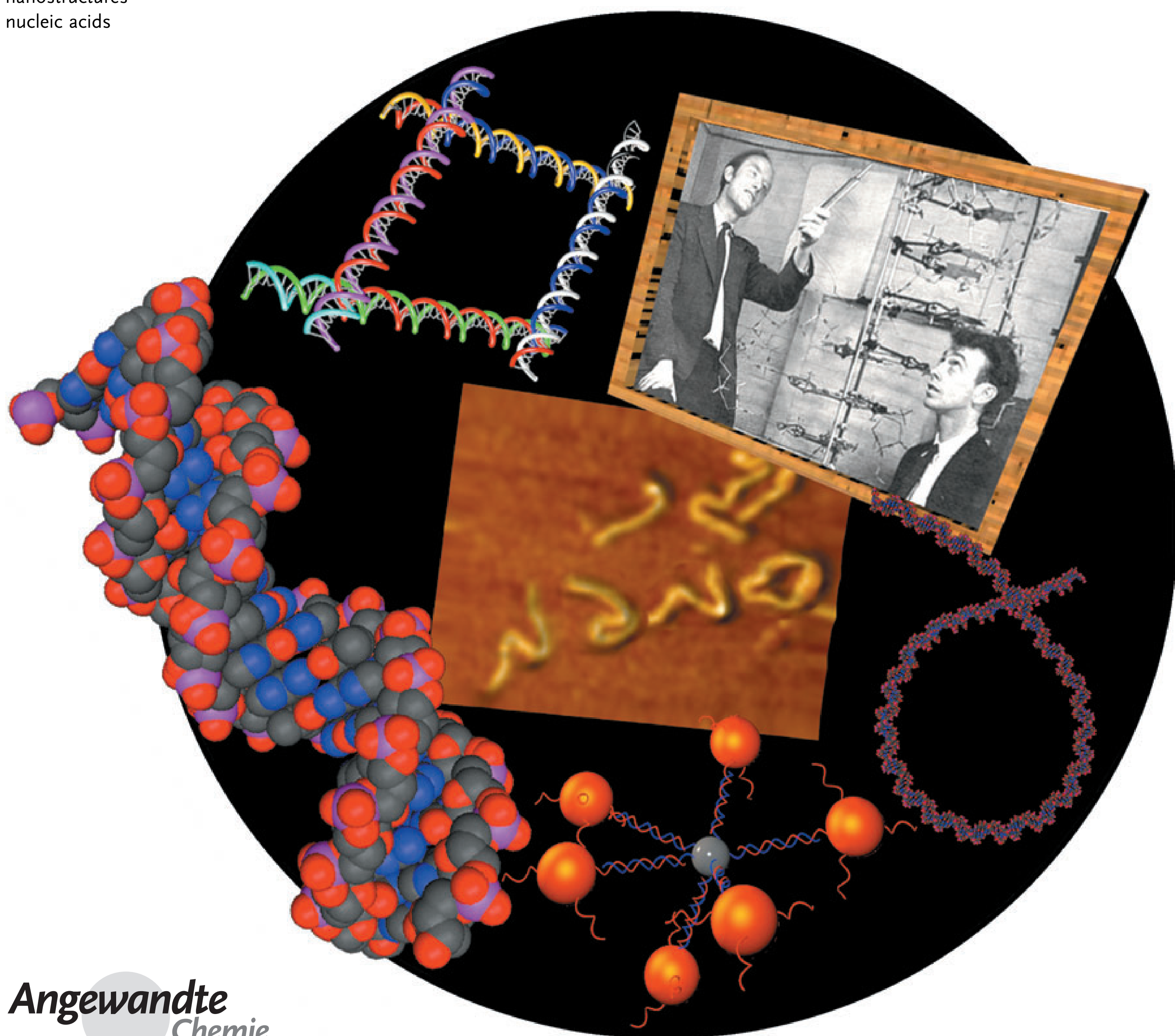
Bionanotechnology

DNA Codes for Nanoscience

Bruno Samorì and Giampaolo Zuccheri*

Keywords:
bionanotechnology · DNA · DNA–protein interactions · molecular recognition · nanostructures · nucleic acids

In memory of Claude Hélène



Angewandte Chemie

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/anie.200400652

Angewandte Chemie

The nanometer scale is a special place where all sciences meet and develop a particularly strong interdisciplinarity. While biology is a source of inspiration for nanoscientists, chemistry has a central role in turning inspirations and methods from biological systems to nanotechnological use. DNA is the biological molecule by which nanoscience and nanotechnology is mostly fascinated. Nature uses DNA not only as a repository of the genetic information, but also as a controller of the expression of the genes it contains. Thus, there are codes embedded in the DNA sequence that serve to control recognition processes on the atomic scale, such as the base pairing, and others that control processes taking place on the nanoscale. From the chemical point of view, DNA is the supramolecular building block with the highest informational content. Nanoscience has therefore the opportunity of using DNA molecules to increase the level of complexity and efficiency in self-assembling and self-directing processes.

From the Contents

1. Introduction 1167

2. The Base-Pairing Code for DNA Recognition on the Atomic Scale 1167

3. Structural Codes for DNA Recognition on the Nanoscale: Shape and Flexibility 1172

4. Summary and Outlook 1178

1. Introduction

In 2003, celebrations were held around the world in commemoration of the 50th anniversary of the discovery of the double-helical structure of DNA. Over these 50 years, the picture that Watson and Crick gave of the DNA structure in their historical paper in Nature[1] has changed significantly. On analyzing DNA fiber diffraction data, they described a chain homogeneous in its canonical B-form structure, the most common in living organisms and in normal solution conditions. They had already envisaged that "the specific base pairing immediately suggests a possible copying mechanism for the genetic material." On the other hand, a homogeneous straight chain could act only as a simple repository of the genetic information, with little function in itself. This picture was implying that the expression and the control of the genes would be delegated to proteins only. A more unified version is now finding more and more supporters: DNA itself also controls the expression of the genes through protein-recognition mechanisms that are based on the modulation of the DNA structure and dynamics along the chain. To date these recognition processes have been studied on small model systems (oligonucleotides) mostly by high-resolution techniques, such as X-ray crystallography and NMR spectroscopy. Research tools and methods common in the nanosciences now make it possible to study DNA interactive processes on the nano- and the microscale with molecules more similar to the substrates of the cellular processes. These methods led to the discovery that the codes contained in the DNA base sequence rule these interactive processes from the atomic scale of the single base-pair level to the nanometer and micrometer scale-lengths of the DNA superstructures.

The term "code" was defined by Trifonov as "any pattern or bias in the sequence which corresponds to one or another specific biological (biomolecular) function or interaction."[2] The codes of DNA are generally chemical in nature, mostly structural: it is the complementarity of the interaction between two aromatic systems that determines the base-pairing specifity. On a larger length scale, the composition of many local chain deformations in space drives, for instance, the wrapping of DNA around the histone proteins in nucleosomes in chromatin.

Herein we present an overview of the information codes embedded in DNA. These can provide a toolbox of DNA recognition processes that can be used to switch self-organization along different length and energy scales. The field of DNA nanotechnology has relied to date exclusively on the base-pairing code but there is room for much more. After a brief survey of some of the major achievements, we will dwell longer on the more complex codes that can exploit multiple hierarchies of information in the self-assembling of more and more complex DNA-based nanostructures.

2. The Base-Pairing Code for DNA Recognition on the Atomic Scale

The pairing of complementary bases between DNA molecules or, similarly, between DNA and RNA or RNA and RNA, drives not only DNA replication, but also other biological functions, such as DNA transcription, translation, and repair. In nanoscience and nanotechnology methods have been developed that more and more extensively use this code to create tools and molecular constructions.

2.1 Base-Pairing in the Cell

The specific pairing of nucleobases is the main repository of genetic information of DNA. The Watson–Crick pattern of the possible hydrogen bonds and their geometry provides the basis for the encoding of genetic information in a robust code present in two copies in any double stranded DNA (dsDNA)

[*] Prof. B. Samorì, Dr. G. Zuccheri
Department of Biochemistry "G. Moruzzi"
University of Bologna and
The National Institute for the Physics of the Matter (INFM)
Via Irnerio 48, 40126 Bologna (Italy)
Fax (+39) 051-209-4387
E-mail: bruno.samori@unibo.it

 DOI: 10.1002/anie.200400652 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

molecule. The fairly regular structure of the double helix requires a proper alignment of the bases and the exclusive pairing of a purinic with a pyrimidinic base, to create a 2-nm-wide helix.

The need of a perfect pairing for the millions of base pairs of genomic DNA requires a complex and robust control of matching, since many instances of short complementary sequences can be present and compete for the target sequence. Many proteins are used during recombination to make sure that any sub-optimal pairing is rejected (by making it unstable). On the other hand, for evolution to proceed, there must be a balance between genetic stability and rearrangements.[3]

Base-pairing errors are also caused by damage in the DNA. DNA-damaging agents are present both exogenously in the environment and also endogenously as by-products of metabolic processes. Cells have evolved a number of ways to deal with breaks in the double-strand.[4] Furthermore, certain normal processes, such as the recombination steps involved in the development of the lymphoid system in vertebrates, have the potential to create a high number of double-strand breaks.

2.2 Use of Base-Pairing for Nanoscience and Nanotechnology

The specific recognition between complementary stretches of nucleotides to form a Watson–Crick double-chain is being used more and more extensively in nanoscience to drive specific recognition processes that are used to assemble nanoscale constructions made of DNA only, of DNA and other components, or mainly of other components, in which the DNA is only used to promote the assembly. The presence of the DNA, with its controllable molecular properties, enables the creation of constructions with a predictable structure, which leads to ordered (regular) systems. The constructions can reach an impressive level of complexity, and, although being made of numerous DNA strands, assemble precisely.

The DNA base-pairing code is also used to perform calculations. Highly parallel solvers for specific problems and algorithms (such as, the Hamiltonian path) can be implemented through the use of DNA. The field of "DNA computing" is very young and very active, and is moving rapidly from simulations to realizations.[5,6]

For all the applications of the DNA base-pairing code, affinity and specificity are particularly important. In biomolecular interactions that are based on shape complementarity, or steric fit between the two counterparts (enzyme–substrate, antigen–antibody, aptamer–small-molecule complexes) both high specificity and high affinity are achieved at the same time. A nonprecise steric fit between two surfaces results in significant energetic penalties. In the case of nucleic acid interactions, however, as the binding affinity for the chosen target sequence increases, the sequence specificity decreases. This situation is due to the fact that the recognition mode and the association between two nucleic acid chains is based on a one-dimensional (1D) nucleation-zipping mechanism.[7,8] Steric fit and nucleation zipping differ dramatically. The slightest change in the shape of a key will totally impair its function. In an interaction that can be described by the three-dimensional (3D) concept of shape-complementarity, both affinity and specificity depend on the extent of steric fit, and are thus correlated. In contrast, a strong zip with one irregular or missing link can still be fastened with high affinity leaving out the small mismatched part. An increase in affinity through lengthening the complementary section of the zip does not compensate for the small loss of free energy arising from the mismatch, because the longer complex with the mismatch will be stabilized comparably. This mechanism results in a gradual decrease of nucleic acid hybridization specificity with increasing binding affinity: a scientist designing a nucleic acid probe or a molecular construction must bear in mind that a longer oligonucleotide is not always better, and that there might be alternative strategies for increasing the specificity and the affinity at the same time (see ref. [7] for an overview of some of these strategies.)

