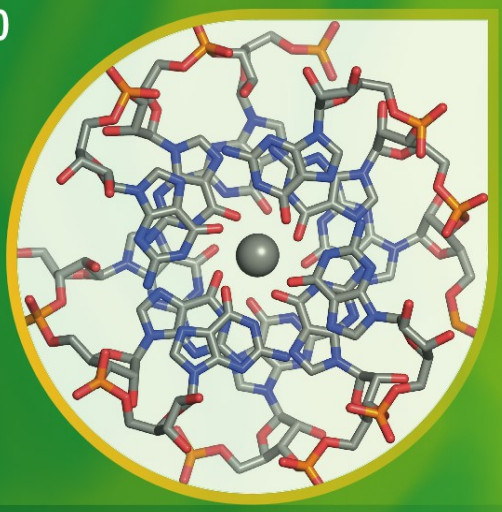
Metal Ions in Life Sciences 10



Astrid Sigel
Helmut Sigel
Roland K.O. Sigel
Editors

Interplay between Metal Ions and Nucleic Acids

Springer

Interplay between Metal Ions and Nucleic Acids

Metal Ions in Life Sciences

Volume 10

Series Editors:

Astrid Sigel, Helmut Sigel, and Roland K.O. Sigel

For further volumes:
http://www.springer.com/series/8385 and http://www.mils-series.com

Astrid Sigel • Helmut Sigel • Roland K.O. Sigel
Editors

Interplay between Metal Ions and Nucleic Acids

Springer

Editors
Astrid Sigel
Department of Chemistry
Inorganic Chemistry
Universität Basel
Spitalstrasse 51
CH-4056 Basel
Switzerland
astrid.sigel@unibas.ch

Helmut Sigel
Department of Chemistry
Inorganic Chemistry
Universität Basel
Spitalstrasse 51
CH-4056 Basel
Switzerland
helmut.sigel@unibas.ch

Roland K.O. Sigel
Institute of Inorganic Chemistry
Universität Zürich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland
roland.sigel@aci.uzh.ch

The figure on the dust cover shows the upper part of Figure 2 (c) of Chapter 1 by M. Pechlaner and Roland K. O. Sigel

ISSN 1559-0836 e-ISSN 1868-0402
ISBN 978-94-007-2171-5 e-ISBN 978-94-007-2172-2
DOI 10.1007/978-94-007-2172-2
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011943624

Historical Development and Perspectives of the Series

Metal Ions in Life Sciences*

It is an old wisdom that metals are indispensable for life. Indeed, several of them, like sodium, potassium, and calcium, are easily discovered in living matter. However, the role of metals and their impact on life remained largely hidden until inorganic chemistry and coordination chemistry experienced a pronounced revival in the 1950s. The experimental and theoretical tools created in this period and their application to biochemical problems led to the development of the field or discipline now known as *Bioinorganic Chemistry*, *Inorganic Biochemistry*, or more recently also often addressed as *Biological Inorganic Chemistry*.

By 1970 *Bioinorganic Chemistry* was established and further promoted by the book series *Metal Ions in Biological Systems* founded in 1973 (edited by H.S., who was soon joined by A.S.) and published by Marcel Dekker, Inc., New York, for more than 30 years. After this company ceased to be a family endeavor and its acquisition by another company, we decided, after having edited 44 volumes of the *MIBS* series (the last two together with R.K.O.S.) to launch a new and broader minded series to cover today's needs in the *Life Sciences*. Therefore, the Sigels new series is entitled

***Metal Ions in Life Sciences*.**

After publication of the first four volumes (2006–2008) with John Wiley & Sons, Ltd., Chichester, UK, and the next five volumes (2009–2011) with the Royal Society of Chemistry, Cambridge, UK, we are happy to join forces now in this still new endeavor with Springer Science & Business Media B.V., Dordrecht, The Netherlands; a most experienced Publisher in the *Sciences*.

The development of *Biological Inorganic Chemistry* during the past 40 years was and still is driven by several factors; among these are (i) the attempts to reveal the interplay between metal ions and peptides, nucleotides, hormones or vitamins, etc., (ii) the efforts regarding the understanding of accumulation, transport, metabolism and toxicity of metal ions, (iii) the development and application of metal-based drugs, (iv) biomimetic syntheses with the aim to understand biological processes as well as to create efficient catalysts, (v) the determination of high-resolution structures of proteins, nucleic acids, and other biomolecules, (vi) the utilization of powerful spectroscopic tools allowing studies of structures and dynamics, and (vii), more recently, the widespread use of macromolecular engineering to create new biologically relevant structures at will. All this and more is and will be reflected in the volumes of the series *Metal Ions in Life Sciences*.

The importance of metal ions to the vital functions of living organisms, hence, to their health and well-being, is nowadays well accepted. However, in spite of all the progress made, we are still only at the brink of understanding these processes. Therefore, the series *Metal Ions in Life Sciences* will endeavor to link coordination chemistry and biochemistry in their widest sense. Despite the evident expectation that a great deal of future outstanding discoveries will be made in the interdisciplinary areas of science, there are still "language" barriers between the historically separate spheres of chemistry, biology, medicine, and physics. Thus, it is one of the aims of this series to catalyze mutual "understanding".

It is our hope that *Metal Ions in Life Sciences* proves a stimulus for new activities in the fascinating "field" of *Biological Inorganic Chemistry*. If so, it will well serve its purpose and be a rewarding result for the efforts spent by the authors.

Astrid Sigel and Helmut Sigel
Department of Chemistry, Inorganic Chemistry,
University of Basel, CH-4056 Basel, Switzerland

Roland K.O. Sigel
Institute of Inorganic Chemistry,
Unversity of Zürich, CH-8057 Zürich, Switzerland

October 2005,
October 2008,
and August 2011

Preface to Volume 10

Interplay between Metal Ions and Nucleic Acids

The preceding Volume 9, *Structural and Catalytic Roles of Metal Ions in RNA*, was to a large part devoted to ribozymes, a vibrant but well defined research area. The present volume, though related to the previous one, is of a much wider character by describing phenomena that are generally due to the interrelations between metal ions and nucleic acids, especially DNA. Yet it should be noted that the role of metal ion-nucleic acid interactions in medication, tumor diagnosis, and anticancer research is not specifically considered because these topics were dealt with in Volumes 41 and 42 of our series *Metal Ions in Biological Systems*. Instead, Volume 10 focuses in 12 chapters on modern developments encompassing the wide range from G-quadruplexes via DNAzymes to peptide nucleic acids (PNAs), topics of relevance, e.g., for chemistry and nanotechnology, but also for molecular biology and genetic diagnostics.

The volume opens with two general chapters. The first one provides an overview on metal ion-nucleic acid interactions and their characterization in solution by describing the identification of metal ion binding sites with chemical and biochemical methods, the application of NMR, etc. It deals with the determination of binding constants and the effects of anions and buffers on metal ion binding to nucleic acids. The corresponding interactions as seen in the solid state by crystal structure analysis are the topic of Chapter 2. Though the interactions of metal ions with nucleic acids and their constituents have attracted much attention over more than five decades, the review focuses mainly on results obtained during the past 15 years leading to tables with over 200 entries. This large body of information is classified and discussed.

Due to the flexible nature of oligonucleotides, they can easily undergo conformational changes which may be observed by various methods. DNA can adopt a number of secondary structures of which the right-handed B-type helix is the predominant form. The possible roles of metal ions regarding the transition from B- to A-DNA (or Z-DNA), the transition from right- to left-handed helices, the denaturation of double-strands, the condensation, and other conformational changes are discussed

in Chapter 3. Special attention deserve polynucleotides containing long repeats of guanosine nucleotides, which have long been known to have a marked tendency to aggregate into gel-like materials as is pointed out in Chapter 4. Analysis has shown that four strands are held together by a hydrogen bonding arrangement of four guanines – the G-quartet. This very stable motif can self-associate and layers of G-quartets may then be formed either from one, two or four strands of DNA or RNA sequences. The roles of metal ions in stabilizing these four-stranded structures are evaluated.

Living organisms from all kingdoms of life need transition metal ions as essential micronutrients as is emphasized in Chapter 5. On the other hand, the intrinsic toxicity of the majority of these metal ions demands a tightly regulated intracellular trafficking that controls their concentration and minimizes the amount of free metal ions. The crucial players in the metal homeostasis networks are specific metal-responsive transcriptional regulators, generally defined as "metal sensors"; they couple specific metal ion binding with a change in their DNA binding affinity and/or specificity, thus translating the concentration of a certain metal ion into a change in genetic expression.

Lanthanide(III) ions have been used as surrogates for alkaline earth ion binding to nucleic acids. Indeed, as demonstrated in Chapter 6, detailed information about the Ln(III) coordination sphere can be obtained from luminescence studies providing insights into metal ion-nucleic acid interactions. Similarly, certain metal ion complexes are well known for their redox chemistry and their properties of oxygen activation as pointed out in Chapter 7. DNA damage and cleavage by metal complexes has attracted much attention since metal complexes may be useful tools for studying molecular DNA lesions due to oxidative stress, which is implicated in many disorders including aging, cancer, neurological diseases, etc. The review focuses on examples illustrating fundamental mechanisms in the chemistry of DNA oxidation by metal ion complexes.

For a long time DNA's only function was perceived as being the genetic material for all organisms. Because of the limitation of functional group diversity in the building blocks of DNA and the mostly invariant double helical structure, DNA was considered as being incapable of catalyzing chemical reactions. This contrasts with the information assembled in Chapter 8: Nowadays we know that DNAzymes may catalyze cleavage of RNA, the ligation of DNA and RNA, the formation of an RNA branch or a lariat, as well as phosphorylation, adenylation, depurination of DNA, and many more reactions. DNAzymes carry out their catalysis with the aid of metal cofactors and some are very selective in this respect. Therefore, DNAzymes have been converted into fluorimetric, calorimetric, and electrochemical sensors for metal ions. Furthermore, especially during the past two decades enantioselective catalysis at the DNA scaffold has emerged, as reviewed in Chapter 9. Thus, DNA turned out to be a highly versatile molecule for applications beyond its natural function, namely in nanotechnology, DNA-templated synthesis, and hybrid catalytic systems.

The interior of the DNA duplex is formed by a parallel stack of pairwise-bonded aromatic nucleobases. Hydrogen bonding, next to π-stacking and shape complementarity, plays an important role in governing the integrity of the double strand.

Replacement of protons from the nucleobases by metal ions leads to base pairs in which hydrogen bonds are formally replaced by coordinative bonds to metal ions. As discussed in Chapter 10, artificial nucleotides have been developed with a large affinity towards metal ions. However, as shown in Chapter 11, metal-mediated base pairs with natural nucleosides as well as artificial purine- or pyrimidine-derived nucleosides may be used as well.

In the terminating Chapter 12 peptide nucleic acids (PNAs) are described; these are non-cyclic pseudo-peptide-nucleic acid structural mimics, in which two nucleobases are linked by a peptide unit instead of a ribosyl-phosphate-ribosyl unit. Consequently, PNA binds more strongly to sequence complementary oligonucleotides than does natural DNA or RNA because there is no electrostatic repulsion between the strands. PNAs have caught the interest in many fields of science from pure chemistry over molecular biology, drug discovery and (genetic) diagnostics to nanotechnology and prebiotic chemistry. The review focuses on metal-complex derivatives of PNA. Such derivatives offer the opportunity to introduce new labels and probes for bioanalytical and diagnostic applications of PNA, but also to modulate or to introduce, e.g., catalytic functions and to create specific biological activities.

Astrid Sigel
Helmut Sigel
Roland K.O. Sigel

Contents

Contributors to Volume 10

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Christopher M. Andolina Department of Chemistry, University at Buffalo, State University of New York, Natural Sciences Complex, Buffalo, NY 14260-3000, USA, christopher.andolina@gmail.com (171)

Katsuyuki Aoki Department of Materials Science, Toyohashi University of Technology, Tempaku-cho, Toyohashi 441-8580, Japan, kaoki@tutms.tut.ac.jp (43)

Nancy H. Campbell CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, UK (119)

Stefano Ciurli Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Viale Giuseppe Fanin 40, I-40127 Bologna, Italy, stefano.ciurli@unibo.it (135)

CERM (Center of Magnetic Resonance), University of Florence, Via Luigi Sacconi 6, I-50019, Sesto Fiorentino (Florence), Italy

Guido H. Clever Institute for Inorganic Chemistry, Georg-August University Göttingen, Tammannstr. 4, D-37077 Göttingen, Germany, gclever@gwdg.de (269)

Almudena García-Fernández Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands (249)

Roland Krämer Inorganic Chemistry Institute, University of Heidelberg, Im Neuenheimer Feld 270, D-69120 Heidelberg, Germany, roland.kraemer@urz.uni-heidelberg.de (319)

Tian Lan Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA (217)

Yi Lu Department of Biochemistry and Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, yi-lu@illinois.edu (217)

Dominik A. Megger Institute for Inorganic and Analytical Chemistry, University of Münster, Corrensstr. 28/30, D-48149 Münster, Germany (295)

Nicole Megger Institute for Inorganic and Analytical Chemistry, University of Münster, Corrensstr. 28/30, D-48149 Münster, Germany (295)

Andriy Mokhir Inorganic Chemistry Institute, University of Heidelberg, Im Neuenheimer Feld 270, D-69120 Heidelberg, Germany, Andriy.Mokhir@urz.uni-heidelberg.de (319)

Janet R. Morrow Department of Chemistry, University at Buffalo, State University of New York, Natural Sciences Complex, Buffalo, NY 14260-3000, USA, jmorrow@buffalo.edu (171)

Jens Müller Institute for Inorganic and Analytical Chemistry, University of Münster, Corrensstr. 28/30, D-48149 Münster, Germany, mueller.j@uni-muenster.de (295)

Kazutaka Murayama Graduate School of Biomedical Engineering, Tohoku University, Aoba, Sendai 980-8575, Japan, kmura@bme.tohoku.ac.jp (43)

Francesco Musiani Laboratory of Bioinorganic Chemistry, Università di Bologna, Viale Giuseppe Fanin 40, I-40127 Bologna, Italy, francesco.musiani@unibo.it (135)

Stephen Neidle CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, UK, stephen.neidle@pharmacy.ac.uk (119)

Maria Pechlaner Institute of Inorganic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland (1)

Geneviève Pratviel Laboratoire de Chimie de Coordination, CNRS, 205, Route de Narbonne, F-31077 Toulouse-Cedex 04, France, genevieve.pratviel@lcc-toulouse.fr (201)

Gerard Roelfes Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands, j.g.roelfes@rug.nl (249)

Mitsuhiko Shionoya Department of Chemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, shionoya@chem.s.u-tokyo.ac.jp (269)

Roland K.O. Sigel Institute of Inorganic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, roland.sigel@aci.uzh.ch (1)

