Bioorganic Chemistry **30,** 62–80 (2002)
doi:10.1006/bioo.2001.1232, available online at http://www.idealibrary.com on IDEAL®

REVIEW

Phosphates, DNA, and the Search for Nonterrean Life: A Second Generation Model for Genetic Molecules[1]

Steven A. Benner[2] and Daniel Hutter

Departments of Chemistry and Anatomy and Cell Biology, University of Florida, P.O. Box 117200, Gainesville, Florida 32611-7200

Received August 24, 2001

Phosphate groups are found and used widely in biological chemistry. We have asked whether phosphate groups are likely to be important to the functioning of genetic molecules, including DNA and RNA. From observations made on synthetic analogs of DNA and RNA where the phosphates are replaced by nonanionic linking groups, we infer a set of rules that highlight the importance of the phosphodiester backbone for the proper functioning of DNA as a genetic molecule. The polyanionic backbone appears to give DNA the capability of replication following simple rules, and evolving. The polyanionic nature of the backbone appears to be critical to prevent the single strands from folding, permitting them to act as templates, guiding the interaction between two strands to form a duplex in a way that permits simple rules to guide the molecular recognition event, and buffering the sensitivity of its physicochemical properties to changes in sequence. We argue that the feature of a polyelectrolyte (polyanion or polycation) may be required for a "self-sustaining chemical system capable of Darwinian evolution." The polyelectrolyte structure therefore may be a universal signature of life, regardless of its genesis, and unique to living forms as well.

Key Words: phosphate-backbone; polyelectrolyte; sulfone; PNA; genetic molecule; nonterrean life.

INTRODUCTION

Ever since the emergence of the Watson–Crick model for DNA (*1*,*2*), chemists and biologists alike have underestimated the complexity of the molecular system that stands behind genetics and inheritance. This was almost certainly due, in part, to the simplicity, elegance, and utility of the model itself. The "first generation model" explicates genetic inheritance using the structure of DNA and just three very simple structural concepts: nucleobase stacking, size complementarity, and hydrogen bonding

[1] This paper is dedicated to Professor Frank H. Westheimer on the occasion of his 90th birthday.

[2] To whom correspondence and reprint requests should be addressed. Fax: (352) 392-7918. E-mail: benner@chem.ufl.edu.

0045-2068/02 $35.00

complementarity. Duplexes are stabilized by base stacking. The specificity of base pairing arises from size and hydrogen bonding complementarity between bases.

The well known Watson–Crick base pairing rules (A pairs with T, G pairs with C) follow from these premises. These rules permit molecular biologists to design molecular recognition systems without much background in organic chemistry. This power eludes more sophisticated chemists attempting to achieve molecular recognition in nonnucleic acid systems.

The simplicity, elegance, and utility of the first generation model has, however, been almost beguiling to scientists, who have often expressed surprise as complexity in nucleic acids, not captured by this model, has emerged from experimental work. When Rich and his coworkers announced the existence of Z-DNA (*3*), for example, the structural biology community was surprised, even though most polymers (indeed, most organic molecules) show a diversity of conformational states. Considerable research activity focuses on "bending DNA" (*4*), in part because it is surprising (under the model) that DNA might bend (even though other biopolymers do). The discovery of catalysis by RNA introns (*5*) and the RNA component of RNase P (*6*) was regarded as surprising, even though "catalysis" by RNA itself had been observed earlier as a normal property of functionalized polymers (*7*), and was suggested for elements of the translation machinery (*8,9*). The ability of RNA to provide general acid–base catalysis was viewed as noteworthy (*10*), even though this ability follows axiomatically from the chemical structure of RNA.

With DNA, however, the simple elegance of the first generation model created in the minds of chemists and biologists alike the expectation that derivatives and analogs of DNA would also behave in a rule-based fashion. Perhaps the most striking example of how the first generation model failed to guide nucleic acid chemistry comes from the "antisense" industry, which rose, flourished, and nearly vanished over a period of only a decade (*11–13*). The antisense concept was simple and entirely reasonable given the first generation model. Many diseases arise through the presence of unwanted DNA that is expressed as mRNA and encodes unwanted proteins. Many drugs work by binding the unwanted target (generally an unwanted protein). What could be simpler than targeting the unwanted mRNA that preceded the unwanted protein? Often, the sequence of the mRNA is known, especially for infectious diseases. The Watson–Crick model suggested simple design principles to create a molecule that binds to the mRNA (A pairs with T (or U), G pairs with C).

The only challenge to implementing the antisense strategy seemed to be the backbone. Nucleic acids have a repeating phosphodiester linkage in their backbone, which join adjacent nucleosides in the chain. The phosphodiester linkages are the targets of nucleases that degrade DNA. Likewise, and perhaps more severely, the repeating phosphodiesters make DNA a polyanion. As reviewed by Westheimer more than a decade ago in his now-famous paper on "Why Nature Chose Phosphates" (*14*), anions (and anionic charges carried by phosphates in particular) prevent molecules from crossing cell membranes (*15*).

These two features, charge and sensitivity to enzymes, of DNA made it unlikely that one could ever use DNA directly as a drug. Most of it would be degraded, and the part that was not would not make it across cell membranes to reach its target.

Fortunately, the first generation model for DNA structure did not propose any

particular role for the polyanionic backbone in the molecular recognition event, although they clearly understood the importance of placing the phosphates in the duplex structure in a way that permitted them to interact with water. The phosphate-sugar backbone otherwise served simply as a scaffold. Perhaps other scaffolds would do? Indeed, Westheimer remarked that perhaps amides might serve as well.

For this reason, we (and others) reasoned, why not make analogs of DNA where the backbone phosphate was replaced? In particular, the replacement would be non-ionic, and stable to enzymatic degradation. The expectation, based on the first generation model, was that virtually any backbone-modified DNA analog would retain the molecular recognition properties of DNA itself. A sample of analogs of DNA where the phosphodiesters are replaced by a neutral linker, following this logic, is shown in Fig. 1 (*16*).

Upon closer consideration from a physical organic perspective, the structure for DNA is not as inevitable as it might seem from the first generation model, however. Indeed, from a physical organic perspective, it does not correspond closely to how a chemist might set out to design two molecules that will bind with great selectivity to each other. Consider just three features of the DNA structure, and contrast them with how chemists might design molecular recognition systems:

(a) DNA strands are floppy. Rigid molecules would seem better for molecular recognition.

(b) DNA uses hydrogen bonding to achieve specificity, according to the model. In water, hydrogen bonds are abundant. It seems curious in water to use hydrogen bonding as the key to molecular recognition.

