

Oligonucleotide Physical Parameters (updated July 30, 2009)

The Custom Oligonucleotide Synthesis (COS) Report accompanying each custom oligonucleotide contains extensive thermodynamic and physical data for each oligo. This technical sheet explains each of the values and the method used for arriving at a particular physical data. The figures may differ for the same oligo sequence by using different software applications, the reason being the selection of particular details and constants for arriving at the calculated value. An example is the salt concentration for calculating T_m ; Gene Link software uses 50 mM monovalent (Na^+ and K^+) salt concentration as a default, which reasonably approximates the total monovalent salt concentration commonly used in PCR reactions.

1. Size

The size is the length of the oligo. For unmodified oligos, this is equal to the number of bases in the oligonucleotide sequence. For modified bases, including 3'- and 5'-modifications, the software assumes these are base, and are represented as such in the size of the oligonucleotide. This will be correct if the modifications are actually modified bases (for example, inosine, dU, etc.), and *incorrect* if the modification is not a modified base (for example, phosphate, amino linker, etc.).

2. MW

MW is the molecular weight of the oligonucleotide. The molecular weight of a substance is the mass of one atom or molecule of that substance, in atomic mass units or Daltons (1 Dalton = exactly $1/12^{\text{th}}$ the mass of an atom of Carbon-12). One mole (6.02×10^{23} particles) of a substance (the molar mass) has a mass in grams equal to its molecular weight.

The molecular weight of an oligonucleotide is equal to the sum of the molecular weights of the individual ribo/deoxyribonucleotides, plus any modifications, it is composed of. For nucleotides, these values are as follows:

Deoxyribonucleotide	Molecular Wt (Da)	Ribonucleotide	Molecular Wt (Da)
dA	313.20	rA	329.20
dC	289.19	rC	305.19
dG	329.21	rG	345.21
dT	304.20	rU	306.17
dI	314.20	rI	330.20

Using these values, the MW of a 20mer DNA oligo with the following sequence:

5'-GAAAGCGTCTAGCCATGGCG-3'

is $(5 \times \text{MW}_A) + (7 \times \text{MW}_G) + (5 \times \text{MW}_C) + (3 \times \text{MW}_T) = (5 \times 313.20) + (7 \times 329.21) + (5 \times 289.19) + (3 \times 304.20) = 1566.00 + 2304.47 + 1445.95 + 912.60 = \mathbf{6229.0 \text{ Da (5 sf)}}$.

3. %GC



Oligonucleotide Physical Parameters (updated July 30, 2009)

The %GC is the percentage of G+C in the oligo. The calculation is the total number of G and C divided by the total number of bases in the oligo.

4. Oligonucleotide Melting Temperature (T_m)

The melting temperature (T_m) of an oligonucleotide duplex is defined as the temperature at which exactly 50% of the double-helix is denatured into single-stranded form (REF). The value of T_m provides a good estimation of the stability of a particular duplex, with a higher T_m implying greater stability. Using T_m to reliably estimate the stability of an oligonucleotide primer-template duplex is critically important for optimizing the reaction conditions (especially the annealing temperature) used in a variety of molecular biology applications, such as PCR, DNA sequencing, DNA/RNA target detection using labeled probes, and RNAi/anti-sense experiments (REF). For oligonucleotides ≤ 60 bases long, we calculate T_m using the salt-corrected nearest neighbor method (see Appendix for the method used prior to July 21, 2009). The utility of the nearest-neighbor method for calculating the T_m of oligonucleotides has been firmly established by several research groups (Breslauer *et al.*, 1986; Sugimoto *et al.*, 1995, 1996; Xia *et al.*, 1998; Santa-Lucia *et al.*, 1998). In combination with the recently published salt correction equations, which adjust nearest-neighbor T_m values based on the concentration of monovalent and divalent ions present (Owczarzy *et al.*, 2004, 2008), the resulting salt-corrected nearest-neighbor method is currently the best available for estimating T_m for oligonucleotides within this size range.