2.2.1 DNA-Based Molecular Constructions

Work done mainly by Ned Seeman's group[9–13] (but also more recently in the Reif and Winfree groups[14,15]) reveals the power of using the code in the nucleotide sequences of single-stranded DNA (ssDNA) to direct self-assembly processes. Large polymeric constructions (1D and 2D)[16] or smaller oligomeric or monomeric objects can be prepared by the assembly of at least six different oligonucleotides into structures based on the blocked Holliday junction (Figure 1a). Computerized methods have been developed to

Bruno Samorì is Professor of Organic Chemistry at the School of Biotechnology of the University of Bologna. Up until 1990, he carried out researches with polarized spectroscopy techniques on the stereochemistry of synthetic organic molecules, (at the University of Bologna, with S. F. Mason at King's College, London, with I. Tinoco at the UC Berkeley). Since then, his research interests moved to the chemistry and biology of DNA superstructures, and, more recently, also to protein folding and mechanochemistry. He has been President of the Division of the Chemistry of the Biological Systems of the Italian Chemical Society and also of the Italian Society of Microscopic Sciences.

Giampaolo Zuccheri is full-time Researcher in the Department of Biochemistry of the University of Bologna. He received his "Laurea" in 1994 from the Faculty of Industrial Chemistry of the University of Bologna and his Ph.D. in Chemistry in 1998 from the University of Calabria, Italy. He has worked with Marcos Maestre (Lawrence Berkeley National Labs) and with Carlos Bustamante (then at the University of Oregon) and he is currently working with Bruno Samorì. His interests focus on the chemistry and biophysics of nucleic acids.

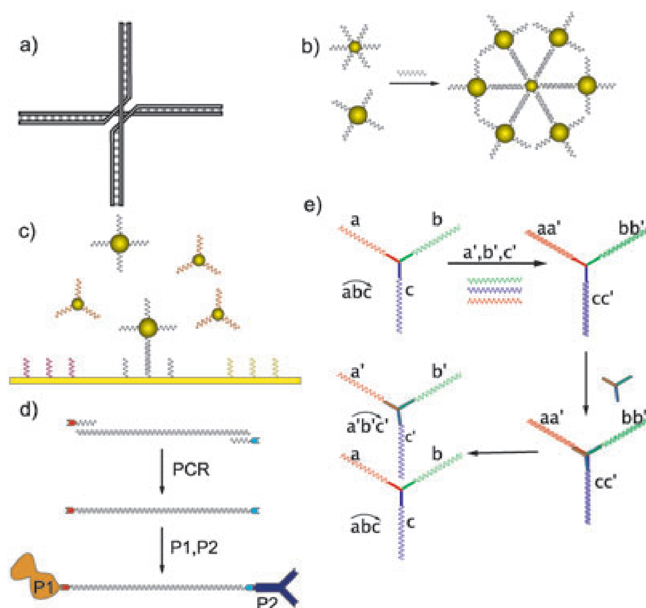


Figure 1. Examples of DNA-based nanoconstructions. a) A Holliday Junction is made by four arms of dsDNA joined at a crossover. In the naturally occurring junction (an intermediate in recombination) the crossover can migrate thanks to the sequence symmetry of the arms; in the blocked junctions, the building block for many of the constructions from the Seeman group, the position of the junction is fixed by breaking the symmetry of the sequence around the crossover point.[16] b) Linking of oligonucleotide functionalized gold nanoparticles by an oligonucleotide motif, each half of which is complementary to one or the other particle-bound oligonucleotide.[23] c) Gold nanoparticles can be specifically anchored on surface spots where oligonucleotides are attached. The assemble signal can be, similarly to (b), an oligonucleotide with each half complementary to the particle-bound or the surface-bound oligonucleotides.[28] d) PCR-mediated introduction of unusual functionalities in a dsDNA segment of desired sequence and length as a method to prepare specifically designed DNA–protein hybrids.[31] e) Self-replication of connectivity among oligonucleotides.[32] Base pairing is used to localize three oligonucleotides in space with termini at the right distance for cross-linking. The base pairing on the preformed tri-linked template regulates the configuration of the assembly on the new tri-linked structure.

choose oligonucleotide sequences that could improve on the self-assembly specificity, by minimizing the possible frameshift errors or the mismatches. These are some of the most serious sources of possible errors in the self-assembly, and which probably determine the upper limit for structural complexity. Atomic force microscopy (AFM) and classical biochemical techniques, such as electrophoresis, are used to characterize the building blocks and the constructed nanoobjects. One of the peculiar properties of the DNA-based constructions born in the Seeman group is their rigidity. Being a fairly stiff polymer, DNA lends itself to the construction of braided structures with at least twice the rigidity of dsDNA (that is, with a persistence length (see Section 3.2) of up to 100 nm).[9]

Shih and co-workers have recently reported the formation of a DNA octahedron, approximately 22 nm in diameter, through the folding of a very long (1699-nucleotide) ssDNA by a simple denaturation–renaturation procedure in the presence of five 40-mer synthetic oligonucleotides.[17] The methodology used to prepare the 1699-nucleotide-ssDNA (originally proposed by Stemmer and co-workers[18]) is also remarkable: starting from synthetic oligodeoxynucleotides, a kilo-base-pair long template can be constructed by a polymerase-chain-reaction(PCR)-based method. The template is amplified in the context of a bacterial plasmid, and then later excised, thus allowing its production by cloning.

In this type of structure DNA does not "simply" have coding and structural functions, it can also have functional properties: at least two types of approaches have been used to create motion in DNA-only objects. DNA can undergo structural transitions in response to changes in its environment. Under appropriate conditions, an alternating GC sequence can undergo the B–Z transition resulting in a reversal of handedness of the double helix. Seeman and co-workers have employed this controlled and induced rotation to change the distance between objects in space.[12] When these objects are labeled with fluorescent dyes, the motion can be easily followed studying their photophysical properties.

If nanoobjects result from the programmed assembly of a number of (long) oligonucleotides, then a second type of approach towards implementing motion can be based on a competitive reconstruction of the objects after the addition of other (more extensively) complementary DNA strands. This strategy has been used by Yurke and co-workers[19] and by Feng and co-workers[15] to drive controlled and cyclic motions in monomeric or polymeric (lattice-like) DNA objects. One of the main features of these motors is that they "burn" DNA, meaning that a full cycle of contraction–retraction requires the addition of two complementary oligonucleotides, and thus produces one double-strand of useless DNA in solution with the motor.

Yurke et al.[19] presented a nanotweezer that very ingeniously could move objects in space if driven by the specific recognition of the base-pairing code. Similarly, the code might also be used to tell which of a series of different tweezers should close or open or where in space a DNA-tethered object should be taken to: all obtained simply by adding a few components in solution.

Feng and co-workers[15] applied the strategy devised by Yurke et al. to the type of arrays invented by Seeman and co-workers. They obtained large flat objects that can be tightened or expanded by the addition of oligonucleotides, so that they might work as movers, as size-specific switchable filters, or in the mechanical release of complexed objects. Other research groups have presented examples where some of the various conformational transitions of DNA have been exploited to move objects around.[6,13,20]

2.2.2 *Exploiting the Code: DNA-Mediated Molecular Construction*

After the pioneering idea of DNA-templated circuits presented by Robinson and Seeman,[21] and the scheme for implementation of DNA-directed electrical wires by Di Mauro and Hollenberg,[22] many examples of DNA objects and constructions have been reported: these are often hybrids

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

of DNA and other types of molecules, which range from small organic molecules to proteins or even synthetic polymers.

In these materials, DNA is mainly used as a structuring element, which drives the self-assembly of molecules that would not interact, or would do so in a disorderly fashion. After the chemical preparation of units which are properly designed to self-assemble into a molecule–oligonucleotide hybrid, it is often only necessary to mix the components together in the right stoichiometry to obtain the adducts.

The group of Mirkin has developed many techniques for arranging nanoparticles exploiting the code in oligonucleotides that are attached to them.[23] The easiest way is to employ thiol-modified oligonucleotides (now available from commercial oligonucleotide providers) that attach strongly to clean gold surfaces.[24] Fritschke and co-workers,[25] Niemeyer and co-workers,[26] and Mirkin and co-workers[27] have provided examples where gold nanoparticles have been attached to flat gold surfaces thanks to the creation of a short stretch of dsDNA between nanoparticle-bound and surface-bound complementary oligonucleotides. In some of the examples, the two types of oligonucleotides are not complementary to each other, but are complementary to either half of a third oligonucleotide that can switch the anchoring of the soluble nanoparticles to the surface, when added to the solution (Figure 1c).[27,28] The same strategy can be used to create ordered aggregates of different nanoparticles in solution, where a number of oligonucleotide-functionalized nanoparticles assemble as a response to the introduction of another oligonucleotide that serves as the "glue", its two halves are complementary to the oligonucleotides anchored on the nanoparticles (Figure 1b).[23]

Niemeyer and co-workers have investigated many DNA–streptavidin constructions: biotinyilated oligonucleotides were obtained commercially, or from the derivatization of other commercially available modified oligonucleotides with hetero-bifunctional linkers.[29] These complexes can also be used to bind other biotinilated products to the already established DNA–streptavidin complex, a virtue of the tetravalency of streptavidin (that could even be expanded). By following this strategy, a wide variety of complexes can be obtained, for instance biotinilated antibodies have been bound to the oligonucleotide-bound streptavidin, so that the antibody complex could then be driven to adsorb to a specific spot on a gold surface where the complementary oligonucleotide was tethered by a gold–sulfur bond.[30]

To date, nanoscientists have mainly employed commercially available modified oligonucleotides that were brought onto the market for totally different purposes (mainly for molecular biology). The full power of synthetic chemistry have still to be unchained to the benefit of this field. A few remarkable examples of this power are already available: Abell and co-workers prepared asymmetric protein dumbbells through the use of custom prepared modified oligonucleotides that could bind proteins specifically and also work as PCR primers[31] (Figure 1d). This result could be interpreted as an initial step towards the organized synthesis of replicas of the multiprotein factories found inside cells.

Another noteworthy example of the use of organic chemistry together with DNA coding has been presented by von Kiedrowski and co-workers.[32] By preparing a tridentate linker, they could assemble a Y-shaped trisoligonucleotide, an uncommon template. The template could link to three soluble complementary oligonucleotides in solution, these oligonucleotides could in turn cross-link to another tridentate linker in a stereochemically controlled fashion (Figure 1e). The entire process can be thought of as the replication of (stereochemically controlled) connectivity, a feat requiring a high degree of control.

Nanoelectronics has been one of the guiding lights for many nanoscience researches. The base-pairing code of DNA has been used for the assembly of nanoelectronic components, as shown, for example, in the sequence-specific single-DNA-molecule lithography of Keren and co-workers.[33] The method involves the use of RecA, a protein that binds to an ssDNA molecule and facilitates its sequence-specific "invasion" inside a long dsDNA molecule. When Keren and co-workers turned a long DNA double-helix into a metallic nanowire by electroless metallization, the presence of the invading nucleoprotein complex (RecA polymerized on the single-stranded DNA) in a section of the chain served as a mask to prevent metallization of that part of the double-helix. Thus, the DNA-templated metal wire had an insulating gap whose precise location and length had been determined by the length and the sequence of the single-stranded section.

