



HAL
open science

Mono-valent salt corrections for RNA secondary structures in the ViennaRNA package

Hua-Ting Yao, Ronny Lorenz, Ivo L Hofacker, Peter F Stadler

► **To cite this version:**

Hua-Ting Yao, Ronny Lorenz, Ivo L Hofacker, Peter F Stadler. Mono-valent salt corrections for RNA secondary structures in the ViennaRNA package. *Algorithms for Molecular Biology*, 2023, 18, pp.8. 10.1186/s13015-023-00236-0 . hal-04062134v2

HAL Id: hal-04062134

<https://hal.science/hal-04062134v2>

Submitted on 19 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH

Open Access



Mono-valent salt corrections for RNA secondary structures in the ViennaRNA package

Hua-Ting Yao^{1*}, Ronny Lorenz¹, Ivo L. Hofacker^{1,2} and Peter F. Stadler^{1,3,4,5,6,7*}

Abstract

Background RNA features a highly negatively charged phosphate backbone that attracts a cloud of counter-ions that reduce the electrostatic repulsion in a concentration dependent manner. Ion concentrations thus have a large influence on folding and stability of RNA structures. Despite their well-documented effects, salt effects are not handled consistently by currently available secondary structure prediction algorithms. Combining Debye-Hückel potentials for line charges and Manning's counter-ion condensation theory, Einert et al. (Biophys J 100: 2745-2753, 2011) modeled the energetic contributions of monovalent cations on loops and helices.

Results The model of Einert et al. is adapted to match the structure of the dynamic programming recursion of RNA secondary structure prediction algorithms. An empirical term describing the salt dependence of the duplex initiation energy is added to improve co-folding predictions for two or more RNA strands. The slightly modified model is implemented in the ViennaRNA package in such way that only the energy parameters but not the algorithmic structure is affected. A comparison with data from the literature show that predicted free energies and melting temperatures are in reasonable agreement with experiments.

Conclusion The new feature in the ViennaRNA package makes it possible to study effects of salt concentrations on RNA folding in a systematic manner. Strictly speaking, the model pertains only to mono-valent cations, and thus covers the most important parameter, i.e., the NaCl concentration. It remains a question for future research to what extent unspecific effects of bi- and tri-valent cations can be approximated in a similar manner.

Availability Corrections for the concentration of monovalent cations are available in the ViennaRNA package starting from version 2.6.0.

Keywords RNA secondary structure, Salt concentration, Debye-Hückel potential

*Correspondence:

Hua-Ting Yao
htyao@tbi.univie.ac.at
Peter F. Stadler
studla@bioinf.uni-leipzig.de

¹ Department of Theoretical Chemistry, University of Vienna, Währinger Straße 17, 1090 Vienna, Austria

² Research Group Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, Währingerstraße 29, 1090 Vienna, Austria

³ Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, Universität Leipzig, Härtelstraße 16–18, 04107 Leipzig, Germany

⁴ Competence Center for Scalable Data Services and Solutions Dresden/Leipzig, Interdisciplinary Center for Bioinformatics, German Centre for Integrative Biodiversity Research (iDiv), and Leipzig Research Center for Civilization Diseases, Universität Leipzig, Augustusplatz 12, 04107 Leipzig, Germany

⁵ Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04109 Leipzig, Germany

⁶ Facultad de Ciencias, Universidad Nacional de Colombia, Sede Bogotá, Ciudad Universitaria, 111321 Bogotá, D.C., Colombia

⁷ Santa Fe Institute, 1399 Hyde Park Rd., NM87501 Santa Fe, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Nucleic acids are highly negatively charged molecules since their phosphate backbone carries one negative charge per nucleotide. Structure formation brings these charges into close proximity and thus incurs substantial electrostatic penalties on the secondary and tertiary structure level. Nucleic acid–ion interactions also provide large interaction energies and therefore contribute decisively to the folding of RNA and DNA and to their interactions with ligands and macromolecule partners [1]. Counter-ions reduce the electrostatic repulsion of the backbone. Cation concentrations determine the extent of this “charge screening” and thus strongly influence RNA folding. Indeed, many functional RNAs will not fold under low salt conditions [2], and experimental investigations of the thermodynamics of RNA folding are mostly confined to high salt conditions. Energy parameters for RNA secondary predictions likewise pertain to 1M sodium concentrations, more precisely 1.021 M while taking all Na^+ ions in the buffer into account [3]. Although the importance of counter-ions for the RNA folding is well known, ion concentration, in contrast to temperature, is not a tunable parameter in most of the currently available RNA secondary structure prediction tools. Although NUPACK [4] and UNAFold [5] offer a corresponding option, they do not handle salt effects in a consistent manner. The empirical salt corrections (derived from DNA for Na^+ [6] and Mg^+ [7]) offered by these tools pertain to stacking energies contributions only and neglect the loop energies due to insufficient empirical data. Loops, however, are subject to salt-dependent effects in a similar energy range as we shall see below. While temperature dependence is conceptually straightforward and can be easily modeled by splitting free energy contributions into enthalpic and entropic contributions [8], the energetics of ion–nucleic acid interactions are much more difficult to understand.

Cations affect RNA structure in two different ways. The electrostatic stabilization of the structure due to charge screening is at least conceptually independent of the chemical nature and charge of the individual cations. In addition, divalent cations and in particular Mg^+ can also strongly bind to specific, chelating sites [9, 10]. Quantitative salt effects on RNA folding have been studied extensively over the last decades, see [11] for a study that summarized much of the pertinent earlier literature. In the absence of a well-founded theoretical model, most authors resorted to describing the salt-dependence of RNA folding by means of simple heuristic function fitting the effects of changes in the sodium concentration on the free energy of folding or a melting temperature. Such empirical fits, however, are limited to handling salt effects close to standard conditions, and an approach

that explains the functional form of salt effects is clearly preferable.

If three-dimensional structures are known, the non-linear Poisson-Boltzmann equation can be solved to obtain electrostatic potentials of RNA molecules in solution [12–14]. This approach, however, appears to be too detailed to derive a practically manageable parametrization of salt effects at the level of secondary structure prediction algorithms. In order to handle counter-ions in RNA secondary structure prediction algorithms, the effects must be attributed to individual bases, base pairs, or loops (including the stacking of two consecutive base pairs). This is necessary because secondary structure prediction algorithms operate on these combinatorial substructures [15]. This, in particular, precludes models that explicitly require a detailed geometric description of three-dimensional structure of an RNA. Ensembles of 3D structure models can be used, however, to estimate cation effects on loops and helices empirically as an alternative to wet-lab experiments [16, 17]. A coarse grained model representing each nucleotide by two virtual bonds ($\text{C}_4\text{-P}$ and P-C_4) [18] and tightly-bound ion theory [19] accounting for strongly correlated multi-valent ions was employed to sample loop conformations in the presence of both Na^+ and Mg^+ . Empirical expressions for electrostatic helix [16] and loop [17] energy contributions were extracted from these simulations.

To derive a suitably simple model, Einert and Netz [20] proposed to represent the RNA backbone as a charged polymer that interacts by means of a Debye-Hückel potential [21] and treats single-stranded regions as freely jointed chains [22]. The non-linear screening effect of monovalent cations is incorporated using Manning’s approach to counter-ion condensation [23]. The formulation for loop contribution was originally developed to understand the salt-dependent modulation of nucleosomal structures [24]. The two strands of a helix are modeled as parallel rods that again interact via a Debye-Hückel potential governed by the screening length. Here, the theory yields a per-position contribution that is independent of sequence features and the position within the helix [20]. While the theory makes several approximations it has been shown by its authors to reproduce experimental data quite well. It also has the advantage that it has no free parameters other than well known generic geometric characteristics of nucleic acid 3D structures such as distances between nucleotides or the planes of stacked pairs. In contrast to empirical salt corrections derived from measurements or simulations, the theory has the advantage of behaving reasonably also outside the regime of available measurements. Of course, the secondary structure based approaches have significant limitations. In particular, they are by design

not suitable to model the site-specific binding of chelated Mg^{+} , which in some cases is known to be crucial for tertiary structure formation and RNA function [25].

Here we implement in the ViennaRNA package [26] the Debye-Hückel/Manning model derived in [20], which captures the energetic effects of electrostatic interactions of mono-valent cations with structured RNA. In the following theory section we briefly review the features of the energy model and show that it can be brought into a form where only the energy parameters but not the dynamic programming recursions are modified. We then evaluate the model on a collection of empirical data from the literature to show that use of this type of salt correction has a significant beneficial effect.

