SMR/111 - 17

SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

THERMODYNAMICS OF NUCLEIC ACIDS SOLUTIONS (I)

F. QUADRIFOGLIO

Istituto di Chimica
Università di Trieste
Trieste
Italy

THERMODYNAMICS OF NUCLEIC ACIDS SOLUTIONS (I)

F. Quadrifoglio - Istituto di Chimica - Università di Trieste

In order to establish the nature of the forces involved in the maintenance
of nucleic acid structure and stability a great number of physico-chemical
studies of conformational transitions have been carried out. The order-disorder
transitions of nucleic acids are characterized by large positive enthalpy and
entropy changes. However interpretation of these thermodynamic changes is not
easy. Hydrophobic forces, hydrogen bonding, base-stacking interactions, elec-
trostatic interactions and conformational entropy are the main contributions
to the observed phenomenon but their relative importance is difficult to assess.
One possible way is to study simple systems in order to derive the single con-
tributions.

a) Hydrophobic forces

The hydrophobic forces, as described by Kauzmann, are measured as the ten-
dency of non-polar residues to associate with one another in aqueous solution.
The primary driving force for the association is the large positive entropy
change as opposed to the negative one connected with the unfavorable interac-
tion with water and the concomitant promotion of solvent structure. The expo-
sition of hydrophobic moieties to water, in addition to a negative ΔS, should
result also in a positive free energy change ($\Delta G > 0$), in a close to zero or
negative enthalpy change ($\Delta H \leq 0$) and in a positive change of the heat capacity
($\Delta C_p > 0$).

Thus these thermodynamic quantities can be used as diagnostic tests for the
participation of hydrophobic forces in the stabilization of ordered structures
exhibited by the nucleic acids in aqueous solution. The model reaction used
to thermodynamically mimic hydrophobic interactions is the transfer of a non-
polar molecule from a nonaqueous solvent to water. For a large variety of mo-
lecules for which this approach has been used the choice of the nonaqeous sol-
vent is not so important, in the sense that the expected changes in the thermo-
dynamic quantities are generally obtained regardless of the nature of the non-
aqueous solvent. In the case of purines and pyrimidines one has to be careful
in this choice as bases possess polar moieties which can make specific inter-
actions with water (for example hydrogen bonding). So the use of alcohols as

nonaqueous solvents should guarantee that some polar interactions present in
water can be maintained. However, since these polar interactions can be quanti-
tatively different in water and in nonaqueous solvents, enthalpy and free ener-
gy changes are unreliable quantities from which to derive the presence of hydro-
phobic interactions in water. More reliable conclusions can be derived from the
entropy and heat capacity changes.

The experimental approach is that of measuring the solubility of the sub-
stance (base) in water and in nonaqueous solvent as a function of the tempera-
ture. Application of the general thermodynamic relationships

$$\Delta G = -RT \ln x(T) \tag{1}$$

$$\Delta H = RT_1T_2 \ln x(T_2) - \ln x(T_1) \tag{2}$$

$$\Delta S = \frac{\Delta H - \Delta G}{T} \tag{3}$$

yields the free energy, enthalpy and entropy of solution in a given solvent,
where $x(T)$ is the solubility of the base expressed in units of mole fraction.
Equation (2) implies the constancy of ΔH with temperature. As this is not
generally the case ($\Delta C_p \neq 0$) a better choice is to measure directly the solu-
tion enthalpy change with a calorimeter at different concentrations and extra-
polating to infinite dilution to cancel the base-base interaction effect.

The change in the thermodynamic quantities connected to the transfer process
from nonaqueous solvent to water can be obtained as difference between the
figures derived from the solubilities in the respective solvents. Before making
this difference one has to be sure that the crystalline solutes in equilibrium
with both saturated solutions are thermodynamically equivalent. One assumption
is that the concentration used in equations (1) and (2) coincides with thermo-
dynamic activity (which is practically true for bases due to low solubility in
both solvents).

Table 1 summarizes the transfer parameters from organic solvents to water
obtained for the bases (with the exception of guanine which is sparingly so-
luble).

If one applies the above mentioned rules for hydrophobic interactions only
thymine seems to agree with the signs expected. However the small change in
entropy of transfer and the zero change in ΔC_{p_t} argue against any major parti-
cipation of a hydrophobic effect even in the case of thymine.

Compound	Solvent from	to	ΔG_t J/mol	ΔH_t J/mol	ΔS_t J/mol K	Δc_{p_t} J/mol K
Adenine	CH_3OH	H_2O	2134	10920	35.1	
Cytosine	"	"	-1225	14477	50.6	
Uracil	"	"	-649	1423	6.7	
Thymine	"	"	1121	-2008	-11.2	0
Adenine	$CHCl_3$	"	-14686	6192	67.4	
Adenine	6M CH_3OH	"	3975	4602	2.1	
Thymine	"	"	1966	2510	2.1	
Adenine	6M 1-Pro- panol	"	5439	9623	13.4	
Thymine	"	"	3013	-418	-10.9	
Adenine	6M Ethylene glycol	"	2761	5439	8.8	
Thymine	"	"	962	2510	5	

b) Stacking interactions

Purinic and pyrimidinic bases, as the corresponding nucleosides and nucleo-
tides, form complexes of various degrees of polymerization in water depending
upon the total concentration. NMR and UV absorption data suggest a stacked
configuration for such complexes.

Thermodynamic informations can be conveniently derived from the spectral
data obtained at different concentration and temperature with the use of class-
ical thermodynamic relationships. However due to practical reasons a better
way to derive thermodynamic quantities is that of using direct calorimetric
measurements.

The basic assumptions for the treatment of the data are the following
i) all species behave ideally. This is certainly true at increasing dilution.
    If molecular species are uncharged (as the bases and the nucleosides at
    neutral pH) they obey dilute solution laws in a wide range of concentration.
ii) association of the monomeric solute proceeds to an indefinite degree (or
stops when dimers are formed)
iii) all association steps are characterized by the same association constant
and association enthalpy.

With these assumptions the various equilibria are characterized by equations

of the form

$$A + A_{n-1} = A_n \; ; \quad K_n = \frac{[A_n]}{[A_{n-1}][A]} \tag{4}$$
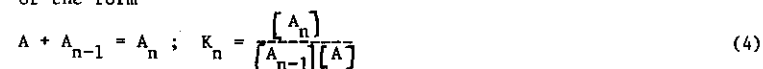
with equilibrium constant $K_n = K$ and enthalpy change $\Delta H_n^\circ = \Delta H^\circ$. If only a
dimer is formed n = 2. In the other cases n can vary from 2 on.

