TBI WINTERSEMINAR

Salt concentration correction in ViennaRNA package

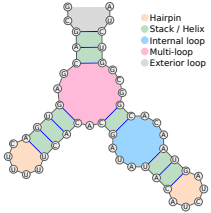
YAO, Hua-Ting

TBI, University of Vienna, Austria

tbi

Bled — Feb 13, 2023

Nearest Neighbor Energy Model

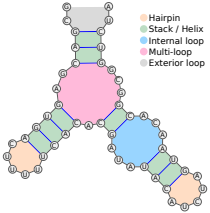


Given a sequence w and a secondary structure S, the strructural free energy is sum of energies for loops in S

$$\mathcal{E}(w,S) = \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \cdots$$

- $\Delta(\ ) : \mathcal{E}_{\text{stack}}(w)$ $\Delta(\ )$, $\Delta(\ ) : \mathcal{E}_{\text{loop}}(m,w)$ $\Delta(\ ) : m \times \mathcal{E}_{\text{MLbase}} + x \times \mathcal{E}_{\text{MLstem}} + \mathcal{E}_{\text{MLclosing}}$
  where $m =$ # unpaired bases and $x =$ # stems $- 1$.
- $\Delta(w, loop)$ measured at salt concentration $\rho_0 = 1.021$ M (Turner & Mathews, 2010)

Nearest Neighbor Energy Model



Given a sequence w and a secondary structure S, the strructural free energy is sum of energies for loops in S

$$\mathcal{E}(w,S) = \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \Delta\Big(\ \Big) + \cdots$$

- $\Delta(\ ) : \mathcal{E}_{\text{stack}}(w)$ $\Delta(\ )$, $\Delta(\ ) : \mathcal{E}_{\text{loop}}(m,w)$ $\Delta(\ ) : m \times \mathcal{E}_{\text{MLbase}} + x \times \mathcal{E}_{\text{MLstem}} + \mathcal{E}_{\text{MLclosing}}$
  where $m =$ # unpaired bases and $x =$ # stems $- 1$.
- $\Delta(w, loop)$ measured at salt concentration $\rho_0 = 1.021$ M (Turner & Mathews, 2010)

What if the concentration differs from 1 M ?

Existing models

- For RNA duplex
  - → Data driven correction from 18 duplexes (Chen & Znosko, 2013)
  - → Optimized thermodynamics parameters (Ferreira *et al.*, 2020)
  - → Mesoscopic models (Ferreira *et al.*, 2021)
- For loop
  - → Tightly bound ion model (Tan & Chen, 2008)
  - → MD simulation (Miner & García, 2018)

Theory for RNA Folding, Stretching, and Melting Including Loops and Salt

Thomas R. Einert[†⋆] and Roland R. Netz[†‡]
[†]Physik Department, Technische Universität München, Garching, Germany; and [‡]Fachbereich Physik, Freie Universität Berlin, Berlin, Germany

- Salt correction for loop and helix (stacking) based on Debye–Hückel work

$$\mathcal{E}^{salt}(w, S, \rho) := \mathcal{E}(w, S) + \mathcal{G}^{salt}(S, \rho)$$

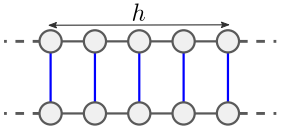
  $\rightarrow$ Salt correction $\mathcal{G}^{salt}$ is context-independent
  $\rightarrow$ Ions-independent
  $\rightarrow$ Depends only on the loop size

- Restricted to monovalent ions

Salt correction of a helix

Given a helix of length h, the salt correction at concentration ρ is

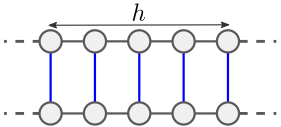
$$\mathcal{G}^{salt}_{\text{helix}}(h,\rho) = h \times \underbrace{(g_{\text{stack}}(\rho) - g_{\text{stack}}(\rho_0))}_{g^{salt}_{\text{stack}}(\rho)}$$

