91. Calculating the Thermodynamics of Weakly Hydrogen-Bonded Complexes from Heteronuclear NMR Data: Base-Pairing Stabilities of a 5-Methyl($^{15}N_2$)[O^2,O^4-$^{17}O_2$]uridine (= ($^{15}N_2$)[O^2,O^4-$^{17}O_2$]Ribosylthymine) Derivative, and Structural Implications

by **Peter Strazewski**

Institut für Organische Chemie der Universität Basel, St. Johanns-Ring 19, CH–4056 Basel

Dedicated to Prof. Dr. *Albert Eschenmoser* on the occasion of his 70th birthday

(8.II.95)

2′,3′-*O*-Isopropylidene-5-methyl($^{15}N_2$)[O^2,O^4-$^{17}O_2$]uridine (= 2′,3′-*O*-isopropylidene ($^{15}N_2$)[O^2,O^4-$^{17}O_2$]-ribosylthymine; **1**) was analyzed by ^{15}N- and ^{17}O-NMR spectroscopy. The ^{15}N and ^{17}O chemical shifts revealed, in the absence and presence of unlabelled 2′,3′-*O*-isopropylideneadenosine (**2**), the formation of thymine-thymine and thymine-adenine base pairs in $CHCl_3$. As expected, cyclic complexes stabilized by two H-bonds occurred at low temperatures, but at elevated temperatures, the data suggest that open complexes involving only one H-bond prevailed. The ^{17}O-NMR data showed the cyclic thymine-adenine pair in a reverse base pair geometry. The open base pair involved contacts to the urea-derived carbonyl O-atom of thymine. The thermodynamics of complex formation of the cyclic and open forms in both homo and hetero pairs were calculated from the temperature and concentration dependence of the ^{15}N-NMR data using a new method. It involves a fitting procedure onto the experimental isotherms using a theoretically derived function with the standard *Gibbs* free energy as a parameter to be optimized. $\Delta H°$ and $\Delta S°$ were derived from a linear regression of $\Delta G°(T)$ *vs.* T. The fitting procedure circumvents the baseline problem and could be automated and used to calculate correct thermodynamics from UV-monitored melting curves of oligonucleotides. Since titrations are not involved, this dilution method should also be a useful alternative for stability studies of supramolecular complexes in H_2O and in organic solvents.

Introduction. – The chemistry of noncovalent molecular interactions has become a rapidly growing research field in organic chemistry. Among these relatively weak but often specific interactions, H-bonded complexes play an important role in higher-order structures of many natural compounds, such as peptides, proteins, and nucleic acids, as well as in large artificial systems commonly circumscribed as supramolecular complexes. The structural aspects of such complexes give insight into the fundamental process of molecular recognition between two or more interacting subunits. The overall 3D structure of supramolecular complexes or the tertiary structure of large biomolecules are very often deduced from molecular-modelling studies in conjunction with more or less extensive nuclear magnetic resonance (NMR) investigations. In addition, X-ray crystal-structure analysis is a highly welcome means of understanding molecular recognition and, among supramolecular chemists, has become a major research goal. The dynamics, or more precisely, the energetic aspects of H-bonded complexes, however, cannot be studied by X-ray crystallography very well. The computation of absolute interaction enthalpies and entropies is still a very difficult task, leaving calorimetric and spectroscopic techniques as the major tools to obtain quantitative statements about H-bonded complexes.

A well working spectroscopic method has been developed for the calculation of the thermodynamics of nucleic acid double-strand formation some time ago. Usually, the

monochromatic UV absorption of the dissolved molecule is monitored at various temperatures, to obtain a 'melting curve' of the denaturing process. For more local aspects of strand or complex formation, NMR monitoring is a very useful alternative. As long as one can completely shift the equilibrium of complex formation from one side to the other by changing the temperature and concentration, virtually any spectroscopic monitoring method will yield reliable results about the thermodynamic stability of the complex. If, however, the accessible temperature or concentration range will not suffice to observe both complexed and uncomplexed states, the calculation of the thermodynamics can become unreliable or even impossible. This might be a reason for the relative scarcity of spectroscopic stability studies in the field of supramolecular chemistry; it also limits quantitative descriptions of the stability of weak complexes of nucleic acids and proteins.

Since we are investigating the formation of secondary and, possibly weak, tertiary structures of ^{15}N-labelled ribonucleic acids ([^{15}N]RNA), we decided to study first the thermodynamics of the thymine-thymine and thymine-adenine base pairs by means of ^{15}N-NMR spectroscopy. It should be emphasized that the adenine-uracil pairing and uracil selfpairing thermodynamics were already calculated from IR experiments in the late sixties [1] [2]. Similarly, nucleobase and nucleoside pairing properties were studied by ^{1}H- [3], ^{15}N- [4], ^{13}C- [5] [6], and ^{17}O-NMR spectroscopy [6] [7], but the thermodynamics were calculated only from a minor part of this data. In particular, no thermodynamics of the adenosine-uridine pairing on the nucleoside level were calculated from ^{15}N-NMR data. On the oligonucleotide level, however, thermodynamics were calculated from ^{15}N-NMR data [8]. This monitoring method proved to be a powerful tool for the determination of local aspects of DNA base pairing. The examples, so far, involved the formation of relatively stable duplexes allowing the thermodynamics to be elucidated from a complete data set, *i.e.*, from a data set where both fully complexed and fully denatured specimens could be monitored within the accessible temperature range.

A method had to be developed allowing the thermodynamics to be calculated from a reduced data set. The least stable canonical base pair on the nucleobase and nucleoside level was shown by many authors to be the A·U pair, hence, this pair should be an ideal model system for our purposes. In addition, we wished to study the same base pair by ^{17}O-NMR spectroscopy, to compare the two monitoring methods and to see whether ^{17}O-NMR could be used to deduce some structural information of the base pair as well. In the following, the heteronuclear NMR analysis of a doubly labelled uracil derivative, 2′,3′-*O*-isopropylidene-5-methyl($^{15}N_2$)[O^2,O^4-$^{17}O_2$]uridine (= 2′,3′-*O*-isopropylidene-($^{15}N_2$)[O^2,O^4-$^{17}O_2$]ribosylthymine; **1**)[1]), in the absence and presence of commercial 2′,3′-*O*-isopropylideneadenosine (**2**) is described.

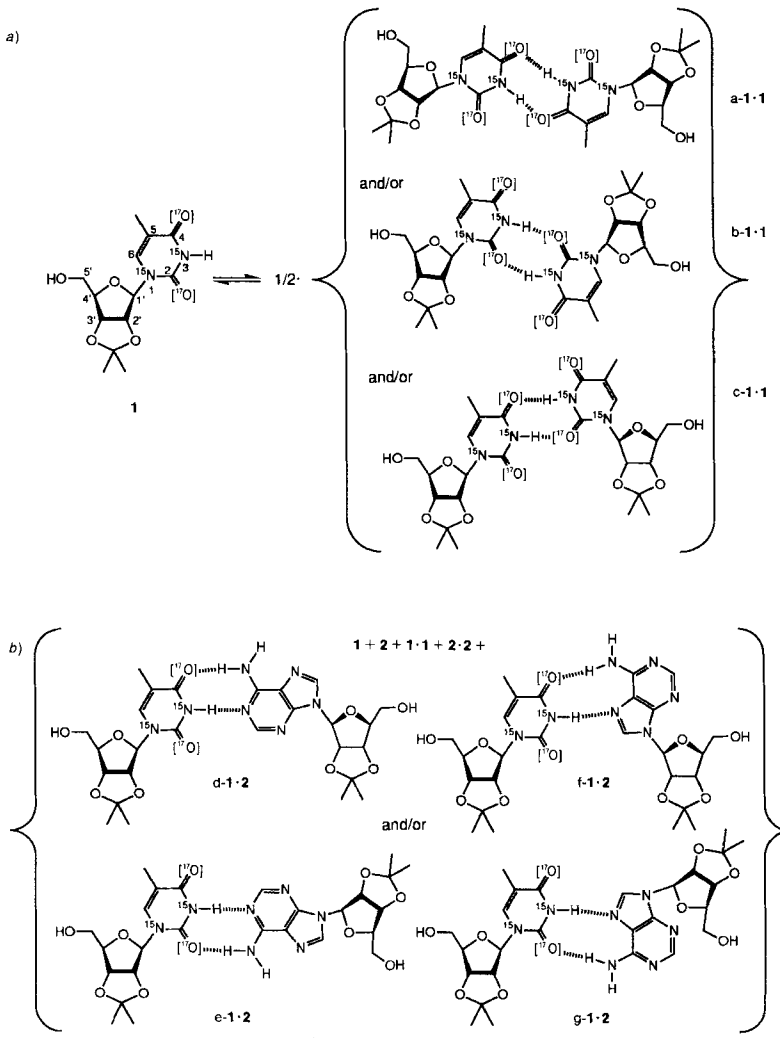
Results. – *^{15}N-NMR Spectroscopy.* To study the base-pairing properties of 5-methyluridine (= ribosylthymine, T) with adenosine (A), H_2O could not be used as a solvent, because monomeric nucleosides would not pair, rather only stack, in this medium [9]. A suitable aprotic and non-H-bonding solvent is $CHCl_3$ [3b] [4a] [10]. The first sign that base pairing occurred, when the synthetic precursor of the nucleoside (see preceding publication), 2′,3′-*O*-isopropylidene derivative **1**, was mixed with 2′,3′-*O*-isopropylideneadenosine (**2**) was the solubility of the compounds. Compound **2** is virtually insolu-

[1]) For convenience, the ^{18}O isotopes are not indicated (*cf.* preceding paper).

ble in $CHCl_3$, **1** is well soluble. A 46 mM suspension of **2** dissolved very rapidly upon addition of 1 equiv. of crystalline **1**. Therefore, **1** was used for a systematic investigation by ^{15}N-NMR spectroscopy in H_2O- and EtOH-free $CDCl_3$.

Compound **1** was measured at 5 and 6 different concentrations (46.13, 23.06, 11.53, 5.77, 2.88, and 1.44 mM from a dilution series) and 12 temperatures (55 to 0° in 5° steps), once alone (*A* series) and once with an equimolar amount of unlabelled **2** (*AB* series). In the *A* series, the monomeric compound is in equilibrium with its selfpaired species

Scheme 1. *Possible Selfpairing and Pairing Geometries between the 2',3'-O-Isopropylidene Derivatives* **1** *and* **2** *of 5-Methyluridine (T) and Adenosine (A), Respectively, in Chloroform.* a) A *Series:* $O^4 \cdot O^4$ *reverse-wobble* (a-**1**·**1**), $O^2 \cdot O^2$ *reverse-wobble* (b-**1**·**1**), *and* $O^2 \cdot O^4$ *wobble pair* (c-**1**·**1**). b) AB *Series:* Watson-Crick (d-**1**·**2**), *reverse*-Watson-Crick (e-**1**·**2**), Hoogsteen (f-**1**·**2**), *and reverse*-Hoogsteen *pair* (g-**1**·**2**)



(*Scheme 1a*). In principle, three pairing geometries are possible: the reverse-wobble arrangement involving both C(4)=O groups (a-**1**·**1**), the reverse-wobble arrangement involving both C(2)=O groups (b-**1**·**1**), and the ordinary wobble arrangement involving one of both C=O groups each (c-**1**·**1**). In the *AB* series, an additional pairing equilibrium with the adenine species occurs (*Scheme 1b*). Here, four geometries are possible: the *Watson-Crick* (d-**1**·**2**), the reverse-*Watson-Crick* (e-**1**·**2**), the *Hoogsteen* (*f*-**1**·**2**), and the reverse-*Hoogsteen* arrangement (g-**1**·**2**). In addition, **2** can selfpair in three geometries involving its *Watson-Crick* and *Hoogsteen* binding sites (not shown). In the ^{15}N-NMR experiment, however, none of the alternative geometries a–c-**1**·**1** and d–g-**1**·**2** are distinguishable, because only the N-atom of **1** is observed. ^{15}N-NMR spectroscopy appears to be ideal for precise measurements at high dilution, because the peaks are narrow and clearly visible down to a concentration corresponding to *ca.* 0.3 mg per ml or 1.4 mM. The shifts of the N signals over a range of up to 3.6 ppm can be observed with a precision of < 0.01 ppm. Base-pair formation is uniformly accompanied by a downfield shift of both $\delta_{N(1)}$ and $\delta_{N(3)}$, as expected for a proton donor and a glycosidic N-atom [4] [8].

Fig. 1 shows the temperature and concentration dependence of the chemical shifts of the glycosidic N(1) signals in the *A* and *AB* series. Although a significant difference between the *A* and *AB* series is visible, it spans at most 0.3 ppm. This was expected because this N-atom is not involved in base-base interactions. N(3), however, shows a marked difference between the *A* and *AB* series as depicted in *Figs. 2* and *3*. The diagrams are identically scaled as the above ones. The differential data (*AB* minus *A* series) in *Figs. 2c* and *3c* show chemical-shift differences that are only due to the effect of added **2**.

Theory. To quantify the measured base-pairing interactions, the theory of thermodynamics must be briefly explored. We are dealing with two kinds of observed chemical equilibria:

$$A + B \rightleftharpoons AB \quad \text{and} \quad A + A \rightleftharpoons AA$$

A corresponds to **1** and *B* to **2**. The 1st equilibrium is a non-selfcomplementary and the 2nd a selfcomplementary system.

For an ideal solution (activity coefficients $\gamma_A = \gamma_B = \gamma_{AA} = \gamma_{AB} = 1.0$), the equilibrium constant of the non-selfcomplementary system is given by $K_{AB} = [AB]/[A]\cdot[B]$. In our case, $[A]$ always equals $[B]$, therefore, *Eqn. 1* holds. The other constant in this experiment is the concentration c_A of the observed molecule (*Eqn. 2*). Let α be defined as the fraction of *A* in the paired state (*Eqn. 3*). α is always between 0 and 1. Thus, the actual concentrations are given by *Eqn. 4*. Replacing the actual concentrations in *Eqn. 1* with *Eqn. 4* yields the quadratic *Eqn. 5*.

$$K_{AB} = [AB]/[A]^{2\,2}) \tag{1}$$

$$c_A = [A] + [AB] \tag{2}$$

---

[2]) Several groups reported a measurable selfpairing of adenine derivatives in organic solvents [2] [3b] [3d–e]. The ethyladenine (9-etAde) selfpairs in $CHCl_3$ with an equilibrium constant half of the selfpairing constant of 1-cyclohexyluracil (1-cyUra) and 1/30 of the mixed pairing constant at 25°. In a 5-mM 9-etAde solution containing an equimolar amount of 1-cyUra, 1.6% 9-etAde and 3% 1-cyUra form cyclic dimers. Therefore, the bias of the $[A]/[B]$ ratio in favor of $[A]$ lies within the experimental uncertainty of c_B and was (as in [2]) neglected for the formulation of K_{AB}.

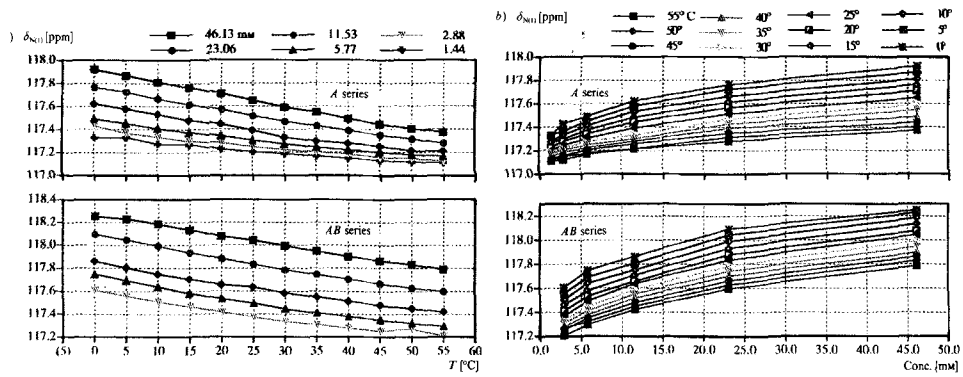


Fig. 1. a) *Temperature and* b) *concentration dependence of the chemical shifts of N(1) of* **1**. Above: *A* series; below: *AB* series; ppm values locked deliberatly relative to external aq. $^{15}NH_4Cl$.

