A General Scheme Based on Empirical Increments for the Prediction of Hydrogen-Bond Associations of Nucleobases and of Synthetic Host–Guest Complexes

Joachim Sartorius and Hans-Jörg Schneider*

Abstract: Association energies ΔG_i in chloroform, in part also in carbon tetrachloride, were determined by NMR titrations of suitably substituted nucleosides and several synthetic analogues. Based on these and on many literature data, two simple free energy increments were derived describing the ΔG_i values of 58 complexes within 1.8 kJ mol^{-1}. With chloroform as solvent the increment for the primary interaction between donor and acceptor is 7.9 kJ mol^{-1}, for the secondary one 2.9 kJ mol^{-1}, irrespective of whether the latter is attractive or repulsive. Addition of only 1% methanol to CCl_4 led to a decrease in association constants by a factor of 25. Calorimetric titrations of G–C nucleoside derivatives in CCl_4 showed substantial contributions from G dimers, in line with NMR titrations, and surprisingly small decreases in entropy. Preliminary NOE measurements allowed us to single out some of the possible association modes; they are also in line with expected self- and triple-association modes of the nucleobases. These modes are generally in accord with nucleobase associations predicted by MM calculations in the literature, which in turn agree with predictions based solely on the increments derived in the present work.

Keywords
hydrogen bonds · LFER · molecular recognition · NMR spectroscopy · nucleobases · thermodynamics

Introduction

Many natural and synthetic supramolecular complexes owe their organization to the presence of neighbouring hydrogen-bond donor and acceptor functionalities at complementary sites.[1] The central role of base pairing in nucleic acids[2] has led to countless theoretical[3] and, to a lesser degree, experimental studies[4] of the association energies involved. It is known that exchange of nucleobases can drastically alter double-helical structures.[5] The ability to understand and predict stability and preferred conformations of such complexes is also important for the rational development of exciting new chemical technologies, such as antisense[6] and triple-helix[7] methods, modification of nucleic acids by new bases,[8] design of antibiotics and peptide nucleic acids,[9] and construction of new materials such as hydrogen-bonded networks leading to thin films,[10] tubular structures[11] or liquid crystals.[12] Zimmerman,[13] Rebek,[14] Hamilton,[15] Bell,[16] Anslyn,[17] and others have demonstrated elegant applications of synthetic host compounds with suitably arranged hydrogen-bond functionalities for selective complexation of a large range of substrates. We wanted to examine experimental association constants and modes, particularly for nucleobase derivatives, and to explore the possibility of predicting stabilities and conformations of such complexes from straightforward, empirically derived increments in the free complexation energy ΔG_{HB}. The extraction of increments for hydrogen bonds from sufficiently large experimental data sets has been quite successful with simple associations held together by single interactions.[18] However, this is the first serious attempt to analyse the association of nucleobase and related systems in terms of linear free energy relationships.

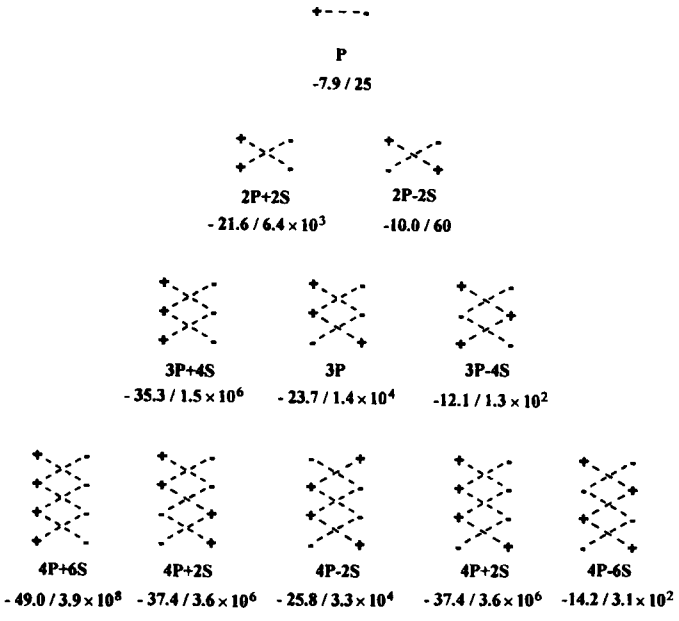
[*] Prof. Dr. H.-J. Schneider, Dr. J. Sartorius
FR Organische Chemie der Universität des Saarlandes
D-66041 Saarbrücken (Germany)
Fax: Int. code +(681)302-4105
e-mail: ch12hs@sbusol.rz.uni-sb.de

Results and Discussion

Synthetic host–guest complexes and nucleobase analogues: Our earlier analyses[19] of host–guest complexes containing amide-type fragments led to a value of $\Delta G_{HB} = 5 \pm 1$ kJ mol^{-1} for each individual hydrogen bond, that is, a single increment was assigned to each of these bonds, independent of the neighbourhood of other donors or acceptors. Jorgensen et al. showed in a series of seminal papers,[3a, b] based on MM calculations on base pairs, that secondary interactions **S** between neighbouring heteroatoms can provide additional stabilization provided that hydrogen-bond donor functions **D** with their positive partial charges are present in one molecule and acceptor functions **A** in the other. Repulsion between diagonally placed functions **A** and **D** will destabilize the primary hydrogen bond **P** if both molecules contain alternating sequences (e.g., **AD**, **DA** etc.). Scheme 1 illustrates the possible arrangements with the predicted stabilities in complexes containing an increasing number of hydrogen bonds.

