

Pseudoknot prevalence and base pair identity distance in prediction of secondary structures

The distance $D^{\zeta, \zeta'}$ between thermodynamic ensembles of secondary structures ζ including pseudoknots and ζ' excluding them is defined as:

$$D^{\zeta, \zeta'} = \frac{\sum_{i,j} |P_{i,j}^{\zeta} - P_{i,j}^{\zeta'}|}{\frac{1}{2} \sum_{i,j} (P_{i,j}^{\zeta} + P_{i,j}^{\zeta'} + |P_{i,j}^{\zeta} - P_{i,j}^{\zeta'}|)}, \text{ where } P_{i,j}^{\zeta} = \frac{\sum_{k=0}^n e^{-\beta E_k}}{Z_{\zeta}} \delta_k^{\zeta}(i,j) \text{ and } Z_{\zeta} = \sum_{k=0}^n e^{-\beta E_k}$$

and $\delta_k^{\zeta}(i,j) = 1$ iff i and j are paired in the k th structure of the thermodynamic ensemble ζ and $\delta_k^{\zeta}(i,j) = 0$ otherwise. With this definition, the distance distribution between the two thermodynamic ensembles of secondary structures (right graphs) would be exactly equivalent to the distribution of pseudoknot proportion (left graphs[19]) if pseudoknots could just be added to secondary structure predictions excluded them. The fact that right graphs are virtually flat (except for sequences having essentially no pseudoknots) reveals the strong cooperativity of secondary structure rearrangements. Statistics is done on thousands of random RNA sequences[19]. Number of lowest free energy structures included in each thermodynamic ensemble: $n=3$.

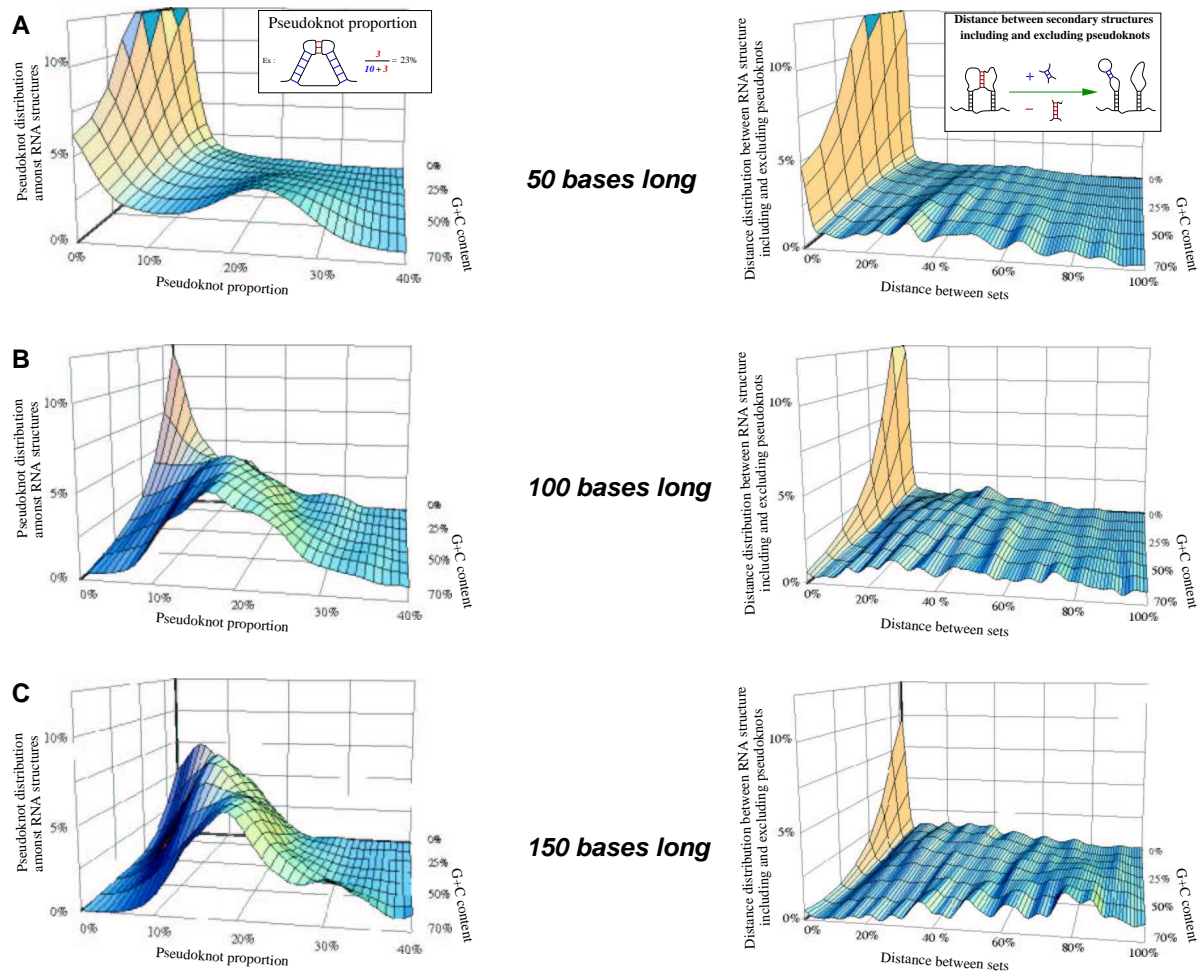
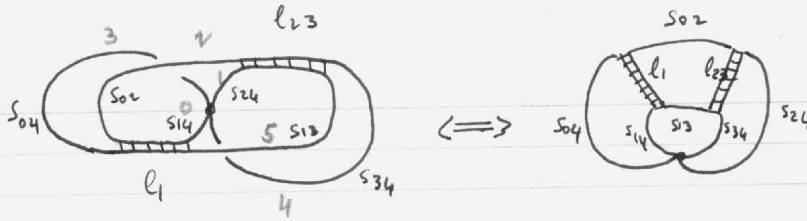


