Nucleic Acids Research, 2005, Vol. 33, Web Server issue ***W605–W610***
doi:10.1093/nar/gki447

Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots

A. Xayaphoummine[1], T. Bucher[1] and H. Isambert[1,2,*]

[1]Laboratoire de Dynamique des Fluides Complexes, CNRS-ULP, Institut de Physique, 3 rue de l'Université, 67000 Strasbourg, France and [2]Physico-chimie Curie, CNRS UMR168, Institut Curie, Section de Recherche, 11 rue P. & M. Curie, 75005 Paris, France

Received February 15, 2005; Revised and Accepted March 31, 2005

ABSTRACT

The Kinefold web server provides a web interface for stochastic folding simulations of nucleic acids on second to minute molecular time scales. Renaturation or co-transcriptional folding paths are simulated at the level of helix formation and dissociation in agreement with the seminal experimental results. Pseudoknots and topologically 'entangled' helices (i.e. knots) are efficiently predicted taking into account simple geometrical and topological constraints. To encourage interactivity, simulations launched as immediate jobs are automatically stopped after a few seconds and return adapted recommendations. Users can then choose to continue incomplete simulations using the batch queuing system or go back and modify suggested options in their initial query. Detailed output provide (i) a series of low free energy structures, (ii) an online animated folding path and (iii) a programmable trajectory plot focusing on a few helices of interest to each user. The service can be accessed at http://kinefold.curie.fr/.

INTRODUCTION

The prediction of RNA secondary structures with no known homologous sequences has been traditionally performed through free energy minimization using efficient 'dynamic programming' algorithms (1–6). These approaches and recent developments combining thermodynamic and homology information have now been interfaced by a number of web servers for RNA/DNA structure prediction (7–11).

Simulations of nucleic acid folding paths (12–19) can complement these primary thermodynamic approaches and provide interesting insights into DNA/RNA folding dynamics. Applications include, in particular, anomalous hybridization kinetics on DNA chips and structural rearrangements of non-coding RNA regions involved in (post)-transcriptional regulation of genetic expression.

The Kinefold web server performs online stochastic folding simulations of nucleic acids using the folding dynamics algorithm introduced by Isambert and Siggia (17) and further enhanced by the 'exactly clustered stochastic simulation' algorithm detailed by Xayaphoummine *et al*. (19). The web server was launched at Institut de Physique (Strasbourg, France) in August 2002 and operated there until August 2004, when it was relocated at Institut Curie (Section de Recherche, Paris, France).

Early features of Kinefold included the prediction of pseudoknots based on a 'physical' constraint model using polymer theory (17). Pseudoknot predictions were successfully tested on a variety of small pseudoknots (17), as well as on longer sequences such as 394 nt *Tetrahymena* group I intron (19). Predictions on folding dynamics and pathways have also been compared with folding experiments, probing natural (17) and designed DNA or RNA sequences (Viasnoff,V., Meller,A. and Isambert,H., submitted for publication and Xayaphoummine,A., Viasnoff,V., Harlepp,S. and Isambert,H., submitted for publication). In addition, the molecular basis of these stochastic folding simulations has been directly tested with detailed analysis of single-molecule micromechanical unfolding/refolding experiments performed on small RNA motifs and on 1524 nt 16S ribosomal RNA of *Escherichia coli* (20).

New features of Kinefold include DNA sequence folding, forced helices and topologically 'entangled' helices (i.e. knots), which have been recently added as simulation parameters.

OVERVIEW OF KINEFOLD FEATURES

Stochastic folding simulation of nucleic acids

Kinefold simulates nucleic acid folding paths at the level of nucleation and dissociation of RNA/DNA helix regions (12,17) (minimum 3 bp, maximum 60 bp), including pseudoknots and topologically 'entangled' helices. These stochastic formation and the removal of individual helices are known to

*To whom correspondence should be addressed. Tel: +33 1 42 34 64 74; Email: herve.isambert@curie.fr

be the time-limiting steps of RNA/DNA folding kinetics from the seminal experimental results (21,22). Single-strand diffusion and individual base-pair stacking/unstacking processes, which have much faster intrinsic dynamics (0.1 μs) than the relevant time scale for nucleic acid folding and unfolding (>10 μs), are assumed to be in a state of quasistatic equilibrium with the rare stochastic transitions between visited secondary structures. The folding path then consists of a discrete series of secondary structures related through successive additions or removals of single helices. Yet, it happens frequently that the trajectory will oscillate repeatedly between a few structures, and only much more slowly transit to visit new configurations. To circumvent these local kinetic traps that can severely impact the efficiency of the code, folding paths are in fact continuously averaged at the level of rapidly exchanging structures using the 'exactly clustered stochastic simulation' approach detailed previously (19).

