MELTING - nearest-neighbor computation of nucleic acid hybridation

Marine Dumousseau, William John Gowers
Nicolas Le Novère
n.lenovere@gmail.com

February 2014

Contents

1 Synopsis

The nearest-neighbor approach is based on the fact that the helix-coil transition works as a zipper. After an initial attachment, the hybridisation propagates laterally. The hybridization process depends on the adjacent nucleotides on each strand (the Crick's pairs). Two duplexes with the same base pairs could have different stabilities, and on the contrary, two duplexes with different sequences but identical sets of Crick's pairs will have the same thermodynamics properties (see Sugimoto et al. 1994). See Wetmur J.G (1991) and Santalucia (1998) for deep reviews on the nucleic acid hybridization and on the different set of nearest-neighbor parameters.

2 Description

MELTING computes, for a nucleic acid duplex, the enthalpy and the entropy of the helix-coil transition, and then its melting temperature. Four types of hybridisation are possible: DNA/DNA, DNA/RNA, RNA/RNA and 2-O-Methyl RNA/RNA. The program uses the method of nearest-neighbors. The set of thermodynamic parameters can be easely changed, for instance following an experimental breakthrough. Melting is a free program in both sense of the term. It comes with no cost and it is open-source. In addition it is coded in Java (1.5) and can be compiled on any operating system.

If you use MELTING, please quote

> Le Novère. MELTING, a free tool to compute the melting temperature of nucleic acid duplex. *Bioinformatics*, 17: 1226-1227.
>
> Dumousseau M, Rodriguez N, Juty N, Le Novère N. MELTING, a flexible platform to predict the melting temperatures of nucleic acids. *BMC Bioinformatics*, 16;13:101, PMID: 22591039.

3 Usage

The options are treated sequentially. If there is a conflict between the value of two options, the latter normally erases the former.

BE AWARE : The option syntax of MELTING 5 is different from the one of MELTING 4. There is a space between the option name and the option value. New option names are available in MELTING 5 to change the default thermodynamic models and default corrections.
There is no interactive mode in MELTING 5 therefore the option '-q' doesn't exist anymore.
You can use the MELTING 4 option syntax, but it doesn't allow the user to change some of the thermodynamic models and corrections. In addition to that, the user can't enter a formamide or DMSO See the README file to choose the adapted executable.
The MELTING 4 option name '-x' is equivalent to the MELTING 5 option name '-am'.
There is no input file option in MELTING 5 (option '-I') but you can use this option

for the compatible executable of MELTING 5.
The MELTING 4 option names '-N', '-t', '-k', 'G' are replaced by the single option '-E' in MELTING 5.
The MELTING 4 option names '-A', '-D', '-M' are respectively equivalent to '-nn', '-sinDE', '-sinMM' in MELTING 5.
The file names to write with the precedent option are replaced by thermodynamic model names (see below).
The MELTING 4 option name '-K' is replaced by '-ion' in MELTING 5.

3.1 Information about MELTING

-h

Displays a short help and quit.

-L

Prints the legal informations and quit.

-V

Displays the version number and quit.

-p

Return the directory supposed to contain the sets of calorimetric parameters and quit. If the environment variable NN_PATH is set, it is returned. Otherwise, the value defined by default during the compilation is returned.

3.2 Mandatory options

-S *sequence*

Sequence of one strand of the nucleic acid duplex, entered 5' to 3'. **Important:** Uridine and thymidine are not considered as identical. The bases can be upper or lowercase.

-C *complementary_sequence*

Enters the complementary sequence, from 3' to 5'. This option is mandatory if there are mismatches, inosine(s) or hydroxyadenine(s) between the two strands. If it is not used, the program will compute it as the complement of the sequence entered with the option **-S**. In case of self complementary sequences, The program can automatically detect the symmetry and deduce the complementary even though there is (are) dangling end(s) and it is not necessary to write the complementary sequence with the option **-C**. Uridine and thymidine are not considered as identical. The bases can be upper or lowercase.

-E *ion1_name=x.xxe-xx:ion2_name=x.xxe-xx:agent1_name=x.xxe-xx...*

Enters the different ion (Na, Mg, Tris, K) or agent (dNTP, DMSO, formamide) concentrations. The effect of ions and denaturing agents on thermodynamic stability of nucleic acid duplexes is complex, and the correcting functions are at

best rough approximations. All the concentrations must be positive numeric values and in M. There are some exceptions for the DMSO concentrations (in %) and the formamide concentrations (in % or M depending on the used correction method). Be aware, the $[Tris^+]$ is about half of the total tris buffer concentration. At least one cation concentration is mandatory, the other agents are optional. See the documentation for the concentration limits. It depends on the used correction.

-P *x.xxe-xx*

Concentration of the nucleic acid strand in excess. It must be a strict positive numeric value and it is mandatory. The oligomer concentration is in mol/L.

-H *hybridisation_type*

Specifies the hybridisation type. Moreover this parameter determines the nature of the sequences entered by the user. Possible values are :

dnadna : DNA sequence (option **-S**) and DNA complementary sequence (option **-C**)

rnarna : RNA sequence (option **-S**) and RNA complementary sequence (option **-C**)

dnarna : DNA sequence (option **-S**) and RNA complementary sequence (option **-C**)

rnadna : RNA sequence (option **-S**) and DNA complementary sequence (option **-C**)

mrnarna : 2-o-methyl RNA sequence (option **-S**) and RNA complementary sequence (option **-C**)

mrnarna : RNA sequence (option **-S**) and 2-o-methyl RNA complementary sequence (option **-C**)

This option is mandatory to select the default equations and methods to use.

3.3 General options

-T *xxx*

Size threshold before approximative computation. The nearest-neighbour approach will be used by default if the length of the sequence is inferior to this threshold, otherwise it is the approximative approach which will be used by default.

-v

Activates the verbose mode, issuing a lot more information about the current run (try it once to see if you can get something interesting).

-nnpath *folder_pathway*

Change the default pathway (Data) where to find the default calorimetric tables (thermodynamic parameters). The program will look for the file in a directory

specified during the installation. However, if an environment variable NN_PATH is defined, melting will search in this one first.

-O *output_file*

The output is directed to this file instead of the standard output. The name of the file must be specified.

-self

To precise that the sequence entered with the option **-S** is self complementary. No complementary sequence is mandatory. The program automatically can detect a self complementary sequence for perfect matching sequences or sequences with dangling ends. In these cases, the option **-self** is not necessary. Otherwise we need to precise that the sequences are self complementary with this option. examples:

```
  Situation 1 : The sequence ATCGCGAT is self
                complementary.
```

The option **-self** is not necessary because the program can automatically detect it.

```
  Situation 2 : The sequence -TCGCGAT is self
                complementary with a single
dangling end.
```

The option **-self** is not necessary because the program can automatically detect it.

```
  Situation 3 :  If the sequence ATCCCGAT is self
                 complementary with a single mismatch
(C/C)
```

The option **-self** is necessary to precise the self complementarity because the program can't detect it.

-F *factor*

This is the correction factor used to modulate the effect of the nucleic acid concentration in the computation of the melting temperature. See section ALGORITHM for details. If the sequences are automatically recognized as self complementary sequences or if the option **-self** is used, the factor correction is automatically 1. Otherwise F is 4 if the both strands are present in equivalent amount and 1 if one strand is in excess. The default factor value is 4.

3.4 Set of thermodynamic parameters and methods (models)

By default, the approximative mode is used for oligonucleotides longer than 60 bases (the default threshold value), otherwise the nearest neighbor model is used.

-am *method_name*

Forces to use a specific approximative formula, based on G+C content. You can use one of the following :

DNA DUPLEXES

ahs01 (from von Ahsen et al. 2001)

che93 (from Marmur 1962,, Chester and al. 1993)

che93corr (from von Ahsen et al. 2001, Marmur 1962, Chester and al. 1993)

schdot (Marmur-Schildkraut-Doty formula)

owe69 (from Owen et al. 1969)

san98 (from Allawi and Santalucia. 1998)

wetdna91 (from Wetmur 1991) (by default)

RNA DUPLEXES

wetrna91 (from Wetmur 1991) (by default)

DNA/RNA DUPLEXES

wetdnarna91 (from Wetmur 1991) (by default)

If there is no formula name after the option **-am**, we will compute the melting temperature with the default approximative formula. This option has to be used with caution. Note that such a calcul is increasingly incorrect when the length of the duplex decreases. Moreover, it does not take into account nucleic acid concentration, which is a strong mistake. examples :

```
command line 1 : "-am"
```

if you want to force the approximative approach with the default formula.

```
command line 2 : "-am ahs01"
```

if you want to use the approximative formula from Ahsen et al. 2001.

-nn *method_name*

Forces to use a specific nearest neighbor model. You can use one of the following :

DNA DUPLEXES

all97 (from Allawi and Santalucia 1997) (by default)

bre86 (from Breslauer et al. 1986)

san04 (from Hicks and Santalucia 2004)

san96 (from Santalucia et al. 1996)

sug96 (from Sugimoto et al 1996)

tan04 (from Tanaka et al. 2004)

RNA DUPLEXES

fre86 (from Freier al. 1986)

xia98 (from Xia et al. 1998) (by default)

DNA/RNA DUPLEXES

sug95 (from Sugimoto et al. 1995) (by default)

MRNA/RNA DUPLEXES

tur06 (from Kierzek et al. 2006) (by default)

If there is no formula name after the option **-nn**, we will compute the melting temperature with the default nearest neighbor model. Each nearest neighbor model uses a specific xml file containing the thermodynamic values. If you want to use another file, write the file name or the file pathway preceded by ':' (-nn [optionalname:optionalfile]). examples:

```
Command line 1 : "-nn"
```

if you want to force the nearest neighbor computation with the default model.

```
Command line 2 : "-nn tan04"
```

if you want to use the nearest neighbor model from Tanaka et al. 2004 with the thermodynamic parameters in the default xml file.

```
Command line 3 : "-nn tan04:fileName"
```

if you want to use the nearest neighbor model from Tanaka et al. 2004 with the thermodynamic parameters in the file fileName.

```
Command line 4 : "-nn :fileName"
```

if you want to use the default nearest neighbor model with the thermodynamic parameters in the file fileName.

-sinMM *method_name*

Forces to use a specific nearest neighbor model to compute the contribution of single mismatch to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

allsanpey (from Allawi, Santalucia and Peyret 1997, 1998 and 1999) (by default)

DNA/RNA DUPLEXES

wat10 (from Watkins et al. 2011) (by default)

RNA DUPLEXES

tur06 (from Lu et al. 2006)

zno07 (from Davis et al. 2007) (by default)

zno08 (from Davis et al. 2008)

To change the file containing the thermodynamic parameters for single mismatch computation, the same syntax as the one for the **-nn** option is used. Single mismatches are not taken into account by the approximative mode.

-GU *method_name*

Forces to use a specific nearest neighbor model to compute the contribution of GU base pairs to the thermodynamic of helix-coil transition. You can use one of the following :

RNA DUPLEXES

tur99 (from Mathews et al. 1999)

ser12 (from Serra et al. 2012) (by default)

To change the file containing the thermodynamic parameters for GU base pair computation, the same syntax as the one for the **-nn** option is used. GU base pairs are not taken into account by the approximative mode.

-tanMM *method_name*

Forces to use a specific nearest neighbor model to compute the contribution of tandem mismatches to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

allsanpey (from Allawi, Santalucia and Peyret 1997, 1998 and 1999) (by default)

RNA DUPLEXES

tur99 (from Mathews et al. 1999) (by default)

To change the file containing the thermodynamic parameters for tandem mismatch computation, the same syntax as the one for the **-nn** option is used. Tandem mismatches are not taken into account by the approximative mode. Note that not all the mismatched Crick's pairs have been investigated.

-intLP ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of internal loop to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

san04 (from Hicks and Santalucia 2004) (by default)

RNA DUPLEXES

tur06 (from Lu et al. 2006) (by default)

zno07 (from Badhwar et al. 2007, only for 1x2 loop)

To change the file containing the thermodynamic parameters for internal loop computation, the same syntax as the one for the **-nn** option is used. Internal loops are not taken into account by the approximative mode.

-sinDE ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of single dangling end to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

bom00 (from Bommarito et al. 2000) (by default)

sugdna02 (from Ohmichi et al. 2002, only for polyA dangling ends)

RNA DUPLEXES

sugrna02 (from Ohmichi et al. 2002, only for polyA dangling ends)

ser08 (from Miller et al. 2008) (by default)

To change the file containing the thermodynamic parameters for single dangling end computation, the same syntax as the one for the **-nn** option is used. Single dangling ends are not taken into account by the approximative mode.

-secDE ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of double dangling end to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

sugdna02 (from Ohmichi et al. 2002, only for polyA dangling ends) (by default)

RNA DUPLEXES

sugrna02 (from Ohmichi et al. 2002, only for polyA dangling ends)

ser05 (from O'toole et al. 2005)

ser06 (from O'toole et al. 2006) (by default)

To change the file containing the thermodynamic parameters for double dangling end computation, the same syntax as the one for the **-nn** option is used. Double dangling ends are not taken into account by the approximative mode.

-longDE ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of long dangling end to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

sugdna02 (from Ohmichi et al. 2002, only for polyA dangling ends) (by default)

RNA DUPLEXES

sugrna02 (from Ohmichi et al. 2002, only for polyA dangling ends)

To change the file containing the thermodynamic parameters for long dangling end computation, the same syntax as the one for the **-nn** option is used. Long dangling ends are not taken into account by the approximative mode.

-sinBU ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of single bulge loop to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

san04 (from Santalucia 2004)

tan04 (from Tanaka et al. 2004) (by default)

RNA DUPLEXES

ser07 (from Blose et al. 2007)

tur06 (from Lu et al. 1999 and 2006) (by default)

To change the file containing the thermodynamic parameters for single bulge loop computation, the same syntax as the one for the **-nn** option is used. Single bulge loops are not taken into account by the approximative mode.

-lonBU ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of long bulge loop to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

san04 (from Hicks and Santalucia 2004) (by default)

RNA DUPLEXES

tur06 (from Mathews et al. 1999 and Lu et al 2006) (by default)

To change the file containing the thermodynamic parameters for long bulge loop computation, the same syntax as the one for the **-nn** option is used. Long bulge loops are not taken into account by the approximative mode.

-CNG ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of CNG repeats to the thermodynamic of helix-coil transition. N represents a single mismatch of type N/N. You can use one of the following :

RNA DUPLEXES

bro05 (from Magdalena et al. 2005) (by default)

To change the file containing the thermodynamic parameters for CNG repeats computation, the same syntax as the one for the **-nn** option is used. CNG repeats are not taken into account by the approximative mode. Be aware : Melting can compute the contribution of CNG repeats to the thermodynamic of helix-coil transition for only 2 to 7 CNG repeats.

-ino ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of inosine bases (I) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

san05 (from Watkins and Santalucia 2005) (by default)

RNA DUPLEXES

zno07 (from Wright et al. 2007, only IU base pairs) (by default)

To change the file containing the thermodynamic parameters for inosine bases computation, the same syntax as the one for the **-nn** option is used. Inosine bases (I) are not taken into account by the approximative mode.

-ha ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of hydroxyadenine bases (A*) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

sug01 (from Kawakami et al. 2001) (by default)

To change the file containing the thermodynamic parameters for hydroxyadenine bases computation, the same syntax as the one for the **-nn** option is used. Hydroxyadenine bases (A*) are not taken into account by the approximative mode.

-azo ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of azobenzenes (X_T for trans azobenzenes and X_C for cis azobenzenes) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

asa05 (from Asanuma et al. 2005)(by default)

To change the file containing the thermodynamic parameters for azobenzene computation, the same syntax as the one for the **-nn** option is used. Azobenzenes (X_T for trans azobenzenes and X_C for cis azobenzenes) are not taken into account by the approximative mode.

-lck ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of single locked nucleic acids (AL, GL, TL and CL) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

mct04 (from McTigue et al. 2004)

owc11 (from Owczarzy et al. 2011) (by default)

To change the file containing the thermodynamic parameters for single locked nucleic acids computation, the same syntax as the one for the **-nn** option is used. Locked nucleic acids (AL, GL, TL and CL) are not taken into account by the approximative mode.

-tanLck ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of consecutive locked nucleic acids (AL, GL, TL and CL) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

owc11 (from Owczarzy et al. 2011) (by default)

To change the file containing the thermodynamic parameters for consecutive locked nucleic acids computation, the same syntax as the one for the **-nn** option is used. Locked nucleic acids (AL, GL, TL and CL) are not taken into account by the approximative mode.

-sinMMLck ***method_name***

Forces to use a specific nearest neighbor model to compute the contribution of single mismatch in consecutive locked nucleic acids (AL, GL, TL and CL) to the thermodynamic of helix-coil transition. You can use one of the following :

DNA DUPLEXES

owc11 (from Owczarzy et al. 2011) (by default)

To change the file containing the thermodynamic parameters for single mismatch in consecutive locked nucleic acids computation, the same syntax as the one for the **-nn** option is used. Locked nucleic acids (AL, GL, TL and CL) are not taken into account by the approximative mode.

-ion ***method_name***

Forces to use a specific ion correction. You can use one of the following corrections :

Sodium corrections

DNA DUPLEXES

ahs01 (from von Ahsen et al. 2001)

kam71 (from Frank-Kamenetskii 2001)

owc1904 (equation 19 from Owczarzy et al. 2004)

owc2004 (equation 20 from Owczarzy et al. 2004)

owc2104 (equation 21 from Owczarzy et al. 2004)

owc2204 (equation 21 from Owczarzy et al. 2004) (by default)

san96 (from Santalucia et al. 1996)

san04 (from Santalucia et al. 1998, 2004)

schlif (from Schildkraut and Lifson 1965)

tanna06 (from Tan et al. 2006)

wetdna91 (from wetmur 1991)

RNA DUPLEXES OR MRNA/RNA DUPLEXES

tanna07 (from Tan et al. 2007) (by default)

wetrna91 (from wetmur 1991)

DNA/RNA DUPLEXES

wetdnarna91 (from wetmur 1991)

Magnesium corrections

DNA DUPLEXES

owcmg08 (from Owczarzy et al. 2008) (by default)

tanmg06 (from Tan et al. 2006)

RNA DUPLEXES OR MRNA/RNA DUPLEXES

tanmg07 (from Tan et al. 2007) (by default)

Mixed Na Mg corrections

DNA DUPLEXES

owcmix08 (from Owczarzy et al. 2008) (by default)

tanmix07 (from Tan et al. 2007)

RNA DUPLEXES OR MRNA/RNA DUPLEXES

tanmix07 (from Tan et al. 2007) (by default)

The effect of ions on thermodynamic stability of nucleic acid duplexes is complex, and the correcting functions are at best rough approximations. By default, the program use the algorithm from Owczarzy et al 2008 : ratio = $[Mg^{0.5}]$ and monovalent = Na + Tris + K

if monovalent = 0, a magnesium correction is used.

if ratio < 0.22, a sodium correction is used.

if 0.22 <= ratio < 6, a mixed Na Mg correction is used.

if ratio >= 6, a magnesium correction is used.

example :

```
Command line : "-ion owcmg08"
```

if you want to force the use of the magnesium correction from Owczarzy et al 2008. This correction will be used independently of the cations present in the solution.