Bernhard Spingler Institute of Inorganic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, spingler@aci.uzh.ch (103)

Barbara Zambelli Laboratory of Bioinorganic Chemistry, Università di Bologna, Viale Giuseppe Fanin 40, I-40127 Bologna, Italy, barbara.zambelli@unibo.it (135)

Titles of Volumes 1-44 in the *Metal Ions in Biological Systems* Series

edited by the SIGELs
and published by Dekker/Taylor & Francis (1973–2005)

Volume 1:	**Simple Complexes**
Volume 2:	**Mixed-Ligand Complexes**
Volume 3:	**High Molecular Complexes**
Volume 4:	**Metal Ions as Probes**
Volume 5:	**Reactivity of Coordination Compounds**
Volume 6:	**Biological Action of Metal Ions**
Volume 7:	**Iron in Model and Natural Compounds**
Volume 8:	**Nucleotides and Derivatives: Their Ligating Ambivalency**
Volume 9:	**Amino Acids and Derivatives as Ambivalent Ligands**
Volume 10:	**Carcinogenicity and Metal Ions**
Volume 11:	**Metal Complexes as Anticancer Agents**
Volume 12:	**Properties of Copper**
Volume 13:	**Copper Proteins**
Volume 14:	**Inorganic Drugs in Deficiency and Disease**
Volume 15:	**Zinc and Its Role in Biology and Nutrition**
Volume 16:	**Methods Involving Metal Ions and Complexes in Clinical Chemistry**
Volume 17:	**Calcium and Its Role in Biology**
Volume 18:	**Circulation of Metals in the Environment**
Volume 19:	**Antibiotics and Their Complexes**
Volume 20:	**Concepts on Metal Ion Toxicity**
Volume 21:	**Applications of Nuclear Magnetic Resonance to Paramagnetic Species**
Volume 22:	**ENDOR, EPR, and Electron Spin Echo for Probing Coordination Spheres**
Volume 23:	**Nickel and Its Role in Biology**
Volume 24:	**Aluminum and Its Role in Biology**
Volume 25:	**Interrelations Among Metal Ions, Enzymes, and Gene Expression**
Volume 26:	**Compendium on Magnesium and Its Role in Biology, Nutrition, and Physiology**

Volume 27: **Electron Transfer Reactions in Metalloproteins**
Volume 28: **Degradation of Environmental Pollutants by Microorganisms and Their Metalloenzymes**
Volume 29: **Biological Properties of Metal Alkyl Derivatives**
Volume 30: **Metalloenzymes Involving Amino Acid-Residue and Related Radicals**
Volume 31: **Vanadium and Its Role for Life**
Volume 32: **Interactions of Metal Ions with Nucleotides, Nucleic Acids, and Their Constituents**
Volume 33: **Probing Nucleic Acids by Metal Ion Complexes of Small Molecules**
Volume 34: **Mercury and Its Effects on Environment and Biology**
Volume 35: **Iron Transport and Storage in Microorganisms, Plants, and Animals**
Volume 36: **Interrelations Between Free Radicals and Metal Ions in Life Processes**
Volume 37: **Manganese and Its Role in Biological Processes**
Volume 38: **Probing of Proteins by Metal Ions and Their Low-Molecular-Weight Complexes**
Volume 39: **Molybdenum and Tungsten. Their Roles in Biological Processes**
Volume 40: **The Lanthanides and Their Interrelations with Biosystems**
Volume 41: **Metal Ions and Their Complexes in Medication**
Volume 42: **Metal Complexes in Tumor Diagnosis and as Anticancer Agents**
Volume 43: **Biogeochemical Cycles of Elements**
Volume 44: **Biogeochemistry, Availability, and Transport of Metals in the Environment**

Contents of Volumes in the *Metal Ions in Life Sciences* Series

edited by the SIGELs

Volumes 1–4
published by John Wiley & Sons, Ltd., Chichester, UK (2006–2008)
<http://www.Wiley.com/go/mils>

Volume 5–9
by the Royal Society of Chemistry, Cambridge, UK (since 2009)
<http://www.rsc.org/shop/metalionsinlifesciences>

and from Volume 10 on
by Springer Science & Business Media BV, Dordrecht, The Netherlands (since 2012)
<http://www.mils-series.com>

Volume 1 Neurodegenerative Diseases and Metal Ions

Volume 3 The Ubiquitous Roles of Cytochrome P450 Proteins

Volume 9 Structural and Catalytic Roles of Metal Ions in RNA

Comments and suggestions with regard to contents, topics, and the like for future volumes of the series are welcome.

Chapter 1
Characterization of Metal Ion-Nucleic Acid Interactions in Solution

Maria Pechlaner and Roland K.O. Sigel

Contents

M. Pechlaner • R.K.O. Sigel (✉)
Institute of Inorganic Chemistry, University of Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland
e-mail: roland.sigel@aci.uzh.ch

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*,
Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_1,

Abstract Metal ions are inextricably involved with nucleic acids due to their polyanionic nature. In order to understand the structure and function of RNAs and DNAs, one needs to have detailed pictures on the structural, thermodynamic, and kinetic properties of metal ion interactions with these biomacromolecules. In this review we first compile the physicochemical properties of metal ions found and used in combination with nucleic acids in solution. The main part then describes the various methods developed over the past decades to investigate metal ion binding by nucleic acids in solution. This includes for example hydrolytic and radical cleavage experiments, mutational approaches, as well as kinetic isotope effects. In addition, spectroscopic techniques like EPR, lanthanide(III) luminescence, IR and Raman as well as various NMR methods are summarized. Aside from gaining knowledge about the thermodynamic properties on the metal ion-nucleic acid interactions, especially NMR can be used to extract information on the kinetics of ligand exchange rates of the metal ions applied. The final section deals with the influence of anions, buffers, and the solvent permittivity on the binding equilibria between metal ions and nucleic acids. Little is known on some of these aspects, but it is clear that these three factors have a large influence on the interaction between metal ions and nucleic acids.

Keywords equilibrium constants • metal ions • methods • ribozymes • RNA

1 Introduction

Due to their polyanionic nature nucleic acids are unthinkable to exist without the close association of cationic counterions. Already the formation of secondary structures depends on the presence of metal ions because it requires the negatively charged backbones to come close to each other [1,2]. Larger RNAs form complicated tertiary structures with patches and cavities of increased negative electrostatic potential [3,4] that can not be stable without charge neutralization by associated

cations. Most of the time, these ions are merely part of a diffuse ion atmosphere, providing the necessary charge neutralization but not interacting directly with the nucleic acid. Such ions can be well described by electrostatic continuum models based on the non-linear Poisson-Boltzmann (NLPB) equation [5]. A very small fraction of metal ions is coordinated tightly in buried binding pockets, where they are forced to release part or all of their hydration shells. In between the two extremes are those metal ions that interact with specific nucleic acid environments via a network of hydrogen bonds formed with their inner-shell water ligands and those, that are localized momentarily at electrostatically favorable sites like the major groove of RNA [3]. The wealth of loosely associated metal ions makes it difficult to single out the few more specifically bound metal ions.

It does not help that affinities even for the more selective and specific metal ion binding sites in nucleic acids are usually low (10^2 and 10^4 M^{-1} [6–8]) compared to those observed in proteins. Typically most of the coordination sites in an RNA will have very similar metal affinities and are therefore filled more or less simultaneously [7,9,10]. In addition, metal ions bound to nucleic acids usually keep at least a part of their hydration shell, which is rather an exception in proteins, and fast exchange with the solvent is a common feature. Thus, the very dynamic nature of nucleic acids is also reflected in the characteristics of their interactions with metal ions. Consequently, approaches that work very well for metal ion-protein binding are often not appropriate or not even possible for nucleic acids. Clearly, a much higher emphasis has to be set on the thermodynamics and kinetics as opposed to the structural features. The verification of conclusions derived from the solid state and crystal structures is of particular importance, if they are to be transferred to the solution state.

Having said the above, a few more words on the importance of metal ion-nucleic acid interactions are necessary. Metal ions have stabilizing roles in DNA and RNA, but can also be responsible for large conformational changes. They can induce bending and unwinding [11,12] of nucleic acids. In addition, they are triggers for the B- to Z-transition in DNA (see also Chapter 3) [13–16]. G-rich sequences in telomeres and in the untranslated regions (UTRs) close to genes form G-quadruplex structures with the help of metal ions (see Chapter 4), for which regulatory roles in replication, transcription, and translation are suspected [17]. Similarly, a Mg^{2+}-sensing riboswitch in the 5'UTR regulates gene expression connected to metal homeostasis [18,19]. Apart from their structural role, coordinated metal ions are also often closely linked to the catalytic function of ribozymes. While it is not easy and not always possible to separate their structural and catalytic roles, metal ions have been identified in the active sites of most studied ribozymes (see e.g., [20–23]) and DNAzymes (see Chapter 8). Furthermore, they may be directly or indirectly involved in catalysis. The limited diversity of functional groups present in the four nucleotide building blocks of nucleic acids can be extended through the influence of metal ions [24], which can lead to a redistribution of electron densities, the shifting of pK_a values [25–30], the favoring of transition state geometries or the stabilization of rare tautomers [29,31].

It becomes thus evident how important a comprehensive knowledge is about the architectural, thermodynamic, and kinetic properties of metal ion-nucleic acid

interactions for the understanding of the structure and function of nucleic acids. After a short introduction of the key players, the involved metal ions (Section 2) and the nucleic acid ligating groups (Section 3), various methods, that have been employed in the characterization of coordination sites, binding kinetics, and binding affinities will be summarized and discussed (Sections 4, 5, and 6).

2 Ligating Sites for Metal Ions in Nucleic Acids

Metal ion coordination is influenced to a good part by the relative softness/hardness of potential metal ions and ligands, a concept that holds qualitative information on the electronegativity and polarizability of the groups [32,33]. Hard metal ions find suitable binding sites in the negatively charged non-bridging phosphate oxygens of the backbone and also in the exocyclic nucleobase oxygens [34], while the endocyclic nitrogens of the nucleobases are a preferred target of softer metal ions (Figure 1) [8,34–38]. Also available are the bridging phosphate oxygens and the 2'-hydroxyl groups of sugar moieties [39,40]. The exocyclic amino groups of the nucleobases are usually not suitable liganding sites because of the delocalization of the lone pair into the aromatic ring.

Steric factors also play a role, since especially in double-stranded helical regions some functional groups become inaccessible. In addition, the high rotameric freedom of the sugar-phosphate backbone, which in principle permits the formation of bi- and tridentate macrochelates, is more restrained in double-stranded regions.

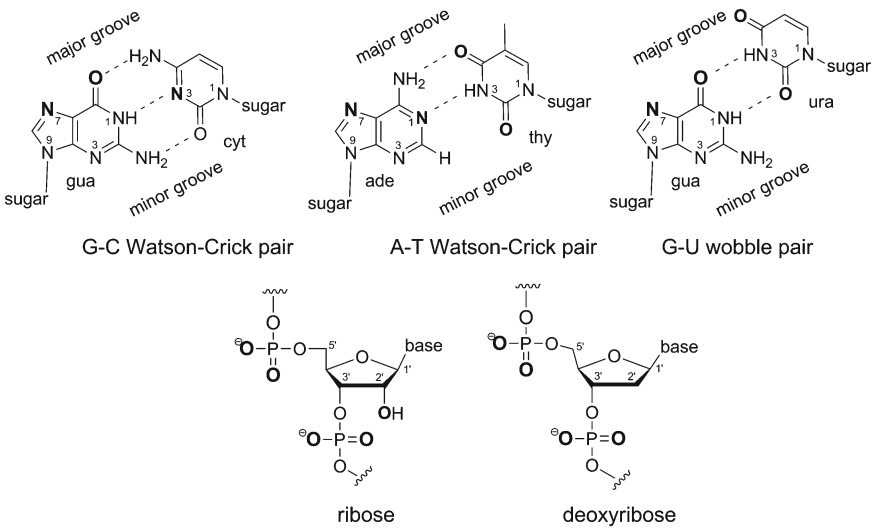


Figure 1 The five most common nucleobases in RNA and DNA: Guanine (gua/G), cytosine (cyt/C), adenine (ade/A), thymine (thy/T) and uracil (ura/U) are drawn as present in canonical Watson-Crick or wobble base pairs (top). The phosphate sugar backbone of DNA and RNA is depicted below. The most common metal ion coordination sites are indicated by bold letters.

3 Metal Ions to Be Considered to Interact with Nucleic Acids

In the cell, kinetically labile metal ion-nucleic acid complexes predominate, which is to a large part due to the nature of the metal ions that are most abundant and freely available in the cellular environment (Table 1) [41–44]. Kinetically stable complexes on the other hand are employed in contexts procured by humans, like the administration of drugs or the development of nano-devices.

3.1 Natural Metal Cofactors

3.1.1 Dominance of Mg^{2+}

Mg^{2+} is the most important cofactor of nucleic acids inside the cell [24,37,45–49]. It is also the most abundant divalent metal ion inside the cell (Table 1), but this is not the only factor that predestines it for its role in nucleic acid structure and function. Mg^{2+} is a relatively hard Lewis acid with a small ionic radius and correspondingly a high charge density. It has a strong preference for oxygen ligands, especially charged ones. In octahedral complexes the distances between the six oxygen ligands are close to their van-der-Waals radii, resulting in an exceptional stability [50]. This leads to reduced solvent exchange rates and increased solvation enthalpy, but at the same time also makes it a most suitable interaction partner for the non-bridging phosphoryl oxygens of the nucleic acid backbone, allowing them to pack particularly closely [51], which is a prerequisite for the physiological folding of large RNAs. In accordance, a nearly perfect correlation between the cleavage rate of the hammerhead ribozyme [52,53] and the metal ion affinity towards phosphate monoesters can be established [34]. Nevertheless, such a correlation cannot be established, for example, for the *glmS* ribozyme [54], indicating a crucial influence of further ligand sites.

Partial dehydration of Mg^{2+} is enthalpically costly and so direct coordination contacts to nucleic acids are only favorable when more than one ligand in a suitable geometry replaces the water molecules [55,56]. Therefore, it is far more common that Mg^{2+} interacts through outersphere contacts of its first hydration shell, especially in the case of the nucleobase sites like the endocyclic nitrogens and exocyclic oxygens [57].

3.1.2 Monovalent Ions

In terms of abundance, Mg^{2+} is surpassed by the monovalent K^+, which is present in the cell at 140 mM, but also by Na^+ present at 10 mM concentration, a value comparable to that of Mg^{2+}. Outside the cell Na^+ and K^+ concentrations are reversed (Table 1). These monovalent ions are far less apt to establish a close packing of the negative charges of the nucleic acid backbone. Usually a largely unspecific role as general charge screeners is attributed to K^+ and Na^+, even if there are also examples of specific monovalent ion binding sites (Figure 2) [58,59].