Overall, Kinefold can simulate physical folding paths over second to minute molecular time scales for nucleic acid sequences up to ~300–400 bases.

Unlike traditional RNA/DNA folding algorithms (1–6), stochastic simulations do not aim, a priori, at finding lowest free energy structures. In practice, however, they usually find those minima rather quickly, provided that the simulated molecular time is long enough. However, this relaxation time may vary significantly from sequence to sequence, even for similar sequence length and G+C-content, and users are advised to run a few independent folding simulations for each sequence submitted to the web server (see Simulated molecular time and Random seed subsections).

Ultimately, long sequences may not have enough time to reach their lowest free energy structures on physiological time scales (typically minutes) (23) and co-transcriptional folding pathways (24) likely to control proper folding into native structures in this case (Xayaphoummine,A., Viasnoff,V., Harlepp,S. and Isambert,H., submitted for publication). Besides, there are indications that co-transcriptional folding pathways of functional RNAs might be encoded into transient non-native helices (17,25). Kinefold stochastic simulations provide an effective approach to predict co-transcriptional folding paths of functional RNAs either into their lowest energy structures or into functional or misfolded kinetic traps.

Pseudoknots and topologically entangled helices

Besides addressing specific issues about folding kinetics, simulating RNA/DNA folding pathways also allows for the formation of helices interior to loops, the so-called 'pseudoknots' (13,17,19). We found that neglecting pseudoknots typically induces extended structural changes in secondary structure predictions (far beyond the missing pseudoknots themselves) owing to the strong cooperativity of RNA structure rearrangements (see discussion below and Supplementary Material 1).

Owing to their non-nested nature, pseudoknots are 'intrinsically non-local in sequence space' (i.e. crossing arcs) and difficult to include in secondary structure predictions (13,17,19,26–31). Yet, the 3D 'physical' constraints they cause to nucleic acid structures can be 'locally modeled in structure space' within substructures called 'nets' (17). The entropic stretching of single strands and the correlation of helix orientation within each elementary net are analytically calculated by modeling RNA/DNA helices as stiff rods and single-stranded regions as polymer springs. This is also complemented with a more global conformational entropy contribution at larger length scale (no entropic stretching of single strands at this scale) (17).

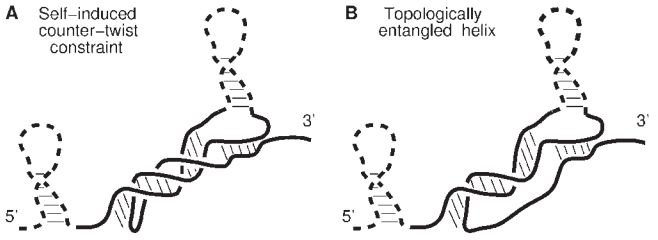

Figure 1. Helical topological constraints caused by pseudoknots. (**A**) Self-induced counter-twist constraints restrict the entropy of stretched single strands, hence resulting in an effective destabilization of long pseudoknots. (**B**) Long, topologically entangled helices leave a residual 'tangle' (32) behind when unpairing from the structure. This results in their effective stabilization until all pseudoknots blocking the helical entanglement are simultaneously removed.

Beyond these geometric constraints, the formation of a pseudoknot may also induce helical topological constraints both on itself (Figure 1A) and on the 'entanglement state' of other non-nested helices in the surrounding (Figure 1B).