(c) Receptor and ligands are both anions. In the design of a molecular recognition

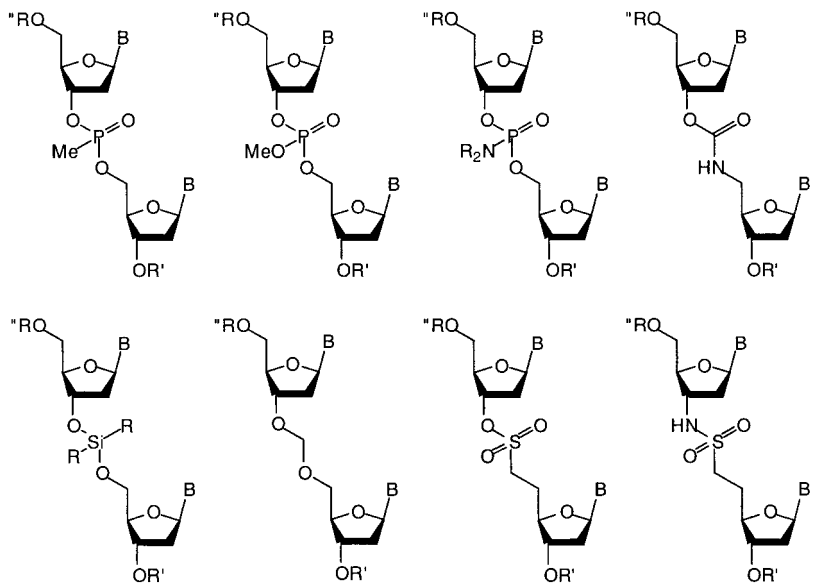

FIG. 1. A selection of analogs of DNA where the phosphate is replaced by an uncharged analog. *R* is alkyl or aryl; *R′* and *R″* are nucleic acid or hydrogen.

system, charge complementarity might be sought. Few chemists would design a polyanion to bind to another polyanion.

The paradigm of physical organic chemistry, which involves systematic synthesis of analogs of DNA and examining their properties, could be used to test the Watson–Crick model. After all, the Watson–Crick model had solved the problem. The solution was elegant. What was there to test?

Perhaps for this reason, prior to 1985, only a few chemists were applying their synthetic skills to DNA. For example, as early as 1970, Pitha *et al.* had attached vinyl groups to uracil and created polymers (*17*). This predated both synthesis and direct sequencing of DNA. In the 1970s, Eckstein had prepared phosphorothioate analogs of DNA (*18*), although with the goal more to develop an understanding of the mechanism of enzymatic reactions involving phosphorus as the electrophile than to test the Watson–Crick model itself.

Starting in Zurich in 1985, we began a research program to analyze the structure of DNA in this light. Strong programs were also developed elsewhere, notably within corporate laboratories at Central Research at Ciba-Geigy, Gilead, and at Isis (*12,16*). Each program was designed to modify systematically the structure of the nucleobases, the sugars, and the phosphates. Each has advanced our knowledge of molecular recognition in DNA. Several have had technological impact as well. For example, our work with the heterocyclic nucleobases has led to diagnostic products that are detecting very low levels of DNA in biological samples, DNA containing a single nucleotide change in a patient, and real time quantitation of mRNA (*19*). In the form of an "account," let us review some of what we have learned from these efforts, especially about the role of phosphates in the molecular recognition event.

A SMALL STEP FROM THE NATURAL PHOSPHODIESTER BACKBONE

Our first effort involved replacing the phosphodiester linkers in DNA by linkers that would be close in structure to the phosphodiesters, but not have their negative charges. Ideally, we wished to "add" a proton to each phosphorus nucleus (in a sense) to "make" the sulfate diester, a unit that would be isoelectronic to the phosphate diester. Simple chemical considerations ruled out the sulfate diester linkage directly, however. The primary 5′-carbon in the sulfate diester would almost certainly be too electrophilic to survive in water. We felt that a sulfonate would be too reactive as well, although heroic efforts by Widlanski and his coworkers (*20*) provided DNA analogs incorporating this structure; their reactivity was used to probe the active site of a polymerase (*21*). The sulfonamide was also considered as a structure that would lack the undesirable reactivity of a sulfonate (*22*).

We chose to address the reactivity problem by making the dimethylenesulfone linker (Fig. 2). A dimethylenesulfone is isoelectronic to a phosphodiester. It is also nonionic, stable to alkaline degradation, and not stereogenic (that is, it does not create diastereoisomers). We were also attracted by the fact that sulfones are intrinsically soluble in water and that Hanahan had shown that dimethylsulfoxide increased the frequency with which transforming DNA entered bacterial cells (*23*).

The preparation of these compounds required solutions to problems encountered generally in large scale organic synthesis. We needed to make four building blocks, in sufficient quantities to permit each to be the starting point of a multistep synthesis

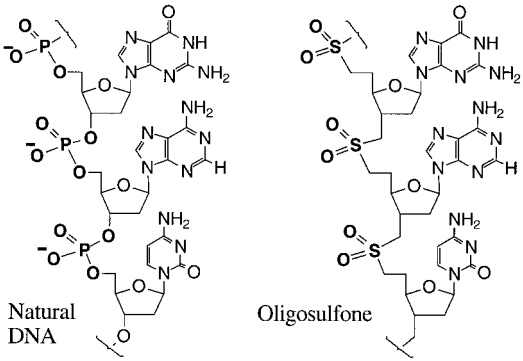

FIG. 2. A dimethylenesulfone-linked analog of DNA.

of an oligomer. All of the oligomers were prepared in solution (not on solid phase) so that they could be fully characterized. Many of the oligomers proved to be challenging analytical targets (see below). The details of synthetic efforts are given elsewhere (*24*,*25*).

The synthetic effort did not go unrewarded. One of the first sulfone analogs of RNA to be made was $G_{SO_2}C$, the sulfone analog of the RNA dinucleotide G_{PO_2}-C, which had been crystallized and studied by Rich and his coworkers (*26*,*27*). Both molecules were self-complementary in the Watson–Crick sense and might be expected to form a duplex. Indeed, the $G_{SO_2}C$ structure (Fig. 3) did so (*28*). The duplex it formed was remarkably similar to the duplex formed by the natural RNA dimer G_{PO_2}-C. In both molecules, the crystallographic unit cell contained two molecules in the form of a right-handed double helix with antiparallel orientation of the strands. In both molecules, the duplexes were joined via canonical Watson–Crick base pairs. In both, the two strands were related via a crystallographic twofold rotation axis. Conformations around the glycosyl bonds (*anti*), ribose puckers (C3′-*endo*-type), and all backbone torsion angles (Tables 1 and 2; Fig. 3) of the $G_{SO_2}C$ analog fell into the same ranges as found in natural RNA duplexes. Further, the overall dimensions of the two duplexes (e.g., relative S(P)⋯S(P) and C1′⋯C1′ distances, Tables 1 and 2) were the same to within 0.3 Å (Fig. 4).

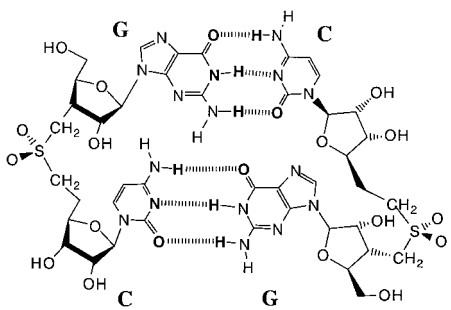

FIG. 3. Schematic of the duplex formed by $G_{SO_2}C$ in the crystal.

TABLE 1

Comparison of Backbone[a] and Glycosyl (χ) Torsion Angles as well as Pseudorotation Phase Angles (P) (in deg) for [$G_{SO_2}C$] (Top) and [$G_{PO_2^-}C$] (Bottom)

	α	β	γ	δ	ε	ζ	χ	P
G(1)			176	93	−140	−66	−178	12
C(2)	−58	−179	44	86			−170	29
G(1)			49	82	−150	−66	−157	7
C(2)	−75	−176	51	75			−152	17

[a] The angular notations for backbone torsions are O5′(C6′)$\underline{\beta}$C5′$\underline{\gamma}$C4′$\underline{\delta}$C3′$\underline{\varepsilon}$O3′ (C3″)$\underline{\zeta}$P(S)$\underline{\alpha}$O5′ (C6′).

