On RNA Triplet Interactions: NMR Study of the Short Intramolecular Duplex Formed by r[GCAm^{1}G-p-O(CH_2CH_2O)$_6$-p-UGCC]

Preliminary Communication

by **Marc-Olivier Ebert**[1]) and **Bernhard Jaun***

Laboratorium für Organische Chemie, Eidgenössische Technische Hochschule Zürich, Universitätstrasse 16, CH-8092 Zürich

and **Werner Pils**[1]) and **Ronald Micura**

Institut für Chemie, Johannes Kepler Universität Linz, Altenbergerstrasse 69, A-4040 Linz

Dedicated to Prof. *Albert Eschenmoser* on the occasion of his 75th birthday

The flexible hexaethylene-glycol linker enhances the stability of the duplex between the two tetranucleotides in compound **1** sufficiently to allow determination of the solution structure by NMR. At 4.8°, two of the three possible imino NH protons are detected as sharp signals and establish the presence of two G·C *Watson-Crick* base pairs. Through assignment of all but one of the non-labile protons and measurement of ^{1}H,^{1}H and ^{1}H,^{31}P coupling constants, as well as NOEs of labile and non-labile protons, it was possible for the first time to derive detailed structural information on such a short duplex. It forms an A-type double helix over the full length, including the dangling nucleotides. Small variations of coupling constants and a broadening of the H–C(8) signal of m^{1}G4 indicate that the two nucleotides connected to the linker are conformationally slightly distorted and/or more flexible than the unlinked end of the duplex.

1. Introduction. – The interaction between complementary RNA nucleotide triplets is of fundamental interest as complexes of this kind reflect codon-anticodon pairing within the ribosome [1]. In aqueous buffer solutions, the pairing strength of an oligoribonucleotide duplex that consists of solely three base pairs is too low to be measured with standard UV- or CD-spectroscopic methods, and the kinetics of the pairing process is so fast that the imino protons exchange too rapidly for observation by NMR spectroscopy. Since observation of the labile imino protons is a prerequisite for the direct proof of pairing by NMR, only studies based on chemical-shift variations [2] but no NMR-derived solution structures of tri- or tetranucleotide duplexes are found in the literature. Information about the conformation of tri- and tetranucleotides bound to large oligonucleotides such as tRNA has been obtained through measurement of transferred nuclear *Overhauser* effects [3].

Recently, it was shown that the unfavorable entropic contribution to the free enthalpy of pairing is substantially reduced if the two short nucleotide sequences are joined on both ends *via* flexible non-nucleotide linkers (*Fig. 1,a*). These cyclic compounds are useful model systems for studying codon-anticodon pairing. They

[1]) Part of the planned Ph.D. thesis.

enable the investigation of sequence-dependent properties of the interacting RNA triplets, including the effect of neighboring unpaired nucleobases [4]. Thereby, it was demonstrated that dangling bases at the 3′-ends of the core double helix highly enhance stability as is known for longer oligoribonucleotides [5]. In contrast to unpaired bases at the 5′-ends, dangling bases at the 3′-ends are structurally relevant with regard to codon-anticodon interactions. In the anticodon strand, they correspond to the usually modified purine base in position 37 of tRNAs and, in the codon strand, to the first (unpaired) base of the following mRNA codon.

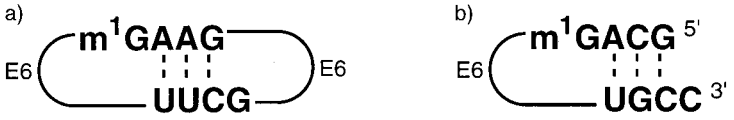

Fig. 1. *Schematic representation of tetranucleotide 'mini duplexes':* a) *linked at both ends*, b) *linked at one end by hexaethylene-glycol linkers* (sequence as in **1**)