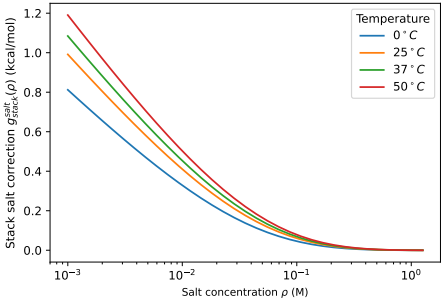


Salt correction of a helix

Given a helix of length h, the salt correction at concentration ρ is

$$\mathcal{G}^{salt}_{\text{helix}}(h,\rho) = h \times \underbrace{(g_{\text{stack}}(\rho) - g_{\text{stack}}(\rho_0))}_{g^{salt}_{\text{stack}}(\rho)}$$

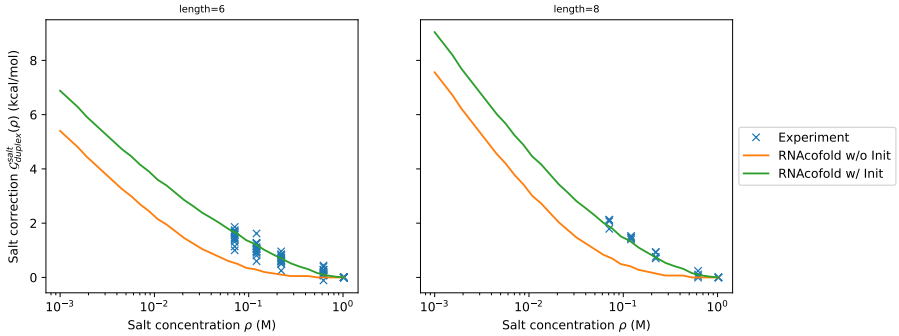




$\Delta($  $) : \mathcal{E}_{\text{stack}}(w) + g^{salt}_{\text{stack}}(\rho)$

Salt correction of duplex initiation

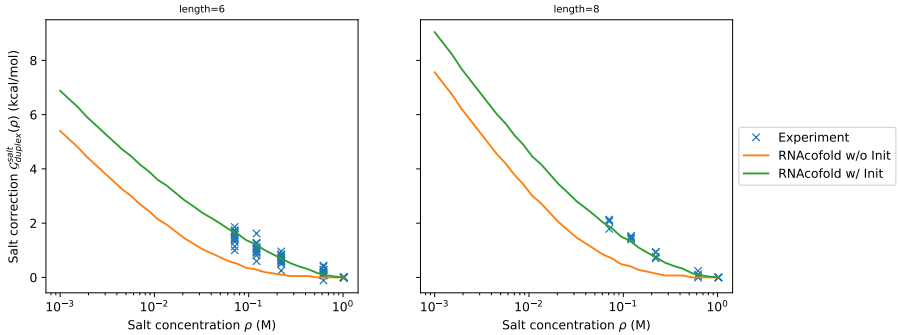
Salt correction of 18 duplexes of length 6 and 8 from experiment (Chen & Znosko, 2013)



Salt correction of a duplex $\mathcal{G}^{salt}_{\text{duplex}}(\rho) : h \times g^{salt}_{\text{stack}}(\rho)$

Salt correction of duplex initiation

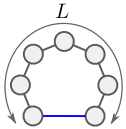
Salt correction of 18 duplexes of length 6 and 8 from experiment (Chen & Znosko, 2013)

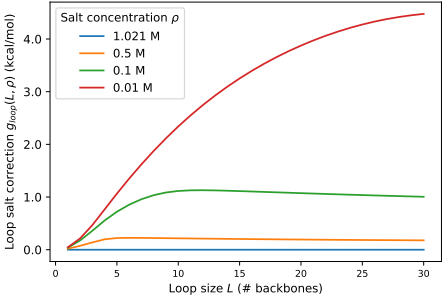


Salt correction of a duplex $\mathcal{G}^{salt}_{\text{duplex}}(\rho) : h \times g^{salt}_{\text{stack}}(\rho) + g_{\text{init}}(\rho)$

Salt correction of a loop

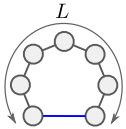
Salt correction $\mathcal{G}^{salt}_{\text{loop}}(L, \rho)$ of a loop at concentration ρ depends on # backbones L.

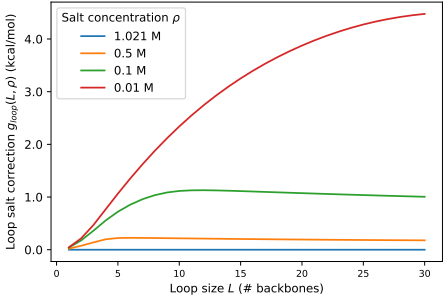




Salt correction of a loop

Salt correction $\mathcal{G}^{salt}_{\text{loop}}(L, \rho)$ of a loop at concentration ρ depends on # backbones L.