-naeq ***method_name***

Forces to use a specific ion correction which gives a sodium equivalent concentration if other cations are present. You can use one of the following :

DNA DUPLEXES

ahs01 (from von Ahsen et al 2001) (by default)

mit96 (from Mitsuhashi et al. 1996)

pey00 (from Peyret 2000)

For the other types of hybridization, the DNA default correction is used but there is no guaranty of accuracy. If there are other cations when an approximative approach is used, a sodium equivalence is automatically computed. The correcting functions are at best rough approximations. example :

```
Command line 1 : "-naeq ahs01"
```

if you want to force the use of the sodium equivalence from Ahsen et al 2001. This sodium equivalence will be used in case of approximative approach. In case of nearest neighbor approach, the sodium equivalence will be used only if a sodium correction is selected by the user.

```
Command line 2 : "-naeq ahs01 -ion san04"
```

it means that the sodium equivalence computed by the method ahs01 (from Ahsen et al 2001) will be combined with the sodium correction san04 (from Santalucia 2004).

-DMSO *method_name*

Forces to use a specific DMSO correction (DMSO is always in %). You can use one of the following :

DNA DUPLEXES

ahs01 (from von Ahsen et al 2001) (by default)

mus81 (from Musielski et al. 1981)

cul76 (from Cullen et al. 1976)

esc80 (from Escara et al. 1980)

For the other types of hybridization, the DNA default correction is used but there is no guaranty of accuracy. If there are DMSO when an approximative approach is used, a DMSO correction is automatically computed. The correcting functions are at best rough approximations. example :

```
Command line : "-DMSO ahs01"
```

if you want to force the use of the DMSO correction from Ahsen et al 2001. This DMSO correction will be used if there is DMSO present in the solutions in case of nearest neighbor approach and approximative approach.

-for *method_name*

Forces to use a specific formamide correction. You can use one of the following :

DNA DUPLEXES

bla96 (from Blake 1996) with formamide concentration in mol/L (by default)

lincorr (linear correction) with a % of formamide volume

For the other types of hybridization, the DNA default correction is used but there is no guaranty of accuracy. If there are formamide when an approximative approach is used, a formamide correction is automatically computed. The correcting functions are at best rough approximations. example :

```
Command line : "-for lincorr"
```

if you want to force the use of the linear formamide correction. This formamide correction will be used if there is formamide present in the solutions in case of nearest neighbor approach and approximative approach.

4 Algorithm

4.1 Thermodynamics of helix-coil transition of nucleic acid

The nearest-neighbor approach is based on the fact that the helix-coil transition works as a zipper. After an initial attachment, the hybridisation propagates laterally. This program first computes the hybridisation enthalpy and entropy for each structure in the duplex. (see later for the different possible structures recognized by Melting). If the sequences are self complementary, a symmetry correction will be added to the initiation energy.

$$\begin{aligned} \Delta H &= \delta h_{\text{initiation}} + \sum \delta h_{\text{structure}} \\ \Delta S &= \delta s_{\text{initiation}} + \sum \delta s_{\text{structure}} \end{aligned}$$

Example :
Sequence with a single mismatch

```
ATCGGCTA
TAGACGAT
```

$$\begin{aligned} \Delta H &= \delta h_{\text{initiation}} + \delta h_{\text{structure1}} + \delta h_{\text{structure2}} + \delta h_{\text{structure3}} \\ \Delta S &= \delta s_{\text{initiation}} + \delta s_{\text{structure1}} + \delta s_{\text{structure2}} + \delta s_{\text{structure3}} \end{aligned}$$

where :
structure1 = perfectly matching sequences ATC/TAG
structure2 = single mismatch G/A
structure 3 = perfectly matching sequences GCTA/CGAT

4.1.1 Perfectly matching sequences

The hybridization process depends on the adjacent nucleotides on each strand (the Crick's pairs). Two duplexes with the same base pairs could have different stabilities, and on the contrary, two duplexes with different sequences but identical sets of Crick's pairs will have the same thermodynamics properties. This program first computes the hybridisation enthalpy and entropy from the elementary parameters of each Crick's pair.

$$\begin{aligned} \Delta h_{\text{perfectly}-\text{matching}} &= \sum \delta h_{\text{Crick}'\text{spair}} \\ \Delta s_{\text{perfectly}-\text{matching}} &= \sum \delta s_{\text{Crick}'\text{spair}} \end{aligned}$$

The initiation computation is not the same for each following model.

model	limits	Article
all97	DNA	Allawi and SantaLucia (1997) Biochemistry 36 : 10581-10594
bre86	DNA	Breslauer et al. (1986) Proc Natl Acad Sci USA 83 : 3746-3750
san04	DNA	Santalucia and Hicks (2004) Annu. Rev. Biophys. Biomol. Struct 33 : 415-440
san96	DNA	SantaLucia et al.(1996) Biochemistry 35 : 3555-3562
sug96	DNA	Sugimoto et al. (1996) Nuc Acids Res 24 : 4501-4505
tan04	DNA	Tanaka et al (2004) Biochemistry 43 : 7143-7150
fre86	RNA	Freier et al (1986) Proc Natl Acad Sci USA 83: 9373-9377
xia98	RNA	Xia et al (1998) Biochemistry 37: 14719-14735
sug95	DNA/RNA	SantaLucia et al.(1996) Biochemistry 35 : 3555-3562
tur06	mRNA/RNA A sodium correction (san04) is automatically applied to the computed entropy to convert the entropy (Na = 0.1M) into the entropy (Na=1M)	Kierzeck et al (2006) Nucleic acids research 34: 3609-3614

Example :

$$\Delta H\begin{pmatrix}\texttt{AGCGA}\\ \texttt{TCGCT}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{AG}\\ \texttt{TC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{GC}\\ \texttt{CG}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{CG}\\ \texttt{GC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{GA}\\ \texttt{CT}\end{pmatrix}$$

(The same computation is performed for ΔS)

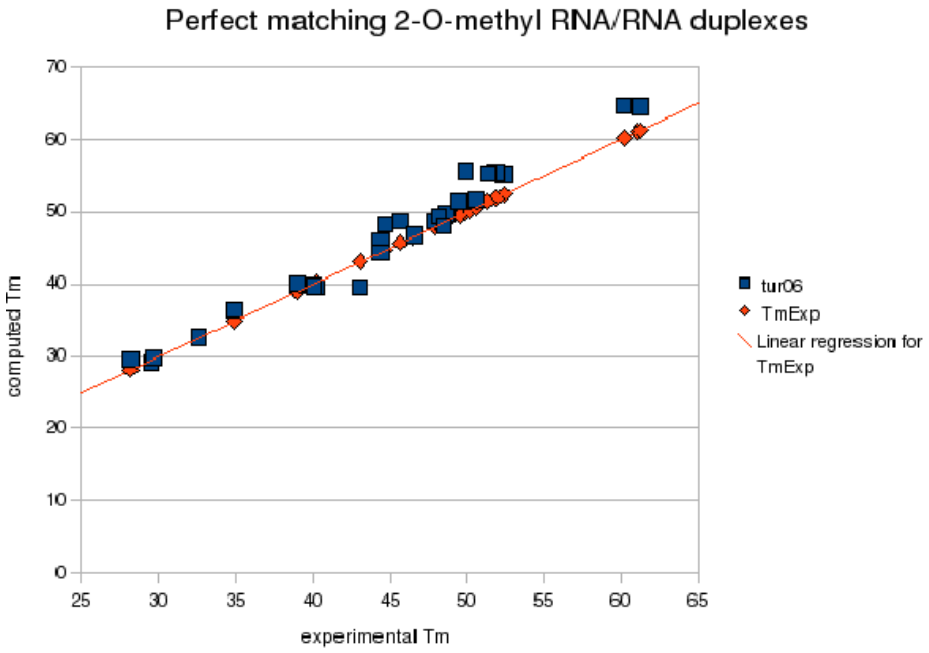

Figure 1: Comparison of experimental and computed Tm for various sets of DNA nearest-neighbor parameters. $[Na^+] = 1$ M, [nucleic acid] $= 4 \cdot 10^{-4}$ M

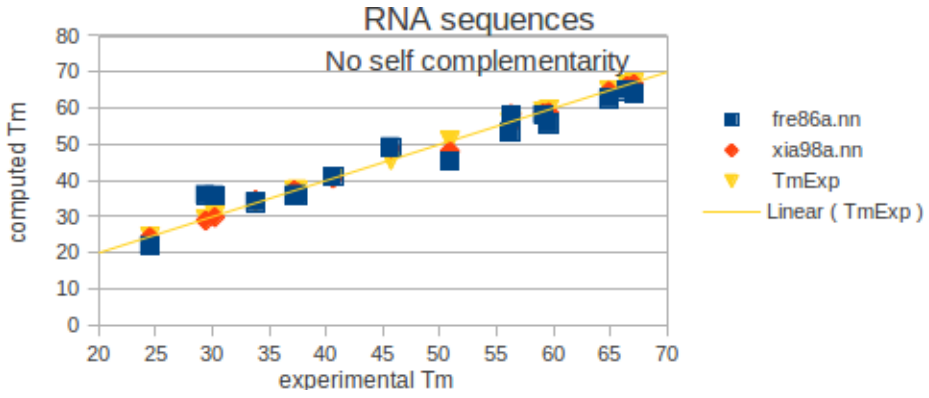

Figure 2: Comparison of experimental and computed Tm for various sets of RNA nearest-neighbor parameters. $[Na^+] = 1$ M, [nucleic acid] $= 2 \cdot 10^{-4}$ M

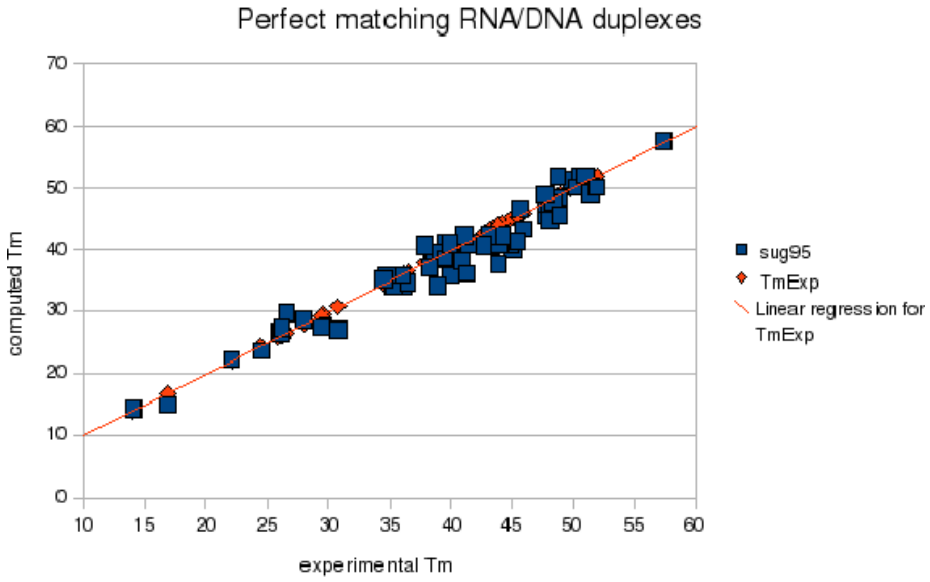

Figure 3: Comparison of experimental and computed Tm for various sets of DNA/RNA nearest-neighbor parameters. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

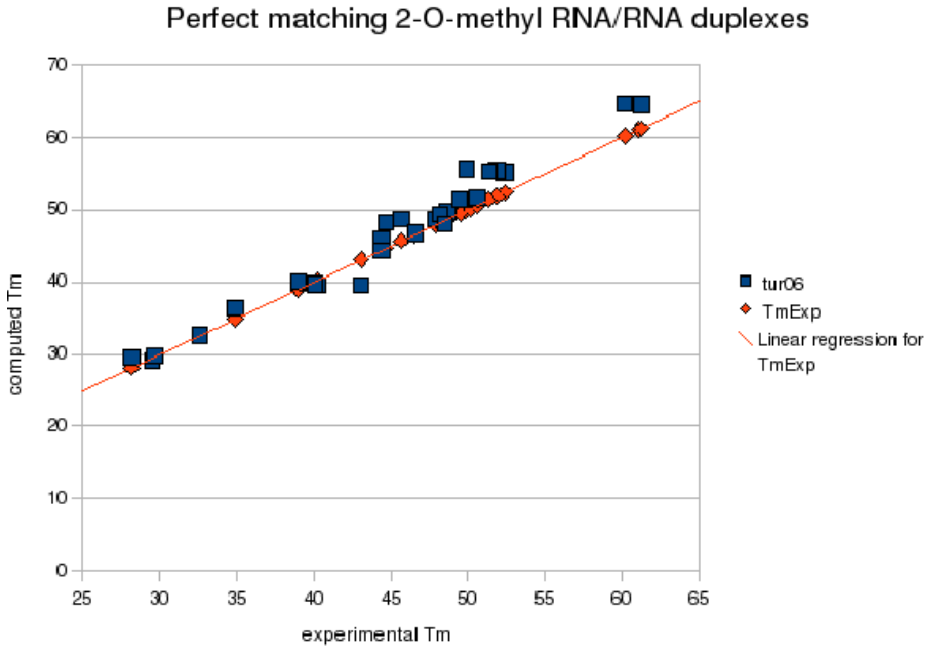

Figure 4: Comparison of experimental and computed Tm for various sets of 2-O-methyl RNA nearest-neighbor parameters. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.2 Sequences composed of CNG repeats

If the sequence (sens 5'3') is a sequence of type `G(CNG)xC` where x is the number of CNG repeats in the sequence and N a unique nucleic acid which will get bound to itself, we can use specific experimental parameters to compute the enthalpy and entropy of the duplex formation. These parameters can be used only for sequences composed from 2 to 7 CNG repeats and the initiation is already included.

$$\Delta H \quad = \quad \Delta h_{\text{sequence-of-type-G(CNG)xC}}$$

For further information, see the referenced article.

model	limits	Article
bro05	RNA Self complementary sequences 2 to 7 CNG repeats	Magdalena et al (2005) Biochemistry 44: 10873-10882

Example :

```
GCAGCAGCAGC
CGACGACGACG
```

$$\Delta H\begin{pmatrix}\texttt{GCAGCAGCAGC}\\ \texttt{CGACGACGACG}\end{pmatrix} \qquad = \qquad \Delta H(\text{3-CAG-repeats})$$

(The same computation is performed for ΔS)

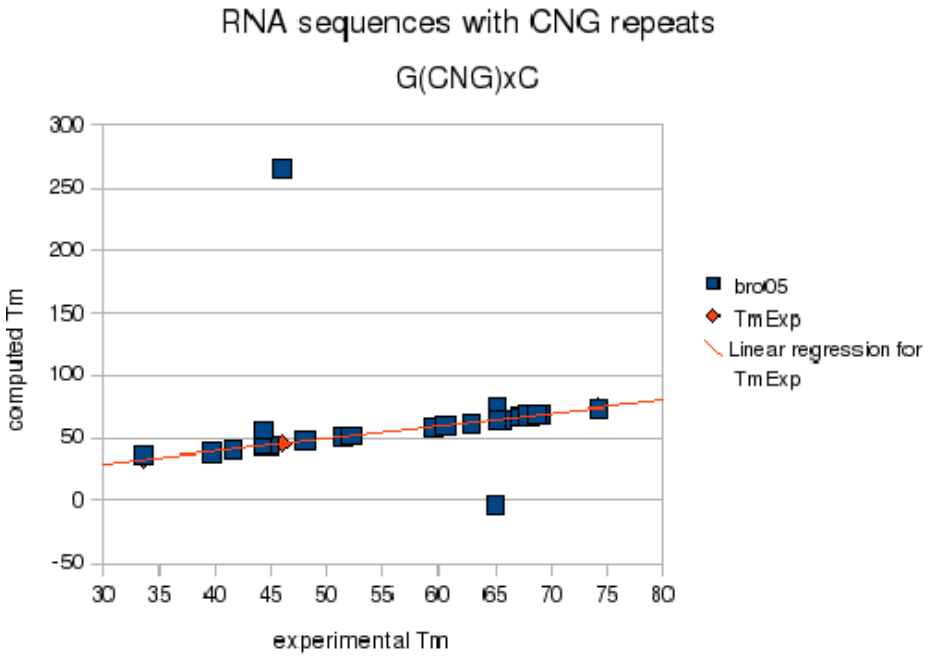

Figure 5: Comparison of experimental and computed Tm for various sets of RNA sequences composed of CNG repeats. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

Be aware : The results for sequences composed of 4 or 5 CCG repeats is not reliable. (the figure shows two values far from the expected temperature). This might be due to a majority of hairpin loop formation. See the article above for further informations.

4.1.3 Single mismatch effect

The single mismatches are taken into account but the two first and positions cannot be mismatched. in such a case, the result is unpredictable, and all cases are possible. for instance (see Allawi and SanLucia 1997), the duplex

```
A           T
 GTGAGCTCAT
 TACTCGAGTG
T           A
```

is more stable than

```
AGTGAGCTCATT
TTACTCGAGTGA
```

For DNA duplexes, this program computes the hybridisation enthalpy and entropy from the elementary parameters of each Crick's pair containing the single mismatch.

$$\Delta h_{\text{single-mismatch}} \quad = \quad \sum \delta h_{\text{Crick's-pair-containing-the-mismatch}}$$

Example :

$$\Delta H\begin{pmatrix}\texttt{ATC}\\ \texttt{TCG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{AT}\\ \texttt{TC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{TC}\\ \texttt{CG}\end{pmatrix}$$

(The same computation is performed for ΔS)

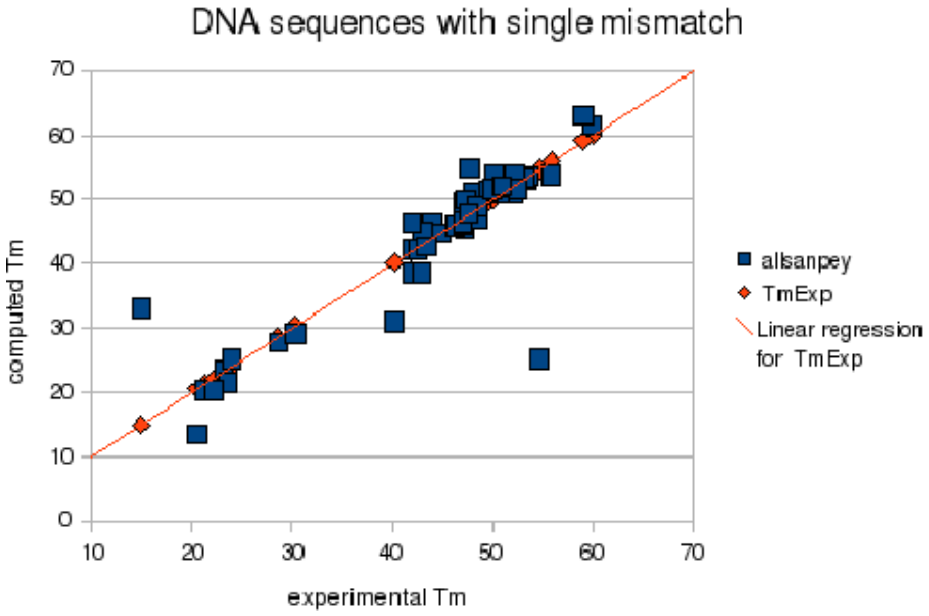

Figure 6: Comparison of experimental and computed Tm for various sets of DNA sequences containing one single mismatch. $[Na^+] = 1$ M, [nucleic acid] $= 4 \cdot 10^{-4}$ M

For DNA/RNA duplexes, the same model is used, taking parameters from Watkins et al. (2011). The only mismatches permitted are `dA/rA`, `dT/rU`, `dC/rC` and `dG/rG`.