For triplets containing weak base pairs (A · U, G · U, I · C, I · U, *etc.*), cyclic bridging is an absolute necessity in order to observe melting transitions in the accessible temperature range in buffer solutions. The above-mentioned studies of cyclic systems suggested, however, that at least for G · C-rich triplets with stabilizing dangling bases, bridging of only one double-helix end might be sufficient to obtain the stability needed for NMR studies (*Fig. 1,b*). We have, therefore, chosen the anticodon sequence motif 5′-GCAm^{1}G-3′ of a tRNA^{Cys} isolated from the archaeon *Haloferax volcanii* [6] and synthesized the corresponding model sequence r(GCAm^{1}G-p-E6-p-UGCC) (**1**; m^{1}G = 1-methylguanosine; E6 = hexaethylene glycol) [7]. The choice of the linker E6 resulted from a preceding systematic optimization of several non-nucleotide linker types as loop replacement in double-helical RNAs [7a]. The melting point of r(GCAm^{1}G-E6-UGCC) was observed at $47 \pm 0.3°$ in 142 mM NaCl, 10 mM Na_2HPO_4, pH 7.0, and is concentration-independent in the investigated range of 5.94 μM to 116 μM, as it is expected for a monomolecular transition. Here, we report the results of the NMR study of this bridged RNA 'mini duplex' and the structural information derived from these data[2]).

2. Results. – *NMR Spectroscopy at 27.3°*. In the ^{1}H-NMR spectrum of **1** in D_2O, 42 mM sodium-arsenate buffer[3]), measured at 27.3°, the signals of five H−C(1′) protons showed the small $^3J(1',2')$ values (< 1.5 Hz) typical for 3′-*endo* sugar conformations [8a,b]. Three H−C(1′) resonances (later assigned to C7, m^{1}G4, and U5, resp.) showed larger couplings of 2, 3, and 4 Hz, indicating small but significant contributions of 2′-*endo* ribose conformations for these nucleotides. The ^{31}P-NMR spectrum was well-dispersed showing eight resolved signals. Measurement of DQF-COSY (with and without ^{31}P-decoupling), TOCSY, ^{1}H-^{31}P-COSY, NOESY, HSQC, and HMBC spectra allowed assignment of all non-exchangeable protons and to determine most of the

---

[2]) A detailed structure calculation based on force-field simulations with NMR-derived distance and angle restraints will be published separately, together with the full experimental data.

[3]) The pH of the sodium-arsenate buffer solution was adjusted to a pH-meter reading of 7.0 in H_2O before lyophilization and redissolution in D_2O or H_2O/D_2O 9:1, and is not corrected for isotope effects.

^{1}H,^{1}H and ^{1}H,^{31}P coupling constants (data not shown). However, no signals of imino protons could be observed at 27.3° in H_2O/D_2O 9:1 (42 mM sodium arsenate).

NMR Spectroscopy at 4.8°. When the temperature was gradually lowered from 27.3 to 0.4°, two sharp signals of imino protons appeared in the ^{1}H-NMR spectrum recorded in H_2O/D_2O 9:1 with solvent suppression through excitation sculpting [9] (*Fig. 2*). Because the chemical shifts in both, the ^{1}H- and the ^{31}P-NMR spectrum, showed individually different variations with temperature, the series of 2D spectra in D_2O (42 mM sodium-arsenate buffer; DQF-COSY with and without ^{31}P-decoupling, TOCSY, ^{1}H,^{31}P-COSY, NOESY, and HSQC) were remeasured at 4.8° in order to securely assign the ^{1}H- and ^{31}P-resonances. Although it was possible to determine the chemical shifts of all but one of the protons of the nucleotide units (signals of the hexaethylene-glycol protons in the E6 linker overlap with the exception of the first and last CH_2 group), the chemical shifts of some ribose protons were almost identical. This made the identification of their cross-peaks in the NOESY spectra (see below) ambiguous and forced us to exclude them from the qualitative analysis of NOEs. The sequence-specific assignment of the sugar spin systems was accomplished with the ^{1}H,^{31}P-COSY spectrum, which also showed the expected two correlations of ^{31}P signals