In another example of use of the DNA codes in nanoelectronics, carbon nanotubes have been derivatized with oligonucleotides. The base-pairing codes of different batches of tubes drive the assembly of families of tubes, to serve as electronic components.[34] This self-assembly could also be interesting for creating nanostructured materials with exceptional mechanical properties, thanks to nanotubes, but that are also water soluble, thanks to the bound DNA.

Recently, DNA has also been coupled with synthetic polymers, in an attempt to realize hybrid materials that could exploit the base-pairing properties of DNA.[35]

2.2.3 *RNA and Synthetic Analogues of Natural Nucleic Acids*

Storage of information through the pairing of bases is not a property of DNA only. The base-pairing of single-stranded RNA with DNA has been used to attach different oligonucleotides (and objects-bearing oligonucleotides) along the RNA chain.[29] The RNA is effectively used as a master program. Alternatively, the interactions amongst RNA molecules themselves can be used for making unusual and complex structures held together by noncovalent bonds. The artificial modular assembly units that are used to form RNA nanoobjects have been termed tecto-RNA.[36] These are self-assembling RNA building blocks that are designed and programmed to generate RNA super-architectures in a highly predictable manner for a wide range of applications.[37]

There is currently much interest, also from the pharmaceutical industry, in small double-stranded RNAs (designated as small interfering RNAs or siRNAs) because they can be used for the sequence-specific silencing of gene expression through RNA interference in eukaryotic cells.[38] Exogenously supplied siRNAs have potent and specific effects in reducing the expression of homologous endogenous genes. This gene-

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

silencing approach based on RNA interference might help to overcome efficiency problems of traditional antisense molecules.

Peptide nucleic acids (PNAs) are DNA or RNA synthetic mimics, in which the nucleobases are attached to a peptide-like (polyamide) uncharged backbone.[39] PNA oligomers can form stable duplex structures with Watson–Crick complementary DNA and RNA (or PNA) oligomers, and they can also bind to targets in duplex DNA by helix invasion.[40] Peter Nielsen, recognized as the inventor of PNA, has recently reviewed the applications of these molecules in chemistry, biology, and medicine, including drug discovery, genetic diagnostics, molecular recognition, and the origin of life.[41]

PNA molecules that function as molecular beacons have been developed for the recognition of specific sequences within dsDNA without denaturation. Molecular beacons are sensitive fluorescent probes which hybridize selectively to designated DNA and RNA targets. With the aid of PNA "openers" (dehybridizing agents that open the double helix) molecular beacons were employed for the detection of a chosen target sequence in dsDNA and, in particular, to discriminate between complementary versus mismatched dsDNA sequences.[42] PNA beacons are advantageous over DNA beacons because they can be used to analyze unpurified or non-deproteinized DNA samples. This feature of PNA beacons may find applications in the emerging area of fluorescent DNA diagnostics.[43]

PNA has several useful characteristics, these include the ability to from triple helices and complexes with other topologies, the possibility of being chiral,[44] and a stronger base-pairing in PNA–DNA hybrids than in regular dsDNA (because of the electric neutrality of the PNA backbone). These properties provide researchers with a wide spectrum of possibilities for the tailoring and the control of the specificity, the affinity, and the steric effects of the base-pair interaction.[7]

Other recently developed DNA analogues, the locked nucleic acids (LNAs) are modified nucleotides that contains a 2′-O, 4′-C methylene bridge. This bridge, locked in 3′-endo conformation, restricts the flexibility of the ribofuranose ring and locks the structure into a rigid bicyclic formation, which confers an enhanced hybridization performance and an exceptional biological stability.[45] An LNA probe forms a more stable hybrid with its target sequence than the corresponding DNA stand would.[46] This increase in hybridization stability allows a significant broadening of the experimental conditions.[47]

Eschenmoser and co-workers synthesized another type of nucleic acids called L-α-*threo*furanosyl oligonucleotides, or TNAs.[48] They found that complementary TNA strands can form stable double helices and are capable of Watson–Crick base pairing with DNA, RNA, and TNA. A variety of DNA polymerases have been screened for activity on a TNA template, and several showed a surprisingly good ability to copy limited stretches of TNA.[49]

Attempts have also been made at extending the concept of DNA base-pairing through the synthesis of a pseudoDNA containing bigger polycyclic bases.[50]

2.2.4 *DNA Chips and DNA Detection: The Reading of the Code and the Interplay Between Affinity and Specificity in the Base Pairing*

One of the most established technical applications in which the base-pairing codes of DNA are exploited is genetic analysis. For research and diagnostic purposes, cells are scanned for the presence of genes, or for the level of expression of peculiar genetic products.[51,52] Parallel genetic testing is currently performed on devices termed "sensing arrays" which consist of an array of "spots" on a surface. A different oligonucleotide is anchored to each spot and can base-pair with the DNA or RNA target or analyte molecules. The read-out of the hybridization is based on the introduction of fluorescent,[51] electroactive,[53] or nanoparticle labels[28,54] to produce a measurable signal. Several attempts at the fully electronic detection of (labeled or unlabelled) nucleic acids and proteins have been performed,[55] but the sensitivity is still far from optimal (more than 10^4 molecules of analyte required). Electrical-detection-based technologies will make the sensing-array technology suitable for simple-to-use, low-cost, point-of-care diagnostics, for applications in environmental analysis, biohazard and bioweapons detection, and for the detection of pathogens.

Many chemical strategies have been employed to immobilize nucleic acids at surfaces including electropolymerization, streptavidin–biotin interactions, gold–thiol links.[56] Fully electronic methods that would detect unlabelled DNA molecules without the need for amplification of the target molecule (that is, that are not based on methods such as PCR) are now required. This important goal strongly depends on the development of 1) nanoscience-based strategies for signal enhancement, 2) methods to increase both the affinity and the specificity of the base-pairing recognition processes. While the former is a new field, much has already been done for the latter, and a particularly clear overview about the very special interplay between affinity and selectivity in nucleic acid interaction can be found in a recent article by Demidov and Frank-Kamenetskii.[7] Some of the approaches used to increase simultaneously the affinity and specificity use oligonucleotides in novel topologies (e.g. circular, dendrimeric, nanoparticle-bound), while others employ newer types of oligonucleotide analogues (LNAs, PNAs).[7]

2.2.5 *The Need to Develop Further Strategies for the Synthesis of Bio-Nanotechnological Building Blocks*

The room for expanding the applications of DNA to nanotechnology is mainly limited by the availability of suitable and efficient chemical strategies for the preparation, purification, and handling of the necessary building blocks.

The available techniques for solid-phase synthesis can prepare milligram amounts of oligonucleotides, much less if high purity, unusual length (more than, say, 40 nucleotides), or subsequent derivatization of the oligonucleotide is called for. The high cost for the preparation of gram amounts of numerous oligonucleotides is prohibitive. For the preparation and the purification of nucleic acids, organic chemists are beginning to employ techniques familiar to the biochemist

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

and the molecular biologist, such as gel electrophoresis and PCR. These techniques that have been developed and optimized for obtaining minute quantities of products, will have to be adapted for the growing needs of the scientific community. Extremely powerful methods based on the use of restriction endonucleases or DNA ligases (methods which are the envy of polymer chemists for their precision), will also need to be made more efficient before they can be used with large quantities of substrates. More proficient methods for the derivatization of surfaces and for cross-linking will be also needed.[57] A great effort is required to adapt the available tools to the needs of a growing bio-nanotechnology community.

3. Structural Codes for DNA Recognition on the Nanoscale: Shape and Flexibility

The base sequence of a DNA segment also encodes the dynamics of the chain. DNA is continuously morphing into shapes and structures alternative to the canonical B-form, it is coiling in the cell nucleus, it is "swilling lazily around in a nourishing molecular soup of transcription factors and other regulatory proteins that are milling around the nucleus."[58]

Atomic force microscopy (AFM) imaging can give a particularly clear view of the apparently chaotic movements of a single DNA molecule. When a drop of solution containing a population of molecules of the same sequence and length is spread on the surface of a freshly-cleaved crystal of mica, and the resulting spread is imaged by AFM, the image can be like the micrograph in Figure 2. No two macromolecules have the same shape and conformation, in spite of being totally identical from any chemical or biochemical point of view. Contrary to the first impression, the apparently chaotic dynamics that leads to such a variety of shapes is not random.

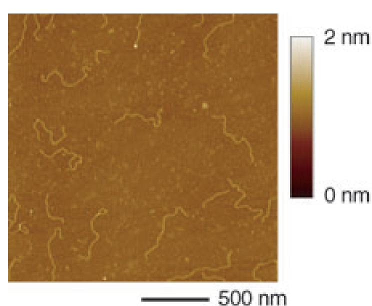


Figure 2. AFM image of a spread of DNA palindromic dimers obtained by dimerizing a restriction fragment excised out of plasmid pBR322 DNA. The 1878 base-pair long DNA molecules were spread on freshly cleaved muscovite mica from a solution of 4 mM HEPES buffer(pH 7.4), 10 mM NaCl, 2 mM $MgCl_2$ containing approximately 1 nM DNA molecules. Only the molecules completely inside the borders of the image and not presenting any anomalous structure (loops, kinks, bound objects) are used for digitally tracking the helical axis and for the subsequent curvature analysis. Smaller DNA fragments (or residual monomers) are recognized from their measured contour lengths, and subsequently neglected.

3.1. The DNA Shape Code: From the Ångström to the Nanometer Scale

The shape assumed in space and in time by a particular DNA molecule has been analyzed in terms of the superposition of the thermal fluctuations of the structure and the intrinsic, lowest energy structure of a chain with that sequence.[59,60] The average structure of dsDNA depends on the sequence: the differences in the spatial arrangement imparted by the different base pairs along the chain give rise to deterministic modulations of the relative orientations of the average planes of the base pairs. These orientations are commonly expressed in terms of the base-step orientation parameters: roll, tilt, and twist (see Figure 3).

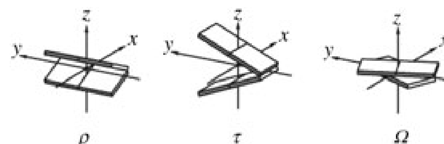


Figure 3. The dinucleotide step orientational parameters: roll (ρ), tilt (τ), and twist (Ω). The composition of the three rotations along the DNA chain gives rise to local and global curvatures.

Considerable effort has gone into defining sets of these parameters corresponding to the lowest energy structures directed by the sequence. Recently, Crothers has reviewed the approaches and the sets of values reported.[61]

Table 1 gives the values that De Santis et al. defined on the basis of conformational energy minimization calculations,[62] and then refined to fit experimental gel-retardation values.[63,64] The roll, tilt, and twist values of the dinucleotide steps significantly deviate from the canonical B-form structure proposed by Watson and Crick[1] and refined by Arnott and Hukins.[65] Sets of angles like those in Table 1 (derived from nearest-neighbor simplifications of the chain properties)

Table 1: Values [°] for the dsDNA base-step orientational parameters.[a]

	3′-end			
5′-end	A	C	G	T
Roll angles (ρ)				
A	−5.400	−2.500	1.000	−7.300
C	6.800	1.300	4.600	1.000
G	2.000	−3.700	1.300	−2.500
T	8.000	2.000	6.800	−5.400
Tilt angles (τ):	A	C	G	T
A	−0.500	−2.700	−1.600	0.000
C	0.400	0.600	0.000	1.600
G	−1.700	0.000	−0.600	2.700
T	0.000	1.700	−0.400	0.500
twist angles (Ω):	A	C	G	T
A	35.975	33.737	34.428	35.260
C	34.078	33.146	33.478	34.428
G	34.647	33.325	33.146	33.737
T	34.450	34.647	34.078	35.975

[a] See Figure 3.[64,88] The 5′-end is in the left column from top to bottom and the 3′-end is from left to right across the row: for example, the roll angle for a 5′-TA-3′ step is 8.0 degrees.

