122. Pyranosyl-RNA ('p-RNA'): Base-Pairing Selectivity and Potential to Replicate

Preliminary Communication[1])

by **Stefan Pitsch, Ramanarayanan Krishnamurthy, Martin Bolli, Sebastian Wendeborn, Armin Holzner, Mark Minton, Catherine Lesueur, Irène Schlönvogt, Bernhard Jaun,** and **Albert Eschenmoser***

Laboratory of Organic Chemistry, Swiss Federal Institute of Technology, ETH-Zentrum, Universitätstrasse 16, CH–8092 Zürich

In memoriam Rolf Scheffold

(1.IX.95)

Base pairing in p-RNA (β-D-ribopyranosyl-(4′ → 2′)-oligonucleotides) is not only stronger than in DNA and RNA, but also more selective in the sense that it is strictly confined to the *Watson-Crick* mode. Homopurine sequences (tested up to decamers) exist as *single strands* under conditions where they undergo *reverse-Hoogsteen* self-pairing in homo-DNA or *Hoogsteen* self-pairing in DNA. This exceptional pairing selectivity is rationalized as hinging on two structural features of p-RNA: the large inclination between backbone axis and base-pair axes in p-RNA duplexes, and the higher rigidity of the p-RNA backbone compared with RNA, DNA, and homo-DNA. The most important consequence of the pairing selectivity refers to the potential of p-RNA to replicate. Replicative copying of sequence information by nonenzymatic template-controlled ligation is not hampered by self-pairing of guanine-rich templates, as it is known to be the case in the RNA series. We have demonstrated two replicative cycles in which G-rich p-RNA-octamer templates induce sequence-selective ligation of tetramer-2′-phosphate derivatives to complementary C-rich octamer sequences, and in which the latter, with comparable efficiency, induce corresponding ligation reactions back to the original G-rich octamers. Ligation is most satisfactorily achieved after pre-activation of the 2′-phosphate groups as 2′,3′-cyclophosphate derivatives; in this version, the process does not proceed as oligo*condensation*, but as a genuine oligo*merization*. This is of considerable promise for the search for potentially natural conditions under which homochiral p-RNA strands might self-assemble and self-replicate.

Introduction. – Our first communication on the chemistry of pyranosyl-RNA [1][1]) described the synthesis of p-RNA-oligonucleotide sequences containing adenine and uracil as nucleobases, as well as our observations on some of their pairing properties. Here, we report, again in preliminary form, on the pairing behavior of p-RNA sequences containing guanine, cytosine, isoguanine (I), 2,6-diaminopurine (Dp), and thymine as nucleobases. We also describe the results of our first successful experiments on the capacity of p-RNA sequences to replicate by template-controlled ligation processes.

Automated solid-phase synthesis of p-RNA oligonucleotides from the β-D-ribopyranosyl phosphoramidite derivatives shown in *Fig. 1* followed the procedure described in [1], with the exception of the final deprotection step (cleavage of the axial 3′-BzO group) which has now been carried out by hydrazinolysis throughout. Careful control of condi-

[1]) Second communication in the series '*Pyranosyl-RNA ('p-RNA')*'; for the first communication (which represented Part VII of the series '*Why Pentose- and Not Hexose-Nucleic Acids?*'), see [1]. The present paper is also communication No. 16 in the series '*Chemistry of α-Amino Nitriles*', for communication No. 15, see [2].

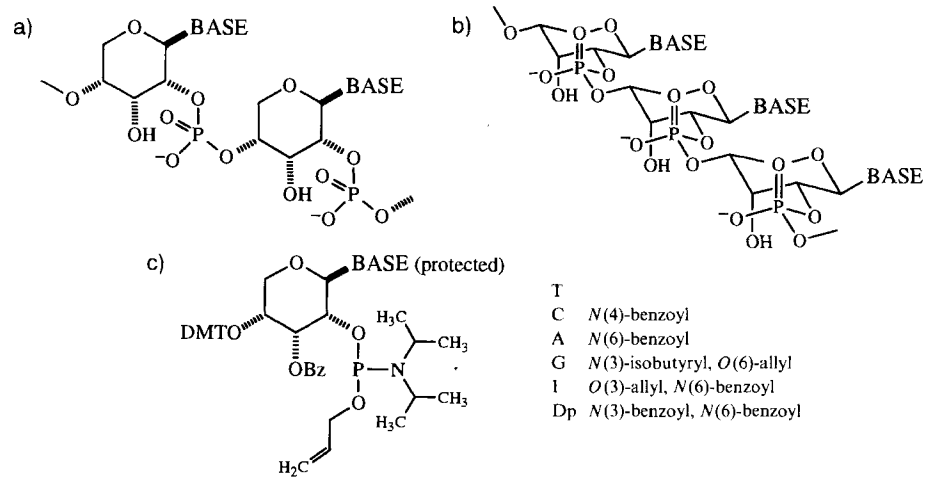

Fig. 1. a) *Constitutional type of β-*D*-ribopyranosyl-(4′→2′)-oligonucleotides (p-RNA).* b) *Pairing conformation* (idealized) *of p-RNA-oligonucleotide strand as deduced by conformational analysis* [1] [12] [16]. c) *Building blocks for automated synthesis of p-RNA-oligonucleotides.* T = thymine, C = cytosine, A = adenosine, G = guanine, I = isoguanine, Dp = 2,6-diaminopurine, DMT = 4,4′-dimethoxytrityl. Procedures and standards adopted in automated solid-phase synthesis (*Pharmacia* Gene-Assembler Plus) as described in [1].

tions (hydrazine hydrate/H_2O 1:4 at 4° during 24–36 h) prevents phosphodiester cleavage (strand scission) and makes the procedure compatible with the (deprotected) nucleobases thymine, cytosine, guanine, isoguanine, and 2,6-diaminopurine (but not uracil and xanthine). Cleavage (using *Noyori*'s method [3]) of allyloxy groups utilized for protection of phosphodiester as well as the carbonyl functions of the guanine and isoguanine bases has to precede the hydrazinolysis step. Phosphoramidite derivatives containing isoguanine and 2,6-diaminopurine were derived from the corresponding 2,6-dichloropurine nucleoside [4] (obtained from *β*-D-ribopyranose tetrabenzoate by the *Vorbrüggen-Hilbert-Johnson* procedure [5]) analogously to reaction sequences developed earlier in the homo-DNA series [6]; guanine member was prepared by nucleosidation using the *N*(2)-isobutyryl derivative of guanine or (preferably) 6-chloro-2-(isobutyrylamino)purine [7], followed by replacement of the Cl by the allyloxy group [8]. The preparation of the thymine- and cytosine-containing members followed standard procedures [1] [9]. Deprotected p-RNA-oligonucleotides were purified by HPLC [9] under denaturing conditions to purities exceeding 95%, and their composition was in each case controlled by matrix-assisted laser-desorption-ionization-time-of-flight (MALDI-TOF) mass spectrometry [10].