Some differences were evident between $G_{SO_2}C$ and $G_{PO_2^-}C$ upon close examination, of course. In the $[G_{SO_2}C]_2$ duplex, for example, the glycosidic torsion angles were larger by 21° and 18° for the G and C residues than in the $[G_{PO_2^-}C]_2$ duplex (Table 1). The torsion angles α, δ, and ε were different by at least 10° in the sulfone duplex. The helical parameters of the duplex were modestly different in the sulfone duplex (Table 2). The rise in the natural duplex was slightly smaller than in the sulfone. The bases were stacked parallel in the sulfone with a spacing of 3.40 Å, instead of the 3.68 Å spacing with a roll of 8° observed in the stacking of bases in the natural RNA duplex. The $[G_{SO_2}C]_2$ duplex was unwound to give ca. 17 residues per helical turn, rather than the 11 residues in standard A-type RNA.

These differences in the overall structures of the RNA and sulfone duplexes can all be explained, however, without invoking the loss of charge. For example, the sulfone places the C6′ methylene group at the site occupied by O5′ in the natural duplex. This methylene group is larger than the oxygen, and may interact sterically with the C6 C-H group of cytosine. This interaction appeared to result in an increased C6′–C5′–C4′ angle, and a slight change in the cytosine ribose conformation (compare P values of residues C(2) of the two duplexes in Table 1). Torsion δ for the C residues appeared to be adjusted to reflect this change. Likewise, the unwinding in the $[G_{SO_2}C]_2$ double helix appeared to result predominantly from the fact that the S–C6′–C5′ bond angle is 111°, instead of the 120° found in RNA for the corresponding P–O5′–C5′ angle.

Thus, the absence of the negative charge in the sulfone-linked RNA analog appeared

TABLE 2

Comparison of Helical Parameters[a] for $[G_{SO_2}C]_2$ and $[G_{PO_2^-}C]_2$ (Angles in Degrees and Distances in Å)

	Rise	Twist	Inclination	Slide	Roll	S⋯S P⋯P	C1′⋯C1′	Buckle	Prop. twist
$G_{SO_2}C$	2.92	20.8	9.4	−3.2	0.7	18.0	10.6	6.0	0.8
GpC	2.59	34.7	28.0	−1.3	7.9	17.7	10.7	5.8	2.7

[a] Atoms selected to determine the helical operator were C1′, N1(9), C2′, O4′.

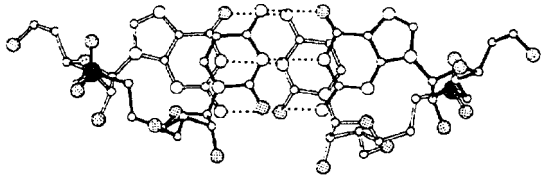

FIG. 4. The crystal structure of $G_{SO_2}C$ (*28*).

to have remarkably little impact on the overall duplex structure. Only in solvation did the change in charge seem to have an impact. Thus, the sulfone unit cell carried 8 water molecules and 1 methanol per strand; the natural duplex carried 9 water molecules. The major and minor groove base functional groups were extensively hydrated in both the sulfone analog and the natural RNA, but the solvation of the sulfone group itself was different. Instead of forming contacts with a sodium ion and three water molecules as seen with the phosphate, the sulfone group formed contacts with one water molecule, a cytosine NH_2 group from a neighboring duplex, and the CH_3 group of a methanol molecule.

The ability of the $G_{SO_2}C$ sulfone dimer to form duplexes suggested that a negative charge is not essential for Watson–Crick duplex formation. But the next round of experiments, with ($A_{SO_2}U$), suggested that the tendency to form Watson–Crick pairing might be weaker in a sulfone analog than in the corresponding phosphodiester. Here, the self-complementary $A_{SO_2}U$ duplex would be joined by only four hydrogen bonds, instead of the six that join the $G_{SO_2}C$ sulfone duplex. The natural $A_{PO_2^-}U$ dimer did form a duplex (*29*). In contrast, $A_{SO_2}U$ crystallized to give a single stranded structure (*30*), with Watson–Crick interactions not dominating the structure. The two structures available for the methylphosphonate linked deoxynucleotide dimers d($G_{POCH_3}C$) and d($A_{POCH_3}T$) found strikingly parallel results (*31,32*). The first formed a duplex joined by six hydrogen bonds involving two Watson–Crick pairs. The second formed a single strand, not a duplex joined by four hydrogen bonds involving two Watson–Crick pairs.

THE BREAKDOWN OF WATSON–CRICK BASE PAIRING IN LONGER UNCHARGED ANALOGS

These results encouraged us to prepare longer oligosulfones. Here, we encountered stronger evidence that the loss of the phosphodiester charge might have consequences. The first evidence came from tetrameric oligosulfone analogs of RNA. For example, the tetramer R-$T_{SO_2}T_{SO_2}T_{SO_2}$C-SR′ (where R and R′ were various protecting groups) formed dimers in tetrahydrofuran. The dimer was disrupted by adding small amounts of methanol, suggesting that hydrogen bonding might be involved. This dimer could not, however, be joined by canonical Watson–Crick pairing, as the system contained no A to pair with T and no G to pair with C.

Still longer sulfone-linked oligonucleotide analogs departed further from expectations formed by analogy to DNA. Particularly remarkable was the sulfone $A_{SO_2}U_{SO_2}G_{SO_2}G_{SO_2}U_{SO_2}C_{SO_2}A_{SO_2}U$. The molecule displayed a remarkable thermal denaturation curve, melting at ca. 80°C. Upon melting, a large hyperchromicity was observed (>200%; 25% is typical for melting of a DNA duplex). The sequence

was not, in the Watson–Crick sense, self-complementary. Further, the denaturation appeared to be unimolecular. Thus, we concluded that this oligosulfone folded to a rather stable conformation, indeed one of the most stable single stranded "RNA" structures known.

The results of these and many other experiments suggested that each oligosulfone has its own unique properties and reactivity. Different sulfones differing (in some cases) by only one nucleobase displayed different levels of solubility, aggregation behavior, folding, and chemical reactivity. Thus, the properties of tetranucleoside sulfone analogs and octanucleoside sulfone analogs was influenced dramatically by adding a single charge to one end of the molecule, or replacing a single nucleobase (*33*).

Further, Watson–Crick pairing seemed to be lost. For example, Huang found evidence for base pairing between d($U_{SO_2}U_{SO_2}U_{SO_2}U_{SO_2}U_{SO_2}U_{SO_2}U_{SO_2}U_{SO_3^-}$) and its DNA complement (*33*). Eschgfaeller found, however, no evidence for the binding of an octamer differing only at one position, and replacing T for U: $T_{SO_2}T_{SO_2}T_{SO_2}C_{SO_2}T_{SO_2}T_{SO_2}T_{SO_2}T_{SO_3^-}$ to its complementary DNA sequence. On the other hand, the second molecule displayed good bioavailability in a mouse model (*34*,*35*).

These are not, of course, behaviors expected from DNA itself. The physical properties of two DNA molecules of the same length are usually nearly identical even when their sequences are quite different. Different DNA sequences of the same length generally move at (nearly) the same position in an electrophoresis experiment. Pairing is not, in general, dramatically altered by changing a single nucleobase. Virtually all DNA sequences are soluble in water, and insoluble in ethanol. Indeed, the constancy in the physical behavior of DNA, despite large change in sequence, is the key to the laboratory manipulation of DNA.