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

make it possible to easily predict the lowest energy chain conformation of a molecule from its sequence.

Positive or negative roll or tilt angles give rise to local bending of the double-helix axis. These local bends might lead to a zigzag pattern of the chain axis, which remains essentially straight, unless the bend occurs in phase with the double-stranded helical repeat. In this latter case, the bend might give rise to extended persistent curvatures that propagate from the Ångström to the nanometer scale.[66] An example of large-scale curvature is the 211 base-pair segment from the kinetoplast DNA of the Trypanosomatidae Protozoan *Crithidia fasciculata*. This is the most highly curved natural DNA known. Its sequence (Figure 4c), is characterized by a periodical recurrence of tracts of 3 to 6 adenine residues; the centers of most of these tracts are separated by 10 or 11 base pairs, that is, the average helical repeat. This distribution of the adenine tracts, perfectly phased with the helical winding, means that this short DNA segment has its lowest conformational energy when wrapped in a circle (Figure 4a,b). Experimental evidence of such a large curvature was first provided by Griffith et al.[67]

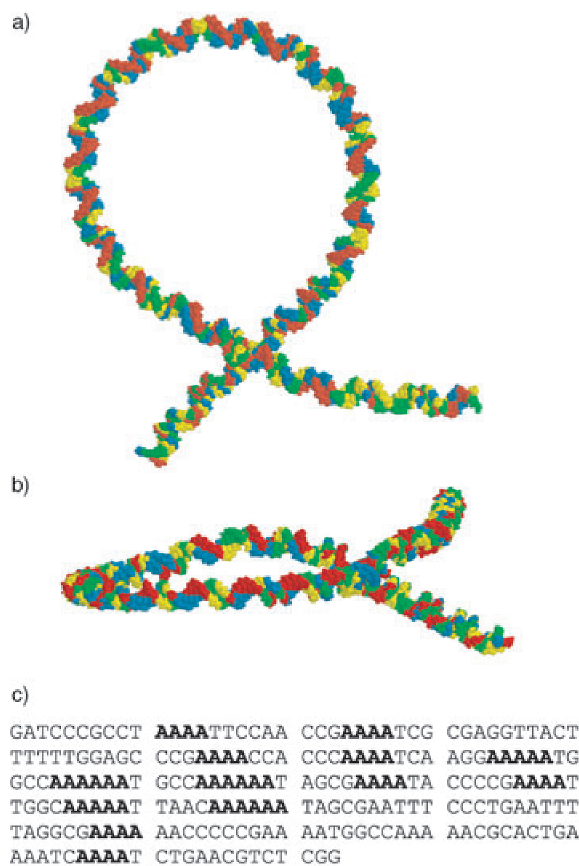


Figure 4. Predicted 3D structure of the curved segment of *Crithidia fasciculata*. (http://archimede.chem.uniroma1.it/webdna.html). a) View along a direction nearly perpendicular to the curvature plane, b) view along a direction almost parallel to the curvature plane, the segregation of adenines and thymines on the two faces of the curvature plane is clearly visible: red adenine, blue thymine, green guanines, yellow citosines. c) The base sequence of the molecule shown in (a) and (b).

Intrinsic curvatures have been monitored and studied by X-ray crystallography on very short double-stranded oligonucleotides.[68] On longer DNA molecules, the curvatures have been studied by gel retardation,[69,70] circularization kinetic,[71,72] electron microscopy (EM),[73] AFM,[74,75] and have been simulated by molecular dynamics.[76] Commonly, these experiments were carried out with dsDNA constructs with 1) anomalous flexibility sites, such as single-stranded stretches,[75] internal loops arising from mismatches,[70] a single nick,[77] a double-stranded linker connecting triple-helix tracts,[78] or 2) segments whose curvature was tailored and controlled by accurate phasing[75] or unphasing,[60,71] of adenine tracts with the helical periodicity. All these approaches to the study of intrinsic curvature are based on the determination of global parameters of the whole chain under investigation such as the persistence length, the end-to-end distance, or the cyclization J factor. Information about the sequence encoding and the molecular mechanisms that drives the formation of extended curvatures was inferred from the variations of the values of these global DNA parameters with respect to their values for reference sequences. A combinatorial approach has been proposed for this kind of study.[71]

The trace of the trajectory of the double-helical axis of individual dsDNA chains deposited on a substrate can be recorded by EM or AFM. The intrinsic curvatures of dsDNA can thus be studied from the single-molecule point of view and it is possible to set up methods to map the intrinsic curvature along the chain of a natural DNA of any sequence. This investigation is carried out by computing the curvature from the angular chain deflections (Figure 5a) along a large number of profiles, averaging the values, and plotting these averaged values as a function of the (fractional) position along the chain (Figure 5b).[74]

The intrinsic DNA curvatures can therefore be predicted theoretically and experimentally evaluated either as an average value for a particular DNA chain or as a localized value, mapped along the same chain. Despite these capabilities and many high-resolution NMR spectroscopy and X-ray studies of local bends in oligonucleotides, the description of the origin of large intrinsic curvatures at the atomic level remains disputed. No one doubts that a long-range curvature requires adenine tracts separated by a defined distance, and that it depends little on the nature of the sequences that separate the adenine tracts (see Figure 4c).[79] It is not clear how the adenine tracts play a dominant role. The discussion has been mostly focused on the composition of the geometry parameters of the base pairs at the junctions of the adenine tracts with those of the rest of the chain.[68,80] It is somewhat surprising that all different theoretical approaches based on nearest-neighbor models,[63,81,82] reach very similar results in their description of the path of the helical axis in the nanoscale despite the different values they utilize for the base-step parameters (see also Table 1 and 2 in reference [61]).

We must therefore conclude that curvature is a long-range superstructural property that is more determined by the way the double helix selects, composes, and phases the local bends over different spatial scales; the individual parameters of the

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

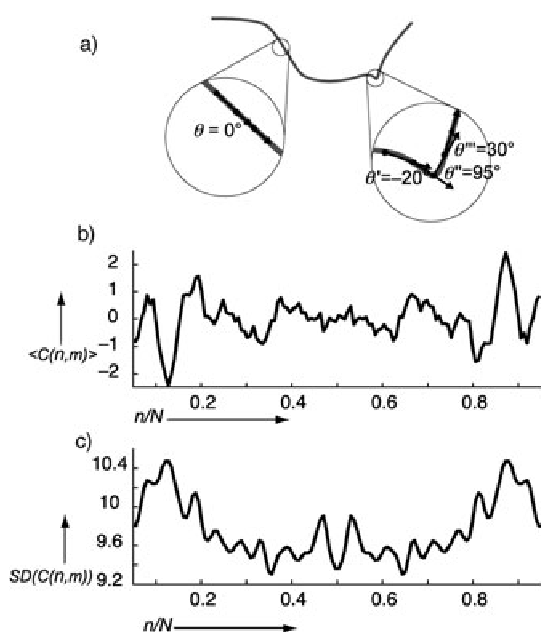


Figure 5. a) Method for quantifying the local DNA curvature from AFM images: the DNA chain axis is digitized in a segmented fashion and the local chain direction (and thus the curvature) is determined numerically; b) Plot of the experimental local DNA curvature (evaluated from the AFM images) along the chain of tail–tail palindromic dimers constructed using the EcoRV and the PstI sites of plasmid pBR322 DNA.[74] The average local curvature $\langle C(n,m)\rangle$, in degrees (where n is the sequence position, $m=2$ helical turns is the width of the averaging window) is plotted against the fractional position along the chain contour, n/N. The experimental plot has been made symmetrical by averaging the two equivalent halves to double the curvature information on the DNA sequence c) Plot of the experimental local DNA flexibility of the same DNA molecule evaluated as the standard deviation (SD) of the local chain curvature.

bends play a much less significant role. The nature of this structural code of DNA is thus not simply determined by the sequence (the words) but more subtly by the way the sequences are arranged along the chain (the language), with the notion that a certain flexibility in the sequence is allowed without producing serious changes in the average shape of the molecule.

3.2. The DNA Flexibility: From the Dinucleotide to the Micrometer Scale

The sequence in a DNA molecule determines not only the lowest energy profile of the molecule but also its local response to the thermal fluctuations. In this way the sequence controls the formation of conformers and superstructures. Note that a conformation, even if poorly populated, can play a crucial biological function. In fact, it can be recognized and selected to switch on processes that the most stable structures might not be able to activate. If our knowledge is limited to the lowest energy profiles, our chance of understanding functions might be limited. One of the experimental observables that gives insight into the accessible conformational space of a chain is its local axial flexibility, that is, the tendency of the long axis of the double helix to deviate both locally and globally from a straight trajectory.

While there is satisfactory agreement on the determinants of DNA curvature (see Section 3.1), the issue of DNA flexibility is still under debate. The considerations that will follow reflect the point of view of the authors of the articles in question, and are certainly bound to require modification as more results become available.

The axial flexibility of the chains is controlled by the spatial arrangement of their sequences, in particular by van der Waals and electrostatic interactions between the adjacent base pairs.[66,82–84] The electrostatic interactions are dominated by a large dipole on G–C pairs which is in contrast with a diffuse distribution of charge on A–T base-pairs.[84,85] Adjacent A–T base-pairs can thus stack without displacement caused by electrostatic forces, whereas the repulsive dipoles in adjacent G–C base pairs lead to a slide displacement which results in a more positive roll angle.