A ^{1}H-, ^{31}P-, and ^{13}C-NMR study of the duplex derived from the self-complementary p-RNA sequence p-Ribo(C-G-A-A-T-T-C-G) fully corroborates the constitutional and configurational assignments based on the programmed automated oligonucleotide-strand synthesis and on *Watson-Crick* pairing being expected in analogy to DNA and homo-DNA. Furthermore, the study fully confirms the *type* of conformation derived by qualitative conformational analysis [1] for the backbone in a p-RNA duplex (*Fig. 1*). A molecular-dynamic study, initiated by *Romain Wolf* (personal communication by Dr. *R. M. Wolf*, *Ciba AG*, Basel), of the same duplex, using the duplex conformation derived by conformational analysis as a starting model, points to a quasi-linear structure which

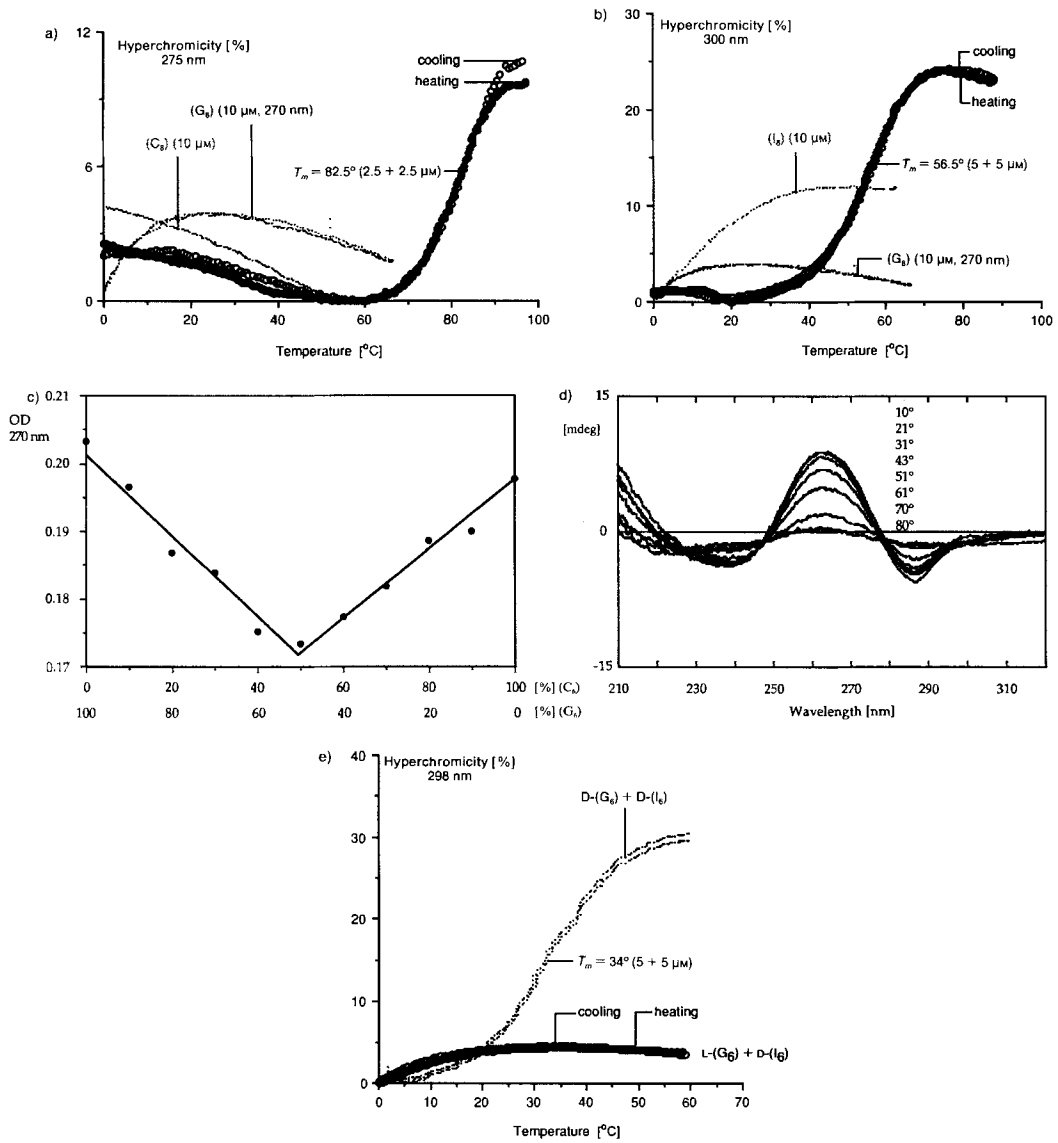

Fig. 2. a) *and* b) *UV Melting curves of p-RNA duplexes* $(G_8)\cdot(C_8)$ *(1:1)* (*Table 1*) *and* $(G_8)\cdot(I_8)$ *(1:1)* (*Table 2*): Included for comparison are the corresponding curves determined for single strands before mixing ($c \approx 10$ μM). Note the non-pairing of these single strands above 0°, and the aggregation of the duplexes at lower temperatures. c) *Determination of stoichiometry* [17] *of p-RNA duplex* $(G_6)\cdot(C_6)$, *UV absorption at 270 nm, 20°, of mixtures* $0\rightarrow100\%$ $(G_6)/100\%\rightarrow0\%$ (C_6), $c(\Sigma) = 7$ μM. d) *Temperature-dependent CD spectra of p-RNA duplex* $(G_8)\cdot(C_8)$ *at 10, 21, 31, 43, 51, 61, 70, and 80°* (top to bottom curve, $\lambda = 260$ nm, $c = 3.5$ μM), *indicating a* T_m *value of 58°*. CD Spectra of single strands do not show sigmaoidal change of ellipticities with rising temperature (not shown). e) *Temperature dependence of UV absorption at 298 nm of 1:1 mixture of homochiral p-RNA strands with opposite sense of chirality* L-(G_6) *and* D-(I_6), *5* μM *each* (included for comparison: UV melting curve of corresponding homochiral 1:1 D/D-mixture). All measurements (except where stated otherwise) in H_2O, 0.15M NaCl, 0.01M Tris·HCl, pH 7.0.

allows for excellent *inter*strand stacking of the bases and shows a weak left-handed twist that would extrapolate to a helical pitch about three times as large as the one in DNA (to be published later in detail in this journal).

Base-Pairing Selectivity. – p-RNA Single strands containing the canonical bases (A, U, T, C, G) in sequences complementary to each other in *antiparallel* strand orientation form duplexes which are uniformly more stable relative to their single strands than corresponding duplexes in the RNA and DNA series and in some cases even more stable than corresponding homo-DNA duplexes (*Table 1*). *In sharp contrast to homo-DNA* [11], *homopurinic p-RNA sequences containing either adenine* (see [1]) *or guanine do not show self-pairing in the reverse-*Hoogsteen *mode under standard conditions* (*ca.* 10 μM in aqueous solution, 0.15M NaCl, Tris buffer, pH 7, tested up to decamers). Specifically and most importantly, oligomers p-Ribo(G_n) with $n = 6$, 8, or 10 exist as single strands at room temperature as shown by their UV and CD spectra as a function of temperature (*Fig. 2*) as well as by gel electrophoresis under non-denaturing conditions. The same is true – again in contrast to homo-DNA [12] [13] – for p-RNA oligomers containing 2,6-diaminopurine or isoguanine (tested up to octamers). On the other hand, the novel type of purine-purine *Watson-Crick* pairing between guanine and isoguanine (*Fig. 3*), observed for the first time in homo-DNA [6] [12] [13], is also found in p-RNA. Purine-purine *Watson-Crick* pairing of this type can be expected to constitute a general property of pairing systems whose duplexes have a quasi-linear (only weakly twisted) ladder structure.

Table 1. *Data of p-RNA Duplexes in Comparison with Corresponding DNA (RNA) and Homo-DNA Duplexes.* T_m Values determined by temperature-dependent UV spectroscopy [11], detection wavelengths 260 nm (A,T,U) and 270 nm (G,C), in *ca.* 10 μM aqueous solution, 0.15M NaCl, 0.01M Tris·HCl, pH 7.0 (when not stated otherwise). Data for DNA and homo-DNA duplexes taken from [11] (data of d(CGAATTCG) duplex newly determined); thermodynamic data determined from concentration dependence of T_m values [15] [11].