Theory

The model of [20] considers loops as freely jointed charged chains and helices as parallel rods interacting via Debye-Hückel potentials and uses Manning’s framework to model counter-ions condensation. This results in distinct types of correction terms from loops and helices, which we describe separately in the following. In either case we focus on how the salt correction terms are incorporated into the dynamic programming schemes for RNA secondary structure prediction. As we shall see, the salt corrections can be phrased in such a way that they exclusively effect the energy parameters. The folding algorithms therefore remain unchanged. Owing to the architecture of the ViennaRNA package, it is therefore possible to handle salt effects exclusively as a pre-processing energy parameter set. The presentation below focuses on RNA, since secondary structure predictions (beyond perfect duplexes) are of practical interest mostly for RNA. The theory, however, applies equally to DNA secondary structures.

Salt corrections for loops

The electrostatic free energy contribution for a loop comprising L backbone bonds can be written, at the level of the Debye-Hückel approximation, in the form [24]:

$$\mathcal{G}_u(L) = RT \frac{\ell_B}{\kappa} \tau_{ss}^2 \Phi(y) \tag{1}$$

with $y = \kappa l_{ss} L$ where l_{ss} is the length of a single stranded RNA backbone bond (phosphate-to-phosphate distance), ℓ_B is the Bjerrum length and κ^{-1} is the Debye screening length, which depends on the ionic strength, and $\tau_{ss} = \min(1/l_{ss}, 1/\ell_B z_c)$ accounts for the nonlinear electrostatic effects. For monovalent ions, the ionic strength equals the salt concentration ρ and thus we have $\kappa = \kappa(\rho)$. Following [24], Φ is given by

$$\begin{aligned} \Phi(y) = & y \ln y + y(\gamma - \ln(\pi/2)) - \frac{y^2}{2} {}_1F_2\left(\frac{1}{2} \middle| \frac{y^2}{4\pi^2}\right) \\ & - \frac{y^3}{2\pi^2} {}_2F_3\left(\frac{1, 1}{3/2, 3/2, 2} \middle| \frac{y^2}{4\pi^2}\right) + 1 - \exp(-y) + y\Gamma(0, y) \end{aligned} \tag{2}$$

Here, ${}_1F_2$ and ${}_2F_3$ are generalized hypergeometric functions and $\Gamma(0, y)$ is the incomplete gamma function. Using ${}_1F_2(\dots, 0) = {}_2F_3(\dots, 0) = 1$, and $y \ln y \rightarrow 0$ for $y \rightarrow 0$, we observe that $\Phi(y) = 0$. Eq. (A8) in [24] gives an expansion for small y of form $\Phi(y)/y = (1 - \ln(\pi/2)) + O(y^2)$, where the $\ln y$ term and the logarithmic divergence of $\Gamma(0, y)$ cancel. Thus, $\Phi(y)$ increases linearly with y . Since both y and κ are proportional to $\sqrt{\rho}$, the ρ -dependence cancels and $\mathcal{G}_u(L)$ approaches L times a constant for $\rho \rightarrow 0$.

Since $\mathcal{G}_u(L)$ describes the salt-dependent electrostatic effects on loops, this term is already included in the empirical energy parameters of the RNA standard model for the standard conditions of $T = 37^\circ C$ and 1 M sodium concentration. Writing $\widehat{\mathcal{G}}_u(L)$ for the values on Eq. (1) at standard conditions, allows us to write the salt correction of a given loop as

$$g_u(L, \rho) := \mathcal{G}_u(L) - \widehat{\mathcal{G}}_u(L) \tag{3}$$

The ViennaRNA package quantifies the length of a loop by the number m of unpaired nucleotides rather than the number of bonds. For hairpin loops we have $L = m + 1$, while $L = m + 2$ for interior loops. In multi-loops we have $L = m + q + 1$, where q is the degree (number of branches) of the multi-loop. Setting $q = 0$ for hairpin loops and $q = 1$ for interior loops, the appropriate salt correction for a loop is therefore $g_u(m + q + 1, \rho)$.

As seen in Fig. 1, the salt correction, as expected, increases while decreasing the salt concentration contributing to the destabilization of the structure in a low salt environment. In the usual temperature range, the plot shows that loop correction is close to a constance with a slight increase at low concentration.

Linear approximation for multi-branch Loop

The free energy of a loop asymptotically depends on the logarithm of its length [27], see also [28] for a recent review. For multi-loops, however, this behavior is usually approximated by a linear function

$$E_{ml} = \alpha + \beta q + \gamma m \tag{4}$$

for the sake of computational efficiency. The parameters α , β , and γ are the energy cost for having, respectively, the closing pair, branch, and unpaired base in a multi-loop. The reason for this linearisation is that recursions

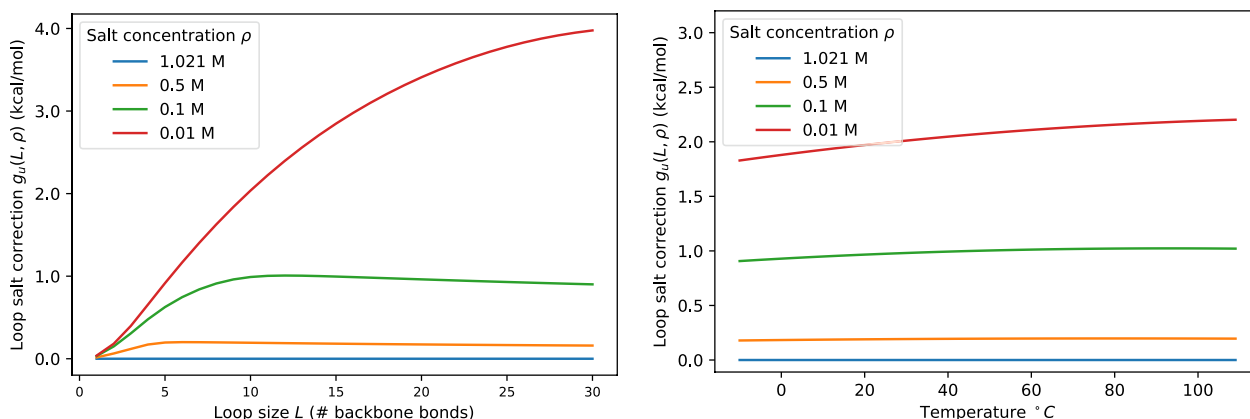


Fig. 1 Loop salt correction $g_u(L, \rho)$ as a function of loop size $L = m + q + 1$ (left) and as a function of temperature (right) for a fixed loop at size $L = 10$ for different salt concentrations ρ

implementing a non-linear dependence of the multi-loop contribution on m or $m + q$ make it necessary to store terms pertaining to substructures delimited by the closing pair of multi-loop of size m . This leads to a cubic-memory and quartic-time algorithm, which is infeasible for larger RNA molecules [29]. The linear approximation is further motivated by the empirical observation that models with logarithmic dependence are outperformed by the linear approximation in terms of accuracy of structure prediction [29].

In order to handle multi-loops without abandoning the memory-efficient multi-loop decomposition of the standard RNA energy model [3], it is necessary to retain the linear form of the multi-loop terms also in the presence of salt corrections. This implies that the salt-correction itself must also be linear in both q and m . Using $L = m + q + 1$ this implies that the salt correction must be of the form

$$g_u(m, q, \rho) \approx a_0(\rho) + a_1(\rho)L \tag{5}$$

where a_0 and a_1 are the parameters of the linear approximation of $g_u(L, \rho)$. The correction therefore amounts to adding $a_0 + a_1$ to the closing pair term α of the multi-loop and a_1 to both, each unpaired base γ and each component β of the multi-loop. The salt-dependent multi-loop model thus reads:

$$E_{ml} = (\alpha + a_0(\rho) + a_1(\rho)) + (\beta + a_1(\rho))q + (\gamma + a_1(\rho))m \tag{6}$$

In order to fit the coefficients a_0 and a_1 appearing in the multi-loop parameters in practice, we first investigated their size distribution in samples of minimum free energy (MFE) structures of 5 000 random sequences for different RNA sizes, as well as in the

natural RNA structures recorded in the RNA STRAND database [30], see Additional file 1: Fig. S1. Although multi-loops with size $L = 3$ exist, very short multi-loops are rare. We argue that inaccuracies in these rare cases are likely acceptable. Very short loops presumably are also subject to specifically constrained three-dimensional structures and thus follow the model only to a crude approximation in the first place. In the current implementation, we use the loop size range $L \in [6, 24]$ to obtain linear fits for a_0 and a_1 from $g_u(L, \rho)$, see Fig. 2 for salt correction and their linear approximations. In general, the fit over-corrects for very small loops and, at low salt concentrations, also for very large loops. The maximal fitting errors are on the order of 1 kcal/mol, which is still within the ball-park of the rather large inaccuracies expected for multi-loop energies.