Another factor that influences the general axial flexibility is the compressibility of both the major and minor grooves which results in the presence of exocyclic groups in the grooves.[86] Recently, De Santis and co-workers showed that the axial flexibility is thermodynamically related to the melting temperature of a DNA tract when a first-order elasticity is assumed. This data can be easily obtained from the sequence by averaging the formal melting temperature assigned to each dinucleotide step. The results obtained by adopting such a dinucleotide flexibility scale satisfactorily explain the static and dynamic curvature dispersion of DNA images and the sequence-dependent thermodynamic stability of nucleosomes as well.[64,74,87,88]

The axial flexibility of a chain in the nanometer length scale can be described in terms of the its persistence length P, a parameter commonly used for defining a polymer bending rigidity. The parameter P is defined as the length over which the polymer axis direction is retained under thermal agitation. A number of techniques including light scattering,[89] rotational diffusion,[90] DNA cyclization kinetics,[91] cryo-electron microscopy,[60] as well as conventional electron microscopy[92,44] and AFM[93] have led to estimates for P of around 50 nm for mixed-sequence B-form DNA. The measured value of P, as determined by most techniques, depends not only on the intrinsic flexibility of the DNA molecule but also on the anisotropy of the axial flexibility which is due to its intrinsic curvature.[60,94]

On the length scale of the dinucleotide steps, the axial flexibility (or bendability) has been estimated from the range of conformations adopted by the specific base steps in crystal structures of either DNA oligomers, or of DNA–protein complexes.[82,95,96] In DNA oligomers, the deformations that force the DNA to bend locally (and make it possible to evaluate its bendability) are due to the lattice structure of the crystal. In DNA–protein complexes, the local bendability of the chain was demonstrated by the ability of the proteins to bend DNA at the binding positions. On this basis, the pyrimidine–purine steps were found to be more easily deformable than the purine–pyrimidine and purine–purine steps. The bendability of the dinucleotide steps decreases in the order CG > CA(=TG) > TA > CC(=GG) > TC(=GA)

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

> GC > TT(=AA) > GT(=AC) > CT(=AG) > AT. Olson and co-workers also deduced harmonic energy functions from the mean value and the dispersion of the base-pair step parameters.[95,97]

Imaging methodologies that enable the trajectories of the DNA molecules under investigation to be visualized, make it possible to map not only the local intrinsic curvatures along the chain but also the local modulation of flexibility that is determined by the sequence. By evaluating the dispersion of the curvature values, flexibility plots, like that in Figure 5 c, were generated for a population of symmetric molecules obtained by the dimerization of a tract of the DNA plasmid pBR322.[74] These plots show that the local flexibility is generally higher where the molecule is more curved. This indicates that the sequence shapes the conformational space of the chain by modulating the intrinsic curvature and the flexibility in the same way. Molecular dynamics simulations (based on all-atom potentials) performed on oligonucleotide duplexes with adenine tracts indicate that they are essentially straight and rigid, and that the more bent and distorted steps (with respect to the canonical B-form) are more dynamically deformable.[98]

There are 16 possible dinucleotides, of which 10 are symmetrically unique. The frequencies of these steps along the chain of the dimer shown in Figure 5 are reported in Figure 6. A good correlation with the flexibility results for the AA(=TT), TA, and AT steps. Anticorrelation is obtained instead for the CG, CA(=TG), and also the GC steps. This result is in contrast with the assignments of the bendability of the dinucleotide steps reported above. In fact, the crystal

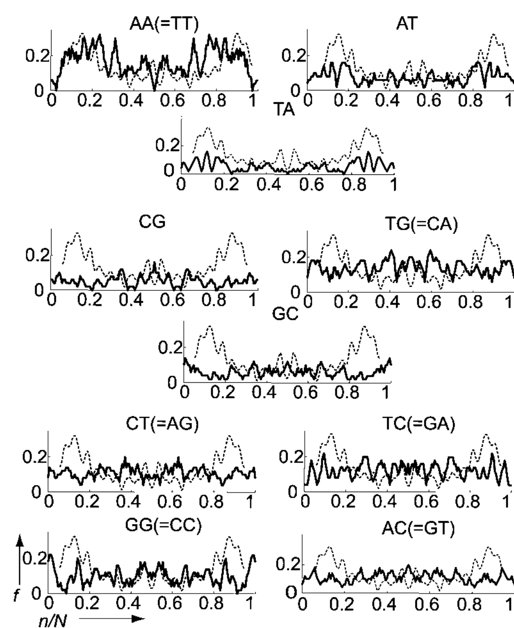


Figure 6. Plots of the dinucleotide steps frequencies along the chain of the dimeric molecules whose curvature and flexibility are plotted in Figure 5. The frequencies are evaluated with a spatial resolution similar to the curvature plots in Figure 5. *f* is the fractional step frequency, *n*/*N* is the fractional position along the chain contour. Superimposed on each plot (dashed trace, on an arbitrary scale) is the profile of the chain flexibility (as evaluated in Figure 5 c), which is shown to enable correlations to be seen.

structures indicated the CG and CA(=TG) steps as the most bendable ones. As in the case of curvature, the flexibility assignments on the nanoscale can hardly be reconciled with those made by different techniques on the atomic scale. The following considerations can be made in an attempt to explain this disagreement: It can be argued that the currently available crystallographic data are not optimal. As the authors of these studies point out, the harmonic energy functions evaluated by them can change as new data accumulate.[97] Furthermore, the oligonucleotides analyzed by X-ray diffraction contained tracts systematically GC-rich at the ends, and AT-rich in the central positions.[99] This composition might have affected the statistical significance of the analysis. Additionally, the spread of the dinucleotide geometric parameters in the set of crystal structures (about a hundred) of the oligonucleotides correlates quite poorly with that of DNA–protein complexes.[61] Nevertheless, both analyses indicate the CG and the CA (=TG) steps as particularly bendable, in marked disagreement with the results obtained on the nanoscale by AFM.[74]

It can reasonably be expected that this disagreement is due to the averaging of the smaller-scale properties along the chain. Okonogi et al. identified the CG step as the most flexible and GC as the most inflexible, while the values of AT and TA were intermediate;[100] the average of the CG and GC flexibility is lower than the average for the combined AT and TA steps. On this basis, sequences containing AT and TA steps were assigned as the most flexible of all dinucleotides.[100] On the other hand, it must be taken into account that the helical structure does not only add and average the different contributions, rather it combines them according to their phase with respect to the helical repeat. In this way, a large local flexibility might be even obtained by appropriately phasing a number of steps that might not be the most bendable. Also the DNA curvature is determined more by the way the double helix selects, composes, and phases the local bends over different spatial scales, than by the extent of the individual deformations (see Section 3.1).

One other possible difference should be taken into account when the dinucleotide bendabilities obtained from X-ray diffraction analyses (as in refs. [82,95]) or from spectroscopic methods (as in ref. [100]) are compared with flexibility data obtained by AFM imaging. The molecules imaged by AFM have been transferred from the 3D space of the solution to the quasi-2D space of the substrate surface. This reduction of the degrees of freedom from a 3D to a 2D space could damp out-of-plane motions and emphasize in-plane flexibility. This effect is another source of disagreement between the measurements made over different length scales. At the same time, results from flattened DNA are particularly valuable in structural biology, since they mimic the restrictions that are expected to occur in DNA–protein complexes, where the DNA chain is confined on the protein surface. This intrinsic limitation of the AFM technique can be useful in shedding light on the topology of these complexes.

Other information about the DNA structural codes on the nanoscale comes from the structures of the complexes of DNA with architectural proteins, such as the histones (see Section 3.3). These complexes are analyzed by focusing on

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

how conformationally rigid and flexible sequences are combined to drive the positioning of the proteins on the DNA chain that wraps them. According to Travers, the results of this analysis agrees with the assignment of AT and TA as fairly flexible steps.[86] The conclusion that TA is the most flexible step is supported by the very low stacking energy of this step, and by the data by Zhang and Crothers[71] and by McConnell and Beveridge.[98] This conclusion fits very well with the AFM data.

As far as the AA step is concerned, there seems to be an energy barrier that prevents it from adopting large distortions. On the other hand, the recent structure of the nucleosome core reported by Richmond and Davey,[101] shows large distortions at the junctions between the AA tracts and the flanking sequences. On this basis, it is likely that the properties of the AA step depend greatly on its sequence content (A. A. Travers, personal communication). The high flexibility assigned to the AA step from AFM measurements might include contributions from the mechanical properties of the flanking positions.

In conclusion, the subtle intertwining of the DNA recognition mechanisms operating at different length scales is teaching us a lot about the structural basis of size-dependent molecular phenomena.

3.3. Indirect Read-Out in DNA–Protein Recognition

DNA-binding proteins can be classified in two groups. The first is that of the proteins, such as histones and histone-like proteins, which maintain the chromosome structure and drive its modification as required by the expression and regulation of the genome. DNA is wrapped around these proteins. The specificity of their binding is strictly related to the sequence-dependent curvature and flexibility of the DNA tracts involved. In particular, the standard free energies of competitive nucleosome-reconstitution experiments are satisfactorily predicted theoretically in terms of curvature and flexibility by adopting a statistical thermodynamic model.[64,87,88]

The proteins of the second group are those involved in the regulation of the gene expression through their interaction with DNA control elements. These proteins can achieve rapid target location by initially binding to a nonspecific site on the DNA and then reaching the specific site by one dimensional diffusion or by intersegment transfer, or both.[102] These proteins recognize their specific binding sites by sampling the specific contacts through a recognition mechanism called "direct read-out."[103,104] This is a recognition process at an atomic level of resolution. This very detailed sampling can take place at each position along the DNA chain or, more efficiently, only at certain positions, such as those in which the DNA can be more easily bent. This latter process implies that during the one dimensional diffusion along the DNA, the protein continuously bends the DNA chain into a particular local conformation and so tests the way in which the DNA conformation can change: this is the so-called "indirect read-out" mechanism that is based on the sampling of conformational and mechanical properties on the nanoscale.[105]

Reporting their study on DNA bending by Cro protein,[106] Bustamante et al. suggested that the dynamics of the one-dimensional diffusion along the DNA might be dominated by the propagation of "bending waves with the protein riding at their vertex." The ease of DNA bending could signal the protein of its arrival at the specific locus. If the specific site requires less energy to distort, it will yield a more stable complex that allows time for the protein to check for sequence-specific contacts. The role of DNA bending in transcriptional regulation has been reviewed.[103] The number of proteins that are recognized to bend DNA is constantly increasing and the idea that all DNA tracking proteins bend DNA is strengthening all the time. The coding for the indirect read-out is more flexible than that for sequence-specific direct read-out: a certain sequence can be replaced by an unrelated one, as long as it is as curved and flexible as the previous one.[107]

Indirect read-out dominates the recognition processes of the architectural proteins of the first group. The same mechanism is just the first component of the recognition between the proteins of the second group and their specific sites. An example of an indirect recognition without the presence of the direct contacts (that is, recognition outside the van der Waals radii) is that of the FIS activator protein with its binding sites.[108]

3.4. The Formation of Loops: An Effect of DNA Structural Codes

As shown by Griffith et al. for the highly curved *Crithidia fasciculata* DNA,[67] the DNA shape and flexibility can facilitate the formation of loops, a superstructure that can control gene expression. In a DNA chain, these loops make it possible to mediate the interactions of units a long way apart in the primary structure.[109] The formation of DNA loops (which can be directly observed by AFM imaging[110]) depends on the appropriate phasing of base-pair deformability with the double-helical repeat. Matsumoto and Olson[97] have simulated a naturally straight chain which contains intrinsically flexible and rigid dimer steps spaced by half-turn increments (5 base pairs apart) and have found that this will bend in a preferred direction, to form loops.

3.5. A Crystal Surface Can Read the Structural Codes of DNA

Macromolecules exert exceptional control over many growth and organization processes, such as the nucleation of inorganic compounds, phase stabilization, assembly, and pattern formation.[111] Much effort is now paid to identifying the appropriate compatibilities and combinations of biological macromolecules with inorganic materials and to understanding how the organizational capabilities of biological molecules can be combined with inorganic systems in self-assembly processes. Examples of this research include, peptides with selectivity for binding to metal surfaces and metal oxide surfaces and those that can recognize and control the growth of an inorganic semiconductor surface, such as that of GaAs.[112]

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

In addition to peptides and proteins, is it possible for a DNA chain to be recognized at the biological–inorganic interface? A straight DNA chain can rotate around its axis on the surface, so many possible orientations are expected to be equally probable and the chemical characteristics are averaged to a cylindrical symmetry. On the other hand, this rotation is somewhat hindered in an intrinsically curved segment.