Figure 4. **Right:** distribution of pseudoknot proportion amongst formed base pairs for 50-nt-long (A), 100-nt-long (B), 150-nt-long (C) random RNA sequences of increasing G+C content. **Left:** distribution of distance between low energy RNA secondary structures folded with and without pseudoknots.

Example of conformational entropy calculation

Contribution to a helix nucleation transition within a « net » substructure

(scanned from the original notebook)



$$e^{-\beta F_{\text{tot}}} = \frac{e^{-A_1 l_1^2 - A_{23} l_2^2}}{\pi^9 C^{3/2}} \left\{ \begin{array}{l} e^{2A_{23} l_2 l_3} - e^{-2A_{23} l_2 l_3} \\ 4A_{23} l_2 l_3 \end{array} \right. \quad \text{if } A_{23} l_2 l_3 = 0$$

$$C = (s_{02} + s_{13}) \left(\overset{2 \ 5}{s_{24} s_{14}} (s_{24} + s_{34}) + \overset{3 \ 0 \ 1 \ 4}{s_{24} s_{34}} (s_{04} + s_{14}) \right) + \overset{2 \ 5}{s_{02} s_{13}} (s_{04} + s_{14}) \left(\overset{3 \ 0 \ 1 \ 4}{s_{24} + s_{34}} \right)$$

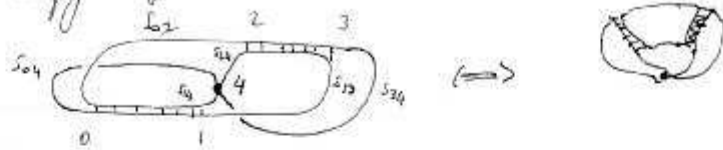
$$A_1 = \overset{1 \ 4}{s_{24} s_{34}} (s_{04} + s_{14} + s_{02} + s_{13}) + \overset{1 \ 4}{(s_{24} + s_{34})} \left[\overset{3 \ 0 \ 2 \ 5 \ 3 \ 5 \ 0 \ 2}{s_{04} s_{14} + s_{02} s_{13} + s_{04} s_{13} + s_{14} s_{02}} \right]$$

$$A_{23} = \overset{0 \ 3}{s_{04} s_{14}} (s_{24} + s_{34} + s_{02} + s_{13}) + \overset{C \ 0 \ 3}{(s_{04} + s_{14})} \left[\overset{1 \ 4 \ 2 \ 5 \ 1 \ 5 \ 2 \ 4}{s_{24} s_{34} + s_{02} s_{13} + s_{04} s_{13} + s_{14} s_{02}} \right]$$

$$A_{123} = \overset{3 \ 0 \ 1 \ 4}{s_{04} s_{14}} (s_{24} + s_{34}) + \overset{C}{s_{24} s_{34}} (s_{04} + s_{14}) + \overset{3}{s_{02} s_{14} s_{34}} + \overset{3}{s_{13} s_{04} s_{24}}$$

①

Free energy of β



$$\begin{aligned}
 & e^{-\frac{\bar{r}_4^2}{s_{04}}} - \frac{(\bar{r}_4 - \bar{r}_1)^2}{s_{14}} - \frac{(\bar{r}_4 - \bar{r}_2)^2}{s_{24}} - \frac{(\bar{r}_4 - \bar{r}_3)^2}{s_{34}} e^{-\frac{(\bar{r}_3 - \bar{r}_1)^2}{s_{13}}} e^{-\frac{s_2^2}{s_{02}}} \\
 & \frac{1}{\pi^4 (s_{04} s_{14} s_{24} s_{34} s_{13} s_{02})^{3/4}} \frac{\delta(|r_1| - l) \delta(|r_3 - r_2| - l_{23})}{4\pi l^2 \quad 4\pi l_{23}^2} \\
 & d\bar{r}_1 d\bar{r}_2 d\bar{r}_3 d\bar{r}_4
 \end{aligned}$$