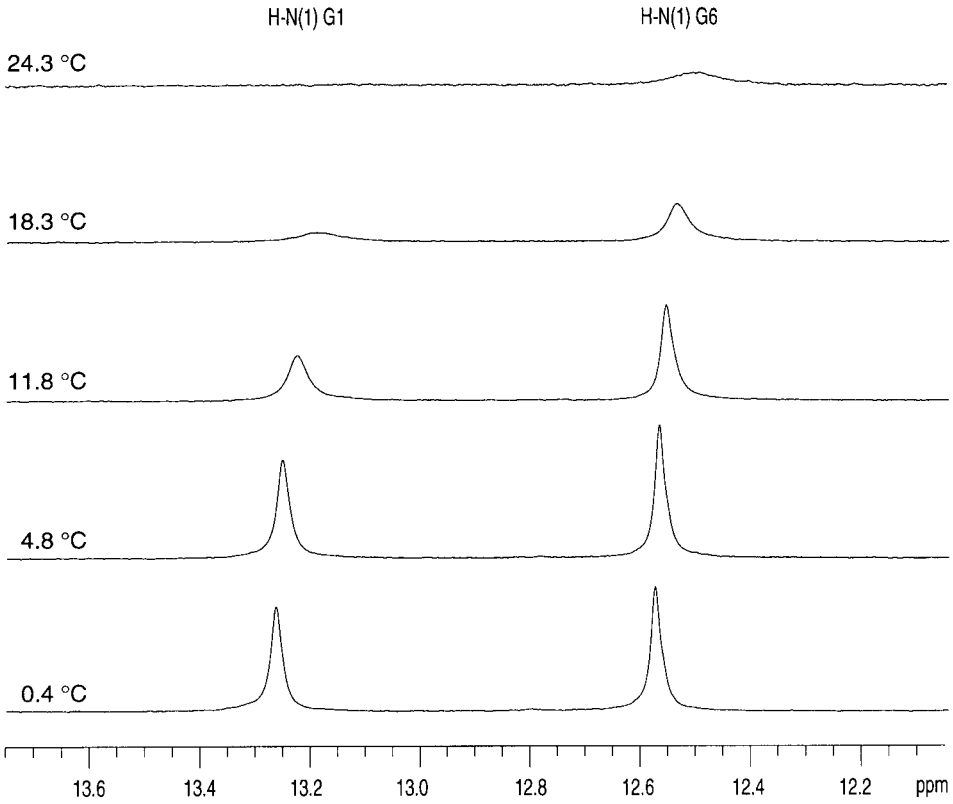

Fig. 2. *^{1}H-NMR Spectrum of* **1** *in H_2O/D_2O, 42 mM sodium-arsenate buffer, pH 7.0 at different temperatures.* The solvent signal was suppressed by excitation sculpting [9]

with the first and last of the linker CH_2 groups. The base protons were assigned based on *intra*-residual NOEs to the corresponding ribose protons. The chemical shifts at 4.8° are listed in *Table 1*, and those $^1H,^1H$ and $^1H,^{31}P$ coupling constants that could be determined by fitting of traces in the ^{31}P-decoupled and undecoupled DQF-COSY spectra are shown in *Table 2*.

Comparison of the ^1H-NMR spectra recorded at 27.3° and 4.8° revealed that, at 4.8°, the $^3J(1',2')$ coupling constants are more uniform than at 27.3°, and all lie below 1.3 Hz (*Fig. 2*). The ^{31}P resonances are less well-dispersed at 4.8° than at 27.3°; the signals for

Table 1. *^1H- and ^{31}P-NMR Chemical Shifts* [ppm] *and Assignments of* **1** *at 4.8°*

	H–C(1′)	H–C(2′)	H–C(3′)	H–C(4′)	H_a–C(5′)[a])	H_b–C(5′)	Me–Ne(1)
G1	5.86[b])	4.84	4.56	4.42	4.12	3.94	
C2	5.61	4.74	4.62	4.48	4.61	4.17	
A3	5.97	4.69	4.59	4.53	4.53	4.18	
m^{1}G4	5.58	4.19	4.51	4.37	4.46	4.05	3.29
U5	5.86	4.78	4.75	4.56	4.33	4.22	
G6	5.93	4.62[c])	4.58[c])	–[d])	4.51	4.37	
C7	5.55	4.35	4.42	4.35	4.57	4.08	
C8	5.71	3.93	4.14	3.93	4.40	4.01	
	H–C(5)	H–C(6)	H–C(2)	H–C(8)	NH_2[e])	H–N(1)	^{31}P[f])
G1				8.10	–	13.25	–0.79
C2	5.28	7.83			6.88, 8.55*		0.10
A3			7.34	7.93	–		–0.52
m^{1}G4				7.37	6.64		0.32
U5	6.03	8.16				–	0.85
G6				8.02	–	12.57	0.10
C7	5.23	7.54			7.02, 8.47*		–0.79
C8	5.62	7.72			–		–0.56