The intrinsic DNA curvature also defines an average plane for the curved segment: the two faces of the DNA chain (on the two sides of the plane) are also chemically different owing to the different spatial distribution of the dinucleotide steps that give rise to the chain curvature.

A high-resolution molecular model of the lowest energy conformation of the *Crithidia fasciculata* fragment shows that the almost planar structure has one face that is A-rich while the other is T-rich. This arrangement is the result of the recurrence of the A-tract that is phased with the helical winding (see Figure 4b). In principle, this structure can deposit on mica on either of these two faces. Its direct observation (with the EM or the AFM) does not allow the preferred face of adsorption, if any, to be identified because it is not possible to read the direction of the sequence.

The strategy used to determine the preferred adsorption face is shown schematically in Figure 7a: the two faces of a square thin object (for example, a paper square), one black and one white, can only be distinguished by their colors, since they have the same shape. Likewise the distinguishing feature of DNA chains, the directions of the sequence, cannot be identified in EM or AFM images of the *Crithidia* segment. On the other hand, with the colored square, if a new object is built by linking two of these squares to make a chiral shape like that in Figure 7b, the two prochiral faces of this thin object can be distinguished on a purely morphological basis: just from their shape. By ligation of two *Crithidia* segments (Figure 7c) either in the tail–tail (PvuII-EcoRI-PvuII, solid trace in Figure 8) or in the head–head (NheI-SalI-NheI, dashed trace in Figure 8) orientation two palindromic dimers were constructed. These DNA palindromic dimers of a strongly curved DNA segment are like the prochiral objects in Figure 7b. In fact, when the curved palindromic molecule in two dimensions assumes an S-like average shape (with its internal dyadic axis perpendicular to the surface), because of the segregation of A and T bases on the two faces of the monomeric curved tracts, its two prochiral faces expose either A-rich or T-rich sequences (Figure 7c). The face exposed to the crystal surface by a population of molecules deposited on mica can thus only be identified by AFM imaging of the palindromic dimers of the segment of interest, not of the segment itself. By analyzing the shape assumed by two large sets of the two palindromic dimers of the *Crithidia fasciculata* fragment (Figure 8) it was discovered that the face that both dimers expose preferentially to the mica is the T-rich one. A statistical analysis of these shapes demonstrated that the preference is such that one face was deposited between five and nine times more frequently than the other.[113]

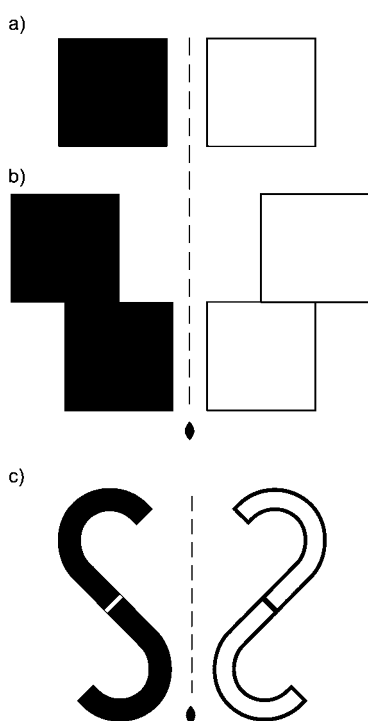


Figure 7. a) The oppositely colored faces of a thin square object can not be distinguished by form alone. b) The two oppositely colored faces of a chiral object obtained by fusing two copies of an object as in (a). In this case, the face (white or black) shown can be distinguished by the morphology of the object, even if the color could not be distinguished. c) A palindromic DNA dimer made with a curved DNA section is expected to have the same property as the model in (b) when flattened on a surface in a S-like shape: the face can be recognized without reading the direction of the base sequence.

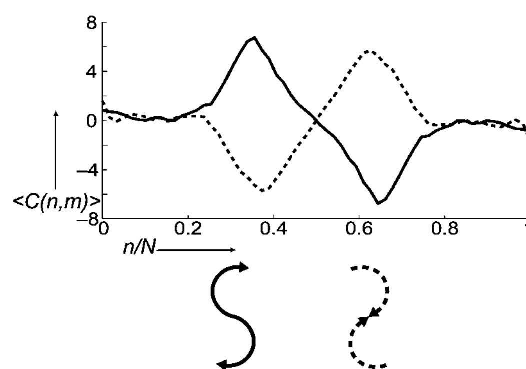


Figure 8. Plot of the experimental local DNA curvature (evaluated from the AFM images) along the chain of tail–tail (PvuII-EcoRI-PvuII, solid trace) and head–head (NheI-SalI-NheI, dashed trace) palindromic dimers constructed using the *Crithidia fasciculata* segment. The average local curvature, in degrees (over windows of approximately 2 helical turns, as in Figure 5b) are evaluated as a function of the fractional position along the chain contour, n/N. According to the signs of the local curvatures, the average shapes preferentially assumed by the tail–tail (S-type, solid line) and head–head (S*-type, dashed line) dimers are sketched.

We believe that this effect should be interpreted as the recognition of a DNA superstructure based on an indirect read-out mechanism. The extent of this recognition effect is not directly controlled by the sequence but by the degree of curvature. In fact, from the analysis of average local curvatures of dimers of pBR322 restriction fragments, with sequences without the extensive phasing found in *Crithidia*

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

and therefore with only moderate curvatures, the same evidence of the presence of a surface recognition phenomenon was found, but with a smaller magnitude (B. Samorì, P. De Santis, unpublished results). The surface of a crystal such as mica does not recognize the base sequence itself but the periodicity of the adenine tracts: if the adenine tracts are not properly phased no extended curvature occurs and thus no significant segregation of complementary bases on well-defined faces is obtained. This recognition effect is therefore expected to be general for any curved dsDNA molecules. Further investigations are requested to discover the structural basis upon which this recognition process is based and, consequently, to decipher the underlying informational code.

We can hypothesize that a recognition process of this kind might have been relevant in the pre-cellular stages of the evolution of life. Inorganic surfaces have served as catalysts for prebiotic syntheses[114] and also as templates for the self-organization of increasingly more complex biostructures.[115]

4. Summary and Outlook

The DNA sequence encodes the nanoscale properties, superstructures, and recognition mechanisms that DNA exhibits. A nanoscientific point of view shows how the DNA double-helical structure acquires different qualities over different spatial scales. DNA achieves these qualities by behaving like an antenna that eliminates the "perturbations" that are out of phase and adds only those in phase.

Looking at biology from a nanoscientific point of view introduces us into the methodologies and the purposes that nature follows in its "engineering" of complex systems. As Horst Störmer says in his lectures: "nanoscale science is raising the lid on the biggest LEGO of the universe."

While the biologists are also beginning to tackle their scientific topics from a nanoscientific point of view, nanoscience and nanotechnology are more and more inspired by biology and interested in DNA. The information content and the scale of complexity of DNA can introduce a higher level of complexity in self-assembling processes that also involve non-biological molecules.

DNA can also provide recognition processes whose selectivity and stringency can be modulated on different length scales, such as in the direct and indirect read-out mechanisms between DNA and proteins. Nanoscience and nanotechnology have already adopted the direct read-out strategy based on the base-pairing code as a handy tool. The other, more complex codes are at the moment just a source of inspiration for nanoscience. Investigations might open the possibility to exploit these codes in a similar fashion to the base pairing code. A toolbox of DNA recognition processes that could switch self-organization among different length and energy scales could be available. The discovery that DNA superstructures can be recognized by a crystal surface is a first step towards this goal: complex DNA-based self-assembling nanostructures[10] could be tailored in a way to be recognized by a crystal surface. Two different hierarchies of information could be exploited at the same time. Further steps towards higher levels of complexity could also be made by using DNA-binding proteins. DNA-based nanotraces on cationic crystal surfaces might be designed by assembling highly curved and straight DNA tracts, that are bound by proteins at consensus sequences.

Addendum

Since the submission of this manuscript, several interesting and relevant papers have been published. Sherman and Seeman showed the performance of a DNA walking device, where a DNA biped can stroll in a controlled manner on a DNA "sidewalk".[116] The fine degree of control of the motion is achieved thanks to the addition of oligonucleotides that bind the legs of the biped to specific locations on the sidewalk.

Mao et al. presented a fully autonomous DNA motor, which exploits an integrated DNA enzyme to produce cyclic motion.[117] The nanomotor continues to cycle undisturbed between its open and closed states as long as the enzyme substrate is present in solution (unless a molecular brake is set). This is a true case of conversion of chemical energy into nanoscale motion in a fully synthetic device.

Surfaces, which play a major role in nanotechnology are also becoming an important part of DNA computation: Su and Smith described a surface DNA computer,[118] while Reif and co-workers are applying DNA codes to cryptography.[119]

The principles behind DNA assembly are becoming clearer as a result of the increasing number of experimental examples (such as the DNA tetrahedron of Goodman and co-workers[120] or the rigid nanotriangle modules and arrays of Mao and co-workers[121]). Winfree and co-workers have tried to define the thermodynamic rules for the correct assembly of nucleic acids sequences.[122]

The very prolific group of Chengde Mao has also shown how to do nanoscale lithography using DNA arrays: once laid on a surface, the holes of the arrays can be filled with metal by high-vacuum evaporation. Removing the metal film, while leaving the DNA on the surface, yields a nanoscale-patterned metal structure that is a negative replica of the DNA array pattern.[123]

On July 28th, 2004, Sir Francis Harry Compton Crick died.

Experimental work was supported by Ministero dell'Istruzione Università Ricerca, Programmi Biotecnologie Legge 95/95 (5 %), and Progetti di Interesse Nazionale 2001–2003, Progetti Pluriennali Università di Bologna, and FISR D.M. 16/10/20 Anno 1999 and ESF Eurocore "SONS" Programme for 2003–2006. We thank Pasquale De Santis (Rome) for helpful discussion and comments about the manuscript. We also thank Andrew Travers (Cambridge) for sharing unpublished work, for critically reading the manuscript, and for very helpful suggestions, especially about the comparative analysis of the flexibility of the TA and AA steps (Section 3.2). We thank Dr. Andrea Giro for help in preparation of the frontispiece graphics.