Duplex base sequences	Oligonucleotide system	T_m° (10 μM)	ΔG (25°) [kcal/mol]	ΔH [kcal/mol]	$T\Delta S$ (25°) [kcal/mol]
-AAAAUUUU	p-RNA[a])	21°	6.2	–44.3	–38.1
UUUUAAAA-	RNA[a])	5°[b])	3.3	–51.0	47.7
-ATATATAT	p-RNA	38°	–9.2	58.7	–49.5
TATATATA-	DNA	7°	–3.8	–48.0	–44.2
-CGAATTCG	p-RNA	60°	–12.6	–54.9	–42.3
GCTTAAGC-	DNA	32°	–8.2	–54.4	46.2
	homo-DNA	51°	–10.6	–55.3	–44.7
-GCGCGC	p-RNA	62°	–11.3	–40.5	–29.2
CGCGCG-	DNA	52° (50 μM)	–10.8	–61.5	–50.3
	homo-DNA	64° (50 μM)	–10.3	–39.7	–29.4
-CGCGCG	p-RNA	65°	–12.5	–47.9	–35.4
GCGCGC-	DNA	53° (39 μM)	–10.2	–53.7	–43.5
	homo-DNA	62° (35 μM)	–10.6	–41.4	–30.8
-GGGGGG	p-RNA	61° (3.5 + 3.5 μM)	–13.5	54.3	–40.8
CCCCCC-	DNA	33° (25 + 25 μM)	–7.1	–61.3	–54.2
	homo-DNA	61° (25 + 25 μM)	–10.9	39.5	–28.0

[a]) 1.00M NaCl, 0.01M Na_2HPO_4·NaOH, 10^{-4}M EDTA.
[b]) Extrapolated from thermodynamic data taken from [14].

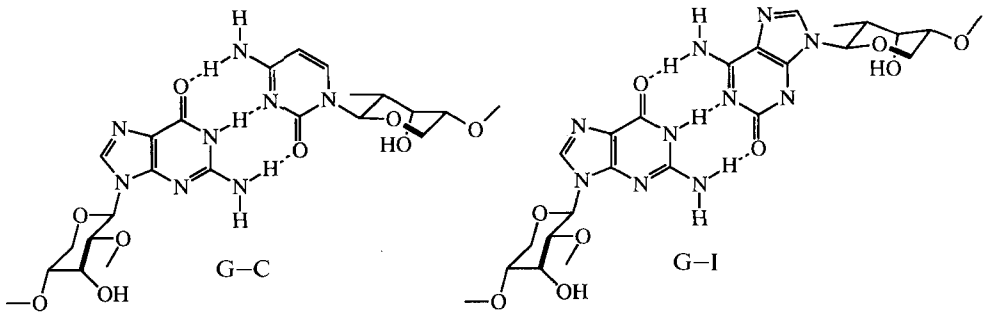

Fig. 3. *The two* Watson-Crick *base pairs involving guanine observed in the p-RNA series: guanine-cytosine and guanine-isoguanine*

Tridentate *Watson-Crick* pairing in p-RNA is *enantioselective*. Homochiral strands of p-Ribo(G_6) and p-Ribo(I_6) with *opposite* sense of chirality do not show any pairing (*Fig. 2*) under conditions where the strands with the *same* sense of chirality form a duplex that melts at 34° (10 μM). In the case of the purine-pyrimidine duplexes p-Ribo(G_n)·p-Ribo(C_n) ($n = 6$ and 8), the strong pairing of the homochiral strands with identical sense of chirality ($T_m = 61°$ and 82° respectively, see *Table 2*) is replaced by a drastically weaker interaction ($T_m \approx 0°$ and *ca.* 15°), presumably of a didentate reverse-*Watson-Crick* type, when the two homochiral strands are opposite in their sense of chirality. An analogous behavior is observed with the strands p-Ribo(Dp_8) and p-Ribo(T_8) (*Table 2*).

The exceptional selectivity of p-RNA with regard to pairing modes as well as the enantioselectivity of its pairing are rationalized as hinging on three specific structural features (see also [1]). First, the pyranosyl-(4′ → 2′)-oligonucleotide backbone is expected to be much less flexible (less adjustable) than that of homo-DNA or DNA. The backbone unit of DNA has *six*, that of homo-DNA *five*, but the backbone unit of p-RNA only *four* flexible bonds: in contrast to the furanose ring in DNA, the pyranosyl chairs of homo-

Table 2. T_m *Values of p-RNA Duplexes* (methods and conditions, see *Table 1*), *Illustrating* Enantioselectivity *of* Watson-Crick *Base Pairing*. Thermodynamic data for duplex D-(G_8)·D-(I_8): ΔG (25°) = −12.5 kcal/mol ($\Delta H = -50.2$ and $T\Delta S = -37.7$ kcal/mol), determined under standard conditions (see *Table 1*).

Duplex base sequences		T_m ($c = 5 + 5$ μM)
Homochiral strands of same sense of chirality (D,D)	Homochiral strands of opposite sense of chirality (L,D)	
D-(G_8)·D-(I_8)		56° (*Fig. 2*)
	L-(G_8)·D-(I_8)	≪ 0°
D-(G_6)·D-(I_6)		34° (*Fig. 2*)
	L-(G_6)·D-(I_6)	≪ 0° (*Fig. 2*)
D-(G_8)·D-(C_8)		82° ($c = 2.5 + 2.5$ μM)
	L-(G_8)·D-(C_8)	~ 15°
D-(G_6)·D-(C_6)		61° ($c = 3.5 + 3.5$ μM)
	L-(G_6)·D-(C_6)	~ 0°
D-(Dp_8)·D-(T_8)		50°
	L-(Dp_8)·D-(T_8)	~ 0°

DNA and p-RNA, with three large substituents in equatorial positions will not allow rotation around the *endocyclic* backbone bonds, since this would imply chair inversion. Second, the phosphodiester groups in p-RNA are *per se* sterically more constrained than in homo-DNA or DNA, being attached on both sides to secondary C-atoms. Third, the axes defined by the p-RNA backbone in its idealized (quasi-linear) pairing conformation