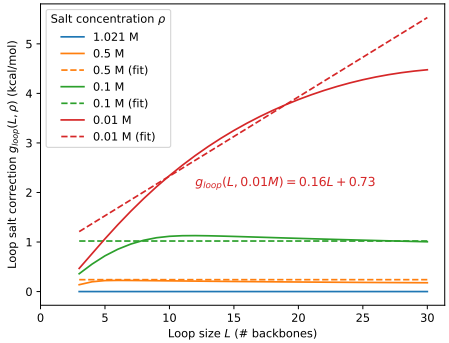




$\Delta($  $), \Delta($  $) : \mathcal{E}_{\text{loop}}(m, w) + \mathcal{G}^{salt}_{\text{loop}}(m+1 \text{ or } 2, \rho)$

$\Delta($  $) : m \times \mathcal{E}_{\text{MLbase}} + x \times \mathcal{E}_{\text{MLstem}} + \mathcal{E}_{\text{MLclosing}}$???

Linear fit of loop salt correction



$$\mathcal{G}^{salt}_{\text{loop}}(L, \rho) \approx m_\rho \times L + c_\rho$$

$$\Delta(\text{  }) : m \times (\mathcal{E}_{\text{MLbase}} + m_\rho) + x \times (\mathcal{E}_{\text{MLstem}} + m_\rho) + (\mathcal{E}_{\text{MLclosing}} + m_\rho + c_\rho)$$

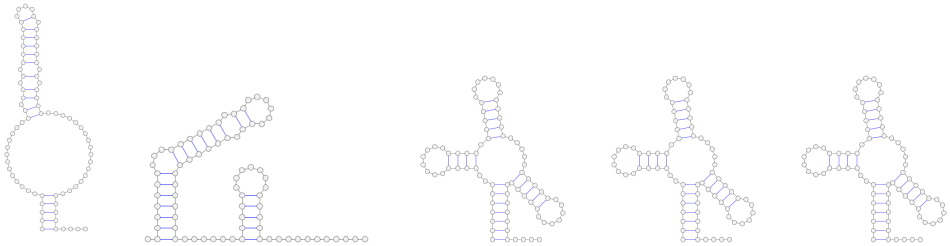
```
import RNA
rho = 0.1
w = RNA.random_string(50, 'ACGU')

md = RNA.md(salt=rho)
fc = RNA.fold_compound(w, md)
```

Upadate the parameter table while creating the fold compound

$$\mathcal{E}_{\text{stack}}(w) \leftarrow \mathcal{E}_{\text{stack}}(w) + g^{salt}_{\text{stack}}(\rho)$$
$$\mathcal{E}_{\text{loop}}(m, w) \leftarrow \mathcal{E}_{\text{loop}}(m, w) + \mathcal{G}^{salt}_{\text{loop}}(m + 1 \text{ or } 2, \rho)$$
$$\mathcal{E}_{\text{MLbase}} \leftarrow \mathcal{E}_{\text{MLbase}} + m_\rho$$
$$\mathcal{E}_{\text{MLstem}} \leftarrow \mathcal{E}_{\text{MLstem}} + m_\rho$$
$$\mathcal{E}_{\text{MLclosing}} \leftarrow \mathcal{E}_{\text{MLclosing}} + m_\rho + c_\rho$$
$$\mathcal{E}_{\text{init}} \leftarrow \mathcal{E}_{\text{init}} + g_{\text{init}}(\rho)$$

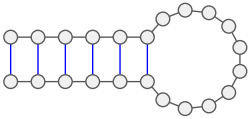
Case study: tRNA

tRNA (76 nts): GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA

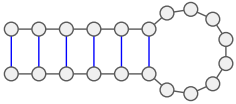


MFE Prediction					
ρ	0.003 M	0.005 M	0.018 M	0.03 M	ρ_0
Free energy (kcol/mol)	-1.65	-3.37	-9.15	-11.45	-22.40