(i) The first topological constraint (Figure 1A) arises because a pseudoknot of n helical turns requires a compensating twist of $-n$ turns on the coil sections. While most helices are shorter than a full turn (10–11 bp), longer pseudoknots are 'topologically destabilized' by this self-induced entropic constraint. Hence, counter-twisting topological constraints associated with long pseudoknots do not create new or enhanced transition barriers for the nucleic acid folding/unfolding dynamics. In practice, the local contribution of self-induced topological constraints to global Boltzman weight are estimated at the level of 'nets' (17) with a longer apparent stretching distance on the single-strand twisting around the pseudoknot (see below).

(ii) Conversely, the effect of pseudoknots on the 'entanglement state' of other non-nested helices in their surrounding can be much more dramatic and lead to new kinetic traps for topologically 'entangled' helices (Figure 1B). Entangled helices are defined as helices longer than half a helix turn (>6 bp) whose unpairing transition is topologically impeded by the presence of at least one pseudoknot; when an entangled helix unpairs from an RNA/DNA structure, the newly unpaired bases cannot freely separate due to the residual 'tangle' [after Conway's classification of knots (32)] left behind by the unpaired helix. Such helical entanglement cannot be generally resolved until all the pseudoknots blocking its release through free rotation are simultaneously removed, hence effectively leaving entangled helices kinetically trapped. Entanglements can prevent, in particular, pseudoknots and long adjacent helices from rearranging during co-transcriptional folding owing to topologically blocked transition intermediates. This may cause long-lived metastable structures of untranslated region (UTR) RNA switches regulating the expression of downstream genes at the level of either transcription (e.g. control of early termination/antitermination) or

translation (e.g. control of Shine–Dalgarno motif sequestration). Kinefold represents base pairs of entangled helices with the same color as the phosphate-ribose backbone to emphasize the apparent strength of these bonds.

The effective helix length l for the calculation of the conformational entropy of 'nets' (Figure 2) depends on individual 'entanglement states',

$$l = \left[d^2 \sin^2\left(\pi n_s/n_p\right) + h^2\left(n_s/n_p\right)^2\right]^{1/2},$$

for short ($n_s \leqslant 6$ bp) or 'entangled' helices, and

$$l = n_s/6\left[d^2 \sin^2\left(\pi 6/n_p\right) + h^2\left(6/n_p\right)^2\right]^{1/2},$$

for long ($n_s > 6$ bp) pseudoknots, where n_s is the number of RNA (resp. DNA) base pairs, $d = 3.3a$ (resp. $2.7a$) is the apparent helix diameter (in base unit $a = 6$ Å), $n_p = 11$ (resp. 10) is the number of base pairs per complete turn, and $h = 5a$ (resp. $5.6a$) is the stem length for one turn (33).

Stochastic folding simulations are restricted to all the RNA/DNA structures that can be locally decomposed into nets with no more than two internal helices of respective length l_1 and l_2. Their conformational entropy and all entropic transition barriers to nucleate a new helix within such net are found to follow a single general form,

$$e^{S/k} = \frac{e^{-A_1 l_1^2 - A_2 l_2^2}}{D^{3/2}} \times \frac{e^{2A_3 l_1 l_2} - e^{-2A_3 l_1 l_2}}{4A_3 l_1 l_2}.$$

A multiplicative factor α, corresponding to the 'confinement' cost each time a loop is formed, is added for each helix on the structure [$\alpha = 0.0068$ has been tuned to best fit the tabulated thermodynamic parameters for short loops (34,35)]. The parameters A_1, A_2, A_3 and D depend on the single-strand lengths ($s_1, s_2, \ldots, s_k$), the net class and for transition barriers, on the actual transition within the net class. The expressions of A_1, A_2, A_3 and D are given in Table 1 for all the included nets, together with the most complex transition barrier for which we also provide a scanned analytic calculus in Supplementary Material 2. Conformational entropies and other transition barriers can be deduced from the later result taking simple limits.