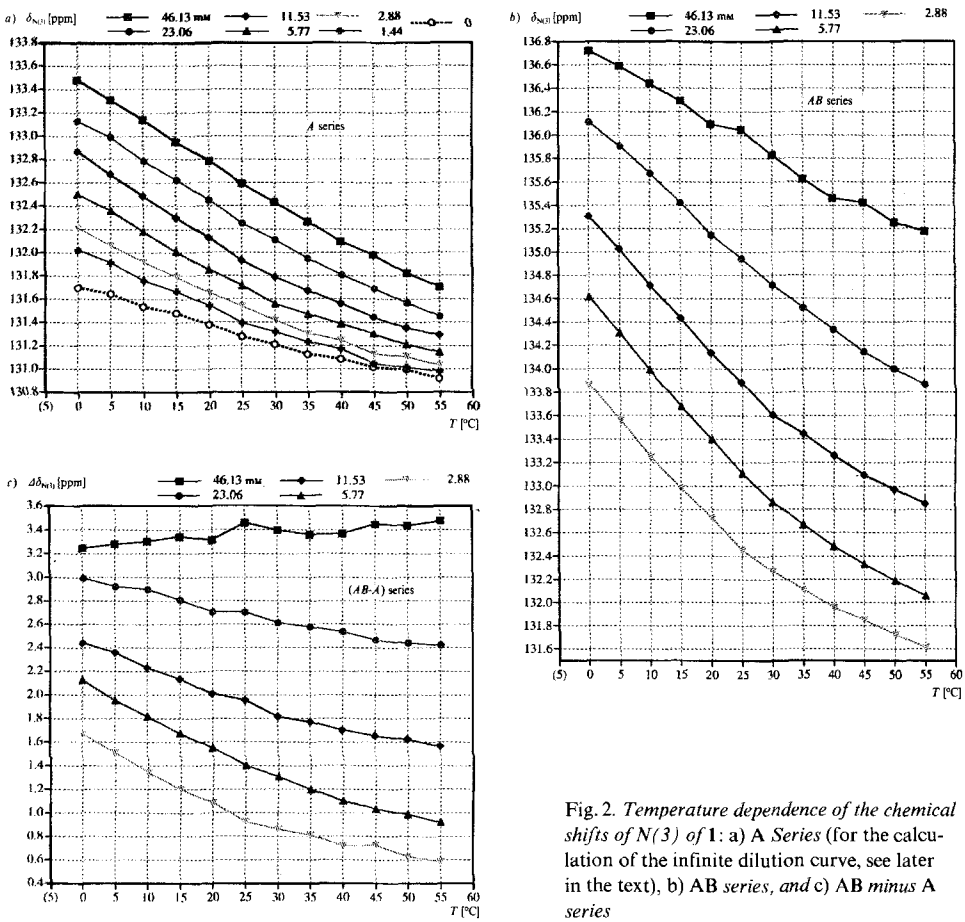


Fig. 2. *Temperature dependence of the chemical shifts of N(3) of* **1**: a) A *Series* (for the calculation of the infinite dilution curve, see later in the text), b) AB *series, and* c) AB *minus* A *series*

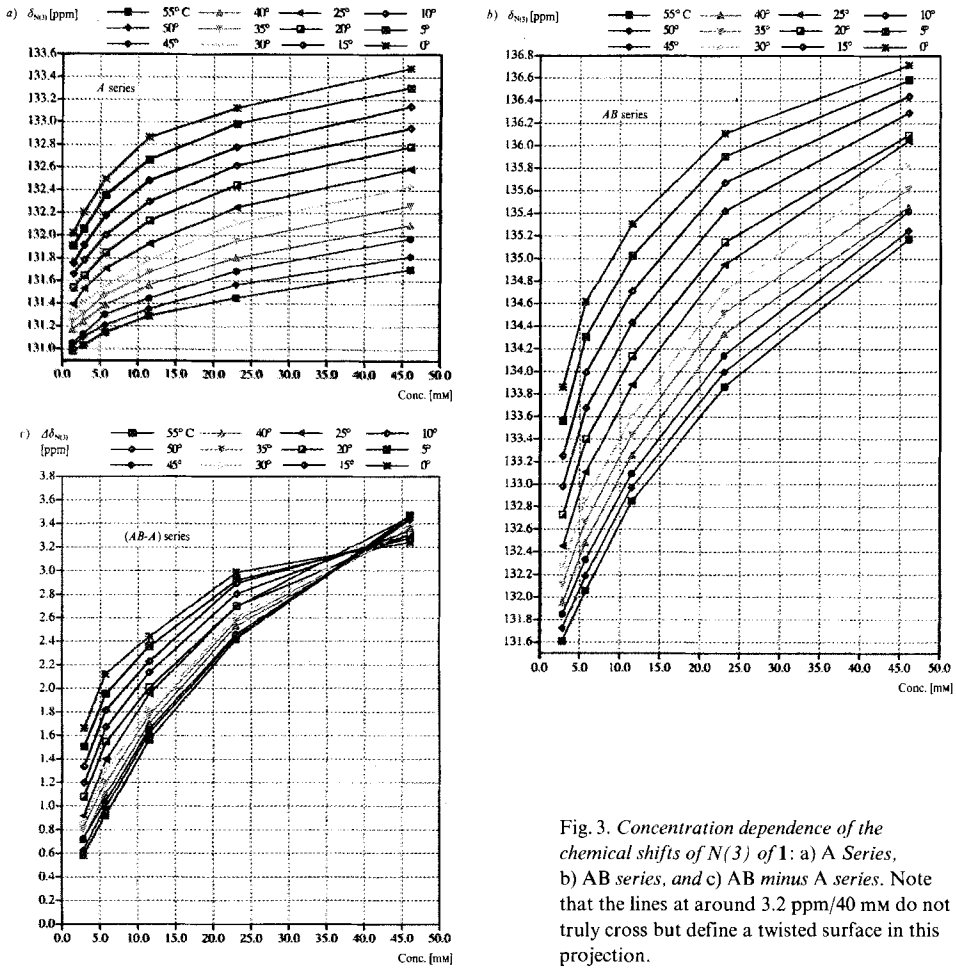


Fig. 3. *Concentration dependence of the chemical shifts of N(3) of* **1**: a) A *Series,* b) AB *series, and* c) AB *minus* A *series.* Note that the lines at around 3.2 ppm/40 mM do not truly cross but define a twisted surface in this projection.

$$\alpha \equiv [AB]/([A] + [AB]) = [AB]/c_A \tag{3}$$

$$[AB] = \alpha \cdot c_A \quad \text{and} \quad [A] = (1 - \alpha) \cdot c_A \tag{4}$$

$$K_{AB} = \frac{\alpha}{c_A \cdot (1 - \alpha)^2} \tag{5}$$

Solving *Eqn. 5* for α gives two solutions, one of which yields an α between 0 and 1 (*Eqn. 6*). Since, for any process at equilibrium, $\Delta G^\circ = -RT\ln K_{eq}$ (R = 1.98586 cal mol^{-1} K^{-1}), we can substitute K_{AB} with $e^{(-\Delta G^\circ/RT)}$. After some rearrangements, we obtain *Eqn. 7*. This expression contains both independent quantities, concentration and temperature, as variables. ΔG° and α are the unknown. Replacement of ΔG° by $\Delta H^\circ - T \cdot \Delta S^\circ$ yields a function showing the dependence of α on the wanted thermodynamic parameters ΔH° and ΔS° (*Eqn. 8*).

$$\alpha(c_A, K_{AB}) = \frac{1 + 2\,c_A\,K_{AB} - \sqrt{1 + 4\,c_A\,K_{AB}}}{2\,c_A\,K_{AB}}\ ^{3}) \tag{6}$$

$$\alpha(c_A, T) = \frac{2\,c_A + e^{\Delta G^\circ/RT} - e^{\Delta G^\circ/2RT}\sqrt{4\,c_A + e^{\Delta G^\circ/RT}}}{2\,c_A} \tag{7}$$

$$\alpha(c_A, T) = \frac{2\,c_A + e^{(\Delta H^\circ - T\Delta S^\circ)/RT} - e^{(\Delta H^\circ - T\Delta S^\circ)/2RT}\sqrt{4\,c_A + e^{(\Delta H^\circ - T\Delta S^\circ)/RT}}}{2\,c_A} \tag{8}$$

For selfcomplementary systems, *Eqns. 9–11* hold. The difference between the two systems becomes apparent in the dependence of $[AA]$ on c_A, because $[A]$ decreases more rapidly upon shifting the equilibrium towards the selfpaired species (*Eqns. 12*), yielding a different equilibrium constant in terms of α (*Eqn. 13*).

$$K_{AA} = [AA]/[A]^2 \tag{9}$$

$$c_A = [A] + 2\cdot[AA] \tag{10}$$

$$\alpha \equiv 2\cdot[AA]/([A] + 2\cdot[AA]) = 2\cdot[AA]/c_A \tag{11}$$

$$[A] = (1 - \alpha)\cdot c_A \quad \text{but} \quad [AA] = \alpha\cdot c_A/2 \tag{12}$$

$$K_{AA} = \frac{\alpha}{2\,c_A\cdot(1 - \alpha)^2} \tag{13}$$

Hence, *Eqns. 14–16* are obtained.

$$\alpha(c_A, K_{AA}) = \frac{1 + 4\,c_A\,K_{AA} - \sqrt{1 + 4\,c_A\,K_{AA}}}{4\,c_A\,K_{AA}} \tag{14}$$

$$\alpha(c_A, T) = \frac{4\,c_A + e^{\Delta G^\circ/RT} - e^{\Delta G^\circ/2RT}\sqrt{8\,c_A + e^{\Delta G^\circ/RT}}}{4\,c_A} \tag{15}$$

$$\alpha(c_A, T) = \frac{4\,c_A + e^{(\Delta H^\circ - T\Delta S^\circ)/RT} - e^{(\Delta H^\circ - T\Delta S^\circ)/2RT}\sqrt{8\,c_A + e^{(\Delta H^\circ - T\Delta S^\circ)/RT}}}{4\,c_A} \tag{16}$$

What is the shape of the *Eqns. 7* and *15*? *Figs. 4* and *5* show some graphical displays of *Eqn. 7*. Since it is a 4-dimensional function, 3-dimensional projections are shown at deliberate concentrations, temperatures, or *Gibbs* free energies. *Fig. 4a* demonstrates the sigmoidal curvature of α *vs.* ΔG° at all temperatures. *Fig. 4b* visualizes that, at $T = 0$ K, α equals 1 and at $T = \infty$ K, α equals 0, irrespective of ΔG° (see *Eqns. 17*). *Fig. 4c* shows that the same sigmoidal curvature of α *vs.* ΔG° is present at all concentrations. Note that while

---

[3]) *Eqn.* 5 corresponds to the common equation $K_{AB}(\alpha)$ (*e.g.* Eqn. 2 in [11]). However, when an equilibrium is monitored by UV spectroscopy, total concentrations $c_{tot} = [A] + [B] + [AB]$ are used. For $[A] = [B]$, $\alpha \equiv 2[AB]/(2[A] + [AB])$ and $K_{AB} = 2\alpha/(c_{tot}(1 - \alpha)^2)$. Solving this eqn. yields $\alpha(c_{tot}, K_{AB}) = (1 + c_{tot}K_{AB} - (1 + 2c_{tot}K_{AB})^{1/2})/c_{tot}K_{AB}$.

the sigmoidal dependence of α on $\Delta G°$ is symmetrical with respect to the lower *vs.* the upper part ($0 < \alpha < 0.5$ *vs.* $0.5 < \alpha < 1$), the sigmoidal dependence of α on T is not, at least not on the isoenergetic surface (*cf. Figs. 4a* and *4b*). *Fig. 4d* visualizes the typical concentration dependence of α within a representative range of *Gibbs* free energies. The curvature is steeper the higher $-\Delta G°$, but at infinite dilution, α always equals 0, and at infinite concentration, α equals 1 (*Eqns. 18*).

$$\lim_{T\to 0} \alpha(c_A, T) = 1 \quad \text{and} \quad \lim_{T\to\infty} \alpha(c_A, T) = 0 \tag{17}$$

$$\lim_{c_A\to\infty} \alpha(c_A, T) = 1 \quad \text{and} \quad \lim_{c_A\to 0} \alpha(c_A, T) = 0 \tag{18}$$

Figs. 5a and *5b* show the concentration and temperature dependence of $\alpha(c_A, T)$, *i.e.*, the isotherms and melting curves of a non-selfcomplementary bimolecular process at a constant *Gibbs* free energy.

Eqns. 7 and *15* are real solutions for a bimolecular reaction; the same procedure can be applied to equilibria of any molecularity. The general form of *Eqns. 5* and *13* are given by *Eqns. 19* and *20*, respectively, where n is the molecularity of the reaction [11].

$$K_{\text{non-self}} = \frac{\alpha}{\left(\frac{c_{\text{tot}}}{n}\right)^{n-1} \cdot (1-\alpha)^n} \tag{19}$$

$$K_{\text{self}} = \frac{\alpha}{n \cdot c_{\text{tot}}^{n-1} \cdot (1-\alpha)^n} \tag{20}$$

A monomolecular equilibrium of the kind $A \rightleftharpoons A'$, such as the hairpin formation of a single strand, is concentration-independent (*Eqns. 21* and *22*).

$$K_{A'} = \alpha/(1-\alpha) \quad \text{and} \tag{21}$$

$$\alpha(K_{A'}) = 1/(1 + K_{A'}^{-1}) \quad \text{or} \quad \alpha(T) = 1/(1 + e^{(\Delta H° - T\cdot\Delta S°)/RT}) \tag{22}$$

A non-selfcomplementary trimolecular reaction equilibrium of the kind $A + B + C \rightleftharpoons ABC$ with equimolar single reaction partners, $c_{\text{tot}} = 3[A] + [ABC]$ and $\alpha \equiv 3\,[ABC]/c_{\text{tot}}$, is defined by *Eqn. 23*, and the selfcomplementary equilibrium constant for the process $A + A + A \rightleftharpoons AAA$ by *Eqn. 24*.

$$K_{ABC} = \frac{9\,\alpha}{c_{\text{tot}}^2 \cdot (1-\alpha)^3} \tag{23}$$

$$K_{AAA} = \frac{\alpha}{3\,c_{\text{tot}}^2 \cdot (1-\alpha)^3} \tag{24}$$

Both cubic equations can be solved for α and yield one real solution each (*Eqns. 25* and *26*).