$$\begin{aligned}
 & e^{-\bar{r}_4^2 \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}} \right) + 2\bar{r}_4 \cdot \left(\frac{\bar{r}_1}{s_{14}} + \frac{\bar{r}_2}{s_{24}} + \frac{\bar{r}_3}{s_{34}} \right)} \\
 & e^{-\frac{\bar{r}_1^2}{s_{14}} - \bar{r}_2^2 \left(\frac{1}{s_{02}} + \frac{1}{s_{24}} \right) - \frac{\bar{r}_3^2}{s_{34}} - \frac{(\bar{r}_3 - \bar{r}_1)^2}{s_{13}}}
 \end{aligned}$$

$$\left[\int \dots \right] \frac{1}{4\pi l^2} \frac{\delta(\dots)}{4\pi l_{23}^2} \frac{1}{\pi^4} d\bar{r}_i$$

$$\int d\bar{r}_4 e^{-\bar{r}_4^2 \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}} \right) + 2\bar{r}_4 \cdot \left(\frac{\bar{r}_1}{s_{14}} + \frac{\bar{r}_2}{s_{24}} + \frac{\bar{r}_3}{s_{34}} \right)}$$

$$= \left[\sqrt{\frac{\pi}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}}} \right]^3 e^{-\frac{\left(\frac{\bar{r}_1}{s_{14}} + \frac{\bar{r}_2}{s_{24}} + \frac{\bar{r}_3}{s_{34}} \right)^2}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}}}$$

(2).

$$e^{-\frac{\left(\frac{\bar{r}_1}{s_{14}} + \frac{\bar{r}_1 + 1}{s_{24} s_{34}}\right) \left(\frac{\bar{r}_2 - \bar{r}_1}{s_{34}}\right)^2}{\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} - \frac{(\bar{r}_2 - \bar{r}_1 + \bar{r}_1)^2}{s_{34}} - \frac{(\bar{r}_2 - \bar{r}_1 + \bar{r}_1 - \bar{r}_1)^2}{s_{13}}}$$

$$e^{-\frac{\bar{r}_1^2}{s_{14}} - \bar{r}_1^2 \left(\frac{1}{s_{24}} + \frac{1}{s_{24}}\right)} \frac{\delta(|\bar{r}_2 - \bar{r}_1| - \ln 2)}{\ln^2} d(\bar{r}_2 - \bar{r}_1)$$

$$e^{-\left(\frac{\bar{r}_2 - \bar{r}_1}{s_{34}}\right)^2 \left[\frac{1}{s_{14}} + \frac{1}{s_{13}} - \frac{1}{s_{34}^2 \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}\right)} \right]}$$

$$e^{-2(\bar{r}_2 - \bar{r}_1) \left[\frac{\bar{r}_1}{s_{34}} + \frac{\bar{r}_2 - \bar{r}_1}{s_{13}} - \frac{\frac{\bar{r}_1}{s_{14}} + \bar{r}_2 \left(\frac{1}{s_{14}} + \frac{1}{s_{34}}\right)}{s_{34} \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}\right)} \right]}$$

$$e^{-\frac{\bar{r}_1^2}{s_{14}} - \bar{r}_1^2 \left(\frac{1}{s_{24}} + \frac{1}{s_{24}}\right) + \left(\frac{\bar{r}_1}{s_{14}} + \bar{r}_2 \left(\frac{1}{s_{24}} + \frac{1}{s_{34}}\right)\right)^2 \frac{1}{\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}}}$$

$$e^{-\frac{\bar{r}_2^2}{s_{34}} - \frac{(\bar{r}_2 - \bar{r}_1)^2}{s_{13}}} \frac{\delta(|\bar{r}_2 - \bar{r}_1| - \ln 2)}{\ln^2} d(\bar{r}_2 - \bar{r}_1)$$

$$\left[\frac{1}{s_{24}} + \frac{1}{s_{13}} - \frac{1}{s_{34}^2 \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}\right)} \right]$$

$$= \frac{\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}}}{s_{34} \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}\right)} + \frac{1}{s_{13}}$$

$$= \frac{s_{24} s_{14} + s_{14} s_{24} + s_{24} s_{24}}{s_{24} s_{14} s_{24} + s_{34} (s_{24} s_{14} + s_{14} s_{24} + s_{24} s_{24})} + \frac{1}{s_{13}}$$

$$= \frac{(s_{24} + s_{13}) \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} \right) + s_{24} s_{14} s_{24}}{s_{13} [s_{24} s_{14} s_{24} + s_{34} (\quad)]} \quad (c)$$

3.