The heat of infinite dilution ($-m\varphi_L$) is defined as the heat obtained when
an $\underline{m}$ molal solution (1 Kg of solvent) is diluted with an infinite amount of
solvent. Thus (for association from m = 0 to m)

$$m\varphi_L = [A_2]\Delta H^\circ + [A_3]2\Delta H^\circ + \ldots \tag{5}$$

by using K

$$m\varphi_L = K[A]^2\Delta H^\circ + 2K^2[A]^3\Delta H^\circ + \ldots \tag{6}$$

The molality of the solution is given by

$$m = [A] + 2[A_2] + 3[A_3] + \ldots$$
$$= [A] + 2K[A]^2 + 3K^2[A]^3 + \ldots \tag{7}$$

Combining equations (6) and (7) yields the following

$$\varphi_L = \Delta H^\circ - (\Delta H^\circ/K)^{\frac{1}{2}}(\varphi_L/m)^{\frac{1}{2}} \tag{8}$$

By plotting φ_L vs $(\varphi_L/m)^{\frac{1}{2}}$ one should obtain a straight line. From the value
of the slope and intercept both K and ΔH° can be determined.
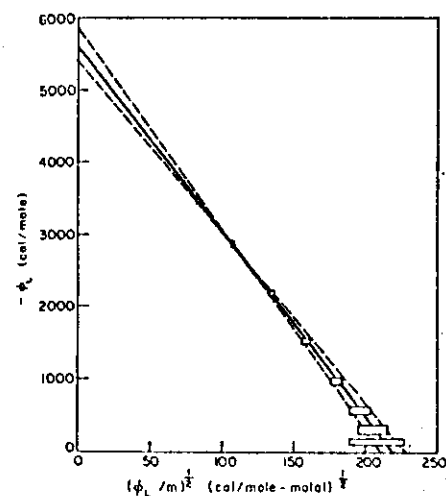

Figure 1.

Fig.1 shows the results obtained with
6-methylpurine by Stoesser and Gill.
If only dimerization occurs then
$$\varphi_L = \Delta H^\circ/2 - \frac{1}{2}(\Delta H^\circ/K)^{\frac{1}{2}}(\varphi_L/m)^{\frac{1}{2}} \tag{9}$$
so that both slope and intercept are
halved.

As it is not possible to distinguish
between the two mechanisms using calori-
metric data, auxiliary informations are
needed, such as those given by osmotic
coefficient determination or by ultra-
centrifugation.

Thermodynamic quantities obtained with
this approach are reported in Table 2.

Table 2

Thermodynamic parameters for a single step in the dissociation of purine and pyrimidine derivatives and analogous compounds

Compound	ΔH^o kJ/mol	ΔG^o kJ/mol	ΔS^o J/mol.K	K, M
Purine	17.6 ± 0.8	1.84	54	0.48
6-Methyl-purine	25.1 ± 1.7	4.69	67	0.15
"	23.3 ± 0.2	5.06 ± 0.2	61	0.13
6-Dimethyl-amino-purine	38.1 ± 0.4	10.21 ± 0.04	94	0.016
N^6,N^9-Dimethyl-adenine	36.4 ± 6.3	9.47	90	0.022
Purine riboside	10.5 ± 0.4	1.6	29	0.5
Adenosine	40.2	3.7	121	0.22
Deoxyadenosine	15.5 ± 2.5	6.3	29	0.08
Caffeine	14.2 ± 0.8	6.3	25	0.08
Cytidine	11.8 ± 0.4	0.33	42	1.14
Uridine	11.3 ± 0.4	1.2	42	1.64
N^6-Dimethyl-adenosine	26.8	7.95	63.2	0.042
Flavin mono-nucleotide	15.6	14.9	2.26	0.025

Inspection of Table 2 reveals a large contribution of enthalpy to the association of bases which is also dependent upon the chemical nature of the base. The sign of ΔS is in agreement with the presence of hydrophobic forces but the question of a real contribution of such interactions cannot be answered positively only from these measurements.

c) Oligonucleotides studies

Oligonucleotides can be valid models for the study of the structural and thermodynamic properties of natural polynucleotides. For many aspects they are better models than the same synthetic polynucleotides. As a matter of fact their optical and NMR spectra are simpler to analyze. In addition their study can be carried out as a function of the concentration in order to distinguish between inter- and intramolecular interactions. Another major advantage is that, since few years, it is possible to synthetize oligonucleotides at will so that the study of the effects of different base sequences and of the presence of hairpins, loops, bulges in the structure can be analyzed. Of course there are some drawbacks, the major ones being the absence in the oligonucleotides of long range interactions (which are important in long chains) and the presence of particular end effects which in long chains are neglectable.

The simplest oligonucleotide is a dinucleotide NpN which can be assumed to exist in only two limit conformations: fully stacked and unstacked forms.

The amount of the two forms is established by an equilibrium constant of stacking

$$K_{st} = \frac{s}{u}; \quad \Delta G^o_{st} = -RT\ln K_{st} = \Delta H^o_{st} - T\Delta S^o_{st} \tag{10}$$

s and u mean the fraction of stacked and unstacked forms, respectively.