Received: January 28, 2004
Revised: April 12, 2004
Published Online: November 5, 2004

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737 – 738.
[2] E. N. Trifonov, *Bull. Math. Biol.* **1989**, *51*, 417 – 432.
[3] D. R. F. Leach in *DNA Recombination and Repair* (Eds.: P. J. Smith, C. J. Jones), Oxford University Press, Oxford, **1999**, pp. 1 – 15.
[4] B. K. Singleton, P. A. Jeggo in *DNA Recombination and Repair* (Eds.: P. J. Smith, C. J. Jones), Oxford University Press, Oxford, **1999**, pp. 16 – 37.
[5] a) L. M. Adleman, *Science* **1994**, *266*, 1021 – 1024; b) C. Mao, T. H. LaBean, J. H. Relf, N. C. Seeman, *Nature* **2000**, *407*, 493 – 496.
[6] a) Y. Benenson, R. Adar, T. Paz-Elizur, Z. Livneh, E. Shapiro, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2191 – 2196; b) Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh, E. Shapiro, *Nature* **2001**, *414*, 430 – 434.
[7] V. V. Demidov, M. D. Frank-Kamenetskii, *Trends Biochem. Sci.* **2004**, *29*, 62 – 71.
[8] V. A. Bloomfield, D. M. Crothers, I. Tinoco, *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito, **2000**.
[9] P. Sa-Ardyen, A. V. Vologodskii, N. C. Seeman, *Biophys. J.* **2003**, *84*, 3829 – 3837.
[10] N. C. Seeman, *Nature* **2003**, *421*, 427 – 431.
[11] N. C. Seeman, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 225 – 248.
[12] C. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144 – 146.
[13] H. Yan, X. Zhang, Z. Shen, N. C. Seeman, *Nature* **2002**, *415*, 62 – 65.
[14] a) E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **1998**, *394*; b) H. Yan, T. H. LaBean, L. Feng, J. H. Reif, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8103 – 8108; c) H. Li, S. H. Park, J. H. Reif, T. H. LaBean, H. Yan, *J. Am. Chem. Soc.* **2004**, *126*, 418 – 419; d) R. M. Dirks, M. Lin, E. Winfree, N. A. Pierce, *Nucleic Acids Res.* **2004**, *32*, 1392 – 1403.
[15] L. Feng, S. H. Park, J. H. Reif, H. Yan, *Angew. Chem.* **2003**, *115*, 4478 – 4482; *Angew. Chem. Int. Ed.* **2003**, *42*, 4342 – 4346.
[16] C. Mao, W. Sun, N. C. Seeman, *J. Am. Chem. Soc.* **1999**, *121*, 5437 – 5544.
[17] W. M. Shih, J. D. Quispe, G. F. Joyce, *Nature* **2004** *427*, 618 – 621.
[18] W. P. Stemmer, A. Crameri, K. D. Ha, T. M. Brennan, H. L. Heyneker, *Gene* **1995**, *164*, 49 – 53.
[19] B. Yurke, A. J. Turberfield, A. P. J. Mills, F. C. Simmel, J. L. Neumann, *Nature* **2000**, *406*, 605 – 608.
[20] a) P. Alberti, J. L. Mergny, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1569 – 1573; b) D. Liu, S. Balasubramanian, *Angew. Chem.* **2003**, *115*, 5912 – 5914; *Angew. Chem. Int. Ed.* **2003**, *42*, 5734 – 5736; c) C. M. Niemeyer, M. Adler, S. Lenhert, S. Gao, H. Fuchs, L. F. Chi, *ChemBioChem* **2001**, *2*, 260 – 264; d) C. M. Niemeyer, M. Adler, *Angew. Chem.* **2002**, *114*, 3933 – 3937; *Angew. Chem. Int. Ed.* **2002**, *41*, 3779 – 3783; e) J. W. J. Li, W. H. Tan, *Nano Lett.* **2002**, *2*, 315 – 318.
[21] B. H. Robinson, N. C. Seeman, *Protein Eng.* **1987**, *1*, 295 – 300.
[22] E. Di Mauro, C. P. Hollenberg, *Adv. Mater.* **1993**, *5*, 384 – 386.
[23] C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **1996**, *382*, 607 – 609.
[24] a) M. Hegner, P. Wagner, G. Semenza, *FEBS Lett.* **1993**, *336*, 452 – 456; b) S.-J. Park, A. A. Lazarides, C. A. Mirkin, P. W. Brazis, C. R. Kannewurf, R. L. Letsinjger, *Angew. Chem.* **2000**, *112*, 4003 – 4006; *Angew. Chem. Int. Ed.* **2000**, *39*, 3845 – 3848; c) C. Bamdad, *Biophys. J* **1998**, *75*, 1997 – 2003.
[25] a) A. Csáki, R. Möller, W. Straube, J. M. Köhler, W. Fritzsche, *Nucleic Acids Res.* **2001**, *29*, E81; b) R. Möller, A. Csáki, J. M. Köhler, W. Fritzsche, *Nucleic Acids Res.* **2000**, *28*, E91.
[26] a) C. M. Niemeyer, B. Ceyhan, P. Hazarika, *Angew. Chem.* **2003**, *115*, 5944-5948; *Angew. Chem. Int. Ed.* **2003**, *42*, 5766 – 5770; b) C. M. Niemeyer, B. Ceyhan, M. Noyong, U. Simon, *Biochem. Biophys. Res. Commun.* **2003**, *311*, 995 – 999.
[27] L. M. Demers, D. S. Ginger, S. J. Park, Z. Li, S. W. Chung, C. A. Mirkin, *Science* **2002**, *296*, 1836 – 1838.
[28] S. J. Park, T. A. Taton, C. A. Mirkin, *Science* **2002**, *295*, 1503 – 1506.
[29] C. M. Niemeyer, W. Bürger, J. Peplies, *Angew. Chem.* **1998**, *110*, 2391 – 2395; *Angew Chem. Int. Ed.* **1998**, *37*, 2265 – 2268.
[30] C. M. Niemeyer, B. Ceyhan, *Angew. Chem.* **2001**, *113*, 3798 – 3801; *Angew. Chem. Int. Ed.* **2001**, *40*, 3685 – 3688.
[31] J. M. Tomkins, B. K. Nabbs, K. Barnes, M. Legido, A. J. Blacker, R. A. McKendry, C. Abell, *ChemBioChem* **2001**, *2*, 375 – 378.
[32] L. H. Eckardt, K. Naumann, W. M. Pankau, M. Rein, M. Schweitzer, N. Windhad, G. von Kiedrowski, *Nature* **2002**, *420*, 286.
[33] K. Keren, M. Krueger, R. Gilad, G. Ben-Yoseph, U. Sivan, E. Braun, *Science* **2002**, *297*, 72 – 75.
[34] K. A. Williams, P. T. Veenhuizen, B. G. de la Torre, R. Eritja, C. Dekker, *Nature* **2002**, *420*, 761.
[35] a) K. J. Watson, S. J. Park, J. H. Im, S. T. Nguyen, C. A. Mirkin, *J. Am. Chem. Soc.* **2001**, *123*, 5592 – 5593; b) L. Zhu, P. S. Lukeman, J. W. Canary, N. C. Seeman, *J. Am. Chem. Soc.* **2003**, *125*, 10178 – 10179.
[36] L. Jaeger, N. B. Leontis, *Angew. Chem.* **2000**, *112*, 2576 – 2580; *Angew. Chem. Int. Ed.* **2000**, *39*, 2521 – 2524.
[37] L. Jaeger, E. Westhof, N. B. Leontis, *Nucleic Acids Res.* **2001**, *29*, 455 – 463.
[38] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, T. Tuschl, *Nature* **2001**, *411*, 494 – 498.
[39] P. E. Nielsen, M. Egholm, O. Buchart, *Bioconjugate Chem.* **1994**, *5*, 3 – 7.
[40] V. V. Demidov, N. E. Broude, I. V. Lavrentieva-Smolina, H. Kuhn, M. D. Frank-Kamenetskii, *ChemBioChem* **2001**, *2*, 133 – 139.
[41] P. E. Nielsen, *Mol. Biotechnol.* **2004**, *26*, 233 – 248.
[42] H. Kuhn, V. V. Demidov, J. M. Coull, M. J. Fiandaca, B. D. Gildea, M. D. Frank-Kamenetskii, *J. Am. Chem. Soc.* **2002**, *124*, 1097 – 1103.
[43] H. Kuhn, V. V. Demidov, B. D. Gildea, M. J. Fiandaca, J. C. Coull, M. D. Frank-Kamenetskii, *Antisense Nucleic Acid Drug Dev.* **2001**, *11*, 265 – 270.
[44] S. Sforza, G. Haaima, R. Marchelli, P. E. Nielsen, *Eur. J. Org. Chem.* **1999**, 197 – 204.
[45] M. Petersen, J. Wengel, *Trends Biotechnol.* **2003**, *21*, 74 – 81.
[46] U. Christensen, N. Jacobsen, V. K. Rajwanshi, J. Wengel, T. Koch, *Biochem. J.* **2001**, *354*, 481 – 484.
[47] D. Latorra, D. Hopkins, K. Campbell, J. M. Hurley, *Biotechniques* **2003**, *34*, 1150 – 1158.
[48] a) K. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347 – 1351; b) A. Eschenmoser, *Science* **1999**, *284*, 2118 – 2124.
[49] J. C. Chaput, J. K. Ichida, J. W. Szostak, *J. Am. Chem. Soc.* **2003**, *125*, 856 – 857.
[50] a) H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, E. T. Kool, *Science* **2003**, *302*, 868 – 871; b) A. Marx, D. Summerer, *Angew. Chem.* **2004**, *116*, 1653 – 1654; *Angew. Chem. Int. Ed.* **2004**, *43*, 1625 – 1626.
[51] E. M. Southern, *Methods Mol. Biol.* **2001**, *170*, 1 – 15.
[52] M. J. Heller, *Annu. Rev. Biomed. Eng.* **2002**, *4*, 129 – 153.
[53] a) E. Palecek, *Talanta* **2002**, *45*, 809 – 819; b) E. Palecek, F. Jelen, *Crit. Rev. Anal. Chem.* **2002**, *32*, 261 – 270; c) K. Kerman, M. Kobayashi, E. Tamiya, *Meas. Sci. Technol.* **2004**, *15*, R1 – R11.