$$-2(\bar{\Gamma}_3 - \bar{\Gamma}_2) \cdot \left[\bar{\Gamma}_2 \left(\frac{1}{S_{34}} + \frac{1}{S_{13}} - \frac{\frac{1}{S_{24}} + \frac{1}{S_{34}}}{S_{24} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} \right) - \bar{\Gamma}_1 \left(\frac{1}{S_{13}} + \frac{1}{S_{14} S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} \right) \right]$$

$$-2(\bar{\Gamma}_3 - \bar{\Gamma}_2) \cdot \left[\bar{\Gamma}_2 \left(\frac{\frac{1}{S_{24}} + \frac{1}{S_{14}}}{S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} + \frac{1}{S_{13}} \right) - \bar{\Gamma}_1 \left(\frac{S_{24} S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right) + \frac{1}{S_{13}}}{S_{13} S_{14} S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} \right) \right]$$

$$-2(\bar{\Gamma}_3 - \bar{\Gamma}_2) \cdot \left[\bar{\Gamma}_2 \left(\frac{\left(\frac{1}{S_{24}} + \frac{1}{S_{14}} \right) \frac{1}{S_{13}} + S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)}{S_{24} S_{13} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} - \bar{\Gamma}_1 \frac{S_{24} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right) + \frac{1}{S_{13} S_{14}}}{S_{13} S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} \right]$$

$$-2(\bar{\Gamma}_3 - \bar{\Gamma}_2) \cdot \left(\frac{\left(\frac{1}{S_{24}} + \frac{1}{S_{14}} \right) \frac{1}{S_{13}} + S_{34} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)}{S_{24} S_{13} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)} \right) \textcircled{A}$$

$$\cdot \left[\bar{\Gamma}_2 - \bar{\Gamma}_1 \left(\frac{1 + \frac{S_{13}}{S_{34}} \frac{1}{S_{14}} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)^{-1}}{1 + \frac{S_{13}}{S_{24}} \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} \right) \left(\frac{1}{S_{24}} + \frac{1}{S_{14}} + \frac{1}{S_{24}} + \frac{1}{S_{34}} \right)^{-1}} \right) \right] \textcircled{B}$$

$$-2(\bar{\Gamma}_3 - \bar{\Gamma}_2) \cdot A \left[\bar{\Gamma}_2 - \bar{\Gamma}_1 B \right]$$

⑤.

$$\int e^{-\tau_1^2} \left[\frac{1}{s_{14}} + B^2 \left(\frac{1}{s_{02}} + \frac{1}{s_{24}} \right) - \frac{1}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \left(B \left(\frac{1}{s_{04}} + \frac{1}{s_{34}} \right) + \frac{1}{s_{14}} \right)^2 + \frac{B^2}{s_{34}} + \frac{(B-1)^2}{s_{13}} \right]$$

$$e^{-2|\bar{\tau}_2 - B\bar{\tau}_1| \cos \theta |\tau_1|} \left[B \left(\frac{1}{s_{02}} + \frac{1}{s_{24}} \right) - \frac{\frac{1}{s_{04}} + \frac{1}{s_{34}}}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \left(B \left(\frac{1}{s_{04}} + \frac{1}{s_{34}} \right) + \frac{1}{s_{14}} \right) \right]$$

$$\left[\frac{B}{s_{34}} + \frac{B-1}{s_{13}} \right] \textcircled{1}$$

$$e^{-(\bar{\tau}_2 - B\bar{\tau}_1)^2} \left[\frac{1}{s_{02}} + \frac{1}{s_{24}} - \frac{\left(\frac{1}{s_{04}} + \frac{1}{s_{34}} \right)^2}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} + \frac{1}{s_{34}} + \frac{1}{s_{13}} \right] \textcircled{2}$$

$$e^{-\ell_{23}^2 C}$$

$$\frac{e^{2A \ell_{23} |\bar{\tau}_2 - B\bar{\tau}_1|} - e^{-2A \ell_{23} |\bar{\tau}_2 - B\bar{\tau}_1|}}{4A \ell_{23} |\bar{\tau}_2 - B\bar{\tau}_1|} \cdot \frac{d(\bar{\tau}_2 - B\bar{\tau}_1)}{d\tau_1} \cdot \frac{\delta(\tau_1 - \ell_1)}{4A \ell_1^2}$$

$$\rightarrow \boxed{\frac{A \cdot \textcircled{1}}{\textcircled{2}} = A_{123}}$$

(6)

$$\textcircled{1} = \left[B \left\{ \frac{1}{s_{02}} + \left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \left[1 - \frac{\frac{1}{s_{24}} + \frac{1}{s_{34}}}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \right] \right\} \right] \frac{1}{s_{13}}$$

$$- \frac{1}{s_{14}} \frac{\frac{1}{s_{24}} + \frac{1}{s_{34}}}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} - \frac{1}{s_{13}}$$

$$= \left[B \left\{ \frac{1}{s_{02}} + \frac{1}{s_{13}} + \frac{\left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} \right)}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \right\} \right]$$

$$- \frac{1}{s_{14}} \frac{\frac{1}{s_{24}} + \frac{1}{s_{34}}}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} - \frac{1}{s_{13}}$$

$$\textcircled{1} A = \left[AB \left\{ \frac{1}{s_{02}} + \frac{1}{s_{13}} + \frac{\left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} \right)}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \right\} \right]$$

$$- A \left(\frac{1}{s_{13}} + \frac{1}{s_{14}} \frac{\frac{1}{s_{24}} + \frac{1}{s_{34}}}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}} \right)$$

$$A = \frac{1}{s_{13}} + \frac{\frac{1}{s_{24}} + \frac{1}{s_{34}}}{s_{14} \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}} \right)}$$

$$AB = \frac{1}{s_{13}} + \frac{\frac{1}{s_{14}}}{s_{14} \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}} \right)}$$

$$\textcircled{2} = \frac{1}{s_{02}} + \frac{1}{s_{13}} + \frac{\left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \left(\frac{1}{s_{04}} + \frac{1}{s_{14}} \right)}{\frac{1}{s_{04}} + \frac{1}{s_{14}} + \frac{1}{s_{24}} + \frac{1}{s_{34}}}$$