If one knows the dependence of K_{st} with temperature one can obtain ΔH^o_{st} and ΔS^o_{st}. The usual procedure is to monitor any parameter which varies in a definite manner with temperature and to plot it as a function of temperature. The result is generally termed "melting curve" (fig.2).
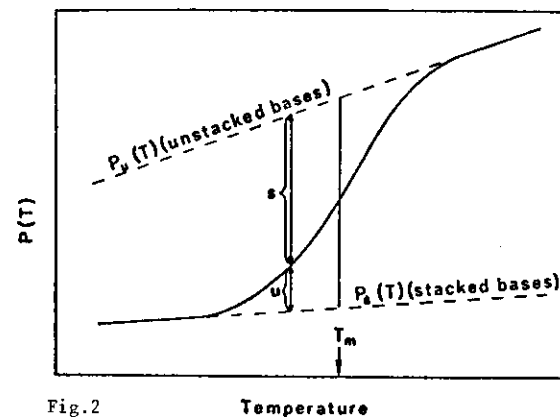

Fig.2   Temperature

P is the property under investigation. T_m is called melting temperature and is defined as the temperature at which $\Delta G^o_{st} = 0$ ($K_{st} =1$). Therefore

$$T_m = \Delta H^o_{st}/\Delta S^o_{st} \tag{11}$$

The difficulties of using this approach for dinucleotides reside in the fact that often the two limiting conformations are attained at very different temperatures so that it is difficult to draw precise base-lines. Some of the results obtained with this approach are shown in Table 3. The high spread of ΔH^o_{st} and ΔS^o_{st} values obtained with ApA (using different optical techniques) demonstrates the high difficulties in choosing the low temperature limiting value of P_s

but experimental evidences have been accumulated showing that a two-state model is not a fully correct description of dimer stacking. Unfortunately no calorimetric results are available for such compounds.

Table 3
Thermodynamic parameters of unstacking of dinucleoside phosphates derived on the basis of the two state model.

Compound	ΔH° kJ/mol	ΔS° J/mol K
ApA	22.2	84
"	27.2	
"	33.5	
"	35.6	
"	39.3	
"	41.8	
"	35.6	119
ApG	20.1	75
ApC	25.9	92
"	25.5	88
ApU	35.1	134
"	28	100
GpA	23.4	84
"	25.5	92
GpC	32.6	117
GpU	28.5	105
CpA	30.5	113
"	29.3	100
CpG	20.1	71
CpC	35.7	124
"	28.9	105
"	31.4	105
CpU	32.6	117
"	28.5	100
UpA	21.3	88
UpG	25.1	96
UpC	25.9	92
UpU	32.6	121

The results shown in Table 3 have been obtained by using optical techniques and therefore low concentrations of material. At these low concentrations no evidence of base-pairing (in addition to base-stacking) exists. To achieve base-pairing in the case of dinucleotides, concentrations of the order of 10^{-1} M are often necessary. At these concentrations a valuable technique to study association of complementary dinucleotides is NMR. While some of the protons show an upfield shift due to ring current effects caused by stacking, few protons show a marked downfield shift due to hydrogen bond formation. The latter shift is very sensitive to concentration and can be used to study the thermodynamic parameters. The analysis of the experimental data can be carried out according to the equation

$$\delta_{obs} = \delta_M + (\delta_{M_2} - \delta_M)\left[(1 + 4KM_o - \sqrt{1 + 8KM_o})/4KM_o\right] \qquad (12)$$

where δ_{obs} is the chemical shift observed at concentration M_o, δ_M is the chemical shift corresponding to the single-strand, δ_{M_2} is the chemical shift corresponding to the duplex and K is the equilibrium constant for the formation of the duplex.

Table 4 shows the results obtained with different dinucleotides. From the data one can infer that A-T and A-U pairs are weaker than G-C pairs.

Table 4
Formation constants of dinucleotide complexes as determined by NMR

Structure	Temperature (°C)	K (M⁻¹)
pdGpdC·pdGpdC	2	7.8 ± 0.7
pdCpdG·pdCpdG	2	3.0 ± 0.5
pdGpdG·pdCpdC	2	9.1 ± 1.1
pdGpdC·pdGpdC	25	< 2
pdGpdT·pdApdC	2	< 3
CpG·CpG	4	14.0 ± 2.5
GpU·ApC	1	≪ 14
GpC·GpC	4	≳ 14

Oligonucleotides containing more than two bases can associate (if complementary) to give miniduplexes also at low concentrations so that optical studies can be easily performed.

In the case of self-complementary oligonucleotides the association constant K_A can be expressed as

$$K_A = \frac{\alpha/2}{(1 - \alpha)^2 C_T} \qquad (13)$$

where α represents the fraction of stacked bases and C_T the concentration of the oligonucleotide. In the case of nonidentical but complementary oligonucleotides

$$K_A = \frac{2\alpha}{(1 - \alpha)^2 C_T} \qquad (14)$$

if concentration of both oligonucleotides is $C_T/2$.

Derivation of van't Hoff enthalpy from (13) and (14) yields

$$\Delta H_{vH} = 6RT_m^2 \left(\frac{\partial \alpha}{\partial T}\right)_{T_m} \qquad (15)$$

Another method to derive ΔH_{vH} is to measure T_m as a function of C_T for complementary oligonucleotides according to the relation

$$\frac{1}{T_m} = \frac{R}{\Delta H_{vH}} \ln C_T + \frac{\Delta S}{\Delta H_{vH}}$$

(16)

Of course one can obtain ΔH_{vH} by plotting, in the usual way, $\ln K_A$ (calculated using measured α values at each temperature) as a function of $1/T$.

All these methods should give the same enthalpy value. In fact they do not in most cases. This is probably due to an inaccurate determination of base-lines before and after the temperature range in which transition occurs (fig.2) more than to an inadequacy of the all-or-none model used. When the base-lines are correctly taken into account a good agreement between the ΔH values derived with different methods is obtained and, more important, a satisfactory correspondence between van't Hoff enthalpy and calorimetric enthalpy is attained. An analysis of the data obtained with the above methods suggests that, in case of disagreement, the more reliable figures are those based on equation (16) because of its much lower sensitivity to base-line manipulation.

Most of the data collected have been obtained with oligoribonucleotides. As an example Table 5 shows the data obtained by Borer et al (1974). The results indicate a strong chain length and base composition dependence of helix stability. Also a strong effect of base sequence on stability is quite evident, which means that the contribution of each base-pair to the stability of the helix depends heavily on its nearest neighbors. If one consider only Watson-Crick base-pairs there are only ten possible base-pairs doublets. If one includes end effects there are two more contributions. The data in Table 5 can be used to obtain the thermodynamic contribution of each of the nearest-neighbors doublets to the stability of the model helix.

When a double-stranded helix is formed the process can be visualized as a two steps reaction. The first step is called nucleation and the second one propagation. Nucleation has a thermodynamic constant κ which takes into account the entropy change of bringing the two strands together and the energy of first base-pair formation. Propagation has a thermodynamic constant (s) which contains the enthalpy and entropy of stacking a base-pair on top of a preformed one in addition to base-pair energy. One expects that propagation constant has a higher value than a nucleation constant.