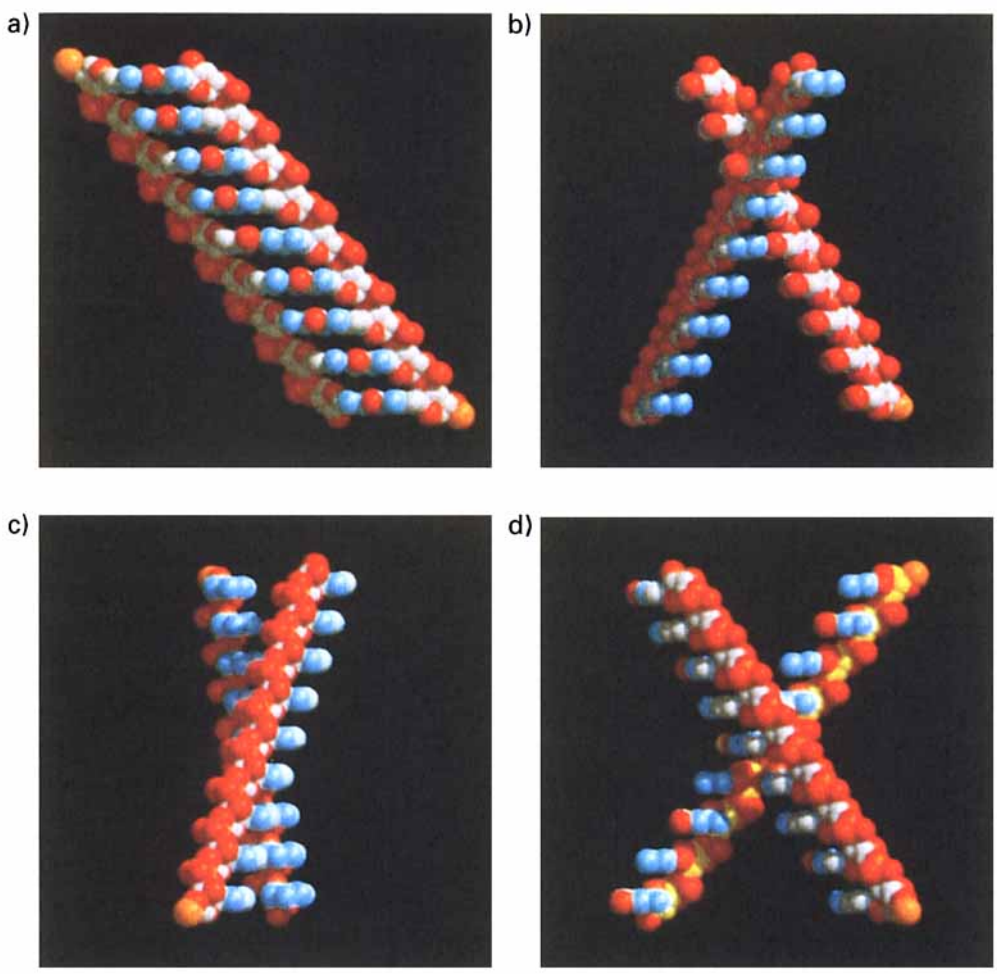

Fig. 4. *With p-RNA single strands in their idealized (linear) repetitive conformation* (all single bonds staggered) *purine-pyrimidine base pairing in antiparallel strand orientation in the* Watson-Cirk *mode requires minimal conformational adjustment of the two strands* (a). *In contrast, purine-pyrimidine pairing in the* Hoogsteen (b) *or purine-purine in the reverse-*Hoogsteen (c) *mode demands extended conformational changes in the backbones for adjusting base pairs. The same is true, if two homochiral purine-pyrimidine strands of opposite chirality* (nucleobases of one strand in *syn*-conformation) *were to pair in the* Watson-Crick *mode with antiparallel strand orientation* (d). The computer drawings show complementary single-strand models in idealized (linear) conformation and antiparallel strand orientation with one base pair adjusted in pertinent pairing mode (see also [1]); note that in Fig. 9 of [1], the drawing of the backbone-base pair shear of a p-RNA duplex has the inverted sense of chirality.

and by the *Watson-Crick* base pairs are strongly inclined toward each other (more so than they are in homo-DNA or, for that matter, in DNA and RNA) by constitutional as well as conformational reasons. Besides providing the spatial relationships for nearly optimal *inter*strand base stacking in p-RNA duplexes, it is this structural feature that is considered to be responsible for the strand orientation in duplexes being strictly antiparallel and – together with the backbone's rigidity – for the preference of p-RNA strands (with identical sense of chirality) to pair in the *Watson-Crick* mode and to avoid both *Hoogsteen* and reverse-*Hoogsteen* pairing (*Fig. 4*).

Replicative Template-Directed Ligation. – Whether a given structural nucleic-acid alternative is relevant to the chemical aetiology of the natural nucleic acids' structure, or, in more general terms, relevant to the problem of the origin of biological information, depends on criteria such as the alternative system's potential for constitutional self-assembly from natural constituents, its pairing properties, and, above all, on its potential for self-replication in the absence of enzymes as we know them today. The results of our first successful experiments carried out on p-RNA's suitability for template-directed copying of sequence information are as follows.

The tetramer-2′-phosphate p-Ribo(C_4)-2′-p is ligated to the octamer p-Ribo(C_8)-2′-p and the dodecamer p-Ribo(C_{12})-2′-p in the *presence*, but *not in the absence* of the template p-Ribo(G_8) by the water-soluble carbodiimide EDC [19] under the reaction conditions given in *Fig. 5*. The reaction is accompanied by formation of the 2′,3′-cyclophosphate derivative of the 2′-phosphorylated oligomers by the carbodiimide reagent. Analogously, the tetramer p-Ribo(G_4)-2′-p is ligated to the octamer p-Ribo(G_8)-2′-p and dodecamer p-Ribo(G_{12})-2′-p (and corresponding 2′,3′-cyclophosphate derivatives) with comparable efficiency in the presence of the complementary template p-Ribo(C_8).

Ligation products were separated by HPLC (*Fig. 5*) and characterized by their molecular weights determined by MALDI-TOF-MS [10] and, in the case of the main products p-Ribo(C_8)-cp, p-Ribo(C_8)-p, p-Ribo(G_8)-cp, and p-Ribo(G_8)-p, by confirming duplex formation with their complementary template strands p-Ribo(G_8) and p-Ribo(C_8), respectively, by temperature-dependent UV and CD spectroscopy (melting and CD curves similar to those of the duplex formed by the 1:1 mixture of the two authentic template sequences; $T_m \approx 82°$, $c = 2.5 + 2.5$ μM). The structure assignments of these ligation products is further based on HPLC identification (co-injection) of the dephosphorylation products, obtained from p-Ribo(C_8)-2′-p and p-Ribo(G_8)-2′-p by treatment with alkaline phosphatase, with authentic p-Ribo(C_8) and p-Ribo(G_8), respectively, as well as by converting p-Ribo(C_8)-2′-p and p-Ribo(G_8)-2′-p into the corresponding 2′,3′-cyclophosphates (1M DEC·HCl in H_2O; 1.5 h; room temperature) and identifying them by HPLC with the ligation-product fractions p-Ribo(C_8)-2′,3′-cp and p-Ribo(G_8)-2′,3′-cp. Finally, the ligation-product fractions p-Ribo(C_8)-2′-p and p-Ribo(C_8)-2′,3′-cp were compared directly (HPLC, co-injection) with authentic material prepared for comparison purposes by automated synthesis using the solid-support methodology applied in the preparation of the tetramer-2′-phosphates p-Ribo(C_4)-2′-p and p-Ribo(G_4)-2′-p (*cf.* the caption of *Fig. 5*).

In control experiments carried out under identical conditions *but in the absence of template*, it was determined by HPLC that ligation in both series – if it had occurred at all – had done so to an extent of less than *ca.* 0.01% yield (*Fig. 5*). In these experiments, the