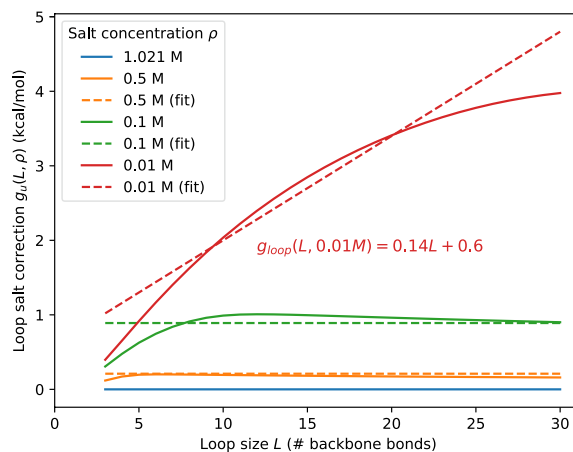


Fig. 2 Loop salt correction (solid) and linear approximation (dashed) in the function of loop size L at different salt concentration

Salt corrections for stacked base pairs

Describing the backbones of stacked pairs as rods with distance of $d = 20 \text{ \AA}$ interacting via a Debye-Hückel potential with screening length $1/\kappa$ yields the following electrostatic energy for the interaction of a nucleotide with the other strand [20]:

$$\mathcal{G}_p = 2RT\tau_{ds}^2 l_{ds} \ell_B K_0(\kappa d) \tag{7}$$

Here the charge density $\tau_{ds} = \min(1/l_{ds}, 1/(\ell_B z_c))$ is again estimated according to Manning’s counterion condensation theory [23]. The length parameter l_{ds} is the helical rise per base pair and $z_c = 1$ is the charge of the cation. K_0 denotes the zeroth-order modified Bessel function of the 2nd kind, see e.g. [31] and Additional file 1: Fig. S2. Since $K_0(z)$ diverges like $-\ln z$ for $z \rightarrow 0$, and $\kappa \sim \sqrt{\rho}$, \mathcal{G}_p diverges logarithmically for vanishing salt concentrations.

As in the case of loops, these electrostatic effects are already included in the empirical energy parameters for standard conditions. The relevant salt correction thus is given by the difference between the values \mathcal{G}_p at the current conditions and standard conditions.

$$g_p := \mathcal{G}_p - \widehat{\mathcal{G}}_p \tag{8}$$

Figure 3 shows the dependence of the position-wise salt correction for stacking energies as function of salt concentration and temperature. Similar to the loop correction, the stack correction is close to a constant in the usual temperature range for a given salt concentration.

Salt corrections for duplex initialization

The formation of a double strand from two RNA molecules in solution is associated with an additional initialization energy E_{init} in the Turner energy model. One expects that duplex formation becomes more difficult

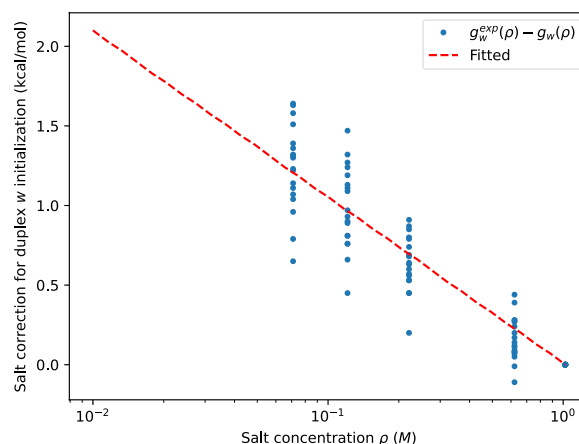


Fig. 4 Salt correction for duplex initialization fitted (red) from the difference between experimental and predicted duplex salt correction (blue)

due to electrostatic repulsion at low salt concentrations. The initialization energy thus should also depend on ρ . In addition, the distance between two single strands changes during formation, which was neglected in the theory as discussed in [20]. Indeed, as we will see later in the experimental data comparison section, the duplex free energies are systematically overestimated compared to the experimental data.

Due to the lack of theoretical support, we propose a salt correction for duplex initialization $g_{init}(\rho)$ derived from the prediction and the experimental data taken from [11]. Let $g_w(\rho)$ and $g_w^{exp}(\rho)$ be the predicted and experimental salt correction at concentration ρ from the standard condition for a give duplex w . Fitting the difference $g_w^{exp}(\rho) - g_w(\rho)$ of 18 duplexes at four non-standard sodium concentrations yields, as plotted in Fig. 4, the salt corrections for duplex initialization

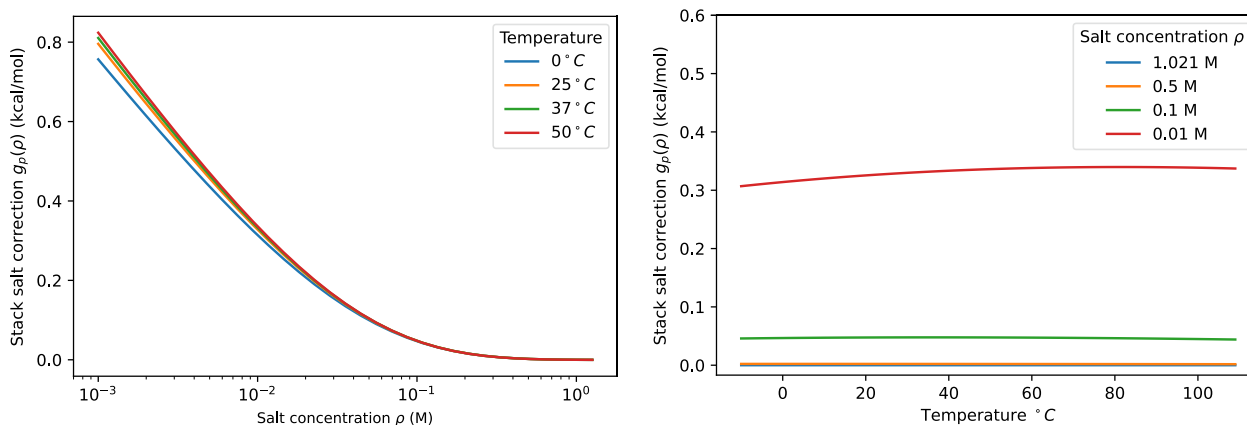


Fig. 3 Salt correction for a stacked pair as a function of salt concentration (left) and temperature (right)

$$g_{\text{init}}(\rho) = a \ln\left(\frac{\rho}{\rho_0}\right) \tag{9}$$

with $a = -0.45324$ kcal/mol for RNA and $a = -0.58389$ kcal/mol for DNA.

The nature of the initialization term a is not quite clear from the literature. On the one hand, it is argued as an entropic contribution in [32]. On the other hand, it is included with a large enthalpic contribution in recent versions of the standard energy model [3]. Given that the temperature dependence of the salt corrections for both stacking and loop is small, we feel justified in assuming a to be temperature-independent. At present, there are no data to test this assumption.

Implementation in ViennaRNA

The extension of ViennaRNA provides access to four user-defined parameters: the concentration of the monovalent cation ρ , the bounds L_1 and L_2 delimiting the interval of loop length that is used to fit the two multiloop parameters a_0 and a_1 for given salt concentration, and the user-provided salt correction g for duplex initialization. Default values are $L_1 = 6$, $L_2 = 24$, salt concentration $\rho = 1.021$ M, and $g = 99999$, indicating no value is provided.

On the command line, a new option `-salt` provides access to the salt corrections in all interactive programs of the ViennaRNA package. Internally, the relevant parameters are appended in the model object `vrna_md_s` as `salt` for ρ , `saltMLLower` and `saltMLUpper` for L_1 and L_2 , and `saltDPXInit` for g . For the salt concentration, ViennaRNA assumes the standard conditions of the Turner energy model, i.e., $\rho = 1.021$ M. Thus no salt corrections apply by default. A detailed description of the API in the ViennaRNA library is available at <https://www.tbi.univie.ac.at/RNA/ViennaRNA/refman/index.html>.