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[54] T. A. Taton, C. A. Mirkin, R. L. Letsinger, *Science* **2000**, *289*, 1757 – 1760.
[55] C. Guiducci, C. Stagni, G. Zuccheri, A. Bogliolo, L. Benini, B. Samorì, B. Riccò, *Biosens. Bioelectron.* **2004**, *19*, 781 – 787.
[56] R. Hölzel, N. Gajovic Eichelmann, F. F. Bier, *Biosens. Bioelectron* **2003**, *18*, 555 – 564.
[57] M. C. Pirrung, *Angew. Chem* **2002**, *41*, 1326 – 1341; *Angew. Chem. Int. Ed.* **2002**, *41*, 1276 – 1289.
[58] H. Pearson, *Nature* **2003**, *421*, 310 – 312.
[59] a) J. A. Schellman, S. C. Harvey, *Biophys. Chem.* **1995**, *55*, 95 – 114; b) C. R. Calladine, H. R. Drew, *J. Mol. Biol.* **1996**, *257*, 479 – 485.
[60] J. Bednar, P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, A. Stasiak, *J. Mol. Biol.* **1995**, *254*, 579 – 594.
[61] D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15163 – 15165.
[62] a) P. De Santis, A. Palleschi, M. Savino, A. Scipioni, *Biophys. Chem.* **1988**, *32*, 305 – 317; b) P. De Santis, A. Palleschi, S. Morosetti, M. Savino in *Structure and Dynamics of Nucleic Acids, Proteins and Membranes* (Eds.: E. Clementi, S. Chin), Plenum, New York, **1986**, pp. 31 – 49.
[63] P. De Santis, A. Palleschi, M. Savino, A. Scipioni, *Biochemistry* **1990**, *29*, 9269 – 9273.
[64] a) D. Boffelli, P. De Santis, A. Palleschi, G. Risuleo, M. Savino, *FEBS Lett.* **1992**, *300*, 175 – 178; b) P. De Santis, A. Palleschi, M. Savino, A. Scipioni, *Biophys. Chem.* **1992**, *42*, 147 – 152.
[65] S. Arnott, D. W. Hukins, *J. Mol. Biol.* **1973**, *81*, 93 – 105.
[66] M. A. El Hassan, C. R. Calladine, *Philos. Trans. R. Soc. London* **1997**, *355*, 43 – 100.
[67] J. Griffith, M. Bleyman, C. A. Rauch, P. A. Kitchin, P. T. Englund, *Cell* **1986**, *46*, 717 – 724.
[68] R. E. Dickerson, D. Goodsell, M. L. Kopka, *J. Mol. Biol.* **1996**, *256*, 108 – 125.
[69] P. R. Hardwidge, R. B. Den, E. D. Ross, L. J. Maher, *J. Biomol. Struct. Dyn.* **2000**, *18*, 219 – 230.
[70] J. D. Kahn, E. Yun, D. M. Crothers, *Nature* **1994**, *368*, 163 – 166.
[71] Y. Zhang, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3161 – 3166.
[72] a) Y. Zhang, D. M. Crothers, *Biophys. J.* **2003**, *84*, 136 – 153; b) P. De Santis, M. Fuà, M. Savino, C. Anselmi, G. Bocchinfuso, *J. Phys. Chem.* **1996**, *100*, 9968 – 9976.
[73] a) J. A. Cognet, C. Pakleza, D. Cherny, E. Delain, E. Le Cam, *J. Mol. Biol.* **1999**, *285*, 997 – 1009; b) G. Muzard, B. Theveny, B. Revet, *EMBO J.* **1990**, *9*, 1289 – 1298.
[74] G. Zuccheri, A. Scipioni, V. Cavaliere, G. Gargiulo, P. De Santis, B. Samorì, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3074 – 3079.
[75] C. Rivetti, C. Walker, C. Bustamante, *J. Mol. Biol.* **1998**, *280*, 41 – 59.
[76] W. K. Olson, V. B. Zhurkin, *Curr. Opin. Struct. Biol.* **2000**, *10*, 286 – 297.
[77] E. Le Cam, F. Fack, J. Menissier-de Murcia, J. A. Cognet, A. Barbin, V. Sarantoglou, B. Revet, E. Delain, G. de Murcia, *J. Mol. Biol.* **1994**, *235*, 1062 – 1071.
[78] T. Akiyama, M. E. Hogan, *Biochemistry* **1997**, *36*, 2307 – 2315.
[79] T. E. Haran, J. D. Kahn, D. M. Crothers, *J. Mol. Biol.* **1994**, *244*, 135 – 143.
[80] a) A. Barbic, D. P. Zimmer, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2369 – 2373; b) D. Strahs, T. Schlick, *J. Mol. Biol.* **2000**, *301*, 643 – 663.
[81] A. Bolshoy, P. McNamara, R. E. Harrington, E. N. Trifonov, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2312 – 2316.
[82] A. A. Gorin, V. B. Zhurkin, W. K. Olson, *J. Mol. Biol.* **1995**, *247*, 34 – 48.
[83] a) C. R. Calladine, *J. Mol. Biol.* **1982**, *161*, 343 – 352; b) K. Yanagi, G. G. Prive, R. E. Dickerson, *J. Mol. Biol.* **1991**, *217*, 201 – 214; c) M. A. Young, G. Ravishanker, D. L. Beveridge, H. M. Berman, *Biophys. J.* **1995**, *68*, 2454 – 2468.
[84] C. A. Hunter, *J. Mol. Biol.* **1993**, *230*, 1025 – 1054.
[85] a) C. A. Hunter, *Bioessays* **1996**, *18*, 157 – 162; b) J. Gallego, A. R. Pascual-Teresa, M. T. Pisabarro, F. Gago in *QSAR and Molecular Modeling. Concepts, Computational Tools and Applications* (Eds.: F. Sanz, J. Giraldo, F. J. R. Manaut), Prous Science, Barcelona, **1995**, pp. 274 – 281.
[86] A. A. Travers, *Philos. Trans. R. Soc. London Ser. A* **2004**, *362*, 1423 – 1438.
[87] C. Anselmi, G. Bocchinfuso, P. De Santis, M. Savino, A. Scipioni, *J. Mol. Biol.* **1999**, *286*, 1293 – 1301.
[88] C. Anselmi, P. De Santis, R. Paparcone, M. Savino, A. Scipioni, *Biophys. Chem.* **2002**, *95*, 23 – 47.
[89] E. S. Sobel, J. A. Harpst, *Biopolymers* **1991**, *31*, 1559 – 1564.
[90] P. J. Hagerman, *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 265 – 286.
[91] a) D. Shore, R. L. Baldwin, *J. Mol. Biol.* **1983**, *170*, 957 – 981; b) W. H. Taylor, P. J. Hagerman, *J. Mol. Biol.* **1990**, *212*, 363 – 376.
[92] C. Frontali, E. Dore, A. Ferrauto, E. Gratton, A. Bettini, M. R. Pozzan, E. Valdevit, *Biopolymers* **1979**, *18*, 1353 – 1373.
[93] T. Berge, N. S. Jenkins, R. B. Hopkirk, M. J. Waring, J. M. Edwardson, R. M. Henderson, *Nucleic Acids Res.* **2002**, *30*, 2980 – 2986.
[94] A. Scipioni, C. Anselmi, G. Zuccheri, B. Samorì, P. De Santis, *Biophys. J.* **2002**, *83*, 2408 – 2418.
[95] W. K. Olson, A. A. Gorin, X. J. Lu, L. M. Hock, V. B. Zhurkin, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11163 – 11168.
[96] a) R. E. Dickerson, *Nucleic Acids Res.* **1998**, *26*, 1906 – 1926; b) R. E. Dickerson, T. K. Chiu, *Biopolymers* **1997**, *44*, 361 – 403.
[97] A. Matsumoto, W. K. Olson, *Biophys. J.* **2002**, *83*, 22 – 41.
[98] K. J. McConnell, D. L. Beveridge, *J. Mol. Biol.* **2001**, *314*, 23 – 40.
[99] W. K. Olson, http://rutchem.rutgers.edu/~olson/Bdna_list.html, **1998**.
[100] T. M. Okonogi, S. C. Alley, A. W. Reese, P. B. Hopkins, B. H. Robinson, *Biophys. J.* **2002**, *83*, 3446 – 3459.
[101] T. J. Richmond, C. A. Davey, *Nature* **2003**, *423*, 145 – 150.
[102] a) P. H. von Hippel, O. G. Berg, *J. Biol. Chem.* **1989**, *264*, 675 – 678; b) R. Fickert, B. Muller-Hill, *J. Mol. Biol.* **1992**, *226*, 59 – 68.
[103] J. Perez-Martin, V. de Lorenzo, *Annu. Rev. Microbiol.* **1997**, *51*, 593 – 628.
[104] A. A. Travers, *DNA–Protein Interactions, 1st Ed.*, Chapman & Hall, London, New York, **1993**.
[105] L. Jen-Jacobson, *Biopolymers* **1997**, *44*, 153 – 180.
[106] D. A. Erie, G. Yang, H. C. Schultz, C. Bustamante, *Science* **1994**, *266*, 1562 – 1566.
[107] B. Revet, S. Brahms, G. Brahms, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7535 – 7539.
[108] I. K. Pemberton, G. Muskhelishvili, A. A. Travers, M. Buckle, *J. Mol. Biol.* **2002**, *318*, 651 – 663.
[109] K. Rippe, P. H. von Hippel, J. Langowski, *Trends Biochem. Sci.* **1995**, *20*, 500 – 506.
[110] a) C. Wyman, E. Grotkopp, C. Bustamante, H. C. Nelson, *EMBO J.* **1995**, *14*, 117 – 123; b) K. Rippe, M. Guthold, P. H. von Hippel, C. Bustamante, *J. Mol. Biol.* **1997**, *270*, 125 – 138; c) M. Barna, T. Merghoub, J. A. Costoya, D. Ruggero, M. Branford, A. Bergia, B. Samorì, P. P. Pandolfi, *Dev. Cell* **2002**, *3*, 499 – 510.
[111] a) N. C. Seeman, A. M. Belcher, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6451 – 6455; b) A. M. Belcher, X. H. Wu, R. J. Christensen, P. K. Hansma, G. D. Stucky, D. E. Morse, *Nature* **1996**, *381*, 56 – 58; c) G. Falini, S. Albeck, S. Weiner, L. Addadi, *Science* **1996**, *271*, 67 – 69.

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[112] a) S. Brown, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8651 – 8655; b) S. Brown, *Nat. Biotechnol.* **1997**, *15*, 269 – 272; c) S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, A. M. Belcher, *Nature* **2000**, *405*, 665 – 668.
[113] B. Sampaolese, A. Bergia, A. Scipioni, G. Zuccheri, M. Savino, B. Samorì, P. De Santis, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13566 – 13570.
[114] R. Saladino, U. Ciambecchini, C. Crestini, G. Costanzo, R. Negri, E. Di Mauro, *ChemBioChem* **2003**, *4*, 514 – 521.
[115] a) N. Lahav, S. Nir, A. C. Elitzur, *Prog. Biophys. Mol. Biol.* **2001**, *75*, 75 – 120; b) P. Szabó, I. Scheuring, T. Czárán, E. Szathmáry, *Nature* **2002**, *420*, 340 – 343; c) S. J. Sowerby, N. G. Holm, G. B. Petersen, *Biosystems* **2001**, *61*, 69 – 78; d) S. J. Sowerby, C. A. Cohn, W. M. Heckl, N. G. Holm, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 820 – 822; e) C. Anselmi, P. De Santis, R. Paparcone, M. Savino, A. Scipioni, *Origins Life Evol. Biosphere* **2004**, *34*, 143 – 149.
[116] W. B. Sherman, N. C. Seeman, *Nano Lett.* **2004**, *4*, 1203 – 1207.
[117] a) Y. Chen, C. Mao, *J. Am. Chem. Soc.* **2004**, *126*, 8626 – 8627; b) Y. Chen, M. Wang, C. Mao, *Angew. Chem.* **2004**, *116*, 3638 – 3641; *Angew. Chem. Int. Ed.* **2004**, *43*, 3554 – 3557.
[118] X. Su, L. M. Smith, *Nucleic Acids Res.* **2004**, *32*, 3115 – 3123.
[119] A. Gehani, T. LaBean, J. Reif in *Aspects of Molecular Computing* (Eds.: N. Jonoska, G. Paun, G. Rozenberg), Springer, Heidelberg, **2004**, pp. 167 – 188.
[120] R. P. Goodman, R. M. Berry, A. J. Turberfield, *Chem. Commun.* **2004**, 1372 – 1373.
[121] D. Liu, M. Wang, Z. Deng, R. Walulu, C. Mao, *J. Am. Chem. Soc.* **2004**, *126*, 2324 – 2325.
[122] R. M. Dirks, M. Lin, E. Winfree, N. A. Pierce, *Nucleic Acids Res.* **2004**, *32*, 1392 – 1403.
[123] Z. Deng, C. Mao, *Angew. Chem.* **2004**, *116*, 4160 – 4162; *Angew. Chem. Int. Ed.* **2004**, *43*, 4068 – 4070.

Quality counts...

The best of chemistry every week

Wiley-VCH
P.O. Box 10 11 61
69451 Weinheim
Germany
Phone +49 (0) 6201 – 606-400
Fax +49 (0) 6201 – 606-184
e-mail: angewandte@wiley-vch.de
www.angewandte.org

Angewandte Chemie International Edition is a journal of the GDCh, the German Chemical Society

GDCh

WILEY-VCH

12180404_gu

 © 2005 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim