# The ViennaRNA Package

**Ronny Lorenz[1] with** *(in alphabetical order)*
**F. Amman, S. H. Bernhart, C. Flamm, A. R. Gruber, C. Höner zu Siederdissen, D. Luntzer, U. Mückstein, R. Neuböck, J. Qin, P. F. Stadler, H. Tafer, A. Tanzer, M. T. Wolfinger, *et al.*, and Ivo L. Hofacker[123]**

[1] Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria.
[2] Research Group Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, Währinger Straße 17 A-1090 Wien, Austria.
[3] Center for non-coding RNA in Technology and Health, University of Copenhagen, Grønnegårdsvej 3, 1870 Frederiksberg, Denmark.

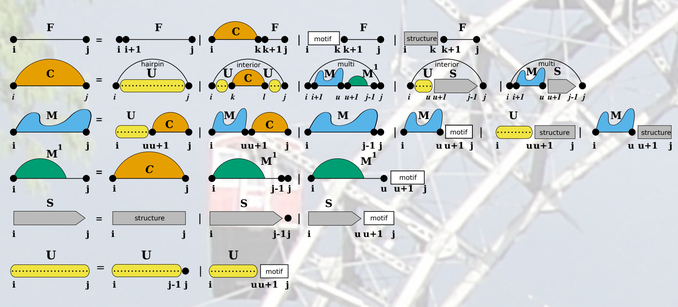
Contact: {rna,ronny,ivo}@tbi.univie.ac.at - http://www.tbi.univie.ac.at/RNA

## High Quality Results

By using the **latest** available Nearest Neighbor **energy parameters** (Turner 2004) our algorithms provide best quality predictions [2]. User-defined free energy parameters can be loaded at runtime to enable adaptations of the prediction to novel measurements. The `ViennaRNA Package` already ships a set of existing parameters for RNA and DNA, derived from both, UV-melting experiments and training on a large set of reference structures.

## RNA Folding Grammar

We use an extended and **unified** recursive structure **decomposition scheme**, the RNA folding grammar. All algorithms adhere to (parts of) it to predict global or local structures for single sequences and sequence alignments [8].



## The Story of Success

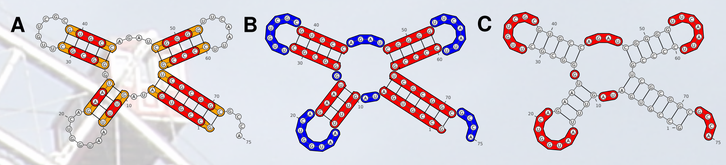
With a total of more than **6,000 citations**, the `ViennaRNA Package` is one of the most widely used tool packages for RNA secondary structure prediction in the world [1, 2, 3, 4]. It combines very good prediction accuracy paired with high performance to make everything in an RNA-bioinformaticians life easier. We constantly increase the **usability** of the command line programs and extend the underlying C-library to yield the best experience for both, **beginners** and **experts**.