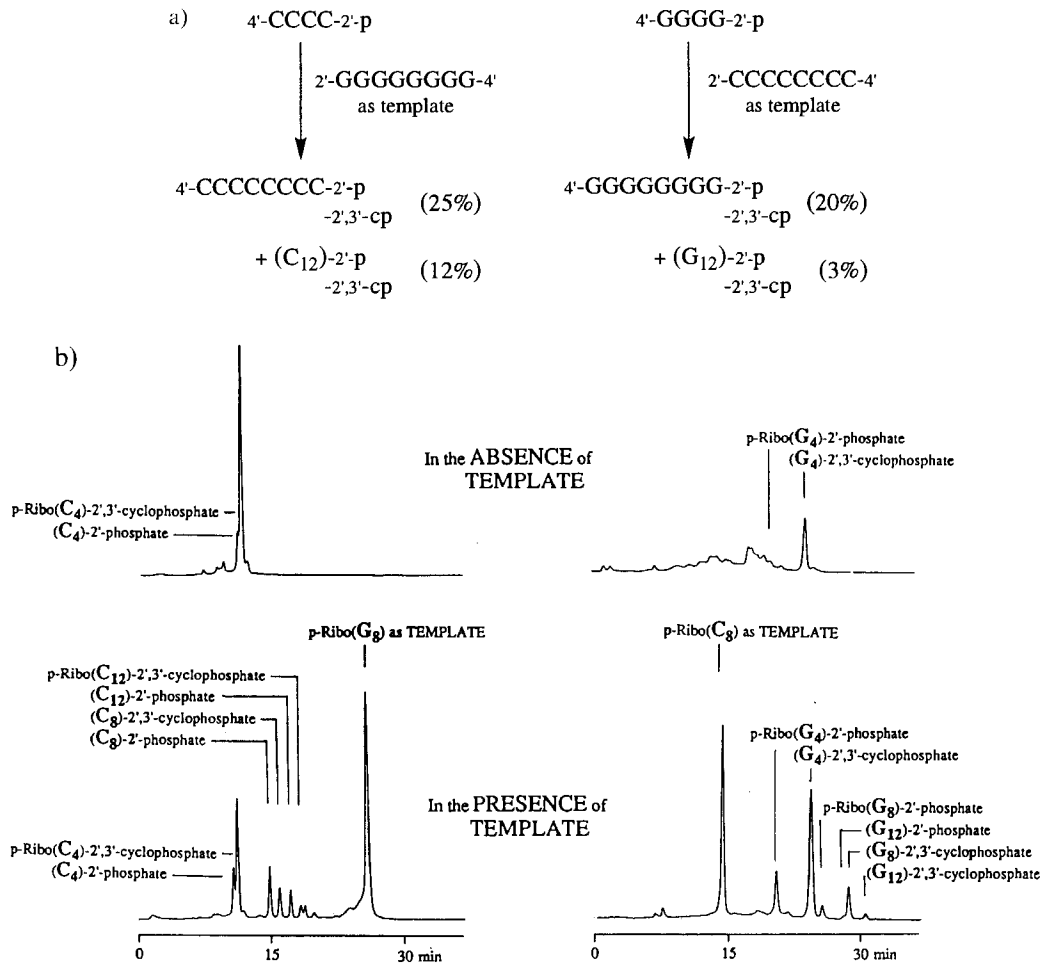

Fig.5. a) *Replicative ligation experiments using the water-soluble carbodiimide DEC · HCl* (*N*-[3-(dimethylamino)-propyl]-*N*-ethylcarbodiimide hydrochloride) *as activating agent*. Reaction conditions: H_2O; 450 μM in tetramer-2′-phosphates, 150 μM in template octamers (free 2′-ends), 1.5M NaCl, 0.125M $MgCl_2$, 0.1M HEPES buffer, pH 6.3; DEC · HCl in large excess, 40°, 36 h. Product isolation by first desalting reaction mixtures using a *Sep-Pak* cartridge (washing with 0.1M triethylammonium hydrogen carbonate buffer and elution of oligonucleotide with H_2O/MeCN 3:1), followed by HPLC (see below) and subsequent desalting of the individual fractions (as above). Material thus isolated (purity > 95%) amounts to 70–90% of products in reaction mixture. HPLC: *Mono-Q-HR 5/5* column (0.5 × 5 cm; *Pharmacia*), elution at pH 11.5 (10 mM sodium phosphate buffer = denaturing conditions) with linear NaCl gradient (0 → 1.0M within 30 min); detection at 270 nm. Tetramer-2′-phosphates prepared by automated solid-phase oligonucleotide synthesis using the solid-support starter group described in [18] (obtained from Dr. *U. Pieles* and Dr. *H. Moser*, *Ciba AG*, Basle). b) *Reaction mixtures of ligation experiments in the presence and in the absence of templates as observed by HPLC* (see above) *after desalting* (for product assignements, see text). Product yields refer to templates and were determined by HPLC with reference to standard solutions of corresponding (unphosphorylated) product oligonucleotides. Ligation products exist as duplexes with templates in the reaction mixture, but as single strands under HPLC conditions.

reactant p-Ribo(C_4)-2′-p remained essentially intact (recovery 86%) apart from being largely converted to the corresponding 2′,3′-cyclophosphate by the carbodiimide, whereas in the case of the reactant p-Ribo(G_4)-2′-p only *ca.* 25% was recovered (as 2′,3′-cyclophosphate), and the remainder observed in HPLC (*Fig. 5*) to consist of a complex mixture of unidentified components, the origin of which is to be sought in modifications of the (single-stranded) tetramer's guanine bases by the carbodiimide. In similar control experiments, but without carbodiimide, p-Ribo(G_4)-2′-p was recovered in 94% yield (HPLC). It is noteworthy that such modifications of guanine nuclei by the carbodiimide do not seem to occur, when the guanines are engaged in duplex pairing, as is the case in the *presence* of the template.

As *Fig. 5, b*, shows, in the ligation experiment described above by-products are observed which, according to their MALDI-TOF-MS data, are the corresponding dodecamers. Their formation is of interest, since it points to the ability of the corresponding templates to act in shifted frames.

The bulk of product formation in the ligations by carbodiimide did not proceed *via* the tetramer-2′,3′-cyclophosphates which are formed under the reaction conditions in a competing side reaction; this was shown by control experiments. However, while 2′,3′-cyclophosphates of the p-RNA series are fairly stable toward hydrolysis under slightly acidic conditions, they have been observed to be hydrolyzed to mixtures of 2′- and 3′-phosphates in slightly basic aqueous solution, provided that, *e.g.*, Mg^{2+} is present. Such

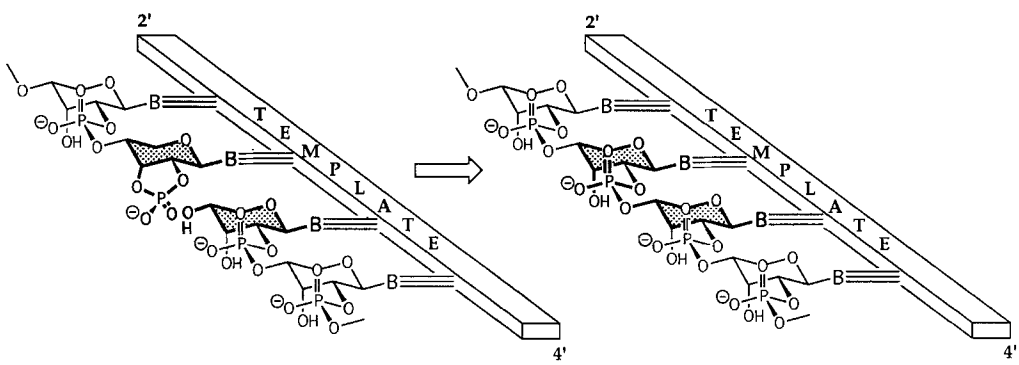

Fig. 6. *Ligation of p-RNA oligonucleotide-2′,3′-cyclophosphate on p-RNA templates.* Note that, due to the large inclination between backbone and base-pair axes, (4′→2′)-opening of the 2′,3′-cyclophosphate group to form a (4′→2′)-phosphodiester junction is expected to be unambiguously favored over (4′→3′)-, (3′→2′)-, or (3′→3′)-opening.

conditions can be kept mild enough to maintain the phosphodiester groups of the p-RNA backbone essentially untouched. Therefore, conditions were searched for under which template-assisted ligation of p-RNA tetramers to octamers could be achieved by simple *isomerization* (transesterification) of 2′,3′-cyclophosphates (*Fig. 6*) without any further activating reagent. This, indeed, is possible, as the experiments summarized in *Fig. 7* illustrate.