If one assumes that ΔH°_{nucl} is zero (which means assuming that hydrogen bonding between the bases in water have an enthalpy change close to zero) then

TABLE 5

Thermodynamic parameters for double-stranded oligonucleotide helices

Molecule†	T_m (°C) at:‡ 10 μM	100 μM	Comcn (μM) at T_m = 25°C	Total ΔH° (kcal/mol)	ΔH° (kcal/mol)	Total ΔG° (kcal/mol) at 25°C	Total ΔS° (kcal/mol deg.)
Helix length = 6							
A₂CG + CGU₂	−13·9	−1·3	13500	−26	−6	−3·4	−0·0766
U₂CGA₂	1·6	11·3	2040	−37		−3·7	−0·112
A₂CGU₂	10·8	22·1	174	−34		−6·1	−0·0969
A₂C₂ + C₂U₄	14·0	22·8	347	−44		−6·6	−0·129
A₂CGU₂	19·6	28·3	40·7	−46		−8·0	−0·134
Helix length = 7							
A₂CU₂ + A₂GU₃	7·1	16·4	1610	−41	−7	−4·7	−0·122
Helix length = 8							
A₄U₄	6·3	11·6	4070	−51	−8	−3·5	−0·160
A₂CU₃ + A₃GU₄	14·6	22·1	436	−52		−5·4	−0·166
A₃CGU₃	28·3	35·1	2·76	−66		−7·6	−0·196
A₃GCU₃	34·7	42·3	0·794	−63		−8·3	−0·160
Helix length = 9							
A₄CU₄ + A₄GU₄	19·1	26·0	135	−68	−9	−6·1	−0·174
Helix length = 10							
A₅U₅	18·2	24·3	129	−67	−10	−6·3	−0·207
A₄UAU₄	21·9	27·7	32·5	−72		−8·1	−0·221
A₄CU₄ + A₄GU₄	25·8	32·6	14·1	−63		−7·4	−0·167
A₄CGU₄	37·0	41·9	0·977	−76		−9·6	−0·223
A₄GCU₄	40·7	46·8	0·0123	−78		−10·8	−0·225
Helix length = 11							
A₅CU₅ + A₅GU₅	30·6	38·4	1·17	−89	−11	−8·9	−0·269
Helix length = 12							
A₆U₆	25·4	31·6	4·79	−84	−12	−7·3	−0·257
Helix length = 14							
A₇U₇	36·2	30·3	0·0210	−107	−14	−10·4	−0·324

† None of the oligomers in this work bear terminal phosphates and the phosphodiester linkage has been omitted. Thus A₂U₂ is identical to (Ap)₄(Up)₃.U.
‡ The T_m values for the non self-complementary helices refer to concentrations of 20 μM and 200 μM for each oligomer. ΔH° is obtained from $\Delta H^\circ = R\,d\ln(1/T_m)$. For self-complementary molecules $\Delta G^\circ = \Delta H^\circ - RT_m \ln c_t$ for non self-complementary molecules $\Delta G^\circ = \Delta H^\circ - RT_m \ln (c_t/4)$, c_t is the total concentration in mols of the oligomers.

$$\Delta G^\circ_{nucl} = -RT \ln k = -T \Delta S^\circ_{nucl} \qquad (17)$$

Superscript zero means standard conditions of one mole/liter strand concentration (initiation is very dependent on strand concentration).

ΔG°_{nucl} depends on whether the helix nucleates with AU or GC base pair. The values estimated for the two nucleations are $k_{CG} = 2.5 \times 10^{-4}$ and $k_{AU} = 4 \times 10^{-5}$ which correspond to $\Delta G^\circ_{nucl,AU} = +6$ Kcal/mol oligomer and $\Delta G^\circ_{nucl,GC} = +5$ Kcal/mol oligomer.

After nucleation had occurred there will be a corresponding ΔG_{gr} of propagation of each base-pair which depends on the sequence. These steps are practically independent of concentration. Thus

$$\Delta G_T = \Delta G^\circ_{nucl} + \sum \Delta G_{gr} \qquad (18)$$

In principle the data shown in Table 5 are sufficient to derive all the 10 nearest neighbor thermodynamic contributions. However, due to errors associated with doublets which occurr rarely in the oligonucleotides examined, the authors preferred to assume that AU/AU and UA/UA stacking have the same thermodynamic parameters. The same has been assumed for AC/GU, CA/UG, AG/CU and GA/UC stackings.

With these assumptions the resolution of 19 linear equations with six unknowns with best least-squares gave the results shown in Table 6.

Of course the study of a greater number of oligomers with different length and sequence would have avoided the assumptions and would have allowed the independent determination of all possible nearest-neighbors combination. However, also with these limitations, the results shown in Table 6 allowed Borer et al. to predict almost correctly the melting temperature of some duplexes not used in obtaining the data of Table 6.

Other limitations inherent to the method are:
a) all the duplexes studied have been assumed to possess the same conformation
b) the end effects have been neglected
c) the values of ΔH and ΔS have been assumed independent of temperature.

d) Imperfect nucleic acid helices

The simple model of perfectly aligned nucleic acid helices does not describe all the naturally occurring structures. Defects characterized by internal loops bulges or hairpin structures have been demonstrated in many natural DNAs and RNAs. Thus a corre ct interpretation of the thermodynamic parameters must take into account the contribution of such defects.

Table 6
Thermodynamics of adding a base pair to a double-stranded helix

Reaction	ΔH° (kcal)	ΔS° (kcal deg^{-1})	ΔG° (kcal) at 25°C
	−8.2	−0.0235	−1.2
	−6.5	−0.0164	−1.6
	−5.9	−0.0127	−2.1
	−13.0	−0.0335	−3.0
	−14.7	−0.0349	−4.3
	−13.7	−0.0298	−4.8
	—	—	0.3
	—	—	0.0

Note: The assumed values for the standard (strand concentration of 1 M) free energy of initiation at 25 C are +6.0 kcal for an A–U base pair and +5.0 kcal for a G·C base pair. In the last reaction, X and Y are any Watson-Crick complementary base pair.

Source: After P. Borer et al. J. Mol. Biol. 86:843 (1974).