[a]) Not stereospecifically assigned; H_a indicates the proton resonating at lower field. [b]) Assignments based on COSY and $^1H,^{31}P$-COSY if not otherwise indicated. [c]) Assignment based solely on NOESY. [d]) Assignment not possible. [e]) The H-bonded NH_2 protons are labeled with *. [f]) The ^{31}P resonances are listed in 5′ to 3′ direction, chemical shifts in ppm *vs.* external 85% H_3PO_4.

Table 2. *Coupling Constants* [Hz] *of* **1** *at 4.8°*[a])

	$^3J(1',2')$[b])	$^3J(2',3')$[b])	$^3J(3',4')$[b])	$^3J(3',P)$[b])	$^3J(5'a,P)$[c])	$^3J(5'b,P)$[c])
G1	0.8	2.3	8.8	8.8		
C2	0.9	2.0	9.3	7.8	3.7	2.3
A3	0.8	2.2	9.2	7.8	3.6	2.1
m^{1}G4	1.3	4.2	–	7.1	3.5	2.0
U5	1.3	4.2	7.1	–	5.8	4.9
G6	–[d])	–	–	–	–	–
C7	–	–	–	–	3.8	0.8
C8	0.9	2.2	–		3.8	1.7

[a]) Estimated error ±0.5 Hz. [b]) Determined by fitting the appropriate number of *Gauss* lines to traces of the COSY spectrum. [c]) Determined by fitting 2 respectively 3 *Gauss* lines to traces of the cross-peaks correlating the geminal 5′-protons in COSY spectra with/without {^{31}P} decoupling. [d]) Not determined because of overlap.

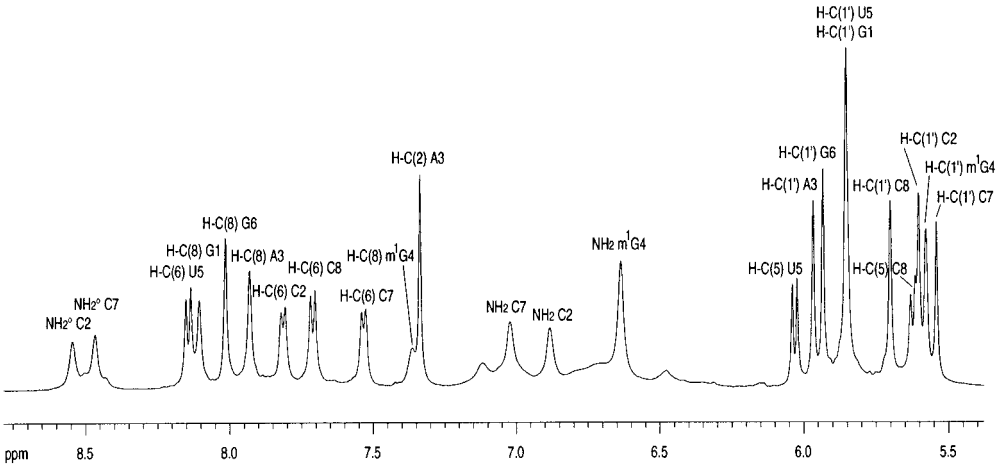

Fig. 3. *Region of the ^{1}H-NMR spectrum of* **1** *in H_2O/D_2O showing the base NH_2, H–C(6/8), and the H–C(1′) protons* (conditions: $c = 1.8$ mM in H_2O/D_2O 9:1, 42 mM sodium arsenate, pH 7.0, T 4.8°, solvent suppression with excitation sculpting [9])

G1-p-C2 and G6-p-C7, as well as C2-p-A3 and U5-p-G6, pairs that correspond to opposite positions in the duplex, have indistinguishable chemical shifts.

NOESY Spectra were measured at 4.8° in D_2O (42 mM sodium-arsenate buffer, pH 7.0) with mixing times of 50, 100, and 150 ms (*Fig. 4*), as well as in H_2O/D_2O 9:1 (50 mM sodium-arsenate buffer, pH 7.0) with mixing times of 70, 130, and 160 ms. From the build-up curves of the NOESY cross-peak volumes and calibration with known distances, the NOEs were qualitatively divided into the categories strong, medium, and weak, and assigned to the distance ranges 2.0–2.5, 2.5–3.5, and 3.5–5.0 Å, respectively. The unambiguously assignable *inter*-residual NOEs are schematically represented in *Fig. 5*.