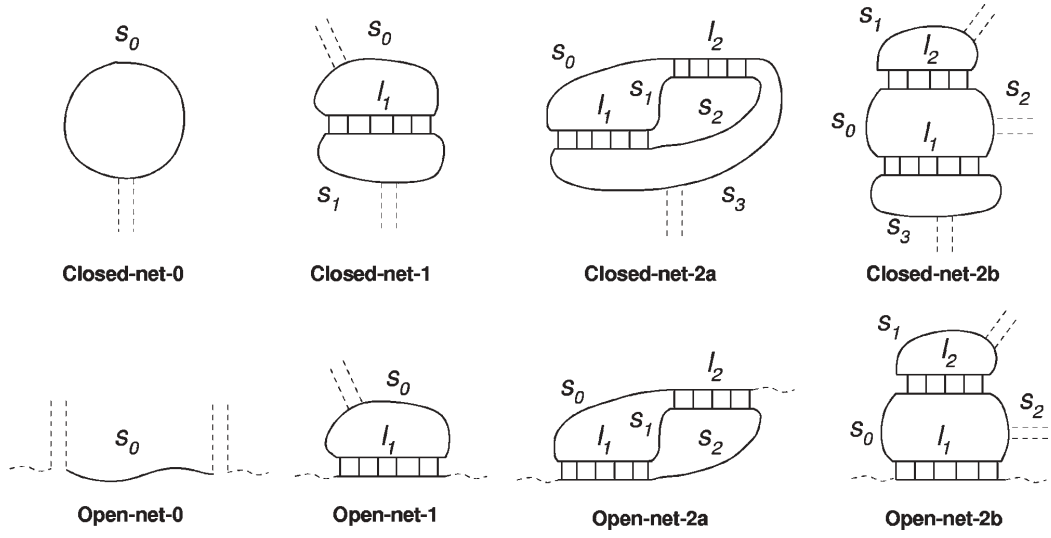

Figure 2. The eight allowed 'nets' enclosing up to two internal helices [see (17) for a detailed description of nets]. Single strands s_0, s_1, etc. are stretched by internal helices whose apparent lengths l_1 and l_2 depend on the 'entanglement state' of each helix, see main text. Conformational entropy parameters for all allowed nets and an example of transition barrier (see Supplementary Material 2) are given in Table 1.

Table 1. Conformational entropy parameters of allowed nets and transition barriers. $\beta = 3/2ab$, with Kuhn length $b = 2.5a$ and $a = 6$ Å

Class of net	D	A_1	A_2	A_3
Open-net-0	1	0	0	0
Open-net-1	s_0	β/D	0	0
Open-net-2a	$s_0s_1 + s_0s_2 + s_1s_2$	$\beta(s_1 + s_2)/D$	$\beta(s_0 + s_1)/D$	$\beta s_1/D$
Open-net-2b	$s_1(s_0 + s_2)$	$\beta s_1/D$	$\beta(s_0 + s_1 + s_2)/D$	$\beta s_1/D$
Closed-net-0	s_0	0	0	0
Closed-net-1	s_0s_1	$\beta(s_0 + s_1)/D$	0	0
Closed-net-2a	$s_0s_3(s_1 + s_2) + s_1s_2(s_0 + s_3)$	$\beta(s_0 + s_3)(s_1 + s_2)/D$	$\beta(s_2 + s_3)(s_0 + s_1)/D$	$\beta(s_0s_2 - s_1s_3)/D$
Closed-net-2b	$s_1s_3(s_0 + s_2)$	$\beta(s_0 + s_2 + s_3)s_1/D$	$\beta(s_0 + s_1 + s_2)s_3/D$	$\beta s_1s_3/D$
A transition barrier in closed-net-2ab (see Supplementary Material 2)	$s_0s_3(s_1 + s_4)(s_2 + s_5) + s_1s_4(s_0 + s_3)(s_2 + s_5) + s_2s_5(s_0 + s_3)(s_1 + s_4)$	$\beta[(s_1 + s_4)(s_0s_3 + s_2s_5) + (s_1 + s_4)(s_3s_5 + s_0s_2) + s_1s_4(s_3 + s_0 + s_2 + s_5)]/D$	$\beta[(s_0 + s_3)(s_1s_4 + s_2s_5) + (s_0 + s_3)(s_1s_5 + s_2s_4) + s_0s_3(s_1 + s_4 + s_2 + s_5)]/D$	$\beta[s_1\, s_4(s_0 + s_3) + s_0s_3(s_1 + s_4) + (s_2s_0s_4 + s_5s_3s_1]/D$

INPUT AND OPTIONS

Input, options and output for the web server are documented online by a help page with a direct link for each of the following items.