For oligonucleotides, the nearest neighbor formula for T_m is the following:

$$T_m(^{\circ}C) = \frac{\Delta H^0}{\Delta S^0 + R \ln C} - 273.15$$

where

R: ideal gas constant = 0.0019865 kcal/mol K

C: the concentration of oligonucleotide primer or probe strand (mol/L)

ΔH^0 (kcal/mole): ΔH^0 is the change in enthalpy, that is, the change in the amount of heat energy contained in a chemical or physical system due to a change (such as a chemical reaction or melting) in the system. Here ΔH^0 is calculated by summing together the enthalpy values of each of the dinucleotide nearest neighbor base pairs, and including the values for solvent-terminal base end effects and symmetry effects.

ΔS^0 (kcal/mole): ΔS^0 is the change in entropy, that is, the change in the amount of disorder a chemical or physical system exhibits due to a change in the system. Here ΔS^0 is calculated by summing together the entropy values of each of the dinucleotide nearest neighbor base pairs, and including values for solvent-terminal base end effects and symmetry effects.

For DNA oligos, we use the dinucleotide nearest-neighbor thermodynamic values reported by Allawi and SantaLucia (1997). For RNA oligos, we use those reported by Xia *et al.* (1998).



Oligonucleotide Physical Parameters (updated July 30, 2009)

Salt Correction ($^{\circ}\text{C}$): The nearest neighbor T_m value needs to be corrected for the actual concentration of salt present in the reaction mixture. This is because the nearest neighbor parameters for ΔH^0 and ΔS^0 were obtained from DNA melting experiments performed in 1M Na^+ buffer. This salt correction accounts for the effect on T_m of different concentrations of free monovalent (Owczarzy *et al.*, 2004) and Mg^{2+} (Owczarzy *et al.*, 2008) cations in solution. The nature of the effect is governed by the ratio R ,

$$R = \frac{\sqrt{[\text{Mg}^{2+}]}}{[\text{Na}^+]}$$

where $[\text{Mg}^{2+}]$ and $[\text{Na}^+]$ are the concentrations of Mg^{2+} and Na^+ , respectively.

If $R < 0.22$, monovalent cations exert a dominant effect on T_m . For this case, the following monovalent cation salt correction equation is used:

$$\frac{1}{T_m(\text{Na}^+)} = \frac{1}{T_m(1\text{M Na}^+)} + [(4.29 f_{GC} - 3.95) \ln[\text{Na}^+] + 0.940 \ln^2[\text{Na}^+]] \times 10^{-5}$$

where f_{GC} is the fraction of GC base pairs in the oligonucleotide.

If $0.22 < R < 6.0$, both monovalent and divalent cations affect the T_m , and so the divalent cation salt correction equation is used (Owczarzy, R. *et al.*, *Biochemistry*, **47**, 5336):

$$\begin{aligned} \frac{1}{T_m(\text{Mg}^{2+})} &= \frac{1}{T_m(1\text{M Na}^+)} + [(a - 0.911 \ln[\text{Mg}^{2+}]) + (f_{GC} \times (6.26 + d \ln[\text{Mg}^{2+}])) \\ &+ \frac{1}{2(N_{bp} - 1)} (-48.2 + 52.5 \ln[\text{Mg}^{2+}] + g \ln^2[\text{Mg}^{2+}])] \times 10^{-5} \end{aligned}$$

where N_{bp} is the number of base pairs and the three parameters a , d , and g show the following dependence on Na^+ concentration:

$$a = 3.92(0.843 - 0.352\sqrt{[\text{Na}^+]} \times \ln[\text{Na}^+])$$

$$d = 1.42(1.279 - 0.00403 \ln[\text{Na}^+] - 0.00803 \ln^2[\text{Na}^+])$$

$$g = 8.31(0.486 - 0.258 \ln[\text{Na}^+] + 0.00525 \ln^3[\text{Na}^+])$$

If $R \geq 6.0$, divalent cations exert a dominant effect on T_m , and so the divalent cation salt correction equation shown above is used, but with the three parameters a , d , and g having the constant values 3.92, 1.42, and 8.31, respectively.