3. Discussion. – In view of the relatively low melting point of 47°, it had to be expected that **1** would be a borderline case with respect to the exchange rate of labile imino protons. Fortunately, lowering the temperature from 27.3° to 4.8° slowed down the exchange process sufficiently to make two of the three possible imino NH resonances appear as sharp signals at low field. They could be assigned to residues C2 and C7 through a chain of NOEs $\{HN(1) \rightarrow NH_2^* \rightarrow NH_2 \rightarrow H-C(5)\}$. A weak sequential NOE between the two imino protons confirms that the bases G1/C7 and C2/G6 form two adjacent *Watson-Crick* pairs.

The ^3J(H,H) coupling constants of the ribose protons, in particular the low values of $^3J(1',2')$ and $^3J(2',3')$, point to predominant 3′-*endo* (N) conformations for all eight sugar rings [8]. Interestingly, the two nucleotides m^{1}G4 and U5 that are joined by the linker show slightly higher values for both $^3J(1',2')$ and $^3J(2',3')$. Together with the slightly smaller $^3J(3',4')$ observed for U5, this indicates either a static distortion of the regular N conformation or a dynamic contribution of S-type conformations for these two ribose units. The dangling nucleotide C8 at the 3′-end and A3, for which the

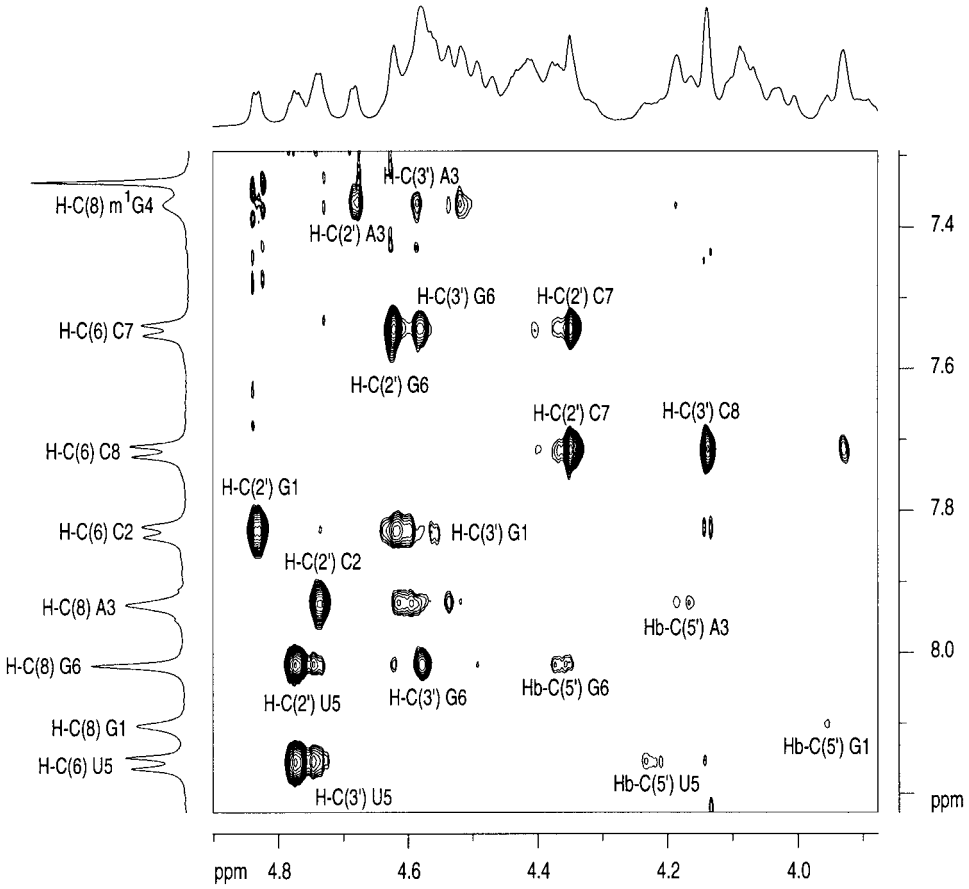

Fig. 4. *Expansion of the NOESY spectrum* (t_m 150 ms) *containing the cross-peaks between aromatic H–C(6/8) protons and ribose protons* (H–C(1′) protons not shown) (conditions: $c = 1.8$ mM in D_2O, 42 mM sodium arsenate, pH 7.0, T 4.8°)