Sequence

The sequence should be a string of unmodified RNA/DNA bases (A, U/T, G, C) and may also include additional (X) bases which do not pair. X can be used to model modified bases or to investigate hybridization dynamics between two sequences by artificially connecting them with a long linker of Xs. The current length limitation is set to 400 bases for both 'renaturation fold' and 'co-transcriptional fold' simulations but, depending on their G+C-content, long sequences might not have relaxed by the maximum duration of the stochastic simulation (~1 day). Hence, new users are strongly advised to start their query with shorter subsequences (e.g. up to 200 nt) to obtain a first understanding of the potentials and the limitations of Kinefold and its various available options described below.

RNA versus DNA

Stochastic folding simulations are performed either on RNA (default) or on DNA sequences using stacking free energy parameters from the Turner laboratory (34,35) and the SantaLucia laboratory (36), respectively. Unpublished terminal mismatches for DNA are estimated by adding 5′ and 3′ dandling end stackings, while helix–helix stacking is estimated by double counting terminal mismatches. These approximations generate smaller errors on global molecular free energy (a few percents) than the intrinsic error bars on Watson–Crick base pair stacking energies (10–15%).

Type of stochastic simulation

RNA/DNA folding pathways largely depend on folding conditions. Two types of folding scenarios are currently available (both at 37°C with 1 M NaCl and no divalent ions) and are as follows.

Renaturation fold. Folding starts from a completely denatured configuration (high temperature, no helix formed) which is then instantaneously cooled to 37°C. Alternatively, users have the possibility to force some initial helices on the structure (see Forcing helices).

Co-transcriptional fold. Folding proceeds while the sequence is being 'synthetized' (from its 5′ to 3′ ends). The transcription rate can be tuned to fit average transcription speeds of different RNA polymerases: a new base is typically added every 3, 20 or 200 ms with T7 phage, prokaryote or eukaryote RNA polymerase, respectively. The nascent transcript starts to fold before the whole sequence is fully available. Depending on actual sequence, co-transcriptional folding may guide folding or mislead RNA molecules into long-lived metastable configurations. An example of animated co-transcriptional folding path visualized with the Kineviewer Java applet can be found at http://kinefold.curie.fr/cgi-bin/viewer1.pl?file=HDVribozyme.coord.

Simulated molecular time

The simulated molecular time corresponds to the actual time experienced by the RNA/DNA molecule, not the time needed for the simulation to run (which can be much longer).

The web server automatically estimates a statistical folding time for each submitted sequence. However, users can also predefine the total simulated molecular time, which should be between 1 and 10 000 000 ms (i.e. ~3 h). The 'Enforced time' option is very convenient to restrict the simulation time for long sequences; Above 300 bases, a good starting point is to first simulate a few independent folding paths (see Random seed) of 10 000 ms to check for early reproducibility of folding paths before requesting longer runs such as 30 000 or 100 000 ms.

Pseudoknots

They are visualized for convenience as colored single-stranded regions connected by two straight lines corresponding to the first and last base pairs of the pseudoknot. In the actual simulation, however, pseudoknots are locally treated as any regular helices and the 3D conformational entropic constraints they cause are evaluated using a coarse-grained geometric approach based on modeling RNA/DNA helices as stiff rods and single-stranded regions as polymer springs (17) (Figure 2 and Table 1). Pseudoknots are included by default in stochastic folding paths but they can also be excluded altogether by users.

We have performed extensive statistics on pseudoknot prevalence in random and natural RNA sequences [see Supplementary Material 1 and (19)]. While the proportion of pseudoknots among base pairs is typically 10–15% on average, the actual variation between different sequences is quite broad (from 0 up to 30% for G+C-rich molecules). Moreover, we found that excluding the formation of pseudoknots on an RNA structure, where they correspond to ~10% of base pairs, typically induces much more dramatic base pair changes (>30%) owing to the strong cooperativity of secondary structure rearrangements. This has also been independently observed [Table 4 in (37)] by comparing the accuracy of secondary structure predictions without pseudoknots to known secondary structures including a small number of pseudoknots (up to 10%). In fact, we have found that the distribution of base pair identity between low energy structures with and without pseudoknots is essentially 'flat' (see Supplementary Material 1); there is virtually no correlation between predictions including and excluding pseudoknots. This suggests some intrinsic limitations for predicting pseudoknots through iterative approaches starting from non-pseudoknotted structures, unless they are also combined to comparative methods using homologous sequences (10,11).