1) Bulges

This type of defect occurs when an extra (or more) base (not paired) is contained in one strand of a perfect helix. Both ribo- and deoxyoligomers which give rise to this structure have been synthesized. Thermodynamic analysis carried out in comparison with parent compounds without bulges has shown that bulges destabilizes the structure by about 1-1.2 Kcal/mol for ribooligomers. The destabilization is more pronounced (about 1.8 Kcal/mol) for deoxyoligomers.

2) Hairpin loops

Table 7 shows the melting temperatures of some hairpin loops formed by intramolecular base-pairing. The same compounds can dimerize in solution to give duplexes with internal loops.

Table 7

Melting temperatures of oligoribonucleotide hairpins as determined by optical techniques

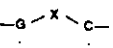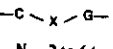| Complex	Presumed helix length	Melting temperature (°C) at oligonucleotide concentration of 10 μM
5' AAAAG^C_C . . . . . 3' UUUUC$_C$	10	20
5' AAAAG^{CC} . . . . . 3' UUUUC$_{CC}$	10	36
5' AAAAAA^C_C . . . . . . 3' UUUUUU$_C$	12	2
5' AAAAAA^{CC} . . . . . . 3' UUUUUU$_{CC}$	12	13
5' AAAAAA$^{CC}_C$ . . . . . . 3' UUUUUU$_{CC}$	12	21
5' AAAAAA$^{CCC}_C$ . . . . . . 3' UUUUUU$_{CCC}$	12	13

By simple mass action considerations one expects that hairpins are more stable at low concentrations whereas duplexes with internal loops are stable at higher concentrations. However the size of the loop is also very important in determining the stability of hairpins. Table 7 shows that the higher stability corresponds to loop size of 6-7 residues. Loop size of 3 or less residues are practically unstable.

As in the case of duplexes, hairpin formation requires nucleation and propagation steps. Differently from the linear duplexes the nucleation free energy not only will depend on the type of base-pairing (AU vs GC) which will start the closure of the loop, but also from the size and the base sequence of the loop.

Equation (18), however, holds also for hairpins. At T_m $\Delta G_T = 0$ because hairpin formation is an intramolecular process, independent of concentration. So, if one knows $\sum \Delta G_{gr}$ (that can be calculated as in the case of linear duplexes) the ΔG°_{nucl} can be obtained. Table 8 shows the available values.

Table 8

Loop free energies at 25°C, 1 M Na$^+$

Loop structure (N = number of unpaired bases)	ΔG° (kcal mole^{-1})
Hairpin	
Closed by A-U (add 1 kcal mole^{-1} if no U in loop)	
$N = 3$	+7
$N = 4$ to 5	+6
$N = 6$ to 7	+5
$N = 8$ to 9	+6
$N \geqslant 10$	$7 + 0.9 \ln(N/10)$
Closed by G-C	
$N = 3$	+8
$N = 4$ to 5	+5
$N = 6$ to 7	+4
$N = 8$ to 9	+5
$N \geqslant 10$	$6 + 0.9 \ln(N/10)$
Internal loop	
$-G \nearrow ^X \searrow C-$ $-C \searrow _X \nearrow G-$	0
$N = 2$ to 6 (except structure above)	+2
$N = 7$	+3
$N \geqslant 8$	$3 + 0.9 \ln(N/7)$
Single-stranded bulge loop	
$N = 1$	+3
$N = 2$ to 3	+4
$N = 4$ to 7	+5
$N \geqslant 8$	$6 + 0.9 \ln(N/8)$

Source: After V. Bloomfield, D. Crothers, and I. Tinoco, Jr., *Physical Chemistry of Nucleic Acids* (New York: Harper & Row, 1974).

An attempt has been made to derive the enthalpic and entropic contributions to ΔG°_{nucl} for hairpins in which both nucleation and growth is obtained with AU pairs. As ΔH_{gr} for AU is known, the difference between calculated $\sum \Delta H_{gr}$ and the van't Hoff enthalpy of melting of a series of different hairpins has been attributed to the enthalpy of nucleation. With this method large and positive values of nucleation have been found. However when these values were used to predict the melting temperature of hairpin containing t-RNA fragments, large difference between calculated and experimental results were found, whereas a good prediction of the melting temperature is obtaine if ΔH°_{nucl} is taken as zero, as in the case of linear duplexes. These aspects, however, deserve more experimental work.

Table 8 shows also some free energy contribution for internal loops.

e) Single-stranded and double-stranded polynucleotides

Table 9 shows the thermodynamic parameters obtained on single-stranded homopolynucleotides in different laboratories. Some of the results were obtained by spectral measurements versus temperature whereas the remaining were obtained by direct calorimetry. The two sets of data are not comparable: those obtained by calorimetry refer to one mole of base, whereas those obtained by optical methods refer to the average unit which cooperatively changes its conformation. the ratio $\Delta H_{vH}/\Delta H_{cal}$ gives an idea of the length of the chain which cooperatively melts. In the case of single-stranded polynucleotides the cooperative length is not too high, as demonstrated also by the large temperature range in which the transition occurs (with the exception of poly rG).

TABLE 9

Thermodynamic parameters for dissociation of synthetic polynucleotides in neutral aqueous solution

Compound	Method	ΔG° kJ/mol	ΔH° kJ/mol	ΔS° J/mol·K
poly(rA)	UV		39.3	123
oligo(rA)	ORD		27.2	90
poly(rA)	ORD	4.6(0°C)	33.1	105
poly(rA)	UV	7.9(0°C)	56.5	167
poly(rA)			37.7	
poly(rA)	hcc		18.8	
poly(rA)	mc		11.5	
poly(rA)	hcc		10.5	
poly(rA)	hcc		12.6	41
poly(rU)	ORD	1.3(0°C)	25.1	88
poly(rU)	hcc		10.9	
poly(rU)	mc		10.8	
poly(rC)	UV	5.6(0°C)	40.2	126
poly(rG)	hcc		9.2	
poly(rI)	hcc		7.9	25.1

As the sample in Table 9 have a single-stranded conformation (with the exception of poly rG and, possibly, of poly rI) the ΔH reported should represent only the base stacking enthalpy. The results appear quite similar.

In the case of double-stranded polynucleotides it is known that thermal stability of ordered structure dependes not only on base composition, but also on nucleotide sequence. Table 10 shows the melting temperature of a series of synthetic polynucleotides with different base composition and sequence. These results have been attributed to different nearest-neighbor base-stacking interactions (as in the case of RNA oligomers previously discussed).