Replacing the phosphodiester groups by sulfone linkers generated a polymer that folded, had different properties, and whose properties changed dramatically upon small changes in sequence. Biological chemists are, of course, familiar with biopolymers that fold, have different properties, and whose properties can change dramatically by substitution of a single building block. These properties are all characteristic of proteins. Thus, we might view the results obtained with nonionic nucleotide analogs as saying that by replacing the anionic phosphodiester groups by uncharged sulfone groups, we have transformed DNA into a kind of "protein."

PHOSPHATES: THE KEY TO MOLECULAR RECOGNITION IN NUCLEIC ACIDS?

These results suggest three hypotheses that propose roles for the phosphodiester linkages, and how they might be important to the molecular recognition that characterizes DNA. They suggest three specific roles for the phosphodiester anions, in addition to the (perhaps obvious) role of making the polymer soluble in water. They all stress the universality of a polyelectrolyte as part of any molecule that needs to perform a genetic role.

Phosphates Force Interactions Between Strands as Far From the Backbone as Possible

As natural as the Watson–Crick hydrogen bonding pattern is, it is not necessarily the preferred one. Upon cocrystallization of equimolar amounts of adenine and thymine,

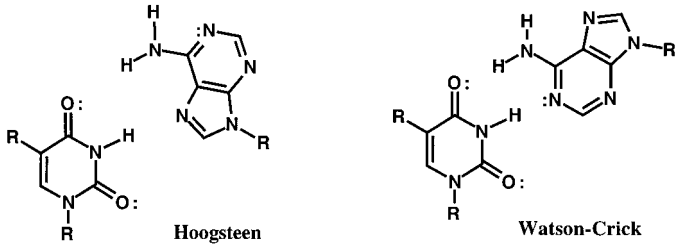

FIG. 5. A representation of the Hoogsteen (left) and Watson–Crick (right) base pairs.

Hoogsteen was surprised to discover that the adenine-thymine mixture paired through hydrogen bonding to N-7 of the adenine, not N-1 (as is seen in the standard DNA duplex structure) (Fig. 5).

From first principles, one expects a richly functionalized molecule to interact with other molecules through any of its functional groups. Proteins interacting with proteins offer many examples suitable for detailed study. What is also expected, however, is that two DNA strands, because of their polyanionic nature, cannot interact with each other from any point on their structure. In particular, the complex between two DNA strands is expected to place the polyanionic backbones as far from each other as possible. Conversely, the two strands should contact each other as far from the phosphate backbone as possible. This is, of course, the Watson–Crick "edge" of the heterocycles (Fig. 6).

The polyanionic nature of the backbone, under this hypothesis, is key to the molecular recognition between two DNA strands. It constrains the interaction of two strands to the regions of the molecule where those interactions are desired, and where those interactions can be described by simple rules.

Consistent with this hypothesis, Steinbeck and Richert (*36*) examined the solution structure of the dimethylenesulfone-linked $U_{SO_2}C$ dinucleotide analog using two-dimensional NMR and restrained molecular dynamics. In $CDCl_3$, the sulfone formed

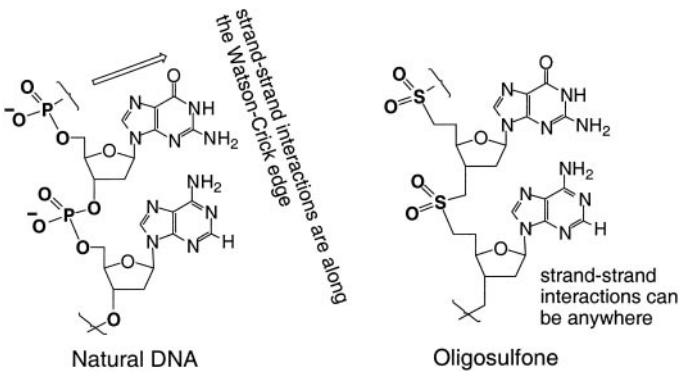

FIG. 6. The repulsion of the backbone charges forces strand–strand interactions in the polyanionic DNA to occur as far from the backbone as possible, along the Watson–Crick "edge" of the nucleobases.

a parallel duplex with a single U:U base pair and roughly antiparallel orientation of the two ribose rings within each strand. A hydrogen bonding network stabilizing this duplex included a two-pronged hydrogen bonding between the uridines, two hydrogen bonds from the ribose hydroxyls of one strand to O2 of the cytosine bases of the opposite strand, and intrastrand hydrogen bonds from the 2′ hydroxyls of the 5′-terminal residues to hydroxyls of the 3′-terminal residues. The melting point of the duplex in tetrachloroethane was determined via NMR to be 91°C (ΔS of -47 cal K^{-1} mol^{-1} and ΔH of -22 kcal mol^{-1}).

Phosphates Keep DNA Single Strands from Folding

One of the key features of DNA is its ability to act as a template. This requires that it *not* fold, but rather that it extend itself. It is well known that polymers, in their "random coil" state, extend themselves more when they are polyanionic. For example, statistical thermodynamics theory of polymers holds that a unidimensional polymer in a random coil will occupy a volume that scales with the length of the polymer raised to the 1.5 power.

This power increases once one considers the excluded volume of the polymer, which reflects the fact that no real polymer is truly unidimensional. For a neutral polymer, the volume scales with the length of the polymer to (typically) the 1.7 power. This power scaling changes dramatically when the polymer is polyanionic, however. For polyglutamate at a pH where it is a polyanion, the power is ca. 2.3 (*37*).

In simpler language, a flexible polyanion is more likely to extend than a flexible neutral polymer (Fig. 7). Conversely, to have a neutral polymer adopt an extended conformation, one needs to add rigidity. Alternatively, one can view this result as simply a coulombic obstacle to folding. It is difficult for a polyanion to form a compact structure. In contrast, in proteins, the repeating unit is a dipole. A repeating dipole easily forms a folded structure. Indeed, the standard secondary structural elements of a polypeptide (the alpha helix and beta sheet) are defined by the interactions between the repeating dipole units of the backbone. There is no comparable way to pack a repeating monopole. Hence, the polyanionic backbone encourages DNA to extend to be a template.

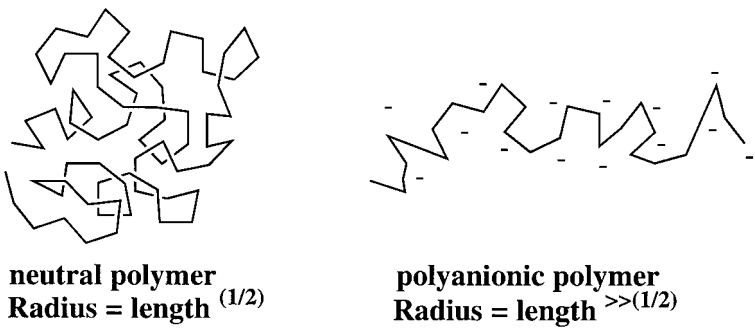

FIG. 7. A polyanionic polymers adopts a more extended conformation than a neutral biopolymer, facilitating its action as a template.

Phosphates Create the Physical Organic Chemistry Needed for a Molecule to Support Darwinian Evolution

Perhaps the most important hypothesis connecting the phosphodiester polyanion to its physiological role relates to the need of a genetic molecule to support Darwinian evolution. This places some unusual chemical demands on the molecular system.