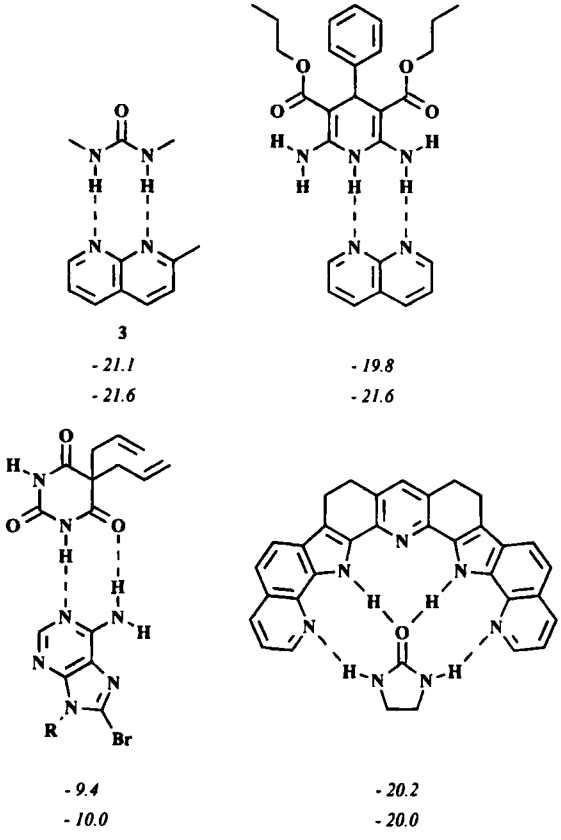
Several host complexes synthesized by Zimmerman,[13] Bell,[16] Anslyn[17] and others have shown that particularly high association constants are indeed attainable with nonalternating **DDD/AAA** sequences. We set out to analyze these and new data on the basis of as few ΔG_{HB} increments as possible. Since reports

0947-6539/96/0211-1446 $ 15.00+.25/0 Chem. Eur. J. 1996, 2, No. 11



Scheme 1. Possible arrangements in hydrogen bonded associations with predicted values for ΔG_t (kJ mol^{-1}) and K (M^{-1}) in chloroform. P: primary interaction ($\Delta G_p = 7.9$ kJ mol^{-1}); S: secondary interaction, repulsive or attractive ($\Delta G_s = 2.9$ kJ mol^{-1}).

of simple **AA/DD** sequences are scarce, we synthesized the 2-methyl-1,8-naphthyridine (**1**) as an **AA** acceptor by literature procedures.[20] Complexation constants K of **1** with N,N'-dimethylurea (**2**) in $CDCl_3$ to give **3** were measured by NMR titrations following established protocols.[21] The corresponding total free energy ΔG_t is shown in Scheme 2 together with selected available literature data on related associations.



Scheme 2. Overview of representative synthetic host–guest complexes used for the analyses, with experimental (top numbers) and calculated (bottom numbers) ΔG_t values in chloroform.

Analysis of 64 association constants for 58 different complexes led to the surprising result that only *one* increment for the primary interaction and *one* for the secondary is needed in order to reproduce the correct ΔG_t values. The difference between the calculated and the experimental total free complexation energies is only ± 1.8 kJ mol^{-1} on average (see Scheme 2). A plot of observed versus calculated ΔG_t (Fig. 1) shows a linear correlation (coefficient $r = 0.913$) with a slope of $m = 0.84$. The value for the primary interaction ($\Delta G_p = 7.9$ kJ mol^{-1}) is larger than that derived earlier ($\Delta G = 5$ kJ mol^{-1}).[19] The earlier analysis neglected the repulsive secondary interactions, which were more frequent in the smaller data set used there. The fact that the increments for the secondary interactions are $\Delta G_s = 2.9$ kJ mol^{-1}, irrespective of whether they are attractive or repulsive, is in line with their purely electrostatic nature; the Coulomb energies is then constant, as long as the geometry is similar within the limits of these "soft" potentials. The underlying geometries are indeed expected to stay within these limits as they are essentially determined by at least two stronger primary interactions in all complexes. We note that *absolute* values of these free energies in solution, which include—obviously close to constant—solvation contributions, were not obtained in the calculations of Jorgensen et al.[3a, b] The increments are expected to be almost twice as large if, for instance, carbon tetrachloride is used as solvent instead of the hydrogen-bond donor chloroform.

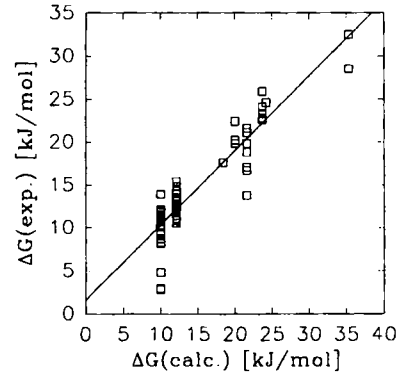


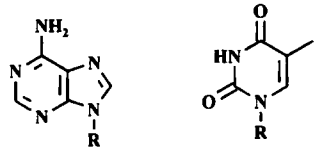
Fig. 1. Experimental vs. calculated free association energies (kJ mol^{-1}) for 58 complexes; correlation coefficient $r = 0.913$, slope $m = 0.84$.

Bell, Zimmerman et al.[22] have recently emphasized the importance of preorganization for effective binding with receptors, leading to remarkable affinities even in DMSO-containing chloroform solutions. The DMSO content makes it difficult to compare the absolute ΔG values with our increments; however, $\Delta G = 30.1$ kJ mol^{-1} in $CDCl_3$/20% DMSO[22] for the best receptor with five primary hydrogen bonds and eight secondary interactions is not very far from the predicted value of $\Delta G = 39.5$ kJ mol^{-1} in pure $CDCl_3$. More importantly, the effectiveness of two similar receptors, one with and the other without two flexible bonds,[22] differs only by the amount predicted based on the difference in the number of interactions in the two complexes. This indicates significant compensation effects are operating, namely, better matching between flexible ligand moieties occurs at the expense of higher loss of entropy on complex formation. The fact that only two increments are sufficient to predict ΔG_t for different functionalities, such as O=C–NHR or N=C–NHR (see Scheme 2), must be the result of other compensation effects. Most of the complexes contain alternating **AD/DA** sequences; for the more polar amide function there is a stronger primary contribution, which is, however,

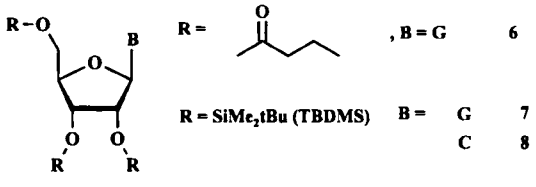
weakened by a larger secondary repulsion. Other compensation effects may be due to changes of bond orders that cancel each other out. Systematic X-ray or, preferably, neutron diffraction studies can be expected to shed light on the different hydrogen-bonding modes in such complexes, as might MO calculations with large basis sets on the different fragments. The merit of an empirical analysis such as the present one is that it provides an experimental basis for predictions in solution *independent* of mechanistic reasoning.