Table 1 Concentration of various metal ions in the ocean, mammalian cells, the extracellular space, as well as bacterial cells. The respective literature source is always given in the title row.

M^{n+}	Seawater [42] [ppm]	Seawater [43] [mM]	Mammalian cell [43] [mM]	Extracellular space [43] [mM]	Bacterial cells [44] [mg/kg]	Bacterial cytosol [41] [mM]
Li^+	0.17					
Na^+	$1.1 \cdot 10^4$	470	10	145	$4.6 \cdot 10^3$	
K^+	$3.9 \cdot 10^2$	10	140	5	$115 \cdot 10^3$	> 10
Rb^+	90					
Cs^+	3					
Be^{2+}	$6 \cdot 10^{-7}$					
Mg^{2+}	$1.35 \cdot 10^3$	50	30[a]	1	$7 \cdot 10^3$	> 10
Ca^{2+}	$4.1 \cdot 10^2$	10	1	4	$5.1 \cdot 10^3$	0.1
Sr^{2+}	375					
Ba^{2+}	425					
Cr^{2+}	$5 \cdot 10^{-4}$				4	
Mn^{2+}	$2 \cdot 10^{-3}$				260	10^{-2}
Fe^{2+}	$3 \cdot 10^{-3}$	$1 \cdot 10^{-4}$			170	0.1
Co^{2+}	$4 \cdot 10^{-4}$	$3.1 \cdot 10^{-6}$			7.5	low
Ni^{2+}	$7 \cdot 10^{-3}$	$1 \cdot 10^{-6}$				low
Cu^+						10^{-2}
Cu^{2+}	$3 \cdot 10^{-3}$	$5 \cdot 10^{-6}$ [b]			150	
Zn^{2+}	$1 \cdot 10^{-2}$	$1 \cdot 10^{-4}$			83	0.1
Ru^{3+}	$1 \cdot 10^{-2}$					
Pd^{2+}	$1 \cdot 10^{-2}$					
Ag^+	$7 \cdot 10^{-2}$					
Cd^{2+}	0.2				31	
Pt^{2+}	$1 \cdot 10^{-2}$					
Au^{3+}	$4 \cdot 10^{-3}$					
Hg^{2+}	$8 \cdot 10^{-2}$					
Tl^+	0.5					
Pb^{2+}	13					

[a] The free concentration of Mg^{2+} is about 1 mM.

[b] Taken from ref. [345].

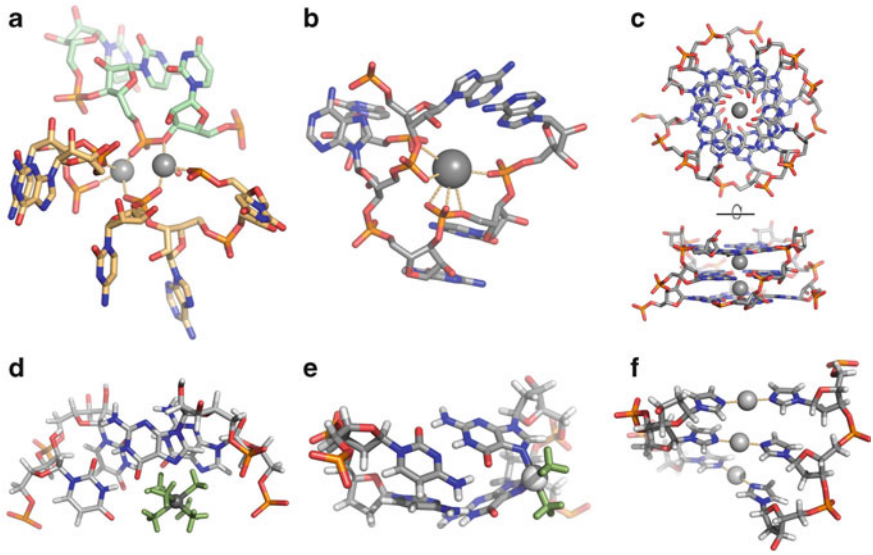


Figure 2 Views of different metal ion binding sites in nucleic acids. (**a**) Two Mg^{2+} ions in the active site of a group II intron (light orange) coordinate the scissile bond of the substrate oligonucleotide (light green) (PDB ID 3G78; [341]). (**b**) Buried K^+ ion in its binding pocket as found in the crystal structure of a 58 nt long rRNA fragment (PDB ID 1HC8; [342]). (**c**) Top and side view of a G-quadruplex structure with two K^+ ions in its central channel (PDB ID 3IBK; [343]). (**d**) $[Co(NH_3)_6]^{3+}$ bound to the major groove of two G·U pairs, the amine ligands are shown in light green (PDB ID 1AJF; [225]). (**e**) Intrastrand *cis*-Pt(II) adduct with the Pt^{2+} coordinated to two guanine-N7 sites and two NH_3 ligands shown in light green (PDB ID 1A84; [344]). (**f**) Three consecutive Ag^+-mediated imidazole base pairs (PDB ID 2KE8; [118]). All panels were prepared using pymol (http://www.pymol.org) and the indicated PDB IDs.

3.1.3 Influences of Metal Ions Other than Mg^{2+} and K^+

Na^+ and Ca^{2+} levels in the cell are lower than those of Mg^{2+} and K^+. Na^+ is expected to share the task of unspecific charge screening together with K^+. Ca^{2+} is used as a signaling molecule in the cell and therefore it is tightly regulated, e.g., by the fast uptake into the endoplasmic reticulum (ER) and mitochondria [60]. Tightly controlled are also typical protein metal cofactors like $Fe^{2+/3+}$, Zn^{2+}, and $Cu^{+/2+}$ [41,61,62], with significant free cellular concentrations only in case of intoxication, disease or drug treatment [63]. Nonetheless, also low levels of a metal ion are able to influence nucleic acids if binding specificity and affinity are favorable, as is shown for example by the allosteric inhibition of a group II intron when 5% Ca^{2+} compete against 95% Mg^{2+} [64]. In the same way other low-abundance metal ions might be involved in the regulation of RNA function [34]. In the following, we will shortly mention some relevant metal ions that are frequently studied in association with nucleic acids, highlighting their distinctive properties with respect to Mg^{2+} or K^+. Those latter two unfortunately are hard to detect and to characterize by most methods due to their low molecular weight and

Table 2 Collection of some physicochemical properties of metal ions, i.e., ionic radius, preferred coordination number (CN), water exchange rate from the first hydration shell (k_{ex} at 298 K), $pK_{a,aq}$ values of the bound water molecules, Pearson hardness (η), free enthalpy of hydration (ΔG_{hydr}), and for coordination number 6 also $\Delta G_{hydrCN6}$. The respective literature sources for the values are given in the top row.

M^{n+}	Ionic radius[a] [Å] [68]	Coordination numbers[b] [43,69]	Preferred ligands [10,43]	k_{ex} [s^{-1}] [70]	$pK_{a,aq}$ [71]	Pearson hardness [72]	ΔG_{hydr} [a] [kJ/mol] [73]	$\Delta G_{hydrCN6}$ [kJ/mol] [74]
Li^+	0.59	4		~10^9	13.64	35.1	–438	–509
Na^+	1.02 (1.18)	6 (8)		~10^9	14.18	21.1	–412 (–383)	–404
K^+	1.38 (1.51)	6 (8)		~10^9	14.46	13.6	–322 (–308)	–330
Rb^+	1.52, 1.61	6, 8		~10^9		11.6	–292, –279	–309
Cs^+	(1.67) 1.74	(6) 8		~10^9		10.6	(–269) –256	–287
Be^{2+}	0.27 (0.45)	4 (6)		730 (CN=4)	5.4	67.8	–2381 (–2218)	–2445
Mg^{2+}	0.72	6	O	5.3-6.6 · 10^5	11.44	32.6	–1858	–1883
Ca^{2+}	(1.06) 1.12	(6) 8	O	~10^9	12.85	19.5	(–1569) –1657	–1558
Sr^{2+}	(1.18) 1.26	(6) 8	O	~10^9	13.29	16.3	(–1385) –1460	–1452
Ba^{2+}	(1.35) 1.42	(6) 8	O	~10^9	13.47		(–1289) –1360	–1289
Cr^{2+}	0.80	6		~10^9		7.2		–1901
Cr^{3+}	0.62	6	O/N	2.4 · 10^{-6}	4	9.1	–4523	–4492
Mn^{2+}	0.83	6	O/N	2.1 · 10^7	10.59	9.0	–1761	–1821
Fe^{2+}	0.61, 0.78	ls 6, hs 6		4.39 · 10^6	9.5	7.2	–1870	–1911
Fe^{3+}	0.55, 0.65	ls 6, hs 6		1.6 · 10^2 / 10^{-3} [c]	0.7 (ls) / 2.19 (hs) [e]	12.1	–4469	–4352
Co^{2+}	0.58, 0.75	4, 6	O/N/S	3.18 · 10^6 (CN=6)	9.65	8.2	–1707, –1887	–1964
Co^{3+}	0.55, 0.61	ls 6, hs 6	O/N	~10^{-1} [e]		8.9		–4572
Ni^{2+}	0.55, 0.49	th 4, sq 4	O/N/S	3.15 · 10^4 (CN=6)	9.86	8.5	–1757 (–1950)	–2045
Cu^+	0.77	6				6.3		–562
Cu^{2+}	0.73	6		4.4 · 10^9	>8	8.3		–2059

Zn^{2+}	0.60, 0.74	4, 6	O/N/S	$4.1 \cdot 10^{8}$ [f]	8.96	10.9	–1690, –1870	–2007
Ru^{3+}	0.68	6		$3.5 \cdot 10^{-6}$		10.7		
Pd^{2+}	0.64	sq 4		560	2.3 [c]	6.8		–1967
Ag^{+}	0.67, 1.02	2, sq 4			12.0	7.0	–335, –381	–468
Cd^{2+}	0.78, 0.95	4, 6	O/N/S	$6.1 \cdot 10^{8}$ [f]	10.08	10.3	–1489, –1640	–1780
Pt^{2+}	0.60	4		$7.1 \cdot 10^{-6}$	~2.5 [c]	8.0		–2059
Au^{3+}	0.68	4				8.4		–4353
Hg^{2+}	1.02	6		$1.8 \cdot 10^{9}$ [f]	3.4	7.7	–1598	–1809
Tl^{+}	1.50	6			13.21	7.2	–329	–337
Pb^{2+}	1.19 (1.29)	6 (8)	O		7.71	8.5	–1431 (–1523)	–1477
Eu^{3+}	1.07, 1.12	8, 9 [d]	O		7.8		–3547 [g]	
Tb^{3+}	1.04	8 [d]	O	$5.6 \cdot 10^{-6}$ (CN=8)	7.9		–3605 [g]	

[a] The value in parentheses refers to the respective coordination number also given in parentheses in column 3. [b] If appropriate, low spin (ls) and high spin (hs) is distinguished. In the case of CN = 4, the square planer (sq) or tetrahedral (th) geometry is indicated. [c] Values taken from [69]. [d] Values taken from [75]. [e] Values taken from ref [76]. [f] Values taken from ref [77]. [g] Values taken from [78].

spectroscopic silence and hence, a variety of metal ions has been used as substitutes in different settings.

The interactions of metal ions in general with nucleic acids are governed by complex stabilities according to the Irving-Williams series [65] and Martin's Stability Ruler [66,67], by the relative affinities for oxygen versus nitrogen ligands [8] and by first shell hydration enthalpies (Table 2) [10,68–78]. Ca^{2+}, for example, shares the preference of Mg^{2+} for phosphate oxygen ligands, but is larger in size and thus compatible with higher coordination numbers and a more loose packing of ligands. Ca^{2+} can inhibit Mg^{2+}-dependent enzymes and ribozymes (e.g., [64,79–82]). Li^{+} has a charge density and ligand preference comparable to Mg^{2+}, but favors tetragonal geometries. Mn^{2+} on the other hand, which is very similar to Mg^{2+} regarding both size and preferred coordination geometry, has a more balanced affinity for nitrogen and oxygen ligands and a ligand exchange rate 100 times faster than Mg^{2+} (Table 2). It is used as a paramagnetic probe in EPR as well as NMR experiments. Cd^{2+} is a much softer metal ion and therefore much more likely to interact through innersphere contacts with nucleobase nitrogens. In addition, it also has a higher inclination to form macrochelates than the other metal ions mentioned so far [37,57,83]. Cd^{2+} is significantly larger than Mg^{2+}, but its thiophilicity [84–86] makes it a popular choice for metal-rescue experiments with phosphorothioates despite a serious caveat [86]. Zn^{2+} is also very thiophilic and more similar in size to Mg^{2+}, but engages in variable coordination geometries [87]. Tl^{+} has been employed as a substitute for K^{+}, making use either of its thiophilicity or of the abundant spin 1/2 nucleus $^{205}Tl^{+}$. The kinetically inert complex $[Co(NH_3)_6]^{3+}$ can be of help to probe $[Mg(H_2O)_6]^{2+}$ binding sites, even if binding affinity and hydrogen bonding capacities differ because of the higher charge and the nature of the ammonium ligands, respectively. $[Co(NH_3)_6]^{3+}$ is, e.g., found to bind to the major groove of tandem G·U wobble pairs (Figure 2d).

Metal ions other than Mg^{2+} have been shown to support catalysis of natural ribozymes, some even at increased reaction rates, like Mn^{2+}, Co^{2+}, Zn^{2+}, and Cd^{2+} in the case of the RzB and *Schistosoma* hammerhead ribozymes [52,53] or Ca^{2+} in the case of the antigenomic form of the hepatitis delta virus (HDV) ribozyme [88]. While in the case of the two Hammerhead ribozymes, the observed cleavage rates correlate with the respective phosphate affinity of the metal ion applied [34], in the case of the *glmS* ribozyme, no such correlation is observed [54]. Although such "more exotic" metal ions do not occur freely in the cell and are usually employed under *in vitro* conditions above their physiological abundances, locally or in secluded compartments of the cell, metal ions like Zn^{2+} might also *in vivo* reach concentrations where the described effects become significant.

3.2 Expanding the Natural Metal Repertoire of Catalytic RNAs and DNAs

While on the one hand ribozyme catalysis is in many cases promoted by a variety of divalent metal ions or even monovalent ions alone, on the other hand so far only a

limited subset of metal ions have been shown to be specifically required in naturally occurring ribozymes, a fact that seems surprising in light of the diverse metal ion binding capabilities and selectivities of various *in vitro* selected ribozymes and DNAzymes. The latter have been shown to discriminate against Mg^{2+}, while selecting for Ca^{2+} [89–91], Cu^{2+} [92–94], Co^{2+} [95], Zn^{2+} [96], Mn^{2+} [94], Pb^{2+} [97], Ni^{2+} [98] or a small subgroup of transition metal ions [99]. By the same means extraordinarily selective DNA biosensors for various metal ions have been engineered [100,101] (see also Chapter 8).