If a different concentration ρ is requested, first the value of $g_p(\rho)$ for the stack and the array of values $g_u(L, \rho)$ for loop of different size $L \in [1, 31]$ is computed. Note that these energy contributions depend on the temperature T and thus are recomputed if the user sets a different

temperature. In addition, the use of duplex initialization salt correction $g_{\text{init}}(\rho)$ for duplex is turned off if users set `saltDPXInit` to zero.

The array g_u is then used to determine $a_0(\rho)$ and $a_1(\rho)$ by linear regression. Subsequently, the free energy parameters are set as sums of the default values $E(0)$ for the given temperature and the salt corrections:

$$\begin{aligned} E^{bp} &= E^{bp}(0) + g_p(\rho) \\ E_m^{HL} &= E_m^{HL}(0) + g_u(m + 1, \rho) \\ E_m^{LL} &= E_m^{LL}(0) + g_u(m + 2, \rho) \\ \alpha &= \alpha(0) + a_0(\rho) + a_1(\rho) \\ \beta &= \beta(0) + a_1(\rho) \\ \gamma &= \gamma(0) + a_1(\rho) \\ E_{\text{init}} &= E_{\text{init}} + g_{\text{init}}(\rho) \end{aligned} \tag{10}$$

Here E^{bp} , E_m^{HL} , E_m^{LL} , and E_{init} refer to *all* parameters for stacked pairs, hairpin loops of length m , interior loops of length m , and duplex initialization, respectively. Dangling end and coaxial stacking contributions, on the other hand, remain unchanged.

The Bessel functions K_0 is computed as in `scipy`, which in turn used the `cephes` mathematical function library described in [33]. The function Φ in loop salt correction can be approximated as given in [20]:

$$\begin{aligned} \Phi(y) &= y \ln y + y(\gamma - \ln(\pi/2)) \\ &+ y \frac{(2\pi)^6}{y^6 + (2\pi)^6} \left(\frac{y^4}{36\pi^4} - \frac{y^3}{24\pi^2} + \frac{y^2}{2\pi^2} - \frac{y}{2} \right) \\ &+ \left(1 - \frac{(2\pi)^6}{y^6 + (2\pi)^6} \right) (y(\log(2\pi) - 1.96351) - y \log y) \\ &+ (1 - e^{-y}) + y\Gamma(0, y) \end{aligned} \tag{11}$$

Parameters

The key physical parameter containing the salt correction terms are the Debye screening length κ^{-1} . It is convenient to express κ^{-1} in terms of Bjerrum length ℓ_B and the ionic strength I :

$$\begin{aligned} \ell_B &= \frac{e^2}{4\pi k_B T \epsilon_0 \epsilon_r(T)} \approx (167092.53 \text{ \AA}) \frac{1}{T \epsilon_r(T)} \\ \kappa^{-1} &= \sqrt{\frac{k_B T \epsilon_0 \epsilon_r(T)}{2N_A e^2 I}} = \frac{1}{\sqrt{8\pi N_A}} \frac{1}{\sqrt{\ell_B I}} \approx \left(8.1285 \text{ \AA}^{\frac{3}{2}} \text{ mol}^{\frac{1}{2}} \text{ L}^{-\frac{1}{2}} \right) \frac{1}{\sqrt{\ell_B I}} \end{aligned} \tag{12}$$

The values of physical constants appearing in these expressions are taken from CODATA [34] and listed in Additional file 1: Table S1. The salt concentration enters only through the ionic strength $I := \frac{1}{2} \sum_i \rho_i z_i^2$, where ρ_i is the concentration and z_i is the charge of ion-species i . For monovalent ions, as in the case of NaCl, the ionic strength reduces to the salt concentration, i.e., $I = \rho$. The salt concentration ρ is expressed as molarity M, i.e., in units of mol/L, whereas the Debye screening length κ^{-1} and the Bjerrum length ℓ_B are conveniently expressed in Å. Note that the conversion factor for the length units L , in liters (0.001 m^3) versus Å, is absorbed into the numerical constant.

The temperature dependence of ϵ_r can be fitted from empirical data. In the current implementation we use the function proposed in [35]:

$$\epsilon_r(T) = 5321T^{-1} + 233.76 - 0.9297T + 1.417 \times 10^{-3}T^2 - 0.8292 \times 10^{-6}T^3 \quad (13)$$

with temperature in Kelvin. The temperature-dependence of ϵ_r ensures that the Bjerrum length ℓ_B is longer than the backbone bond length l_{ss} in the entire temperature range, see Additional file 1: Fig. S3. The nonlinear electrostatic effects on unpaired bases thus become $\tau_{ss} = 1/\ell_B$ for the monovalent cations considered here.

The model of Einert et al. [20] is equally applicable to RNA and DNA. The only difference are the two geometric parameters and the empirically determined slope for the salt correction of the duplex initialization energy, which are summarized in Table 1. The values of two geometric parameters, helical rise per base and backbone bond length (phosphate-to-phosphate distance), are obtained from [17, 36–38]. Fitting data reported in [39] yields the slope for the salt correction for DNA duplex initialization, see Additional file 1: Fig. S4 for more detail. Our implementations allow for adapting the values of the two geometry parameters l_{ss} and l_{ds} .

Comparison with experimental data

Available data sets

Even though the dependence of RNA structures on salt concentrations is of considerable practical interest, systematic data sets suitable for benchmarking salt correction models are by no means abundant. Most of the

direct evidence for the salt dependence derives from studies of short duplexes and hairpins. Here, we analyzed datasets of melting experiments from four publications that reported melting temperature T_m and/or free energy as well as corresponding enthalpy and entropy. The melting temperature is defined as the temperature at which half of the RNA molecules form duplexes or hairpins, respectively.

- 18 RNA self-complementary perfect duplexes of length 6 or 8 at the 5 different sodium concentrations, 0.071, 0.121, 0.221, 0.621, and 1.021 M, with different species concentration c were reported in [11, 40]. Free energies at 37°C were obtained from $1/T_m$ vs. $\ln c$ plots. The data were also used by the same lab [40] to obtain optimised thermodynamic parameters, entropy and enthalpy, at a given salt concentration.
- A different set of 8 self-complementary duplexes, including two imperfect duplexes, of length 10, 12, and 14 was reported in [41]. The data set covers two different species concentration c , 100 μM and 2 μM and two sodium concentrations, 1.0002 M and 0.0102 M.
- Two hairpins at sodium concentrations of 0.021, 0.036, 0.061, 0.111, 0.211, and 1.011 M were reported in [42]. The hairpins consisted of a helix of length 5 enclosing a hairpin loop of size $m = 8$ or 10.
- 14 hairpins at two salt concentrations, 1.02 and 0.03 M Na^+ , measured in [43]. One of the hairpins consists of a helix of length 8 and a hairpin loop of size $m = 4$, while in the remaining 13 hairpins the helix is interrupted by a 1 nt bulge.

The values of melting temperatures, enthalpies, and entropies for datasets 2, 3, and 4 were obtained by means of fitting melting curves in the respective publications.

Comparison of duplex free energies

Free energies were computed using RNAcofold [45, 46] with a self-complementary correction of $RT \ln 2$ added for all sequences that coincided with their reverse complement. Since the predicted free energy at the standard condition differs from the experimental values, we are interested in comparing the salt correction $g_w(\rho) = E_w(\rho) - E_w(\rho_0)$ at concentration ρ from the standard concentration $\rho_0 = 1.021 \text{ M}$, where $E_w(\rho)$ is the free energy of duplex w at concentration ρ . Let l_w and g_c be the length and the fraction of GC of duplex w . Then the salt correction for RNAcofold is then given by

$$g_w^{\text{RNAcofold}}(\rho) = (l_w - 1)g_p(\rho) + g_{\text{init}}(\rho). \quad (14)$$

The salt correction proposed by Chen & Znosko [11] is

Table 1 RNA and DNA specific parameters

Parameter	Symbol [units]	RNA	DNA
Helical rise length	l_{ds} [Å]	2.8	3.4
Backbone bond length	l_{ss} [Å]	6	6.76
Slope of duplex init	a [kcal/mol]	-0.45324	-0.58389