Example :

$$\Delta H\begin{pmatrix}\texttt{CAT}\\ \texttt{GAU}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{CA}\\ \texttt{GA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AT}\\ \texttt{AU}\end{pmatrix}$$

(The same computation is performed for ΔS)

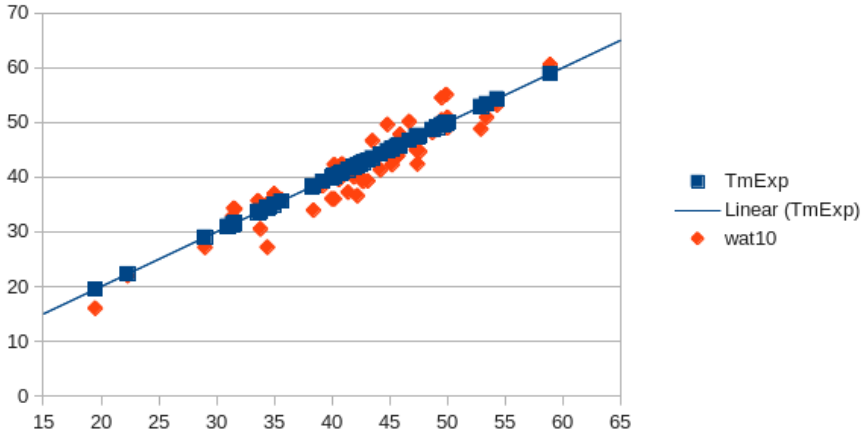

Figure 7: Comparison of experimental and computed Tm for various sets of DNA/RNA sequences containing one single mismatch. $[Na^+] = 1$ M, [nucleic acid] $= 4 \cdot 10^{-4}$ M

For RNA duplexes, the different models to computes the thermodynamic contribution of single mismatch to the helix coil stability are more complex.

Model from Amber R. Davis and Brent M Znosco, 2007-2008

$$\Delta h(\text{single-mismatch}) = \delta h_{\text{mismatch}-\text{nucleotides}} + \delta h_{\text{mismatch}-\text{NN}-\text{interaction}} + \delta h_{\text{AU/GU}}$$

Where :

$\delta h_{\text{mismatch}-\text{nucleotides}}$ accounts for the identity of the single mismatch nucleotides.

$\delta h_{\text{mismatch}-\text{NNinteraction}}$ accounts for the interaction between the mismatch nucleotides and the nearest neighbors. (R purine, Y pyrimidine)

$\delta h_{\text{AU/GU}}$ accounts for AU or GU nearest neighbors.

Example :

$$\Delta H \begin{pmatrix} \texttt{AUC} \\ \texttt{UUG} \end{pmatrix} = \Delta H \begin{pmatrix} \texttt{U} \\ \texttt{U} \end{pmatrix} + 1 \times \Delta H\,\texttt{AU} + \Delta H \begin{pmatrix} \texttt{RYY} \\ \texttt{YYR} \end{pmatrix}$$

(The same computation is performed for ΔS)

Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\Delta h(\text{single-mismatch}) = \delta h_{\text{initiation}-\text{loop}-\text{of}-2} + \delta h_{\text{per}-\text{AU/GU}} + \delta h_{\text{GG}} + \delta h_{\text{RU/YU}}$$

Where :
$\delta h_{\text{initiation}-\text{loop}-\text{of}-2}$ accounts for the initiation of a single non canonical pair.
δh_{GG} accounts for a GG single mismatch.
$\delta h_{\text{RU/YU}}$ accounts for a 5'RU/3'YU stack with R a purine and Y a pyrimidine.
$\delta h_{\text{per}-\text{AU/GU}}$ accounts for AU or GU nearest neighbors.
Example :

$$\Delta H\begin{pmatrix}\texttt{AUC}\\ \texttt{UUG}\end{pmatrix} = \Delta H\text{initiation-loop-of-2} + 1 \times \Delta H\texttt{per-AU} + \Delta H\begin{pmatrix}\texttt{RU}\\ \texttt{YU}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
allsanpey	DNA	Allawi and SantaLucia (1997) Biochemistry 36: 10581-10594 Allawi and SantaLucia (1998) Biochemistry 37: 2170-2179 Allawi and SantaLucia (1998) Nuc Acids Res 26: 2694-2701 Allawi and SantaLucia (1998) Biochemistry 37: 9435-9444 Peyret et al. (1999) Biochemistry 38: 3468-3477
wat11	DNA/RNA	Watkins et al. (2011)
tur06	RNA	Lu et al (2006) Nucleic Acids Research 34: 4912-4924
zno07	RNA	Davis and Znosko (2007) Biochemistry 46: 13425-13436
zno08	RNA at least one adjacent GU base pair	Davis and Znosko (2008) Biochemistry 47: 10178-10187

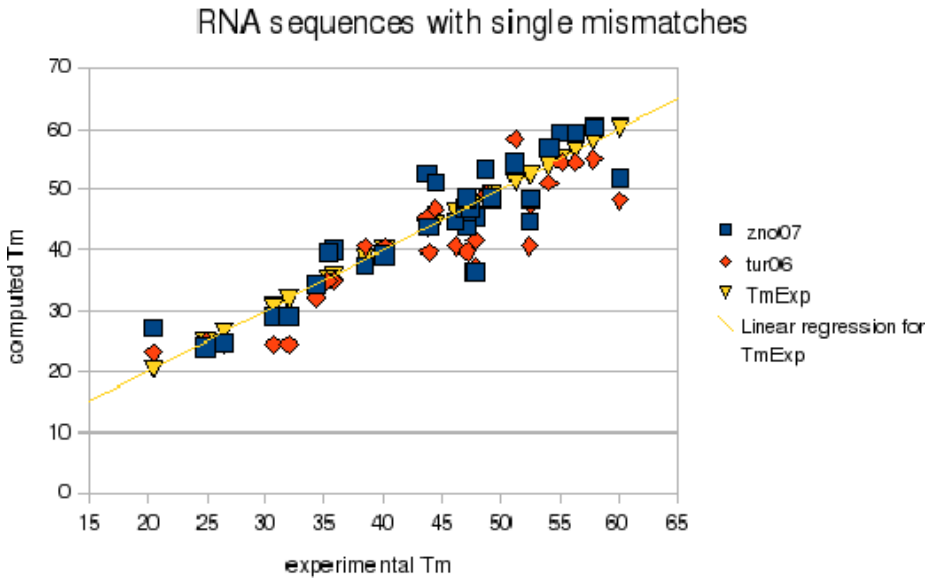

Figure 8: Comparison of experimental and computed Tm for various sets of RNA sequences containing one single mismatch. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.4 Tandem mismatches effect

The tandem mismatches (two adjacent mismatches) are taken into account but the two first and positions cannot be mismatched. Moreover the thermodynamic parameters are still not available for every possible cases. In such a case, the program, unable to compute any relevant result, will quit with a warning.

For DNA duplexes, this program computes the hybridisation enthalpy and entropy from the elementary parameters of each Crick's pair containing the mismatch(es).

$$\Delta h_{\text{tandem-mismatch}} = \delta h_{\text{Crick's-pair-containing-tandem-mismatch}} + \sum \delta h_{\text{Crick's-pair-containing-single-mismatch}}$$

Example :

$$\Delta H\begin{pmatrix}\texttt{ATGC}\\ \texttt{TCAG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{AT}\\ \texttt{TC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{TG}\\ \texttt{CA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{GC}\\ \texttt{AG}\end{pmatrix}$$

(The same computation is performed for ΔS)

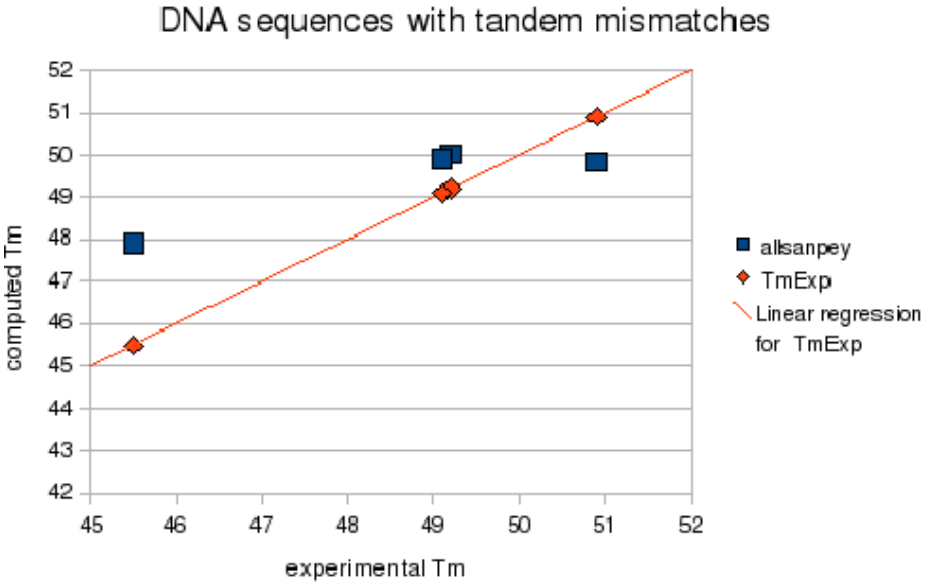

Figure 9: Comparison of experimental and computed Tm for various sets of DNA sequences containing one tandem mismatch. $[Na^+] = 1$ M, [nucleic acid] $= 4 \cdot 10^{-4}$ M

For RNA duplexes, the different models to computes the thermodynamic contribution of tandem mismatch to the helix coil stability are more complex.

Symmetric tandem mismatches : Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\Delta h(\text{tandem-mismatch}) = \delta h_{\text{tandem}-\text{mismatch}+\text{closing}-\text{base}-\text{pairs}}$$

Where :

$\delta h_{\text{mismatch}-\text{nucleotides}}$ accounts for the identity of the double mismatch nucleotides and the identity of the base pairs adjacent to the tandem mismatches.

Example :

$$\Delta H\begin{pmatrix}\texttt{G AC C}\\ \texttt{C CA G}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{AC}\\ \texttt{CA}\end{pmatrix} - ad\,jacent - to - GC$$

(The same computation is performed for ΔS)

Asymmetric tandem mismatches : Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\Delta h(\text{tandem-mismatch}) = (\delta h_{\text{symmetric}-\text{duplex}-1} + \frac{\delta h_{\text{symmetric}-\text{duplex}-2)}}{2} + \delta h_{\text{GG}} + \delta h_{\text{p}}$$

Where :

$\delta h_{\text{symmetric}-\text{duplex}-1}$ accounts for the enthalpy of a symmetric tandem mismatch composed of the first closing base pair and the first mismatch nucleotides.

$\delta h_{\text{symmetric}-\text{duplex}-2}$ accounts for the enthalpy of a symmetric tandem mismatch composed of the second closing base pair and the second mismatch nucleotides.

δh_{GG} accounts for a GG pair adjacent to a AA pair or any non canonical pair containing a pyrimidine.

δh_{p} accounts for an AG or GA pairs adjacent to a UC, CC or CU pair and a UU pair adjacent to an AA pair .

Example :

$$\Delta H\begin{pmatrix}\texttt{A GC C}\\ \texttt{U AU G}\end{pmatrix} = (\Delta H\begin{pmatrix}\texttt{A GA U}\\ \texttt{U AG A}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{G UC C}\\ \texttt{C CU G}\end{pmatrix} + \Delta H\texttt{GA} - ad\,jacent - to - \texttt{CU}$$

(The same computation is performed for ΔS)

For further information, see the referenced articles.

model	limits	Article
allsanpey	DNA only GT mismatches and TA/TG mismatches	Allawi and SantaLucia (1997) Biochemistry 36: 10581-10594 Allawi and SantaLucia (1998) Biochemistry 37: 2170-2179 Allawi and SantaLucia (1998) Nuc Acids Res 26: 2694-2701 Allawi and SantaLucia (1998) Biochemistry 37: 9435-9444 Peyret et al. (1999) Biochemistry 38: 3468-3477
tur99	RNA no adjacent GU or UG base pairs	Mathiews et al (1999) J.Mol.Biol. 288: 911-940

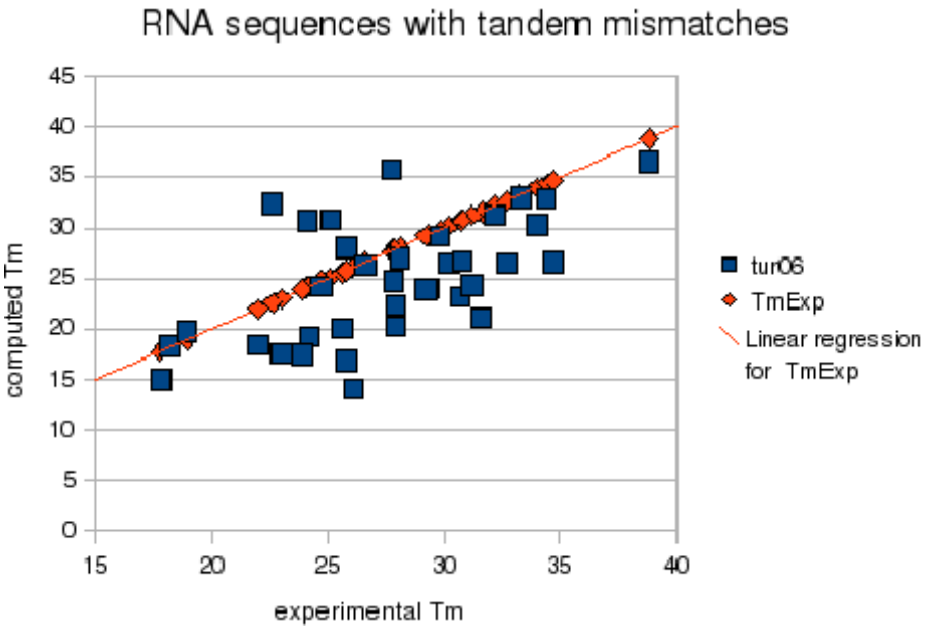

Figure 10: Comparison of experimental and computed Tm for various sets of RNA sequences containing one tandem mismatch. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.5 Internal loop effect

The internal loops (more than two adjacent mismatches) are taken into account but the two first and positions cannot be mismatched. Moreover the thermodynamic parameters are still not available for every possible cases. In such a case, the program, unable to compute any relevant result, will quit with a warning. Moreover, the thermodynamics of the nucleic acids within the internal loop are salt independent and no salt correction will be applied to it. However, the thermodynamics of the terminal mismatches are salt dependent and a salt correction will be applied to them. The thermodynamic model for DNA and RNA duplexes are similar.

DNA duplexes :Model from John Santalucia, Jr. and Donald Hicks, 2004

$$\begin{aligned}\Delta h(\text{internal-loop(n)}) = \delta h_{\text{asymmetry}} + \delta h_{\text{left-terminal-mismatch}} \\ + \delta h_{\text{right-terminal-mismatch}} \\ \Delta s(\text{internal-loop(n)}) = \delta s_{\text{loop(n)}} + \delta s_{\text{asymmetry}} + \delta s_{\text{left-terminal-mismatch}} \\ + \delta s_{\text{right-terminal-mismatch}}\end{aligned}$$

Where :

$\delta h_{\text{internal-loop(n)}}$ accounts for the internal loop of n nucleotides.

$\delta h_{\text{asymmetry}}$ accounts for the internal loop asymmetry (when the number of nucleic acid within the internal loop is higher in one of the strand).

$\delta h_{\text{left-terminal-mismatch}}$ accounts for the identity of the first mismatch nucleotides of the loop.

$\delta h_{\text{right-terminal-mismatch}}$ accounts for the identity of the last mismatch nucleotides of the loop.

Example : Symmetric internal loop

$$\Delta H\begin{pmatrix}\texttt{G ACCG C}\\ \texttt{C CATA G}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GA}\\ \texttt{CC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{GC}\\ \texttt{AG}\end{pmatrix}$$
$$\Delta S\begin{pmatrix}\texttt{G ACCG C}\\ \texttt{C CATA G}\end{pmatrix} = \Delta S\text{loop of 8} + \Delta S\begin{pmatrix}\texttt{GA}\\ \texttt{CC}\end{pmatrix} + \Delta S\begin{pmatrix}\texttt{GC}\\ \texttt{AG}\end{pmatrix}$$

RNA duplexes :Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\begin{aligned}\Delta h(\text{internal-loop(n)}) = \delta h_{\text{initiation-loop(n)}} + \delta h_{\text{per-AU/GU}} + (n1 - n2)\delta h_{\text{asymmetry}} \\ + \delta h_{\text{first-non-canonical-pairs}}\end{aligned}$$

Where :

$\delta h_{\text{initiation-loop(n)}}$ accounts for the internal loop of n nucleotides.

$\delta h_{\text{asymmetry}}$ accounts for the internal loop asymmetry (when the number of there is an unequal numbers of nucleotides on each side) with n1 and n2 the number of nucleotides on each strand..

$\delta h_{\text{per}_\text{A}\text{U/GU}}$ accounts for each AU or GU base pair adjacent to the internal loop.

$\delta h_{\text{first}-\text{non}-\text{canonical}-\text{pairs}}$ accounts for each sequence specific first mismatch (bonus). It is not applied to loops of the form 1 x (n-1) with n > 2.