## Fast folding and Comparison

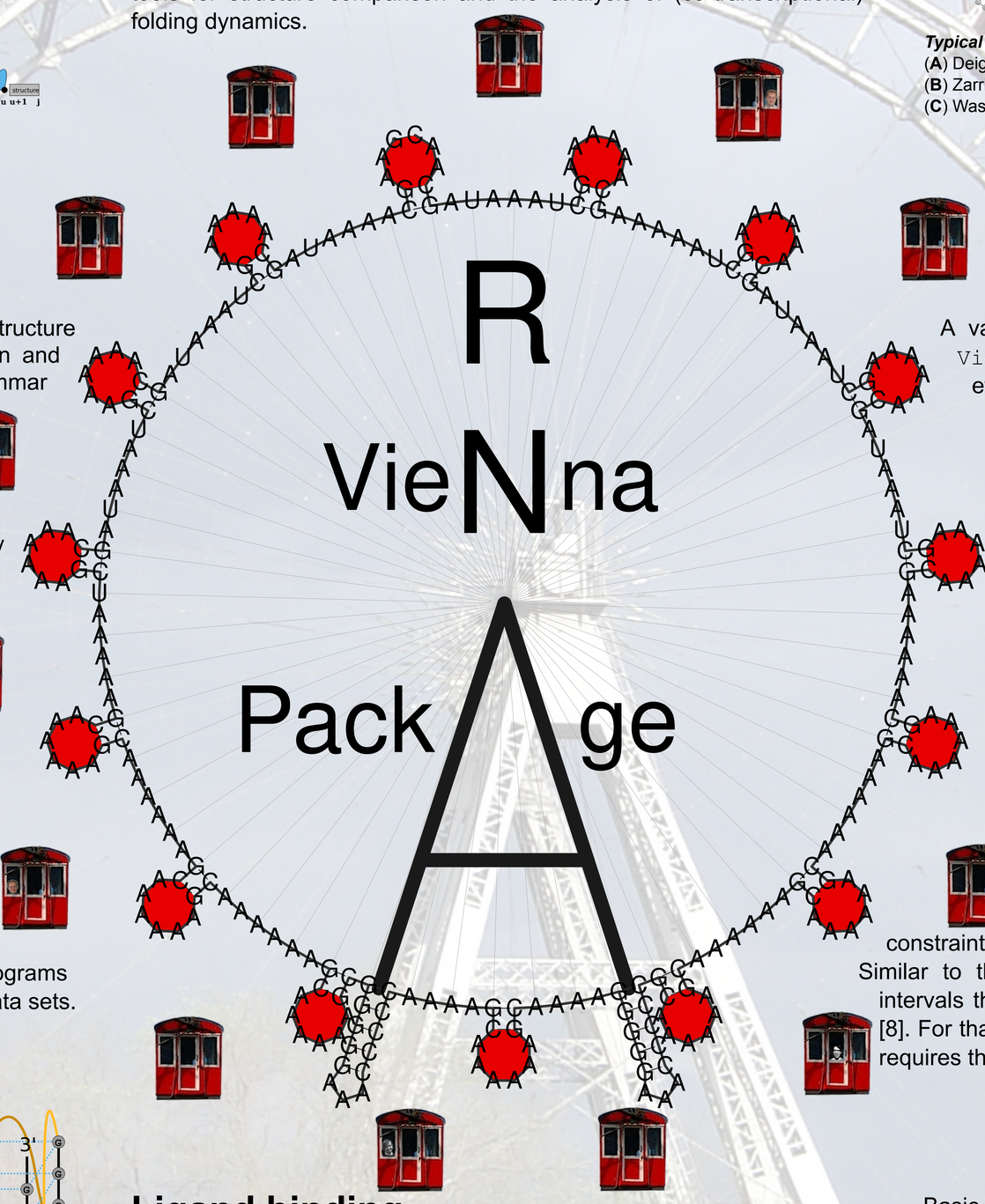
The `ViennaRNA Package` is a toolbox comprised of about 30 different instruments. It not only enables one to predict secondary structures of single sequences, intermolecular RNA-RNA interactions, and consensus structures for multiple sequence alignments, but also offers tools for structure comparison and the analysis of (co-transcriptional) folding dynamics.

## Experimental Structure Probing

Chemical and enzymatic probing of RNA structures reveals at nucleotide resolution if a nucleotide is more likely to be paired or unpaired. Especially the coupling of probing methods with high-throughput Next Generation Sequencing generates massive amounts of data suitable to guide structure prediction. For instance, most of our tools accept normalized **SHAPE reactivity data** as input which in turn is simply converted into pseudo-energy terms and added to particular conformations [6, 7].




***Typical configurations SHAPE reactivity data is applied to.***
(**A**) Deigan et al. (2009) method, stacked nucleotides
(**B**) Zarringhalam et al. (2012) method, paired and unpaired nucleotides
(**C**) Washietl et al. (2012) method, unpaired nucleotides




## Structure Constraints

We offer a generic yet systematic way to augment structure prediction. A fully **transparent access** to the derivation and energy evaluation of the implemented RNA folding grammar renders our approach most flexible [8]. Here, we distinguish two kinds of conceptually different constraints:
(i) **Hard constraints** that limit the candidate space by pruning particular derivation trees, and
(ii) **Soft constraints** acting on the evaluation level by adding "*bonus energies*".
Particular application scenarios such as guiding predictions by experimental SHAPE probing data or RNA-ligand binding are already available through high-level convenience functions.