(P)

$$\frac{\textcircled{1} A}{\textcircled{2}} = AB - \frac{A}{\textcircled{2}} \left(\frac{1}{s_3} + \frac{1}{s_4} \frac{\frac{1}{s_{22}} + \frac{1}{s_{34}}}{\frac{1}{s_{22}} + \dots + \frac{1}{s_{34}}} \right)$$

$$A \begin{bmatrix} \\ \end{bmatrix} = \frac{\frac{1}{s_3} () + \frac{1}{s_{24}} \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} \right)}{()} \cdot \frac{\frac{1}{s_3} () + \frac{1}{s_4} \left(\frac{1}{s_{22}} + \frac{1}{s_{34}} \right)}{()}$$

$$\frac{1}{\textcircled{2}} = \frac{()}{\left(\frac{1}{s_{22}} + \frac{1}{s_{13}} \right) () + \left(\frac{1}{s_{22}} + \frac{1}{s_{14}} \right) \left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right)}$$

$$\frac{A \begin{bmatrix} \\ \end{bmatrix}}{\textcircled{2}} = \frac{\left[\frac{1}{s_3} () + \frac{1}{s_{24}} \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} \right) \right] \left[\frac{1}{s_3} () + \frac{1}{s_4} \left(\frac{1}{s_{22}} + \frac{1}{s_{34}} \right) \right]}{() \left[\left(\frac{1}{s_{22}} + \frac{1}{s_{13}} \right) () + \left(\frac{1}{s_{22}} + \frac{1}{s_{14}} \right) \left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \right]}$$

$$AB = \frac{\frac{1}{s_3} () + \frac{1}{s_4} \frac{1}{s_{24}}}{()}$$

$$AB - \frac{A \begin{bmatrix} \\ \end{bmatrix}}{\textcircled{2}} = - \left[\frac{1}{s_3} () + \frac{1}{s_{24}} \left(\frac{1}{s_{24}} + \frac{1}{s_{14}} \right) \right] \left[\frac{1}{s_3} () + \frac{1}{s_4} \left(\frac{1}{s_{22}} + \frac{1}{s_{34}} \right) \right] \\ + \left[\frac{1}{s_3} () + \frac{1}{s_4} \frac{1}{s_{24}} \right] \left[\left(\frac{1}{s_{22}} + \frac{1}{s_{13}} \right) () + \left(\frac{1}{s_{22}} + \frac{1}{s_{14}} \right) \left(\frac{1}{s_{24}} + \frac{1}{s_{34}} \right) \right]$$
$$() \left[\begin{matrix} \text{par} & \text{top} \\ \text{par} & \text{top} \end{matrix} \right]$$

... 8

$$AB - \frac{AEI}{(2)} = \frac{1}{S_{13}} \frac{1}{S_{02}} ()^2$$
$$\frac{1}{S_{13}} () \left\{ -\frac{1}{S_{14}} \left(\frac{1}{S_{24}} + \frac{1}{S_{34}} \right) - \frac{1}{S_{24}} \left(\frac{1}{S_{04}} + \frac{1}{S_{14}} \right) + \left(\frac{1}{S_{04}} + \frac{1}{S_{14}} \right) \left(\frac{1}{S_{24}} + \frac{1}{S_{34}} \right) + \frac{1}{S_{14}} \frac{1}{S_{24}} \right\}$$

$$\frac{1}{S_{02}} () \frac{1}{S_{14}} \frac{1}{S_{34}}$$

$$- \frac{1}{S_{14}} \frac{1}{S_{24}} \left(\frac{1}{S_{04}} + \frac{1}{S_{14}} \right) \left(\frac{1}{S_{24}} + \frac{1}{S_{34}} \right)$$

$$+ \frac{1}{S_{14}} \frac{1}{S_{24}} \left(\frac{1}{S_{04}} + \frac{1}{S_{14}} \right) \left(\frac{1}{S_{24}} + \frac{1}{S_{34}} \right)$$

$$() \left[\frac{1}{S_{04}} + \frac{1}{S_{14}} \right]$$

$$= \frac{1}{S_{13}} \frac{1}{S_{02}} ()^2 + \frac{1}{S_{13}} \frac{1}{S_{04}} \frac{1}{S_{24}} () + \frac{1}{S_{24}} \frac{1}{S_{14}} \frac{1}{S_{02}} ()$$

$$() \left[\frac{1}{S_{04}} + \frac{1}{S_{14}} \right]$$

$$= \frac{1}{S_{13}} \frac{1}{S_{02}} () + \frac{1}{S_{13}} \cdot \frac{1}{S_{04}} \frac{1}{S_{24}} + \frac{1}{S_{02}} \cdot \frac{1}{S_{14}} \cdot \frac{1}{S_{24}}$$

$$\left(\frac{1}{S_{02}} + \frac{1}{S_{13}} \right) () + \left(\frac{1}{S_{04}} + \frac{1}{S_{14}} \right) \left(\frac{1}{S_{24}} + \frac{1}{S_{34}} \right)$$