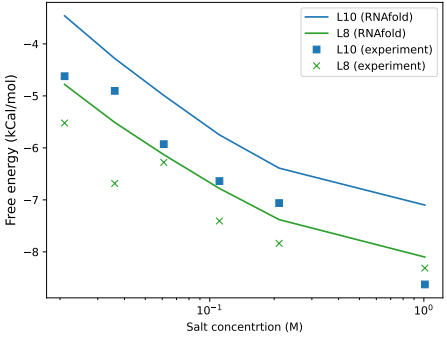
Case study: Hairpins (Williams *et al.*, 1996)



L10



L8



Duplex - Melting temperature

Melting temperature T_m: Temperature at which half of monomer A form dimer AA,

$$2A \rightleftharpoons AA$$

$$[\mathsf{AA}]_{T=T_m} = \frac{C_t}{4}$$

where C_t is the total species (A) concentration

Duplex - Melting temperature

Melting temperature T_m: Temperature at which half of monomer A form dimer AA,

$$2A \rightleftharpoons AA$$

$$[\mathsf{AA}]_{T=T_m} = \frac{C_t}{4}$$

where C_t is the total species (A) concentration

Van't Hoff analysis:

$$\frac{1}{T_m} = \frac{R}{\Delta H} \ln C_t + \frac{\Delta S}{\Delta H}.$$

Duplex - Melting temperature

Melting temperature T_m: Temperature at which half of monomer A form dimer AA,

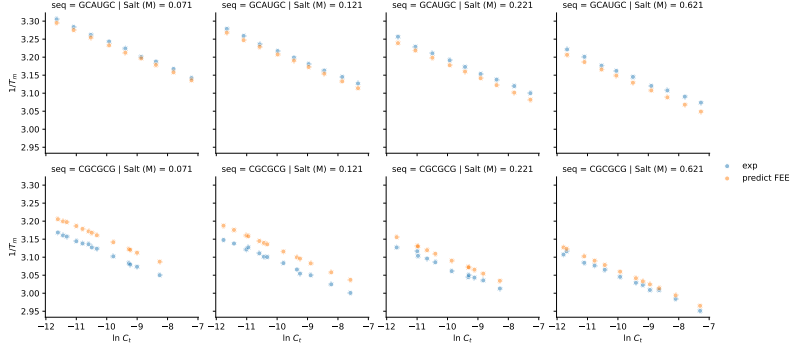
$$2A \rightleftharpoons AA$$

$$[\mathsf{AA}]_{T=T_m} = \frac{C_t}{4}$$

where C_t is the total species (A) concentration

Van't Hoff analysis:

$$\frac{1}{T_m} = \frac{R}{\Delta H} \ln C_t + \frac{\Delta S}{\Delta H}.$$

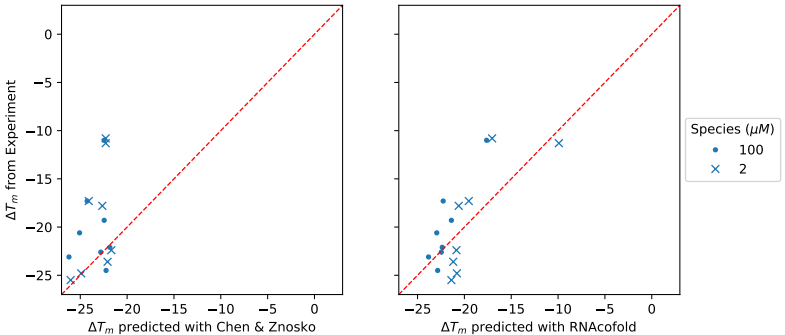


Long duplexes (Nakano *et al.*, 2007)

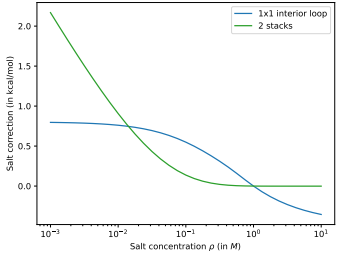
- 8 duplexes of length 10, 12, and 14 with $C_t = 2\mu M$, $100\mu M$ at $\rho = 0.01M$, $1M$
- Melting temperature correction

$$\Delta T_m(\rho, C_t) = T_m(\rho, C_t) - T_m(\rho_0, C_t)$$

- $\Delta T_m^{C\&Z}(\rho) = (-1.842\ \mathrm{gc}_w + 2.675)\ln(\rho/\rho_0) - 0.7348(\ln(\rho/\rho_0))^2$



- Take mismatch into account



- Incorporate ion binding model