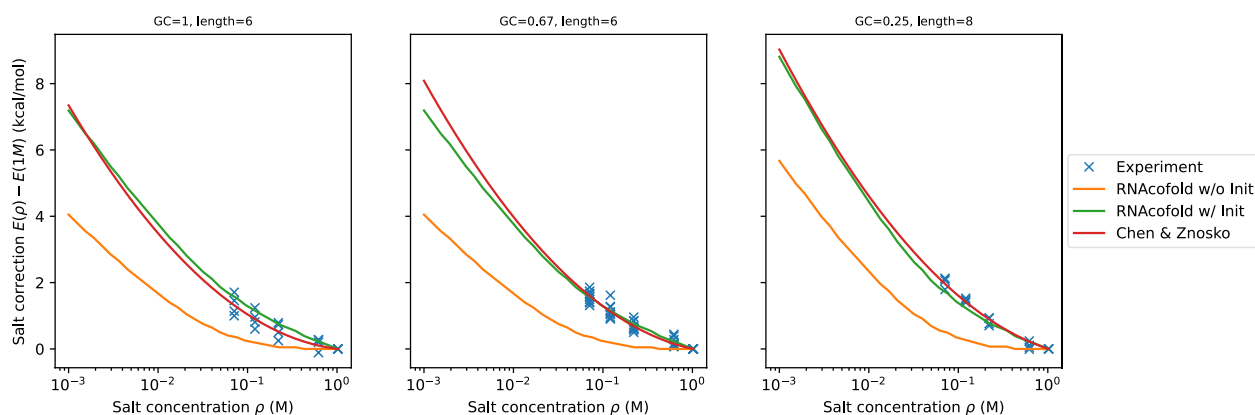


Fig. 5 Duplex salt correction at different salt concentration from experiment (blue), Chen & Znosko (red), and ViennaRNA prediction with (green) and without (orange) salt correction for duplex initialization. The experimental free energy is derived from van 't Hoff plots of $1/T_m$ versus $\ln c$. Most of the experimental values at 1 M are taken from [44]

$$g_w^C Z(\rho) = (0.324g_{c_w} - 0.486) \ln(\rho/\rho_0) + 0.133(\ln(\rho/\rho_0))^2. \quad (15)$$

The computational results are summarized in Fig. 5. Not surprisingly, the Chen & Znosko salt correction provides a slightly better fit to the data because the empirical formula for $g_w^{C\&Z}$ was obtained by fitting to the same data set. In contrast, only the duplex initialization $g_{\text{init}}(\rho)$ is fitted in our model. The largest deviations are observed for GC-only sequences.

Comparison of duplex melting temperatures

Let A be a self-complementary RNA sequence and AA the corresponding dimer. The dimerization reaction is $2A \rightleftharpoons AA$. The corresponding concentrations are denoted by $[A]$ and $[AA]$, respectively. In equilibrium, we have

$$\frac{[AA]}{[A][A]} = \frac{Z_{AA}}{Z_A Z_A} = e^{(2G_A - G_{AA})/RT} \quad (16)$$

where Z_A and Z_{AA} are the partition functions of the monomer and the dimer, respectively. Here $G_A = -RT \ln Z_A$ and $G_{AA} = -RT \ln Z_{AA}$ are the ensemble free energies of A and AA , respectively. Note that in a pure two-state system, we can replace G_A and G_{AA} by the corresponding minimum free energies E_A and E_{AA} , respectively. We define the melting temperature T_m as the temperature at which half of A forms the dimer AA , i.e., where $[AA] = c/4$ and $[A] = c/2$. Eq. (16) then yields

$$T_m = \frac{2G_A(T_m) - G_{AA}(T_m)}{-R \ln c}, \quad (17)$$

where we write $G_A(T_m)$ and $G_{AA}(T_m)$ to emphasize the temperature dependence of the ensemble free energies.

The ensemble free energy G_{AA} for the pure dimer state is accessible directly via the function `vrna_pf_dimer()` within the ViennaRNA library. The correction for self-complementary is already taken into account during the computation of the partition function [46]. Since the ensemble free energy is also a function of temperature, we use a binary search to find the melting temperature T_m .

For the data set comprising 18 short duplexes, experimental melting temperatures are available for several distinct species concentrations c . Van 't Hoff's equation

$$\frac{1}{T_m} = \frac{R}{\Delta H} \ln c + \frac{\Delta S}{\Delta H} \quad (18)$$

implies a linear relationship between changes in $1/T_m$ and changes in RNA concentration. We compare the predicted and experimental "van 't Hoff" plots, i.e., diagrams of $1/T_m$ versus $\ln c$, see Additional file 1: Fig. S5. Overall, we observe an excellent agreement on the slope between RNACofold prediction and experiment. Predicted and experimental intercepts are slightly more different for a few of the duplexes. To further investigate, we performed linear regression on each van 't Hoff plot and obtained the entropy and the enthalpy of each duplex at different salt concentration using Eq. 18. Figure 6 shows the difference of these thermodynamic values per base pair stack computed from the predictions and from the experiments. For both enthalpy and entropy, the predicted values are in general larger than the experimental one, but are in the same order of magnitude. The differences in enthalpy and entropy largely compensate in the free energy $\Delta G = \Delta H - T\Delta S$.

For the second dataset consisting of longer duplexes, we are interested in melting temperature correction

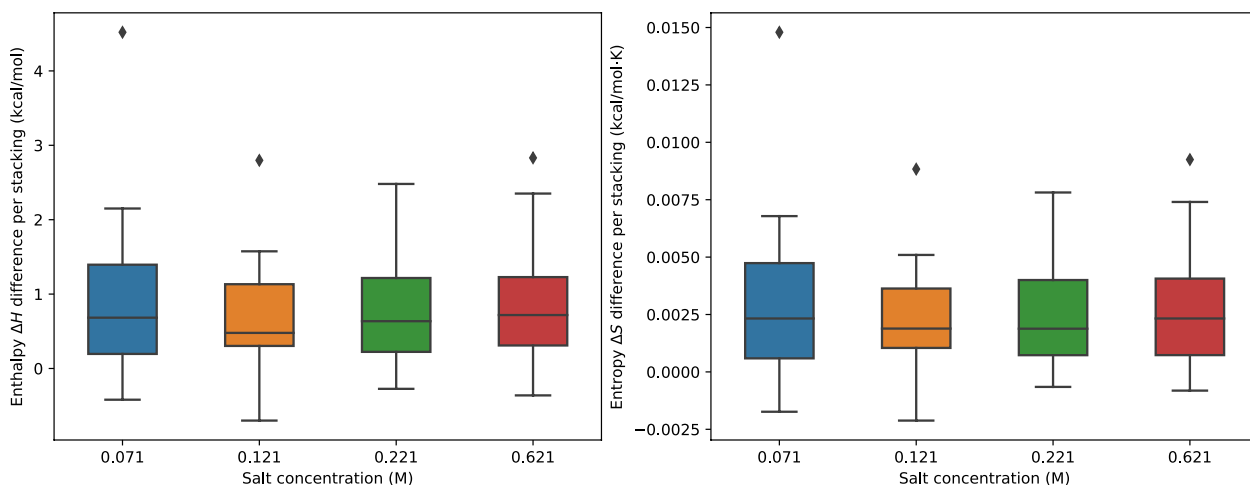


Fig. 6 Boxplot of differences between experimental and predicted salt effects for 18 short duplexes measured in [11]. Enthalpies and entropies were estimated from linear regression according to van 't Hoff's Eq. (18) from the data plotted in Additional file 1: Fig. S5