3.3 Kinetically Stable Metal Ion Complexes

In Nature, the structural integrity and function of nucleic acids is maintained by metal ions with intermediate to fast water exchange rates (Table 2). Consequently, these ions also form kinetically rather labile complexes with RNA and DNA. The action of metal-based drugs, on the other hand, usually depends on the irreversible association of metal ion complexes to DNA. Following the historical discovery of the anticancer properties of *Cisplatin, cis*-$(NH_3)_2PtCl_2$ (Figure 2e) [102], many other Pt^{2+}, but also Pt^{4+}, Ru^{3+}, Ga^{3+} and Ti^{4+} complexes have been investigated (see also Chapters 2 and 7) with regard to their anticancer activity [103].

3.4 Metallated Nucleic Acids for Nanotechnology

DNA, as a self-assembling molecule, has long been of interest for nanotechnological applications. The addition of metal ions can switch nucleic acids between two conformations (e.g., from single- to double-strand, or from hairpin to duplex), which has been taken advantage of for designing biosensors for metal ions (e.g., [104–107] and Chapter 8) and recently even logical AND and OR gates [108]. Furthermore, metal ions can modify the physico-chemical properties of nucleic acids and thereby extend their natural functional repertoire. While DNA alone possesses only marginal conductivity, one idea is that the incorporation of metal ions to natural or artificial nucleotides in double-stranded nucleic acid structures might yield useful molecular wires or magnets [109].

Incorporation of metal ions such as Zn^{2+}, Co^{2+}, or Ni^{2+} into DNA yielding so-called M-DNA has been proposed to make DNA more conductive [110–112]. There is still some controversy about this [113] as well as about the definite structure of M-DNA and the situation of metals therein [114–116].

In contrast, the structure of Hg^{2+}- and Ag^{+}-mediated base pairs, where the metal replaces the imino protons at the N3 position in U-U/T-T base pairs or stabilizes C-C mismatches, respectively, has been well established by NMR measurements and UV melting studies [104,117–119]. The binding of an Au^{3+} ion in the center of a G-C base pair has been observed in the crystal structure of an RNA [120].

The incorporation of artificial nucleotides that can serve as metal ion binding sites can help to fine-tailor the conformation and stabilities of nucleic acid structures. For instance, Müller, Sigel, and coworkers have demonstrated the incorporation of Ag^+ ions into imidazole and triazole base pairs (Figure 2f) [118,121]. Clearly, with artificial nucleotides one can also further extend the range of possible metal binding sites. Examples are Cu^{2+}-hydroxypyridone base pairs [122] and salen metal base pairs (Cu^{2+}, $Mn^{2+/3+}$, Fe^{3+}, Ni^{2+}, and VO^{2+}) [123,124]. In a few cases artificial metal ion binding nucleotides have also been incorporated in polynucleotides with modified backbones like PNA (peptide nucleic acid; see also Chapter 12) and GNA (glycol nucleic acid) (see [125] and references therein). For a more detailed discussion on artificial base pairs we refer the reader to Chapters 10 and 11 of this volume and recent reviews by Müller and Clever and Shionoya [125,126].

4 Structural Characterization of Metal Ion Binding Sites

As summarized in the above section, the range of metal ions that have been observed to interact with nucleic acids in one or another specific context is very wide. The structural characterization of their binding sites, their precise localization, and the number and type of associated ligands in solution have been studied by a combination of chemical, biochemical and spectroscopic methods that will be discussed in the following.

4.1 Chemical and Biochemical Methods

Well-established biochemical and chemical methods are used to map metal ion binding sites through backbone cleavage. However, information on structural features of the interaction sites can be inferred only indirectly by studying the effect on ribozyme kinetics or conformation of mutated functional groups or the substitution of hard ligands or metals by soft ones. Nonetheless, many metal ion binding sites could be predicted by these methods, which only later on were confirmed in X-ray crystal structures. The strength of chemical and biochemical methods lies in the identification and characterization of catalytically important metal ion binding sites in ribozymes and their applicability also to very large molecules.

4.1.1 Metal Ion-Induced Hydrolytic Cleavage

Many metal ions catalyze the cleavage of phosphodiester bonds. The cleavage pattern in a large RNA thus holds information about the location of metal ion binding sites. The cleavage depends on local geometry and is mainly thought to take place through an “in-line” nucleophilic attack of the metal at a 2’-hydroxyl group

that is followed by transesterification (Figure 3) [127,128]. A 2',3'-cyclic phosphate and a 5′-hydroxyl group result of this cleavage, analogous to the reaction catalyzed by the small phosphodiester-cleaving ribozymes. The cleavage rates have been shown to be pH dependent, thus indicating the involvement of the metal hydroxides [129]. Mg^{2+} has a very low cleavage capacity at neutral pH (the pK_a of $Mg(H_2O)_6^{2+}$ is 11.4; Table 2) and consequently, transition metal and lanthanide ions are normally used as probes. Pb^{2+} is much more efficient than Zn^{2+} and Mn^{2+} [130], but all three have been used to localize metal-RNA interaction sites [131–134]. Lanthanide(III) ions are especially suitable and employed frequently as probes because their charge makes them bind rather tightly and pK_a values close to neutral facilitate efficient cleavage [135–137].

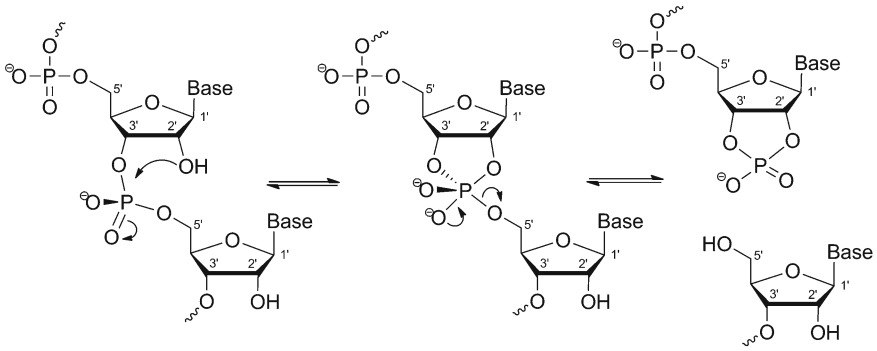


Figure 3 Mechanism of RNA cleavage by a classical in-line attack. The nucleophillic 2'-OH attacks the adjacent phosphodiester and the opposite 5'-OH is liberated upon formation of 2',3'-cyclic phosphate in this S_{N2} reaction. The attacking nucleophile and the leaving group are positioned ideally in a 180° angle to achieve a maximum cleavage rate.

Cleavage assays require 5'- or 3'-^{32}P-end-labeling of the RNA. The cleavage products are separated by denaturing PAGE and the cleavage patterns subsequently used to determine the binding sites. Conclusions, however, are not straightforward because different factors determine the sites of cleavage. Firstly, favorable backbone geometries for cleavage are more common in single-stranded and loop regions than in double-stranded helices [128,138]. On the other hand, strong relevant binding sites in a more rigid environment with an unfavorable geometry for in-line attack might be missed. Secondly, while lanthanide and transition metal ions often seem to bind at the same sites as Mg^{2+} [136], they still are able to adopt diverse coordination geometries or have divergent ligand preferences. Hence, one might miss the most specific binding sites while non-Mg^{2+} binding sites might be detected. Competition experiments with Mg^{2+} can help to identify physiologically relevant binding sites, but they are not able to distinguish effects of structural changes induced by a Mg^{2+} binding non-competitively at a different site from effective lanthanide displacement by Mg^{2+}. Cleavage experiments are thus prone to false negative as well as false positive results, but in combination with other experiments can give valuable information.

4.1.2 Metal Ion-Induced Radical Cleavage

Alternatively, metal ion binding sites can be probed with Fe^{2+} [139], taking advantage of the Fenton reaction [140,141]. In the presence of H_2O_2 Fe^{2+} is oxidized to Fe^{3+} generating short-lived hydroxyl radicals that will lead to backbone cleavage in the near proximity of the $Fe^{2+/3+}$ binding sites. Fe^{3+} is then reduced again to Fe^{2+} by sodium ascorbate. Fe^{2+} compares well to Mg^{2+} in terms of radius and coordination geometry (Table 2) [142] and can compete for Mg^{2+} binding sites [139]. However, it has to be kept in mind that their different Pearson hardness will in many cases lead to divergent binding preferences.

4.1.3 Mutational Approaches to Determine Metal Ligands

Once putative binding sites have been determined, various nucleotide analogs [143] can help to identify more details about a coordination site. Functional groups of a specific nucleobase, sugar or phosphate moiety are modified or removed and the effect can be monitored, e.g., by looking at the metal ion cleavage pattern. Alternatively, the metal ion binding sites that are important for catalysis or essential for structural integrity can be inferred from the catalytic competence or overall conformation of the mutated nucleic acid as determined in enzymatic or electrophoretic mobility shift assays, respectively.

Nucleotide analogue interference mapping (NAIM) [144] is an efficient way to systematically probe the importance of RNA functional groups [10]. In combination with metal ion switch experiments this method is also useful in establishing putative metal ion binding sites (see Figure 4 and Section 4.1.4.) [145–148]. T7 RNA polymerase is used to randomly introduce phosphorothioate nucleotides or nucleotide analogs (NTPαS) at levels of less than one substitution per transcript. Afterwards those mutants, in which activity or another screenable characteristic has been abolished, are separated from the pool of transcripts. Iodine cleavage and subsequent analysis on the gel reveal the positions of the mutated nucleotides [144].

The method can, for example, be used to identify metal-coordinating phosphate groups through metal rescue as described below, with the restriction that only R_P phosphorothioates can be studied as T7 polymerase selects against the other diastereomer [149]. Basu et al. have used a combination of X-ray crystallography and NAIM to identify a monovalent metal ion binding site in the *Tetrahymena* ribozyme P4-P6 domain [58]. They showed the rescue of a 6-thioguanosine mutation by Tl^+, a thiophilic K^+ substitute.

4.1.4 Metal Ion Switch Experiments

The divergent ligand preferences of different kinds of metal ions are the basis of metal ion switch experiments. Mg^{2+} and Mn^{2+} prefer the harder oxygen ligands, while Cd^{2+}, Zn^{2+} and Pb^{2+} display a pronounced preference for the softer aromatic-nitrogen

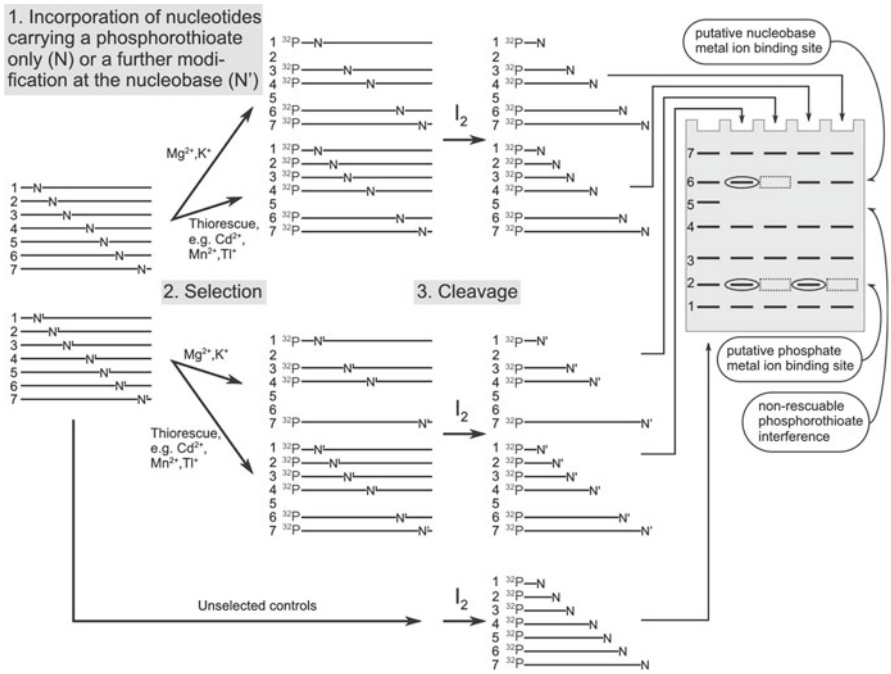


Figure 4 The use of NAIM analysis combined with thiorescue to identify metal ion binding sites in RNA. The example shown visualizes a phosphorothioate interference at position 2 and a nucleobase mutation at position 6 that both can be rescued by a thiophilic metal. The introduced thiophosphate group at position 5 cannot be rescued by the addition of a more thiophilic metal ion, indicating that this group undergoes another kind of crucial (but unknown) interaction within the three-dimensional architecture of the RNA. This Figure is adapted from [144].

and especially sulfur sites [84–86,150,151]. Mutations that exchange oxygen for sulfur or nitrogen groups can significantly suppress Mg^{2+} binding, but the adverse effects can often be rescued by Cd^{2+} or Mn^{2+}. The method is especially useful in identifying catalytically involved metal ions (see also recent reviews [152] and [10]) and is usually thought to be limited to the analysis of directly coordinated metal ions. However, Basu and Strobel report the rescue of outerspherely coordinated binding sites by Mn^{2+} [153]. Since Mg^{2+} preferentially binds to phosphoryl oxygens, the most common mutation in this kind of experiment is the exchange of phosphates for phosphorothioates, but also 2'-OH substitution by NH_2 has been used [154].

Evaluation of metal rescue experiments is complicated by several possible adverse effects. Firstly, there are the inherent physico-chemical differences (e.g., size, hydrogen bonding, coordination geometry, electronegativity) not only of the switched metal ions but also of the switched ligands. Sulfur, for example, is larger than oxygen and thus the introduction of a phosphorothioate by itself can already

alter local structure and/or distort a metal ion binding site. Secondly, ribozymes usually present many other possible interaction sites for the rescuing metal, at which binding can influence the structure and function significantly. To exclude them as well as possible, experiments are normally performed against a background of Mg^{2+}. Evaluation is usually based on relative rate constants that take into account the catalytic rates of all four combinations of the unmodified and modified ribozyme with Mg^{2+} and with the rescuing metal ion (equation 1) [152,155].

$$k^{\text{rel}} = \left(\frac{k_{M^{2+},S/N}}{k_{Mg^{2+},S/N}}\right)\left(\frac{k_{M^{2+},O}}{k_{Mg^{2+},O}}\right)^{-1} \quad (1)$$

$k_{M^{2+},S/N}$ and $k_{Mg^{2+},S/N}$ are the rate constants of the modified ribozyme in the presence of the substituting metal ion or Mg^{2+}, respectively, and $k_{M^{2+},O}$ and $k_{Mg^{2+},O}$ the corresponding rates in the wild-type ribozyme. Using an approach termed thermodynamic fingerprint analysis (TFA) that combines metal ion rescue with kinetic analysis, Herschlag, Piciriilli, and coworkers were able to establish how many metal ions are involved in catalysis by the *Tetrahymena* group I intron ([154] and more detailed descriptions in [152,156]). By examining the influence of mutations at several sites in the RNA substrate and of the cofactor GTP both in the absence and presence of Mn^{2+}, the authors were able to determine three sites with different Mn^{2+} binding affinities, thereby confirming the presence of three different metal ions. In the case of the hammerhead ribozyme, a thio-rescue experiment gave valuable clues about the structure of the transition state active site: The analysis of enzyme kinetics in single and double mutants in a thio-rescue experiment strongly supports a bridging metal ion between a previously described metal ion binding site and the cleavage site [157].