(9) :

$$\times s_{13} s_{12} s_{04} s_{14} s_{24} s_{34}$$

$$AB - A \begin{matrix} \rightarrow \\ \rightarrow \\ \rightarrow \end{matrix} = \frac{\{s_{14} s_{24} s_{34} + s_{04} s_{24} s_{34} + s_{04} s_{14} s_{34} + s_{04} s_{14} s_{24}\} + s_{02} s_{14} s_{34} + s_{13} s_{04} s_{24}}{(s_{02} + s_{13}) \{ \} + s_{02} s_{13} (s_{04} + s_{14}) (s_{24} + s_{34})}$$

$$= \frac{\{ \} + s_{02} s_{14} s_{34} + s_{13} s_{04} s_{24}}{(s_{02} + s_{13}) (s_{04} + s_{14}) (s_{24} + s_{34})} \left[s_{02} // s_{13} + s_{04} // s_{14} + s_{24} // s_{34} \right]$$

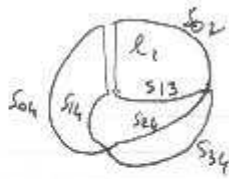
$$C = (s_{02} + s_{13}) \left\{ s_{14} s_{24} s_{34} + s_{04} s_{24} s_{34} + s_{04} s_{14} s_{34} + s_{04} s_{14} s_{24} \right\} + s_{02} s_{13} (s_{04} + s_{14}) (s_{24} + s_{34})$$

$$A \begin{matrix} \rightarrow \\ \rightarrow \\ \rightarrow \end{matrix} = \left\{ s_{14} s_{24} s_{34} + s_{04} s_{24} s_{34} + s_{04} s_{14} s_{34} + s_{04} s_{14} s_{24} \right\} + s_{02} s_{14} s_{34} + s_{13} s_{04} s_{24}$$

$$C = (s_{02} + s_{13}) (s_{04} + s_{14}) (s_{24} + s_{34}) \left[s_{02} // s_{13} + s_{04} // s_{14} + s_{24} // s_{34} \right]$$

(b) ...

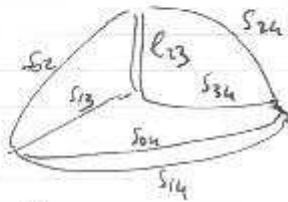
$l_{23} \rightarrow 0$



$$\text{Fzt.} \left\{ \begin{array}{l} s_{12} \rightarrow s_{13} \\ s_{34} \rightarrow s_{24} // s_{34} \end{array} \right.$$

$$A_1 = \frac{s_{24} s_{34} (s_{04} + s_{14} + s_{02} + s_{13}) + (s_{24} + s_{34}) (s_{04} s_{14} + s_{02} s_{13} + s_{04} s_{13} + s_{14} s_{02})}{(s_{04} s_{14} (s_{24} + s_{34}) + (s_{04} + s_{14}) s_{24} s_{34}) (s_{02} + s_{13}) + s_{02} s_{13} (s_{04} + s_{14}) (s_{24} + s_{34})}$$

$l_1 \rightarrow 0$



As previous with $s_{04} \leftrightarrow s_{24}$
 $s_{14} \leftrightarrow s_{34}$

$$A_{23} = \frac{s_{04} s_{14} (s_{24} + s_{34} + s_{02} + s_{13}) + (s_{04} + s_{14}) (s_{24} s_{34} + s_{02} s_{13} + s_{24} s_{13} + s_{34} s_{02})}{C}$$

Output of the Kinefold Server

OUTPUT PAGE

The output page of Kinefold provides (1) a series of *low free energy structures*, (2) an *online animated folding path* and (3) a *programmable trajectory plot* focussing on a few helices of interest to each user.

Lowest free energy structures Low free energy structures visited along the simulated folding path are displayed as small images which can be enlarged through a simple click (Fig. 3, main manuscript). These structures are ordered by increasing free-energies, starting with the lowest free energy structure ever found during the simulation. Some of these structures may not be shown amongst the selected structures displayed in the online animation (*.coord* and *.rnm* files), see below.

The free-energies are estimated using the stacking interaction tables from the Turner lab for RNA[34,35] and the SantaLucia lab for DNA[36] (although the mismatch contributions used are somewhat less detailed). In addition, we also estimate the physical contribution of pseudoknots and entangled helices to global free energy of the molecule, as described earlier.