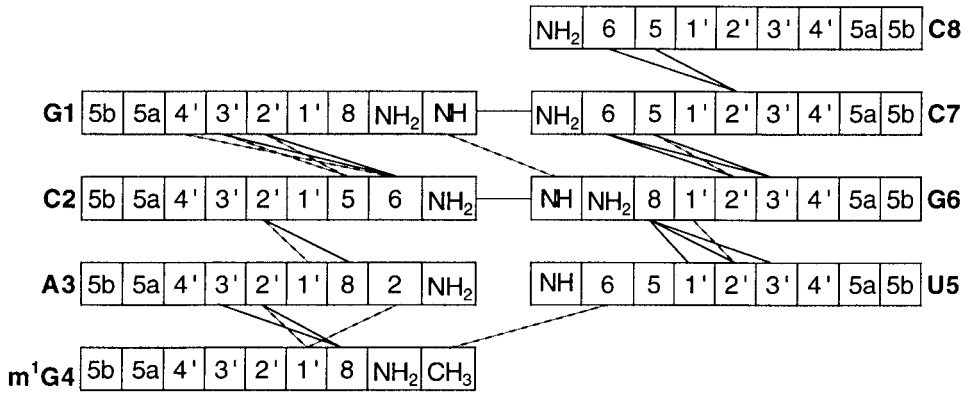

Fig. 5. *Graphical representation of the* inter-*residual NOEs observed for* **1** *at 4.8°* (– – –: weak; —: medium)

potential base pairing with U5 could not be proved directly, show regular N-type conformations.

The measured $^{1}H,^{31}P$ coupling constants, although not available for all nucleotides, allow deduction of qualitative information about the backbone angles β, γ, and ε. The small couplings between PO5′ and both geminal $CH_2(5')$ protons are only consistent with $\beta = t$. The significant $^{4}J(PO5',H-C(4'))$ four-bond couplings observed in the $^{1}H,^{31}P$-COSY and DQF.COSY spectra confirm $\beta = t$, and additionally restrict γ to g^{+} [10]. The vicinal coupling constants $^{3}J(PO3',H-C(3')) \approx 8$ Hz would be consistent with either $\varepsilon = g^{-}$ or t. Since, for $\varepsilon = g^{-}$, cross-peaks originating from long-range couplings $^{4}J(PO3',H-C(2'))$ would be expected in the $^{1}H,^{31}P$-COSY spectra, their absence points to $\varepsilon = t$. Hence, the available $^{1}H,^{31}P$-coupling data are consistent with the backbone-angle characteristics of A-RNA [11].

The intensities of *intra*-residual and *inter*-residual NOEs between the pyrimidine/purine H–C(6)/H–C(8) and the ribose protons (see *Fig. 4*) follow the pattern typically observed in regular A-double-helical domains of larger RNA molecules and indicate glycosidic torsion angles χ in the *anti* range for all bases [8b–d]. Because of the extreme overlap of the resonances of the ethylene-glycol units, no conformational information on the linker itself could be deduced from the NMR data.

Since units A3 and, with slight deviations, U5 assume a conformation that would allow them to form a *Watson-Crick* pair, the question arises why the corresponding imino proton of U5 exchanges too fast to be detected even at 0.4°. In NMR studies of DNA duplexes with ten to fourteen base pairs, it is commonly observed that the imino protons of one or two base pairs at both ends of the duplex are not detectable at temperatures near or slightly below room temperature [12]. Furthermore, this 'fraying' of the terminal nucleotides usually manifests itself in coupling constants that indicate mixed conformations for the corresponding ribose units.