TABLE 10
T_m, Chain Lengths, and T_m Length-Dependent Increments for DNA Duplexes

Duplex Number	Duplex	Wells and Coworkers		Ratliff and Coworkers	
		T_m (°C)	Chain Length (base pairs)	T_m (°C)	Chain Length (base pairs)
1	d(G-C)$_n$·d(G-C)$_n$	105	457		
2	d(G)$_n$·d(C)$_n$	95			
3	d(A-G-C)$_n$·d(G-C-T)$_n$			91.5	3497
4	d(G-G-T)$_n$·d(A-C-C)$_n$			87.0	3612
5	d(A-C)$_n$·d(G-T)$_n$	83	292	79.5	3420
6	d(G-A-T)$_n$·d(A-T-C)$_n$	76	354	74.0	4431
7	d(A-A-C)$_n$·d(G-T-T)$_n$	75	226	70.5	3098
8	d(A-G)$_n$·d(C-T)$_n$	75	667	71.5	4019
9	d(A-G-T)$_n$·d(A-C-T)$_n$	72	304	68.0	3587
10	d(T-T-C)$_n$·d(G-A-A)$_n$	70	196		
11	d(A)$_n$·d(T)$_n$	61	5000	61.8	1000
12	d(A-A-T)$_n$·d(A-T-T)$_n$			56.7	1000
13	d(A-I-C)$_n$·d(I-C-T)$_n$			51.0	3474
14	d(A-T)$_n$·d(A-T)$_n$	54	2800	54.5	1000
15	d(A-C)$_n$·d(I-T)$_n$			50.5	3511
16	d(A-I-T)$_n$·d(A-C-T)$_n$			49.4	3888
17	d(I-C)$_n$·d(I-C)$_n$	49	457		
18	d(A-I)$_n$·d(C-T)$_n$			45.0	3425
19	d(I-I-T)$_n$·d(A-C-C)$_n$			44.0	5004
20	d(I)$_n$·d(C)$_n$	39			

Recent calorimetric measurements performed on polymers of different sequences gave the results shown in Table 11. The enthalpy data are reported in terms of average values for the indicated nearest-neighbor interactions. Despite differences in melting temperatures, these average enthalpy values can be compared directly, since none of the transitions exhibited a significant heat capacity change.

TABLE 11

Compound	Interaction	ΔH (kcal/stack of base pairs)	T_m (°C)
poly[d(AT)]·poly[d(AT)]	AT/TA, TA/AT	7.1	48.5 (0.01M) 74.5 (1.0M)
poly(dA)·poly(dT)	AA/TT	8.6	56 (0.01M)
poly[d(AC)]·poly[d(TG)]	AC/TG, CA/GT	5.6	77 (0.01M)
d(GCGCGGC)	GC/CG, CG/GC	11.9	65 (LCM)

Table 12 shows the comparison of the data with those obtained by Borer et al. in the case of RNA. The similarity is striking.

TABLE 12
Comparison of DNA and RNA Base-Stacking Enthalpies

Stacking Interaction		Stacking Enthalpies	
DNA	RNA	DNA	RNA
AT/TA; TA/AT	AU/UA; UA/AU	7.1	6.5
AA/TT	AA/UU	8.6	8.2
AC/TG; TG/AC	AC/UG; UG/AC	5.6	5.9
GC/CG; CG/GC	GC/CG; CG/GC	11.9	13.8

In order to test the goodness of the data the authors tried to predict the melting enthalpy values of a series of self-complementary oligomers. The results of the comparison of the predicted and measured enthalpies are shown in Table 13. The agreement is indeed good.

TABLE 13
Comparisons of Predicted and Calorimetric Transition Enthalpies

Deoxyoligomer	ΔH_{pred} (kcal/duplex)	ΔH (kcal/duplex)
ATGCAT	37.3	33.4
GGAATTCC	59.3	58.8
CGCGAATTCGCG	106.9	102

THERMODYNAMICS OF NUCLEIC ACIDS SOLUTIONS (II)

F. Quadrifoglio - Istituto di Chimica - Università di Trieste

A series of experiments have demonstrated that for alkali cations the effect of ionic strength on the melting temperature of double-stranded DNA is independent of base sequence and composition.
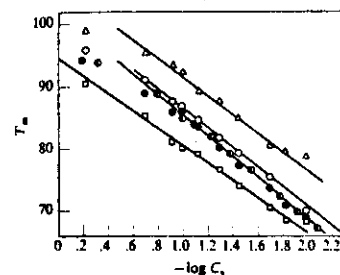

Figure 1

Dependence of DNA melting temperature on salt concentration. Points shown are for D. pneumoniae DNA in KCl (□), E. coli DNA in KCl (O), Ps. aeruginosa DNA in KCl (△), E. coli DNA in sodium phosphate buffer (●), and E. coli DNA in sodium citrate buffer (◍). The deviations from linear behavior at high salt concentrations are real. In fact, at even higher salt concentrations, T_m can decrease with increasing molarity. [After C. Schildkraut and S. Lifson. Biopolymers 3:195 (1968).]

The equation which fits the experimental data (at least in the range 10^{-2}-10^{-1}M uni-univalent salts) is

$$T_m = 16.6 \log C_s + 41 \, x_{GC} + 81.5 \qquad (1)$$

x_{GC} being the fraction of GC base-pairs.

The linear relationship between T_m and $\log C_s$ can be theoretically explained by postulating a different binding of salts to double-stranded with respect to single-stranded DNA. One can write the equilibrium between the two states as

$$D° \rightleftharpoons S_1° + S_2° + \Delta n \, M \qquad (2)$$

where $D°$ is the double-stranded polymer with a definite number of thermodynamically associated positive ions (reference state) and $S_1°$ and $S_2°$ are the single-stranded chains with their thermodynamically associated positive ions. Δn is the number of positive ions "released" in the process.