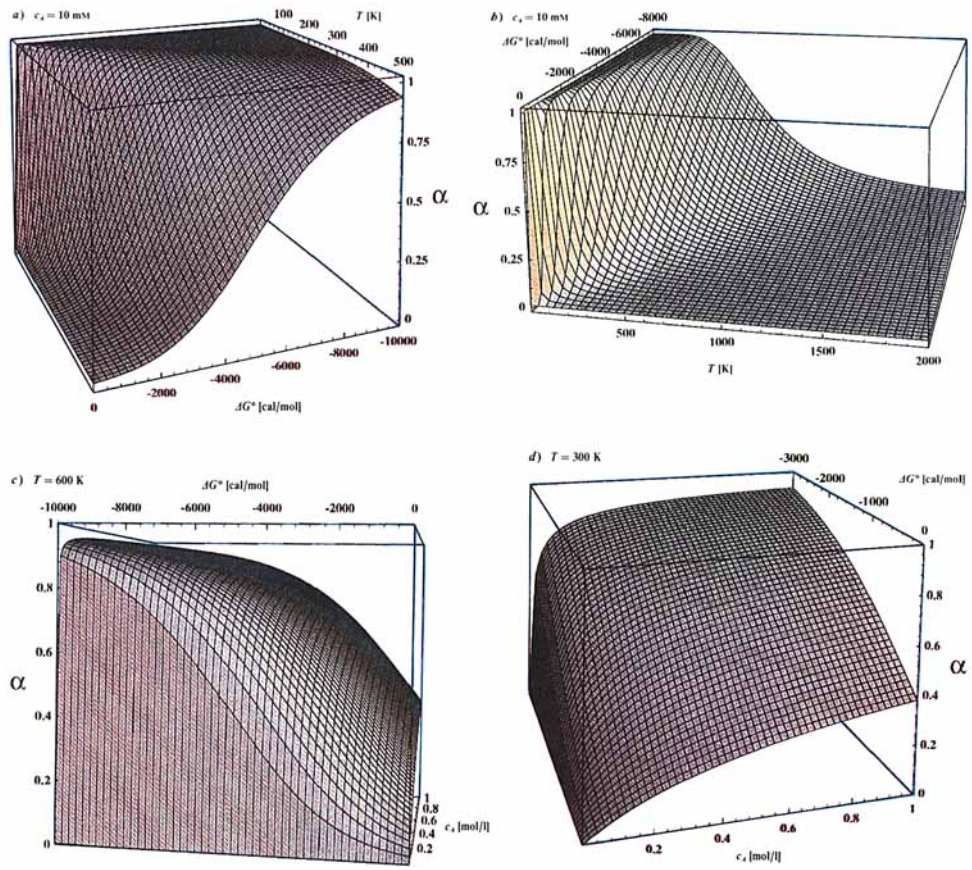


Fig. 4. *Graphical displays of* Eqn. 7 *showing the* $\Delta G°$ *dependence of* α *over a wide* T *and* c_A *range*

$$\alpha(c_{tot}, K_{ABC}) = 1 - \frac{\sqrt[3]{18}}{c_{tot}\sqrt[2]{K_{ABC}}\sqrt[3]{\sqrt[2]{3}\sqrt[2]{4 + 3\,c_{tot}^2\,K_{ABC}} - 3\,c_{tot}\sqrt[2]{K_{ABC}}}} + \frac{\sqrt[3]{3/2}\sqrt[3]{\sqrt[2]{3}\sqrt[2]{4 + 3\,c_{tot}^2\,K_{ABC}} - 3\,c_{tot}\sqrt[2]{K_{ABC}}}}{c_{tot}\sqrt[2]{K_{ABC}}} \tag{25}$$

$$\alpha(c_{tot}, K_{AAA}) = 1 - \frac{\sqrt[3]{2}}{3\,c_{tot}\sqrt[2]{K_{AAA}}\sqrt[3]{\sqrt[2]{4 + 81\,c_{tot}^2\,K_{AAA}} - 9\,c_{tot}\sqrt[2]{K_{AAA}}}} + \frac{\sqrt[3]{\sqrt[2]{4 + 81\,c_{tot}^2\,K_{AAA}} - 9\,c_{tot}\sqrt[2]{K_{AAA}}}}{3\sqrt[3]{2}\,c_{tot}\sqrt[2]{K_{AAA}}} \tag{26}$$

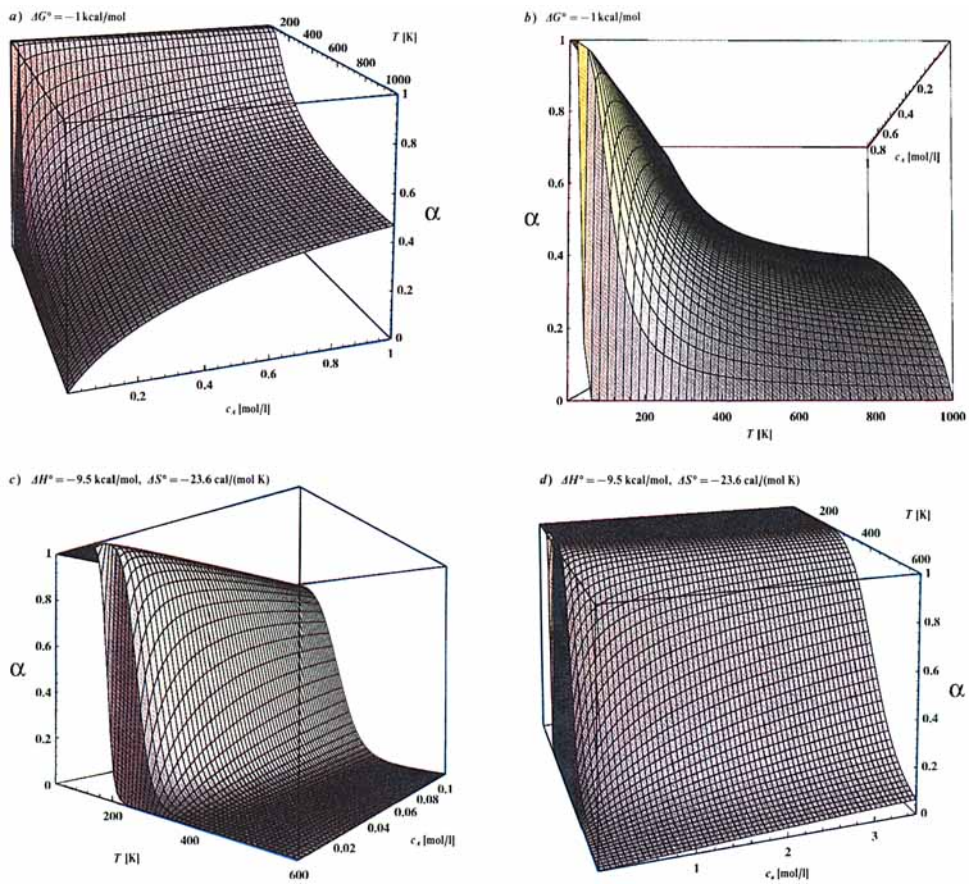


Fig. 5. a) b) *Concentration and temperature dependence of* α(ΔG°) *according to* Eqn. 7. c) d) *Concentration and temperature dependence of* α(ΔH°, ΔS°) *according to* Eqn. 8

Despite the complexity of *Eqns. 25* and *26*, or the more so after they had been replaced by the functions $\alpha(c_{tot}, T)$, the curvature of α *vs.* $\Delta G°$, T, and c_{tot} is much the same as in the bimolecular case. Tetramolecular equilibria of the kind $A + B + C + D \rightleftharpoons ABCD$ or $4\,A \rightleftharpoons AAAA$ are described by the corresponding equations for K_{ABCD} or K_{AAAA} and yield similar solutions for $1 \geq \alpha(c_{tot}, T) \geq 0$ (not shown).

Returning to the bimolecular reaction, the statistical difference between non-selfcomplementary and selfcomplementary but otherwise identical reaction partners formally concerns only $\Delta S°$, not $\Delta H°$. In the former system, each molecule A that pairs with B forms one complex AB, whereas in the latter it needs two molecules A to form one complex AA. Thus, for a given amount of complexes AB, the amount of unpaired species A is larger at equilibrium than for the same amount of complexes AA of identical stability. Therefore, since $S°$ is proportional to the number of molecules involved: $|S°_{AB} - S°_{A(\text{non-self})}| > |S°_{AA} - S°_{A(\text{self})}|$. The reduced entropic penalty for the selfcomplementary system becomes apparent when the α *vs.* T or c_A curves are calculated for both

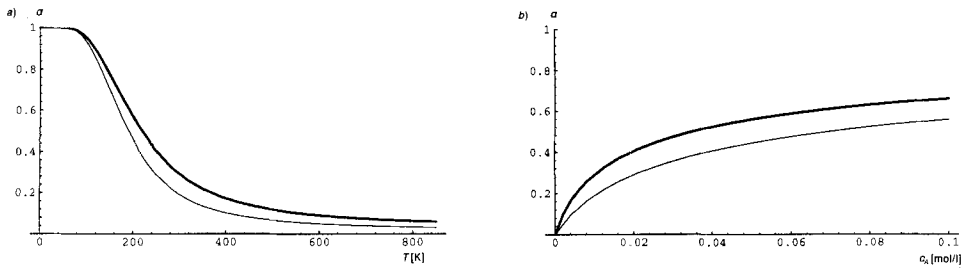


Fig. 6. *Temperature and concentration dependencies of* α *of a selfcomplementary and a non-selfcomplementary system at the same* Gibbs *free energy:* a) *Melting curves at* $c_A = 10$ mM *and* $\Delta G° = -2.0$ *kcal mol*$^{-1}$ *and* b) *isotherms at* T *300 K and* $\Delta G° = -2.0$ *kcal mol*$^{-1}$. —— Selfcomplementary; —— non-selfcomplementary.

systems at the same *Gibbs* free energy: the selfcomplementary system appears to be more stable (*Fig. 6*).

So far, $\alpha(c_A, T)$ was only depicted at the same $\Delta G°$ for all temperatures or concentrations. This implies that we only observed systems where no temperature dependence of $\Delta G°$, *i.e.*, no entropy term $T \cdot \Delta S°$ occurred. This might be realistic for isomerizations, such as certain tautomeric equilibria, but has little relevance to base-pair or other complex formation. If we calculate an α *vs.* T or c_A curve at a given $\Delta G°$, *e.g.* -1 kcal mol^{-1}, with no $\Delta S°$ involved, a relatively flat and unsymmetrical sigmoidal curvature $\alpha(T)$ results, where the temperature for A close to 0 (lower baseline) can only be extrapolated (*Fig. 7a*). Through stepwise doubling the interaction enthalpy and, at the same time, through

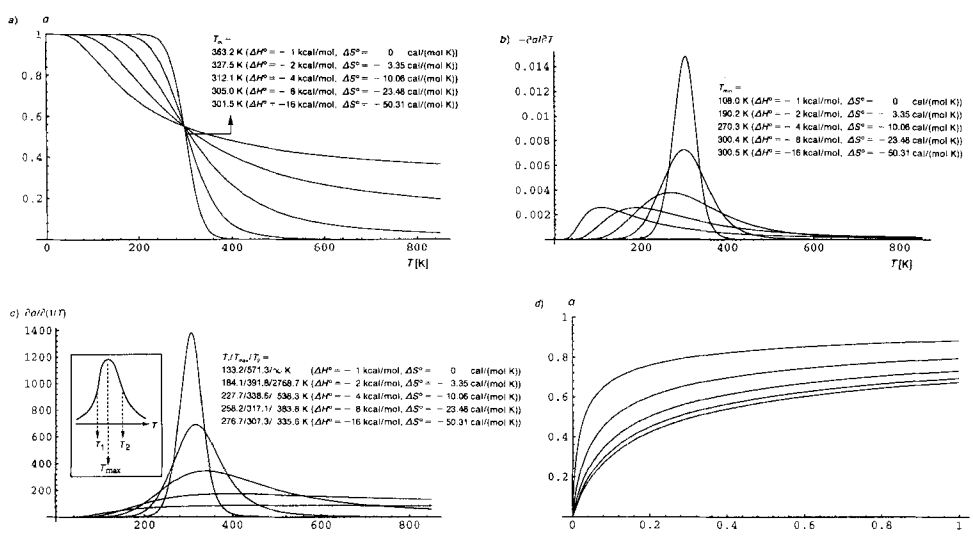


Fig. 7. *Temperature and concentration dependencies of* α *and partial derivatives of* α *at the same* Gibbs *free energy at 298 K but with increasing enthalpy and entropy terms* ($\Delta G°$ (298 K) $= -1.0$ kcal mol$^{-1} = \Delta H° - 298 \cdot \Delta S°$): a) *Melting curves at* $c_A = 500$ mM, *melting temperatures* T_m *at* $\alpha = 0.5$; b) *same as in* a), *but differential melting curves* $\partial\alpha(T)/\partial T$; T_{min} *at* $\partial^2\alpha(T)/\partial T^2 = 0$; c) *same as in* a), *but differential melting curves* $\partial\alpha(T)/\partial(1/T)$; T_{max} *at* $\partial^2\alpha(T)/\partial(1/T)^2 = 0$, T_1 *and* T_2 *at* $0.5 \cdot [\partial\alpha(T)/\partial(1/T)]$; d) *isotherms at* T *273 K*

concomittantly increasing the entropy difference so that at one temperature, *e.g.* 298 K, the resulting *Gibbs* free energy is identical ($\Delta G^\circ = -1$ kcal mol$^{-1} = \Delta H^\circ - 298 \cdot \Delta S^\circ$), the curvature becomes increasingly steeper and symmetric with respect to both baselines.

The melting temperatures T_m, the characteristic stability values for complex formations at $\alpha = 0.5$, were calculated to shift from 363.2 to 301.5 K with an increasing entropy term (*Fig. 7a*). Some common software packages that were designed to calculate melting temperatures from UV melting curves determine the inflection points of α, $[\partial\alpha(T)/\partial T]_{min}$, of the experimental melting curves. In *Fig. 7b*, the corresponding differential melting curves $\partial\alpha(T)/\partial T$ *vs.* T are depicted along with the calculated T_{min} values. Note that, while the inflection point becomes more and more apparent as the entropy term rises, the inflection-point temperatures T_{min} do not correspond to the true melting temperatures T_m at all (compare with *Fig. 7a*). The only T_{min} values that are fairly close to the real melting temperatures T_m are the ones involving high-entropy terms. These differences are derived from the asymmetry of the melting curves involving low-entropy terms with respect to the baselines $\alpha = 1$ and 0. The inverse way of constructing differential melting curves is to plot $\partial\alpha(T)/\partial(1/T)$ *vs.* T (*Fig. 7c*). Here T_{max}, T_1, and T_2 are characteristic values defining the maximum temperature and the half-width of the transition. Again the T_{max} values do not agree with the melting temperatures. The advantage of the inverse differential plot is a smaller temperature range that fully characterizes a melting curve, when compared to baseline temperatures at, *e.g.*, $\alpha = 0.999$ and 0.001 in the normal plot. In *Fig. 7d*, the corresponding isotherms at 273 K are shown. Note that they visibly converge towards $c_A \rightarrow \infty$, whereas in *Fig. 5a* (on the isoenergetic surface), they appear parallel between 250 and 300 K. The effect of the entropy term is also visualized in *Figs. 5c* and *5d* where *Eqn. 8* is plotted. It shows $\alpha(c_A, T)$ on a non-isoenergetic surface with a separate enthalpy and entropy term. The values for ΔH° and ΔS° correspond to the T·A base pair under investigation (*vide infra*). *Fig. 5c* suggests α to be 0 at high temperatures but, as the concentration range is extrapolated to irrealistic values (crystalline state around 3M), the stability of the base pair increases showing the expected curvature (*Fig. 5d*). In reality, the base pair would be even more stable at very high concentrations owing to aggregation effects (base stacking).

Calculating Thermodynamics. The thermodynamics of nucleic acid double strand formation are calculated from melting curves usually monitored by UV spectroscopy at one wavelength (see *e.g.* [12a]), although ^{1}H- [12b] or ^{15}N-NMR [8] detection is also possible. The sigmoidal melting curves are transformed into $\alpha(c_{tot}, T)$ by determining the ratio of the change in UV absorption or chemical shift at a given temperature (relative to the low-temperature baseline) to the UV-absorption or chemical-shift difference between the complexed state (low-temperature baseline) and the single strand (high-temperature baseline). If only one melting curve was measured at one concentration, the obtained values for $\alpha(T)$ are used to detertmine $K_{non\text{-}self}$ or K_{self} using *Eqns. 19* and *20*, respectively. The thermodynamics are then obtained from a *van't Hoff* analysis, *i.e.*, by means of a linear regression of ln K *vs.* T^{-1} based on *Eqn. 27*. A more reliable method is to measure several melting curves at different concentrations (dilution method). After conversion of the measured data points into $\alpha(c_{tot}, T)$, T_m at $\alpha = 0.5$ is determined for each concentration, and the parameters ΔH° and ΔS° are determined from a linear regression of $1/T_m$ *vs.* ln c_{tot} based on *Eqns. 28* and *29* for non-selfcomplementary and selfcomplementary bimolecular systems, respectively [11].

$$\ln K = \Delta S^\circ/R - \Delta H^\circ/RT \tag{27}$$

$$\frac{1}{T_m} = \frac{R}{\Delta H^\circ} \ln c_{tot} + \frac{\Delta S^\circ - R\ln 4}{\Delta H^\circ} \quad \text{and} \quad \frac{1}{T_m} = \frac{R}{\Delta H^\circ} \ln c_{tot} + \frac{\Delta S^\circ}{\Delta H^\circ} \tag{28, 29}$$

Eqns. 27–29 imply two assumptions, that of a two-state transition mechanism where only single and double strands are present with no contribution from intermediate states, and a temperature-independent enthalpy and entropy of complex formation. For long DNA strands, the monomolecular intermediate helix growth steps become dominant, thereby producing an artificially reduced concentration dependence or a *pseudo*-first-order equilibrium for which the melting temperature is concentration-independent. For such systems, it is better to determine ΔH° and ΔS° using the calorimetric method where the heat capacity is measured, and the transition enthalpy does not depend on the nature of transition [11].