4.1.5 Kinetic Isotope Effect

Compared to the modifications mentioned in the previous paragraphs, substitution of an atom with a heavier isotope of the same element is a very small, but nonetheless detectable intrusion in a system. The substitution can have an enhancing or decreasing effect on a reaction and thereby reveal information about the transition state structure [158,159]. Recently, this method has been applied to the transition state of RNase P [160,161]. The detected $H_2{}^{18}O$ heavy isotope effect agrees perfectly with the results obtained from the Mg^{2+}-catalyzed cleavage of the model compound 5′-*p*-nitrophenylphosphate (T5PNP). This indicates that the cleavage mechanism in RNase P involves the same transition state as in T5PNP, where the $H_2{}^{18}O$ is directly coordinated to the metal ion [160]. In the future the better availability of specifically ^{18}O-enriched nucleotides might open the field for a wider application [162,163].

4.2 Spectroscopic Methods

4.2.1 Electron Paramagnetic Resonance

Electron paramagnetic resonance (EPR) spectroscopy can be very useful for the characterization of metal ion-nucleic acid interaction in several ways. However, most EPR methods require the sample to be frozen. In solution at room temperature, EPR can provide information on the populations of free and bound metal ions and thereby binding affinities as well as cooperativities but structural details are not accessible [164–166]. At low temperature in a frozen sample, hyperfine interactions between the unpaired electron spin of a paramagnetic metal ion and the spin of the nuclei are able to hint at the coordination environment of the metal ion. Two more advanced applications, ENDOR (electron-nuclear double resonance spectroscopy) and ESEEM (electron spin echo envelope modulation) can even reveal details about the number and type ($^{14}N/^{15}N$, $^{1}H/^{2}H$, ^{31}P) of the immediate coordinating partners within a radius of 6–7 Å. For a more detailed description of the techniques we refer the interested reader to recent expert reviews by DeRose and coworkers [167,168].

Since EPR is limited to paramagnetic metal ions, Mn^{2+} is a popular object of study due to its similarities with Mg^{2+} (Table 2). Both its electron and nuclear spin are 5/2, but degeneration leads to a spectrum with only 6 characteristic main lines. The exchange of metal ion-coordinated water molecules for RNA ligands is accompanied by subtle perturbations of the spectrum. As signals from differently bound metal ions will always overlay, an unambiguous analysis usually requires sample conditions with only one prominent metal ion binding site. A single tightly bound Mn^{2+} ion can, for example, be observed in a background of monovalent ions in the hammerhead ribozyme: ESEEM spectroscopy allowed the precise localization of the metal ion at a site with specifically ^{15}N-labeled guanine, the determination of further liganding sites and also of the hydration level [169]. In a different approach, multiple Mn^{2+} binding sites in the Diels-Alder ribozyme were gradually silenced by Cd^{2+}, allowing thus their individual characterization [166].

4.2.2 Lanthanide(III) Luminescence

The luminescence of lanthanide(III) metal ions is sensitive to the direct coordination environment and thus can yield information on the metal ion binding pocket. This method is mentioned here for the sake of completeness but we refer the reader to Chapter 6 of this volume for a detailed description.

4.2.3 X-ray Absorption Spectroscopy

The coordination environment of tightly bound transition metal ions can be characterized in detail in solution by X-ray absorption spectroscopy (XAS) methods like

XANES and EXAFS. XANES (X-ray absorption near edge structure) is sensitive to the average oxidation state of the metal in the sample, while EXAFS (extended X-ray absorption fine structure) allows the deduction of the number and type of ligands as well as coordination geometry and metal-ligand atom distances. Detection is possible in dilute solution and metal ions that are classified as spectroscopically silent (like Na^+, K^+, Mg^{2+}, or Ca^{2+}, Cu^+, and Zn^{2+}) are accessible [170].

Nucleic acids are not optimally suited for XAS investigations, considering the predominantly weak interactions and the coexistence of many metal ion binding sites in most constructs. Nonetheless, higher-affinity binding sites, like the ones in the G-quadruplex channel [171] or in a short RNase P helix P4 model [172] can be characterized in remarkable detail, not to forget complexes of kinetically more inert metal ions [173,174].

4.2.4 Vibrational Spectroscopies

Metal ion binding is reflected in changes of the vibrational bands of nucleobase, sugar, and phosphate constituents, which can be measured by infrared (IR) and Raman spectroscopy. While IR depends upon oscillating dipole moments and is influenced by all types of non-symmetrically bonded atoms, Raman signals derive from the inelastic scattering of photons and occur when there are changes in polarizability. The latter are therefore especially sensitive to electron-rich or multiply bonded groups. In addition, water absorbs in the IR range but is not observed in Raman spectra. Thus, the two methods can yield useful complementary information on a fast timescale and in all physical states [175,176]. Discrete vibrational bands for base, sugar and phosphate groups can be observed, albeit only as an average signal of all conformational states in the sample. Isotopic labeling is an approach to partly alleviate this problem [177,178].

FT-IR (Fourier-transform infrared) spectroscopy and Raman spectroscopy have been employed to semiquantitatively follow metal ion-induced conformational changes [13,179,180] and identify primary binding sites for a variety of metal ions in RNA and DNA [181–187]. By Raman spectroscopy the largest metal ion-dependent changes are observed in the phosphodiester signals. The symmetric stretching of non-bridging phosphate oxygens has recently been proposed to contain quantitative information about the degree of innersphere coordinated metal ions in RNA [188]. Christian et al. [188] find that while electrostatic and hydrogen interactions yield only an attenuation, innersphere coordination is accompanied by a shift of the Raman signal, an effect that could in addition also be roughly correlated to the electronegativity and the Pearson hardness of the respective metal ion. Raman spectroscopy in solution requires rather concentrated and highly pure samples, restricting this method to the analysis of smaller constructs. Raman crystallography or microscopy is an approach that can significantly increase signal intensity and reduce background signals. It allowed the identification of inner- and outersphere coordinated metal ions in the HDV ribozyme [189–191].

4.2.5 Nuclear Magnetic Resonance Methods

Nuclear magnetic resonance (NMR) spectroscopy has provided the three-dimensional structures of hundreds of RNA and DNA molecules. Its restriction in terms of molecular size – compared to X-ray crystallography – is outweighed by the singular capacity to reveal not only structure and conformation, but also local and global dynamics of macromolecules in solution in a quantitative manner. The big advantage compared to other spectroscopic methods in solution is the resolution of the individual nuclei at almost every single position in a polynucleotide chain, allowing the simultaneous site-specific characterization of multiple metal ion binding sites.

Most approaches are based on indirect observations of metal ion induced changes in the nucleic acid binding sites. The most abundant isotopes of hydrogen and phosphorus (^{1}H and ^{31}P) both have nuclear spins of 1/2 as it is the case with ^{15}N and ^{13}C. The latter are widely used in isotopically enriched nucleotides instead of the natural isotopes ^{14}N and ^{12}C. But also the direct observation of NMR-active metal nuclei can help in the structural and thermodynamic characterization of metal binding sites [192,193]. Metal ion-nucleic acid interactions can have repercussions on chemical shifts, relaxation properties and scalar couplings of involved NMR active nuclei [194]. In addition, NOEs (nuclear Overhauser effects) to the protons of popular metal ion mimics, $[Co(NH_3)_6]^{3+}$ and NH_4^+, can denote coordination sites [195].

4.2.5.1 Chemical Shift Perturbations:

Chemical shifts are the most straightforwardly measurable factor in a NMR experiment (Figure 5). Binding of metal ions can influence the chemical shift in two ways: Either through direct deshielding of a nucleus or through shielding/deshielding effects upon conformational changes of the neighborhood that result from a binding event. The latter effect is especially prevalent for ^{1}H shifts and reduces the accuracy with which a binding site can be localized, but on the other hand it is also a highly sensitive flag to define the binding pocket. In the case of smaller structures that do not undergo significant conformational changes upon metal ion interaction, ^{1}H shift perturbations can be used as good indicators of metal ion binding sites [6,196,197]. Imino proton shifts are mostly well resolved and can often be monitored satisfactorily in 1D spectra. However, the observable ones are usually part of a Watson-Crick hydrogen bonding pattern and then not in close proximity to the metal ion binding atom. In addition, the chemical shift of imino protons is very sensitive to temperature and accessibility of bulk water, which may falsify results.

In this case, homo- or heteronuclear 2D NMR spectra can often yield more significant shift perturbations from the non-exchangeable protons. Mg^{2+}-induced ^{1}H chemical shift changes are usually not larger than 0.1–0.2 ppm [6,7,197] at limiting metal concentrations of ~10–20 equivalents and perturbations of ^{13}C chemical shifts are similarly small in general [197–200]. Chemical shifts induced by $[Co(NH_3)_6]^{3+}$ are slightly more pronounced due to the higher charge and associated 10 times stronger binding affinity [201]. Unfortunately, exchange rates of

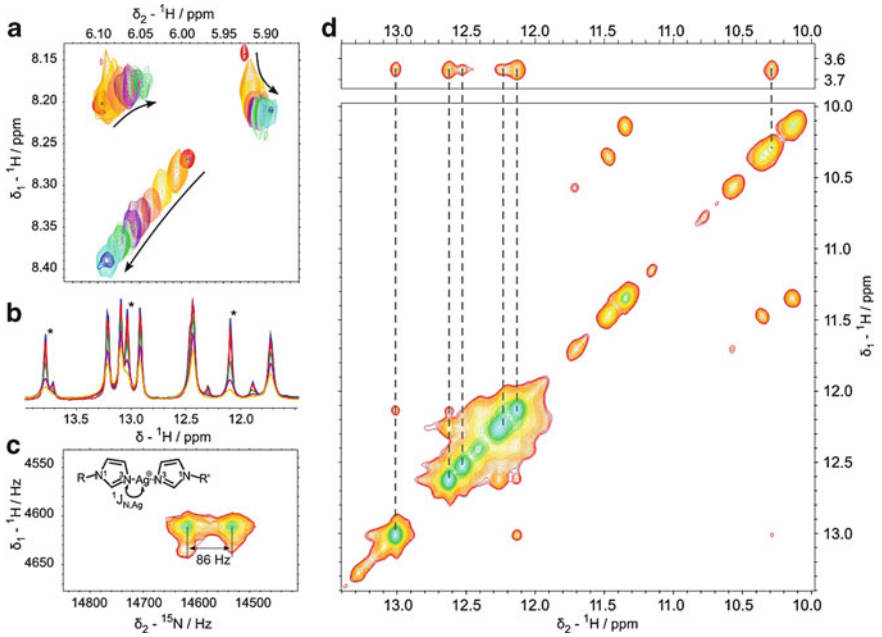


Figure 5 Signatures of metal ion binding to nucleic acids as observed in NMR spectra. (**a**) Mg^{2+} interaction is indicated by chemical shift changes (titration from 0 mM, red, to 12 mM Mg^{2+}, dark blue). Line broadening indicates direct Mg^{2+} coordination. The panel is adapted from [7] and [252]. (**b**) Selective line broadening is observed in a ^{1}H NMR spectrum with increasing amounts of Mn^{2+} (titration from 0 μM, blue, to 210 μM, yellow). (**c**) Coordination of a Ag^{+} ion causes a 86 Hz splitting of the imidazole N3 signal in a [^{1}H,^{15}N]-HSQC. Panel adapted from [118]. (**d**) Binding of $[Co(NH_3)_6]^{3+}$ can be observed by NOE crosspeaks between the protons of its ammine ligands (bulk resonance at 3.65 ppm) and the imino protons of a short RNA hairpin.

commonly employed metal ions are often in the intermediate regime on the NMR time scale [84] leading to a general broadening of the lines already at lower concentrations that can impede the analysis of chemical shift changes.

With ^{31}P, nucleic acids contain a second highly sensitive and abundant spin 1/2 nucleus in addition to protons, whose only disadvantage is the low signal dispersion. Most ^{31}P resonances cluster in the small region between –3 to –1 ppm and only a few with non-standard backbone torsion angles can be resolved from the rest [202]. ^{31}P resonances can be unambiguously assigned by the site-specific incorporation of a non-bridging ^{17}O that efficiently broadens the signal of the adjacent ^{31}P [203], an approach which has been employed, e.g., by Hansen et al. [204] to pinpoint a binding site in a hammerhead ribozyme. Alternatively, ^{31}P resonances from the crowded bulk region can be resolved by the incorporation of phosphorothioates, which have a ^{31}P resonance shifted by ~60 ppm [203,205], allowing to observe effects of metal ion binding to selected sites in the phosphate backbone [206]. Coordination of Cd^{2+} to a phosphorothioate, for example, will lead to an upfield shift of the ^{31}P resonance by a few ppm in both R_P and S_P diastereomers, as has been observed for both the A9/

G10.1 and scissile phosphate locations in the hammerhead ribozyme [207,208] in agreement with the results of biochemical studies by Wang et al. [157]. However, such experiments have to be interpreted with care because the introduction of a sulfur atom might also create a new or specially shifted binding site.

Considering the unfavorable NMR characteristics of oxygen isotopes, ^{15}N is the only nucleus in nucleic acids that can serve as a probe for the detection of an inner-sphere coordination site. In addition, ^{15}N is straightforwardly observable by NMR spectroscopy. This explains why with ^{15}N, compared to ^{1}H, ^{13}C and ^{31}P, much higher chemical shift perturbations can be observed. Direct coordination to nucleobase nitrogens is confirmed in Hg^{2+}-mediated thymine base pairs in DNA by 30 ppm ^{15}N downfield shifts [117] as well as by Ag^{+} coordination to imidazole nucleotides [118]. Also the more labile coordination of Cd^{2+} and Zn^{2+} ions to N7 induces 20 ppm shifts in 1D ^{15}N experiments [209–211]. An initial broadening of the ^{15}N resonance is attributed to the exchange between free and bound states. At higher, saturating Cd^{2+} concentrations the peak grows sharper again. Resonances for the adjacent purine $^{13}C8$ and $^{1}H8$ are perturbed less but still significantly [172], which is not observed in outersphere coordination [212,213]. Since direct observation of the low sensitivity nucleus ^{15}N can be lengthy, the recording of 2D [^{1}H,^{15}N]-HSQC can be advantageous, while at the same time enhancing the resolution. On the basis of ^{2}J-[^{1}H,^{15}N]-HSQC experiments Erat et al. have proposed that the combined information of ^{1}H and ^{15}N shifts and characteristic broadening can be used to identify inner- and outersphere binding events [214]. The empirically collected knowledge agrees well with theoretical calculations, confirming that ^{15}N chemical shifts are a valid means to determine the coordination mode [215,216].