Natural nucleobases: Association constants with natural base pairs have previously been measured in chloroform[4] or other hydrophobic media,[2, 23] with lipophilic derivatives containing mostly aliphatic or cyclic alkyl chains at the heterocycle. As noted by Jorgensen et al.,[3a, b] some of these data are probably not very accurate in view of the methods available at the time, which also often neglected the occurrence of complexes other than simple 1:1 adducts of the Watson–Crick type. We have used high-field NMR titrations with modern accumulation techniques and nonlinear curve-fitting methods, which overcome some of the earlier restrictions. Furthermore, some equilibria were also checked by calorimetric titrations, in collaboration with the group of Professor Blandamer in Leicester (UK) and of Professor Raevsky and Dr. Solov'ev in Chernogolovka (Russia).

As chloroform-soluble substrates we used alkyl-substituted nucleobases like **4** and **5** and, for the first time, ribonucleosides **6**–**8**, which are lipophilic owing to the *tert*-butyldimethylsilyl (TBDMS) or butanoyl substituents at all sugar OH positions (Scheme 3). The syntheses of these compounds, which were also



4 R = $-(CH_2)_5-CH_3$ **5**



R = , B = G **6**

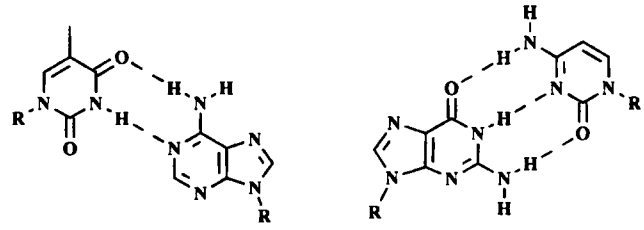
R = SiMe$_2$tBu (TBDMS) B = G **7**

C **8**

Scheme 3. Structures of the derivatives of natural nucleobases considered in this study.

found to be remarkably stable towards traces of water, are described in the Experimental Section. Such ribonucleoside derivatives have, to the best of our knowledge, not previously been used, although they have the advantage 1) being closer to the natural systems, 2) preventing some of the many different association modes,[1, 3] and 3) providing additional signals for NMR analyses. Additional *structural* studies by NOE methods were found to be necessary for the nucleobases, beyond those performed on the synthetic systems dealt with above, as only the latter are designed to give only one complex conformation in solution.

NMR titrations[24] of A–T base-pair derivatives **4** and **5** in $CDCl_3$ yielded K values that are close to literature values,[2] although, in view of the perfect curve fit to a 1:1 model, self-association of the single nucleobases (see below) was neglected; the ΔG is essentially in line with the prediction for an **AD–DA** pair (Scheme 4). Titrations with the G and C derivatives **7** and **8** were carried out with various concentration ranges and solvents (Table 1). As usual the NH signals of both bases were



AT4 **GC1**

-8.7 **-24.0**

-10.0 *-23.7*

Scheme 4. Observed (bold) and calculated (italics) ΔG_i values for the Watson–Crick base pairs A–T and G–C.

affected most by association; they could, however, be followed only within concentration ranges too small for a nonlinear curve fit. The reason for this is a substantial broadening for both the C and G NH_2 signals, owing to association and exchange with residual water as well as to the known restriction of free rotation around the C–NH_2 bond upon complexation.[25] The only signal that could be observed over the desirable concentration ranges of between about 20 and 80% complexation was H-8 of the G derivative (Fig. 2). The ΔG_i derived from curve fitting to a 1:1 association model is again similar to literature values[4a] with other derivatives, and is again close to the predicted value for a **DDA–AAD** pair (Scheme 4). The limiting CIS (complexation induced shift) values for the NH protons (Table 1), calculated from the equilibrium constant K, which are derived from the titration with the G H-8 signal, show deshielding of up to +1.6 ppm at the end of the titration, which is typical for the involvement of such protons in hydrogen bonding. Measure-

Table 1. NMR titrations [a] with the G–C derivatives **7** and **8** and the A–T derivatives **4** and **5**: equilibrium constants K [M^{-1}], ΔG_i [kJ mol^{-1}] and CIS values [ppm].

Complex	Solvent	Proton	K (σ) [k]	CIS	ΔG
A_2	$CDCl_3$	NH_2	2.4 (0.2)	(+1.0) [j]	−2.2
T_2	$CDCl_3$	NH	3.5 (1.2)	(+1.4) [j]	−3.1
GC [b,e]	$CDCl_3$	NH	– [d]	+1.61	–
		H8	4.7×10^4 [f] (1.6×10^4)	−0.25	−26.7
		H8	1.6×10^4 [g]	−0.13	−24.0
GC [c,e]	$CDCl_3$	NH	– [d]	+1.4	–
		H8	3.5×10^4 [f] (8.9×10^3)	−0.11	−25.9
		H8	1.6×10^4 [g]	−0.23	−24.0
C_2	$CDCl_3$	H6	40 [f] (12)	−0.15	−9.1
G_2	$CDCl_3$	NH_2	3×10^2 [f] (130)	+1.5	−14.1
		NH	– [h]	<0.06	–
		H8	2.9×10^2 [f] (120)	−0.04	−14.0
GC	CCl_4	H8	5×10^4 [i]	−0.21	−26.9
GC	CCl_4 +1 vol% MeOH	H8	2×10^3 [i]	−0.14	−18.8

[a] TMS in $CDCl_3$ as external reference; $T = 298\pm0.1$ K; water content 5 to 7.5 mM in $CDCL_3$, 5 mM in CCl_4 (against internal standard CH_3CN). [b] $[G]_0 = 9.5\times10^{-5}$ M, $[C]_0 = 9.4\times10^{-6}$ M, $[G]_{end} = 3\times10^{-5}$ M, $[C]_{end} = 2.12\times10^{-4}$ M. [c] $[G]_0 = 1$ mM, $[G]_{end} = 0.56$ mM, $[C]_{end} = 4.4$ mM. [d] Only observed with excess G; CIS value determined with 1.2×10^4 scans. [e] G NH_2 broadened, only observed with excess G. [f] Fitting for 1:1 model. [g] Fitting for 3 coupled equilibria (G_2, C_2, GC), no standard deviation given by the used program. [h] Fitting not convergent. [i] Estimated for ≥90% complexation. [j] Maximum of measured shift change; fitted values out of reasonable range. [k] σ: standard deviation.