Darwinian evolution requires that the molecule be replicable. Replicatability may be enhanced by enzymes, cofactors, and other adjuvants in an advanced biochemical organism. It presumably must also be an intrinsic property of the molecular system, however.

Many molecular systems make copies of themselves, including some in quite trivial ways. For example, a crystal of sodium chlorate ($NaClO_3$), if fragmented, will nucleate (or "seed") the formation of more crystals of $NaClO_3$. The surface of the crystal is a template for the self-assembly of freely wandering ions into organized matter. In this particular case, the system spontaneously generates homochirality, a property often regarded as unique to living systems (*38*). Either right or left handed prisms of $NaClO_3$ form from solutions of the achiral ions. If the solution is stirred, only right-handed or only left-handed crystals are formed (*39*). If the solution is *not* stirred, crystals of both enantiomorphic forms precipitate in equal amounts. This highly reproducible result is a consequence of a "self-reproduction." Stirring breaks apart the first crystal formed with random chirality, whose fragments then seed the formation of many "daughter" crystals with the same chirality.

Templated replication is widespread in chemistry, and many examples are now known in bioorganic chemistry as well. Rebek ((*40*), but see the challenge by Menger and coworkers over the validity of Rebek's interpretation (*41*) and Rebek's rebuttal (*42*)). Ghadiri (*43*) and von Kiedrowski (*44*) have shown how small molecules, peptides, and oligonucleotides undergo template-directed synthesis. It seems not to be excessively difficult to design such systems, and many are being examined in many laboratories.

Mutation is the second property that a molecular system must display to partake in Darwinian evolution. In some usage, "mutation" denotes any structural change. Thus, a molecular system of $NaClO_3$ might be said to "mutate" if the sodium is replaced by potassium to yield $KClO_3$. In practice, however, mutation cannot be so drastic if it is to be effectively coupled with natural selection. At the very least, mutation in a molecular system that can self-replicate must not change the physical properties of the molecules involved in a way that obviates mechanisms used for self-replication. The replacement of Na by K in the example above does.

To support a self-sustaining chemical system capable of undergoing Darwinian evolution, a biopolymer must be capable of changing its structure, and therefore changing its encoded information, without changing its bulk properties, in particular, properties essential for replication. This property of a molecular system lends itself to a convenient acronym: COSMIC-LOPER (capable of surviving modifications in constitution without loss of properties essential for replication) (*45*).

Very few organic molecular systems are COSMIC-LOPER, as is well known to synthetic organic chemists. Non-COSMIC behavior is commonly encountered in natural product synthesis. Synthetic efforts on natural products are, for example, often

preceded by exploratory work with a "model system," a simpler molecule that is more accessible than the natural product itself, but that is presumed to be representative of the kinds of problems that will be encountered when the natural product is synthesized. As often as not, reactivities of the model and the real natural product are sufficiently different that chemistry developed on the first is defeated by the second. A single methyl group difference can be sufficient to alter the physical properties of an organic molecular system.

In this respect, DNA is unusual. Changing the sequence of the molecule generally does not change the overall physical properties of the molecule, or its general reactivity. This is not the case with the sulfone analogs of DNA; in these analogs, simply changing a nucleobase can dramatically change the behavior of the molecule.

We can hypothesize that the repeating negative charge provided by the phosphodiester linkage is what confers COSMIC-LOPER behavior upon DNA and RNA. Here, the repeating monopole (charge) dominates the variable dipoles (the hydrogen bonding units; the heterocycles) in determining the physical behavior of the molecule. One can change the nucleobase sequence without changing the physical properties of the oligonucleotide.

Proteins, the "other" encoded biopolymer, are not COSMIC-LOPER, even in cases where they can direct template-based replication (*46*). The physical properties of proteins can change dramatically upon point mutation within the mutation space allowed by the 20 standard amino acids. The textbook case is sickle cell hemoglobin, where a single amino acid substitution creates a variant that undergoes partial precipitation.

Precipitation is a common property of peptides; a peptide of random sequence chosen from the 20 standard amino acid is most likely to be insoluble, a property also known to those who boil eggs (after heating to escape the metastable state where the proteins are soluble). Thus, it is not remarkable that sickle-cell hemoglobin precipitates; what is remarkable is that native hemoglobin is soluble at the nearly gram-per-milliliter concentrations found in red blood cells. One of the most characteristic features of the divergent evolution of the hemoglobin gene family is the extent to which substitution has been constrained to avoid insolubility.

The non-COSMIC properties of peptides are also encountered in peptide design. Focusing on peptides that form helix bundles, Johnsson *et al.* successfully designed a small (14 amino acid) polypeptide that formed nearly exclusively a four helix bundle (*47,48*). They then determined its structure in solution (*49*). The identical peptide was then prepared, but with an N-terminal acetyl group. The small change was sufficient to cause the peptide to aggregate. Still other point mutations caused the helix not to form. If solubility and/or helix formation are essential to the replicability of a peptide template, a large range of plausible mutations would destroy it, at least in this system.

A biopolymer with the properties of a peptide would have a difficult time supporting Darwinian evolution. Selection would operate too strongly at the level of the encoding molecule. Indeed, natural DNA and RNA are not entirely COSMIC-LOPER. Best known to the molecular biologist are RNA molecules that have G-rich sequences. These adopt tertiary structures around a G-quartet, and these structures often disrupt the templating ability of an RNA sequence (*50*); this property too came as a surprise.

This means that if an RNA molecule searches G-rich regions of sequence space, it runs the risk that it will lose certain properties essential for replication. In comparison with virtually every other class of organic molecule, however, nucleic acids are the most COSMIC-LOPER.

A UNIVERSAL STRUCTURAL FEATURE OF A GENETIC MOLECULE?

NASA has defined life to be "a self-sustaining chemical system capable of Darwinian evolution" (*51*). It has proven to be difficult to identify objective chemical markers that can be used experimentally to detect evidence of life on NASA planetary missions. Chirality can be a consequence of nonbiological processes. Adenine and several amino acids are generated under nonbiological conditions. NASA missions must be supported by a proposal for a chemical signature that is unique to life (not generated by nonliving processes) and universal to life (found in all life forms regardless of genesis).

A polyelectrolyte (polyanion or polycation) may well be a structural feature of genetic molecules regardless of their genesis. As Westheimer noted (*14*), phosphates are an easy way to get a polyelectrolyte; there are few other linking units that are as easy to create that confer the same property. It is conceivable, that the anionic or cationic charge need not be carried by the linking unit (*52*).

Polyelectrolytes are, in practice, easy to detect, as they adsorb on to polyelectrolytes of the opposite charge. This makes it easy to conceive of instruments that would detect them in planetary missions. A simple schematic is shown in Fig. 8.

POLYAMIDE-LINKED NUCLEIC ACID ANALOGS (PNAs): THE EXCEPTION THAT "PROVES" (TEST) THE RULE

Virtually all of the nonionic linkers that have been introduced into DNA over the past two decades have diminished rule-based molecular recognition when compared to that displayed by natural DNA itself. Many of these are summarized in Ref. (*16*). Each of these confirms the notion that a polyanion is a key to DNA-like molecular recognition in water.