Animated visualization of folding paths The simulated stochastic folding paths can be visualized *either* directly online with the Kineviewer java applet (Fig. 5) *or* offline with either the standalone Kineviewer java application (input format *.coord* files) or the RNAMovies software upgraded to display pseudoknots and entangled helices (input format *.rnm* files)[38]. The *.rnm* files are in fact used by RNAMovies to generate the *.coord* files, which can then be visualized online or offline using Kineviewer. The animation displays each new lowest free energy structure visited along the stochastic folding path (which is not equivalent to all lowest free-energy structures in reverse order). The total number of structures visited or revisited in the meantime is dynamically displayed in the comment field of the Kineviewer (renaturation fold only). The current 3'-end of the sequence is also indicated (co-transcriptional fold only).

Trajectory plot. The trajectory plot (Fig. 6) follows the lowest free energy during RNA folding (plotted below time-axis), while monitoring which helices are formed amongst the 5 most stable helices (plotted above time-axis). The helix presence is averaged over time and displayed graphically between 0% (not formed) and 100% (always formed). Alternatively, users can choose which helices to follow along the folding path by specifying so in the request form. (see *Tracing and forcing helices*). Time-axis on the trajectory plot is automatically adapted to each stochastic simulation type:

- For a renaturation fold it is in log scale, to display both the rapid initial collapse of the molecules and their much slower structural rearrangements thereafter.
- For a co-transcriptional fold (Fig. 6), the time-axis is first in linear scale during transcription (which proceeds at constant rate) and in log scale after transcription.

Users can download the multicolumn trajectory plot file (*.e*), if they would like to plot the graph differently from the available one (*.jpg*). The first lines of the file correspond to the coordinates of the traced helices and the following lines and columns are in order : the simulated molecular time, the lowest free energy reached so far and the presence of each traced helices.

ZIP PAGE

The zip page provides a variety of result files (images and raw files) which can be individually selected for automatic zipping and downloading.

IMPLEMENTATION

The web service is implemented in static HTML pages and CGI scripts in PERL. KineViewer, which runs on the user's machine, uses a JAVA Applet for portability; only Java Runtime Environment is required (JRE is available on most platform) and users also need to activate JAVA in their browser parameters. The server manager, which takes care of the batch queuing system, and the actual Kinefold program, which predicts RNA/DNA folding paths and structures, are both implemented in ANSI C.