Entanglements

Kinefold freezes the unpairing transitions of 'entangled' helices (Figure 1) while they remain entangled. Meanwhile terminal zipping/unzipping adjustments still occur and 'detailed balance' is perfectly maintained throughout the simulation (as it should be). This new option is particularly recommended when users suspect 'kinetically trapped alternative structures' with their sequence. For clarity, the base pair connections of entangled helices are colored in green (Figure 3).

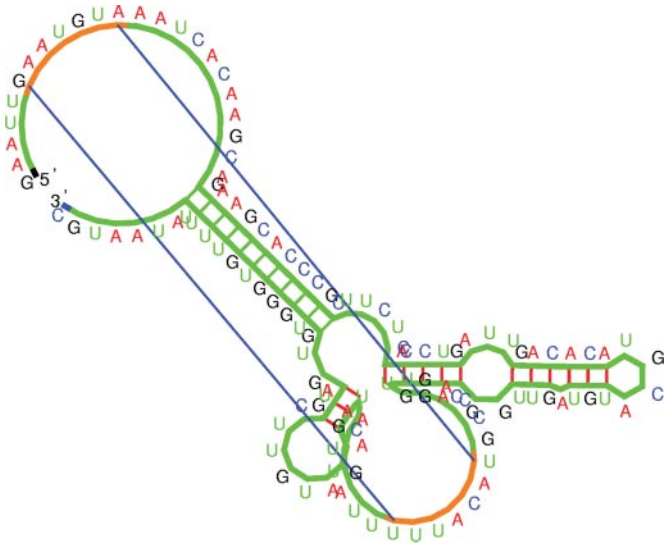

Figure 3. Illustration of predicted structure including a 'pseudoknot' (the helix is visualized by two straight lines corresponding to the first and last base pairs) and an 'entangled helix' (base pairs are drawn with the same color as the phosphate-ribose backbone, see main text). The image is generated by the RNAMovies software upgraded to display pseudoknots and entangled helices (38). The structure shown is a co-transcriptional folding intermediate of the 5′-UTR of the *infC* gene from *Bacillus subtilis*. The RNA Pol at the nascent 3′ end prevents disentanglement of the entangled helix.

Tracing and forcing helices

There can be maximum of five helices (minimum 3 bp, maximum 60 bp).

Tracing helices. Users can follow whether certain helices of special interest to them are formed or not on the structure as the sequence folds. This can be helpful, in particular, to identify transient helices (minimum of three stacking base pairs) that are eventually removed in the native structure. Helix traces are graphically displayed on the trajectory plot.

Enter $T\ i\ j\ k$, with $k > 2$ on one line to trace when any of the consecutive base pairs $i{\cdot}j, i+1{\cdot}j-1, \ldots, i+k-1{\cdot}j-k+1$ is formed.

Forcing helices. In 'renaturation fold' mode, users may also force $k > 2$ consecutive base pairs on the RNA structure (minimum of three stacking base pairs). During folding, the corresponding helices will not be removed, shrunk nor extended by additional matching base pairs, while the rest of the molecule will fold normally. Forced helices, which should be non-overlapping and compatible with the pseudoknot requirements, are also graphically displayed on the trajectory plot. Improper specifications are automatically dismissed.

Enter $F\ i\ j\ k$, with $k > 2$ on one line to force the consecutive base pairs $i{\cdot}j, i+1{\cdot}j-1, \ldots, i+k-1{\cdot}j-k+1$ (and prevent additional extension).