Not all possible analogs have been examined, of course, and Westheimer himself wondered whether polyamide linkers might perform as well as phosphates in an encoding molecule (*14*). Four years later, an example of these came in the form of the "peptide nucleic acid," perhaps better named a "polyamide-linked nucleic acid analog," or PNA, invented by Nielsen, Egholm, and their collaborators (*53*). PNA is a DNA/RNA mimic in which the phosphate deoxyribose backbone has been replaced by uncharged *N*-(2-aminoethyl)glycine linkages with the nucleobases attached through methylene carbonyl linkages to the glycine amino group (Fig. 9).

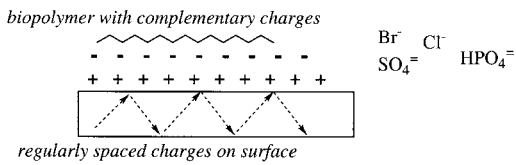

FIG. 8. A schematic of an instrument to detect a polyanion in the subsurface water of Mars or in the oceans of Europa.

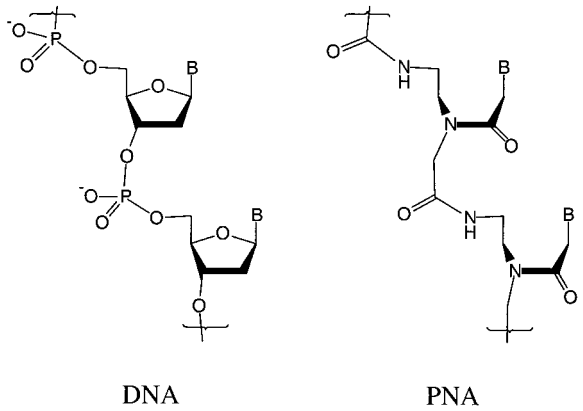

FIG. 9. The structure of the peptide nucleic acid (PNA) backbone compared to the natural backbone.

PNAs again display rule-based molecular recognition, at least in short oligomers (*54–56*). Many PNA:PNA duplexes and PNA:DNA(RNA) duplex hybrids are more stable than DNA:DNA or DNA:RNA duplexes, with a difference in melting temperature (ΔT_m) of approximately 1–1.5°C per base pair (*57*). The hybridization of PNAs to complementary sequences is characterized by good mismatch discrimination (*58*). PNAs also possess high rates of association for duplex DNA (*59*) and a remarkable propensity for the invasion of double-stranded structure. PNAs are not hydrolyzed by nuclease or proteases (*60*).

This phenomenon is so striking that PNAs are being examined closely for a variety of technological applications (54–56), including examining telomeres (*61*), inhibiting human telomerase (*62–64*), capturing nucleic acids (*65,66*), screening genetic mutations (*67, 68*), detecting specific sequences in unamplified DNA (*69*), and labeling plasmids with fluorophores (*70*). Less well documented, but also deserving of exploration, are nonionic morpholino analogs of DNA (*71*).

As for PNA, the range of its applications has encouraged Corey and coworkers to call PNA "one of the most successful designed macromolecules" (*72*). Orgel and his coworkers have done experiments to explore PNA as possibly the primordial DNA (*73*). Relevant to this review, PNAs may be the example that disproves the hypothesis requiring a polyelectrolyte in the universal genetic molecule.

We ourselves have examined PNA as a conjugate with DNA for its suitability as a primer for DNA polymerases (*74,75*). Other unpublished considerations have raised the following issues, however. Short sulfone analogs and short methylphosphonate analogs support Watson–Crick pairing (see above). Longer sulfone and longer methylphosphonate analogs do not. We may expect that different non-ionic backbones will sustain Watson–Crick rule-based molecular recognition up to different lengths, depending on the backbone, its interaction with the solvent, and its potential to interact with itself and the heterocycles that it carries. But we expect that *all* nonionic analogs must arrive at a length where they prefer to fold, aggregate, and precipitate rather than template.

Even though the PNA backbone supports Watson–Crick molecular recognition for

sequences of modest length, it does not seem to be able to do so for indefinitely long sequences. This recognition does not disappear with increasing length as rapidly as it disappears in oligosulfones. PNA derivatives as long as 20 nucleotides have been reported that continue to bind to complementary DNA using Watson–Crick rules (*58*). But their propensity to self-aggregate is recorded in the literature (*76*), as is self-structure that appears to interfere with PNA–DNA duplex formation (*77*). Longer PNA molecules suffer aggregation and other physical behaviors that interfere with their ability to recognize complementary DNA. PNAs also change their physical properties substantially (and largely unpredictably) with small changes in sequences (*78,79*), although adding charged appendages helps (*80*) (as it does with sulfones). Much of this is captured in the informal literature on PNAs. Perhaps the best place to find this information is on the web pages of suppliers who offer custom-designed PNAs for sale.

PNAs display other peculiar properties. For example, Bergstrom and his coworkers noted that the universal base 3-nitropyrrole, which causes relatively little destabilization of a DNA–DNA duplex, forms a bad mismatch when supported on a PNA backbone (*81*). They concluded that the movement of the base is more tightly coupled to the movement of the backbone in PNA than in DNA due to the "flexibility of the ribose ring." Alternative explanations might invoke solvation, although this has little experimental basis.

We cannot solve the PNA problem here. It is sufficient perhaps to note only that PNA is as remarkable a molecule as DNA, if only because it carries molecular recognition much farther than do many other nonionic analogs, which lose this power at the dimer, tetramer, or hexamer stage. Perhaps the success of PNA is due to an interesting interaction that the backbone displays with water. But it seems clear that like all other nonionic analogs of DNA, PNA will fold, aggregate, and precipitate once it becomes long enough. It is not likely for this reason to be a molecular system that supports genetics in any advanced form.

PHYSICAL ORGANIC PARADOXES RELATING TO THE ORIGIN OF LIFE

This discussion creates a paradox concerning origins. In terms of its macromolecular chemistry, life on Earth is a "two-biopolymer" system. Nucleic acid is the genetic biopolymer, storing information within an organism, passing it to its descendants, and suffering the mutation that makes evolution possible. Nucleic acids also direct the biosynthesis of the second biopolymer, proteins. Proteins generate most of the selectable traits in contemporary life, from structure to motion to catalysis.

The two-biopolymer strategy evidently works well. It has lasted on Earth for billions of years, adapting to a remarkable range of environments, surviving formidable efforts by the cosmos to extinguish it, and generating intelligence capable of exploring beyond Earth.

The terrestrial version of two-biopolymer life contains a paradox, however, relating to its origins. It is difficult enough to envision a nonbiological mechanism that would allow either proteins or nucleic acids to emerge spontaneously from nonliving precursors. But it seems highly improbable that both biopolymers arose simultaneously and spontaneously, and even more improbable that both arose spontaneously, simultaneously, and as an encoder-encoded pair.

Accordingly, "single-biopolymer" models have been proposed for life that may have preceded the two-biopolymer system that we know on contemporary Earth (*82*). Such models postulate that a single biopolymer can perform both the catalytic and genetic roles and undergo the Darwinian evolution that defines life (*51*). RNA was proposed some time ago as an example of such a biopolymer (*8,9,83,84*). This proposal became more credible after Cech, Altman, and their coworkers (*5,6,85*) showed that RNA performs catalytic functions in contemporary organisms. The notion of an "RNA World," an episode in natural history where RNA served both genetic and catalytic roles, is now part of the culture of molecular biology (*86*).

The discussion above suggests that catalysis on one hand and information storage on the other place competing and contradictory demands on molecular structure. Specifically:

1. A biopolymer specialized to be a catalyst must have many building blocks, so that it can display a rich versatility of chemical reactivity. A biopolymer specialized to store information must have few building blocks, as a way of ensuring faithful replication (*87,88*).