The association constant is therefore

$$K° = \frac{[D°]}{[S_1°][S_2°][M]^n} = \exp\left(-\frac{\Delta G°}{RT}\right) \qquad (3)$$

from which one can derive

$$\frac{\partial \ln K^\circ}{\partial \ln[M]} = -\Delta n \qquad (4)$$

Because it is much easier to measure T_m (the melting temperature of the chain) than K° one can derive

$$\frac{\partial \ln K^\circ}{\partial T_m} = \frac{\Delta H^\circ}{RT_m^2} \qquad (5)$$

Equation (5) holds for the piece of chain which melts cooperatively (cooperative unit).If ΔH° and ΔS° do not change with temperature

$$\frac{\partial \ln K^\circ}{\partial T_m} = \frac{\partial \ln K^\circ}{\partial \ln[M]} \frac{\partial \ln[M]}{\partial T_m} \qquad (6)$$

Substituting (4) and (5) in (6) one obtains

$$\frac{\partial(1/T_m)}{\partial \ln[M]} = \Delta n \frac{R}{\Delta H^\circ} \qquad (7)$$

If Δn is positive and ΔH° is negative, equation (7) shows a negative linear dependence of $1/T_m$ with $\ln[M]$. (ΔH° is the enthalpy of formation of the double-stranded polymer from the single strands and contains also the enthalpy contribution of ion association to the chains. The predominant contribution is, however, the enthalpy of base-pairing and base stacking which is largely negative).

Equation (7) is not in contraddiction with the experimental results of fig.1 because the experimental data have been collected in a limited range of salt concentration (and therefore in a limited range of T_m). In this range the scale linear in T_m is almost linear also in $1/T_m$ (note the deviation from linearity at higher salt concentration shown in fig. 1).

However while the experimental results of T_m vs. log C_s agree qualitatively with the theory they do not quantitatively. In order to obtain a better agreement it is necessary to make some general considerations on the thermodynamics of linear polyelectrolytes.

If a macromolecule A can be modelled as a linear array of univalent negative charges with average spacing b (measured on the axis of the macromolecule or on its contour length) then a dimensionless structural parameter ξ is defined (Manning, 1969)

$$\xi = \frac{e^2}{\varepsilon k T b} \qquad (8)$$

where e is the protonic charge, ε is the bulk dielectric constant of the solvent, k the Boltzmann's constant and T the Kelvin temperature.

$\frac{e^2}{\varepsilon k T}$ has a value of 7.14 Å in water at 25°C. Because of the practical independence of the product εT on temperature, the dependence of ξ on temperature is generally neglected (for a fixed conformation of the polymer). It is the structural (or charge density) parameter ξ which determines the extent of the counterion association to polyion in solution. When $\xi < 1$ there is no condensation on the polyion. However if $\xi > 1$ counterion condensation takes place.

$\xi > 1$ means that the average distance between the charges (measured on the axis) is lower than 7.14 Å. For native, B-DNA the average distance of negative phosphates, measured on helical axis, is 1.7 Å (two phosphates every 3.4 Å). Thus

$$\xi_{\text{B-DNA}} = \frac{7.14}{1.7} = 4.2$$

Manning has calculated that ion condensation occurs until leaving on the polyion a charge fraction close to $(N\xi)^{-1}$, where N is the valence of the counterion. In the case of sodium B-DNA the charge fraction is therefore 0.24 and in the case of magnesium B-DNA is 0.12. This means that an average of 0.76 sodium and of 0.44 magnesium ions are condensed per phosphate group.

For single stranded DNA at neutral pH b has been estimated to be around 4.3 Å (see below). Thus $\xi = 1.7$ Å and the charge fraction is 0.61 and 0.3 for sodium and magnesium counterions, respectively.

About the physical nature of ion condensation there is a variety of situations which depends upon the nature of the charged groups on the polyion and of the counterion species. The two extreme cases may be visualized as a) fully dehydrated and directly coordinated ions, and b) completely hydrated and delocalized ions moving on the surface of the polyelectrolyte. In the case of DNA experimental data clearly show that alkali metal and alkaline earth cations condense on DNA mostly with the second mechanism.

After condensation both double and single stranded DNA are still highly charged polyelectrolytes with an average of one uncompensated charge per 7.14 Å (ξ^{-1} charges per phosphate). There is therefore an additional screening effect of low molecular weight ions. This screening effect is thermodynamically equivalent to the binding of an additional fraction $(2\xi)^{-1}$ of counterion per nucleotide (Record et al.,1976).

The thermodynamic counterion binding parameter ψ takes into account both condensation and screening effects

$$\psi = \bar{\theta}_M + (2\xi)^{-1} = 1 - \xi^{-1} + (2\xi)^{-1} = 1 - (2\xi)^{-1} \quad \text{(for univalent counterions)}$$

where $\bar{\theta}_M$ is the fraction of condensed counterion

ψ results 0.88 for B-DNA and 0.70 for single-stranded DNA.

Condensation of ions onto polyelectrolytes has been verified experimentally for a number of chains. For DNA experiments carried out with sodium-23 NMR have shown that in the concentration range of sodium ion 0.005-0.5 M the $\bar{\theta}_M$ value is constant and equal to 0.75 ± 0.1, in excellent agreement with the theoretical predictions (Anderson et al.,1978).

Going back to the melting equilibrium of DNA chains one can operationally define the process as

$$D \rightleftarrows S_1 + S_2 \quad ; \quad K_{obs} = \frac{[D]}{[S_1][S_2]} \quad ; \quad \Delta G_{obs} = -RT \ln K_{obs} \quad (9)$$
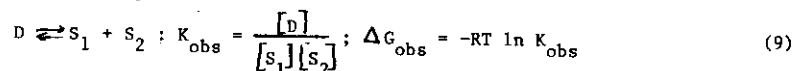
The true thermodynamic equilibrium constant (K°) has been already reported (2). The difference between K° and K_{obs} is due to two effects:
a) the free energy change involved in the uptake of counterions when double-strands are formed (ΔG_c) and
b) the free energy difference due to the different ion atmosphere shielding between double- and single-strands (ΔG_{el}).