0947-6539/96/0211-1448 $ 15.00+.25/0 Chem. Eur. J. 1996, 2, No. 11

ments in CCl_4 showed an increase in association (Table 1) that is smaller than the one observed earlier with related systems,[19] perhaps because of the presence of water traces. The extreme effect of competiting protic solvents is obvious from one titration in which the presence of only 1 vol% of methanol leads to a decrease of the equilibrium constant by an order of magnitude (Table 1).

Calorimetric titrations with nucleobase derivatives in lipophilic solvents like carbon tetrachloride have until now rarely been carried out. They suffer from the small heat capacity of solvent and other technical problems requiring, for example, long measuring and equilibration times. Microcalorimetric titrations with **7** and **8** under the conditions described in the Experimental Section yielded a ΔG_t value of $27 \pm 1.0\ kJ\,mol^{-1}$, which compares well with the NMR-derived value of $26.9\ kJ\,mol^{-1}$ (Table 1). The free enthalpy value $\Delta H = 26 \pm 1.0\ kJ\,mol^{-1}$ was almost the same as ΔG (similar results were obtained in independent titrations with different equipment by Dr. O. Solov'ev, with $\Delta G = 26.5 \pm 0.5\ kJ\,mol^{-1}$ and $\Delta H = 30 \pm 1.0\ kJ\,mol^{-1}$, evaluated on the basis of models taking into account the formation of G and C dimers, as discussed below). The fact that the corresponding entropy contributions are below $T\Delta S = 1\ kJ\,K^{-1}\,mol^{-1}$ is surprising for complex formation in view of the absence of substantial solvation effects of the inert carbon tetrachloride. A possible explanation is the occurrence of self-aggregates before base-pair formation, as well as of many additional aggregates between the nucleobases (see below).

Self-associations and higher aggregates: The fit of the H8 NMR shifts (Fig. 2) already shows a systematic deviation, which is the result of neglecting self-associations of the monomers. Using the

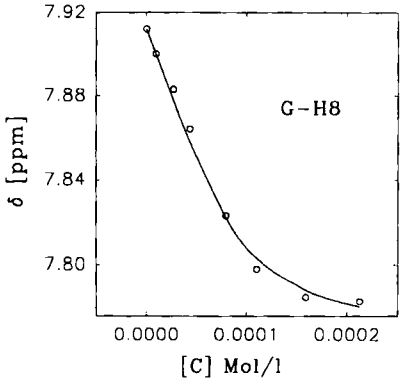


Fig. 2. NMR titration with least-squares fit for the G H-8 signal of the complex GC (GC derivatives see Scheme 3); in $CDCl_3$, 298 K.

program CHEMEQUI[26] a nonlinear least-squares fit based on a model involving the dimers GG and CC, in addition to the G–C base pair, yielded an equilibrium constant for the G–C association (Table 1) which, although somewhat lower than the one obtained based on the 1:1 model, is in excellent agreement with the prediction for the **DDA–AAD** combination. Self-association studied by dilution shifts yielded in the case of **8** (CC), after a suitable curve fit (Fig. 3), values close to the one obtained form the G+C titration. Dilution experiments with **4** (A) and **5** (T) were less conclusive as the simple 1:1 model used did not sufficiently describe the complex aggregation behaviour under the necessary experimental conditions (Fig. 4, Table 1). Estimates of K, however, based on the measured shift changes and simulated nomograms for self-association, were again in line

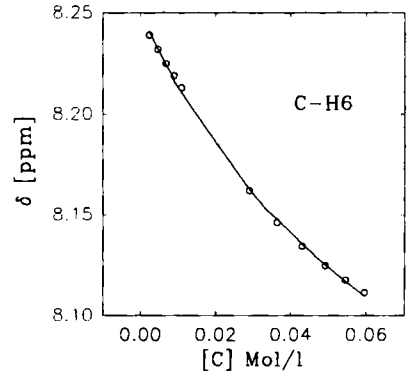


Fig. 3. NMR titration for the self-association of C (CC) followed by means of the C H-6 signal; other conditions see Fig. 2.

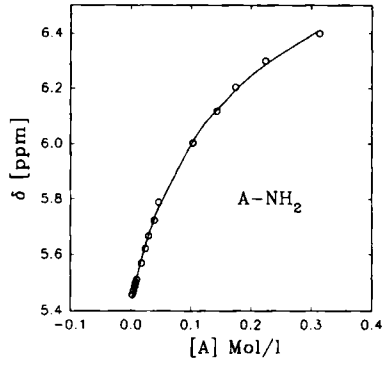


Fig. 4. NMR titration for the self-association of A (AA) followed by means of the A NH_2 signal; other conditions see Fig. 2.

with the predicted values. The available ΔG values obtained for the CC dimer (9.0 kJ) are also in accord with the predicted affinities (10 kJ) for an **AD–DA** combination. The value for the dimer of **7** (GG) and the measured shift change (especially for the NH proton, $< +0.06$ ppm) does not agree with the structure for the postulated most stable dimer with an **AA–DD** pattern (Fig. 5).

