} \begin{array}{|c|} \hline -2.4 \\ \hline \end{array} + \begin{array}{c} \text{G} \quad \text{C} \\ \text{U} \quad \text{G} \end{array} \begin{array}{|c|} \hline 2.5 \\ \hline \end{array} + \\ \\ \begin{array}{c} \text{G} \\ \text{U} \quad \text{G} \\ \text{C} \end{array} \begin{array}{|c|} \hline -2.1 \\ \hline \end{array} + \begin{array}{c} \text{U} \\ \text{A} \quad \text{G} \\ \text{U} \end{array} \begin{array}{|c|} \hline -1.0 \\ \hline \end{array} + \begin{array}{c} \text{C} \\ \text{G} \quad \text{A} \\ \text{A} \end{array} \begin{array}{|c|} \hline -2.1 \\ \hline \end{array} + \begin{array}{c} \text{U} \quad \text{U} \\ \text{A} \quad \text{C} \\ \text{U} \quad \text{G} \end{array} \begin{array}{|c|} \hline 3.8 \\ \hline \end{array} + \\ \\ \begin{array}{c} \text{C} \\ \text{C} \quad \text{G} \\ \text{G} \end{array} \begin{array}{|c|} \hline -3.3 \\ \hline \end{array} + \begin{array}{c} \text{C} \\ \text{C} \quad \text{G} \\ \text{G} \quad \text{A} \end{array} \begin{array}{|c|} \hline 0.5 \\ \hline \end{array} + \begin{array}{c} \text{C} \\ \text{C} \quad \text{G} \\ \text{G} \end{array} \begin{array}{|c|} \hline -3.3 \\ \hline \end{array} + \begin{array}{c} \text{A} \\ \text{C} \quad \text{U} \\ \text{G} \end{array} \begin{array}{|c|} \hline 5.4 \\ \hline \end{array} \\ \\ = -6.50 \text{ kcal/mol} \end{array}$$

- Secondary structures can be uniquely decomposed into loops.
- Contributions of a base pair only depends on neighboring pairs.
- Each loop is assigned a free energy contribution<sup>[2]</sup>.

Turner et al., "NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure.", 2009, NAR <sup>[2]</sup>

- [1] Frank Jühling et al. **“tRNAdb 2009: compilation of tRNA sequences and tRNA genes”**. In: *Nucleic Acids Research* 37.suppl\_1 (2009), pp. D159–D162. DOI: [10.1093/nar/gkn772](https://doi.org/10.1093/nar/gkn772).
- [2] Douglas H Turner and David H Mathews. **“NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure”**. en. In: *Nucleic Acids Research* 38.suppl\_1 (Oct. 2009), pp. D280–282. ISSN: 0305-1048, 1362-4962. DOI: [10.1093/nar/gkp892](https://doi.org/10.1093/nar/gkp892).
- [3] Ronny Lorenz et al. **“ViennaRNA Package 2.0”**. In: *Algorithms for Molecular Biology* 6.1 (2011). DOI: [10.1186/1748-7188-6-26](https://doi.org/10.1186/1748-7188-6-26).
- [4] Christian Lorenz, Christina Lünse, and Mario Mörl. **“tRNA modifications: Impact on structure and thermal adaptation”**. In: *Biomolecules* 7.4 (2017), p. 35. DOI: [10.3390/biom7020035](https://doi.org/10.3390/biom7020035).

- [5] Pietro Boccaletto et al. **“MODOMICS: a database of RNA modification pathways. 2021 update”**. In: *Nucleic Acids Research* 50.D1 (2022), pp. D231–D235.
- [6] Yuliia Varenyk et al. **“Modified RNAs and predictions with the ViennaRNA Package”**. In: *Bioinformatics* 39.11 (Nov. 2023), btad696. DOI: [10.1093/bioinformatics/btad696](https://doi.org/10.1093/bioinformatics/btad696).