For our COS Report, we use the following default values for calculating the reported oligonucleotide T_m :

$$C = 0.25 \mu\text{M}$$



Oligonucleotide Physical Parameters (updated July 30, 2009)

$$[\text{Na}^+] = 50. \text{ mM}$$

$$[\text{Mg}^{2+}] = 0 \text{ } \mu\text{M}$$

These values were chosen because they are fairly common primer and sodium ion concentrations used in routine PCR reactions. Using these values, the calculated T_m of a 20mer DNA oligo with the following sequence:

5'-GAAAGCGTCTAGCCATGGCG-3'

is

Salt-corrected nearest-neighbor method: $T_m = 60.59^\circ\text{C}$.

By comparison, the corresponding calculated T_m values using the salt-adjusted method and a default $[\text{Na}^+] = 0.1\text{M}$ (previously used by Gene Link prior to July 21, 2009—see Appendix for the equation) and the classical 4+2 Rule ($(4^\circ\text{C} \times (\#\text{G}+\text{C})) + (2^\circ\text{C} \times (\#\text{A}+\text{T}))$) rule are

Salt-adjusted method: $T_m = 59.50^\circ\text{C}$

4+2 Rule: $T_m = (4^\circ\text{C} \times (12)) + (2^\circ\text{C} \times (8)) = 64^\circ\text{C}$

5. Extinction Coefficient (EC_{260} or ϵ_{260})

The extinction coefficient (ϵ_λ) is defined as the fraction of light of a given wavelength (λ) lost to both scattering and absorption per unit distance traveled in a specific medium (for example, a dilute aqueous solution of DNA). Extinction coefficient is an intrinsic property of the medium. The relationship between the extinction coefficient and absorbance (A_λ) of a medium for a given wavelength of light is given by the Beer-Lambert Law:

$$A_\lambda = \epsilon_\lambda \times b \times c$$

where b = path length (the distance the light travels through the medium), c = concentration of the medium, and $A = -\log_{10}(I/I_0)$, where I_0 = initial intensity of light just before entering the medium, and I = final intensity of light just after passing through the medium.

The absorbance of a particular sample is typically determined using a spectrometer sensitive to the desired wavelength range (such as UV-Visible) and a sample cuvette having a path length of 1 cm. For a particular wavelength λ , the absorbance of a set of concentration standards of the medium is measured, and an absorbance vs. concentration calibration curve plotted, with the slope of the curve being equal to the extinction coefficient ϵ_λ . Once its extinction coefficient is known, the particular concentration of any sample of that medium can be calculated from the measured absorbance and the path length using the Beer-Lambert Law.



Oligonucleotide Physical Parameters (updated July 30, 2009)

When there is more than one absorbing species in a solution, the total absorbance is the sum of the individual absorbances of each species:

$$A_T = (\epsilon_1 c_1 + \epsilon_2 c_2 + \epsilon_3 c_3 + \dots) \times b$$

For an oligonucleotide, the absorbing species are not simply its individual bases, but actually its nearest-neighbor base pairs, since the individual bases are close enough together to significantly interact with each other. In other words, for an oligonucleotide, both its base composition and its base order effect the value of its extinction coefficient. Consequently, the extinction coefficient for an oligonucleotide of length n turns out to be the sum of the extinction coefficients of all the nearest-neighbor base pairs minus a factor that corrects for multiple countings of internal bases that occur while summing the nearest-neighbor pairs.

$$\epsilon_{\text{oligo}} = \sum_1^{n-1} \epsilon_{\text{nearestneighbor}} - \sum_2^{n-1} \epsilon_{\text{individualbase}}$$

For calculating the extinction coefficient of DNA/RNA oligos, we use previously published ϵ_{260} dinucleotide nearest-neighbor values (Warshaw and Tinoco, 1966), and individual ϵ_{260} nucleotide values (Cantor and Warshaw, 1970; Cantor, *et al.*, 1970), using $\text{mM}^{-1}\text{cm}^{-1}$ for units. For DNA, the values of $\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$) are