Example : asymmetric internal loop

$$\Delta H \begin{pmatrix} \texttt{A ACCG C} \\ \texttt{U C-UA G} \end{pmatrix} = \Delta H\text{loop initiation}(7) + 1 \times \Delta H\texttt{per-AU} + (4-3)\Delta H\text{asymmetry}$$

(The same computation is performed for ΔS)

For further information, see the referenced articles.

model	limits	Article
san04	DNA missing asymmetry penalty, not tested with experimental results	Santalucia and Hicks (2004) Annu. Rev. Biophys. Biomol. Struct 33 : 415-440
tur06	RNA not tested with experimental results	Lu et al (2006) Nucleic Acids Research 34: 4912-4924

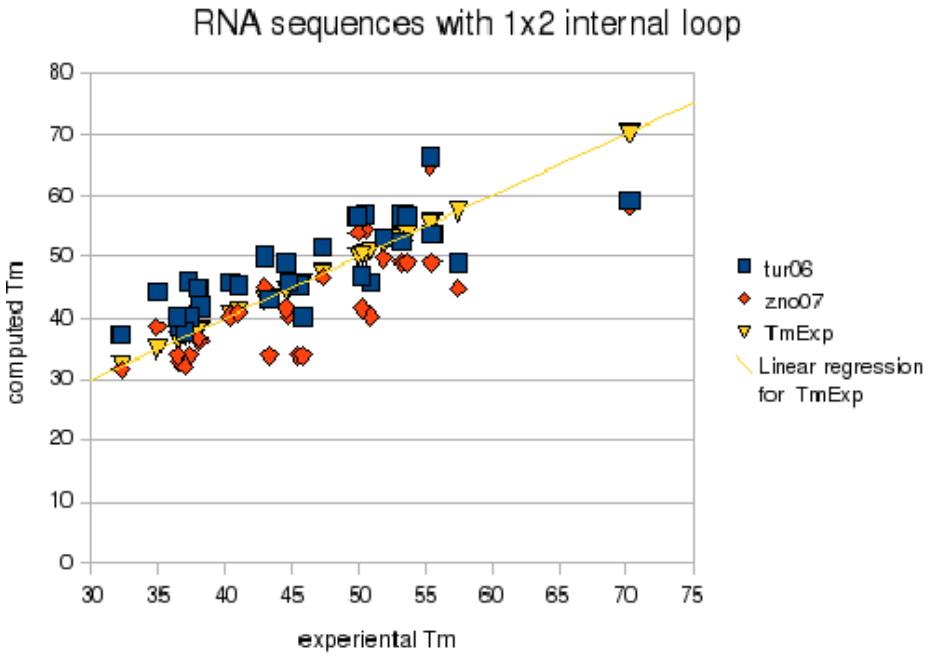

Figure 11: Comparison of experimental and computed Tm for various sets of RNA sequences containing one 1x2 internal loop. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.6 GU wobble base pairs effect

The wobble GU base pairs are taken into account. This pairing is a non-Watson-Crick base pairing between two nucleotides in RNA molecules, but the thermodynamic stability of a wobble base pair is comparable to that of a Watson-Crick base pair. Melting can also compute the thermodynamic of patterns with several adjacent GU base pairs. This program computes the hybridisation enthalpy and entropy from the elementary parameters of each Crick's pair containing the GU base pairs.

$$\Delta h_{\text{pattern-composed-of-GU-base-pairs}} = \sum \delta h_{\text{Crick'spair-containing-GU-base-pairs}}$$

Examples : One GU base pair

$$\Delta H\begin{pmatrix}\texttt{GUC}\\ \texttt{CGG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GU}\\ \texttt{CG}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{UC}\\ \texttt{GG}\end{pmatrix}$$

Examples : Two adjacent GU base pairs

$$\Delta H\begin{pmatrix}\texttt{GUGC}\\ \texttt{CGUG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GU}\\ \texttt{CG}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{UG}\\ \texttt{GU}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{UC}\\ \texttt{GG}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	**limits**	**Article**
tur99	RNA	Mathiews et al (1999) J.Mol.Biol. 288: 911-940

model	**limits**	**Article**
ser12	RNA	Serra et al (2012) Biochemistry 51: 3508-3522

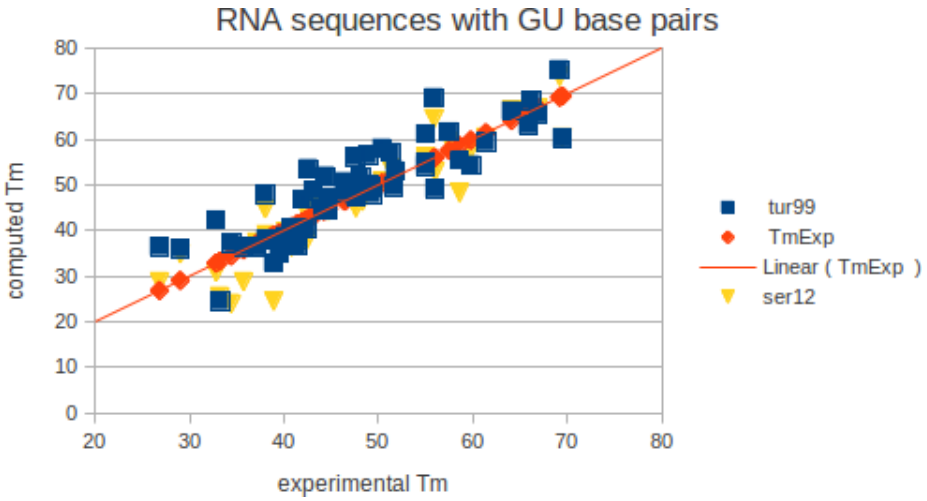

Figure 12: Comparison of experimental and computed Tm for various sets of RNA sequences containing GU base pairs. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.7 Single dangling end effect

The single dangling ends, that is the unmatched terminal nucleotides, can be taken into account, but all the thermodynamic parameters are not available. In such a case, the result is unpredictable, and all cases are possible.

For DNA and RNA duplexes, this program computes the hybridisation enthalpy and entropy from the elementary parameters of the Crick's pair containing the single dangling end.

$$\Delta h_{\text{single-dangling-end}} = \delta h_{\text{Crick's-pair-containing-the-dangling-end}}$$

Example : If the duplex is :

```
GCTAG-
CGATCA
```

$$\Delta H\begin{pmatrix}\texttt{GCTAG-}\\ \texttt{CGATCCA}\end{pmatrix} = \Delta H\text{perfectly-matching-sequence} + \Delta H\text{single-dangling-end}$$

$$\Delta H\begin{pmatrix}\texttt{GCTAG-}\\ \texttt{CGATCCA}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GCTAG}\\ \texttt{CGATCC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{G-}\\ \texttt{CA}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	**limits**	**Article**
bom00	DNA	Bommarito et al. (2000) Nuc Acids Res 28: 1929-1934
sugdna02	DNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372
sugrna02	RNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372
ser08	RNA only 3' UA, GU and UG terminal base pairs only 5' UG and GU terminal base pairs	O'tool et al. (2006) Nucleic Acids research 34: 3338-3344 Miller et al. (2008) Nucleic Acids research 36: 5652-5659

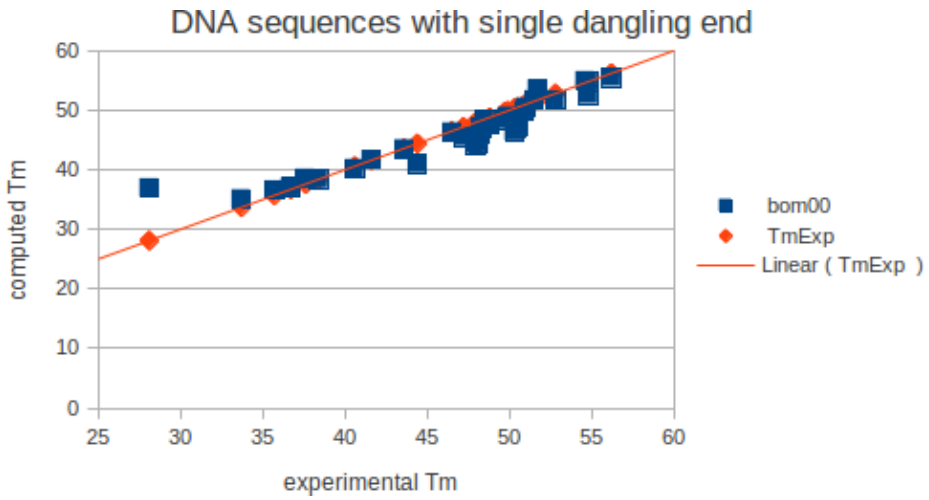

Figure 13: Comparison of experimental and computed Tm for various sets of DNA sequences containing single dangling ends. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

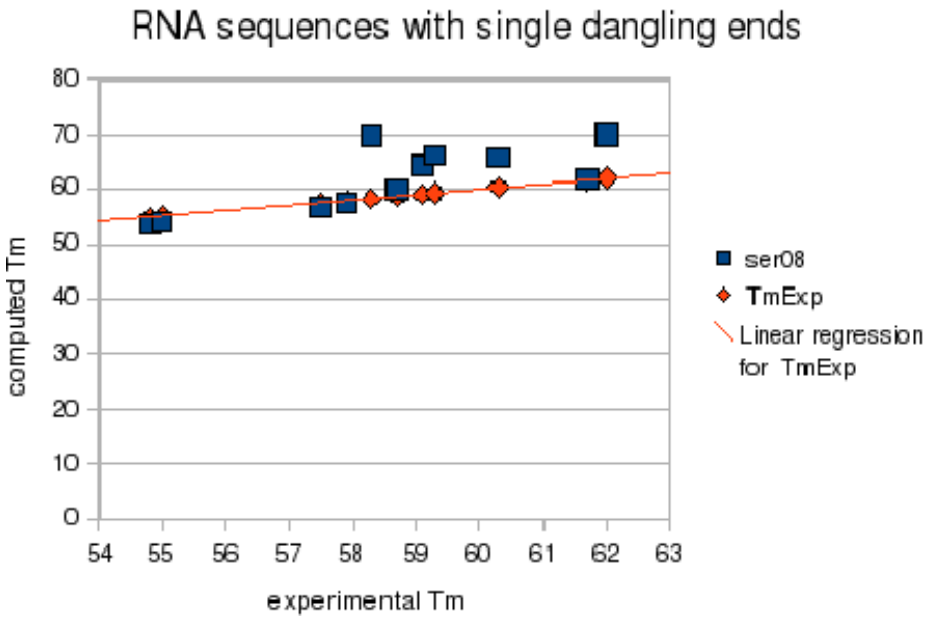

Figure 14: Comparison of experimental and computed Tm for various sets of RNA sequences containing single dangling ends. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.8 Double dangling end effect

The double dangling ends, that is the two adjacent unmatched terminal nucleotides, can be taken into account (mostly for RNA sequences). This program computes the hybridisation enthalpy and entropy in two times : First, it computes the energy from the single dangling end as if the duplex contained only a single danging end and then, it adds a bonus for the second dangling end if it is necessary.

$$\Delta h_{\text{double-dangling-end}} = \delta h_{\text{single-dangling-end}} + \delta h_{\text{bonus-second-dangling-end}}$$

Example :

$$\Delta H \begin{pmatrix} \texttt{UAC} \\ \texttt{A-} \end{pmatrix} = \Delta H\text{UA} + \Delta H\text{bonus-pyrimidine-purine-pyrimidine}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
sugdna02	DNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372
sugrna02	RNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372
ser05	RNA depends on the available thermodynamic parameters for single dangling ends	O'toole et al. (2005) RNA 11: 512-516
ser06	RNA	O'toole et al. (2006) Nucleic Acids research 34: 3338-3344

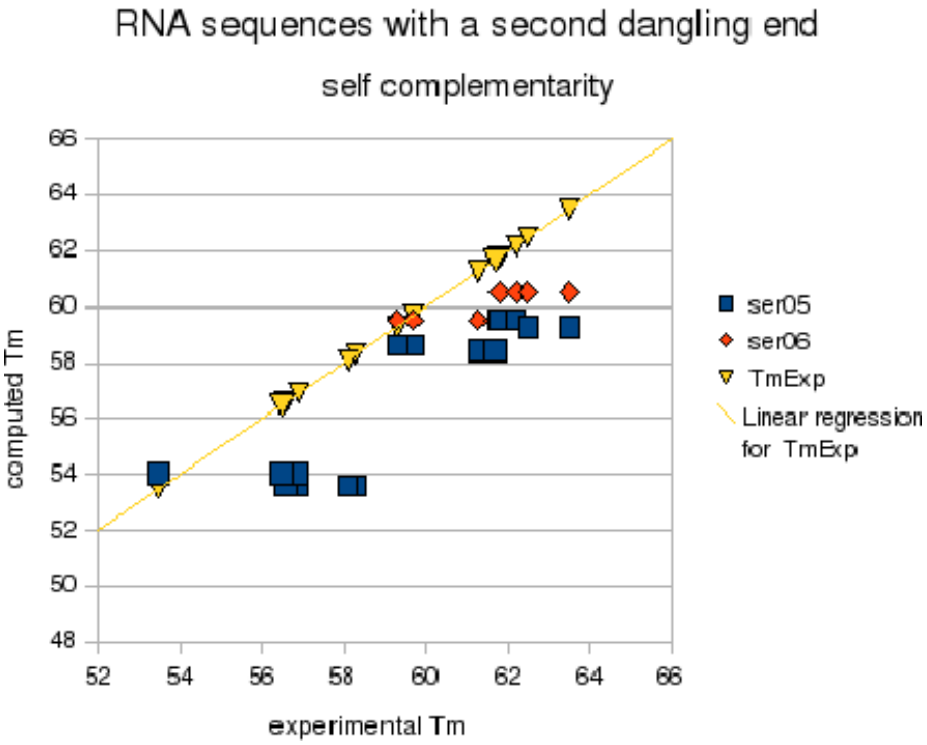

Figure 15: Comparison of experimental and computed Tm for various sets of RNA sequences containing double dangling ends. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.9 Long dangling end effect (poly A)

The long dangling ends, that is all the adjacent unmatched terminal nucleotides, can be taken into account (only for polyA dangling ends for the moment). It is possible to compute the thermodynamic form one to four poly A dangling end. This program computes the hybridisation enthalpy and entropy from the parameters of the long dangling end with the adjacent terminal base pair.

$$\Delta h_{\text{long-dangling-end}} = \delta h_{\text{adjacent-terminal-base-pair+polyA}}$$

Example :
If the duplex is :

```
GCTAG---
CGATCAAA
```

$$\Delta H\begin{pmatrix}\texttt{GCTAG--}\\ \texttt{CGATCCAAA}\end{pmatrix} = \Delta H\text{perfectly-matching-sequence} + \Delta H\text{long-dangling-end}$$

$$\Delta H\begin{pmatrix}\texttt{GCTAG--}\\ \texttt{CGATCCAAA}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GCTAG}\\ \texttt{CGATCC}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{G--}\\ \texttt{CAAA}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
sugdna02	DNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372
sugrna02	RNA only terminal poly A self complementary sequences	Ohmichi et al. (2002) J. Am. Chem. Soc. 124: 10367-10372

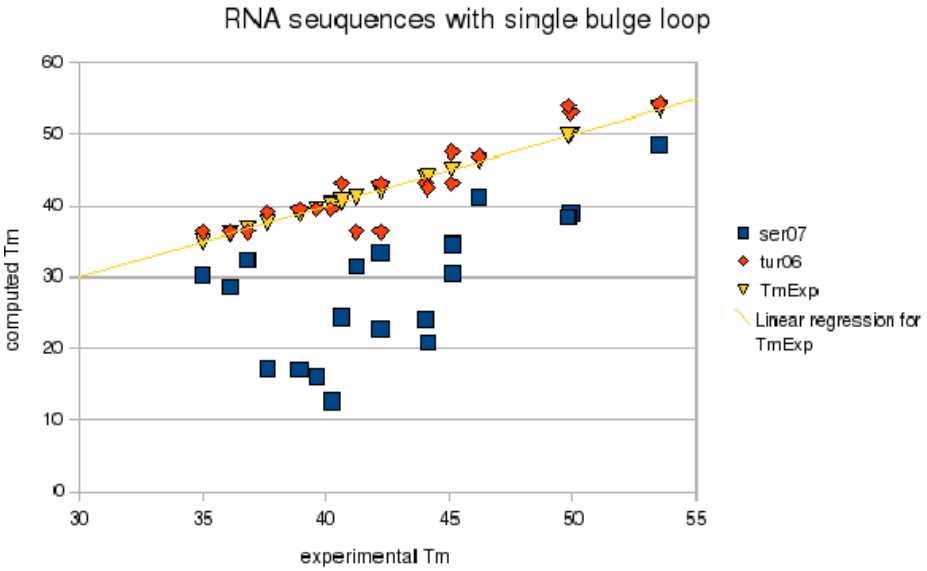

Figure 16: Comparison of experimental and computed Tm for various sets of DNA sequences containing long polyA dangling ends. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

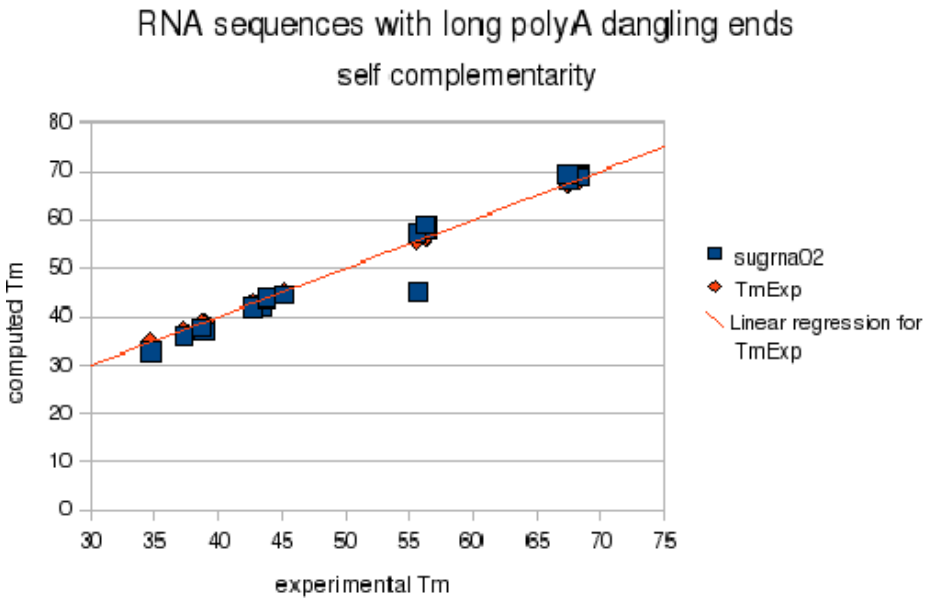

Figure 17: Comparison of experimental and computed Tm for various sets of RNA sequences containing long polyA dangling ends. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.10 Single bulge loop effect

The single bulge loops, that is the single unmatched internal nucleotides, can be taken into account. , but all the thermodynamic parameters are not available. In such a case, the result is unpredictable, and all cases are possible. There are several different models to compute the thermodynamic of single bulge loop:

DNA and RNA duplexes :nearest neighbor model "NNN"

$$\Delta h(\text{single-bulge-loop}) = \delta h_{\text{unpaired}-\text{nucleotid}+\text{adjacent}-\text{base}-\text{pairs}}$$

Example : If the duplex is :

```
GCTTAGGC
CGA-TCCG
```

$$\Delta H\begin{pmatrix}\texttt{GCTTAGGC}\\ \texttt{CGA-TCCG}\end{pmatrix} = \Delta H\text{perfectly-matching-sequence-1} + \Delta H\text{single-bulge-loop} + \Delta H\text{perfectly-matching-sequence-2}$$

$$\Delta H\begin{pmatrix}\texttt{GCTTAGGC}\\ \texttt{CGA-TCCG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GCT}\\ \texttt{CGA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{TTA}\\ \texttt{A-T}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AGGC}\\ \texttt{TCCG}\end{pmatrix}$$

(The same computation is performed for ΔS)

However, some types of single bulge loop can't be only modelled with a NNN nearest neighbor model and the following models can give more reliable and accurate results (mostly for RNA single bulge loops.)