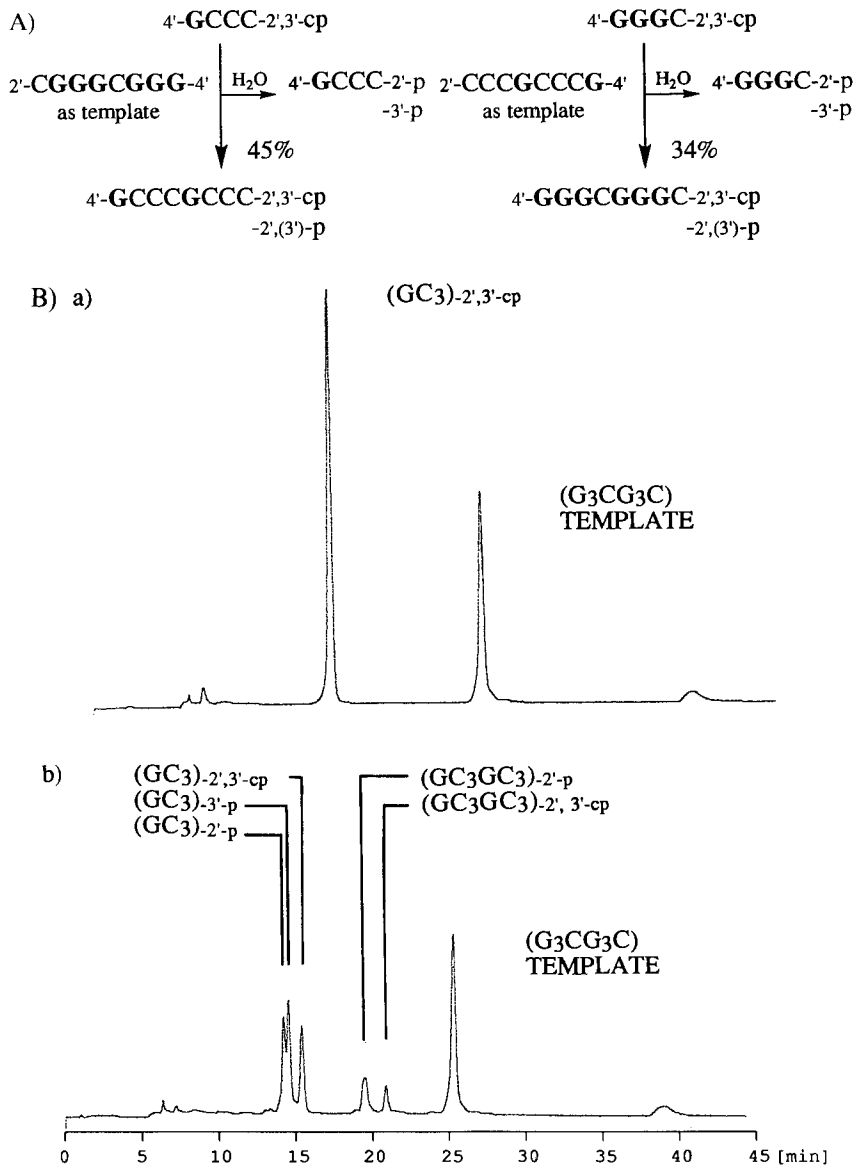

Fig. 7. A) *Replicative ligation experiments using 2',3'-cyclophosphates as activated form of ligand oligonucleotides.* Reaction conditions: H_2O; 450 μM in tetramer-2',3'-cyclophosphates, 150 μM in template octamers (with free 2'-ends); 1.0M NaCl, 0.025M $MgCl_2$, 0.1M HEPES buffer, pH 8.0; *ca.* 25°, 1–4 weeks. Product isolation and HPLC, see caption of *Fig. 5*. Tetramer-2',3'-cyclophosphates were prepared from corresponding 2'-phosphates (see caption of *Fig. 5*) by treatment with DEC·HCl (2 h, r.t.). B) *Reaction mixture at* a) *the beginning and* b) *the end of a ligation experiment in the presence of the G-rich template referred to above* (for product assignments, see text). C) a) *HPLC Traces of ligation experiment referred to in* B *followed during 4 weeks, yields of ligation products* (see B) *with reference to template.* b) *Same experiment in the* absence of template. Note progressive hydrolysis of ligand-2',3'-cyclophosphates to a mixture of 2'- and 3'-phosphates. c) *Same experiment in the* presence of template only. Note the stability of the template over 4 weeks under the conditions of the ligation.

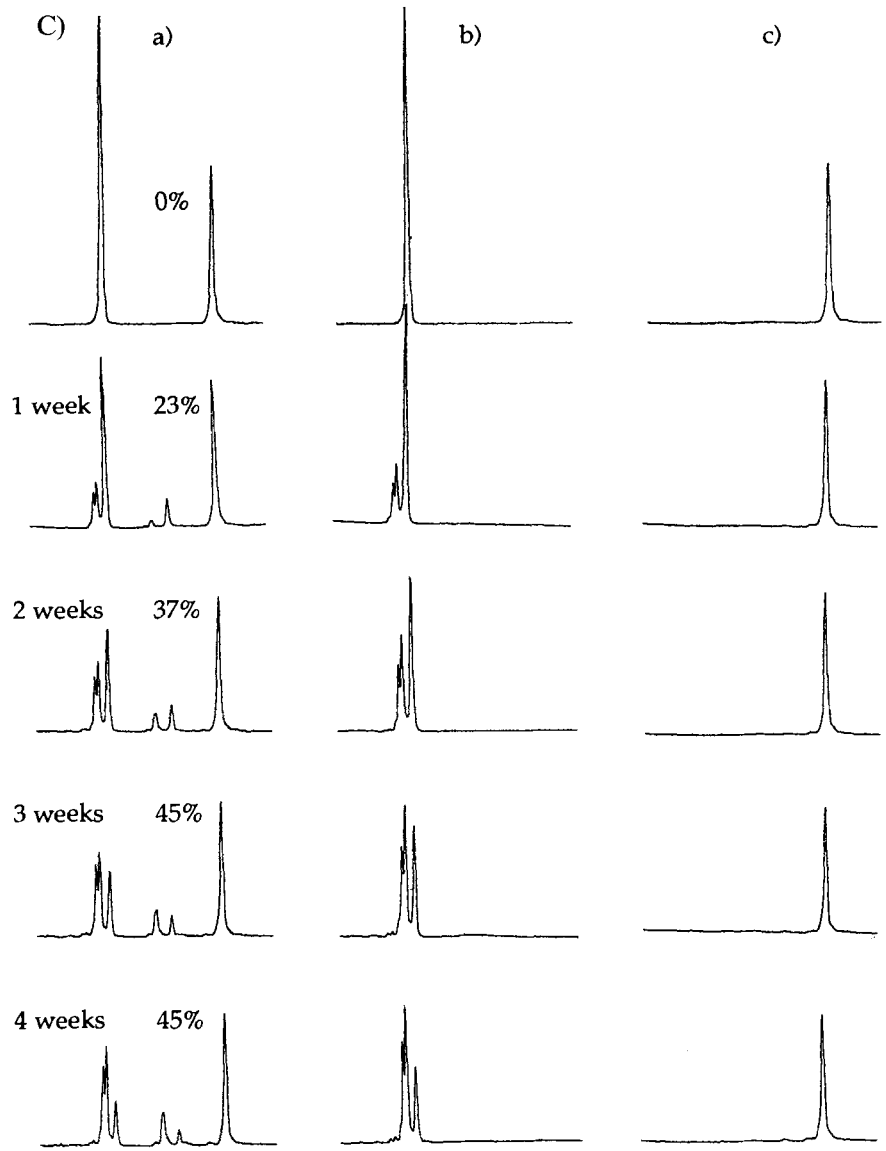

Fig. 7 (cont.)