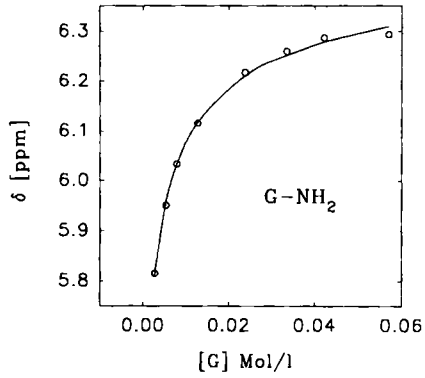
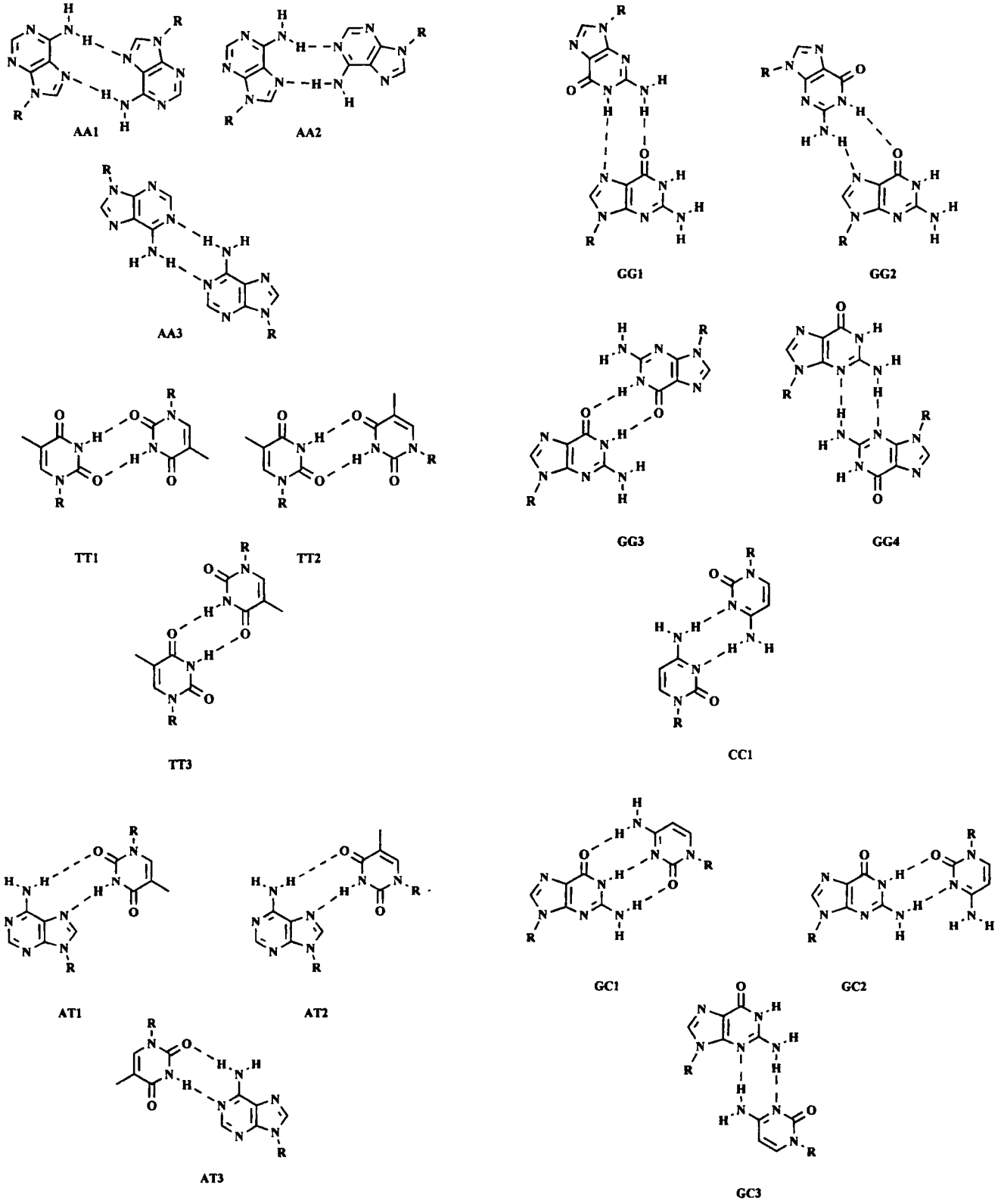


Fig. 5. NMR titration for the self-association of G (GG) followed by means of the G NH_2 signal; other conditions see Fig. 2.

The question of which structures correspond to the G_2 and C_2 dimers, and to higher aggregates must be discussed against the background of all possible nucleobase di- and oligomer structures. The expected aggregates with the predicted ΔG_t values are listed in Scheme 5. It should be noted that they are essentially the same as obtained from MM calculations.[3c–i] Experi-



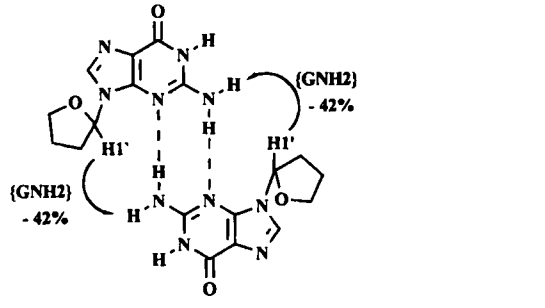
Scheme 5. Important association modes [1,3] of natural nucleobase derivatives.

mental evidence for the different association modes was until now largely restricted to X-ray studies[1] in the solid state, where extended hydrogen-bond networks containing many single bases are of course much more likely than in solution, where they are disfavoured for statistical reasons. For this reason and in view of some obvious deviations between experimental and predicted ΔG_t for the natural nucleobase derivatives, some preliminary NOE measurements were carried out in the present study. Large relaxation times $T1$ and saturation transfer effects, also to the solvent, led us to apply only steady-state NOEDIFF techniques,[27] which do, however, already allow some of the association modes to be singled out. Most of the observed intermolcular NOEs were either very small or negative as consequence of the molecular weight of the aggregates and

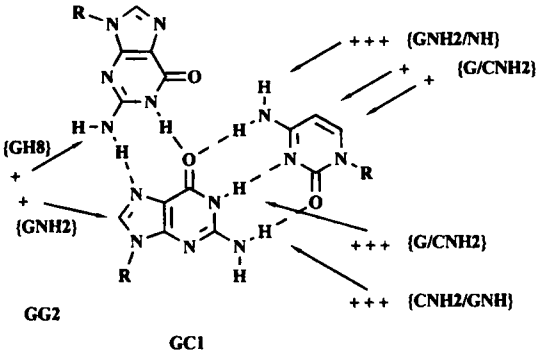
0947-6539/96/0211-1450 $ 15.00+.25/0 Chem. Eur. J. 1996, 2, No. 11

their slow tumbling time. It is possible that the size of the NOE does not reflect the populations of the different possible aggregates; in particular, those with short intermolecular distances may be overemphasized.