Random seed

Each folding request simulates one single folding pathway. As RNA folding kinetics is a stochastic process during which certain helices are successively formed and removed somewhat randomly (see Type of stochastic simulation subsection), folding pathways may vary significantly from run to run. Hence, before drawing any positive or negative conclusions for a specific sequence, users are advised to try a few runs to look for possible variations. Independent runs are generated with different random seeds, which are integers between 1 and 10 000. One is automatically generated each time the request page is refreshed. Alternatively, it is also possible to choose a particular random seed in order to redo the very same simulation provided that all other parameters remain unchanged too with the exception of the 'simulated molecular time' and 'Traced heliced'.

- In 'renaturation fold' mode, three independent folding pathways leading to the same structure(s) is a good indication (although not a proof) that the lowest free energy structures have been reached. Alternative scenarios may suggest long-lived metastable configurations of potential interest.
- In 'co-transcriptional fold' mode, robust features over a few independent runs (like for instance similar 'trajectory plots' for strong specific helices) are also good indications of well-defined folding pathways either toward a native structure or toward a long-lived kinetically trapped configuration.

OUTPUT

The output page provides (i) a series of *low free energy structures*, (ii) an *online animated folding path* and (iii) a *programmable trajectory plot* focusing on a few helices of interest to each user. Owing to space limitation, these are presented and discussed in Supplementary Material 3.

ACKNOWLEDGEMENTS

We thank A. Condon, D. Evers, C. Lemoine, C. W. Pleij, H. Putzer, J. Robert, M. Springer, F. Thalmann, H. Touzet, S. Tsacas and V. Viasnoff for helpful feedback and discussions. Funding to pay the Open Access publication charges for this article was provided by Fonds National de la Science, Ministere de la Recherche, France.

Conflict of interest statement. None declared.

REFERENCES

1. Waterman,M.S. and Smith,T.F. (1978) RNA secondary structure: a complete mathematical analysis. *Math. Biosci.*, **42**, 257–266.
2. Nussinov,R., Pieczenik,G., Griggs,J.R. and Kleitman,D.J. (1978) Algorithms for loop matchings. *SIAM J. Appl. Math.*, **42**, 68–82.
3. Nussinov,R. and Jacobson,A.B. (1980) Fast algorithm for predicting the secondary structure of single-stranded RNA. *Proc. Natl Acad. Sci. USA*, **77**, 7826–7830.
4. Zuker,M. and Stiegler,P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.*, **9**, 133–148.
5. McCaskill,J.S. (1990) The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*, **29**, 1105–1119.
6. Hofacker,I.L., Fontana,W., Stadler,P.F., Bonhoeffer,L.S., Tacker,M. and Schuster,P. (1994) Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.*, **125**, 167–188.
7. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
8. Hofacker,I. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.*, **31**, 3429–3431.
9. Andronescu,M., Aguirre-Hernandez,R., Condon,A. and Hoos,H.H. (2003) RNAsoft: a suite of RNA secondary structure prediction and design software tools. *Nucleic Acids Res.*, **31**, 3416–3422.