On the other end of DNA lengths, there is the so-called baseline problem. The way of transforming a melting curve into a function $\alpha(T)$ requires linearly sloped baselines. Short strands having low stabilities, or strands that are mispaired, either do not fully pair at the freezing point of H_2O or even melt below this temperature ($T_m < 0°$), particularly at high dilutions. In such systems, it is difficult or impossible to determine correct K's or T_m's, because the lower baseline of the melting curves corresponding to $\alpha = 0$ cannot be determined. If the error is systematic in a series of measurements involving various concentrations, the slope of *Eqns. 28* or *29*, *i.e.* ΔH°, will be correct, but ΔS° is likely to be underestimated.

The lower baseline problem is partly circumvented through the use of an alternative method involving inverse differentiated melting curves $\partial\alpha/\partial(1/T)$ as shown in *Fig. 7c*. As long as the temperature window is within the range of T_1 and T_{max}, *i.e.*, within the 'upper half' of the melting curve, the determination of T_{max} and T_1 from a corresponding plot allows the *van't Hoff* transition enthalpy to be calculated [11] [13] (see *Eqn. 30*). $B'(n)$ is a constant that depends on the molecularity n of the process; *e.g.* $B'(2) = -4.38$. If, however, the equilibrium does not involve large entropy terms, as often found in short oligonucleotides or, particularly, in supramolecular complexes consisting of relatively rigid monomeric species, T_1 may be too different from T_{max} for both values to be measurable within a realistic temperature window. *Fig. 7c* demonstrates how sensitive the difference between T_1 and T_{max} is upon variation of the entropy term. Therefore, if the baseline problem is such that neither the upper nor the lower baseline can be reliably determined and the system under investigation involves rather small entropy terms, then the methods presented so far are bound to yield unreliable results. For the analysis of weak nucleobase pairings and in the growing field of supramolecular chemistry, $K(\alpha, T)$ must be fitted onto the monitored data.

$$\Delta H_{VH} = B'(n)/((1/T_{max}) - (1/T_1)) \tag{30}$$

In most of the studies involving spectroscopic monitoring of weak complexation equilibria, one partner is titrated against the other [14]. The *Gibbs* free energy of complexation is usually derived from a nonlinear least-squares curve-fitting procedure based on a *Benesi-Hildebrand* analysis. It operates with chemical shifts or extinction coefficients for fully monomeric and fully complexed states, thus, correlating titration curves with equi-

librium constants. If the titrations are carried out at several temperatures, a subsequent *van't Hoff* analysis reveals enthalpy and entropy of complex formation.

An alternative to the titration method is a fitting procedure that correlates equilibrium constants with melting instead of titration curves. This was done with UV- and ^{1}H-NMR-monitored melting data of short selfcomplementary RNA oligomers that all showed a clean sigmoidal curvature but not always clearly visible lower baselines [12]. The authors fitted *Eqn. 13* onto their melting curves with the assumption that $(1-\alpha)$ and α were linearly dependent on the extinction coefficients or the chemical shifts of the single and double strands, respectively. $K(\alpha) = \exp((-\Delta H°/RT) + \Delta S°/R)$ was directly fitted by the *Marquart* least-squares method (minimizing $\chi^2 = \Sigma[y_i - f(x_i)]^2$; $\{x_i, y_i\}$ are data points and $f(x_i)$ is the applied function). The parameters to be optimized were $\Delta H°$, $\Delta S°$, and four constants: two slopes and two intercepts determining the assumed linear realtionship between α, $(1-\alpha)$, and the corresponding extinction coefficients or chemical shifts, respectively. Interestingly, calculating the UV data revealed that, although the average fitted thermodynamic parameters from the individual melting curves agreed well with the values from the dilution method (*Eqn. 29*; T_m's from fitted curves), a slight temperature dependence of $\Delta H°$ and $\Delta S°$ could be detected, suggesting that single- to double-strand transition was not purely two-state.

Fitting Isotherms. In the following, a similar fitting method for the calculation of transition enthalpies and entropies independent of T_m is presented. It differs from the above method in that it does not fit $K(\alpha)$ onto the actual melting curves, but rather the concentration dependence of $\alpha(c_A)$ separately for each measured temperature, *i.e.*, the isotherms. The data points are $\{c_A, \Delta\delta_{N(3)}(AB-A)\}$ (from *Fig. 3c*), and the function is *Eqn. 7* multiplied by a factor x for the relationship between α and Δppm (see *Eqn. 31*, $\alpha \cdot x = \Delta ppm$).

$$\Delta ppm = x \cdot \frac{2c_A + e^{\Delta G°/RT} - e^{\Delta G°/2RT}\sqrt{4c_A + e^{\Delta G°/RT}}}{2c_A} \qquad (31)$$

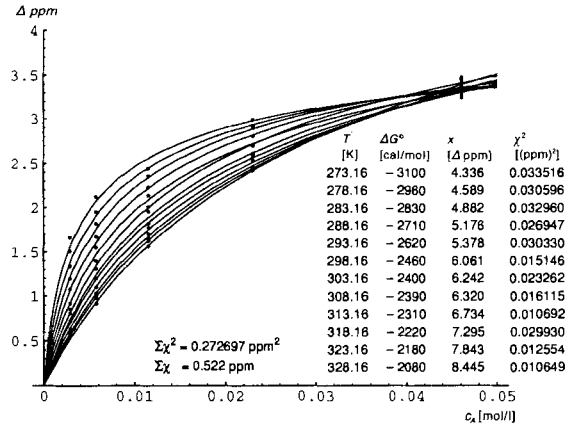
$$T = \{328.16, 323.16, \ldots, 273.16\ \mathrm{K}\}$$

In this form, x may be a constant, a linear, or a nonlinear function of T. With the boundary condition that

$$\lim_{T\to 0} x\cdot\alpha(c_A, T) = \lim_{c_A\to\infty} x\cdot\alpha(c_A, T) = x\ \mathrm{ppm} \quad \text{(from } \textit{Eqns. 17} \text{ and } \textit{18}\text{)}$$

an imaginable temperature dependence of x can be seen and might perhaps be formulated. The apparent slightly negative temperature dependence of Δppm at the highest concentration c_A (*Fig. 2c*) and the way how the experimental isotherms not only converge but 'cross' at high concentrations (*Fig. 3c*) shows that the T·A pairing data involves more than a significant entropy term: a significant temperature dependence of x.

The variable is c_A, and the parameters to be optimized are $x(T)$ and $\Delta G°(T)$. Δppm was fitted by a linear combination of α with x; no separate constant as an intercept was added because of the use of differential data points ($\lim_{c_A\to 0} x\cdot\alpha(c_A, T) = \lim_{[AB]\to 0} x\cdot\alpha(c_A, T) = 0$ ppm). $\Delta G°(T)$ was optimized by minimizing $\chi^2 = \Sigma(\Delta ppm - \Delta\delta_{N(3)})^2$ with a precision of ± 5 cal mol^{-1} for $\Delta G°$, which appeared to be sufficiently accurate with respect to the experimental error range of $\Delta\delta_{N(3)}$. *Fig. 8* depicts the fitted curves with the respective



T [K]	ΔG° [cal/mol]	x [Δ ppm]	χ^2 [(ppm)²]
273.16	−3100	4.336	0.033516
278.16	−2960	4.589	0.030596
283.16	−2830	4.882	0.032960
288.16	−2710	5.176	0.026947
293.16	−2620	5.378	0.030330
298.16	−2460	6.061	0.015146
303.16	−2400	6.242	0.023262
308.16	−2390	6.320	0.016115
313.16	−2310	6.734	0.010692
318.16	−2220	7.295	0.029930
323.16	−2180	7.843	0.012554
328.16	−2080	8.445	0.010649

Fig. 8. *Fitted concentration dependence of* Δppm *for the T·A pairing*

ΔG° [cal/mol] vs. T [273.16–303.16 K]	
Count	7
Num. missing	5
r	0.997
r^2	0.993
Adjusted r^2	0.992
r.m.s. Residual	23.129

ΔG° [cal/mol] vs. T [303.16–328.16 K]	
Count	6
Num. missing	6
r	0.984
r^2	0.969
Adjusted r^2	0.962
r.m.s. Residual	24.708

ΔG° [cal/mol] vs. T [273.16–303.16 K]

	Coefficient	Std. error	Std. coeff.	t Value	P Value
Intercept	−9538.738	252.061	−9538.738	−37.843	< 0.0001
T [273–303 K]	23.643	0.874	0.997	27.046	< 0.0001

ΔG° [cal/mol] vs. T [303.16–328.16 K]

	Coefficient	Std. error	Std. coeff.	t Value	P Value
Intercept	−6448.083	373.013	−6448.083	−17.286	< 0.0001
T [303–328 K]	13.257	1.181	0.984	11.223	0.0004

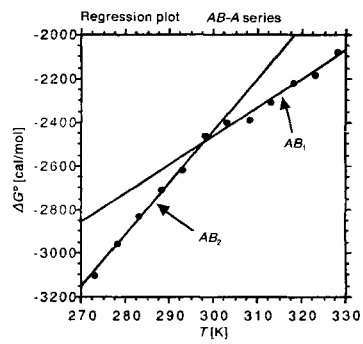


Fig. 9. *Linear regression of* ΔG° (T) vs. T (*AB* – *A* series)

optimized parameters ($\Delta G°$ and x), temperature, and χ^2. The optimized parameters $\Delta G°(T)$ were subsequently submitted to a linear regression *vs.* T directly revealing $\Delta H°$ by the intercept and $\Delta S°$ by the slope (*Fig. 9*).

The correlation coefficient r of the linear regression is a direct measure for the temperature (in)dependence of $\Delta H°$ and $\Delta S°$ and, therefore, a test for the assumption of a pure two-state model. In this system, no monomolecular helix growth contribution could devaluate the two-state model, of course. However, significant base and/or base-pair stacking would show in a monotone temperature dependence of $\Delta H°$ and $\Delta S°$ [3d]. Since the compound was measured in $CDCl_3$ where aggregation is thought to be suppressed by solvation, a simple two-state model was expected to be applicable [3b].

The linear regression over the whole measured temperature range furnishes a coefficient r of 0.985, a residual root mean square (r.m.s.) deviation for $\Delta G°$ of *ca.* 60 cal mol^{-1}, and an intercept and slope corresponding to the enthalpy and entropy change of $\Delta H° = -7.8 \pm 0.3$ kcal mol^{-1} and $\Delta S° = -17.7 \pm 1.0$ cal mol^{-1} K^{-1}, respectively (the

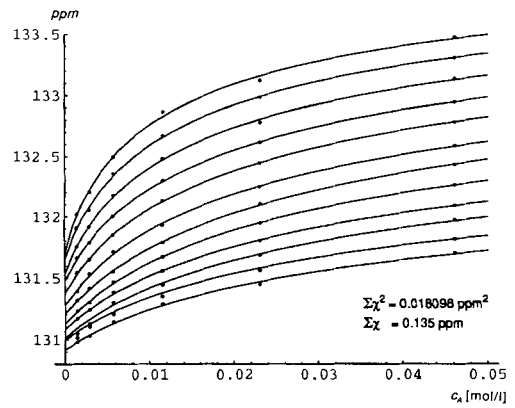
uncertainties are standard deviations). A more critical look at the regression plot in *Fig. 9*, however, suggests that the temperature dependence of $\Delta G°$ might not be linear over the whole temperature range, but biphasic showing two different linear dependencies. A linear regression within the temperature range between 273 and 303 K results in a substantially better fit producing different enthalpy and entropy differences: $\Delta H_2° = -9.5 \pm 0.3$ kcal mol^{-1}, $\Delta S_2° = -23.6 \pm 0.9$ cal mol^{-1} K^{-1}. The corresponding correlation coefficient r amounts to 0.997 and the r.m.s. residual for $\Delta G°$ is only 23 cal mol^{-1}. A linear regression of the high-temperature range between 303 and 328 K produces a fit of a similar quality as the regression over the whole temperature range ($r = 0.984$), but with an expected smaller r.m.s. residual of 24 cal mol^{-1} and smaller negative enthalpy and entropy differences: $\Delta H_1° = -6.4 \pm 0.4$ kcal mol^{-1}, $\Delta S_1° = -13.3 \pm 1.2$ cal mol^{-1} K^{-1}. Apparently, a two-step pairing process was observed each of which predominated within a certain temperature range.

For the calculation of the thermodynamics of the selfpair **1**·**1**, a function from *Eqn. 15* was used to fit onto the data points $\{c_A, \delta_{N(3)}(A)\}$ of the A series by the described procedure, but with an additional parameter $ppm°$, a temperature-dependent constant for the intercept of the linear combination of α and x ($\alpha \cdot x + ppm° = ppm$; see *Eqn. 32*)

$$ppm = ppm° + x \cdot \frac{4\,c_A + e^{\Delta G°/RT} - e^{\Delta G°/2RT}\sqrt{8\,c_A + e^{\Delta G°/RT}}}{4\,c_A} \qquad (32)$$

$$T = \{328.16, 323.16, \ldots, 273.16\ \text{K}\}$$

The fit is better with respect to $\Sigma\chi^2$ (*Fig. 10*) than the fit from the $AB - A$ series. The optimized intercepts $ppm°(T)$ were used to construct the temperature dependence of the chemical shifts of totally unpaired **1**, namely at infinite dilution ($c_A = 0$, dotted line in *Fig. 2a*). The optimized $\Delta G°(T)$ from this fit are plotted against T for a linear regression (*Fig. 11*). When all data points are included, the regression is rather bad resulting in an unacceptable correlation coefficient, $r = 0.887$ (not shown). Again, the regression plot suggests the selfpaired data to express two pairing equilibria of quite different stabilities producing a biphasic temperature dependence of $\Delta G°$. A linear regression within the



T [K]	ΔG° [cal/mol]	ppm° [ppm]	χ [Δppm]
273.16	−1620	131.697	2.830
278.16	−1640	131.645	2.836
283.16	−1610	131.533	2.832
288.16	−1810	131.479	2.830
293.16	−1620	131.381	2.861
298.16	−1570	131.277	2.791
303.18	−1710	131.210	2.917
308.16	−1760	131.126	2.660
313.16	−1830	131.083	2.417
318.16	−1950	131.014	2.320
323.16	−2000	130.994	1.997
328.16	−2100	130.917	1.951

Fig. 10. *Fitted concentration dependence of* ppm *for the T·T pairing*

$\Delta G°$ [cal/mol] *vs.* *T* [308.16–273.16 K]

Count	7
Num. missing	5
r	0.994
r^2	0.988
Adjusted r^2	0.985
r.m.s. Residual	22.168

$\Delta G°$ [cal/mol] *vs.* *T* [328.16–308.16 K]

Count	5
Num. missing	7
r	0.387
r^2	0.150
Adjusted r^2	–
r.m.s. Residual	13.038

$\Delta G°$ [cal/mol] *vs.* *T* [273.16–303.16 K]

	Coefficient	Std. error	Std. coeff.	*t* Value	*P* Value
Intercept	– 6703.269	241.588	– 6703.269	– 27.747	< 0.0001
T [K]	16.857	0.838	0.994	20.119	< 0.0001

$\Delta G°$ [cal/mol] *vs.* *T* [273.16–303.16 K]

	Coefficient	Std. error	Std. coeff.	*t* Value	*P* Value
Intercept	– 6703.269	241.588	– 6703.269	– 27.747	< 0.0001
T [308–328 K]	16.857	0.838	0.994	20.119	< 0.0001

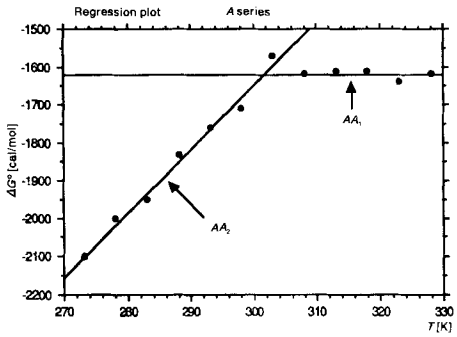


Fig. 11. *Linear regression of* $\Delta G°(T)$ vs. T (*A* series)

temperature range between 273 and 308 K results in a coefficient $r = 0.994$ and an r.m.s. residual for $\Delta G°$ of 22 cal mol^{-1}, showing an enthalpy and entropy difference of $\Delta H_2° = -6.7 \pm 0.2$ kcal mol^{-1} and $\Delta S_2° = -16.9 \pm 0.8$ cal mol^{-1} K^{-1}. The regression in the temperature range between 308 and 328 K suggests the *Gibbs* free energy to be essentially temperature-independent: $\Delta G° = \Delta H_1° = -1.4 \pm 0.3$ kcal mol^{-1} ($\Delta S_1° \approx 0$)[4]). The r.m.s. residual of $\Delta G°$ is only 13 cal mol^{-1} (the correlation coefficient is very low due the flat slope).