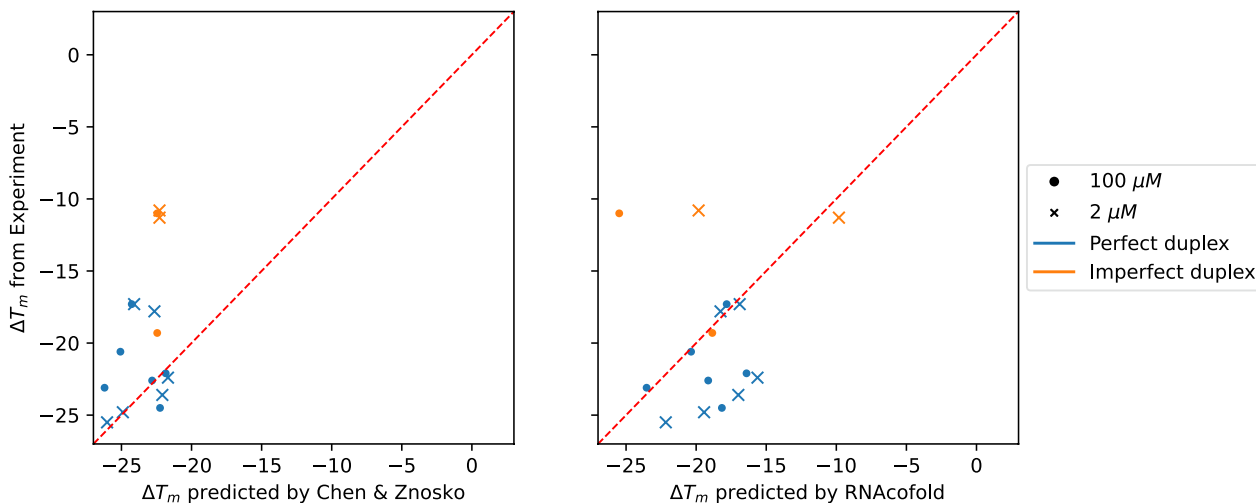


Fig. 7 Comparison of experimental and predicted melting temperature corrections ΔT_m using the empirical fit by Chen & Znosko [11] (left) and RNAfold with the salt corrections terms described in the present contribution (right). The data of perfect duplex is plotted in blue while the one of imperfect duplex is in orange. The red dashed line draws $x = y$ meaning the predicted value matches the experimental one

$\Delta T_m(0.01M)$ from the standard condition 1 M since the data is only available for two species concentrations. Chen & Znosko [11] proposed the following fit

$$\Delta T_m^{C\&Z}(\rho) = (-1.842gc_w + 2.675) \ln \rho - 0.7348(\ln \rho)^2 \tag{19}$$

where gc_w is the GC fraction of duplex w . Figure 7 shows the experimental melting temperature correction compared with $\Delta T_m^{C\&Z}$ and the one computed by RNAfold. Overall, the salt corrections described above show similar agreement (Root mean squared deviation $r = 5.638$) with the experiment as the empirical fit ($r = 5.862$) from [11].

Comparison of hairpin free energies

For the hairpins HP8 and HP10 described in [42], entropy, enthalpy, and free energy were derived from melting experiments. For comparison, we computed the free energy using RNAfold at 37°C for different salt concentrations ρ . Figure 8 summarizes the experimental and predicted free energy corrections from 0.211 M. They are in very good agreement in particular for lower salt concentrations.

For the 14 hairpins measured in [43] we used RNAfold to compute the free energy difference between salt concentration 0.03 M and 1.02 M at 37°C. For the hairpin with helix length 8, the free energy difference is

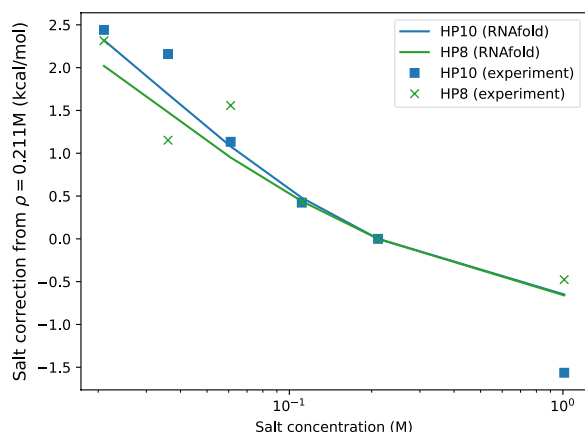


Fig. 8 Comparison of free energy correction from experiment for hairpin HP8 (green) and HP10 (blue) with RNAfold prediction

2.9 kcal/mol from experiment and 1.94 kcal/mol from computation. For the remaining 13, structurally identical, hairpins, the average experimental free energy difference is 3.68 ± 0.30 kcal/mol compared to a predicted value at 2.15 kcal/mol. In general, the model underestimates the salt correction by 1.5 kcal/mol.

Discussion

Salt concentrations significantly influence folding and thermodynamics of nucleic acids. In this contribution we report on the implementation of an approximate physical model proposed by Einert and Netz [20] that represent the RNA backbone as a charged polymer interacting by means of a Debye-Hückel potential. It is worth noting that model has no parameter that captures individual properties of the monovalent cation. It therefore applies equally to Na^+ , for which experimental data were available for comparison, and other monovalent cations. The model was adapted here to preserve the linear multi-loop model required for computational efficiency and extended by an empirical initiation term for duplex

formation. While not perfect, the model reproduces experimental thermodynamic data on the NaCl dependence of folding energies and melting temperatures with reasonable accuracy. We note that the salt-dependent energies for stacks diverge logarithmically for vanishing salt concentrations ρ . The model thus cannot be used if cations are virtually absent.

The approach taken here has the practical advantage that it does not require any changes in the folding algorithms. The modification of the energy parameters is sufficient. The ViennaRNA packages handles this step during preprocessing. As a consequence, the computational performance of the folding routines remain unchanged. Moreover, the salt corrections are consistently applied to all variants of the folding algorithms, e.g. minimum energy and partition function computations, consensus computations from alignments, and co-folding of two or more components.

Changes in salt concentration also affect the predicted secondary structures, see Fig. 9 for an example. A quantitative analysis, Fig. 10, confirms that the number of base pairs increases with salt concentration. Interestingly, the number of hairpin loops as measure for the overall “branched-ness” of the secondary structure, also increases. To our knowledge, there are no detailed experimental data that document structural changes as a function of NaCl concentration so that a directed validation of structures predicted for low salt concentrations remains a topic for future research.

Additional file 1: Fig. S6 shows the salt correction comparison with the coarse grained model of Tan & Chen for duplex [16] and hairpin loops [17]. Both models give similar predictions for duplexes of length 6 to 10. However, the Tan & Chen correction is not linear in length making it hard to be integrated into the current folding grammar. On the other hand, salt corrections for hairpin loops computed using two models have different behaviors, where the Tan & Chen correction diverges with increasing loop size.

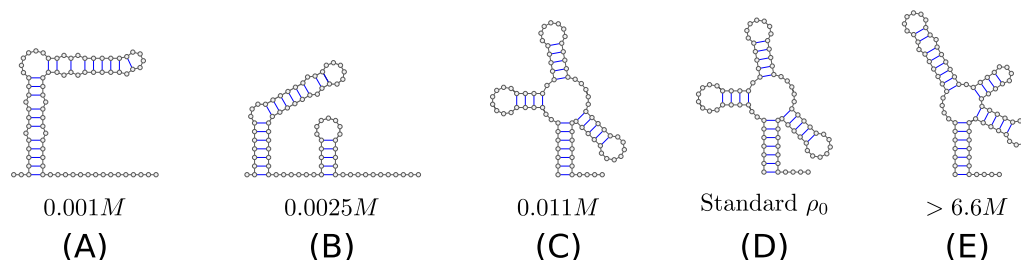


Fig. 9 Examples of structural transitions. MFE structures of a tRNA sequence at different salt concentrations are predicted with RNAfold. Within the concentration range from 0.011 to 6.6 M, the MFE structure is same as the one at the standard condition (C, D). The denaturation is observed at low concentration (A, B), while at high concentration (> 6.6 M, corresponding to a saturated saline solution), E become the MFE structure. The tRNA sequence used is GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA

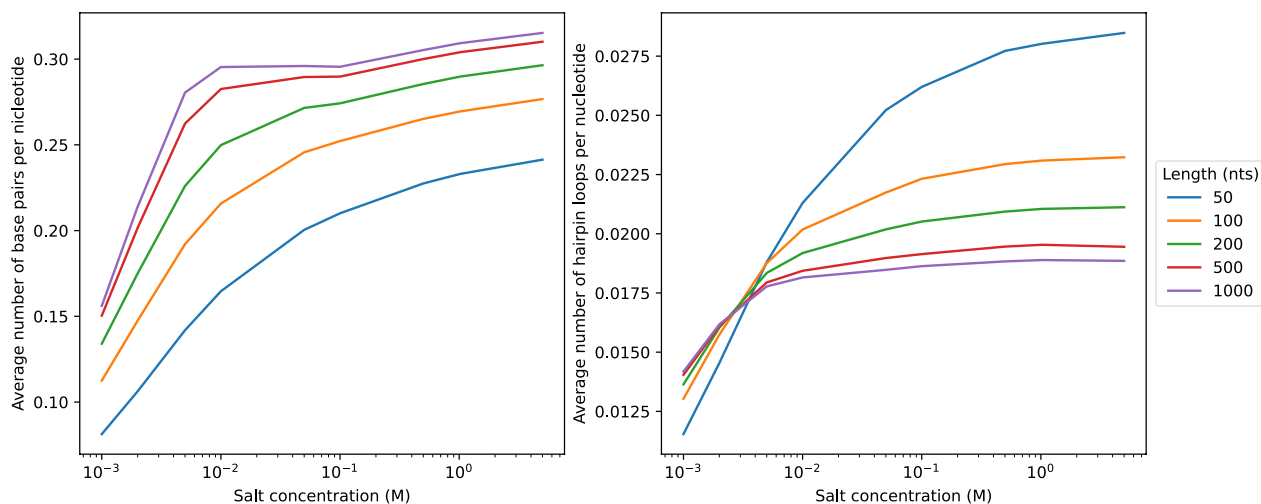


Fig. 10 Average number of base pairs (left) and hairpin loops (right) per nucleotides in MFE structures at different salt concentrations. For each length n , 5 000 RNA sequences are uniformly and randomly selected from $\{A, C, G, U\}^n$. Each sequence is then folded using `RNAfold` at different salt concentrations