Nearest-neighbor pair	$\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$)	Nearest-neighbor pair	$\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$)
AA	27.4	AG	25.0
CA	21.2	CG	18.0
GA	25.2	GG	21.6
TA	23.4	TG	19.0
AC	21.2	AT	22.8
CC	14.6	CT	15.2
GC	17.6	GT	20.0
TC	16.2	TT	16.8

And the values of $\epsilon_{\text{individualbase}}$ ($\text{mM}^{-1}\text{cm}^{-1}$) are

$$A = 15.4; C = 7.4; G = 11.5; T = 8.7$$

For RNA, the values of $\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$) are

Nearest-neighbor pair	$\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$)	Nearest-neighbor pair	$\epsilon_{\text{nearestneighbor}}$ ($\text{mM}^{-1}\text{cm}^{-1}$)
AA	27.4	AG	25.0
CA	21.0	CG	17.8
GA	25.2	GG	21.6
UA	24.6	UG	20.0



Oligonucleotide Physical Parameters (updated July 30, 2009)

AC	21.0	AU	24.0
CC	14.2	CU	16.2
GC	17.4	GU	21.2
UC	17.2	UU	19.6

And the values of $\epsilon_{\text{individualbase}}$ ($\text{mM}^{-1}\text{cm}^{-1}$) are

A = 15.4; C = 7.2; G = 11.5; U = 9.9

Using these values, the ϵ_{260} of a 20mer DNA oligo with the following sequence:

5'-GAAAGCGTCTAGCCATGGCG-3'

is $(\epsilon_{GA} + \epsilon_{AA} + \epsilon_{AA} + \epsilon_{AG} + \epsilon_{GC} + \epsilon_{CG} + \epsilon_{GT} + \epsilon_{TC} + \epsilon_{CT} + \epsilon_{TA} + \epsilon_{AG} + \epsilon_{GC} + \epsilon_{CC} + \epsilon_{CA} + \epsilon_{AT} + \epsilon_{TG} + \epsilon_{GG} + \epsilon_{GC} + \epsilon_{CG}) - (\epsilon_A + \epsilon_A + \epsilon_A + \epsilon_G + \epsilon_C + \epsilon_G + \epsilon_T + \epsilon_C + \epsilon_T + \epsilon_A + \epsilon_G + \epsilon_C + \epsilon_C + \epsilon_A + \epsilon_T + \epsilon_G + \epsilon_G + \epsilon_C) = (25.2 + 27.4 + 27.4 + 25.0 + 17.6 + 18.0 + 20.0 + 16.2 + 15.2 + 23.4 + 25.0 + 17.6 + 14.6 + 21.2 + 22.8 + 19.0 + 21.6 + 17.6 + 18.0) - (15.4 + 15.4 + 15.4 + 11.5 + 7.4 + 11.5 + 8.7 + 7.4 + 8.7 + 15.4 + 11.5 + 7.4 + 7.4 + 15.4 + 8.7 + 11.5 + 11.5 + 7.4) = (392.8) - (197.6) = \mathbf{195.2 \text{ mM}^{-1}\text{cm}^{-1} (1 \text{ dp})}$

6. $\mu\text{g}/A_{260}$

The $\mu\text{g}/A_{260}$ is the concentration of the oligo in micrograms per OD at 260 nm. The formula is

$$\mu\text{g}/A_{260} = \frac{MW}{\epsilon_{260}}$$

7. nmol/A_{260}

The nmol/A_{260} is the concentration of the oligo in number of nanomoles (10^{-9} mol) per OD at 260 nm. The formula is

$$\text{nmol}/A_{260} = \frac{1000}{\epsilon_{260}}$$

Appendix: Comparison of Current Salt-Corrected Nearest-Neighbor Method with Previous Methods for Calculating T_m .