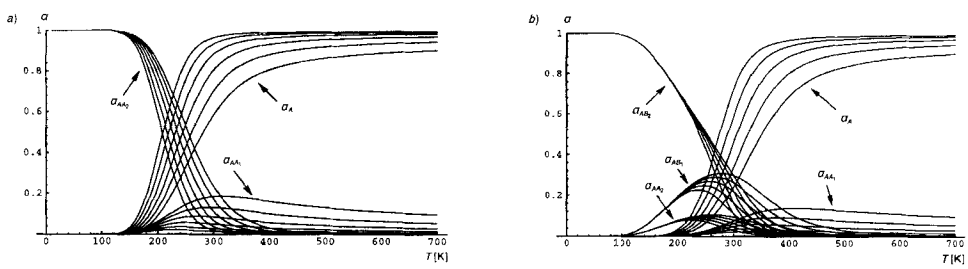


Fig. 12. *Mole fractions of unpaired T, the weak and strong T·T and T·A pairs* a) *in the* A b) *and in the* AB *series at the measured concentrations.*

a) $c_A = [A] + 2[AA_1] + 2[AA_2]$; $[A]/c_A = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})/C$; $[AA_1]/c_A = \alpha_{AA1}\cdot(1-\alpha_{AA2})/2C$; $[AA_2]/c_A = (1-\alpha_{AA1})\cdot\alpha_{AA2}/2C$; $C = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2}) + 0.5\cdot\alpha_{AA1}\cdot(1-\alpha_{AA2}) + 0.5\cdot(1-\alpha_{AA1})\cdot\alpha_{AA2}$. Functions α_{AA1} and α_{AA2} according to *Eqn. 16* with the optimized values for $\Delta H_1°$, $\Delta S_1°$, $\Delta H_2°$, and $\Delta S_2°$ (*A* series). b) $c_A = [A] + 2[AA_1] + 2[AA_2] + [AB_1] + [AB_2]$; $[A]/c_A = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2})/C$; $[AA_1]/c_A = \alpha_{AA1}\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2})/2C$; $[AA_2]/c_A = (1-\alpha_{AA1})\cdot\alpha_{AA2}\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2})/2C$; $[AB_1]/c_A = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot\alpha_{AB1}\cdot(1-\alpha_{AB2})/C$; $[AB_2]/c_A = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot\alpha_{AB2}/C$; $C = (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2}) + 0.5\cdot\alpha_{AA1}\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2}) + 0.5\cdot(1-\alpha_{AA1})\cdot\alpha_{AA2}\cdot(1-\alpha_{AB1})\cdot(1-\alpha_{AB2}) + (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot\alpha_{AB1}\cdot(1-\alpha_{AB2}) + (1-\alpha_{AA1})\cdot(1-\alpha_{AA2})\cdot(1-\alpha_{AB1})\cdot\alpha_{AB2}$. Functions α_{AA1} and α_{AA2} as in a); α_{AB1} and α_{AB2} according to *Eqn. 8* with the optimized values for $\Delta H_1°$, $\Delta S_1°$, $\Delta H_2°$, and $\Delta S_2°$ ($AB - A$ series).

[4]) Theoretically, this unexpected low pairing entropy could have resulted from the neglection of a significant activity coefficient $\gamma_{AA} < 1.0$ at higher concentrations c_A. However, a fit and regression involving the high-temperature data points from the *A* series (308.16–328.16 K) at low concentrations (1.44–11.53 mM) only did not produce a more negative $\Delta S_1°$ value within its standard deviation.

With the thermodynamics of all pairing equilibria at hand, the mole fractions of monomeric **1** ($[A]/c_A$) and of both **1**·**1** selfpairs ($[AA_1]/c_A$ and $[AA_2]/c_A$) in the *A* series can be calculated using $\alpha_{AA_1}(c_A, T)_{\Delta H°_1,\Delta S°_1}$ and $\alpha_{AA_2}(c_A, T)_{\Delta H°_2,\Delta S°_2}$ from *Eqn. 16* (*Fig. 12a*). Note how the weak pairing AA_1 predominates at elevated temperatures owing to its insignificant entropic penalty. Similarly, the mole fractions of **1**·**2** ($[AB_1]/c_A$ and $[AB_2]/c_A$ using $\alpha_{AB}(c_A, T)_{\Delta H°_1,\Delta S°_1}$ and $\alpha_{AB_2}(c_A, T)_{\Delta H°_2,\Delta S°_2}$ from *Eqn.8*), **1**·**1** ($[AA_1]/c_A$ and $[AA_2]/c_A$), and monomeric **1** ($[A]/c_A$) in the *AB* series can be calculated (*Fig. 12b*). The comparison of the plots in *Fig. 12a* and *12b* demonstrates how effectively the more stable hetero pairing AB_2 competes with selfpairing AA_2.

^{17}O-NMR Spectroscopy. The ^{15}N-NMR analysis is a reliable method for the calculation of the thermodynamics of the observed system. However, because only one N-atom was observed, it does not allow for any structural conclusions. According to the literature, ^{17}O-NMR spectroscopy is a very sensitive method for the detection of H-bonds involving various compounds (see *e.g.* [6] [7] [15] [16]). The removal or addition of an O-bound proton is accompanied by a significant change in the chemical shift δ(O); so is the deprotonation of an O-containing functional group where the proton was not necessarily bound to the O-atom (*e.g.* lactams). Generally, the higher the π-bond order is, the larger are the shifts. Thus, ether O-atoms produce small, phosphate O-atoms intermediate, and carbonyl O-atoms the largest shifts upon H-bonding or protonation. While deprotonations of lactam groups induce upfield shifts of 40–100 ppm, the hydration of a carbonyl group usually involves upfield shifts of *ca.* 25 ± 10 ppm per H-atom, but can be higher in intramolecular cases. In some cases, however, H-bond formation was found to direct not upfield but downfield shifts of a smaller magnitude. The formation of a H-bond to the OH group of H_2O or MeOH was measured to induce a downfield shift of 12 and 6 ppm depending on whether the OH group acted as a H-donor or H-acceptor, respectively [17]. The H-bonding of the NH_2 group of adenine produced downfield shifts in the complexed thymine carbonyl O-atom of up to 11 ppm [7]. In any case, a linear correlation between the degree of H-bonding and ^{17}O-NMR chemical shift is legitimate.

Since compound **1** is doubly labelled, the involvement of the O-atoms of **1** paired with **2** and with itself could be studied by ^{17}O-NMR spectroscopy under the same conditions as by ^{15}N-NMR. From the published data on the selfassociation of 2′,3′-*O*-isopropylideneuridine in MeCN [6] and on adenine-thymine pairing in DMSO [7] monitored by ^{17}O-NMR, small but significant shifts of the involved base O-atoms were expected upon decreasing temperatures in either direction, depending on whether the H-donor was a lactam (upfield), an OH (upfield), or an NH_2 group (downfield). Larger downfield shifts should occur upon addition of **2**.

The solvents used for this investigation were non-deuterated, H_2O and EtOH-free MeCN, and $CHCl_3$. MeCN was used because of its low viscosity. The line width w of NMR signals derived from quadrupolar nuclei is related to the viscosity η of the medium through the linear dependence of w on the rotational correlation time τ which, in turn, is proportional to η at a given temperature [15g]: $w \propto \tau \propto \eta/T$. Hence, well resolved and relatively narrow peaks were expected from the spectra in MeCN. $CHCl_3$ was used to be able to compare the results with those from the ^{15}N-NMR-spectroscopic investigation. Neat 1,4-dioxane served as an external standard, because its chemical shift was measured to be essentially temperature-independent (not shown). As usual in ^{17}O-NMR spectroscopy, no deuterium lock was applied, because the ^{2}H frequency is too close to the ^{17}O

frequency (61.4 *vs.* 54.2 MHz in a 9.3948-T field of a ^{1}H-400-MHz magnet [16]). Regular controls with the standard showed that the ppm scale hardly ever shifted more than 1 ppm during prolonged measurements. In each solvent, two temperature-dependent measurements were carried out, one containing **1** and one containing an equimolar mixture of **1** and **2**. The concentration of **1** and **2** was 46 mM each, the highest concentration in the ^{15}N-NMR measurements. The temperatures spanned from 70 to 25° in MeCN and from 55 to 0° in $CHCl_3$. *Fig. 13* shows two representative spectra of **1** and **1/2** in MeCN (*Fig. 13a*) and all spectra in $CHCl_3$ (*Figs. 13b* and *13c*). In *Figs. 14* and *15*, all chemical shifts and half-intensity widths, respectively, are plotted *vs. T*.

First of all, ^{17}O chemical shifts cannot be determined nearly as accurately as ^{15}N chemical shifts. This is only partly due to the fact that the measurements were carried out with no lock frequency. A certain systematic error between the measurements of two different solutions can occur owing to baseline rolling, a phenomenon that makes it difficult to find the correct phases for the *Fourier* transformation. However, much more disturbing is the increasingly bad signal-to-noise ratio owing to signal broadening at lower temperatures. Therefore, the chemical shifts depicted in *Fig. 14* are the less reliable, the lower the temperature is. The half-intensity widths of the signals (*Fig. 15*) can be measured more accurately, but the relative errors increase with the line widths for the same reason.

Despite the very well-resolved signals in MeCN (*Fig. 13a*), the spectra show disappointingly small differences in both chemical shift and signal half-intensity widths at even lowest temperature. With decreasing temperatures, an upfield shift of 1.5 to 2 ppm for O^2

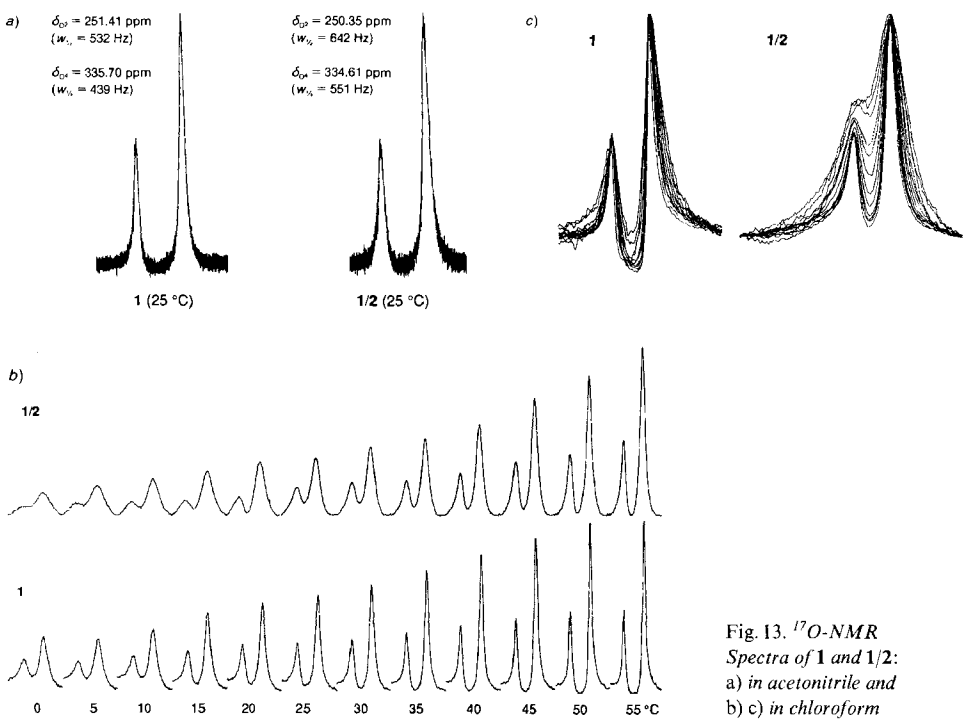


Fig. 13. *^{17}O-NMR Spectra of* **1** *and* **1/2**: a) *in acetonitrile and* b) c) *in chloroform*

and O^4 is observed over the temperature range of 45°. The insignificant upfield (!) shift of *ca.* 1 ± 1.0 ppm upon addition of **2** is hardly temperature-dependent suggesting that practically no base pairing between **1** and **2** occurred in this medium. The difference between both measurements might be due to baseline rolling. The present results including the ^{15}N-NMR investigation suggest that, if no pairing with **2** is observed under the applied conditions, a selfpairing can be safely excluded as well. *Fig. 15* shows that the half-intensity widths of both O-atoms change only very slightly within 45° ($\Delta w_{1/2}$ = 145–248 Hz), irrespective of the presence or absence of **2**. These differences, as well as the difference of *ca.* 50 to 100 Hz between both solutions (*A* and *AB* series), are attributed to the increasing viscosity of the medium with decreasing temperatures and in the presence of **2**, respectively [15g]. All in all, not much happens in MeCN. The solvent is a too good H-bond acceptor being able to effectively compete with **1** and **2**.

In contrast, $CHCl_3$ once again proved to be the ideal solvent for the system, as can be immediately seen from *Fig. 13b*. The substantial temperature-dependent line broadenings in both solutions are witnesses of the intermolecular interactions that gave rise to significant changes in the average rotational correlation time τ of the complexes involved (τ is also proportional to the molecular volume V_m [15g]). Furthermore, the more pronounced broadening in the *AB* series demonstrates the effect of added **2**. From a first glance, the data seem to be sufficiently significant to be used for a fit as described for the ^{15}N-NMR data. However, *Fig. 13c* shows all corresponding spectra overlapped in such a way that not the integrals – as in *Fig. 13b* – but the intensities remain roughly constant. Now, one sees the difficulties in determining the exact chemical shifts and signal widths, especially in the low-temperature range, which is the reason why only a qualitative rather than quantitative analysis is anticipated. Despite these difficulties, important tendencies are visible in the corresponding plots in *Fig. 14*.

The O^2 signal in the *A* series (T, $CHCl_3$) is shifted upfield by *ca.* 5 to 6 ppm between 50 and 0°. Three reasons are imaginable: a selfpairing in the $O^2 \cdot O^4$ wobble, in the $O^2 \cdot O^2$ reverse-wobble geometry, and a contribution from the intramolecular H-bonding of O^2 to the 5′-OH group. The apparent temperature independence of the corresponding O^4 signal is rather puzzling. However, the signal does seem to shift upfield between 55 and 20°, *i.e.*, in the more reliable range, by *ca.* 1 ppm. Below that temperature, the errors in the chemical shifts are too large to be safely interpreted as temperature-independent.

The selfpairing of uracil and thymine in DMSO and added MeOH or H_2O was previously shown to be governed by the $O^4 \cdot O^4$ and $O^2 \cdot O^2$ reverse-wobble geometries with no contribution from the $O^2 \cdot O^4$ wobble arrangement [7]. In addition, several studies involving nucleoside or alkylnucleobase derivatives rather support a preference for the $O^4 \cdot O^4$ reverse-wobble geometry for this pair [5] [6] [18]. In one particular investigation, the existence of an intramolecular complexation equilibrium between O^2 and OH–C(5′) was deduced from ^{17}O-NMR spectra of 2′,3′-*O*-isopropylideneuridine and 5′-deoxy analogs thereof ([6] and refs. cit. therein). Therefore, it seems likely that the intramolecular H-bonding may be a major contributor to the large shift difference for O^2. The shifts of both signals indicate that the selfpairing of **1** is strongly overlapped by this interaction making it difficult to distinguish between the possible wobble geometries.