4.2.5.2 Paramagnetic Effects:

The large magnetic moment of an unpaired electron of a paramagnetic metal ion species efficiently relaxes nuclei in its immediate environment (Figure 5b) [172,217,218]. The effect is strictly distance-dependent (relative to r^{-6}) and leaves less room for ambiguity than chemical shift perturbations. Paramagnetic line broadening information can even be included as weak distance restraints in molecular dynamics calculations [197]. Mn^{2+} exchanges very fast between the free and bound form and can therefore effectively relax a whole molecule at substoichiometric concentrations. µM amounts are usually enough to detect site-specific line-broadening effects. Mn^{2+} is most widely used because of its likeness to Mg^{2+} (e.g., [219–221]), but also Co^{2+} and Ni^{2+} are suitable paramagnetic probes [98,222].

4.2.5.3 NOE Crosspeaks to $[Co(NH_3)_6]^{3+}$ and NH_4^+:

$[Co(NH_3)_6]^{3+}$ is a mimic for $[Mg(H_2O)_6]^{2+}$ [223] and has been used in a wide range of studies [172,200,201,224–226]. Similarly, NH_4^+ can be used as a substitute for monovalent metal ions [195]. Both compounds possess protons that are amenable for direct

observation of inter-molecular NOE crosspeaks (Figure 5d) to nucleic acid protons in a radius smaller than 6 Å. In the most common case tumbling of the cation in the binding site and exchange with the unbound state are fast on the chemical shift time scale and all $[Co(NH_3)_6]^{3+}$/NH_4^+ protons resonate at a single frequency. G-quadruplex structures are an exception, displaying a binding site with on-off rates slow enough to allow observation of separate resonances for free and bound ammonium [227,228]. The tight binding even significantly slows down the usually fast proton exchange of ammonium [195]. It is advantageous to employ ^{15}N-labeled NH_4^+ because it not only avoids the quadrupolar ^{14}N nucleus, but also allows for heteronuclear HSQC experiments that detect the exchange of ions between binding sites and the solvent [229]. Structural information, however, can be inferred in both fast and slow exchanging cases. From the NOE cross-peaks of $[Co(NH_3)_6]^{3+}$ or NH_4^+ weak distance restraints can be extracted and integrated in molecular dynamics simulations to infer binding sites [45].

4.2.5.4 Direct Detection of NMR-Active Metal Isotopes:

A good number of the metal ions that have been studied in association with nucleic acids have at least one NMR-active isotope. Unfortunately many of them, including the biologically most relevant $^{23}Na^+$, $^{39}K^+$, and $^{25}Mg^{2+}$, have half-integer spins > 1/2. Fast quadrupolar relaxation usually restricts their use in solution NMR experiments to the study of kinetic and thermodynamic features by line shape analysis (see also Sections 5 and 6) [230–234] and pushes them more into the field of solid state NMR, which is better suited to handle the large quadrupolar effects.

Spin 1/2 isotopes are available in ^{52}Fe, ^{107}Ag, ^{109}Ag, ^{111}Cd, ^{113}Cd, ^{195}Pt, ^{199}Hg, ^{203}Tl, ^{205}Tl, and ^{207}Pb. Many have wide chemical shift ranges and can give useful information about coordination geometry and ligand atom identity, as has been shown in proteins [235,236]. In nucleic acids, however, their use is most often restricted by the problem of weak binding and fast exchange. An exception is again the high-affinity binding of monovalent metal ions to G-quadruplex structures that even permitted the direct observation of separate peaks for the bound and unbound species of $^{23}Na^+$, $^{39}K^+$, and $^{85}Rb^+$, which was not thought possible so far [192,237].

$^{205}Tl^+$ is a very useful NMR substitute for K^+ and Na^+ due to its high natural abundance (70%) and high sensitivity [238]. Basu et al. observe two well separated $^{205}Tl^+$ peaks for the bound and unbound species in a G-quartet [239]. In a subsequent work Gill et al. even managed to measure small scalar ^1H-$^{205}Tl^+$ couplings of < 1 Hz between the ion and imino and aromatic protons, which allow the precise localization of metal ions in a G-quadruplex structure [193].

Such direct couplings could not be observed between $^{113}Cd^{2+}$ and ^{15}N upon coordination of the metal ion to an N7 in the tandem G·A base-pair motif of the hammerhead ribozyme [209,212], which is attributed to the fast exchange of Cd^{2+}. The coordination of Ag^+ ions between imidazole nucleotide analogs in a DNA duplex on the other hand is stable enough to observe a 86 Hz $^1J(^{15}N,^{107/109}Ag)$ coupling indirectly in the splitting of the nitrogen signal (Figure 5c) [118].

5 Determination of Binding Kinetics

Metal ions interact with nucleic acids on a very wide range of time scales. Fast exchange (ms to μs range) dominates the dynamic interactions of the majority of alkali and alkaline earth metal ions. On the other extreme, metal complexes effective in cancer treatment are characterized by particularly slow kinetics that can be in the order of hours or days [240]. Mg^{2+} is special with regard to its water-ligand exchange rate of ~$2 \cdot 10^5$ s^{-1} being four orders of magnitude slower than that of most of the main group metal ions, but also significantly slower than many transition metal ions [241].

5.1 Nuclear Magnetic Resonance

The appropriate approach for establishing binding kinetics is determined by the association-dissociation rate of the metal ion, which can be fast, intermediate or slow on the respective time scale accessible to the employed method. For example, what is considered fast on the NMR (~10^{-1}–10^{-9} s) or EPR (~10^{-4}–10^{-8} s) time scale will still be slow in IR (~10^{-13} s) or UV (~10^{-15} s) spectroscopy [42]. In the case of slow exchange, separate signals can be observed for each exchanging species, while fast exchange only displays one averaged signal (Figure 6). NMR has proven the most versatile tool to determine thermodynamic and kinetic parameters on different time scales. Most of the relevant exchange processes in nucleic acids happen in fast exchange (> 10^3 s^{-1}) relative to the NMR chemical shift time scale. In this regime, the NMR signal has a chemical shift that corresponds to the population-weighted average of the chemical shifts of the bound and unbound forms. A concentration-dependent movement of the signal can be used to deduce affinity constants (see Section 5.2). If the exchange is moderately fast (too fast would be $k_{-1} > 40\,|\delta_{bound} - \delta_{unbound}|$ [242]) a full line-shape analysis can give on- and off-rates of metal binding as well as equilibrium dissociation constants [242,243]. Relaxation rates and line-shapes of the quadrupolar nuclei ^{23}Na, ^{25}Mg, and ^{43}Ca are modulated by their interactions with larger molecules and have been used in a number of studies to characterize binding kinetics, cooperativity and co-solute dependence [230–232,234,244–246].

Magnetic relaxation dispersion (MRD) experiments allow the probing of molecular motions over a wide range of time scales by recording autorelaxation rates of quadrupolar nuclei at various different field strengths [247,248]. In this way the residence times of small molecules like H_2O (through ^2H or ^{17}O) [249] or metal ions (especially well suited is $^{23}Na^+$) on the surface of a biomolecule can be determined under fast exchange conditions. By MRD experiments, Halle and colleagues confirmed the binding of $^{23}Na^+$ in the minor groove of DNA A-tracts by showing that binding is blocked by the minor groove binding drug netropsin [250,251]. Results support the existence of relatively long-lived $^{23}Na^+$ (50 ns [250]; 10 ns to 100 μs [251]) and $^{58}Rb^+$ (0.2 ± 0.1 μs [251]) ions.

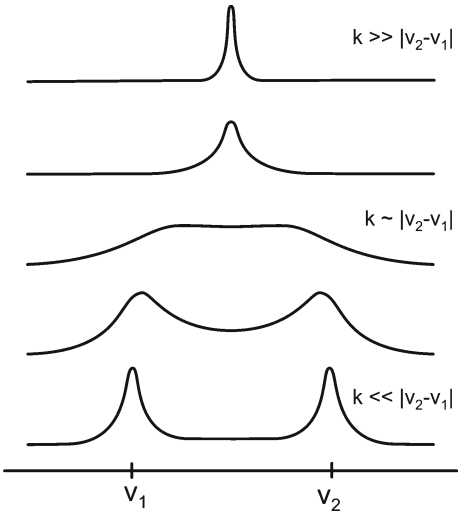


Figure 6 Depending on the exchange rate between two states and the respective resonances in their NMR spectrum, the two peaks fall together (top) or are clearly separated (bottom) as the two extremes.

When NH_4^+ or $[Co(NH_3)_6]^{3+}$ are used as metal ion mimics, cross-relaxation experiments on the basis of the nuclear Overhauser effect (NOESY, ROESY) can be used to observe fast exchange between different binding sites and the solvent [11]. Together, the peak magnitudes and the absence of an observable change in the dominant bulk chemical shift of ammonium protons allow the deduction of upper and lower limits for the exchange rate, but no exact quantification.

Intermediate exchange is characterized by severely broadened resonances. As long as the resonances are still detectable, temperature-dependent changes and lineshape analysis can yield exchange rates. Mg^{2+} interactions with RNA often occur on an intermediate time scale leading to a gradual broadening of all resonances in the course of a metal titration [6,7,10,196,252].

In the region of slow exchange on the chemical shift time scale each exchanging species gives a separate signal that can be monitored individually. Only very few metal binding motifs have been characterized so far that are able to slow down exchange so significantly. Among them are the strong binding sites in the stem of G-quadruplex structures, for which lifetimes can be exceptionally long on the NMR chemical shift time scale [227,228] and which have been studied extensively due to their compact size.

In slowly exchanging systems, the frequency differences between bound and unbound species directly indicate lower limits for residence times, even if the fully bound state is usually not reached. In this way directly observed G-quadruplex channel-bound and free $^{23}Na^+$ and $^{85}Rb^+$ resonances gave conservative lower residence limits of 90 μs and 17 μs, respectively [237]. Similarly, the shifts of the K^+-substitute $^{205}Tl^+$ allows estimation of exchange rates in a G-quartet, where also one especially slowly exchanging ion could be identified through the exceptionally small line width [239]. Slow exchange is also amenable to saturation transfer experiments. Monitoring the efficiency of magnetization transfer from free $^{205}Tl^+$ to bound $^{205}Tl^+$ for varying saturation times, bound lifetimes between 80 ± 10 ms and

155 ± 65 ms for different binding sites on a G-quadruplex were found [193]. Employing $^{15}NH_4^+$, ^{15}N–^1H N_z exchange HSQC experiments at incremented mixing times allow the observation of different ammonia species – free and bound via inner- and outersphere binding in the channel permitting the quantification of the exchange rates between the individual species [229,253,254]. A considerably slower exchange for $^{15}NH_4^+$ than for Na^+ is observed.

The binding rates of the anticancer drug *cis*-diamminedichloroplatinum(II) and its *trans* isomer, however, are on a completely different time scale than any of the slowly-exchanging examples above. To illustrate this: binding kinetics in the range of 10^{-5} s^{-1} are determined from the gradually growing peaks of the mono- and bifunctional Pt^{2+}-DNA adducts over many hours [255].

5.2 Further Methods

NMR is the by far most used method to study metal ion binding kinetics in nucleic acids and there are only few examples where a different approach has led to the determination of association-dissociation rates. Dynamics in the µs-range are fast on the NMR, but not necessarily on the EPR time scale and thus exchange processes that are averaged out in the former, might be resolved in the latter [167].

Labuda and Pörschke have used the intrinsic fluorescence of the wyosine (Wye) nucleobase, which is a natural modification of guanine, to establish binding kinetics of Mg^{2+} and Ca^{2+} to the anticodon loop in tRNA [256]. By a temperature jump method, relaxation parameters in the absence and presence of metal ions were determined. Interestingly, the rate constants for Ca^{2+} binding in general and Mg^{2+} innersphere binding particularly were found to be in the ms range.

Where experimental methods reach their limits in terms of resolution and exchange rates, molecular dynamics (MD) simulations can give complementary information on the metal binding to known nucleic acid structures. However, the force field description of di- and multivalent cations is much more demanding than that of monovalent ions [257]. In addition, at present, simulations are limited to below µs time scales, which is too short to observe, e.g., the dehydration of Mg^{2+} [258]. What is possible in terms of kinetics, however, is the evaluation of the residence times of the more easily dehydrated K^+ and Na^+ with nucleic acid surfaces that happen on the ps to ns time scale [259–262] and for longer-lived bound metal species the estimation of lower boundaries of residence times [263].

6 Determination of Binding Affinities

Two challenges are faced in the thermodynamic analysis of metal ion binding to nucleic acids. Firstly, metal ion binding is inextricably interwoven with the folding and structural stability of nucleic acids and the one cannot be studied without affecting the other [264]. Secondly, usually there are only few specific binding sites with

higher affinity and they coexist with a large background of weak, transient electrostatic interactions that are not easily accounted for in simple binding polynomials [265]. Both problems can sometimes be alleviated by determining apparent affinity constants under moderately high monovalent salt conditions, assuming that then the nucleic acid conformation will not significantly change any more upon addition of the metal ion to be analyzed and that all unspecific interaction sites will be saturated by the monovalent ion.

Metal ion affinities have been inferred from a range of different observables, like thermodynamic stabilities of the nucleic acid fold measured by UV melting curves [266–269] or the thermodynamics of folding of the hammerhead ribozyme by isothermal titration calorimetry (ITC) [270]. Qualitative and quantitative assessment of sequence specific ion binding to DNA A-tracts has been undertaken based on the reduced effective charge that can be detected by free solution capillary chromatography [271,272]. Apparent metal affinity constants have been determined from observed rate constants of ribozyme catalysis in many cases (e.g., [154,157,273–275]). An approach based on the gas-phase fragmentation of metal-nucleic acid complexes in ESI-MS (electrospray ionization mass spectrometry) was recently applied to determine binding affinities to the thrombin binding aptamer [276]. In addition, the ratio of bound and unbound metal ions, as determined by fluorescent indicators, AES (atomic emission spectroscopy), NMR, EPR or lanthanide luminescence, in dependence of metal ion concentration, yields information on binding affinities (see Sections 6.1–6.3).