The p-RNA-tetramer-2′,3′-cyclophosphate p-Ribo(GCCC)-2′,3′-cp in the presence of 0.33 mol-equiv. of the octamer template p-Ribo(GGGCGGGC) and 0.025M $MgCl_2$, at pH 8 and room temperature undergoes two reactions: ligation to the octamer p-Ribo(GCCCGCCC)-2′,3′-cp and, alternatively, cyclophosphate hydrolysis to the p-Ribo(GCCC)-2′-p besides smaller amounts of what is assumed to be the corresponding 3′-phosphate. Again, in the *absence* of template, but under otherwise identical conditions,

there is no ligation detectable by HPLC (less than *ca.* 0.01%). The two competing reactions, ligation and hydrolysis, proceed with comparable rate and, under the conditions used, very slowly (days to weeks), but essentially without side reactions, the Mg^{2+} concentration being kept low enough to avoid any strand scission by phosphodiester hydrolysis. As the reaction proceeds, the ligated octamer accumulates as the 2′-phosphate (besides small amounts of – presumably – the 3′-phosphate). In the complementary system, ligation of p-Ribo(GGGC)-2′,3′-cp to p-Ribo(GGGCGGGC)-2′-p and -3′-p on the template p-Ribo(GCCCGCCC) behaves analogously, except that the competition between ligation and hydrolysis is slightly shifted toward cyclophosphate hydrolysis. Product characterizations and structure assignments were carried out similarly to the carbodiimide experiments by MALDI-TOF-MS, UV, and CD characterization of duplex formation between (isolated) ligation products and template strands, as well as by HPLC identification of products after interconversion by alkaline phosphatase (of 2′-phosphate fraction) and cyclophosphate formation (from 2′-phosphate fraction) by carbodiimide.

Ligation in the p-RNA system is *regioselective*: no indication has been found in the experiments referred to above that ligation would proceed by any of the formally possible regioisomeric pathways leading either to a (4′ → 3′)-phosphodiester junction (isomeric opening of the cyclophosphate group by the 4′-OH function), or one of the junctions 3′ → 2′ or 3′ → 3′ which would result, if the 3′-OH instead of the 4′-OH group acted as the nucleophile. In our view, the identity of the product properties observed (UV melting curves and CD spectra of product-template duplexes, HPLC of product single strands after dephosphorylation) with those of the corresponding authentic products precludes the possibility that the ligation products could contain a phosphodiester junction other than 4′ → 2′ at the site of ligation. Consideration of models [1] of p-RNA ligation complexes predict the regioselectivity: as a consequence of the strong inclination between backbone and base-pair axes, the equatorial 4′-OH group seems perfectly positioned for a co-linear backside substitution [20] at the cyclophosphate group with the 3′-O-atom (and not the 2′-O-atom) as the leaving group (*Fig. 6*), whereas the adjacent (axial) 3′-OH group would seem to have no chance of reaching the cyclophosphate substitution center. Correspondingly, duplex models that are given a (4′ → 3′)-, (3′ → 2′)- or (3′ → 3′)-phosphodiester junction at the ligation site are markedly strained.

The conspicuous *absence* of ligation in the *absence* of templates in the experiments described above strongly suggests that the template-induced ligation of tetramer-2′,3′-cyclophosphates, under the conditions used, will be *sequence-selective*. This expectation is corroborated by observations made in, so far, two control experiments: *no* ligation (but partial cyclophosphate hydrolysis) was observed by HPLC, when, under otherwise identical conditions, the tetramer 4′-GCCC-2′,3′-cp was incubated with the 'wrong' template sequence 4′-GCCCGCCC-2′, and 4′-GGGC-2′,3′-cp with 4′-GGGCGGGC-2′, respectively (4 mismatches in product octamers).

So far, we have carried out one experiment to test our expectation that template-controlled ligation in the p-RNA series will be *enantioselective* and not inhibited by the presence of enantiomers of the ligation components. The ligation of D-p-Ribo(G_4)-2′-p mediated by the template D-p-Ribo(C_8) under the conditions of the carbodiimide experiments referred to above (*Fig. 5*) is observed to proceed equally well in the *presence* (of equimolar amounts), as it does in the *absence* of the enantiomeric building block L-p-Ribo(G_4)-2′-p.

Discussion. – Of the two versions of template-controlled ligation described above, the one *via* isomerization of 2′,3′-cyclophosphates is clearly the more interesting. There are two factors (only) that limit the ligation yield in this system under the reaction conditions used: the hydrolysis of the cyclophosphate group of the tetramer, occurring competitively with ligation and in comparable rate, and the consumption of the template due to non-dissociation, under the reaction conditions used, of the duplexes formed between ligation product and template. The first of these is expected (and, in preliminary experiments, observed) to be largely compensated by an excess of tetramer cyclophosphate, taking advantage of the remarkable fact that there are no tetramer-consuming processes occurring under the conditions used, other than ligation and cyclophosphate hydrolysis. Both the rate of ligation and of cyclophosphate hydrolysis depend strongly on the Mg^{2+} concentration. Preliminary tests beyond those described above indicate that lowering the Mg^{2+} concentration even further (under otherwise identical conditions) slows down further the ligation *rate*, but markedly increases the ligation *yield* (with respect to template consumption) which, therefore, can be expected to become the higher, the slower the reaction is made to proceed (by manipulating the Mg^{2+} concentration). It should be noted that in the aetiological context in which this work is carried out, a slow rate of ligation is not necessarily to be considered a shortcoming, provided that competing and consuming side reactions proceed even more slowly. Yet, the observations made so far stipulate a search for selective catalysis of the ligation reaction as well as for selective inhibition of cyclophosphate hydrolysis, this primarily in view of the challenge to design conditions that will promote the ligation processes from their present status of template-directed replicative steps to the level of autocatalytic self-replicative growth.

Thanks to the pioneering studies of *Orgel* and coworkers [21–25], much is known today about nonenzymatic template-directed oligonucleotide synthesis in the RNA series. In exploring mainly the most demanding of the conceivable versions of template-controlled sequence copying, namely, condensation of appropriately activated *mono*-nucleotide units, *Orgel* and coworkers observed that replicative copying of RNA sequences, while working in the pyrimidine → purine direction, can be seriously hampered in the purine → pyrimidine direction by self-deactivation of G-rich template sequences through self-pairing of guanine. *The absence of template self-deactivation in the p-RNA series, as deduceable from, and partially shown in, the present investigation, is one of the remarkable and important features by which the pyranosyl- and furanosyl-RNA isomers differ from each other chemically.*

Internucleotide bond formation by cyclophosphate → phosphodiester isomerization would appear to offer the simplest *as well as a potentially natural* pathway for nonenzymatic assembly of oligonucleotide chains. This possibility had also been examined by *Orgel* and coworkers [26] in their early attempts in the RNA series to polymerize adenosine-2′,3′-cyclophosphate on a poly(U) (triplex) template, with limited success, however. Unfavorable hydrolysis/coupling ratios and almost exclusively (2′ → 5′)-internucleotide junctions in the *di*nucleotide(s) formed were observed (see also the work of *Usher* [27] and *Usher* and *McHale* [28]). Our observations in the p-RNA series seem to support our expectation that the driving force of ligation through cyclophosphate isomerization has to come in important measure from the energetics of base pairing and base stacking in the ligand-template and product-template complexes.