In the *AB* series, the shifts are more significant. As expected from the H-bonding of an NH_2 group to a carbonyl group, the O^2 signal is shifted downfield upon addition of **2** by up to 4 to 6 ppm (at 0°). The temperature dependence of the shift is less pronounced but,

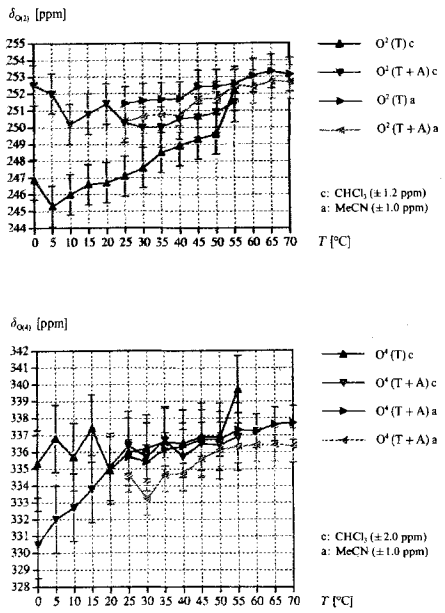


Fig. 14. *Temperature dependence of the ^{17}O chemical shifts of the O^2 and O^4 atoms of* **1**

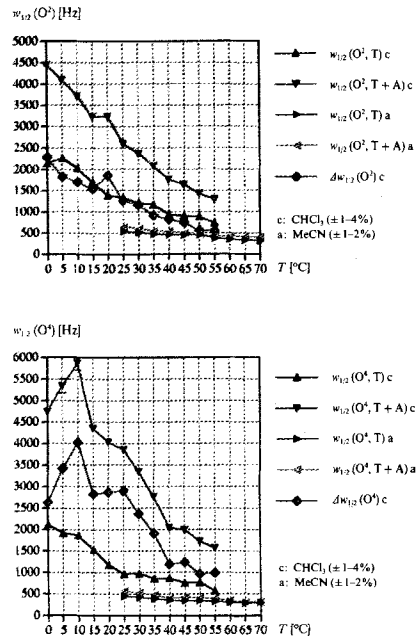


Fig. 15. *Temperature dependence of the ^{17}O half-intensity widths of the O^2 and O^4 atoms of* **1**

within the uncertainty limits, the shift difference is 1–2 ppm between 55 and 0°. This experimental evidence supports the notion that the pairing between a pyrimidine- and a purine-nucleoside derivative favors a (O^2-bound) reverse-*Watson-Crick* or reverse-*Hoogsteen* geometry over the normal (O^4-bound) one. The reverse pairing induces an upfield shift of possibly (see *Fig. 14*, $CHCl_3$, *AB* series) up to 6 ppm in the O^4 signal. This indirect effect (owing to the N–H···N H-bond) additionally excludes the possibility of a normal *Watson-Crick* or *Hoogsteen* pairing, because it points into the upfield direction.

Most interestingly, both chemical shifts in the *AB* series show a biphasic temperature dependence. The tendency changes somewhere between 35 and 25°, in agreement with the results from the ^{15}N-NMR analysis (*Fig. 9*). Upon cooling, the O^2 signal is shifted first slightly upfield (possibly due to N–H···N), then downfield (O···H_2N), but is always significantly downfield from the corresponding signal in the *A* series. The O^4 signal, in contrast, hardly changes between 55 and 25° as in the *A* series, but, upon further cooling, is shifted into the mentioned upfield direction. This behavior is consistent with the assumption that base-pair formation occurs in two steps. The higher temperatures merely allow for one H-bond to be formed. In such an 'open base pair', the NH_2 group preferentially contacts the O^2 atom. A fraction of imino proton-bound open base pair might contribute to the slight initial upfield shift of $\delta(O^2)$ (if real). At lower temperatures, the cyclic reverse base-pairing geometry predominates.

The signal half-intensity widths in both homo- and hetero-paired systems seem to be a better means of monitoring the pairing equilibrium than the chemical shifts (*cf. Figs. 14* and *15*). The curvatures obtained from $w_{1/2}$ are reminiscent of the melting curves obtained

from the ^{15}N chemical shifts and, if not for the fact that the ^{15}N data were more accurate, could be used to calculate the thermodynamics by fitting $\alpha(c_A, T)$ onto $\{c_A, w_{1/2}(T)\}$. Like the chemical-shift differences, the half-intensity width differences seem to be more sensitively monitored by the O^4 than the O^2 signal, albeit less reliably at low temperatures in this case. In the *A* series, $w_{1/2}$ ranges from *ca.* 600 to 2200 Hz for both signals ($\Delta w_{1/2} = 1600$ Hz). In the *AB* series, the O^2 half-intensity width shifts from *ca.* 1300 to 4500 Hz ($\Delta w_{1/2} = 3200$ Hz), whereas the O^4-signal half-width shifts from 1600 to (probably) *ca.* 5800 Hz ($\Delta w_{1/2} \approx 4200$ Hz). The enhanced sensitivity shows in the significantly higher and steeper differential melting curve for O^4 (1000 to *ca.* 4000 Hz, $\Delta w_{1/2} = 3000$ Hz) than for O^2 (500 to 2000 Hz, $\Delta w_{1/2} = 1500$ Hz). This difference may be related to the different electronic-field gradients at the O^4 and O^2 nuclei in the hetero base pair, that define the oxygen quadrupole coupling constant (*QCC*), *i.e.*, the proportionality factor between $w_{1/2}$ and τ [15g].

Discussion. – *Fitting Procedure.* The presented fitting procedure is based on the optimization of the parameter $\Delta G°$ within a function $\Delta ppm = x(T) \cdot \alpha(\Delta G°, c_A, T)$ or $ppm = ppm°(T) + x(T) \cdot \alpha(\Delta G°, c_A, T)$ to such a degree that the mean square difference χ^2 between the function and the experimental data points is minimized (first-degree-polynomial fit). Function α is linearly scaled by factor x and, in the analysis of the selfpair T·T, positioned onto the experimental absolute ppm values by an intercept *ppm*°. Since $\Delta G°$ is temperature-dependent, the fittings are performed with each experimental isotherm separately. An inspection of the theoretical functions depicted in *Figs. 4* and *5* shows that, by fitting isotherms, the relative order of data points at one temperature is fully characterized by its *Gibbs* free energy of interaction, *i.e.*, only the *curvature* of every single isotherm is relevant for the fitting, irrespective of the relation of the isotherms relative to each other. Therefore, the fitting is independent on the actual concentration of a particular base pair (or complex) under investigation. This allows us to subtract contributions from other base pairs present, the data points of which are derived from a separate experiment. In our example, the data points of the selfpairing could be subtracted from the data points of mixed pairing, although the actual concentration of the selfpair in the *A* series was different from the same pair in the *AB* series, as can be clearly seen in *Fig. 12*. The difference in the actual concentrations merely manifests itself in the assymptotes $x(T)$ of the isotherms of the reduced data set (*Fig. 3c*), not in their curvature. Therefore, this second parameter $x(T)$ not only accounts for the translation of the chemical-shift differences into α, it is also dependent on the chemical nature of the H-bond (O···H–N *vs.* N···H–N) *and* the differential actual concentrations of base pairs that are stabilized by different kinds of H-bonds. Hence, no assumption with respect to the temperature dependence of x was made.

It is interesting how x depends on T. As already mentioned, x accounts – albeit not exclusively – for the chemical-shift difference between monomeric and fully complexed molecules and can by derived from $\lim_{c_A \to \infty} ppm - ppm°$ and $\lim_{c_A \to \infty} \Delta ppm$. The selfpairing data from the *A* series suggest $x(A)$ to be essentially constant between 0 and 30°; it decreases steadily at higher temperatures (*Fig. 16*, lower data points). The constant value of somewhere around 2.83 ppm originates from the AA_2 pairing and, given the fact that the assymptote represents an extrapolated value at infinite concentrations, shows only small deviations. $x(A)$ decreases steadily at higher temperatures where the contribution from

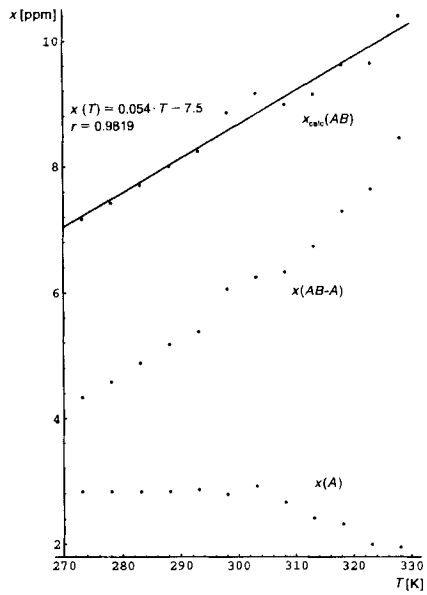


Fig. 16. *Temperature dependence of* x

the AA_1 pairing becomes dominant. The $(AB - A)$ data that were used to calculate the T·A pairing produced $x(AB - A)$ values that are far from constant, in fact they range from *ca.* 4.3 to 8.4 ppm (*Fig. 16*, middle data points). The positive temperature dependence is probably linear in the lower temperature range where the AB_2 pairing predominates, but drifts away from linearity in the temperature range where AB_1 takes over.

The reason, why $x(AB - A)$ is a different function of T (and $\Delta G°$, not shown) than $x(A)$ is based on the fact that this differential data set is inherently different from the selfpaired data set. Subtracting selfpairing from mixed pairing $(AB - A)$ involves the subtraction of a concentration-dependent curvature from another one, for each isotherm separately. From the addition of the x values from both fitted data sets it seems that the temperature dependence of x derived from the original mixed AB data, $x_{calc}(AB) = x(A) + x(AB - A)$, is linearly dependent on T. A linear regression produces a straight line through the combined data points for $x_{calc}(AB)$ with $r = 0.982$ revealing a temperature dependence of $x_{calc}(AB)$ of 0.054 ppm K^{-1} (*Fig. 16*, upper data points). Indeed, fitting *Eqn. 7* to the data points from the AB series results in $x(AB)$ values that show a linear, albeit higher temperature dependence over the whole temperature range ($x_{fit}(AB) = 0.096 \cdot T - 19.7$, $r = 0.985$, not shown). The difference between $x_{calc}(AB)$ and $x_{fit}(AB)$ originates from the fact that fitting the $AB - A$ data points involved a no-intercept expression (*Eqn. 31*). The constraint of forcing the isotherms through the origin narrows down the range of $x(AB - A)$ values. Releasing this constraint by fitting the $AB - A$ data with an additional intercept parameter *ppm*° (as in *Eqn. 32*, but using *Eqn. 7* instead of *Eqn. 15* for α) produces a similarly shaped but steeper temperature dependence of $x'(AB - A)$ (not shown). Addition of these values to the original $x(A)$ assymptotes results in new $x'_{calc}(AB)$ values which lie within the uncertainty limits of $x_{fit}(AB)$

($x'_{calc}(AB) = x(A) + x'(AB - A) = 0.093 \cdot T - 18.9, r = 0.983$). Hence, the assymptotes x not only relate to the physical and chemical nature of the measured system, they also depend on the details of the fitting procedure.

Direct fittings using *Eqns. 8* and *16* instead of *Eqns. 7* and *15*, respectively, *i.e.* using directly $\Delta H°$ and $\Delta S°$ instead of $\Delta G°$ as the parameters to be optimized, would presumably produce the same results, provided that x were truly temperature-independent. On the other hand, the temperature dependence of $\Delta G°$ is precisely what is asked for, if one wishes to calculate $\Delta H°$ and $\Delta S°$. The subsequent linear regression is a correct procedure in the statistical sense, because no logarithmic or other transformations are required. Therefore, the standard deviations are legitimate error ranges for $\Delta H°$ and $\Delta S°$, and the coefficient r represents a sensitive measure for the temperature independence of $\Delta H°$ and $\Delta S°$. In addition, the division into an isotherm fit and a linear regression renders more transparency to the whole procedure.

The presented fitting procedure relies on dilution causing equilibrium shifts rather than temperature differences[5]) or molar ratios (titrations)[6]). It also omits the *van't Hoff* analysis (*cf.* next *Sect.*). Therefore, it should be an alternative to known calorimetric and spectroscopic titration methods, particularly to the differential calorimetric method, when the thermodynamics of relatively weak complexes are measured (short and/or mispaired DNA/RNA fragments) or when complexation is accompanied by a small entropic penalty (supramolecular complexes made of rigid, structurally and conformationally 'pre-formed' monomers). The differential calorimetric method generates reliable enthalpy differences (if sufficiently large) but less reliable, because often ill-defined, entropy differences (few data points define the 'slope' of a transition). The linear regression of $\Delta G°$ *vs.* T generates entropy differences through well defined slopes (in the measured temperature range) but enthalpy changes through the extrapolation of the data points to zero K.

Thermodynamics. The regression plots of $\Delta G°$ *vs.* T revealed in both pairing systems two equilibria involving H-bonded complexes that were stabilized by a weaker and a stronger *Gibbs* free energy of interaction. It seems obvious that the thermodynamics of the stronger pairings, designated $\Delta H°_2$ and $\Delta S°_2$, correspond to the base pairs as they are known from many studies: cyclic dimers involving two H-bonds each. Despite the lower entropic penalty for the cyclic T·T dimer, the stability of the selfcomplementary base pair at 25° is lower than the cyclic A·T dimer by 0.80 ± 0.41 kcal mol^{-1} ($\Delta G°$ (25°) $= -1.66 \pm 0.3$ *vs.* -2.47 ± 0.1 kcal mol^{-1}). This is due to the substantially decreased enthalpy of formation of a thymine-thymine pair.

The thermodynamics of the weak pairings ($\Delta H°_1$ and $\Delta S°_1$) derived from the high-temperature isotherms suggest that they correspond to more open and disordered complexes than the strong ones. The entropic penalty is more than halved in the A·T pair, $\Delta S°_1 = -13.3 \pm 1.2$ *vs.* $\Delta S°_2 = -23.6 \pm 0.9$ cal mol^{-1} K^{-1}, whereas it virtually vanishes in the T·T pair: $\Delta S°_1 = 0.6 \pm 0.8$ *vs.* $\Delta S°_2 = -16.9 \pm 0.8$ cal mol^{-1} K^{-1}. Interestingly, the

---

[5]) All attempts to fit melting curves $\alpha(T)_{c_A}$ – instead of the isotherms $\alpha(c_A)_T$ – onto the data points failed because of the existence of two equilibria over the whole temperature range. To fit two equilibria would mean to introduce more parameters ($ppm = ppm° + x \cdot \alpha(\Delta G°_{AA1}) \cdot \alpha(\Delta G°_{AA2})$).

[6]) Titrations of weak complexes may involve too high concentrations of one or both compounds which could cause a significant departure from ideal-solution conditions.

enthalpy changes less dramatically in the A·T pair, $\Delta H^\circ_1 = -6.4 \pm 0.4$ *vs.* $\Delta H^\circ_2 = -9.5 \pm 0.3$ kcal mol^{-1}, than in the T·T pair: $\Delta H^\circ_1 = -1.4 \pm 0.3$ *vs.* $\Delta H^\circ_2 = -6.7 \pm 0.2$ kcal mol^{-1}.