If the cooperatively melting unit contains Z phosphates the amounts of counterions bound to double helix and to single-strands are $Z(1 - \xi_D^{-1})$ and $Z(1 - \xi_S^{-1})$, respectively. If at melting the concentration of counterions is $[M]$

$$\frac{\Delta G_c}{RT} = -\Delta n \ln[M] = -Z(\xi_S^{-1} - \xi_D^{-1}) \ln[M] \quad (10)$$

As for DNA $\xi_S^{-1} > \xi_D^{-1}$, Δn is positive: counterions are released in the denaturation process. This has been proved experimentally (Rix-Montel et al.,1974). On the other hand the free energy difference due to different shielding (activity coefficient effect) has been calculated (Manning, 1978) and results

$$\frac{\Delta G_{el}}{RT} = -Z(\xi_D^{-1} \ln \kappa b_D - \xi_S^{-1} \ln \kappa b_S) \quad (11)$$

where κ is the Debye-Huckel screening parameter which is equal to $0.33[M]^{\frac{1}{2}}$ for uni-univalent ions. Thus

$$\frac{\Delta G_{el}}{RT} = -\frac{Z}{2}(\xi_D^{-1} - \xi_S^{-1}) \ln[M] - Z(\xi_D^{-1} \ln 0.33 b_D - \xi_S^{-1} \ln 0.33 b_S) \quad (12)$$

Thus, using (10) and (12)

$$\ln K_{obs} = \ln K^\circ + \frac{\Delta G_c}{RT} + \frac{\Delta G_{el}}{RT} = \ln K^\circ - \frac{Z}{2}(\xi_S^{-1} - \xi_D^{-1}) \ln[M] - Z(\xi_D^{-1} \ln 0.33 b_D$$
$$- \xi_S^{-1} \ln 0.33 b_S) \quad (13)$$

Thus, the effect of salt concentration on the <u>observed</u> equilibrium constant is, from (13)

$$\frac{\partial \ln K_{obs}}{\partial \ln[M]} = -\frac{Z}{2}(\xi_S^{-1} - \xi_D^{-1}) = -\frac{\Delta n}{2} \quad (14)$$

which is one-half the effect predicted earlier (4). With the same procedure used to obtain equation (7) one derives

$$\frac{\partial (1/T_m)}{\partial \ln[M]} = -\frac{\Delta n}{2} \frac{R}{\Delta H^\circ} \quad (15)$$

which is in quantitative agreement with the experimental results. The halved effect is due to the fact that whereas the ion condensation is higher on double-stranded than on single-stranded DNA,so that an increase in ionic strength will favor the formation of double-stranded structures, the single-stranded form has a greater net charge and will be more shielded. The two effects oppose each other, being condensation effect prevalent.

If one expresses the melting properties per phosphate group

$$\frac{\partial (1/T_m)}{\partial \ln[M]} = \frac{\Delta \xi^{-1}}{2} \frac{R}{\Delta H_p^\circ} \quad (16)$$

from which

$$\Delta \xi^{-1} = -2 \frac{\Delta H_p^\circ}{R} \frac{\partial (1/T_m)}{\partial \ln[M]} = -2 \frac{\Delta H_p^\circ}{RT_m^2} \frac{\partial T_m}{\partial \ln[M]} \quad (17)$$

Equation (17) can be used to evaluate the average spacing between phosphate groups in the single strand. For example for T4DNA $RT_m^2/\Delta H_p^\circ$ measured calorime-

trically, is 50 degrees, whereas $\partial T_m / \partial \ln[M]$ is 8.9. Thus $\Delta \xi^{-1} = 0.36$. But $\xi_D^{-1} = 0.24$, therefore $\xi_S^{-1} = 0.60$ and $\xi = 1.7$. It follows that $b = 7.4/1.7 = 4.3$ Å.

When equation (17) is applied to the system poly A – poly U which shows four distinct transitions (fig.2) it is possible to derive four ξ^{-1} measuring

Figure 2

Phase diagram of the poly A + poly U system, at neutral pH in the absence of divalent ions. Melting transitions are shown as a function of Na⁺ concentration. The effect of the stoichiometry of mixtures in each of the four zones is described in the text. [After H. Krakauer and J. Sturtevant, Biopolymers 6:491 (1968).]
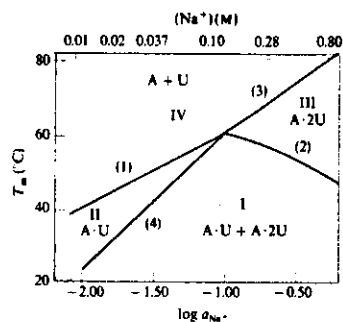
the corresponding variation of T_m with $\log[M]$ and the corresponding ΔH_p°. Because the structures of poly A – poly U and of poly A – 2 poly U are known from X-ray fiber diffraction one can calculate the relative ξ parameters. From a combination of $\Delta \xi^{-1}$ derived from equation (17) and ξ for the ordered structures one can obtain the corresponding parameters ξ for the single-strands poly A and poly U (Tables I and II)

From Table II it results that poly U and denatured DNA appear similar in structure whereas poly A shows an average spacing of 3.1 Å which is in agreement with the high stacking of bases in this polymer.

TABLE 1. *Thermodynamic parameters for polynucleotide transitions* *(NaCl, pH 7)*

Transition	$\dfrac{dT_m}{d \ln a_{\pm}}$	β	$\Delta\psi$	$\Delta\theta_{\text{a}^+}$
AU → A + U	8·5	55	0·16	0·33
AU₂ → A + 2U	11·3	54	0·22	0·44
AU₂ → AU + U	15·6	159	0·10	0·20
2AU → AU₂ + A	−6·5	230	−0·03	−0·06

TABLE 2. *Ion association parameters of nucleic acids*

Structure	ψ	θ_{a^+}	b (Å)
Poly A·2poly U	0·93	0·85	1·0
Poly A·poly U	0·89	0·78	1·5
Helical DNA	0·88	0·76	1·7
Poly A	0·78	0·56	3·1
Denatured DNA	0·70	0·39	4·3
Poly U	0·68	0·36	4·5

REFERENCES

General

Cantor,Ch.R. & Schimmel,P.R. in "Biophysical Chemistry", Part III, Freeman, San Francisco

Specific

Manning,G.S. (1969), J.Chem.Phys. 51,924-933.

Record,M.T.,Lohman,T.M. & deHaseth,P.L. (1976), J.Mol.Biol. 107,145-158.

Anderson,C.F.,Record,M.T. & Hart,P.A. (1978), Biophys.Chem. 7,301-316.

Manning,G.S. (1978), Q.Rev.Biophys. 11,179-246.

Rix-Montel,M.A.,Grassi,H & Vasilescu,D. (1976), Nucleic Acids Res. 3,1001-1011.