At or below 4.8°, the base pair G1·C7 at the unlinked end of the duplex exhibits slow exchange of the imino proton and regular N-type/*anti* conformations for the paired nucleotides and their dangling neighbor (C8). In contrast, the imino proton of the potential base pair A3·U5 at the linked end is too labile to be detected, and both, U5 and the dangling $m^{1}G4$, show coupling constants that deviate slightly from the regular A-RNA values. Either this 'fraying' of the linked end is due to the inherently shorter lifetime of A·U *vs.* G·C base pairs [12], or it has to be attributed to strain or solvation changes induced by the linker. Further studies with 'mini duplexes' that have linkers at both ends and different sequences will help to answer this question.

In conclusion, it has been demonstrated that the solution structure of short duplexes with only three base pairs and one dangling base at each of the 3′-ends can be investigated by NMR, if the two strands are connected by at least one flexible linker. This opens the possibility to systematically study the influence of the sequence and neighbouring bases not only on the stability but also on the structure, an important question in the context of codon-anticodon pairing and the occurrence of frame shifts.

The *Austrian Academy of Science* is acknowledged for an *APART* fellowship (Austrian Programme for Advanced Research and Technology) to *R. M.* This work was supported by the *Austrian Science Fund*, Vienna (P13216-CHE). *M.-O. E.* thanks the Laboratorium für Organische Chemie, ETH-Zürich, for financial support.

REFERENCES

[1] a) B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J. D. Watson, 'Molecular Biology of the Cell', Garland, New York, 1994; b) J. H. Cate, M. M. Yusupov, G. Z. Yusupova, T. N. Earnest, H. F. Noller, *Science* **1999**, *285*, 2095.

[2] A. Sinclair, D. Alkema, R. A. Bell, J. M. Coddington, D. W. Hughes, T. Neilson, P. J. Romaniuk, *Biochemistry* **1984**, *23*, 2656; P. J. Romaniuk, T. Neilson, D. W. Hughes, R. A. Bell, *Can. J. Chem*. **1978**, *56*, 2249; P. J. Romaniuk, D. W. Hughes, R. J. Gregoire, T. Neilson, R. A. Bell, *J. Chem. Soc., Chem. Commun*. **1979**, 559.

[3] G. M. Clore, A. M. Gronenborn, L. W. McLaughlin, *J. Mol. Biol*. **1984**, *174*, 163.

[4] a) R. Micura, W. Pils, K. Grubmayr, *Angew. Chem.*, *Int. Ed.* **2000**, *39*, 922; b) R. Micura, *Chem. Eur. J.* **1999**, *5*, 2077.

[5] T. Xia, D. H. Mathews, D. H. Turner in 'Comprehensive Natural Product Chemistry', Eds.: D. Söll, S. Nishimura, P. B. Moore, Elsevier, Amsterdam, 1999, Vol. 6, pp. 21 – 47.

[6] M. Sprinzl, C. Horn, M. Brown, A. Ioudovitch, S. Steinberg, *Nucleic Acids Res.* **1998**, *26*, 148.

[7] a) W. Pils, R. Micura, *Nucleic Acids Res.* **2000**, *28*, 1859; b) S. Pitsch, P. A. Weiss, X. Wu, D. Ackermann, T. Honegger, *Helv. Chim. Acta* **1999**, *82*, 1753.

[8] a) F. A. A. M. de Leeuw, C. Altona, *J. Chem. Soc., Perkin Trans. 2*, **1982**, 375; b) S. S. Wijmenga, B. N. M. van Buuren, *Progr. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287; c) G. Varani, F. Aboul-ela, F. H.-T. Allain, *Progr. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51; d) K. Wüthrich, 'NMR of Proteins and Nucleic Acids', John Wiley & Sons, New York, 1986.

[9] T.-L. Hwang, A. J. Shaka, *J. Magn. Reson. A*. **1995**, *112*, 275.

[10] M. J. J. Blommers, D. Nanz, O. Zerbe, *J. Biomol. NMR*, **1994**, *4*, 595.

[11] S. S. Wijmenga, M. M. W. Mooren, C. W. Hilbers in 'NMR of Macromolecules: A Practical Approach', Ed. G. C. K. Roberts, IRL Press, Oxford, 1993, p. 225.

[12] M. Leijon, A. Gräslund, *Nucleic Acids Res.* **1992**, *20*, 5339, and refs. cit. therein.

Received June 9, 2000