N–H···O bonds are generally thought to be more stable than N–H···N bonds. A closer look at published standard enthalpies of H-bond formation, however, does not confirm this general assumption. The standard enthalpy of formation of an indole-pyridine, a pyrrole-pyridine H-bond in CCl_4, and an aniline-pyridine H-bond in hexane were reported to be -3.6 ± 1.2, -3.2, and -3.43 kcal mol^{-1}, respectively. In contrast, γ-butyrolactam and pyridin-2(1*H*)-one in CCl_4 were reported to selfpair with a standard enthalpy of -3.5 ± 0.4 and -4.4 ± 0.4 kcal mol^{-1} involving two N–H···O=C bonds each (p. 20–122 in [19]). IR-Monitored selfpairing and pairing enthalpies of 1-cyclohexyluracil and 9-ethyladenine in $CHCl_3$ were reported to be -4.3 ± 0.4 (U·U) and -6.2 ± 0.6 kcal mol^{-1} (U·A), respectively [2]. The corresponding entropy changes were -11.0 ± 1.0 and -11.8 ± 1.2 cal mol^{-1} K^{-1}. The enthalpy difference between homo and hetero pairing was 1.9 ± 1.0 kcal mol^{-1} *vs.* 2.8 ± 0.6 kcal mol^{-1} in this study. Moreover, the corresponding entropy difference was 0.8 ± 2.2 cal mol^{-1} K^{-1} *vs.* 6.7 ± 1.9 cal mol^{-1} K^{-1} in this study. The latter might be due to a significant entropy difference between alkyl and ribosyl substituents. Other published selfassociation thermodynamics of uracil derivatives in $CHCl_3$ and the calculated association energies of thymine and adenine using quantum-mechanical methods agree with our values to different degrees. The 1-cyclohexyluracil was reported to self-associate with a pairing enthalpy of -5.3 kcal mol^{-1} and a pairing entropy of -14.7 cal mol^{-1} K^{-1} as measured by ^{1}H-NMR spectroscopy [3e]. From a combined ^{13}C- and ^{17}O-NMR spectroscopic study investigating the hydration of 2′,3′-*O*-isopropylideneuridine in wet MeCN [6], a standard selfpairing enthalpy of -10.1 ± 1.2 kcal mol^{-1} was calculated (extrapolated to zero H_2O content). Using the atomic-dipole approximation, the selfpairing energy of thymine in the $O^4 \cdot O^4$ and $O^2 \cdot O^2$ reversed-wobble arrangement was calculated to be -5.21 and -3.73 kcal mol^{-1}, respectively, the adenine-thymine pairing energy in the *Watson-Crick* geometry -7.00 kcal mol^{-1} [18c].

Hence, the quoted ΔH° values among themselves and compared to the ones derived from the low-temperature isotherms measured in this study (ΔH°_2) differ by roughly ± 30–50%. Our selfpairing enthalpy ($\Delta H^\circ_{AA_2}$) (-6.7 kcal mol^{-1}) is somewhere between the extremes (-4.3 and -10.1 kcal mol^{-1}), while the hetero pairing ($\Delta H^\circ_{AB_2} = -9.5$ kcal mol^{-1}) appears somewhat more stable than in other studies. Neglection of the biphasic nature of the ΔG°_{AB} *vs.* *T* plot (*Fig. 9*) would have resulted in a lower (averaged) pairing enthalpy of -7.8 kcal mol^{-1} [7]). Recent calorimetric titration studies involving strong neutral H-bonds in organic solvents, *e.g.* pairings between diamines and diols in benzene, seem to confirm an enthalpy change of *ca.* 4–5 kcal mol^{-1} per H-bond [20].

The strong concentration dependence of the high-temperature isotherms indicates that both observed 'weak' equilibria (*AB* and *A* series) are at least bimolecular. Therefore, neither conformational changes owing to intramolecular interactions (first-order equilibrium), nor solvation-desolvation equilibria (*pseudo*-first-order at the applied concen-

---

[7]) Neglection of the influence of selfpairing in the *AB* series, *i.e.*, an intercept fitting of *Eqn. 7* onto the *AB* data points (*Fig. 3b*) would suggest pairing enthalpies of $\Delta H^\circ_2 = -8.5 \pm 0.4$ and $\Delta H^\circ_1 = -4.4 \pm 0.9$ kcal mol^{-1} and entropies of $\Delta S^\circ_2 = -21.3 \pm 0.4$ and $\Delta S^\circ_1 = -7.7 \pm 2.8$ cal mol^{-1} K^{-1}.

trations) could account for the shift differences. A bimolecular pairing equilibrium involving only one H-bond and base-base stacking equilibria of possibly even higher molecularity could, in principle, give rise to the observed high-temperature isotherms. The former was suggested by *Rich* and coworkers to be the most likely process in $CHCl_3$ ('open base pairs' [2]). The authors mentioned in their IR study that the linearity of their plots of ln K *vs.* $1/T$ for the 1-cyclohexyluracil · 1-cyclohexyluracil and the 1-cyclohexyluracil · 9-ethyladenine pair indicated that only cyclic dimers formed, but the same plot for the 9-ethyladenine dimer showed some degree of a biphasic behavior comparable to this study (Fig. 3 in [2]). A two-step process involving an open and cyclic 9-ethyladenine dimer could not be ruled out.

It is conceivable that ln K *vs.* $1/T$ regression plots are less sensitive to the departure from linearity than those of $\Delta G°$ *vs.* T, if not the plots themselves then so the resulting correlation coefficients. A comparative regression analysis of the fitted data derived from the $AB - A$ series, ln $K = -\Delta G°_{AB}/RT$ *vs.* $1/T$ including the whole temperature range, produced the same thermodynamics within the standard deviation, but a seemingly better correlation coefficent than the $\Delta G°$ *vs.* T regression ($r = 0.993$ *vs.* 0.985). If an unnoticed biphasic dependence was analyzed as one straight line, the outcoming enthalpy and entropy changes would correspond to averaged values for open and cyclic forms, thus, would underestimate the thermodynamics of cyclic-complex formation.

In addition, one has to bear in mind that the enthalpy and entropy changes obtained are not simply differences resulting from the formation of H-bonds. They express the entire heat of formation of solvated dimer from solvated monomers. Therefore, different enthalpies of H-bond formation derived from different molecules should be compared with caution.

Structural Interpretations. The structures of the base pairs depicted in *Scheme 1* are based an general assumptions. *Rich* and coworkers [2] concluded from their quantitative IR-spectroscopic studies that U · U selfpairing was dominated by one particular geometry rather than a mixture of the three (as in a-, b-, and c-**1** · **1**; most probably the one corresponding to a-**1** · **1**). Furthermore, A · U pairing was similarly dominated by one of the four possible geometries, yet the authors could not determine which one. Although the *Watson-Crick* arrangement is the predominant one for the A · T or A · U pair in DNA and RNA oligomers, the monomeric compounds are not as constrained in their pairing geometry. Cocrystal structures of 1:1 mixtures of related compounds gave no conclusive indication, since many possible contacts were found (refs. cit. in [2]).

A subsequent quantification of the IR-monitored association of 1-cyclohexylthymine and 9-ethyladenine in $CHCl_3$ revealed association constants under the same conditions: $K^{IR}_{2T\rightleftharpoons T\cdot T} = 3.2$ and $K^{IR}_{T+A\rightleftharpoons T\cdot A} = 130$ [18a]. Later, the thermodynamics of these pairings were calculated from ^{13}C-NMR spectroscopic data [5]. The concentration-dependent shifts of both carbonyl groups of 1-cyclohexylthymine were used to calculate the corresponding association constants under the same conditions: $K^{13C}_{2T\rightleftharpoons T\cdot T}(O^4\cdot O^4) = 4.2 \pm 0.2$, $K^{13C}_{2T\rightleftharpoons T\cdot T}(O^2\cdot O^2) = 2.2 \pm 0.1$, $K^{13C}_{T+A\rightleftharpoons T\cdot A}(O^4) = 60 \pm 5$ and $K^{13C}_{T+A\rightleftharpoons T\cdot A}(O^2) = 73 \pm 4$. These values agreed fairly well with the IR-derived values. They were also the first experimental data that quantified the preference of the $O^4\cdot O^4$ over the $O^2\cdot O^2$ reverse-wobble pairing in the selfassociation of a thymine derivative and the preference of the reverse (O^2-bound) over normal (O^4-bound) thymine-adenine pairing, albeit, again, not in which geometry, *Watson-Crick* or *Hoogsteen*.

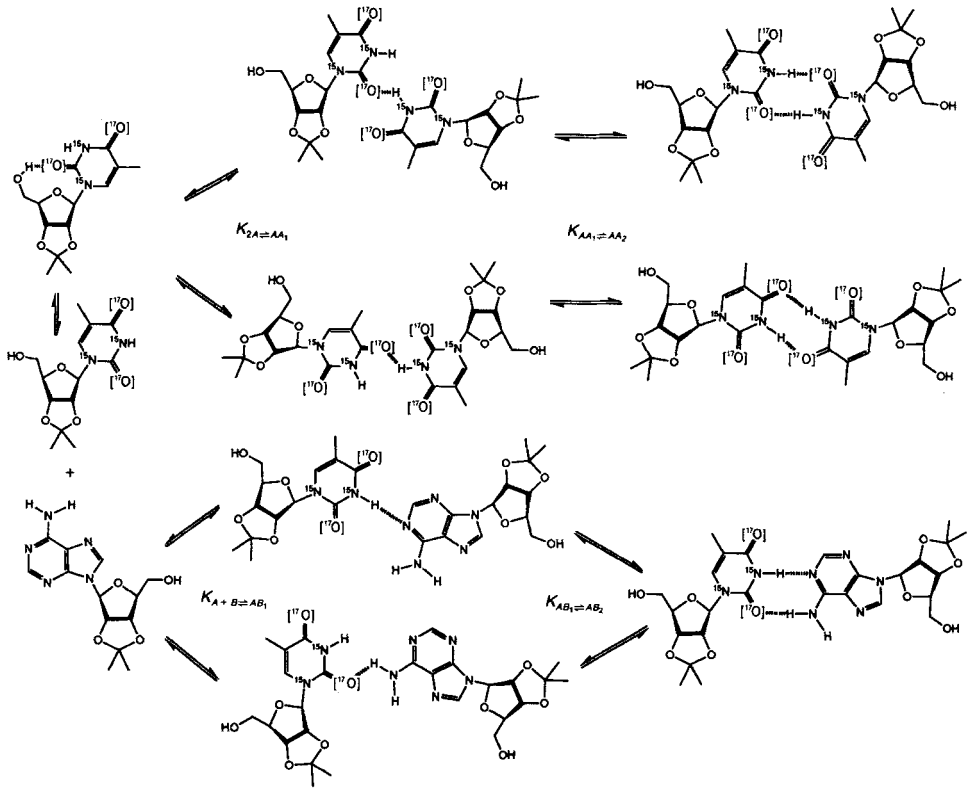
The $O^4 \cdot O^4$ reversed-wobble geometry of the selfpair does agree with the solid-state IR spectrum of 1-cyclohexyluracil, with several related crystal structures – among them 3′,5′-di-*O*-acetylthymidine [18b] – and with the ^{13}C-NMR spectroscopic study of 2′,3′-*O*-isopropylideneuridine in MeCN [6]. Yet, a correlation between ^{13}C and ^{17}O chemical shifts of the C(4)=O group of the latter compound revealed that the ^{13}C shifts were not related to the degree of H-bonding in a linear fashion, suggesting that quantitative conclusions based on ^{13}C-NMR-derived data may be subject to error. An indication for which geometry of the adenine-thymine pair is dominant may be obtained from additional investigations: experiments involving the protonation of adenosine revealed that the *Watson-Crick* site of the purine nucleoside is more basic than the *Hoogsteen* site with N(1) being the most basic N-atom [21]. Several calculations suggested that N(3) of adenine and adenosine might be a second protonation site leaving N(7), the '*Hoogsteen* atom' to the end of the hierarchy [22]. These studies suggest that the *Watson-Crick* or reverse-*Watson-Crick* pairs (d- and e-**1** · **2**) are more likely candidates than the *Hoogsteen* or reverse-*Hoogsteen* pairs (f- and g-**1** · **2**).

In this study, the temperature dependence of the ^{17}O chemical shifts and the half-intensity widths of the O^2- and O^4-atoms of **1** were measured in the presence (*AB* series) and absence (*A* series) of an equimolar amount of unlabelled **2** (46 mM each). Comparably inaccurate as the chemical shift measurements were, they nevertheless allowed us to draw some interesting conclusions. In $CHCl_3$, the temperature-dependent chemical-shift differences of the O^2 and O^4 signals suggest that, at higher temperatures, the T · A pair forms surprisingly stable open base pairs involving only one H-bond between the NH_2 group of adenine and the O^2 rather than O^4 carbonyl O-atom of thymine. Open base pairs that are stabilized by N–H · · · N bonds cannot be ruled out. At lower temperatures, a cyclic reverse base pair, most probably in the reverse-*Watson-Crick* geometry, dominates.

The geometry of neither the weak nor the strong T · T selfpair in the *A* series can be deduced from the ^{17}O chemical shifts. Other studies ([6] and refs. cit. therein) suggest that it could be the $O^2 \cdots HO$–C(5′) intramolecular interaction in monomeric **1** that overlaps any base-pairing effects. The interaction forces the base into the *syn*-glycosidic conformation. Although imaginable, it is impossible to determine from the ^{17}O-NMR data whether this first-order equilibrium also exists in the open and cyclic base pairs. In analogy to the weak hetero pairing, the weak selfpairing could be stabilized by one H-bond between NH and O^2 and O^4, respectively, forming rather labile but entropically favored selfpairs. At lower temperatures, the base pairs form cyclic dimers, both, in the $O^2 \cdot O^2$ and $O^4 \cdot O^4$ reverse-wobble geometry, as suggested by several other studies. *Scheme 2* summarizes these interpretations.

The apparently higher sensitivity of the ^{17}O chemical shifts and half-intensity widths of the C(4)=O group upon H-bonding, as compared to the C(2)=O group, is usually explained by its higher ground-state π-bond order and other electronic factors ([6] [15a] and ref. cit. therein). This physical property is also the reason for O^4 to be the favored protonation site in uridine derivatives [22a, b]. Therefore, one wonders why the adenosine derivative favors the complexation of the C(2)=O group? Moreover, if the enthalpies of formation of $O^2 \cdot O^2$ and $O^4 \cdot O^4$ reverse-wobble pairs are roughly equal or at least similar, why do $O^2 \cdot O^4$ wobble selfpairs not occur? The reason might be based on the difference in symmetry of the reverse *vs.* normal base pairs.

Scheme 2. *Selfpairing and Pairing Equilibria between the 2',3'-O-Isopropylidene Derivatives of 5-Methyluridine (T) and Adenosine (A) in Chloroform*



The relation between symmetry and the stability of molecules can be described by the theory of statistical thermodynamics. The statistical entropy S of a macroscopic ensemble can be calculated from the quantum-mechanical degrees of freedom of a single molecule by *Eqn. 33*, k being the *Boltzman* constant, ΔU the internal energy, and Q the canonical (*in vacuo*) partition function of the system under investigation. $Q = q^N$ for a mixture of distinguishable molecules, and $Q = q^N/N!$ for non-distinguishable species. N is the number of molecules ($N \cdot k = n \cdot R$, n being the number of mol and R the universal gas constant), and q is the molecular partition function. The *Gibbs* free-energy difference ΔG of the ensemble is then given by *Eqns. 34*.

$$S = \Delta U/T - k \ln Q \tag{33}$$

$$\Delta G = -nRT \ln q \text{ (distinguishable) and } \Delta G = -nRT \ln (q/N) \text{ (non-distinguishable)} \tag{34}$$

The molecular partition function q is dependent on the energy contributions from all possible molecular degrees of freedom. It is calculated from the translational, rotational, vibrational, and electronic contributions to the internal energy of the molecule:

$q = q^{trans} \cdot q^{rot} \cdot q^{vib} \cdot q^{elec}$. The translational, vibrational, and electronic terms are symmetry-independent, they are only dependent on the ensemble volume, the molecular mass ($V/m^{3/2}$), the vibrational frequencies, and the degeneracy of the electronic ground state, respectively. The rotational contribution is derived from *Eqn. 35* where A, B, and C are the rotation constants (being inversely proportional to the principal moments of inertia $I_{A,B,C}$) and $\sigma = 1,2,3,\ldots$ is the symmetry number. The temperature dependence cancels out with the same inverse-temperature dependence of the translational component. The symmetry number σ, however, relates through *Eqns. 35* and *33* the symmetry group to which the molecule belongs with its contribution to entropy S of the ensemble (*e.g.* $\sigma(C_1) = 1$, $\sigma(C_2) = 2$, *etc.*). The internal-energy difference ΔU itself, in principle, does depend on $[\partial \ln Q/(\partial(1/kT)]_V$, but at temperatures that are well above 0 K, it becomes constant with respect to q^{rot}, since all rotational energy levels are then equally available (loss of quantum-mechanical effects). Hence, if we wish to calculate the statistical entropy of a selfpair *in vacuo* (*e.g.* **1**·**1**) in two particular geometries, we assume that all energetic contributions to the entropy of the pair remain constant except for the contribution from rotational freedoms (*Eqn. 36*). If we further assume that the difference in the rotation constants between both geometries is negligible for S^{rot}, we are able to estimate the effect of symmetry on the entropy. The entropy difference between a molecular complex belonging to one particular symmetry group and the 'same' complex belonging to another one is proportional to $\ln \sigma_1 - \ln \sigma_2$.

$$q^{rot} = \frac{1}{\sigma}\sqrt{\left(\frac{kT}{hc}\right)^3 \frac{\pi}{ABC}} \tag{35}$$

$$S^{rot} = nR\left(\text{const.} + \ln(\sigma\sqrt{ABC})\right) \quad \text{or} \quad \Delta S^{rot} = nR \ln\left(\sigma\sqrt{ABC}\right) \tag{36}$$

The selfpair **1**·**1** can either be a wobble or a reverse-wobble pair. The selfpair in the wobble geometry (c-**1**·**1** in *Scheme 1*) belongs to the symmetry group C_1 while the same selfpair in the reverse-wobble geometry (a- and b-**1**·**1** in *Scheme 1*) belongs to the C_2 group. According to *Eqn. 36*, the entropy difference *in vacuo* between both geometries ΔS^{rot} equals $nR \ln 2 = 1.38$ cal mol^{-1} K^{-1}, a symmetry-derived contribution to the overall stability rendering the reverse-wobble geometry more stable owing to its smaller entropic penalty for base-pair formation. The equilibrium constant of selfpairing is twice as high for the reverse geometry and, at room temperature, the *Gibbs* free-energy difference amounts to ΔG (298 K) = 0.41 kcal mol^{-1} which is *ca.* 25% of the corresponding *Gibbs* free-energy difference for the selfassociation of **1**.