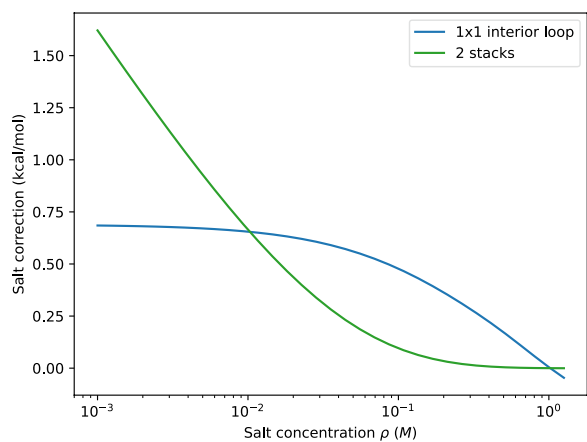


Fig. 11 Free energy correction of an 1×1 interior loop (blue) and a helix of two stackings (green) at different salt concentration

In the current model, a basepair mismatch in a helix is treated as an 1×1 interior loop and thus is associated with the salt correction for loops at non-standard salt concentrations. However, such a mismatch is likely to result in a slightly distorted helix that could still be seen as two parallel charged rods. One could therefore argue, that the salt correction for 2 stacked pairs rather than for a loop should be applied. Figure 11 shows the difference of these two cases as a function of salt concentration. To our knowledge, there are no experimental data that could be used to decide which approach is more appropriate.

The approach taken here does not account for all effects of ions on RNA folding. Most importantly, it covers only unspecific interactions and thus does not

describe specific interactions e.g. of Mg^{2+} with specific binding sites. Even for unspecific interactions, the validity of the model can be argued stringently only for mono-valent ions [20].

We note that the empirical salt-correction formula for stacking energies proposed in [7] incorporates the combined salt concentrations as a linear combination $[Na^+] + 3.3[Mg^{2+}]$. In essence this term expresses the cationic contribution to the ionic strength (up to the empirically determined coefficient 3.3 instead of the theoretical value 4). This may be taken as a hint that the model implemented here may also serve as reasonable approximation for more mixtures ions. At present, available data, e.g. [47], are not sufficient to test whether replacing ρ by the ionic strength is sufficient to reasonably account for mixtures of mono-valent and di-valent cations.

The model described here does not account for salt effects of RNA structures that are neither loops nor stacked base pairs. In particular it does not apply to G quadruplexes [48], which optionally can be included in secondary structure predictions [49]. Separate models for the ion dependencies of such features will need to be derived that account e.g. for the K^+ -dependent stabilization of RNA quadruplexes.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13015-023-00236-0>.

Additional file 1: Figure S1. Length distribution of multi-loops. Distribution of multi-loop size L , number of backbone bonds, among MFE structures of 5000 uniformly selected sequences at varied length. **Figure S2.**

Approximation Error for K_0 . In [20] an approximation for the difference of K_0 at a given concentration and 1 M was proposed. However, we noticed that this approximation yields a non-vanishing salt correction at 1 M. We therefore used the Cephes library to compute K_0 directly. The panel shows the salt correction of base pair stack at 37 °C in the function of salt concentration using the approximation (blue) and the precise computation implemented in ViennaRNA (orange). **Figure S3.** Nonlinear electrostatic effects tss. In [20], the permittivity (relative dielectric constant) ϵ_r of water $\epsilon_r \approx 80$ is assumed to be temperature independent. This assumption results in a discontinuity of tss at around 53.3 °C. Incorporating the empirical temperature dependence of ϵ_r in eq. (13) [35] results in $1/\epsilon B < 1/lss$. **Figure S4.** Comparison of experimental and predicted melting temperature corrections ΔT_m of DNA duplexes. In [39], the authors performed melting experiment of DNA duplexes at different salt concentrations (Table 2) and collected an independent set of DNA melting temperatures (Table 5). The former one is used to fit the salt correction for DNA duplex initialization and the later one is used for validation. Only the duplexes with length smaller than 11 are used in both datasets. **Figure S5.** Van t'Hoff plots for 18 duplexes. Plotting $1/T_m$ versus $\ln c$ shows a generally good agreement of between predictions and the experimental data from [11]. **Figure S6.** Salt correction comparison with Tan & Chen model. Comparison of salt correction predicted with ViennaRNA for hairpin loop and duplex of different size with the ones predicted with Tan & Chen model [16, 17]. **Table S1.** Numeric values of physical constants.

Author contributions

All authors contributed to the design of the study and the writing of the manuscript. HTY implemented the model and conducted the computational evaluation, RL integrated the new features into the ViennaRNA package.

Funding

Open access funding provided by Austrian Science Fund (FWF). This research was funded in part by the Austrian Science Fund (FWF), Grant No. I 4520 and F 80, as well as the German Research Foundation (DFG), Grant No. STA 850/48-1.

Availability of data and materials

The implementation is available in the ViennaRNA Package starting with version 2.6.0 available at <https://www.tbi.univie.ac.at/RNA>. The tutorial-like notebook and the data to produce all figures are available at <https://github.com/ViennaRNA/salty>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 4 April 2023 Accepted: 10 July 2023