DNA duplexes :Model from John Santalucia, Jr. and Donald Hicks, 2004

$$\Delta h(\text{single-bulge-loop}) = \delta h_{\text{intervening}-\text{NN}} + \delta h_{\text{closing}-\text{AT}-\text{penalty}}$$
$$\Delta s(\text{single-bulge-loop}) = \delta s_{\text{bulge}-\text{loop}-\text{of}-1} + \delta s_{\text{intervening}-\text{NN}} + \delta s_{\text{closing}-\text{AT}-\text{penalty}}$$

Where :
$\delta h_{\text{bulge}-\text{loop}-\text{of}-1}$ accounts for the bulge loop of 1 nucleotide.
$\delta h_{\text{intervening}-\text{NN}}$ accounts for the intervening base pair stack.
$\delta h_{\text{closing}-\text{AT}-\text{penalty}}$ accounts for each AT base pair adjacent to the single bulge loop.
Example :

$$\Delta H\begin{pmatrix}\texttt{GAC}\\ \texttt{C-G}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GC}\\ \texttt{CG}\end{pmatrix}$$

$$\Delta S\begin{pmatrix}\texttt{GAC}\\ \texttt{C-G}\end{pmatrix} = \Delta S\text{bulge-loop-of-1} + \Delta S\begin{pmatrix}\texttt{GC}\\ \texttt{CG}\end{pmatrix}$$

(The same computation is performed for ΔS)

RNA duplexes :Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\Delta h(\text{single-bulge-loop}) = \delta h_{\text{initiation}-\text{bulge}-\text{loop}-\text{of}-1} + \delta h_{\text{intervening}-\text{NN}}$$

Where :
$\delta h_{\text{initiation}-\text{bulge}-\text{loop}-\text{of}-1}$ accounts for the initiation of bulge loop of 1 nucleotide.
$\delta h_{\text{intervening}-\text{NN}}$ accounts for the intervening base pair stack.
Example :

$$\Delta H\begin{pmatrix}\texttt{GAC}\\ \texttt{C-G}\end{pmatrix} = \Delta H\text{initiation-bulge-loop-of-1} + \Delta H\begin{pmatrix}\texttt{GC}\\ \texttt{CG}\end{pmatrix}$$

(The same computation is performed for ΔS) For further information, see the referenced articles.

model	**limits**	**Article**
tan04	DNA	Tanaka et al (2004) Biochemistry 43 : 7143-7150
san04	DNA missing closing AT penalty	Santalucia and Hicks (2004) Annu. Rev. Biophys. Biomol. Struct 33 : 415-440
ser07	RNA les reliable results some missing parameters	Blose et al (2007) Biochemistry 46 : 15123-15135
tur06	RNA	Lu et al (2006) Nucleic Acids Research 34: 4912-4924

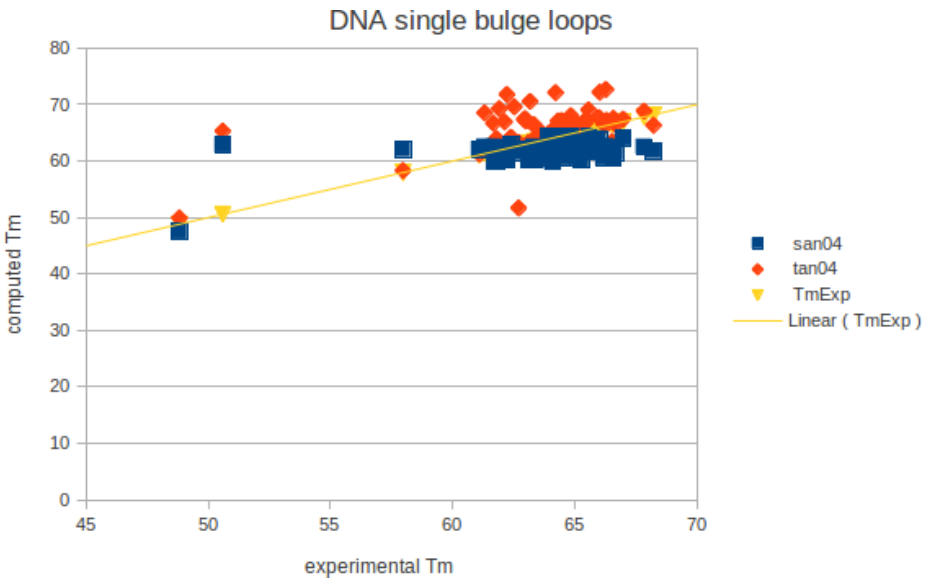

Figure 18: Comparison of experimental and computed Tm for various sets of DNA sequences containing one single bulge loop. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

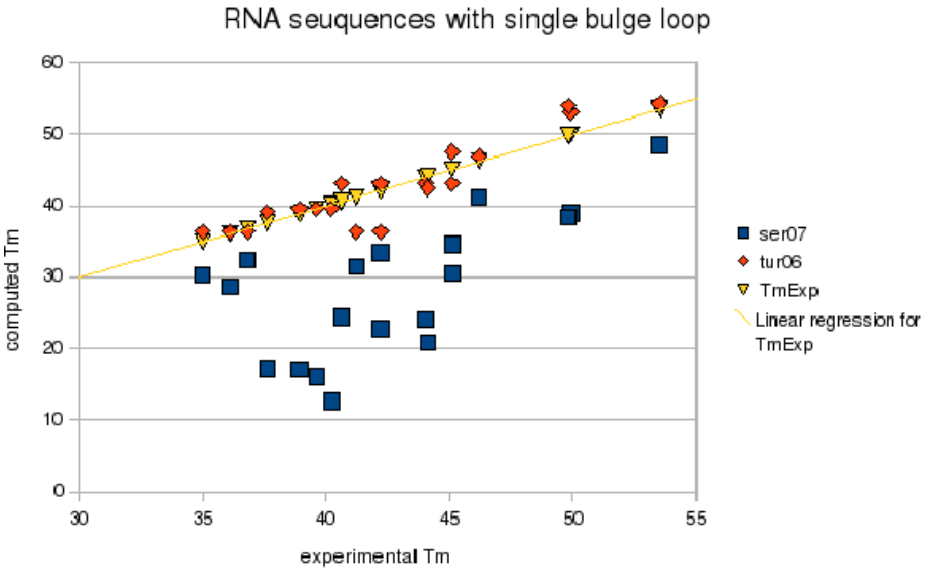

Figure 19: Comparison of experimental and computed Tm for various sets of RNA sequences containing one single bulge loop. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.11 long bulge loop effect

The long bulge loops, that is all the adjacent unmatched internal nucleotides, can be taken into account. , but all the thermodynamic parameters are not available. In such a case, the result is unpredictable, and all cases are possible. The RNA and DNA thermodynamic models are similar :

DNA duplexes :Model from John Santalucia, Jr. and Donald Hicks, 2004

$$\Delta h(\text{long-bulge-loop}) = \delta h_{\text{closing-AT-penalty}}$$
$$\Delta s(\text{single-bulge-loop}) = \delta s_{\text{bulge-loop-of-n}} + \delta s_{\text{closing-AT-penalty}}$$

Where :
$\delta h_{\text{bulge-loop-of-n}}$ accounts for the bulge loop of n nucleotides.
$\delta h_{\text{closing-AT-penalty}}$ accounts for each AT base pair adjacent to the long bulge loop.
Example :

$$\Delta H\begin{pmatrix}\texttt{GACGC}\\ \texttt{C--G}\end{pmatrix} = 0\Delta S\begin{pmatrix}\texttt{GACGC}\\ \texttt{C--G}\end{pmatrix} = \Delta S\text{bulge-loop-of-3}$$

RNA duplexes : Model from Zhi Johm Lu, Douglas H. Turner and David H. Mathews, 2006

$$\Delta h(\text{long-bulge-loop}) = \delta h_{\text{initiation-bulge-loop-of-n}} + \delta h_{\text{per-AU/GU-penalty}}$$

Where :
$\delta h_{\text{initiation-bulge-loop-of-n}}$ accounts for the initiation of the bulge loop of n nucleotides.
$\delta h_{\text{per-AU/GU-penalty}}$ accounts for each AU or GU base pair adjacent to the long bulge loop.
Example :

$$\Delta H\begin{pmatrix}\texttt{AACGC}\\ \texttt{U--G}\end{pmatrix} = \Delta H\text{initiation-bulge-loop-of-3} + 1 \times \Delta H\text{per-AU-penalty}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
san04	DNA missing closing AT penalty not tested with experimental results	Santalucia and Hicks (2004) Annu. Rev. Biophys. Biomol. Struct 33 : 415-440
tur06	RNA not tested with experimental results	Lu et al (2006) Nucleic Acids Research 34: 4912-4924

4.1.12 Inosine bases effect

The inosine bases (I) are taken into account, but all the thermodynamic parameters are not available. In such a case, the result is unpredictable, and all cases are possible, so the program quit with a warning. For the RNA duplexes, only the thermodynamic parameters for IU base pairs are available for the moment. This program computes the hybridisation enthalpy and entropy from the elementary parameters of each Crick's pair containing the inosine base.

$$\Delta h_{\text{pattern-containing-inosine-bases}} = \sum \delta h_{\text{Crick'spair-containing-inosine-bases}}$$

Examples : One inosine base

$$\Delta H\begin{pmatrix}\texttt{AIC}\\ \texttt{TAG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{AI}\\ \texttt{TA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{IC}\\ \texttt{AG}\end{pmatrix}$$

Examples : Two adjacent base pairs containing inosine

$$\Delta H\begin{pmatrix}\texttt{GIAC}\\ \texttt{CAIG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GI}\\ \texttt{CA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{IA}\\ \texttt{AI}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AC}\\ \texttt{IG}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
san05	DNA missing parameters for tandem base pairs containing inosine bases	Watkins and Santalucia (2005) Nucleic acids research 33 : 6258-6267
zno07	RNA only IU base pairs	Wright et al. (2007) Biochemistry 46 : 4625-4634

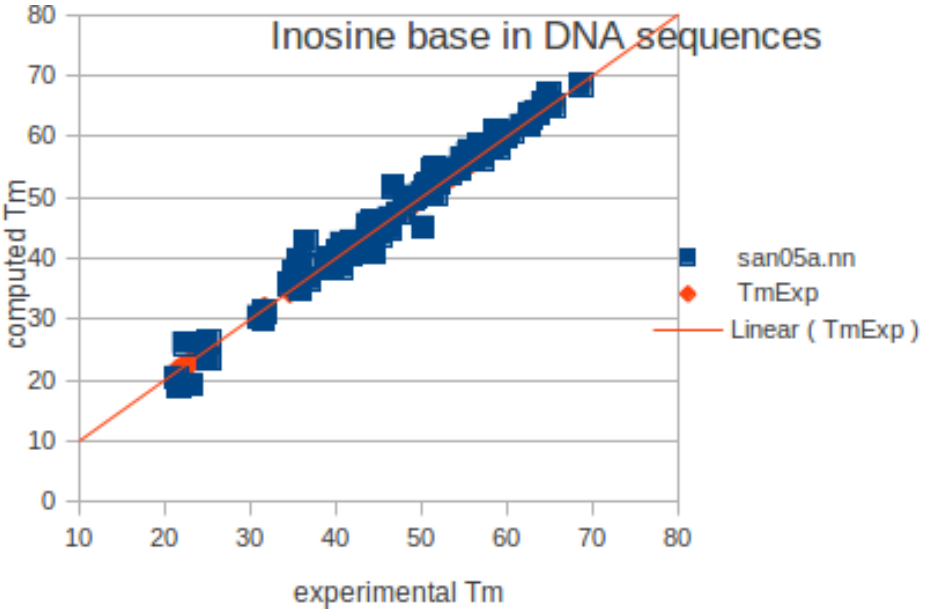

Figure 20: Comparison of experimental and computed Tm for various sets of DNA sequences containing inosine. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

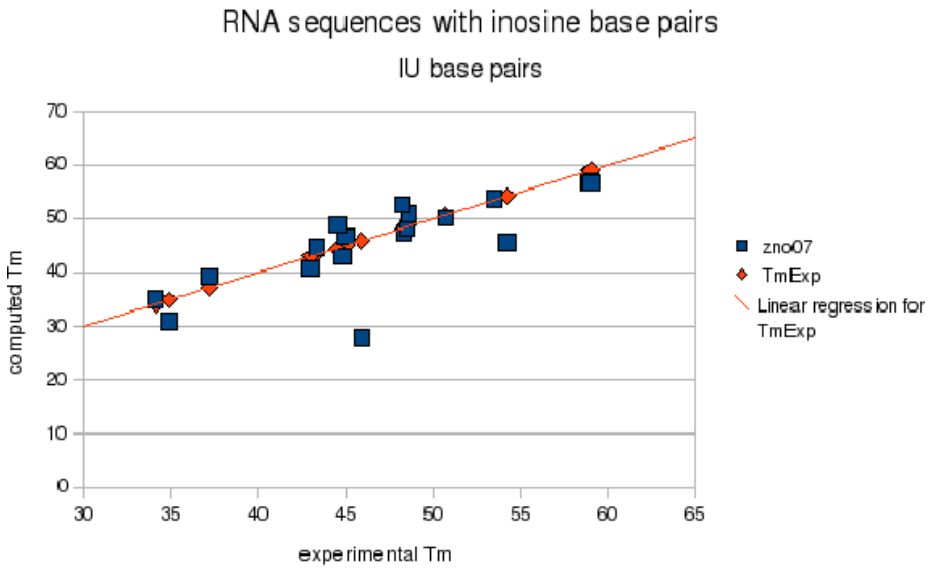

Figure 21: Comparison of experimental and computed Tm for various sets of RNA sequences containing inosine. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.13 Azobenzenes effect

The trans azobenzenes (X_T) and cis azobenzenes (X_C) in DNA duplexes are taken into account. Be aware : when the number of cis azobenzenes increases in the sequence, the predictions are less accurate and less reliable.

$$\Delta h_{\text{pattern-containing-azobenzene}} = \delta h_{\text{Crick'spair-containing-azobenzene+adjacent-base-pairs}}$$

Example : If the duplex is :

```
GCTX_CAGGC
CGATCCG
```

$$\Delta H\begin{pmatrix}\texttt{GCTX\_CAGGC}\\ \texttt{CGATCCG}\end{pmatrix} = \Delta H\text{perfectly-matching-sequence-1} + \Delta H\text{azobenzene} + \Delta H\text{perfectly-matching-sequence-2}$$

$$\Delta H\begin{pmatrix}\texttt{GCTX\_CAGGC}\\ \texttt{CGATCCG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GCT}\\ \texttt{CGA}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{TX\_CA}\\ \texttt{AT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AGGC}\\ \texttt{TCCG}\end{pmatrix}$$

(The same computation is performed for ΔS)
For further information, see the referenced articles.

model	limits	Article
asa05	DNA less reliable results when the number of cis azobenzene increases	Asanuma et al. (2005) Nucleic acids Symposium Series 49 : 35-36

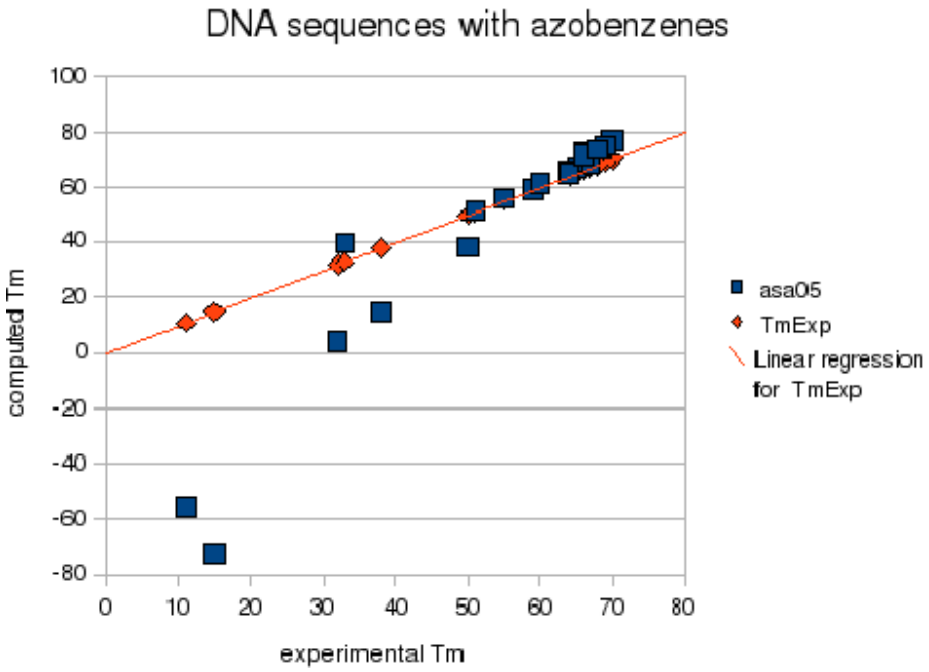

Figure 22: Comparison of experimental and computed Tm for various sets of DNA sequences containing azobenzene. $[Na^+] = 1$ M, [nucleic acid] $= 2 \cdot 10^{-6}$ M

4.1.14 2-Hydroxyadenine bases effect

The 2-hydroxyadenine bases (A*) in DNA duplexes are taken into account, but only in this two different sequence contexts : 5' GA*C 3' and 5' TA*A 3'. The program computes the enthalpy and the entropy in two times : first it computes the enthalpy and entropy of the two Crick's pairs containing the hydroxyadenine as if the base pair containing the hydroxyadenine was a simple AT base pair, and then it computes the hydroxyadenine increments.

$$\Delta h_{\text{pattern-containing-hydroxyadenine}} = \sum \delta h_{\text{Crick'spair-with-AT-base-pair}} + \delta h_{\text{hydroxyadenine-increment}}$$