The NOE spectrum of the GC base pair showed clear evidence not only for the expected strong Watson–Crick base pairing, but also for the presence of a trimer. This is the consequence of Hoogsteen pairing with another G (Scheme 6), and is visible



GG4



GG2 GC1

Scheme 6. Examples of NOE in the aggregates GG4 and the triplex G·GC involving the Watson–Crick GC base pair.

in the effects of irradiating G NH_2 on the G H8 proton. The observed NOEs for the GG dimer (Scheme 6) unexpectedly gave no evidence for a predominating GG1 structure (Scheme 5), but rather for GG4. Thus, irradiation at G NH_2 leads to a strong negative effect on H1'. In line with this, the NH protons in **7**, which would be part of the hydrogen-bond system in GG1, were barely shifted (< + 0.06 ppm). For A and T associations the observed NOEs were all consistent with the expected[1,3] Watson–Crick (AT1, AT2) and/or Hoogsteen (AT3, AT4) pairs. Additional intermolecular NOEs indicate formation of triplexes,[7] such as A·AT or T·AT, depending on the corresponding concentrations.

Conclusions

The synthetic host compounds have been designed to give unique complex structures with high and selective binding. Complexation mode and strength can be predicted in a straightforward manner by using only two increments. The nucleobases selected by nature allow many more binding modes including self-association. Even so, the same increments derived from the synthetic complexes also describe Watson–Crick base pairs with high accuracy, and, to a lesser degree, the Hoogsteen-type of self-associations and of higher aggregates that are predicted to be energetically similar[1,3] in lipophilic media. Interaction modes of natural nucleobases are unambiguous only in the double helix, as a consequence of the chelate effect with the many simultaneously interacting bases and stacking forces. This also protects the hydrogen bonds from competition by water, which—as the experiment with only 1% methanol in CCl_4 shows—would totally disrupt recognition. Groove binding of the Hoogsteen-type is therefore much weaker, but the results indicate that our empirical increments after suitable factorization may apply here too. The preference of triplex formation particularly with G·GC components in purine-rich parts of nucleic acids[7] is also predicted from the secondary interaction terms; the same holds for the melting point increase found on replacing natural nucleobases by synthetic analogues.[8]

Experimental Section

General: NMR titrations were performed with Bruker DRX 500 and AM 400 systems. Solvents were $CDCl_3$ (99% D content) or CCl_4 at 298 ± 0.1 K; external reference: TMS in $CDCl_3$. Least-squares fit of the titration curves were performed for the 1:1 complexes with suitable equations within the program Sigmaplot 5.0 (Jandel Scientific) using Levenberg–Marquardt algorithms, for other complexes with the program Chem-Equi (Dr. V. Solovev, IPAC) using Simplex or MonteCarlo routines.

*T*1 measurements were carried out with degassed samples (freeze–thaw cycle with nitrogen) by using published inversion–recovery routines (Bruker programs). NOE measurements were performed with the steady-state NOEDIF protocol on degassed samples. Integrals, where measurable, were corrected following literature recommendations [27].

Calorimetric measurements were carried out (in Leicester, laboratory of Prof. M. Blandamer) with a microcalorimeter system and the program Origin 3.18 from Microcal; the solutions (CCl_4) were degassed by operation under vacuum.

Materials: Compounds were either commercially available or prepared according to literature procedures, adapted as described below: **1** [20], **5** [28], **4** [29], **6** [30], **7** [31], **8** [31]. Preparative TLC ($CHCl_3$/methanol 9:1; Macherey–Nagel SIL-G-25 $UV_{254+366}$, R_f = 0.54) was used to purify **7**. The workup procedure for **10** was modified according to Köhler [32].

2-Methyl-1,8-naphthyridine (1): ^{1}H NMR ($CDCl_3$, TMS): δ = 9.06–9.04 (m, 1H, H7), 7.42 (dd, 1H, H6), 8.13 (dd, 1H, H4), 7.37 (d, J = 8.3 Hz, H3), 8.06 (d, J = 8.3 Hz, 1H, H5), 2.80 (s, 3H, CH_3).

9-Hexyladenine (4): ^{1}H NMR ([D_6]DMSO, TMS): δ = 8.13 (s, 1H, H8), 1.83–1.77 (m, 2H, β-CH_2), 8.12 (s, 1H, H2), 1.35–1.18 (m, 6H), 7.12 (brs, 2H, NH_2), 0.83 (t, J = 6.8 Hz, 3H, CH_3), 4.12 (t, J = 7.1 Hz, 2H, α-CH_2); $C_{11}H_{17}N_5$ (219.3): calcd C 60.25, H 7.81, N 31.94; found C 60.32, H 7.83, N 31.57.

1-Hexylthymine (5): ^{1}H NMR ([D_6]DMSO, TMS): δ = 11.16 (brs, 1H, NH), 1.58–1.53 (m, 2H, β-CH_2), 7.52 (s, 1H, H6), 1.30–1.15 (m, 6H), 3.60 (t, J = 7.2 Hz, 2H, α-CH_2), 0.85 (t, J = 6.6 Hz, 3H, CH_3), 1.75 (s, 3H, CH_3); $C_{11}H_{18}N_2O_2$ (210.3): calcd C 62.83, H 8.63, N 13.32, found C 63.48, H 8.58, N 12.57.