Hence, statistical thermodynamics merely involving *in vacuo* partition functions support a predominance of selfpairing geometries of higher symmetry. The symmetry-related enhanced stability of the reverse-wobble pair indicates why, despite of a possibly identical enthalpy of formation, a homobase pair like **1**·**1**, uracil·uracil, or thymine·thymine chooses to adopt the mixed $O^4 \cdot O^4$ and $O^2 \cdot O^2$ reverse-wobble rather than the $O^2 \cdot O^4$ wobble geometry. Note that the same conclusion is valid for polymeric homo-paired nucleic acids such as, presumably, double-stranded polyadenylic acid at acidic pH. X-Ray fibre-diffraction data suggested that the strands were oriented parallel [23].

However, in homo and hetero pairs involving the *Watson-Crick* or *Hoogsteen* geometries such as **1**·**2**, both reverse and normal geometries belong to the C_1 symmetry

group containing only a C_2 pseudosymmetry axis (in an idealized pairing geometry: C_2 with respect to the ribose moieties, C_1 with respect to the bases). Therefore, if there is a difference in rotational entropy between both geometries, ΔS^{rot} equals $0.5 \cdot R \cdot (\ln (A_2B_2C_2) - \ln (A_1B_1C_1))$ (according to *Eqn. 36*), a presumably negligible contribution to the overall entropy difference. It follows that the apparently strong preference of **1**·**2** for the reverse over the normal pairing geometry must either originate in differential enthalpic rather than entropic contributions, which agrees with the favored hydration but not protonation site to be O^2 [22a, b], or in a significant entropic effect due to the higher symmetry of the ribose moieties of the reverse base pair. The role of pseudosymmetry in statistical thermodynamics seems unclear.

^{17}O-NMR Spectroscopy of isotope-enriched nucleosides and other compounds has proven to be a valuable method for the elucidation of local structural properties of H-bonded systems. In this study, some structural information was obtained from the temperature dependence of the ^{17}O chemical shifts. The corresponding half-intensity widths are, by virtue of their accuracy, potentially useful for the calculation of the thermodynamics of base pairing. Further studies with higher-molecular weight compounds under aqueous conditions are planned, to learn whether this nucleus could be used in RNA strands as a local marker for the monitoring of secondary- and tertiary-structure formation.

The financial support from the *Swiss National Science Foundation* is gratefully acknowledged.

Experimental Part

Solns. of **1** and **2** were pre-dried over activated molecular sieves (4 Å). $CDCl_3$ and $CHCl_3$ were filtered over neutral Al_2O_3 (act. I) prior to use. NMR Spectra: *Varian-VXR-400*. ^{15}N-NMR (41 MHz): see preceding paper; post-acquisition delay during temp.-dependent measurements: 500 s. ^{17}O-NMR (54 MHz): see preceding paper; post-acquisition delay during temp.-dependent measurements: 500 s; line broadening: 50–200 MHz; $w_{1/2}$ was measured directly when the signal separation was sufficiently large, otherwise the downfield half-widths and the upfield half-widths of the O–C(4) and O–C(2) resonances at half their intensities, resp., were doubled.

Calculations, fitting procedures, and graphical plots were perfomed with a *Macintosh*™ version of *Mathematica*® from *Wolfram Research, Inc.* In the $AB - A$ series, function $\alpha(c_A, T, \Delta G°)$ from *Eqn. 7* was first separately solved for each measured temp. The resulting functions $\alpha(c_A, \Delta G°)_T$ were then separately solved for several $\Delta G°$ values in steps of 10 cal mol^{-1} to give $\alpha(c_A)_{T,\Delta G°}$. The experimental isotherms were fitted with the input expression

Table 1. $\delta_{N(3)}$ [ppm] *in* AB *Series*

T [°C]	46.13 mM	23.06 mM	11.53 mM	5.77 mM	2.88 mM
55	135.177	133.867	132.849	132.058	131.611
50	135.253	134.000	132.971	132.189	131.724
45	135.422	134.146	133.093	132.331	131.849
40	135.458	134.335	133.257	132.482	131.956
35	135.622	134.521	133.441	132.665	132.111
30	135.827	134.716	133.603	132.860	132.272
25	136.046	134.946	133.884	133.109	132.452
20	136.095	135.147	134.137	133.397	132.728
15	136.292	135.423	134.432	133.676	132.981
10	136.441	135.674	134.712	133.991	133.250
5	136.585	135.908	135.026	134.307	133.562
0	136.718	136.111	135.305	134.615	133.863

Table 2. $\delta_{N(3)}$ [ppm] *in* A *Series*

T[C°]	46.13 mM	23.06 mM	11.53 mM	5.77 mM	2.88 mM	1.44 mM	extr. →0
55	131.704	131.450	131.288	131.142	131.029	130.977	130.895
50	131.819	131.562	131.350	131.205	131.101	131.012	130.957
45	131.976	131.686	131.443	131.298	131.126	131.046	131.033
40	132.094	131.804	131.557	131.383	131.240	131.167	131.097
35	132.264	131.947	131.671	131.468	131.302	131.230	131.175
30	132.432	132.107	131.788	131.555	131.414	131.316	131.243
25	132.589	132.248	131.930	131.713	131.532	131.391	131.324
20	132.785	132.447	132.130	131.850	131.648	131.540	131.480
15	132.950	132.618	132.297	132.005	131.785	131.661	131.580
10	133.140	132.779	132.483	132.176	131.914	131.755	131.679
5	133.308	132.987	132.668	132.355	132.055	131.908	131.778
0	133.477	133.123	132.865	132.495	132.201	132.017	131.882

'Fit[$\{c_{Ai}, y_i\}$, $\{\alpha(c_A)_{T,\Delta G°}\}$, c_A]' for the no-intercept fitting involving the data points $y_i = \Delta\delta_{N(3)}(AB - A)$. The output expression '$x \cdot \alpha(c_A)_{T,\Delta G°}$' was followed by the input expression 'Limit[%, $c_A \rightarrow$ Infinity]' for the confirmation of $x_{T,\Delta G°}$. The other fitting (A series) was performed with the input expression 'Fit[$\{c_{Ai}, y_i\}$, $\{1, \alpha(c_A)_{T,\Delta G°}\}$, c_A]' involving the data points $y_i = \delta_{N(3)}(A)$ and function $\alpha(c_A)_{T,\Delta G°}$ derived accordingly from *Eqn. 15*. Following output expression '$ppm° + x \cdot \alpha(c_A)_{T,\Delta G°}$', $x_{T,\Delta G°}$ was calculated from the difference between 'Limit[%, $c_A \rightarrow$ Infinity]' and $ppm_T°$, the corresponding assymptote and intercept, resp. The $\Delta G°$ values that produced the smallest $\chi^2 = \Sigma[y_i - x \cdot \alpha(c_{Ai})_{T,\Delta G°}]^2$ and $\Sigma[y_i - (ppm° + x \cdot \alpha(c_{Ai})_{T,\Delta G°})]^2$ for each isotherm were chosen for a linear regression *vs.* T.

REFERENCES

[1] E. Kuechler, J. Derkosch, *Z. Naturforsch., B* **1966,** *21,* 209.

[2] Y. Kyogoku, R.C. Lord, A. Rich, *J. Am. Chem. Soc.* **1967,** *89,* 496.

[3] a) R.R. Shoup, H. Todd Miles, E.D. Becker, *Biochem. Biophys. Res. Commun.* **1966,** *23,* 194; b) L. Katz, S. Penman, *J. Mol. Biol.* **1966,** *15,* 220; c) R.A. Newmark, C.R. Cantor, *J. Am. Chem. Soc.* **1968,** *90,* 5010; d) S.M. Wang, N.C. Li, *ibid.* **1968,** *90,* 5069; e) G. Lancelot, C. Hélène, *Nucleic Acids Res.* **1979,** *6,* 1063.

[4] a) C.D. Poulter, C.L. Livingston, *Tetrahedron Lett.* **1979,** 755; b) C. Dyllick-Brenzinger, G.R. Sullivan, P.P. Pang, J.D. Roberts, *Proc. Natl. Acad. Sci. U.S.A.* **1980,** *77,* 5580; c) G. Barbarella, A. Bertoluzza, V. Tugnoili, *Magn. Reson. Chem.* **1987,** *25,* 864.

[5] H. Iwahashi, Y. Kyogoku, *J. Am. Chem. Soc.* **1977,** *99,* 7761.

[6] H.M. Schwartz, M. MacCoss, S.S. Danyluk, *J. Am. Chem. Soc.* **1983,** *105,* 5901.

[7] M.I. Burgar, D. Dhawan, D. Fiat, *Org. Magn. Reson.* **1982,** *20,* 184.

[8] X. Gao, R.A. Jones, *J. Am. Chem. Soc.* **1987,** *109,* 3169; B. Goswami, B.L. Gaffney, R.A. Jones, *ibid.* **1993,** *115,* 3832.

[9] M.P. Schweizer, S.I. Chan, P.O.P. Ts'o, *J. Am. Chem. Soc.* **1965,** *87,* 5241.

[10] R. Hamlin, R.C. Lord, A. Rich, *Science* **1965,** *148,* 1734; Y. Kyogoku. R.C. Lord, A. Rich, *ibid.* **1966,** *154,* 518.

[11] L. Marky A., K.J. Breslauer, *Biopolymers* **1987,** *26,* 1601.

[12] a) M. Petersheim, D.H. Turner, *Biochemistry* **1983,** *22,* 256; b) M. Petersheim, D.H. Turner, *ibid.* **1983,** *22,* 269.

[13] J. Gralla, D.M. Crothers, *J. Mol. Biol.* **1973,** *73,* 497.

[14] B.E. Fischer, H. Sigel, *J. Am. Chem. Soc.* **1980,** *102,* 2998; I. Horman, B. Dreux, *Helv. Chim. Acta* **1984,** *67,* 754; M.D. Cowart, I. Sucholeiki, R.R. Bukownik, C.S. Wilcox, *J. Am. Chem. Soc.* **1988,** *110,* 6204; K. Kobayashi, Y. Asakawa, Y. Kato, Y. Aoyama, *ibid.* **1992,** *114,* 10307; Y. Kato, M.M. Cann, J. Rebek, *ibid.* **1994,** *116,* 3279; T.R. Kelly, M.H. Kim, *ibid.* **1994,** *116,* 7072; Y. Kato, M.M. Conn, J. Rebek, *Proc. Natl. Acad. Sci. USA* **1995,** *92,* 1208; J. Cuntze, L. Owens, V. Alcázar, P. Seiler, F. Diederich, *Helv. Chim. Acta* **1995,** *78,* 367.

[15] a) H.M. Schwartz, M. MacCoss, S.S. Danyluk, *Tetrahedron Lett.* **1980,** *21,* 3837; b) A. Steinschneider, M.I. Burgar, A. Buku, D. Fiat, *Int. J. Pept. Protein Res.* **1981,** *18,* 324; c) B. Valentine, A.T.E. St, D. Fiat, *Org. Magn. Reson.* **1984,** *22,* 697; d) S. Chandrasekaran, W.D. Wilson, D.W. Boykin, *J. Org. Chem.* **1985,** *50,* 829; e) A. Spisni, E.D. Gotsis, D. Fiat, *Biochem. Biophys. Res. Commun.* **1986,** *135,* 363; f) H. Eckert, D. Fiat, *Int. J. Pept. Protein Res.* **1986,** *27,* 613; g) G. Jaccard, J. Lauterwein, *Helv. Chim. Acta* **1986,** *69,* 1469; h) G. Jaccard, P.A. Carrupt, J. Lauterwein, *Magn. Reson. Chem.* **1988,** *26,* 239; i) E. Ponnusamy, H. Eckert, D. Fiat, *Int. J. Pept. Protein Res.* **1988,** *32,* 21; j) D.W. Boykin, D.W. Sullins, N. Pourahmady, E.J. Eisenbraun, *Heterocycles* **1989,** *29,* 307; k) J.A. Wilde, P.H. Bolton, A. Mazumder, M. Manoharan, J.A. Gerlt, *J. Am. Chem. Soc.* **1989,** *111,* 1894; l) C. Sakarellos, I.P. Gerothanassis, N. Birlirakis, T. Karayannis, D.M. Sakarellos, M. Marraud, *Biopolymers* **1989,** *28,* 15; m) M. Schumacher, J. Lauterwein, *J. Magn. Reson.* **1989,** *83,* 97; n) H. Dahn, P. Péchy, T.V. Van, *Angew. Chem.* **1990,** *102,* 681; o) J.C. Zhuo, H. Wyler, P. Péchy, H. Dahn, *Helv. Chim. Acta* **1994,** *77,* 317.

[16] 'NMR-Spektroskopie von Nichtmetallen: Grundlagen, ^{17}O-, ^{33}S- und ^{129}Xe-NMR-Spektroskopie', Eds. S. Berger, S. Braun, and H.-O. Kalinowski, Thieme Verlag, Stuttgart, 1992, Vol. 1.

[17] J. Reuben, *J. Am. Chem. Soc.* **1969,** *91,* 5725.

[18] a) Y. Kyogoku, R.C. Lord, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **1967,** *57,* 250; b) C.C. Wilson, J.N. Low, P. Tollin, H.R. Wilson, *Acta Crystallogr., Sect. C* **1984,** *40,* 1712; c) Z.G. Zudritskaya, V.I. Danilov, *J. Theor. Biol.* **1979,** *59,* 303.

[19] S.N. Vinogradov, R.H. Linnell, 'Hydrogen Bonding', Van Nostrand Reinhold Company, New York, 1971.

[20] J.-N. Aebischer, N. Ghoneim, E. Haselbach, 'Calorimetric Investigation of Groundstate Diastereotopic Interactions in H-Bonded Complexes' in Abstracts of '2nd CHiral2 Workshop' – Gwatt (March 95).

[21] H. Lönnberg, in 'Biocoordination Chemistry: Coordination Equilibria in Biologically Active Systems', Ed. K. Burger, Ellis Horwood, London, 1990, pp. 284.

[22] a) J. Del Bene, *J. Phys. Chem.* **1983,** *87,* 367; b) S. Miertus, M. Trebaticka, *Collect. Czech. Chem. Commun.* **1983,** *48,* 3517; c) C. Glemarec, Y. Besidsky, J. Chattopadhyaya, J. Kusmierek, M. Lahti, M. Oivanen, H. Lönnberg, *Tetrahedron* **1991,** *47,* 6689.

[23] A.M. Michelson, 'The Chemistry of Nucleosides and Nucleotides', Academic Press, London–New York, 1963, pp. 465.