Examples

$$\Delta H\begin{pmatrix}\texttt{GA*C}\\ \texttt{CCG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GA}\\ \texttt{CT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AC}\\ \texttt{CG}\end{pmatrix} + \Delta H\text{increment-for-GA*C/CCG}$$

(The same computation is performed for ΔS)

For further information, see the referenced articles.

model	limits	Article
sug01	DNA only in 5' GA*C 3' and 5' TA*A contexts	Kawakami et al.(2001) Nucleic acids research 29 : 3289-3296

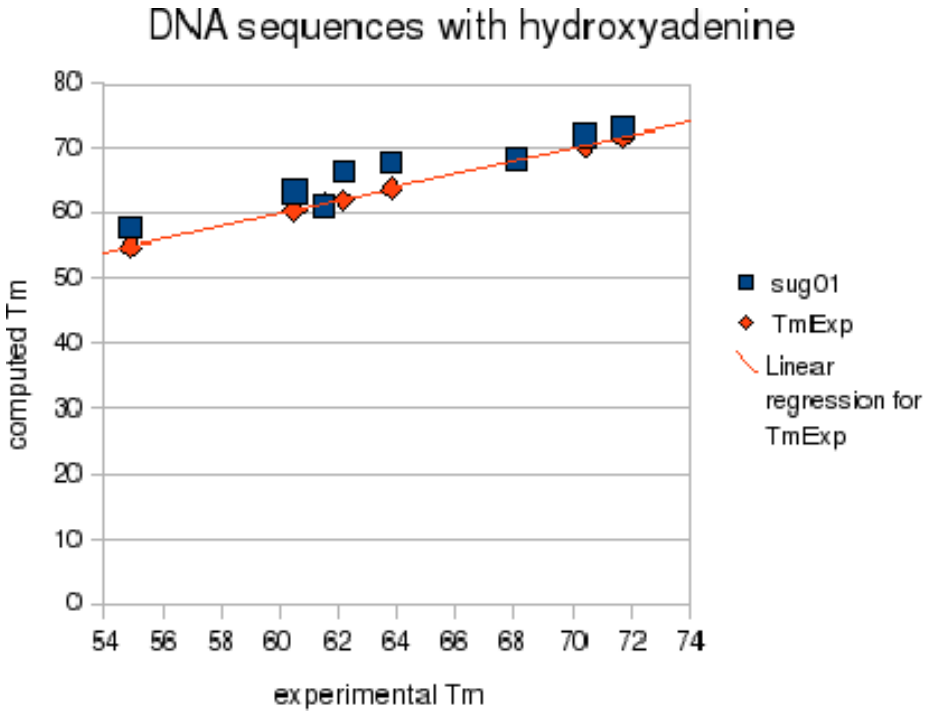

Figure 23: Comparison of experimental and computed Tm for various sets of DNA sequences containing hydroxyadenine. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.1.15 Single Locked nucleic acid effect

The locked nucleic acids (AL, GL, CL, TL) in DNA duplexes are taken into account.

DNA duplexes : Model from McTigue et al., 2004 The program computes the enthalpy and the entropy in two times : first it computes the enthalpy and entropy of the two Crick's pairs containing the locked nucleic acid as if the locked nucleic acid was a simple nucleic acid, and then it computes the locked nucleic acid increments for each Crick's base pair containing the locked nucleic acid.

$$\Delta h_{\text{pattern-containing-Locked-Nucleic-Acid}} = \sum \delta h_{\text{Crick'spair-without-Locked-Nucleic-Acid}} + \sum \delta h_{\text{increment-Crick'spair-with-Locked-Nucleic-Acid}}$$

Examples

$$\Delta H\begin{pmatrix}\texttt{GALC}\\ \texttt{CTG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GA}\\ \texttt{CT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{AC}\\ \texttt{CG}\end{pmatrix} + \Delta H\text{increment-for-GAL/CT} + \Delta H\text{increment-for-ALC/TG}$$

(The same computation is performed for ΔS)

DNA duplexes : Model from Owczarzy et al., 2011 The program computes the enthalpy and the entropy following the same nearest-neighbor formula as for perfectly matching Crick's pairs.

Examples

$$\Delta H\begin{pmatrix}\texttt{GALC}\\ \texttt{CTG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GAL}\\ \texttt{CT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{ALC}\\ \texttt{CG}\end{pmatrix}$$

For further information, see the referenced articles.

model	limits	Article
mct04	DNA	McTigue et al.(2004) Biochemistry 43 : 5388-5405
owc11	DNA	Owczarzy et al.(2011) Biochemistry 50 : 9352-9367

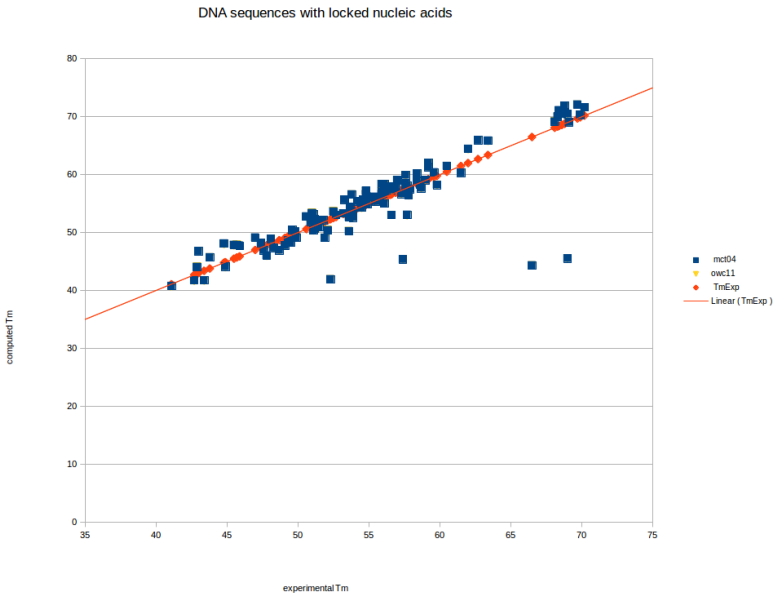

Figure 24: Comparison of experimental and computed Tm for various sets of DNA sequences containing single Locked Nucleic Acid. $[Na^+] = 1$ M, [nucleic acid] $= 5 \cdot 10^{-6}$ M

4.1.16 Consecutive Locked nucleic acids effect

The consecutive locked nucleic acids (AL, GL, CL, TL) in DNA duplexes are taken into account. The program computes the enthalpy and the entropy following the same nearest-neighbor formula as for perfectly matching Crick's pairs.

Examples

$$\Delta H\begin{pmatrix}\texttt{GALCLC}\\ \texttt{CTGG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GAL}\\ \texttt{CT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{ALCL}\\ \texttt{TG}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{CLC}\\ \texttt{GG}\end{pmatrix}$$

For further information, see the referenced articles.

model	limits	Article
owc11	DNA	Owczarzy et al.(2011) Biochemistry 50 : 9352-9367

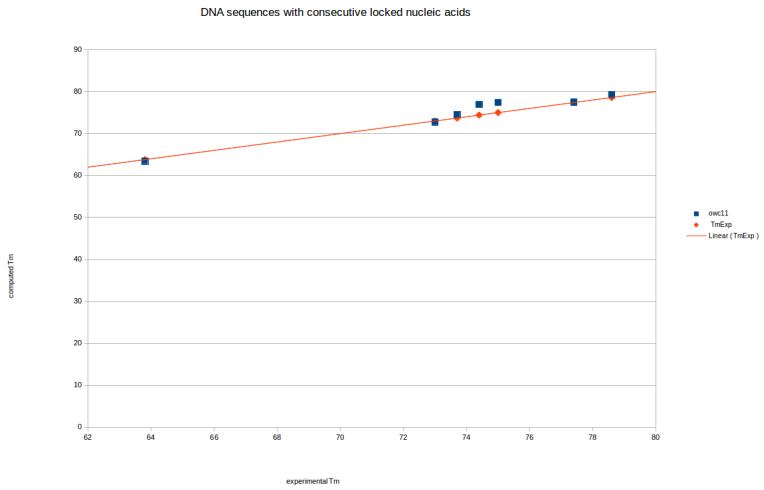

Figure 25: Comparison of experimental and computed Tm for various sets of DNA sequences containing consecutive Locked Nucleic Acids. $[Na^+] = 1$ M, [nucleic acid] $= 2 \cdot 10^{-6}$ M

4.1.17 Consecutive Locked nucleic acids with a single mismatch effect

The consecutive locked nucleic acids with a single mismatch (AL, GL, CL, TL) in DNA duplexes are taken into account. The program computes the enthalpy and the entropy following the same nearest-neighbor formula as for perfectly matching Crick's pairs.

Examples

$$\Delta H\begin{pmatrix}\texttt{GALCLGLC}\\ \texttt{CTTCG}\end{pmatrix} = \Delta H\begin{pmatrix}\texttt{GAL}\\ \texttt{CT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{ALCL}\\ \texttt{TT}\end{pmatrix} + \Delta H\begin{pmatrix}\texttt{CLGL}\\ \texttt{TC}\end{pmatrix}\Delta H\begin{pmatrix}\texttt{GLC}\\ \texttt{CG}\end{pmatrix}$$

For further information, see the referenced articles.

model	limits	Article
owc11	DNA	Owczarzy et al.(2011) Biochemistry 50 : 9352-9367

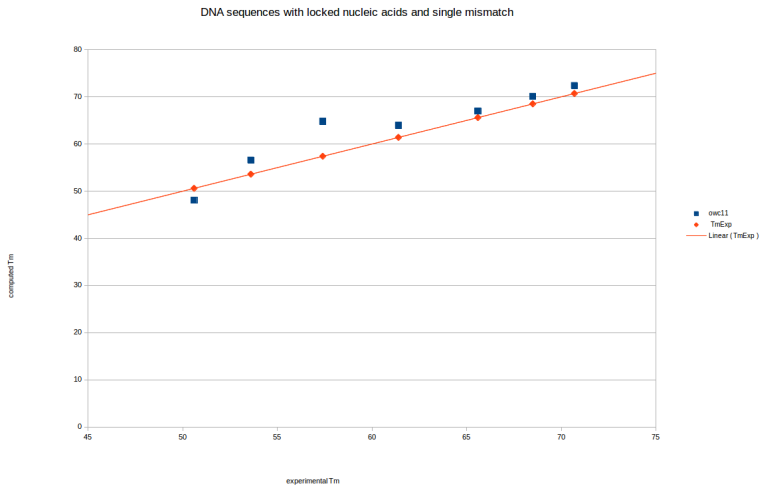

Figure 26: Comparison of experimental and computed Tm for various sets of DNA sequences containing consecutive Locked Nucleic Acids and a single mismatch. $[\mathrm{Na}^+] = 1$ M, $[\text{nucleic acid}] = 2 \cdot 10^{-6}$ M

4.2 The melting temperature

Then the melting temperature is computed by the following formula:

$$\text{Tm} = \frac{\Delta H}{\Delta S + R \ln(C_T/F)} - 273.15$$

Tm in K (for $[\mathrm{Na}^+]$ = 1 M)

In case of self complementary sequences, if the sequence (5' 3') is a sequence of type G(CNG)xC and x > 4, the sequence mainly turns into hairpin loops and this program will compute the melting temperature with this formula :

$$\text{Tm} = \frac{\Delta H}{\Delta S} - 273.15$$

Tm in K

Moreover, no ion correction will be applied to this formula.

4.3 Correction for the concentration of nucleic acid

F is 1 in the case of self-complementarity oligonucleotides. If the ODNs are not self-complementary, F is 4 if both strands are present in equivalent amount and F is 1 if one strand is in excess (for instance in PCR experiments). As a matter of facts, when the oligonucleotides are not self-complementary, the formula in the denominator is $C_{\text{max}} - C_{\text{min}}/2$ where C_{max} is the concentration of the strand in excess and C_{min} the concentration of the other strand. If the excess is sufficient, the total concentration is equivalent to the concentration of the strand in excess. Therefore, if one strand is in excess, the actual formula is effectively $C_{\text{max}} - C_{\text{min}}/2$ but if $C_{\text{max}} \gg C_{\text{min}}$,

$C_{\max} - C_{\min}/2$ is equivalent to the total concentration C_T. If $C_{\max}$ is close to $C_{\min}$, $C_{\max} - C_{\min}/2$ is equivalent to $C_T/4$, which is the default correction.

F is 4 by default but note that MELTING can detect self complementary sequences for perfectly matching sequences even though there is(are) dangling end(s). In this case, the program will automatically change F to 1. In addition to that, the computation takes an entropic term to correct for self-complementarity. In case of other self complementary sequences which doesn't match perfetcly, the option *-self* must be used to inform the program of the self complementarity.

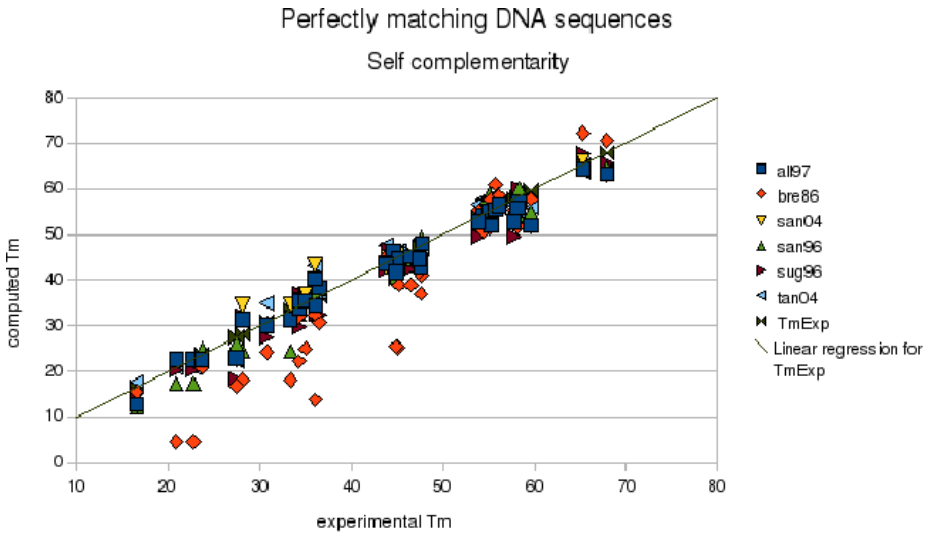

Figure 27: Comparison of experimental and computed Tm for various sets of DNA self complementary sequences. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

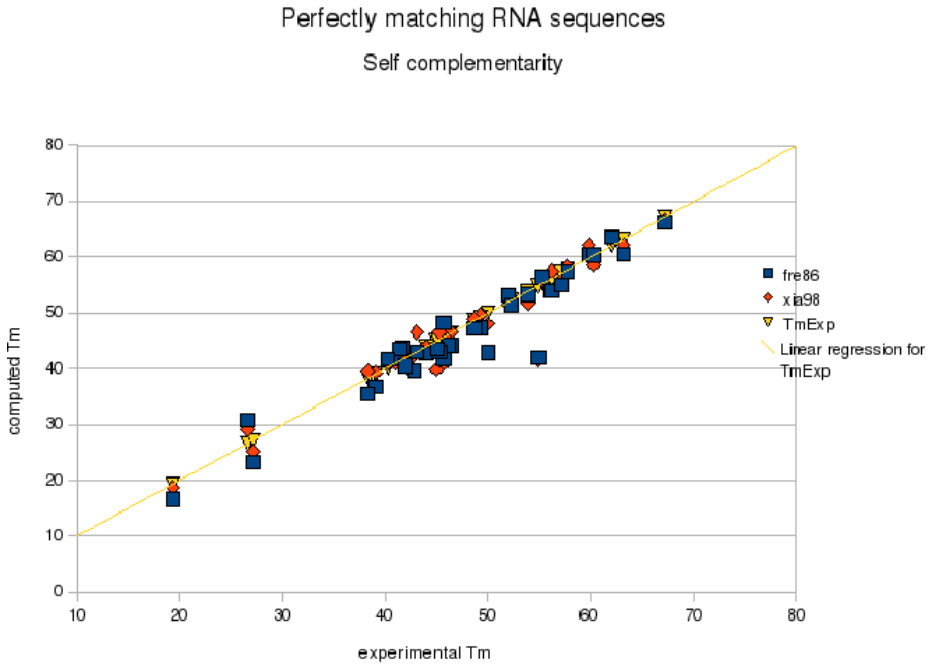

Figure 28: Comparison of experimental and computed Tm for various sets of RNA self complementary sequences. $[Na^+] = 1$ M, [nucleic acid] $= 1 \cdot 10^{-4}$ M

4.4 Correction for the concentration of cations

After the program computed the melting temperature for $[Na^+]$=1, an ion correction wille be applied either directly on the computed melting temperature or on the computed entropy. In the last case, the melting temperature is computed using the first formula of the *Melting temperature* section. We must enter at least one of the following ion concentrations : $[Na^+]$, $[K^+]$, $[Tris^+]$ or $[Mg^{2+}]$ and several ion corrections are proposed (see the reference table to have more information):

4.4.1 Sodium corrections

- *ahs01*

$$\Delta S = \Delta S_{[Na^+]=1\ M} + 0.847 \times (N-1) \times \log[Na^+]$$

  Where N is the length of the duplex.

- *kam71*

$$Tm = Tm_{[Na^+]=1\ M} + (7.95 - 3.06 \times \chi_{GC}) \times \ln[Na^+]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *marschdot*

$$Tm = Tm_{[Na^+]=1\ M} + (8.75 - 2.83 \times \chi_{GC}) \times \ln[Na^+]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *owc1904*

$$Tm = Tm_{[Na^+]=1\ M} + (-3.22 \times \chi_{GC} - 6.39) \times \ln[Na^+]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *owc2004*

$$\frac{1}{Tm} = \frac{1}{Tm_{[Na^+]=1\ M}} + (3.85 \times \chi_{GC} - 6.18) \times \frac{1}{100000} \times \ln[Na^+]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *owc2104*

$$Tm = Tm_{[Na^+]=1\ M} + (-4.62 \times \chi_{GC} + 4.52) \times \ln[Na^+]$$

$$-0.985 \times \ln[(Na^+)^2]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *owc2204*

$$\frac{1}{Tm} = \frac{1}{Tm_{[\mathrm{Na}^+]=1\ \mathrm{M}}} + (4.29 \times \chi_{GC} - 3.95)\times$$

$$\frac{1}{100000} \times \ln[\mathrm{Na}^+] + 9.40 \times \frac{1}{1000000} \times \ln[\mathrm{Na}^+]$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex.

- *san96*

$$12.5 \log[\mathrm{Na}^+]$$

- *san04*

$$\Delta S = \Delta S_{[\mathrm{Na}^+]=1\ \mathrm{M}} + 0.368 * (N-1) \times \ln[\mathrm{Na}^+]$$

  Where N is the length of the duplex.