2′,3′,5′-Tri-*O*-Butanoylguanosine 6: ^{1}H NMR ([D_6]DMSO, TMS): δ = 5.52–5.34 (m, 1H, H2′), 10.73 (brs, 1H, NH), 4.40–4.26 (m, 3H, H4′, H5′/5″), 7.91 (s, 1H, H8), 2.37–2.27 (m, 6H, 3-CH_2-CO-), 6.52 (brs, 2H, NH_2), 1.58–1.48 (m, 6H, 3-CH_2-), 5.97 (d, J = 6.0 Hz, 1H, H1′), 0.92–0.80 (m, 9H, 3-CH_3), 0.81 (t, J = 6.0 Hz, H3′); $C_{22}H_{28}N_5O_8$ (490.5): calcd C 53.87, H 5.75, N 14.28, found C 56.18, H 6.89, N 12.94.

2′,3′,5′-Tri-*O*-Silylguanosine 7: ^{1}H NMR ($CDCl_3$, TMS): δ = 12.0 (brs, 1H, NH), 4.5–3.83 (m, 5H, H2′, H3′, H4′, H5′/5″), 7.96 (s, 1H, H8), 1.02, 0.99, 0.92 (s, 27H, 9CH_3 of *t*Bu), 6.23 (brs, 2H, NH_2), 0.23, 0.22, 0.20, 0.19 (s, 12H, 4CH_3), 5.89 (d, J = 4.2 Hz, 1H, H1′), 0.06 (s, 6H, 2CH_3); ^{13}C NMR ($CDCl_3$, TMS): C=O not observed, δ = 70.9 (C2′), 158.7, 152.8, 151.1 (quart guanine-C), 61.7 (C5′), 135.7 (C8), 29.2, 18.0, 17.6 (3quart *t*Bu), 87.6 (C1′), 25.5, 25.3, 25.2 (9CH_3 of *t*Bu), 84.2 (C4′), −3.4, −4.0, −4.2, −4.8, −5.1, −5.5 (6CH_3), 75.6 (C3′); $C_{25}H_{55}N_5O_5Si_3$ (590.0): calcd C 50.89, H 9.40, N 11.87, found C 53.21, H 8.52, N 11.17.

2′,3′,5′-Tri-*O*-Silylcytidine 8: ^{1}H NMR ($CDCl_3$, TMS): δ = 8.13 (d, J = 9.3 Hz, 1H, H6), 0.93, 0.86, 0.84 (s, 27H, 9CH_3 of *t*Bu), 7.10 (brs, 2H, NH_2), 0.18, 0.11, 0.09, 0.07 (s, 12H, 4CH_3), 5.77 (d, J = 1.8 Hz, 1H, H1′), 0.033, 0.027 (s, 6H, 2CH_3), 5.64 (d, J = 9.3 Hz, 1H, H5), 4.08–4.01, 3.77–3.73 (m, 5H, H2′ to H5′/5″); ^{13}C NMR ($CDCl_3$, TMS): δ = 165.2 (C6), 83.1 (C4′), 25.6, 18.5, 18.0 (3quart *t*Bu C), 155.6 (C2), 76.3 (C3′), −3.6, −4.1, −4.2, −5.0, −5.3, −5.6 (6CH_3), 141.7 (C4), 69.6 (C2′), 93.8 (C5), 61.1 (C5′), 90.1 (C1′), 26.1, 25.9 (9CH_3 of *t*Bu); $C_{24}H_{55}N_3O_5Si_3$ (550.0): calcd C 52.41, H 10.08, N 7.64, found C 54.40, H 9.15, N 7.54.

Acknowledgement. This work was financially supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. We thank Professor Blandamer (Leicester), Professor Raevsky and Dr. Solov'ev (Chernogolovka near Moscow) for calorimetric measurements, and acknowledge support by NATO CRG and DAAD travel grants.

Received: March 25, 1996 [F329]

[1] G. A. Jeffrey, W. Saenger *Hydrogen Bonding in Biological Structures*, Springer, Heidelberg, **1994**.

[2] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, Heidelberg, **1984** and references therein.

[3] a) W. L. Jorgensen, J. Pranata, *J. Am. Chem. Soc.* **1990**, *112*, 2008; b) J. Pranata, S. G. Wierschke, W. L. Jorgensen, *ibid.* **1991**, *113*, 2810; c) K. I. Trollope, I. R. Gould, I. H. Hillier, *Chem. Phys. Lett.* **1993**, *209*, 113; d) P. E. Young, I. H. Hillier, I. R. Gould, *J. Chem. Trans.* **1994**, *2*, 1717; e) A. R. Leach, P. A. Kollman, *J. Am. Chem. Soc.* **1992**, *114*, 3675; f) P. Cieplak, P. A. Kollman, *ibid.* **1988**, *110*, 3734; g) I. R. Gould, P. A. Kollman, *ibid.* **1994**, *116*, 2493; h) A. Pohorille, S. K. Burt, R. D. MacElroy, *ibid.* **1984**, *106*, 402; i) L. X. Dang, P. A. Kollman, *ibid.* **1990**, *112*, 503.

[4] a) Y. Kyogoku, R. C. Lord, A. Rich, *Biochim. Biophys. Acta* **1969**, 179; b) L. D. Williams, B. Chawla, B. R. Shaw, *Biopolymers* **1987**, *26*, 591; c) A. Amantea, M. Henz, P. Strazewski, *Helv. Chim. Acta* **1996**, *79*, 244; d) A. Amantea, M. Weber, K. Sequin, P. Strazewski, *ibid.* **1995**, *78*, 1106; e) P. Strazewski, *ibid.* **1995**, *78*, 1112.

[5] J. L. Mergny, L. Lacroix, X. Han, J. L. Leroy, C. Helene, *J. Am. Chem. Soc.* **1995**, *117*, 8887.

[6] A. Demesmaeker, R. Haner, P. Martin, H. E. Moser, *Acc. Chem. Res.* **1995**, *28*, 366.

[7] a) Y. K. Cheng, B. M. Pettitt, *Prog. Biophys. Mol. Biol.* **1992**, *58*, 225; b) N. T. Thuong, C. Hélène, *Angew. Chem.* **1993**, *105*, 697; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666. c) W. D. Wilson, H. P. Hopkins, S. Mizan, D. D. Hamilton, *J. Am. Chem. Soc.* **1994**, *116*, 3607; d) W. A. Greenberg, P. B. Dervan, *ibid.* **1995**, *117*, 5016.