## Complementary Software

A variety of complementary applications build upon the `ViennaRNA Package` to make use of the fast and efficient implementations of secondary structure prediction algorithms. Among them are tools for **structural sequence alignment**, such as `LocARNA`, several **RNA sequence design** approaches, e.g. `RNAblueprint`, and tools to simulate the **dynamics of the folding process** such as `Barriers/Treekin` and `DrTransformer`. Furthermore, the `ViennaRNA Package` is an essential element in many analysis pipelines for **miRNA/sRNA target detection** and numerous machine learning approaches.

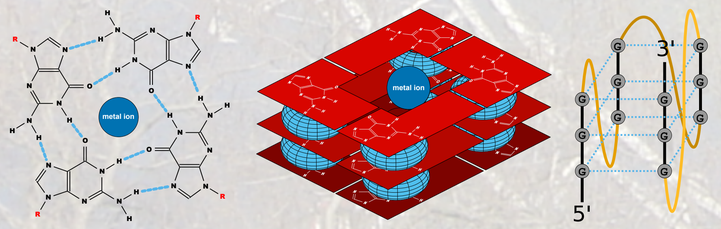
## High Speed Predictions

The `ViennaRNA Package` always was among the fastest in RNA secondary structure prediction. The latest version speeds up computations by automatically detecting **SIMD** features of the host CPU. For batch jobs (*multiple input sequences*) on **multi-core CPUs** many programs also allow for parallel processing of the individual input data sets.

## Unstructured Domains

The constraints framework alone can't treat arbitrary sequence intervals as a binding site, e.g. for **Single Strand Binding (SSB) proteins**. Such interval constraints are, however, available as additional production rules. Similar to the incorporation of G-Quadruplexes, base pair free intervals that interact with external factors are modeled explicitly [8]. For that purpose, the default implementation in `RNAfold` only requires the sequence motif and a binding free energy.

## RNA G-Quadruplexes

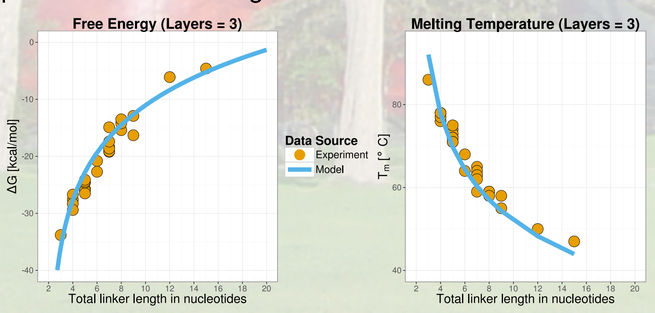



G-rich nucleic acid sequences are known to form highly stable stacks of G-quartes, also known as G-Quadruplexes. For sequence patterns that follow the canonical form


we explicitly include their potential to form a quadruplex through an extension of the RNA folding grammar [5]. Energy contributions are evaluated using a **simplified energy model**

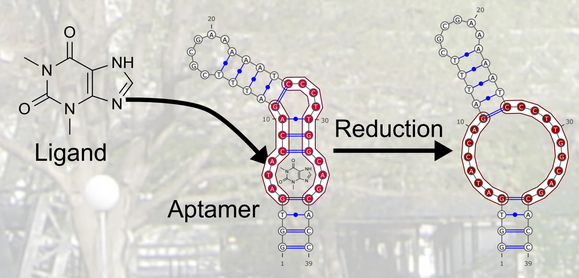
$$E(L,l,T) = a(T)(L-1) + b(T)\ln(l-2), \text{with}$$
$$a(T) = H_a + TS_a \quad \text{and} \quad b(T) = H_b + TS_b$$

that only depends on the number of stacked layers and the total length of linker sequences. Parameters have been fitted to experimental UV- melting data.