- *schlif*

$$Tm = Tm_{[\mathrm{Na}^+]=1\ \mathrm{M}} + 16.6 \times \log[\mathrm{Na}^+]$$

- *tanna06*

$$\Delta S = \Delta S_{[\mathrm{Na}^+]=1\ \mathrm{M}} - 3.22 \times (N-1) \times g1$$

  Where N is the length of the duplex.

$$g1 = a1 + \frac{b1}{N}$$

$$a1 = -0.07 \times \ln[\mathrm{Na}^+] + 0.012 \times (\ln[\mathrm{Mg}^{2+}])^2$$
$$b1 = 0.013 \times (\ln[\mathrm{Mg}^{2+}])^2$$

  item *tanna07*

$$\Delta S = \Delta S_{[\mathrm{Na}^+]=1\ \mathrm{M}} - 3.22 \times (N-1) \times g1$$

  Where N is the length of the duplex.

$$g1 = a1 + \frac{b1}{N}$$

$$a1 = -0.075 \times \ln[\mathrm{Na}^+] + 0.012 \times (\ln[\mathrm{Mg}^{2+}])^2$$
$$b1 = 0.018 \times (\ln[\mathrm{Mg}^{2+}])^2$$

- *wet91*

$$Tm = Tm_{[\mathrm{Na}^+]=1\ \mathrm{M}} + 16.6 \log \frac{[\mathrm{Na}^+]}{1 + 0.7[\mathrm{Na}^+]} + 3.85$$

correction	contexts	Article
ahs01	DNA Na>0	von Ahsen et al ,2001 Clinical Chemistry, 47, 1956-1961.
kam71	DNA Na>=0.069 Na<=1.02	Frank-Kamenetskii et al. 1971 Biopolymers 10, 2623-2624.
marschdot	DNA Na>=0.069 Na<=1.02	Marmur, J., and Doty, P. (1962) J. Mol. Biol. 5, 109-118. Blake and Delcourt. (1998) Nucleic Acids Res. 26, 3323-3332 and corrigendum.
owc1904	DNA Na>0	Owczarzy et al.,2004 Biochemistry,43, 3537-3554.
owc2004	DNA Na>0	Owczarzy et al., 2004 Biochemistry, 43, 3537-3554.
owc2104	DNA Na>0	Owczarzy et al., 2004 Biochemistry, 43, 3537-3554.
owc2204	DNA Na>0	Owczarzy et al., 2004 Biochemistry, 43, 3537-3554.
owc2204	DNA Na>0	Owczarzy et al., 2004 Biochemistry,43, 3537-3554.
san96	DNA Na>=0.1	SantaLucia et al.(1996) Biochemistry 35 : 3555-3562
san04	DNA Na>=0.05 Na<=1.1 oligonucleotides inferior to 16 bases	Santalucia and Hicks (2004) Annu. Rev. Biophys. Biomol. Struct 33 : 415-440 John Santalucia, Jr., 1998 Proc. Natl. Acad. Sci. USA, 95, 1460-1465
schlif	DNA Na>=0.07 Na<=0.12	Schildkraut, C., and Lifson, S. (1965) Biopolymers 3, 195-208.
tanna06	DNA Na>=0.001 Na<=1	Tan et al. 2006, Biophysical Journal, 90, 1175-1190.
tanna07	RNA Na>=0.003 Na<=1	Tan et al, 2007 Biophysical Journal, 92, 3615-3632.
wet91	RNA, DNA and RNA/DNA Na>0	Wetmur 1991 Critical reviews in biochemistry and molecular biology, 26, 227-259

4.4.2 Magnesium corrections

- *owcmg08*

$$\frac{1}{Tm_{[\mathrm{Mg}^{2+}]}} = \frac{1}{Tm_{[\mathrm{Na}^{+}]=1\ \mathrm{M}}} + a - b(\ln[\mathrm{Mg}^{2+}]) + \chi_{GC}(c + d\ln[\mathrm{Mg}^{2+}]) + \frac{1}{2(Nbp-1)}$$

$$(-e + f\ln[\mathrm{Mg}^{2+}] + g(\ln[\mathrm{Mg}^{2+}])^2])$$

Where χ_{GC} is the frequence of GC base pairs in the duplex. *Nbp* is the number of base pairs and a, b, c, d, e, f, g fixed to :

a = 3.92 x 10^{-5}
b = 9.11 x 10^{-6}
c = 6.26 x 10^{-5}
d = 1.42 c 10^{-5}
e = 4.82 x 10^{-4}
f = 5.25 x 10^{-4}
g = 8.31 x 10^{-5}.

- *tanmg06*

$$\Delta S = \Delta S_{[\mathrm{Na}^{+}]=1\ \mathrm{M}} - 3.22 \times (N-1) \times g2$$

Where *N* is the length of the duplex.

$$g2 = a2 + \frac{b2}{(N)^2}$$

$$a2 = 0.02 \times \ln[\mathrm{Mg}^{2+}] + 0.0068 \times (\ln[\mathrm{Mg}^{2+}])^2$$
$$b2 = 1.18 \times \ln[\mathrm{Mg}^{2+}] + 0.344 \times (\ln[\mathrm{Mg}^{2+}])^2$$

item *tanmg07*

$$\Delta S = \Delta S_{[\mathrm{Na}^{+}]=1\ \mathrm{M}} - 3.22 \times (N-1) \times g2$$

Where *N* is the length of the duplex.

$$g2 = a2 + \frac{b2}{(N)^2}$$

$$a2 = \frac{-0.6}{N} + 0.025 \times \ln[\mathrm{Mg}^{2+}] + 0.0068 \times (\ln[\mathrm{Mg}^{2+}])^2$$
$$b2 = \ln[\mathrm{Mg}^{2+} + 0.38 \times (\ln[\mathrm{Mg}^{2+}])^2$$

correction	limits	Article
oxcmg08	DNA Mg>=0.0005 Mg<=0.6	Owczarzy et al.,2008 Biochemistry, 47, 5336-5353.
tanmg06	DNA Mg>=0.0001 Mg<=1 oligomer length superior to 6 base pairs	Tan et al. 2006 Biophysical Journal, 90, 1175-1190.
tanmg07	RNA Mg>=0.1 Mg<=0.3	Tan et al, 2007 Biophysical Journal, 92, 3615-3632.

4.4.3 Mixed Na Mg corrections

- *owcmix08*

$$\frac{1}{Tm_{[\mathrm{Mg}^{2+}]}} = \frac{1}{Tm_{[\mathrm{Na}^{+}]=1\ \mathrm{M}}} + a - b(\ln[\mathrm{Mg}^{2+}]) + \chi_{GC}(c + d\ln[\mathrm{Mg}^{2+}]) + \frac{1}{2(Nbp-1)}$$
$$(-e + f\ln[\mathrm{Mg}^{2+}] + g(\ln[\mathrm{Mg}^{2+}])^2])$$

  Where χ_{GC} is the frequence of GC base pairs in the duplex. *Nbp* is the number of base pairs. b, c, e, f are fixed as in the magnesium correction owcmg08.

$$a = 3.92 \cdot 10^{-5}(0.843 - 0.352[\mathrm{Mon}^{+}]^{0.5}\ln[\mathrm{Mon}^{+}])$$
$$d = 1.42 \cdot 10^{-5}(1.279 - 4.03 \cdot 10^{-3}\ln[\mathrm{Mon}^{+}] - 8.03 \cdot 10^{-3}\ln[\mathrm{Mon}^{+}]^2)$$
$$g = 8.31 \cdot 10^{-5}(0.486 - 0.258\ln[\mathrm{Mon}^{+}] + 5.25 \cdot 10^{-3}\ln[\mathrm{Mon}^{+}]^3$$

- *tanmix07*

$$\Delta S = \Delta S_{[\mathrm{Na}^{+}]=1\ \mathrm{M}} - 3.22 \times ((N-1) \times (x1 \times g1 + x2 \times g2) + g12))$$

  Where N is the length of the duplex.

$$g12 = -0.6 \times x1 \times x2 \times \ln[\mathrm{Na}^{+}] \times \frac{\ln[(\frac{1}{x1} - 1) \times \mathrm{Na}^{+}]}{N}$$

  See what is g1 and g2 in the sodium corrections tanna06 and tanna07 (g1) and magnesium corrections tanmg06 and tanmg07 (g2). Formula representing the fractional contribution of Na+ ions.

$$x1 = \frac{[\mathrm{Na}^{+}]}{(\mathrm{Na}^{+} + (\frac{8.1-32.4}{N}) \times (5.2 - \ln[\mathrm{Na}^{+}]) \times \mathrm{Mg}^{2+})}$$

  Formula representing the fractional contribution of Mg2+ ions.

$$x2 = 1 - x1$$

correction	limits	Article
oxcmix08	DNA Mg>=0.0005 Mg<=0.6 Na+K+Tris/2>0	Owczarzy et al.,2008 Biochemistry, 47, 5336-5353.
tanmix07	DNA and RNA Mg>=0.1 Mg<=0.3 Na+K+Tris/2>=0.1 Na+K+Tris/2<=0.3	Tan et al, 2007 Biophysical Journal, 92, 3615-3632.

If the user doesn't enter any ion correction, the algorithm from Owczarzy et al. (2008) will be used by default :

$$[\text{Mon}^+] = [\text{Na}^+] + [\text{k}^+] + [\text{Tris}^+]$$

Where $[\text{Tris}^+]$ is equal to half of total tris buffer concentration. (in the option -t, it is the Tris buffer concentration which is entered).

- if $[\text{Mon}^+]$ = 0, a default sodium correction will be used.
- if $[\text{Mg}^{2+}]^0.5$ / $[\text{Mon}^+]$ < 0.22, a default sodium correction is used. Monovalent ion influence is dominant, divalent cations can be disregarded.
- if $[\text{Mg}^{2+}]^0.5$ / $[\text{Mon}^+]$ >= 0.22 and $[\text{Mg}^{2+}]^0.5$ / $[\text{Mon}^+]$ < 6, a default mixed Na Mg correction is used. We can have a competitive DNA or RNA binding between monovalent and divalent cations.
- if $[\text{Mg}^{2+}]^0.5$ / $[\text{Mon}^+]$ >= 6, a default magnesium correction is used. Divalent cation influence is dominant, monovalent cations can be disregarded.

Moreover, if the user wants to use a sodium correction but also enters a potassium, Tris buffer and/or a magnesium concentration, a sodium equivalent concentration which takes into account the other ion concentrations is computed before applying the sodium correction. Several sodium equivalence ready to use are proposed by this program :

- *ahs01*

$$[\text{NaEq}^+] = [\text{Na}^+] + [\text{K}^+] + \frac{[\text{Tris}^+]}{2} + 3.79\sqrt{[\text{Mg}^{2+}] - [\text{dNTP}]}$$

- *mit96*

$$[\text{NaEq}^+] = [\text{Na}^+] + [\text{K}^+] + \frac{[\text{Tris}^+]}{2} + 4\sqrt{[\text{Mg}^{2+}] - [\text{dNTP}]}$$

- *pey00*

$$[\mathrm{NaEq}^{+}] = [\mathrm{Na}^{+}] + [\mathrm{K}^{+}] + \frac{[\mathrm{Tris}^{+}]}{2} + 3.3\sqrt{[\mathrm{Mg}^{2+}] - [\mathrm{dNTP}]}$$

For further information, see the referenced articles :

correction	limits	Article
ahs01	DNA	von Ahsen et al. 2001 Clinical Chemistry, 47, 1956-1961.
mit96	DNA	Mitsuhashi. et al, 1996 J. Clin. Lab. Anal, 10, 277-284.
pey00	DNA	Peyret, 2000 Ph.D Thesis, Section .5.4.2, 128, Wayne State University, Detroit, MI

4.5 Correction for the concentration of denaturing agents

MELTING is currently accurate when the hybridisation is performed at pH 7 ± 1, but some temperature corrections for the formamide and DMSO concentrations exists and can be applied. However, these corrections are rough approximations and results accuracy may be lost.

4.5.1 DMSO corrections, DMSO in %

- *ahs01*

$$Tm = Tm(DMSO = 0) - 0.75 \times DMSO$$

- *cul76*

$$Tm = Tm(DMSO = 0) - 0.5 \times DMSO$$

- *esc80*

$$Tm = Tm(DMSO = 0) - 0.675 \times DMSO$$

- *mus81*

$$Tm = Tm(DMSO = 0) - 0.6 \times DMSO$$

For further information, see the referenced articles :

correction	limits	Article
ahs01	DNA not tested with experimental values	von Ahsen et al. 2001 Clinical Chemistry, 47, 1956-1961.
cul76	DNA not tested with experimental values	Cullen et al., 1976 3, 49-62.
esc80	DNA not tested with experimental values	Escara et al., 1980 19, 1315-1327.
mus80	DNA not tested with experimental values	Musielski et al., 1981 Z allg Microbiol 1981; 21, 447-456.

4.5.2 formamide corrections

- *bla96*

$$Tm = Tm(formamide = 0) + (0.453 \times \chi_{GC} - 2.88) \times formamide$$

  Where χ_{GC} is the frequence of GC base pairs in the sequence. formamide is in mol/L

- *lincorr*

$$Tm = Tm(formamide = 0) - 0.65 \times formamide$$

  Where formamide is in %.

For further information, see the referenced articles :

correction	limits	Article
bla96	DNA not tested with experimental values formamide in mol/L	Blake and Delcourt, 1996 Vol. 24, No. 11 2095-2103
lincorr	DNA not tested with experimental in % values Formamide in %	McConaughy et al., 1969 Biochemistry 8, 3289-3295. Record, M.T., Jr, 1967 Biopolymers, 5, 975-992. Casey et al, 1977 Nucleic acids research, 4, 1539-1532. Hutton, 1977 Nucleic acids research, 4, 3537-3555.

4.6 Long sequences

It is important to realise that the nearest-neighbor approach has been established on small oligonucleotides. Therefore the use of MELTING in the non-approximative mode is really accurate only for relatively short sequences (Although if the sequences are two short, let's say < 6 bp, the influence of extremities becomes too important and the reliability decreases a lot). For long sequences an approximative mode has been designed. This mode is launched if the sequence length is higher than the value given by the option -T (the default threshold is 60 bp).

The melting temperature can be computed by one of the following formulas:

- *ahs01*

$$Tm = 80.4 + 0.345 \times \%GC + \log[\mathrm{Na}^+] \times (17.0 - 0.135 \times \%GC) - \frac{550}{size}$$

- *che93*

$$Tm = 69.3 + 0.41 \times \%GC - \frac{650}{size}$$

- *che93corr*

$$Tm = 69.3 + 0.41 \times \%GC - \frac{535}{size}$$

- *marschdot*

$$Tm = 81.5 + 16.6 \times \log[\mathrm{Na}^+] + 0.41 \times \%GC - \frac{675}{size}$$

- *owe69*

$$Tm = 87.16 + 0.345 \times \%GC + \log[\mathrm{Na}^+] \times (20.17 - 0.066 \times \%GC)$$

- *san98*

$$Tm = 77.1 + 11.7 \times \log[\mathrm{Na}^+] + 0.41 \times \%GC - \frac{528}{size}$$

- *wetdna91*

$$Tm = 81.5 + 16.6 \log \frac{[\mathrm{Na}^+]}{1 + 0.7[\mathrm{Na}^+]} + 0.41\%GC - \frac{500}{size} - \%Mismatching$$

- *wetrna91*

$$Tm = 78 + 16.6 \log \frac{[\mathrm{Na}^+]}{1 + 0.7[\mathrm{Na}^+]} + 0.7\%GC - \frac{500}{size} - \%Mismatching$$

- *wetdnarna91*

$$Tm = 67 + 16.6 \log \frac{[\mathrm{Na}^+]}{1 + 0.7[\mathrm{Na}^+]} + 0.8\%GC - \frac{500}{size} - \%Mismatching$$

For further information, see the referenced articles :

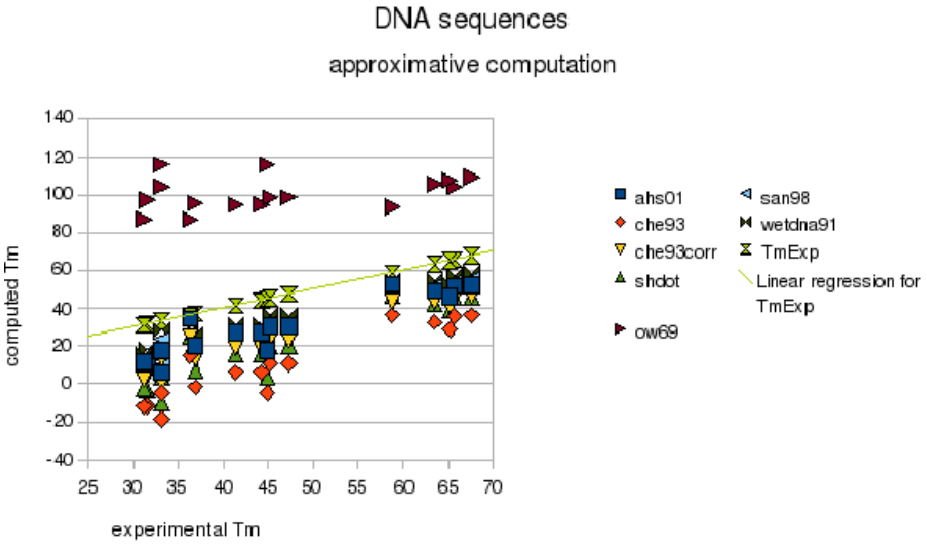

Figure 29: Comparison of experimental and computed Tm for various sets of DNA approximative formulas. $[\mathrm{Na}^+] = 1$ M

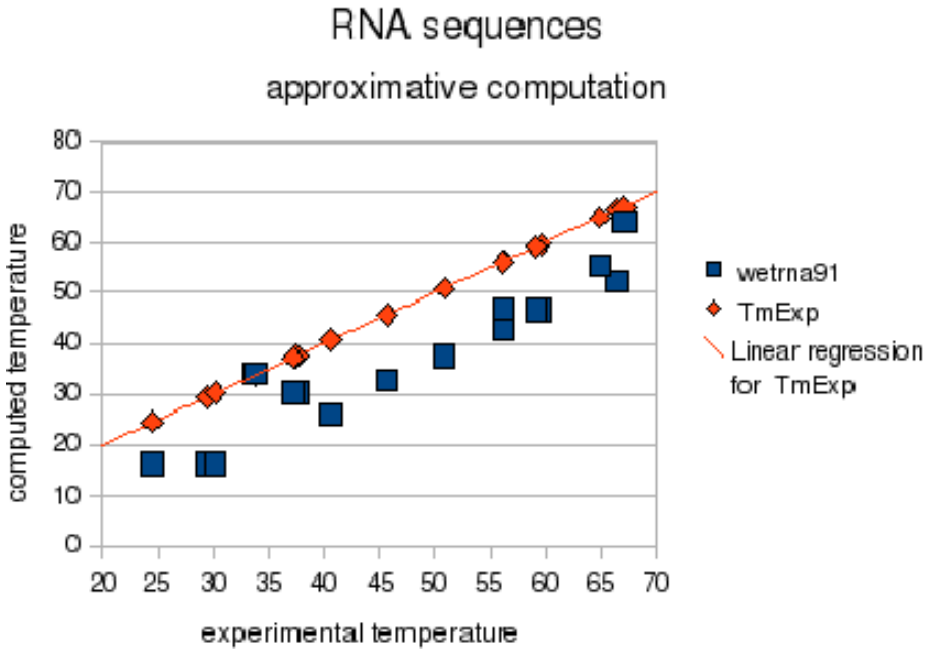

Figure 30: Comparison of experimental and computed Tm for various sets of RNA approximative formulas. $[Na^+] = 1$ M