The strategy of template-directed ligation of short oligonucleotides (instead of mononucleotides) as building blocks [29] [30] has recently provided the basis for experimental demonstrations of autocatalytic, informational replication [31][2]) of modified oligonucleotides in the DNA series by *v. Kiedrowski et al.* [34–36] (see also the reports of *Zielinski* and *Orgel* [37], and of *Li* and *Nicolaou* [38], and, in this context, also [39] [13]). In view of the widely propounded concept of an 'RNA world' [40] [41] [31], an extension of these studies from the DNA series to the chemically demanding and aetiologically more relevant RNA series is eagerly awaited.

Comparison of the chemical properties so far observed for p-RNA and RNA oligonucleotides elicits a comprehensive experimental inquiry into p-RNA's potential for nonenzymatic self-replication as well as for constitutional self-assembly in homochiral form. Such studies are part of a chemical aetiology of the RNA structure, the quest for a chemical rationalization of Nature's evolution of the structure of nucleic acids [12] [42].

This work was supported by *Ciba-Geigy AG*, Basel, *Firmenich & Co.*, Geneva, and the Schulleitung of the ETH-Zürich. Mass spectra were determined in the laboratories of *Ciba AG*, Basel (Dr. *H. Moser*, Dr. *U. Pieles*, *T. Schneeberger*). We thank Drs. *H. Moser* and *U. Pieles* for a generous sample of modified solid support for the synthesis of 2′-phosphates, Dr. *P. Lubini* (ETH) for providing the drawings of *Fig. 3*, and Prof. *C. Wintner*, Haverford College, USA, for his help.

REFERENCES

[1] S. Pitsch, S. Wendeborn, B. Jaun, A. Eschenmoser, *Helv. Chim. Acta* **1993,** *76,* 2161.
[2] S. Pitsch, A. Eschenmoser, B. Gedulin, S. Hui, G. Arrhenius, *Origins Life Evol. Biosphere* **1995,** *25,* 297.
[3] Y. Hayakawa, S. Wakabayashi, H. Kato, R. Noyori, *J. Am. Chem. Soc.* **1990,** *112,* 1691.
[4] Y. H. Pan, R. K. Robins, L. B. Townsend, *J. Heterocycl. Chem.* **1967,** *4,* 246.
[5] H. Vorbrüggen, B. Bennua, *Chem. Ber.* **1981,** *114,* 1279, and ref. cit. therein.
[6] K. Groebke, 'Über Purin-Purin-Paarungen bei Hexopyranose-Nukleinsäuren', thesis No. 10149, ETH-Zürich, 1993.
[7] P. Garner, J. U. Yoo, R. Sarabu, *Tetrahedron* **1992,** *48,* 4259.
[8] Y. Hayakawa, M. Hirose, R. Noyori, *J. Org. Chem.* **1993,** *58,* 5551.
[9] M. Böhringer, H.-J. Roth, J. Hunziker, M. Göbel, Ravichandran Krishnan, A. Giger, B. Schweizer, J. Schreiber, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1992,** *75,* 1416.
[10] U. Pieles, W. Zürcher, M. Schär, H. E. Moser, *Nucleic Acids Res.* **1993,** *21,* 3191.
[11] J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, Ravichandran Krishnan, B. Jaun, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993,** *76,* 259.
[12] A. Eschenmoser, 'Toward a Chemical Etiology of the Natural Nucleic Acids' Structure', in Proc. Robert A. Welch Found. Conf. Chem. Res. XXXVII. '40 Years of the DNA Double Helix', Robert A. Welch Foundation, Houston TX, 1993, p. 201.
[13] A. Eschenmoser, *Nova Acta Leopold., Neue Folge* **1992,** *67,* Nr. 281, 201; English version: A. Eschenmoser, E. Loewenthal, *Chem. Soc. Rev.* **1992,** *21,* 1.
[14] P. N. Borer, B. Dengler, I. Tinoco, Jr., O. C. Uhlenbeck, *J. Mol. Biol.* **1974,** *86,* 843.
[15] L. A. Marky, K. J. Breslauer, *Biopolymers* **1987,** *26,* 1601.
[16] A. Eschenmoser, *Pure Appl. Chem.* **1993,** *65,* 1179.
[17] P. Job, *Ann. Chim. (Paris), 10ème série* **1928,** *9,* 113.
[18] V. A. Efimov, A. A. Buryakova, S. V. Reverdatto, O. G. Chakhmakhcheva, Yu. A. Ovchinnikov, *Nucleic Acids Res.* **1983,** *11,* 8369.
[19] R. Naylor, P. T. Gilham, *Biochemistry* **1966,** *5,* 2722.
[20] F. H. Westheimer, *Acc. Chem. Res.* **1968,** *1,* 70.

---

[2]) For examples of *non-informational* molecular replication, see the recent work of *Rebek* and coworkers [32], and for synthetic 'replicating vesicles', the work of *Luisi* and coworkers [33].

[21] J. Sulston, R. Lohrmann, L. E. Orgel, H. T. Miles, *Proc. Natl. Acad. Sci. U.S.A.* **1968,** *59,* 726.
[22] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974,** *7,* 368.
[23] T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1982,** *162,* 201.
[24] G. F. Joyce, L. E. Orgel, *J. Mol. Biol.* **1986,** *188,* 433.
[25] A. R. Hill Jr., L. E. Orgel, T. Wu, *Origins Life Evol. Biosphere* **1993,** *23,* 285.
[26] M. Renz, R. Lohrmann, L. E. Orgel, *Biochim. Biophys. Acta* **1971,** *240,* 463.
[27] D. A. Usher, *Nature (New Biology)* **1972,** *235,* 207.
[28] D. A. Usher, A. H. McHale, *Proc. Natl. Acad. Sci. U.S.A.* **1976,** *73,* 1149.
[29] J. Ninio, L. E. Orgel, *J. Mol. Evol.* **1978,** *12,* 91.
[30] R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1979,** *14,* 243.
[31] L. E. Orgel, *Nature (London)* **1992,** *358,* 203.
[32] E. A. Wintner, M. M. Conn, J. Rebek, Jr., *J. Am. Chem. Soc.* **1994,** *116,* 8877, and ref. cit. therein.
[33] P. A. Bachmann, P. L. Luigi, J. Lang, *Nature (London)* **1992,** *357,* 57.
[34] G. v. Kiedrowski, *Angew. Chem. Int. Ed.* **1986,** *25,* 932.
[35] G. v. Kiedrowski, B. Wlotzka, J. Helbing, *Angew. Chem. Int. Ed.* **1989,** *28,* 1235.
[36] D. Sievers, G. v. Kiedrowski, *Nature (London)* **1994,** *369,* 221.
[37] W. S. Zielinski, L. E. Orgel, *Nature (London)* **1987,** *327,* 346.
[38] T. Li, K. C. Nicolaou, *Nature (London)* **1994,** *369,* 218.
[39] K. J. Luebke, P. B. Dervan, *J. Am. Chem. Soc.* **1991,** *113,* 7447.
[40] G. F. Joyce, L. E. Orgel, in R. F. Gesteland, J. F. Atkins, Eds., 'The RNA World', Cold Spring Harbor Laboratory Press, 1993, p. 1.
[41] J. A. Doudna, S. Couture, J. W. Szostak, *Science* **1991,** *251,* 1605.
[42] A. Eschenmoser, *Origins Life Evol. Biosphere* **1994,** *24,* 389.