These various approaches differ in their ability to derive quantitative information, their ability to distinguish between multiple coexisting binding sites and their limits in terms of macromolecule size. All of the above methods yield signals that are an overlay of all the binding sites present in the system. Analysis of the dependence of the bound and free metal ion concentrations on the total metal ion concentration by, e.g., Hill or Scatchard plots can help to disentangle them to some extent to determine classes of binding sites with similar affinities and their occupancies and give information about binding cooperativity. However, one should be aware that reciprocal plots, like Scatchard or Eadie-Hofstee plots can lead to wrong values, as some experimental data points can obtain too much weight depending on their distribution [277]. In addition, this approach is in many cases challenged by the interdependence of RNA folding and metal ion binding [264]. Aside from that, there is usually also a restriction to simpler model molecules with one or only few strong binding sites. While the size restriction is also true for NMR, this method has one advantage over all the others: It allows to separately monitor binding events at several individual binding sites in a molecule and to determine intrinsic binding affinities (see Section 6.3).

6.1 Stoichiometric Methods – "Ion Counting"

So-called "ion counting" methods detect the free metal ion concentration in a sample, from which the number of ions strongly bound to the nucleic acid can be

deduced. This can be useful to establish the number of strong divalent metal ion binding sites in a high-salt background [143], but also to evaluate the affinities of different metal ions with respect to each other. Here, fluorescent indicators have been widely used to determine free metal ion concentrations [278–280]. However, such dyes are usually restricted to the detection of one specific kind of metal ion. Recent methodological progress has made AES a very valuable tool in this regard, able to detect a far wider range of mono- and divalent cations and also anions, thereby more completely accounting for the ion cloud around polynucleotides [281,282]. Irrespective of the way of detection, those methods are based on the equilibration of the nucleic acid-containing solution (by dialysis or ultrafiltration spin columns) against a buffer solution with well defined metal content. After the equilibrium is reached, metal ion concentrations in both solutions are measured and the number of metal ions retained per nucleic acid molecule can be determined [278,282].

6.2 Relative Affinities by Competition Experiments

EPR, lanthanide luminescence or NMR experiments can be used to directly sense the bound metal ions. All these methods, however, are restricted to certain metal ions each: EPR mainly to Mn^{2+}, which can differ significantly from Mg^{2+} in binding affinities for nucleic acid ligands [167], luminescence measurements to lanthanide(III) ions, mainly Tb^{3+}, and NMR to NMR-active nuclei. Competition experiments, where the observable ion is displaced stepwise by increasing concentration of another (silent) metal ion, can yield relative apparent affinities for a wider range of ions [275,283–285]. The same method can also be applied, for example, in Tb^{3+} cleavage experiments: The change in cleavage intensity with increasing amounts of competing Mg^{2+} can in principle be used to calculate Mg^{2+} binding affinity at specific sites [9,10].

6.3 Calculating Site-Specific Intrinsic Binding Affinities from NMR Chemical Shifts

While slow exchange can make things easier in the study of binding kinetics, for the determination of affinity constants by NMR fast exchange can be advantageous. Observable nuclei that are affected by the fast exchange between the metal-bound and metal-free state exhibit one resonance at the population-weighted average position. Titration with a metal ion will shift the free-to-bound equilibrium and thereby the position of the peak. Similarly, the peak width of a resonance will be affected when the titrated metal ion is paramagnetic. If a resonance is influenced only by one single association-dissociation event, the chemical shift change or change in peak width, respectively, plotted against metal ion concentration can be fitted by a 1:1 binding isotherm using e.g. a Levenberg-Marquardt [7,196,286,287]

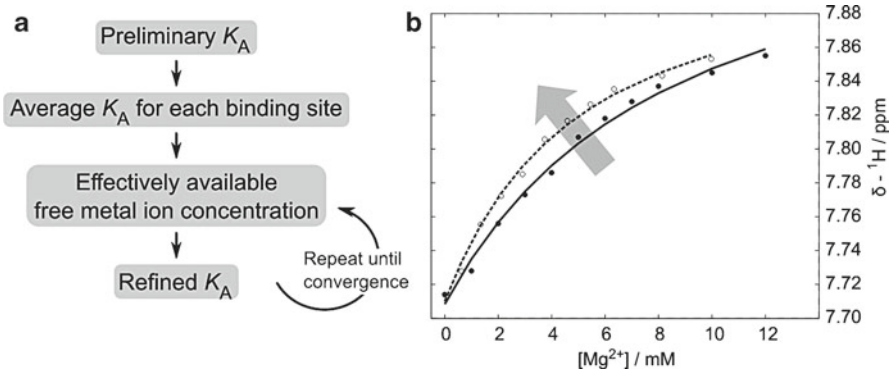


Figure 7 Calculation of intrinsic affinity constants as deduced from chemical shift changes upon Mg^{2+} addition to a large RNA [7,9,10]. (**a**) Scheme of the iterative calculation procedure. (**b**) Fit of the experimental data from H6 of U21 in a group II intron domain 6 after the first and the fifth iteration [7]. The data points shift to the left, as the available Mg^{2+} concentration for each site becomes smaller. Simultaneously the calculated intrinsic affinity constant increases from log K_{A1} = 2.11 ± 0.09 to log K_{A5} = 2.52 ± 0.04 [7].

or a Newton-Gauss [288] algorithm and binding constants can be deduced. This has been widely used for small complexes [25,27,40,86,289–291], as well as for more complicated nucleic acids [7].

However, if there is more than one binding site all having similar affinities, as is usually always the case for nucleic acids, the sites will compete for the metal ions. Therefore the effective metal ion concentration that is available for each binding site at each titration step is lower than what was actually added and calculated affinities will always be underestimated. An iterative approach proposed by Erat, Sigel et al. circumvents this caveat (Figure 7) [7,9,10]. The first step is the grouping of resonances, according to their initial affinity constants and additional information like Mn^{2+} line broadening, into probable binding sites. The average binding constants for those sites are used to determine a refined free metal ion concentration. The newly determined metal concentration is again inserted into the calculation of binding constants. These steps are repeated until the calculated affinities do not significantly change any more.

Proton resonances lend themselves to this kind of study because of their easy and fast acquisition, comparably good signal dispersion in 2D experiments and high sensitivity to changes in their environment. However, there are some factors that can hamper the determination of binding constants in this way. Firstly, bound Mg^{2+} exchanges with the solvent on the intermediate NMR chemical shift time scale and lines can get too broad to follow at higher concentrations. Secondly, the presumption that each binding site is affected by a single binding event is often not true, especially for protons. A second binding event close-by or a structural rearrangement, even if it is small, can make the data unusable. And thirdly, of course, the mapping of binding sites requires that resonances have been assigned.

7 Effects of Other Factors

7.1 *Effect of Anions*

The stabilities of commonly employed metal salts have to be kept in mind when characterizing and quantifying cation-nucleic acid interactions. Adverse effects are possible especially in the high, non-physiological concentrations of certain salts that are part of many setups and can interfere with the collection of quantitative information. For example, the group of DeRose observes significant coordination of Mn^{2+} to Cl^- at high Cl^- concentrations [167]. Similarly, Cd^{2+} is known to form a rather stable $CdCl^+$ complex in aqueous solution: Consequently, an increase in Cd^{2+} affinity towards RNA of 0.8 log unit for log K_A values has been observed for M^{2+} binding to the bulge region within the catalytic domain 5 of a group II intron ribozyme when using 100 mM $KClO_4$ instead of 100 mM KCl [292].

Compared to cations, information on anion binding to nucleic acids is rather scarce, even though they too have been shown to be present in the first coordination shell of nucleic acids. Molecular dynamics simulations and a thorough evaluation of crystal structures [293–295] have identified anions (Cl^-, SO_4^{2-}, ClO_4^-, and PO_4^{3-} as the most frequently employed anions in crystallization buffers and MD simulations). These anions replace water molecules close to nucleobase amino groups as well as to endocyclic nitrogens and 2'-OH groups. The same authors [295] also speculated that anions occupy sites that are also possible binding sites for the negatively charged functional groups of proteins, drugs or phosphate backbone groups from a distant nucleic acid strand.

Anions are also found in close association with metal ions [293,295]. In some cases local arrangements where the anion is situated between two cations suggest an interdependence of binding, where the anion facilitates the close approach of two positive charges (Figure 8) [295]. Close contact of two Mg^{2+} ions is observed in many ribozyme active sites. In some cases, based on abnormally short metal-metal distances and molecular dynamics simulations, the existence of a bridging hydroxide has been suspected [296–298].

Nucleic acids are commonly studied in solutions containing Cl^-, SO_4^{2-}, ClO_4^- or PO_4^{3-}, yet little is known about the differential effects those anions might have on their physical properties. A study of ^{25}Mg relaxation revealed a different behavior in the presence of Cl^- and SO_4^{2-}, which is attributed to a lyotropic effect [234]. A difference between these two anions was also observed in the first step of splicing in a group II intron [299]. While ammonium chloride stimulated branching readily, ammonium sulfate did so only after a long lag phase in reactivity.

7.2 *Effect of Buffers*

Metal ion complexation by a buffer is often neglected, but in fact rather common and can severely hamper the quantitative analysis of binding events (e.g., [300,301]).

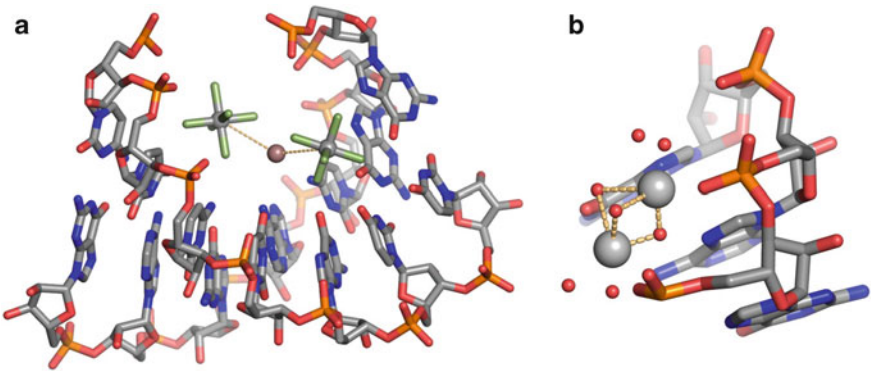


Figure 8 Anions bound to metal ion within nucleic acid structures. (**a**) A chloride ion (pink) bridges two $[Co(NH_3)_6]^{3+}$ complexes in the crystal structure of the *E. coli* signal recognition particle (SRP) RNA (PDB ID 3LQX; [295]). (**b**) A binuclear Mg^{2+} complex in the loop E motif is bridged by three water or hydroxide molecules, respectively (PDB ID 354D; [298]). The panels were prepared using pymol (http://www.pymol.org) and the indicated PDB IDs.

Many buffers commonly employed in biological applications interact significantly with metal ions [302-309]. The preparation of RNA or DNA samples for further experiments usually involves high amounts of Tris (2-amino-2-hydroxymethyl-propane-1,3-diol) and Hepes [N-(2-hydroxyethyl)-piperazine-N′-2-ethanesulfonic acid] [310], e.g., in transcription and electrophoresis running buffers, that have to be removed carefully before any quantitative investigation of metal ion binding is attempted [307] (Figure 9). When applying Mg^{2+}, buffers like Tris or Bistris [bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane] have relatively little influence, but this changes when Ca^{2+} [39,311], 3d metal ions or Cd^{2+} are applied [39,57,307,308], which show an increased affinity, e.g., by about 2 log units for Cd^{2+}, if compared with Mg^{2+} [307]. Phosphate, which is frequently employed as a buffer in spectroscopic or biochemical experiments of nucleic acids, also interacts with metal ions strongly enough to falsify measurements of intrinsic affinity constants [36] or of reaction rates [312–315], an observation that holds for other buffers as well [312,316]. Consequently, one is often forced to work in buffer-free media [312,317,318].

Not only metal ions engage in interactions with buffer molecules, but also nucleic acids have been shown to form complexes with certain buffers, which can lead to complex formation or affect the conformation [319,320]. Thus also the formation of nucleic acid ternary complexes with buffers and metal ions, like they have been observed for nucleotides [57,307,308,311,321] and in human serum albumin [322], is a possibility that should be kept in mind when studying the interactions of metal ions and nucleic acids.

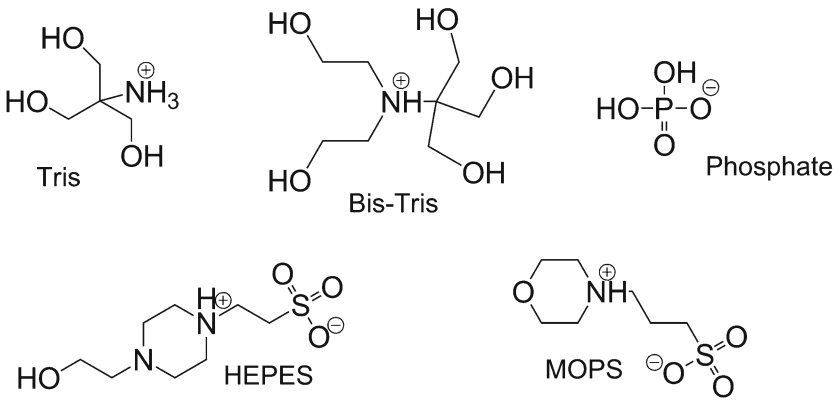


Figure 9 Buffers commonly employed in experiments of nucleic acids. All buffers are shown in their protonated form.

7.3 Effects of Solvent Permittivity and Co-solutes to Mimic Macromolecular Crowding

Two more factors are of importance when evaluating the significance of *in vitro* findings for metal ion nucleic acid interactions of biological systems: The dilute aqueous reaction conditions under which most *in vitro* investigations of nucleic acids take place are quite unrealistic with regard to the molecular crowding present in the cytoplasm and inside cellular organelles [323,324]. Also dielectric constants at the surface and even more in the interior of proteins and nucleic acids differ significantly from the value of about 80 for water. Depending on the local environment values between 20 to 70 can occur [325–330].

Studies in an aqueous solution of 1,4-dioxane have been used to simulate the reduced permittivity of such conditions and found an overall increased Mg^{2+} affinity to a short RNA hairpin [331]. Such an increase was also observed for the binding of Cu^{2+} to single nucleotides in dioxane-enriched solutions [332]. To simulate the crowded cellular environment, Nakano et al. investigated the stability and catalytic activity of a minimal hammerhead ribozyme with PEG8000 as co-solute [333,334]. The authors find that large neutral co-solutes stabilize the tertiary structure and enhance catalytic rates especially at low Mg^{2+} concentrations. In general, molecular crowding is expected to reduce the high concentrations of Mg^{2+} that are usually required for the correct folding of RNA molecules under *in vitro* conditions, when considering the situation in a living cell [335].