Prior to July 21, 2009, the value of the oligonucleotide T_m appearing in the COS Report was calculated using the salt-adjusted formula (Rychlik *et al.*, 1990):

$$T_m = 81.5 + 16.6 \log_{10}([\text{Na}^+] + [\text{K}^+]) + 0.41(\%GC) - (600/N)$$

where N = size of oligo, $[\text{Na}^+] + [\text{K}^+] = 0.1 \text{ M}$ (default).



Oligonucleotide Physical Parameters (updated July 30, 2009)

The following table show the comparable calculated T_m values (°C) of four oligonucleotides of different length for three methods: (a) salt-corrected nearest-neighbor (new method), (b) salt-adjusted method (old method), (c) 4+2 Rule (T_m = ((4°C x (#G+C)) + (2°C x (#A+T))). For a) and b), the total concentration of monovalent cation is set to the previous default value of 0.1 M, in order to establish a more direct comparison between these methods.

Sequence (5'-->3')	Size (mer)	T _m (new)	T _m (old)	T _m (4+2)
GAAAGCGTCTAGCCATGGCG	20	65.56	59.50	64
GAGGGTTTCCCTGCCACAGTCGAGCTCGAC	30	76.06	70.90	98
ATCTTTCACAAATTTTGTAATCCAGAGGTTGATTGTCGAC	40	66.83	64.30	108
TTGTTGTTGTTGTTTACTGGCCGTCGTTTACAGCTATGCTGTAAAACGA	50	73.09	69.20	140

The 4+2 Rule was originally developed to estimate the T_m of probe-target duplexes formed during membrane hybridization experiments. Thus, it is generally not recommended for estimating T_m of duplexes formed in PCR reactions, which is a solution-based method. Even for membrane experiments, the above results indicate that the 4+2 Rule is not a reliable way to estimate T_m for probe-target duplexes longer than about 20 mer.

The salt-corrected nearest-neighbor method consistently yields a T_m value that is a few degrees higher than that generated by the simpler salt-adjusted method.

References

Allawi, H.T., SantaLucia, J. (1997) Thermodynamics and NMR of Internal GT Mismatches in DNA. *Biochemistry* **36**: 10581-10594.

Breslauer K.J., Frank R., Blocker H., Marky L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc.Natl. Acad. Sci. USA* **83**: 3746-3750.

Cantor, C.R. and Warshaw, M.M. (1970). Oligonucleotide interactions. IV. Conformational differences between deoxy- and ribodinucleoside phosphates. *Biopolymers*, **9**: 1079-1103.

Cantor, C.R., Warshaw, M.M., and Shapiro, H. (1970) Oligonucleotide Interactions. III. Circular Dichroism Studies of the Conformation of Deoxyoligonucleotides. *Biopolymers* **9**: 1059-1077.

Owczarzy R., You Y., Moreira B.G., Manthey J.A., Huang L., Behlke M.A., Walder J.A. (2004) Effects of Sodium Ions on DNA Duplex Oligomers: Improved Predictions of Melting Temperatures. *Biochemistry* **43**: 3537-3554.

Owczarzy, R., Moreira, B.G., You, Y., Behlke, M.A., Walder, J.A. (2008) Predicting Stability of DNA Duplexes Containing Magnesium and Monovalent Cations. *Biochemistry* **47**: 5336-5353.



Oligonucleotide Physical Parameters (updated July 30, 2009)

Rychlik W., Spencer, W.J., Rhoads R.E. (1990) Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Res.* **18**: 6409-6412 (adapted from Baldino, F., Chesselet, M.F., and Lewis, M.E. (1989) High-resolution *in situ* hybridization histochemistry. *Meth. Enzymol.* **168**:761-777).

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

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

Sugimoto N., Nakano S., Yoneyama M., Honda K. (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.* **24**: 4501-4505.

Warshaw, M.M. and Tinoco, I. (1966) Optical properties of sixteen dinucleoside phosphates. *J. Mol. Biol.*, **20**: 29-38.

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