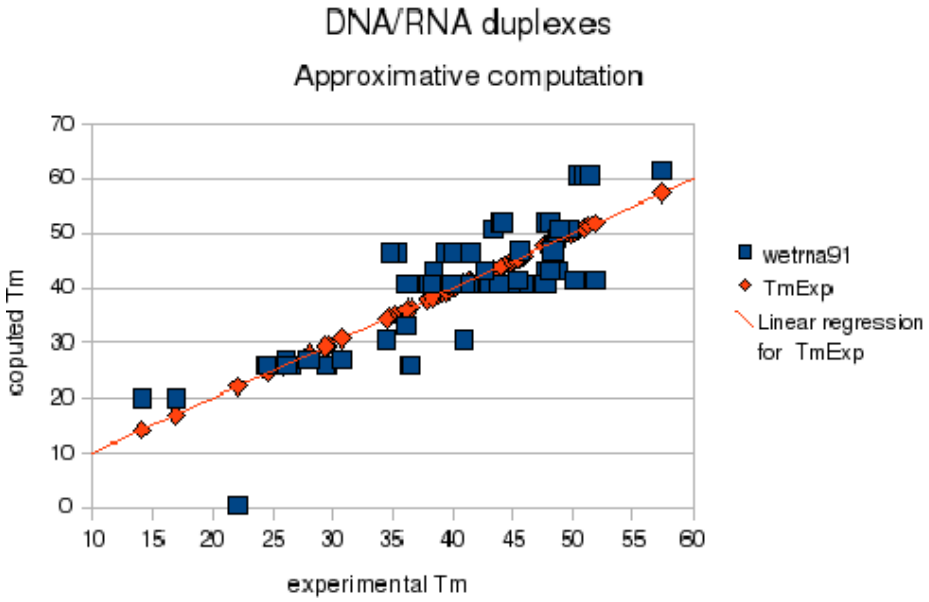

Figure 31: Comparison of experimental and computed Tm for various sets of DNARNA approximative formulas. $[Na^+] = 1$ M, [nucleic acid] $= 4 \cdot 10^{-4}$ M

formula	limits	Article
ahs01	DNA no mismatch	von Ahsen et al. 2001 Clinical Chemistry, 47, 1956-1961.
che93	DNA no mismatch Na=0 Mg=0.0015 Tris=0.01 K=0.05	Marmur et al., 1962 Journal of molecular biology, 5, 109-118. Chester N et al. 1993 Analytical Biochemistry, 209, 284-290.
che93corr	DNA no mismatch Na=0 Mg=0.0015 Tris=0.01 K=0.05	Marmur et al., 1962 Journal of molecular biology, 5, 109-118. Chester N et al. 1993 Analytical Biochemistry, 209, 284-290. Nicolas Von Ahsen et al. 2001 Clinical Chemistry, 47, 1956-1961.
marschdot	DNA no mismatch	Wetmur,1991 Critical reviews in biochemistry and molecular biology, 26, 227-259 Marmur et al., 1962 Journal of molecular biology, 5, 109-118. Chester et al., 1993 Analytical Biochemistry, 209, 284-290. Schildkraut et al., 1965 Biopolymers, 3, 95-110. Wahl et al., 1987 Methods Enzymol;152:399 - 407. Britten et al.,1974 Methods Enzymol ;29:363-418. Hall et al., 1980 J Mol Evol ;16:95-110.
owe69	DNA no mismatch	Owen et al., 1969 Biopolymers, 7:503-16. Frank-Kamenetskii,1971 Biopolymers;10:2623- 4. Blake, 1996 Encyclopedia of molecular biology and molecular medicine, Vol. 2., :1-19. Blake et al.,1998 Nucleic Acids Res;26:3323-32.
san98	DNA no mismatch	Santalucia, 1998 Proc Nacl Acad Sci USA 95, 1460-1465. von Ahsen et al. 2001, Clinical Chemistry, 47, 1956-1961.
wetdna91	DNA	Wetmur,1991, Critical reviews in biochemistry and molecular biology, 26, 227-259
wetrna91	RNA	Wetmur,1991, Critical reviews in biochemistry and molecular biology, 26, 227-259
wetdnarna91	DNA/RNA	Wetmur,1991, Critical reviews in biochemistry and molecular biology, 26, 227-259

References

[1] H.T. Allawi and J. SantaLucia. Thermodynamics and NMR of internal G-T mismatches in DNA. *Biochemistry*, 36:10581–10594, 1997.

[2] H.T. Allawi and J. SantaLucia. Nearest neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry*, 37:9435–9444, 1998.

[3] H.T. Allawi and J. SantaLucia. Nearest neighbor thermodynamics parameters for internal G.A mismatches in DNA. *Biochemistry*, 37:2170–2179, 1998.

[4] H.T. Allawi and J. SantaLucia. Thermodynamics of internal C.T mismatches in DNA. *Biochemistry*, 26:2694–2701, 1998.

[5] Hiroyuki Asanuma, Daijiro Matsunaga, and Makoto Komiyama. Clear-cut photoregulation of the formation and dissociation of the DNA duplex by modified oligonucleotide involving multiple azobenzenes. *Nucleic acids Symposium Series*, 49:35–36, 2005.

[6] Jaya Badhwar, Saradasri Karri, Cody K. Cass, Erica L. Wunderlich, and Brent M. Znosco. Thermodynamic characterization of RNA duplexes containing naturally occuring 1x2 nucleotide internal loops. *Biochemistry*, 46:14715–14724, 2007.

[7] R. D. Blake and S. G. Delcourt. Thermal stability of DNA. *Nucleic Acids Res*, 26:3323–3332 and corrigendum, 1998.

[8] R. D. Blake and Scott G. Delcourt. Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Research*, 24, 11:2095–2103, 1996.

[9] RD. Blake. Denaturation of DNA. *In: Meyers RA, ed. Encyclopediaof molecular biology and molecular medicine*, 2:1–19, 1996.

[10] Joshua Blose, M. Michelle L. Manni, Kelly A. Klapec, Yukiko Stranger-Jones, Allison C. Zyra, Vasiliy Sim, Chad A. Griffith, Jason D. Long, and Martin J. Serra. Non-nearest-neighbor dependence of stability for RNA bulge loops based on the complete set of Group I single nucleotide bulge loops. *Biochemistry*, 46:15123–15135, 2007.

[11] S. Bommarito, N. Peyret, and J. SantaLucia. Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res*, 28:1929–1934, 2000.

[12] K.J. Breslauer, R. Frank, H. Blöcker, and Marky L.A. Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci USA*, 83:3746–3750, 1986.

[13] RJ Britten, DE Graham, and BR Neufeld. Analysis of repeating DNA sequences by reassociation. *Methods Enzymol*, 29:363–418, 1974.

[14] N. Casey, J.and Davidson. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic acids research*, 4:1539–1532, 1977.

[15] Jonathan L. Chen, Abigael L. Dishler, Scott D. Kennedy, Ilyas Yildirim, Biao Liu, Douglas H. Turner, and Martin J. Serra. Testing the nearest neighbor model for canonical RNA base pairs: Revision of GU parameters. *Biochemistry*, 51:3508–3522, 2012.

[16] N Chester and DR Marshak. dimethyl sulfoxide-mediated primer tm reduction : a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Analytical Biochemistry*, 209:284–290, 1993.

[17] Br Cullen and Md Bick. Thermal denaturation of DNA from bromodeoxyuridine substitued cells. *Nucleic acids research*, 3:49–62, 1976.

[18] Amber R. Davis and Brent M. Znosko. Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry*, 46:13425–13436, 2007.

[19] Amber R. Davis and Brent M. Znosko. Thermodynamic characterization of naturally occurring RNA single mismatches with G-U nearest neighbors. *Biochemistry*, 47:10178–10187, 2008.

[20] JF Escara and Jr Hutton. Thermal stability and renaturation of DNA in dimethyl sulfoxidesolutions: acceleration of the renaturation rate. *Biopolymers*, 19:1315–1327, 1980.

[21] M. D. Frank-Kamenetskii. Simplification of the empirical relationship between melting temperature of DNA, its GC content and concentration of sodium ions in solution. *Biopolymers*, 10:2623–2624, 1971.

[22] S.M. Freier, R. Kierzek, J.A. Jaeger, N. Sugimoto, M.H. Caruthers, T. Neilson, and D.H. Turner. *Biochemistry*, 83:9373–9377, 1986.

[23] TJ Hall, JW Grula, EH Davidson, and RJ Britten. Evolution of sea urchin nonrepetitive DNA. *J Mol Evol*, 16:95–110, 1980.

[24] JR Hutton. Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic acids research*, 4:3537–3555, 1977.

[25] Junji Kawakami, Hiroyuki Kamiya, Kyohko Yasuda, Hiroyoshi Fujiki, Hiroshi Kasai, and Naoki Sugimoto. Thermodynamic stability of base pairs between 2-hydroxyadenine and incoming nucleotides as a determinant of nucleotide incorporation specificity during replication. *Nucleic acids research*, 29:3289–3296, 2001.

[26] Elzbieta Kierzek, David H. Mathews, Anna Ciesielska, Douglas H. Turner, and Ryszard Kierzek. Nearest neighbor parameters for Watson Crick complementary heteroduplexesformed between 2-O-methyl RNA and RNA oligonucleotides. *Nucleic acids research*, 34:3609–3614, 2006.

[27] Zhi John Lu, Douglas H. Turner, and David H. Mathews. A set of nearest neighbor parameters for predicting the enthalpy change of RNA secondary structure formation. *Nucleic Acids Research*, 34:4912–4924, 2006.

[28] Broda Magdalena, Elbieta Kierzek, Zofia Gdaniec, Tadeusz Kulinski, and Ryszard Kierzek. Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats. implication for prediction of RNA structure. *Biochemistry*, 44:10873–10882, 2005.

[29] J. Marmur and P. Doty. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.*, 5:109–118, 1962.

[30] David H. Mathews, Jeffrey Sabina, Michael Zucker, and Douglas H Turner. Expanded sequence dependence of thermodynamic parameters improves predictionof RNA secondary structure. *J. Mol. Biol*, 288:911–940, 1999.

[31] B.L. McConaughy, C.D. Laird, and B.I. McCarthy. Nucleic acid reassociation in formamide. *Biochemistry*, 8:3289–3295, 1969.

[32] Patricia M. McTigue, Raymond J. Peterson, and Jason D. Kahn. Sequence-dependent thermodynamic parameters for locked nucleic acid (LNA)DNA duplexformation. *Biochemistry*, 43:5388–5405, 2004.

[33] Stacy Miller, Laura E. Jones, Karen Giovannitti, Dan Piper, and Martin J. Serra. Thermodynamic analysis of 5 and 3 single- and 3 double-nucleotide overhangs neighboring wobble terminal base pairs. *Nucleic Acids research*, 36:5652–5659, 2008.

[34] M. Mitsuhashi. Technical report: Part 1. basic requirements for designing optimal oligonucleotide probe sequences. *J. Clin. Lab. Anal*, 10:277–284, 1996.

[35] H. Musielski, W Mann, R Laue, and S Michel. Influence of dimethylsulfoxide on transcription by bacteriophage T3-induced RNA polymerase. *Z allg Microbiol*, 21:447–456, 1981.

[36] Tatsuo Ohmichi, Shu-ichi Nakano, Daisuke Miyoshi, and Naoki Sugimoto. Long RNA danglingend has large energetic contribution to duplex stability. *J. Am. Chem. Soc.*, 124:10367–10372, 2002.

[37] Amanda S. O'toole, Stacy Miller, Nathan Haines, M. Coleen Zink, and Martin J Serra. Comprehensive thermodynamic analysis of 3' double-nucleotide overhangs neighboring Watson-Crick terminal base pairs. *Nucleic Acids research*, 34:3338–3344, 2006.

[38] Amanda S. O'toole, Stacy Miller, and Martin J Serra. Stability of 3' double nucleotide overhangs that model the 3' ends of siRNA. *RNA*, 11:512–516, 2005.

[39] R. Owczarzy, B.G. Moreira, Y. You, M.B. Behlke, and J.A. Walder. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry*, 47:5336–5353, 2008.

[40] Richard Owczarzy, Yong You, Christopher L. Groth, and Andrey V. Tataurov. Stability and mismatch discrimination of locked nucleic acid DNA duplexes. *Biochemistry*, 50:9352–9367, 2011.

[41] Richard Owczarzy, Yong You, Bernardo G. Moreira, Jeffrey A. Manthey, Lingyan Huang, Mark A. Behlke, and Joseph A. Walder. Effects of sodium ions on dna duplex oligomers: Improved predictions of melting temperatures. *Biochemistry*, 43:3537–3554, 2004.

[42] RJ Owen, LR Hill, and SP. Lapage. Determination of DNA base compositions from melting profiles in dilute buffers. *Biopolymers*, 7:503–516, 1969.

[43] N. Peyret. Prediction of nucleic acid hybridization : parameters and algorithms. *Ph.D Thesis 4.2, 128, Wayne State University, Detroit, MI.*, 5, 2000.

[44] N. Peyret, P.A. Seneviratne, H.T. Allawi, and J. SantaLucia. Nearest neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G and T.T mismatches. dependence and pH effects. *Biochemistry*, 38:3468–3477, 1999.

[45] M.T. Record, Jr. Electrostatic effects on polynucleotide transitions. I. Behavior at neutral pH. *Biopolymers*, 5:975–992, 1967.

[46] J. SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA*, 95:1460–1465, 1998.

[47] J. SantaLucia and Donald Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Struct*, 33:415–440, 2004.

[48] J. SantaLucia, Jr, H.T. Allawi, and P.A. Seneviratne. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, 35:3555–3562, 1996.

[49] C. Schildkraut and S. Lifson. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers*, 3:195–208, 1965.

[50] N. Sugimoto, M. Katoh, S. Nakano, T. Ohmichi, and M. Sasaki. RNA/DNA hybrid duplexes with identical nearest-neighbor base-pairs have identical stability. *FEBS Letters*, 354:74–78, 1994.

[51] N. Sugimoto, S. Nakano, M. Katoh, A. Matsumura, H. Nakamuta, T. Ohmichi, M. Yoneyama, and M. Sasaki. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*, 34:11211–11216, 1995.

[52] N. Sugimoto, S. Nakano, M. Yoneyama, and K. Honda. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nuc Acids Res*, 24:4501–4505, 1996.

[53] Zhi-Jie Tan and Shi-Jie Chen. Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophysical Journal*, 90:1175–1190, 2006.

[54] Zhi-Jie Tan and Shi-Jie Chen. Rna helix stability in mixed Na+/Mg2+ solutions. *Biophysical Journal*, 92:3615–3632, 2007.

[55] Fumiaki Tanaka, Atsushi Kameda, Masahito Yamamoto, and Azuma Ohuchi. Thermodynamic parameters based on a nearest neighbor model for DNA sequences with a single bulge loop. *Biochemistry*, 43:7143–7150, 2004.

[56] Nicolas von Ahsen, Carl T Wittwer, and Ekkehard Schutz. Oligonucleotide melting temperatures under PCR conditions : deoxynucleotide triphosphate and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clinical Chemistry*, 47:1956–1961, 2001.

[57] GM Wahl, SL Berger, and AR. Kimmel. Molecular hybridization of immobilized nucleic acids: theoretical concepts and practical considerations. *Methods Enzymol*, 152:399–407, 1987.

[58] N.E. Watkins and J. Santalucia, Jr. Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. *Nuc Acids Res*, 33:6258–6267, 2005.

[59] Norman E. Watkins, Jr, William J. Kennely, Mike J. Tsay, Astrid Tuin, Lara Swenson, Hyung-Ran Lee, Svetlana Morosyuk, Donald A. Hicks, and John Santalucia, Jr. Thermodynamic contributions of single internal rA.dA, rC.dC, rG.dG and rU.dT mismatches in RNA/DNA duplexes. *Nucleic Acids Research*, 39:1894–1902, 2011.

[60] J.G. Wetmur. DNA probes: applications of the principles of nucleic acid hybridization. *Crit Rev Biochem Mol Biol*, 26:227–259, 1991.

[61] D.J. Wright, J.L. Rice, D.M. Yanker, and B.M. Znosko. Nearest neighbor parameters for inosine-uridine pairs in RNA duplexes. *Biochemistry*, 46:4625–4634, 2007.

[62] T. Xia, J. SantaLucia, M.E. Burkard, R. Kierzek, S.J. Schroeder, X. Jiao, C. Cox, and D.H. Turner. Thermodynamics parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*, 37:14719–14735, 1998.

5 See Also

New versions and related material can be found at

`http://www.pasteur.fr/recherche/unites/neubiomol/meltinghome.html`

`https://sourceforge.net/projects/melting/`

`http://www.ebi.ac.uk/compneur-srv/melting/`

You can use MELTING through a web server at `http://bioweb.pasteur.fr/seqanal/interfaces/melting.html` `http://www.ebi.ac.uk/compneur-srv/melting/melt.php`

6 Copyright

7 Acknowledgements

Thanks to Richard Owczarzy for reporting typos in several thermodynamic parameters and reporting new public parameters. Nicolas Joly is an efficient and kind debugger and advisor. Catherine Letondal wrote the HTML interface to melting. Thanks to Nirav Merchant, Taejoon Kwon, Leo Schalkwyk, Mauro Petrillo, Andrew Thompson, Wong Chee Hong, Ivano Zara for their bug fixes and comments. Thanks to Richard Owczarzy for his magnesium correction. Thanks to Charles Plessy for the graphical interface files. Thanks to Daniel McGreal for the SOAP interface of MELTING at the EMBL-EBI. Thanks to Nicolas Rodriguez for the current web interface and finally thanks to the usenet helpers, particularly Olivier Dehon and Nicolas Chuche.

8 Authors

Nicolas Le Novère, Marine Dumousseau and William John Gowers,
EMBL-EBI, Wellcome-Trust Genome Campus Hinxton Cambridge, CB10 1SD, UK
n.lenovere@gmail.com

9 History

The Java version has been rewriten from the beginning. See the file ChangeLog for the changes of the versions 4 and more recent.