2. A biopolymer specialized to be a catalyst must fold easily so that it can form an active site. A biopolymer specialized to store information should not fold easily, so that it can serve as a template.

3. A biopolymer specialized for catalysis must be able to change its physical properties rapidly with few changes in its sequence, enabling it to explore "function space" during divergent evolution. A biopolymer specialized to encode information must have physical properties largely unchanged even after substantial change in its sequence, so that the polymer remains acceptable to the enzymes required for replication.

At the very least, a single biopolymer attempting to support Darwinian evolution must reflect some sort of compromise between these goals. But no law requires chemistry to deliver a polymeric system that makes this compromise in a satisfactory way. The demands for functional diversity, folding, and rapid search of function space might be so stringent, and the demands for few building blocks, templating ability, and COSMIC-LOPER ability so stringent, that no biopolymer structure achieves a suitable compromise. A biopolymer may not exist that may support robust catalysis at the same time as it enables robust Darwinian evolution. If so, the single-biopolymer model for the origin of life would be unavailable as a solution to the "chicken-or-egg" paradox in the origin of two-biopolymer systems. Life would be scarce in the universe. And if a single biopolymer system did arise, it would be poorly adaptable and easily extinguished. Conversely, if many polymeric systems exist that make an acceptable compromise between the demands of catalysis and the demands of information storage, life would have emerged rapidly via single-biopolymer forms and be abundant in the universe.

SUMMARY

It is remarkable that modern science has now extended from the world of physical organic chemistry to the level of biological genetics. This accomplishment is due to the work of many who have been trained in the first discipline, and applied synthetic chemical skills to address problems in the second. Nature has chosen phosphate for

genetic molecules on Earth. While other polyanionic structures are conceivable (as are many polycationic structures as well (*52*)), there is an air of universality to a species that can support a polyanionic structure on a linker that is stable to chemical degradation but easily cleaved with catalysis. For these reasons, we would not be surprised if phosphates themselves are found widely, and perhaps universally in genetic molecules, regardless of their genesis, however human exploration might uncover them.

ACKNOWLEDGMENTS

We are indebted to the NASA Astrobiology program for partial support of this work. This work was also supported by the NIH through Grant GM 54048.

REFERENCES

1. Watson, J. D., and Crick, F. H. C. (1953) *Nature* **171,** 737–738.
2. Watson, J. D., and Crick, F. H. C. (1953) *Nature* **171,** 964–967.
3. Wang, A. H. J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., Vanboom, J. H., Vandermarel, G., and Rich, A. (1979) *Nature* **282,** 680–686.
4. Harvey, S. C., Dlakic, M., Griffith, J., Harrington, R., Park, K., Sprous, D., and Zacharias, W. (1995) *J. Biomol. Struct. Dyn.* **13,** 301–307.
5. Cech, T. R., Zaug, A. J., and Grabowski, P. J. (1981) *Cell* **27,** 487-496.
6. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) *Cell* **35,** 849–857.
7. Usher, D. A., and McHale, A. H. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1149–1153.
8. Woese, C. R. (1967) Origins of the Genetic Code. Harper & Row, New York.
9. Crick, F. H. C. (1968) *J. Mol. Biol.* **38,** 367–379.
10. Cech, T. R. (2000) *Science* **289,** 878–879.
11. Uhlmann, E., and Peyman, A. (1990) *Chem. Rev.* **90,** 543–584.
12. Mesmaeker, A., Haner, R., Martin, P., and Moser, H. (1995) *Acc. Chem. Res.* **28,** 366–374.
13. Crooke, S. T. (2000) *Methods Enzymol.* **313,** 3–45.
14. Westheimer, F. H. (1987) *Science* **235,** 1173–1178.
15. Davis, B. D. (1958) *Arch. Biochem. Biophys.* **78,** 497–509.
16. Freier, S. M., and Altmann, K. H. (1997) *Nucleic Acids Res.* **25,** 4429–4443.
17. Pitha, J., Pitha, M., and Ts'o, P. O. P. (1970) *Biochem. Biophys. Acta* **204,** 39–48.
18. Eckstein, F., Armstrong, V. W., and Sternbach, H. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2987–2990.
19. Collins, M. L. (1997) *Nucleic Acids Res.* **25,** 2979–2984.
20. Huang, J., McElroy, E. B., and Widlanski, T. S. (1994) *J. Org. Chem.* **59,** 3520–3521.
21. Perrin, K. A., Huang, J., McElroy, E. B., Iams, K. P., and Widlanski, T. S. (1994) *J. Am. Chem. Soc.* **116,** 7427–7428.
22. Mcelroy, E. B., Bandaru, R., Huang, J. X., and Widlanski, T.S. (1994) *Bioorg. Med. Chem. Lett.* **4,** 1071–1076.
23. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
24. Huang, Z., Schneider, K. C., and Benner, S. A. (1991) *J. Org. Chem.* **56,** 3869–3882.
25. Richert, C., Roughton, A. L., and Benner, S. A. (1996) *J. Am. Chem. Soc.* **118,** 4518–4531.
26. Day, R. O., Seeman, N. C., Rosenberg, J. M., and Rich, A. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 849–853.
27. Rosenberg, J. M., Seeman, N. C., Day, R. O., and Rich, A. (1976) *J. Mol. Biol.* **104,** 145–167.
28. Roughton, A. L., Portmann, S., Benner, S. A., and Egli, M. (1995) *J. Am. Chem. Soc.* **117,** 7249–7250.
29. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P., and Rich, A. (1976) *J. Mol. Biol.* **104,** 109–144.
30. Hyrup, B., Richert, C., Schulte-Herbrueggen, T., Benner, S. A., and Egli, M. (1995) *Nucleic Acids Res.* **23,** 2427–2433.
31. Chacko, K. K., Lindner, K., Saenger, W., Miller, P. S. (1983) *Nucleic Acids Res.* **11,** 2801–2814.
32. Han, F., Watt, W., Duchamp, D. J., Callahan, L., Kézdy, F. J., and Agarwal, K. (1990) *Nucleic Acids Res.* **18,** 2759–2767.