Published online: 29 July 2023

References

- Lipfert J, Doniach S, Das RD, Herschlag D. Understanding nucleic acid-ion interactions. *Annu Rev Biochem.* 2014;83:813–41. <https://doi.org/10.1146/annurev-biochem-060409-092720>.
- Brion P, Westhof E. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct.* 1997;26:113–37. <https://doi.org/10.1146/annurev.biophys.26.1.113>.
- Turner DH, Mathews DH. NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucl Acids Res.* 2010;38:280–2. <https://doi.org/10.1093/nar/gkp892>.
- Fornace ME, Huang J, Newman CT, Porubsky NJ, Pierce MB, Pierce NA. NUPACK: Analysis and design of nucleic acid structures, devices, and systems. Technical report, chemRxiv. <https://doi.org/10.26434/chemrxiv-2022-xv981>
- Markham NR, Zuker LS, Zuker M. The UNAFold Web Server – Quickfold. <http://www.unafold.org/Dinamelt/applications/quickfold.php>. Accessed 07 May 2023.
- SantaLucia J Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA.* 1998;95:1460–5. <https://doi.org/10.1073/pnas.95.4.1460>.
- Koehler RT, Peyret N. Thermodynamic properties of DNA sequences: characteristic values for the human genome. *Bioinformatics.* 2005;21(16):3333–9. <https://doi.org/10.1093/bioinformatics/bti530>.
- Lu ZJ, Turner DH, Mathews DH. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Res.* 2006;34:4912–24. <https://doi.org/10.1093/nar/gkl472>.
- Draper DE. A guide to ions and RNA structure. *RNA.* 2004;10:335–43. <https://doi.org/10.1261/rna.5205404>.
- Draper DE. RNA folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophysical J.* 2008;95:5489–95. <https://doi.org/10.1529/biophysj.108.131813>.
- Chen Z, Znosko BM. Effect of sodium ions on RNA duplex stability. *Biochemistry.* 2013;52(42):7477–85. <https://doi.org/10.1021/bi4008275>.
- Chin K, Sharp KA, Honig B, Pyle AM. Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function. *Nat Struct Biol.* 1999. <https://doi.org/10.1038/14940>.
- Misra VK, Draper DE. The interpretation of Mg(2+) binding isotherms for nucleic acids using Poisson-Boltzmann theory. *J Mol Biol.* 1999;294:1135–47. <https://doi.org/10.1006/jmbi.1999.3334>.
- Misra VK, Draper DE. The linkage between magnesium binding and RNA folding. *J Mol Biol.* 2002;317:507–21. <https://doi.org/10.1006/jmbi.2002.5422>.
- Zuker M, Stiegler P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* 1981;9:133–48. <https://doi.org/10.1093/nar/9.1.133>.
- Tan Z-J, Chen S-J. RNA helix stability in mixed Na⁺/Mg²⁺ solution. *Biophys J.* 2007;92(10):3615–32. <https://doi.org/10.1529/biophysj.106.100388>.
- Tan Z-J, Chen S-J. Salt dependence of nucleic acid hairpin stability. *Biophys J.* 2008;95(2):738–52. <https://doi.org/10.1529/biophysj.108.131524>.
- Cao S, Chen S-J. Predicting RNA folding thermodynamics with a reduced chain representation model. *RNA.* 2005;11:1884–97. <https://doi.org/10.1261/rna.2109105>.
- Tan Z-J, Chen S-J. Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. *J Chem Phys.* 2005;122: 044903. <https://doi.org/10.1063/1.1842059>.
- Einert TR, Netz RR. Theory for RNA folding, stretching, and melting including loops and salt. *Biophys J.* 2011;100:2745–53. <https://doi.org/10.1016/j.bpj.2011.04.038>.
- Debye E, Peter und Hückel: Zur Theorie der Elektrolyte. *Physik Z.* 1923;24:185–206. https://doi.org/10.1007/978-3-642-94260-0_9.
- Mazars M. Statistical physics of the freely jointed chain. *Phys Rev E.* 1996;53:6297–319. <https://doi.org/10.1103/PhysRevE.53.6297>.
- Manning GS. Limiting laws and counterion condensation in polyelectrolyte solutions I. Colligative properties. *J Chem Phys.* 1969;51:924–33. <https://doi.org/10.1063/1.1672157>.
- Kunze K-K, Netz RR. Complexes of semiflexible polyelectrolytes and charged spheres as models for salt-modulated nucleosomal structures. *Phys Rev E.* 2002;66: 011918. <https://doi.org/10.1103/PhysRevE.66.011918>.
- Grilley D, Soto AM, Draper DE. Mg²⁺-RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc Natl Acad Sci USA.* 2006;103:14003–8. <https://doi.org/10.1073/pnas.0606409103>.
- Lorenz R, Bernhart SH, Höner zu Siederdisen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Alq Mol Biol* 2011;6:26. <https://doi.org/10.1186/1748-7188-6-26>
- Jacobson H, Stockmayer WH. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J Chem Phys.* 1950;18:1600–6. <https://doi.org/10.1063/1.1747547>.

28. Levene SD, Giovan SM, Hanke A, Shoura MJ. The thermodynamics of DNA loop formation, from J to Z. *Biochem Soc Trans*. 2013;41(2):513–8. <https://doi.org/10.1042/BST20120324>.
29. Ward M, Datta A, Wise M, Mathews DH. Advanced multi-loop algorithms for RNA secondary structure prediction reveal that the simplest model is best. *Nucleic Acids Res*. 2017;45:8541–50. <https://doi.org/10.1093/nar/gkx512>.
30. Andronescu M, Bereg V, Hoos HH, Condon A. RNA STRAND: the RNA secondary structure and statistical analysis database. *BMC Bioinformatics*. 2008;9:340. <https://doi.org/10.1186/1471-2105-9-340>.
31. Yang Z-H, Chu Y-M. On approximating the modified Bessel function of the second kind. *J Inequal Appl*. 2017;2017:41. <https://doi.org/10.1186/s13660-017-1317-z>.
32. Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH. Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA*. 1986;83:9373–7. <https://doi.org/10.1073/pnas.83.24.9373>.
33. Moshier SL. *Methods and programs for mathematical functions*. Hoboken, NJ: Prentice-Hall; 1989.
34. Tiesinga E, Mohr PJ, Newell DB, Taylor BN. CODATA recommended values of the fundamental physical constants: 2018. *Rev Mod Phys*. 2021;93:025010. <https://doi.org/10.1103/RevModPhys.93.025010>.
35. Catenaccio A, Daruich Y, Magallanes C. Temperature dependence of the permittivity of water. *Chem Phys Lett*. 2003;367:669–71. [https://doi.org/10.1016/S0009-2614\(02\)01735-9](https://doi.org/10.1016/S0009-2614(02)01735-9).
36. Taylor P, Rixon F, Desselberger U. Rise per base pair in helices of double-stranded rotavirus RNA determined by electron microscopy. *Virus Res*. 1985;2(2):175–82. [https://doi.org/10.1016/0168-1702\(85\)90247-3](https://doi.org/10.1016/0168-1702(85)90247-3).
37. Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE. Crystal structure analysis of a complete turn of B-DNA. *Nature*. 1980;287(5784):755–8. <https://doi.org/10.1038/287755a0>.
38. Chi Q, Wang G, Jiang J. The persistence length and length per base of single-stranded DNA obtained from fluorescence correlation spectroscopy measurements using mean field theory. *Physica A Stat Mech Appl*. 2013;392(5):1072–9. <https://doi.org/10.1016/j.physa.2012.09.022>.
39. Owczarzy R, You Y, Moreira BG, Manthey JA, Huang L, Behlke MA, Walder JA. Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures. *Biochemistry*. 2004;43(12):3537–54. <https://doi.org/10.1021/bi034621r>.
40. Ferreira I, Jolley EA, Znosko BM, Weber G. Replacing salt correction factors with optimized RNA nearest-neighbour enthalpy and entropy parameters. *Chem Phys*. 2019;521:69–76. <https://doi.org/10.1016/j.chemphys.2019.01.016>.
41. Nakano S-I, Kirihata T, Fujii S, Sakai H, Kuwahara M, Sawai H, Sugimoto N. Influence of cationic molecules on the hairpin to duplex equilibria of self-complementary DNA and RNA oligonucleotides. *Nucleic Acids Res*. 2007;35(2):486–94. <https://doi.org/10.1093/nar/gkl1073>.
42. Williams DJ, Hall KB. Thermodynamic comparison of the salt dependence of natural RNA hairpins and RNA hairpins with non-nucleotide spacers. *Biochemistry*. 1996;35(46):14665–70. <https://doi.org/10.1021/bi961654g>.
43. Groebe DR, Uhlenbeck OC. Thermal stability of RNA hairpins containing a four-membered loop and a bulge nucleotide. *Biochemistry*. 1989;28(2):742–7. <https://doi.org/10.1021/bi00428a049>.
44. Xia T, SantaLucia J Jr, Burkard ME, Kierzek R, Schroeder SJ, Jiao X, Cox C, Turner DH. Thermodynamic parameters for an expanded nearest-neighbour model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*. 1998;37(42):14719–35. <https://doi.org/10.1021/bi9809425>.
45. Bernhart SH, Tafer H, Mückstein U, Flamm C, Stadler PF, Hofacker IL. Partition function and base pairing probabilities of RNA heterodimers. *Algorithms Mol Biol*. 2006;1:3. <https://doi.org/10.1186/1748-7188-1-3>.
46. Lorenz R, Flamm C, Hofacker IL, Stadler PF. Efficient algorithms for co-folding of multiple RNAs. In: Ye X, Soares F, De Maria E, Gómez Vilda P, Cabitza F, Fred A, Gamboa H. (eds.) *Biomedical Engineering Systems and Technologies. BIOSTEC 2020. Communications in Computer and Information Science*, Cham: Springer, p. 1400.
47. Nguyen HT, Hori N, Thirumalai D. Theory and simulations for RNA folding in mixtures of monovalent and divalent cations. *Proc Natl Acad Sci*. 2019;116:21022–30. <https://doi.org/10.1073/pnas.1911632116>.
48. Bhattacharyya D, Arachchilage GM, Basu S. Metal cations in G-quadruplex folding and stability. *Front Chem Sec Chem Biol*. 2016;4:38. <https://doi.org/10.3389/fchem.2016.00038>.
49. Lorenz R, Bernhart SH, Qin J, Höner zu Siederdisen C, Tanzer A, Amman F, Hofacker IL, Stadler PF. 2D meets 4G: G-quadruplexes in RNA secondary structure prediction. *IEEE Trans Comp Biol Bioinf*. 2013;10:832–844. <https://doi.org/10.1109/TCBB.2013.7>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