## Ligand binding

A convenience function that utilizes the soft constraints feature enables one to model **binding** of small ligands **to specific aptamer structures**. Thus, aptamers that resemble hairpin- or internal loops only require a few ingredients, (i) the binding free energy and a (ii) sequence and (iii) structure motif. Once these parameters have been suppied, the `ViennaRNA Package` will consider the bound (*Holo*) and unbound (*Apo*) conformations during MFE and partition function computations [8].




## ViennaRNA Web Services

Basic analysis and structure prediction with the most commonly used tools and most popular parameter options can be easily done **without even installing any software**. Simply submit your analysis at the ViennaRNA Web Services
**`http://rna.tbi.univie.ac.at`**
next to more than 1,500 requests by other users per day.

## Command Line Tools

The `ViennaRNA Package` includes a steadily increasing number of useful applications and scripts. While global structures can be predicted for sequences up to about 30kb, local structure prediction methods allow one to **screen entire genomes**. Most approaches are available for both, single sequences and multiple sequence alignments and the corresponding programs support common input file formats such as `FASTA`, `ClustalW`, `Stockholm 1.0`, and `MAF`.

## Programming Interface

The algorithms implemented in the `ViennaRNA Package` are available as a ***C-library*** that can be easily used by your own programs. It exclusively grants access to functionalities that might not be directly available through the command line programs. Finally, the `Perl` and `Python` wrappers allow for ***rapid prototyping*** and direct interfacing to the efficient implementations from ***analysis pipelines.***

## Acknowledgements

We want to **thank** our users for their **feed-back** and **support!** Work on the most recent parts of the `ViennaRNA Package` hast been supported in parts by the Austrian Science Fund FWF project "*RNA regulation of the transcriptome*" (F43) and the Austrian/French project "*RNA-Lands*" (FWF-I-1804-N28 and ANR-14-CE34-0011).

We also want to thank the participants of our annual group retreats in Bled and Doubice for their great ideas and input discussions.



[1] "Fast folding and comparison of RNA secondary structures", I.L. Hofacker, W. Fontana, P.F. Stadler, L.S. Bonhoeffer, M. Tacker, and P. Schuster (1994), Monatshefte f. Chemie, Volume 125, Issue 2, Pages 167–188
[2] "ViennaRNA Package 2.0", R. Lorenz, S.H. Bernhart, C. Höner zu Siederdissen, H. Tafer, C. Flamm, P.F. Stadler, and I.L. Hofacker (2011), Alg. Mol. Biol., Volume 6, Pages 26
[3] "Vienna RNA secondary structure server", I.L. Hofacker (2003), Nucleic Acids Research, Volume 31, Issue 13, Pages 3429–3431
[4] "The RNA Websuite", A.R. Gruber, R. Lorenz, S.H. Bernhart, R. Neuböck, I.L. Hofacker (2008), Nucleic Acids Research, Volume 36, Issue suppl_2, Pages W70–W74
[5] "2D meets 4G: G-Quadruplexes in RNA Secondary Structure Prediction", R. Lorenz, S.H. Bernhart, J. Qin, C. Höner zu Siederdissen, A. Tanzer, F. Amman, I.L. Hofacker, and P.F. Stadler (2013), IEEE/ACM Transactions on Computational Biology and Bioinformatics, PP:99
[6] "SHAPE directed RNA folding", R. Lorenz, D. Luntzer, I.L. Hofacker, P.F. Stadler, and M.T. Wolfinger (2016), Bioinformatics 32, Pages 145-14.
[7] "Predicting RNA Structures from Sequence and probing Data", R. Lorenz, M.T. Wolfinger, A. Tanzer, and I.L. Hofacker (2016), Methods.
[8] "RNA folding with hard and soft constraints", R. Lorenz, I.L. Hofacker, and P.F. Stadler (2016), Algorithms for Molecular Biology 11:1, Pages 1-13.