33. Huang, Z. (1993) Dissertation No. 10429, Eidgenoessische Technische Hochschule, Zurich, Switzerland.
34. Eschgfaeller, B. (1998) Dissertation No. 12582, Eidgenoessische Technische Hochschule, Zurich, Switzerland.
35. Eschgfaeller, B., Koenig, M., Boess, F., Boelsterli, U. A., and Benner, S. A. (1998) *J. Med. Chem.* **41,** 276–283.
36. Steinbeck, C., and Richert, C. (1998) *J. Am. Chem. Soc.* **120,** 11576–11580.
37. Brant, D. A., and Flory, P. J. (1965) *J. Am. Chem. Soc.* **87,** 2788–2791.
38. Macdermott, A. J., Barron, L. D., Brack, A., Buhse, T., Drake, A. F., Emery, R., Gottarelli, G., Greenberg, J. M., Haberle, R., et al. (1996) *Planetary Space Science* **44,** 1441–1446.
39. McBride, J. M., and Carter, R. L. (1991) *Angew. Chem. Int. Ed. Engl.* **30,** 293–295.
40. Wintner, E. A., Conn, M. M., and Rebek, Jr., J. (1994) *Acc. Chem. Res.* **27,** 198–203.
41. Menger, F. M., Eliseev, A. V., Khanjin, N. A., and Sherrod, M. J. (1995) *J. Org. Chem.* **60,** 2870–2878.
42. Wintner, E. A., Tsao, B., and Rebek, J. (1995) *J. Org. Chem.* **60,** 7997–8001.
43. Lee, D. H., Severin, K., Yokobayashi, Y., and Ghadiri, M. R. (1997) *Nature* **390,** 591–594.
44. von Kiedrowski, G., Wlotzka, B., Helbing, J., Matzen, M., and Jordan, S. (1991) *Angew. Chem.* **30,** 423–426.
45. Benner, S. A., and Switzer, C. Y. (1999) *in* Simplicity and Complexity in Proteins and Nucleic Acids (Frauenfelder, H., Deisenhofer, J., Wolynes, P. G., Eds.), pp. 335–359, Dahlem Workshop Report, Dahlem University Press, Berlin.
46. Saghatelian, A., Yokobayashi, Y., Soltani, K., and Ghadiri, M. R. A. (2001) *Nature* **409,** 797–801.
47. Johnsson, K., Allemann, R. K., and Benner, S. A. (1990) *in* Molecular Mechanisms in Bioorganic Processes (Bleasdale, C., and Golding, B. T., Eds.), pp. 166–187, Roy. Soc. Chem., Cambridge.
48. Johnsson, K., Allemann, R. K., Widmer, H., and Benner, S. A. (1993) *Nature* **365,** 530–532.
49. Allemann, R. K. (1989) Dissertation No. 8804, Eidgenoessische Technische Hochschule, Zurich, Switzerland.
50. Wang, Y., and Patel, D. J. (1994) *Structure* **2,** 1141–1156.
51. Joyce, G. F. (1994) *in* Origins of Life: The Central Concepts (Deamer, D. W., and Fleischaker, G. R., Eds.), Jones and Bartlett, Boston.
52. Dempcy, R. O., Almarsson, O., and Bruice, T. C. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7864–7868.
53. Nielsen, P. G., Egholm, M., Berg, R. H., and Buchardt, O. (1991) *Science* **254,** 1497–1500.
54. Corey, D. R. (1997) *Trends Biotech.* **15,** 221–224.
55. Larson, H. J., Bentin, T., and Nielsen, P. E. (1999) *Biochim. Biophys. Acta* **1489,** 159–166.
56. Ray, A., and Norden, B. (2000) *FASEB J.* **14,** 1041–1060.
57. Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) *Nature* **365,** 566–568.
58. Ratilainen, T., Holmen, A., Tuite, E., Nielsen, P. E., and Norden, B. (2000) *Biochemistry* **39,** 7781–7791.
59. Smulevitch, S. V., Simmons, C. G., Norton, J. C., Wise, T. W., and Corey, D. R. (1996) *Nat. Biotechnol.* **14,** 1700–1704.
60. Demidov, V. V., Potaman, V. N., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., Sonnichsen, S. H., and Nielsen, P. E. (1994) *Biochem. Pharmacol.* **48,** 1310–1313.
61. Hamilton, S. E., Simmons, C. G., Kathriya, I., and Corey, D. R. (1999) *Chem. Biol.* **6,** 343–351.
62. Lansdorp, P. M., Verwoerd, N. P., van de Rijke, F. M., Dragowska, V., Little, M.-T., Dirks, R. W., Raap, A. K., and Tanke, H. J. (1996) *Hum. Mol. Gen.* **5,** 658–691.
63. Norton, J. C., Piatyszek, M. A., Wright, W. E., Shay, J. W., and Corey, D. R. (1996) *Nat. Biotech.* **14,** 615–620.
64. Herbert, B. S., Pitts, A. E., Baker, S. I., Hamilton, S. E., Wright, W. E., Shay, J. W., and Corey, D. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14276–14281.
65. Boffa, L. C., Carpaneto, E. M., and Allfrey, V. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1901–1905.
66. Seeger, C., Batz, H. G., and Orum, H. (1997) *Biotechniques* **23,** 512–516.
67. Carlsson, C., Jonsson, M., Norden, B., Dulay, M. T., Zare, R. N., Noolandl, J., Nielsen, P. E., Tsui, L. C., and Zielenski, J. (1996) *Nature* **380,** 207–207.
68. Orum, H., Nielsen, P. E., Egholm, E., Berg, R. H., Buchardt, O., and Stanley, C. (1993) *Nucleic Acids. Res.* **21,** 5332–5336.

69. Castro, A., and Williams, J. G. K. (1997) *Anal. Chem.* **69,** 3915–3920.
70. Zelphati, O., Liang, X., Nguyen, C., Barlow, S., Sheng, S., Shao, Z., and Felgner, P. L. (2000) *Biotechniques* **28,** 304–316.
71. Wages, J. M., Wages, G. M., Matthews, P., Weller, D., and Summerton, J. (1997) *Biotechniques* **23,** 1116.
72. Doyle, D. F., Braasch, D. A., Simmons, C. G., Janowski, B. A., and Corey, D. R. (2001) *Biochemistry* **40,** 53–64.
73. Schmidt, J. G., Christensen, L., Nielsen, P. E., and Orgel, L. E. (1997) *Nucleic Acids Res.* **25,** 4792–4796.
74. Lutz, M. J., Benner, S. A., Hein, S., Breipohl, G., and Uhlmann, E. (1997) *J. Am. Chem. Soc.* **119,** 3177–3178.
75. Lutz, M. J., Will, D. W., Breipohl, G., Benner, S. A., and Uhlmann, E. (1999) *NucleosidesNucleotides* **18,** 393–401.
76. Egholm, M., Buchardt, O., Nielsen, P. E., and Berg, R. H. (1992) *J. Am. Chem. Soc.* **114,** 1895–1897.
77. Dueholm, K. L., Peterson, K. H., Jensen, D. K., Nielsen, P. E., Egholm, M., and Buchardt, O. (1994) *Biomed Chem. Lett.* **4,** 1077–1081.
78. Bergmann, F., Bannwarth, W., and Tam, S. (1995) *Tetrahedron Lett.* **36,** 6823–6826.
79. Gildea, B. D., Casey, S., MacNeill, J., Perry-O'Keefe, H., Sorensen, D., and Coull, J. M. (1998) *Tetrahedron Lett.* **39,** 7255–7258.
80. Gangamani, B. P., Kumar, V. A., and Ganesh, K. N. (1997) *Biochem. Biophys. Res. Commun.* **240,** 778–782.
81. Peiming, Z., Egholm, M., Paul, N., Pingle, M., and Bergstrom, D. E. (2000) *Methods* **10,** 132–140.
82. Joyce, G. F., Schwartz, A. W., Miller, S. L., and Orgel, L. E. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4398–4402.
83. Goodman, H. M., and Rich, A. (1962) *Proc. Natl. Acad. Sci. USA* **48,** 2101.
84. Orgel, L. E. (1968) *J. Mol. Biol.* **38,** 381.
85. Zaug, A. J., and Cech, T. R. (1986) *Science* **231,** 470–475.
86. Watson, J. D., *et al.* (1987) Molecular Biology of the Gene, Benjamin/Cummings, Menlo Park, CA.
87. Szathmary, E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2614–2618.
88. Lutz, M. J., Held, H. A., Hottiger, M., Hübscher, U., and Benner, S. A. (1996) *Nucleic Acids Res.* **24,** 1308–1313.