10. Touzet,H. and Perriquet,O. (2004) CARNAC: folding families of related RNAs. *Nucleic Acids Res.*, **32**, 142–145.
11. Ruan,J., Stormo,G. and Zhang,W. (2004) ILM: a web server for predicting RNA secondary structures with pseudoknots. *Nucleic Acids Res.*, **32**, 146–149.
12. Mironov,A.A., Dyakonova,L.P. and Kister,A.E. (1985) A kinetic approach to the prediction of RNA secondary structures. *J. Biomol. Struct. Dyn.*, **2**, 953–962.
13. Gultyaev,A.P. (1991) The computer simulation of RNA folding involving pseudoknot formation. *Nucleic Acids Res.*, **19**, 2489–2494.
14. Gultyaev,A.P., van Batenburg,F.H. and Pleij,C.W. (1995) The influence of a metastable structure in plasmid primer RNA on antisense RNA binding kinetics. *Nucleic Acids Res.*, **23**, 3718–3725.
15. Gultyaev,A.P., van Batenburg,F.H. and Pleij,C.W. (1998) Dynamic competition between alternative structures in viroid RNAs simulated by an RNA folding algorithm. *J. Mol. Biol.*, **276**, 43–55.
16. Nagel,J.H., Gultyaev,A.P., Gerdes,K. and Pleij,C.W. (1999) Metastable structures and refolding kinetics in hok mRNA of plasmid R1. *RNA*, **5**, 1408–1418.
17. Isambert,H. and Siggia,E.D. (2000) Modeling RNA folding paths with pseudoknots: application to hepatitis delta virus ribozyme. *Proc. Natl Acad. Sci. USA*, **97**, 6515–6520.
18. Flamm,C., Fontana,W., Hofacker,I.L. and Schuster,P. (2000) RNA folding at elementary step resolution. *RNA*, **6**, 325–338.
19. Xayaphoummine,A., Bucher,T., Thalmann,F. and Isambert,H. (2003) Prediction and statistics of pseudoknots in RNA structures using exactly clustered stochastic simulations. *Proc. Natl Acad. Sci. USA*, **100**, 15310–15315.
20. Harlepp,S., Marchal,T., Robert,J., Léger,J.-F., Xayaphoummine,A., Isambert,H. and Chatenay,D. (2003) Probing complex RNA structures by mechanical force. *Eur. Phys. J. E Soft Matter*, **12**, 605–615.
21. Craig,M.E., Crothers,D.M. and Doty,P. (1971) Relaxation kinetics of dimer formation by self complementary oligonucleotides. *J. Mol. Biol.*, **62**, 383–401.
22. Pörschke,D. (1974) Thermodynamic and kinetic parameters of an oligonucleotide hairpin helix. *Biophys. Chem.*, **1**, 381–386.
23. Morgan,S. and Higgs,P. (1996) Evidence for kinetic effects in the folding of large RNA molecules. *J. Chem. Phys.*, **105**, 7152–7157.
24. Kramer,F. and Mills,D. (1981) Secondary structure formation during RNA synthesis. *Nucleic Acids Res.*, **9**, 5109–5124.
25. Meyer,I.M. and Miklós,I. (2004) Co-transcriptional folding is encoded within RNA genes. *BMC Mol. Biol.*, **5**, 10.
26. Abrahams,J.P., van den Berg,M., van Batenburg,E. and Pleij,C.W.A. (1990) Prediction of RNA secondary structure, including pseudoknotting, by computer simulation. *Nucleic Acids Res.*, **18**, 3035–3044.
27. Rivas,E. and Eddy,S.R. (1999) A dynamic programming algorithm for RNA structure prediction including pseudoknots. *J. Mol. Biol.*, **285**, 2053–2068.
28. Akutsu,T. (2000) Dynamic programming algorithms for RNA secondary structure prediction with pseudoknots. *Discrete Appl. Math.*, **104**, 45–62.
29. Lyngsø,R. and Pedersen,C. (2000) RNA pseudoknot prediction in energy-based models. *J. Comput. Biol.*, **7**, 409–427.
30. Dirks,R. and Pierce,N.A. (2003) A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J. Comput. Chem.*, **24**, 1664–1677.
31. Ruan,J., Stormo,G. and Zhang,W. (2004) An iterated loop matching approach to the prediction of RNA secondary structures with pseudoknots. *Bioinformatics*, **20**, 58–66.
32. Conway,J.H. (1967) An enumeration of knots and links, and some of their algebraic properties. In Leech,J. (ed.), *Computation Problems in Abstract Algebra*. Pergamon Press, Oxford, pp. 329–358.
33. Bloomfield,V.A., Crothers,D.M. and Tinoco,I.,Jr (2000) *Nucleic Acids: Structures, Properties, and Functions*. University Science books, Sausalito, CA.
34. Walter,A.E., Turner,D.H., Kim,J., Lyttle,M.H., Müller,P., Mathews,D.H. and Zuker,M. (1994) Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl Acad. Sci. USA*, **91**, 9218–9222.
35. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
36. SantaLucia,J.,Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
37. Mathews,D.H., Disney,M.D., Childs,J.L., Schroeder,S.J., Zuker,M. and Turner,D.H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl Acad. Sci. USA*, **101**, 7287–7292.
38. Evers,D. and Giegerich,R. (1999) RNAMovies: visualizing RNA secondary structure spaces. *Bioinformatics*, **15**, 32–37.