[8] a) J. J. Voegel, S. A. Benner, *J. Am. Chem. Soc.* **1994**, *116*, 6929; b) C. Y. Switzer, S. E. Moroney, S. A. Benner, *Biochemistry* **1993**, *32*, 10489; c) J. A. Piccirilli, S. E. Moroney, S. A. Benner, *ibid.* **1991**, *30*, 10350; d) J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature (London)* **1990**, *343*, 33; e) C. Y. Switzer, S. E. Moroney, S. A. Benner, *J. Am. Chem. Soc.* **1989**, *111*, 8322; f) B. A. Schweitzer, E. T. Kool, *ibid.* **1995**, *117*, 1863; g) T. L. Trapane, P. O. P. Tso, *ibid.* **1994**, *116*, 10437.

[9] a) C. Bohler, P. E. Nielsen, L. E. Orgel, *Nature* **1995**, *376*, 578; b) M. Leijon, A. Graslund, P. E. Nielsen, O. Buchardt, *Biochemistry* **1994**, *33*, 9820; c) B. Hyrup, M. Egholm, M. Rolland, P. E. Nielsen, R. H. Berg, *J. Chem. Soc. Chem. Comm.* **1993**, *6*, 518; d) S. K. Kim, P. E. Nielsen, M. Egholm, O. Buchardt, R. H. Berg, *J. Am. Chem. Soc.* **1993**, *115*, 6477; e) M. Egholm, P. E. Nielsen, O. Buchardt, R. H. Berg, *ibid.* **1992**, *114*, 9677.

[10] J. C. MacDonald, G. M. Whitesides *Chem. Rev.* **1994**, *94*, 2383.

[11] C. Y. Huang, V. Lynch, E. V. Anslyn, *Angew. Chem.* **1992**, *104*, 1259; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1244.

[12] a) C. M. Paleos, D. Tsiourvas, *Angew. Chem.* **1995**, *107*, 1839; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1696; b) T. Kato, H. Kihara, U. Kumar, T. Uryu, J. M. J. Fréchet, *ibid.* **1994**, *106*, 1728 and **1994**, *33*, 1644.

[13] a) T. J. Murray, S. C. Zimmerman, S. V. Kolotuchin, *Tetrahedron* **1995**, *51*, 635; b) S. C. Zimmerman, T. J. Murray, *Tetrahedron Lett.* **1994**, *35*, 4077; c) S. C. Zimmerman, T. J. Murray in *Computational Approaches in Supramolecular Chemistry* (Ed.: G. Wipff), NATO ASI Series C, Vol. 426, Kluwer, **1994**; d) S. C. Zimmerman, T. J. Murray, *Phil. Trans. R. Soc. Lond. A* **1993**, *345*, 49; e) S. C. Zimmerman, W. Wu, Z. Zeng, *J. Am. Chem. Soc.* **1991**, *113*, 196.

[14] R. S. Meissner, J. Rebek, Jr., J. de Mendoza, *Science* **1995**, *270*, 1485.

[15] a) A. D. Hamilton, D. J. Van Engen, *J. Am. Chem. Soc.* **1987**, *109*, 5035; b) S. K. Chang, A. D. Hamilton, *ibid.* **1988**, *110*, 1318.

[16] D. A. Bell, E. V. Anslyn, *Tetrahedron* **1995**, *51*, 7161.

[17] C.-Y. Huang, L. A. Cabell, E. V. Anslyn, *J. Am. Chem. Soc.* **1994**, *116*, 2778.

[18] a) O. A. Raevsky, *Russ. Chem. Rev.* **1990**, *59*, 219; b) M. H. Abraham, *Chem. Soc. Rev.* **1993**, *22*, 73 and references therein.

[19] H.-J. Schneider, R. K. Juneja, S. Simova, *Chem. Ber.* **1989**, *112*, 1211.

[20] E. M. Hawes, D. G. Wibberley, *J. Chem. Soc. (C)* **1966**, 315.

[21] C. S. Wilcox, in *Frontiers on Supramolecular Organic Chemistry and Photochemistry* (Eds.: H.-J. Schneider, H. Dürr), VCH, Weinheim, **1991**; H.-J. Schneider, R. Kramer, S. Simova, U. Schneider, *J. Am. Chem. Soc.* **1988**, *110*, 6442.

[22] T. W. Bell, Z. Hou, S. C. Zimmerman, P. A. Thiessen, *Angew. Chem.* **1995**, *107*, 2321; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2163.

[23] R. T. West, L. A. Garza, W. R. Winchester, J. A. Walmsley, *Nucleic Acids Res.* **1994**, *22*, 5128.

[24] H.-D. Junker, Diploma thesis, Universität des Saarlandes, **1994**.

[25] a) N. G. Williams, L. D. Williams, B. R. Shaw, *J. Am. Chem. Soc.* **1990**, *111*, 7205; b) H. Iwahashi, H. Sugeta, Y. Kyogoku, *Biochemistry* **1982**, *21*, 631.

[26] V. Solov'ev, unpublished results.

[27] D. Neuhaus, M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, Weinheim, **1989**.

[28] D. T. Browne, J. Eisinger, N. J. Leonard, *J. Am. Chem. Soc.* **1968**, *90*, 7302 and references therein.

[29] N. J. Leonard, K. L. Carraway, J. P. Helgeson, *J. Heterocycl. Chem.* **1965**, *2*, 296.

[30] A. Matsuda, M. Shinozaki, M. Suzuki, K. Watanabe, T. Miyasaki, *Synthesis* **1986**, 385.

[31] a) K. K. Ogilvie, S. L. Beaucage, A. L. Schifman, N. Y. Theriault, K. L. Sadana, *Can. J. Chem.* **1978**, *56*, 2768; b) K. K. Ogilvie, A. L. Schifman, C. L. Penney, *ibid.* **1979**, *57*, 2230; c) G. H. Hakimelahi, Z. A. Proba, K. K. Ogilvie, *ibid.* **1982**, *60*, 1106.

[32] W. Köhler, thesis, Konstanz, **1978**.