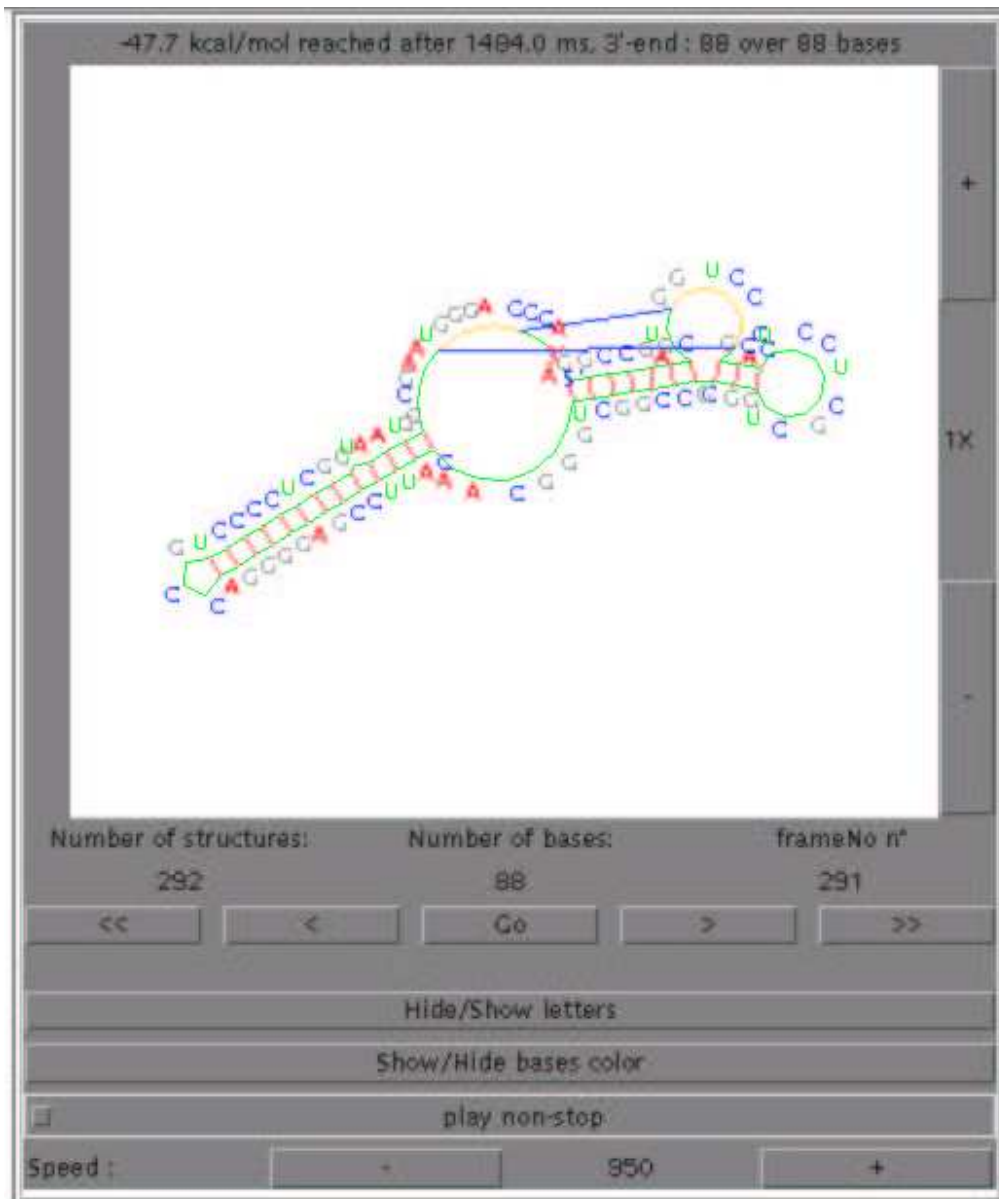


Figure 5. **Snapshot of the Kineviewer java applet** used to visualize animated RNA/DNA folding pathways online. Dynamic comments are displayed above the animation which can be zoomed up to 5X (right buttons) and visualized as a loop and at different speeds with hidden or shown letters and base colors (bottom buttons).

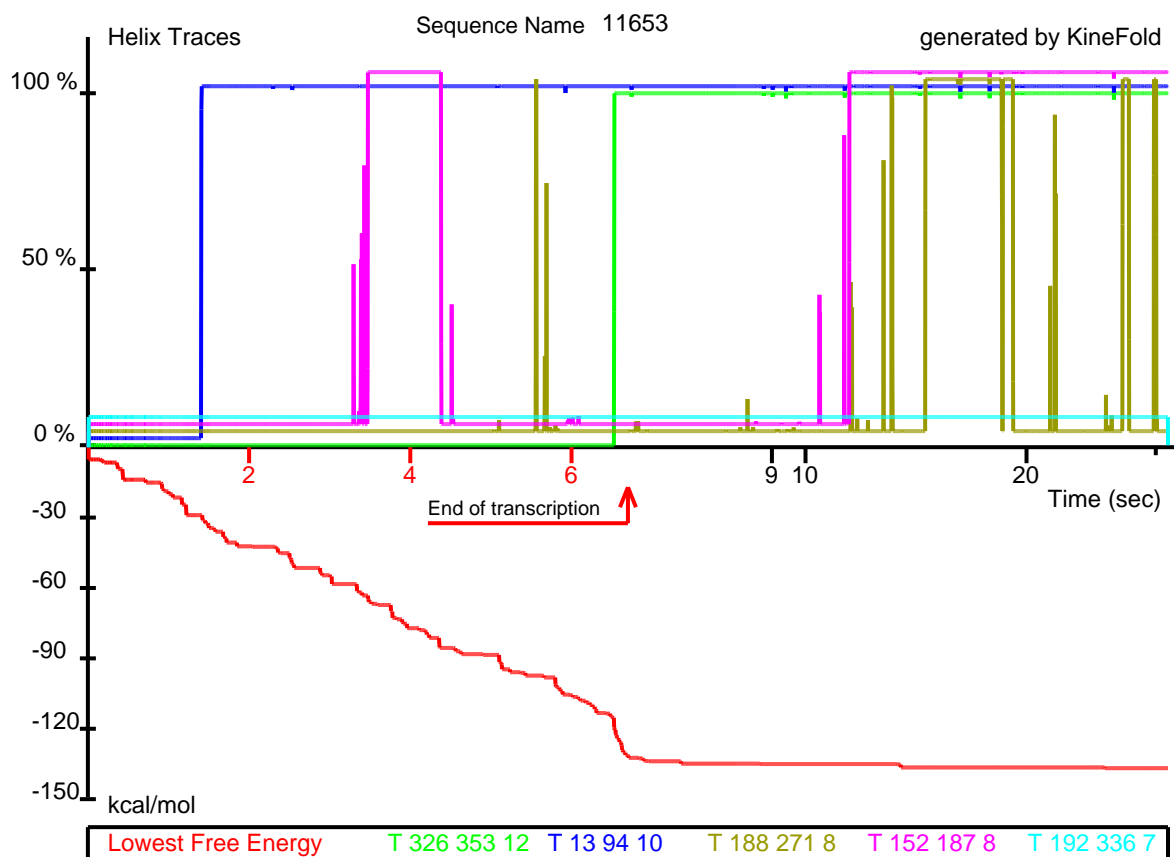


Figure 6. **Illustration of a trajectory plot** following the lowest free energy during RNA/DNA folding (plotted below time-axis) and monitoring (above time-axis) which helices amongst the 5 most stable ones (default) or amongst helices of particular interest to each user are formed along the co-transcriptional or renaturation folding path (see *Tracing helices* section).