8 Concluding Remarks

Compared to the situation in most metalloproteins, metal ion binding to nucleic acids is rather weak and of a highly dynamic nature. The reason for such true binding equilibria are the nature of the ligands – mainly phosphates, aromatic ring nitrogen atoms, and bridging water molecules – and of the metal ions, i.e. alkaline and alkaline earth metal ions. In addition, the metal ions of interest are mostly spectroscopically silent, which makes their detection at the coordinating sites very difficult. Taken together, the detailed characterization of metal ion binding to nucleic acids is rather challenging. In this review we summarize the methods applied nowadays with the aim to understand metal ion binding in solution and compiled the literature for the more interested reader. It is obvious that we are not only far from understanding this fascinating interaction and thus also the folded structure and mechanism of action of nucleic acids, but also that the methods applied carry numerous caveats. Most studies rely on the application of *other* metal ions than, e.g., Mg^{2+}, which automatically means that the system under investigation has been altered. Every kind of metal ion has its distinct coordinating properties and thus also the functional RNA will slightly (or possibly more seriously) change its structure and other properties. In order to draw conclusions how the system works in its wild-type form, i.e., in the presence of Mg^{2+}, one has to know exactly how the coordinating properties of the different metal ions change according to their position in the periodic table. Some efforts in this direction have been made and some progress was achieved in the past few years [8,34,37]. However, to make things more complicated, it also turned out that there seems to be no general rule: While for two different hammerhead ribozymes a clear correlation was found between cleavage rate and phosphate affinities of specific metal ions [34], no such correlation could be established for the *glms* ribozyme, which is about the same size and uses the same mechanism of cleavage [54]. Consequently, it looks like as if every ribozyme, sometimes even of the same class but from different organisms, has its own specificities for metal ion binding.

Most methods applied today find their origin in applications in the metalloprotein field and certainly many more will be transferred, adjusted, and applied to nucleic acids. There is a great need for the development of new methods to investigate metal ion binding to nucleic acids. Recent novel approaches include the use of electrospray ionization mass spectrometry (ESI-MS) [276]: Metal ion-nucleic acid complexes of the thrombin-binding aptamer of different concentration were fragmented in the gas-phase and such the binding affinities determined. Another recent approach suggests to use a specifically placed nitroxide spin label in order to measure the distance to coordinated paramagnetic metal ions by EPR [168]. This latter method has been used repeatedly with proteins (e.g., [336–340]). It will be fascinating to see what other methods will come up in the near and also more distant future to help to solve the "mysteries" of metal ion binding to nucleic acids.

Abbreviations and Definitions

5'UTR	5' untranslated region
AES	atomic emission spectroscopy
Bistris	bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane
ENDOR	electron-nuclear double resonance spectroscopy
EPR	electron paramagnetic resonance
ER	endoplasmic reticulum
ESEEM	electron spin echo envelope modulation
ESI-MS	electrospray ionization mass spectrometry
EXAFS	extended X-ray absorption fine structure
FT-IR	Fourier-transform infrared
glmS	glucosamine-6-phosphate activated ribozyme
GNA	glycol nucleic acid
GTP	guanosine 5'-triphosphate
HDV	hepatitits delta virus
Hepes	N-(2-hydroxyethyl)-piperazine-N′-2-ethanesulfonic acid
HSQC	heteronuclear single quantum coherence
IR	infrared
ITC	isothermal titration calorimetry
MD	molecular dynamics
MOPS	3-(N-morpholino)propanesulfonic acid
MRD	magnetic relaxation dispersion
NAIM	nucleotide analog interference mapping
NLPB	non-linear Poisson-Boltzmann
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NTPαS	nucleoside 5'-α-thiotriphosphate
PAGE	polyacrylamide gel electrophoresis
PNA	peptide nucleic acid
RNase	ribonuclease
ROESY	rotating frame Overhauser effect spectroscopy
rRNA	ribosomal RNA
RzB	trans cleaving derivative of the hammerhead ribozyme from the peach latent mosaic viroid (PLMVd)
SRP	signal recognition particle
T5PNP	5′-*p*-nitrophenylphosphate
TFA	thermodynamic fingerprint analysis
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
tRNA	transfer RNA
Wye	wyosine, a natural analog of guanine
XANES	X-ray absorption near edge structure
XAS	X-ray absorption spectroscopy

Acknowledgments Financial support by the European Research Council (ERC Starting Grant 2010 to RKOS), the Swiss National Science Foundation (RKOS), the Swiss State Secretariat for Education and Research within the COST Action D39 (RKOS) and the University of Zürich is gratefully acknowledged.

References

1. D. E. Draper, D. Grilley, A. M. Soto, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 221–243.
2. G. S. Manning, *Acc. Chem. Res.* **1979**, *12*, 443–449.
3. K. Chin, K. A. Sharp, B. Honig, A. M. Pyle, *Nat. Struct. Biol.* **1999**, *6*, 1055–1061.
4. S. A. Woodson, *Curr. Opin. Chem. Biol.* **2005**, *9*, 104–109.
5. V. K. Misra, R. Shiman, D. E. Draper, *Biopolymers* **2003**, *69*, 118–136.
6. R. K. O. Sigel, D. G. Sashital, D. L. Abramovitz, A. G. Palmer III, S. E. Butcher, A. M. Pyle, *Nat. Struct. Mol. Biol.* **2004**, *11*, 187–192.
7. M. C. Erat, R. K. O. Sigel, *Inorg. Chem.* **2007**, *46*, 11224–11234.
8. R. K. O. Sigel, H. Sigel, *Acc. Chem. Res.* **2010**, *43*, 974–984.
9. M. C. Erat, J. Coles, C. Finazzo, B. Knobloch, R. K. O. Sigel, *Coord. Chem. Rev.* **2011**, doi:10.1016/j.ccr.2011.08.009.
10. M. C. Erat, R. K. O. Sigel, *Met. Ions Life Sci.* **2011**, *9*, 37–100.
11. N. V. Hud, V. Sklenár, J. Feigon, *J. Mol. Biol.* **1999**, *286*, 651–660.
12. L. McFail-Isom, C. C. Sines, L. D. Williams, *Curr. Opin. Struct. Biol.* **1999**, *9*, 298–304.
13. T. Theophanides, H. A. Tajmir-Riahi, *J. Biomol. Struct. Dyn.* **1985**, *2*, 995–1004.
14. M. Guéron, J. P. Demaret, M. Filoche, *Biophys. J.* **2000**, *78*, 1070–1083.
15. B. Spingler, *Chimia* **2009**, *63*, 153–156.
16. B. Spingler, C. da Pieve, *Dalton Trans.* **2005**, 1637–1643.
17. B. A. Armitage, *Nature Chem. Biol.* **2007**, *3*, 203–204.
18. M. J. Cromie, Y. X. Shi, T. Latifi, E. A. Groisman, *Cell* **2006**, *125*, 71–84.
19. C. E. Dann III, C. A. Wakeman, C. L. Sieling, S. C. Baker, I. Irnov, W. C. Winkler, *Cell* **2007**, *130*, 878–892.
20. J.-H. Chen, R. Yajima, D. M. Chadalavada, E. Chase, P. C. Bevilacqua, B. L. Golden, *Biochemistry* **2010**, *49*, 6508–6518.
21. N. Toor, K. S. Keating, S. D. Taylor, A. M. Pyle, *Science* **2008**, *320*, 77–82.
22. N. J. Reiter, A. Osterman, A. Torres-Larios, K. K. Swinger, T. Pan, A. Mondragon, *Nature* **2010**, *468*, 784–789.
23. M. R. Stahley, S. A. Strobel, *Science* **2005**, *309*, 1587–1590.
24. R. K. O. Sigel, A. M. Pyle, *Chem. Rev.* **2007**, *107*, 97–113.
25. R. K. O. Sigel, E. Freisinger, B. Lippert, *J. Biol. Inorg. Chem.* **2000**, *5*, 287–299.
26. R. K. O. Sigel, M. Sabat, E. Freisinger, A. Mower, B. Lippert, *Inorg. Chem.* **1998**, *38*, 1481–1490.
27. B. Knobloch, R. K. O. Sigel, B. Lippert, H. Sigel, *Angew. Chem., Int. Ed.* **2004**, *43*, 3793–3795.
28. B. Lippert, *Prog. Inorg. Chem.* **2005**, *54*, 385–443.
29. B. Lippert, *Chem. Biodivers.* **2008**, *5*, 1455–1474.
30. R. Griesser, G. Kampf, L. E. Kapinos, S. Komeda, B. Lippert, J. Reedijk, H. Sigel, *Inorg. Chem.* **2003**, *42*, 32–41.
31. H. Sigel, *Pure Appl. Chem.* **2004**, *76*, 1869–1886.
32. R. G. Pearson, *J. Chem. Edu.* **1968**, *45*, 643–648.
33. R. G. Pearson, *J. Chem. Edu.* **1968**, *45*, 581–587.
34. J. Schnabl, R. K. O. Sigel, *Curr. Opin. Chem. Biol.* **2010**, *14*, 269–275.
35. R. B. Martin, *Met. Ions Biol. Syst.* **1996**, *32*, 61–89.
36. R. K. O. Sigel, H. Sigel, *Met. Ions Life Sci.* **2007**, *2*, 109–180.
37. E. Freisinger, R. K. O. Sigel, *Coord. Chem. Rev.* **2007**, *251*, 1834–1851.

38. C. P. Da Costa, H. Sigel, *Inorg. Chem*. **2000**, *39*, 5985–5993.
39. F. M. Al-Sogair, B. P. Operschall, A. Sigel, H. Sigel, J. Schnabl, R. K. O. Sigel, *Chem. Rev.* **2011**, *111*, 4964–5003.
40. A. Mucha, B. Knobloch, M. Jeżowska-Bojczuk, H. Kozłowski, R. K. O. Sigel, *Dalton Trans.* **2008**, 5368–5377.
41. L. A. Finney, T. V. O'Halloran, *Science* **2003**, *300*, 931–936.
42. S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, Ca, 1995.
43. *The Biological Chemistry of Magnesium*, Ed J. A. Cowan, VCH Publishers, Inc., New York, 1995.
44. T. Pan, D. M. Long, O. C. Uhlenbeck, in *The RNA World*, Eds R. Gesteland, J. Atkins, Cold Spring Harbor Press, Cold Spring Harbor, 1993, pp. 271–302.
45. R. L. Gonzalez, Jr., I. Tinoco, Jr., *Methods Enzymol*. **2001**, *338*, 421–443.
46. I. Tinoco, Jr., C. Bustamante, *J. Mol. Biol*. **1999**, *293*, 271–281.
47. A. M. Pyle, *Nature* **1996**, *381*, 280–281.
48. A. M. Pyle, *J. Biol. Inorg. Chem*. **2002**, *7*, 679–690.
49. R. K. O. Sigel, *Eur. J. Inorg. Chem*. **2005**, *12*, 2281–2292.
50. C. Hsiao, in *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, Royal Society of Chemistry, Cambridge, UK, 2009, pp. 1–38.
51. A. S. Petrov, J. C. Bowman, S. C. Harvey, L. D. Williams, *RNA* **2011**, *17*, 291–297.
52. M. Roychowdhury-Saha, D. H. Burke, *RNA* **2006**, *12*, 1846–1852.
53. J. L. Boots, M. D. Canny, E. Azimi, A. Pardi, *RNA* **2008**, *14*, 2212–2222.
54. K. Klawuhn, J. A. Jansen, J. Souchek, G. A. Soukup, J. K. Soukup, *ChemBioChem* **2010**, *11*, 2567–2571.
55. V. K. Misra, D. E. Draper, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12456–12461.
56. K. Juneau, E. Podell, D. J. Harrington, T. R. Cech, *Structure* **2001**, *9*, 221–231.
57. H. Sigel, R. Griesser, *Chem. Soc. Rev*. **2005**, *34*, 875–900.
58. S. Basu, R. P. Rambo, J. Strauss-Soukup, J. H. Cate, A. R. Ferré-D'Amaré, S. A. Strobel, J. A. Doudna, *Nature Struct. Biol*. **1998**, *5*, 986–992.
59. J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell* **1989**, *59*, 871–880.
60. T. Pozzan, R. Rizzuto, *Nature Cell Biol*. **2000**, *2*, E25–E27.
61. C. E. Outten, T. V. O'Halloran, *Science* **2001**, *292*, 2488–2492.
62. T. D. Rae, P. J. Schmidt, R. A. Pufahl, V. C. Culotta, T. V. O'Halloran, *Science* **1999**, *284*, 805–808.
63. B. Lippert, in *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, Royal Society of Chemistry, Cambridge, UK, 2009, pp. 39–74.
64. M. C. Erat, R. K. O. Sigel, *J. Biol. Inorg. Chem*. **2008**, *13*, 1025–1036.
65. H. Irving, R. J. P. Williams, *J. Chem. Soc.* **1953**, 3192–3210.
66. R. B. Martin, *Met. Ions Biol. Syst*. **1986**, *20*, 21–65.
67. R. B. Martin, *J. Chem. Edu*. **1987**, *64*, 402–402.
68. R. D. Shannon, *Acta Crystallogr*. **1976**, *A32*, 751–767.
69. *The Chemistry of Aqua Ions*, Ed D. T. Richens, John Wiley & Sons, Hoboken, 1997, pp. 592.
70. S. F. Lincoln, A. E. Merbach, *Adv. Inorg. Chem*. **1995**, *42*, 1–88.
71. C. F. Baes, Jr., R. E. Mesmer, *The Hydrolysis of Cations*, Krieger Publishing Co., Malabar, Florida, 1976, pp. 1–496.
72. R. G. Pearson, *Inorg. Chem*. **1988**, *27*, 734–740.
73. W. E. Morf, W. Simon, *Helv. Chim. Acta* **1971**, *54*, 794–810.
74. K. B. Yatsimirsky, *Theor. Exp. Chem*. **1994**, *30*, 1–11.
75. L. Helm, A. E. Merbach, *Chem. Rev*. **2005**, *105*, 1923–1959.
76. J. E. Huheey, *Inorganic Chemistry*, Harper and Row, New York, 1983.
77. Y. Inada, A. M. Mohammed, H. H. Loeffler, S. Funahashi, *Helv. Chim. Acta* **2005**, *88*, 461–469.
78. A. F. Holleman, E. Wiberg, *Lehrbuch der Anorganischen Chemie*, Walter de Gruyter, Berlin, New York, 1985, pp. 1265–1279.

79. Y. Peeraer, A. Rabijns, J.-F. Collet, E. Van Schaftingen, C. De Ranter, *Eur. J. Biochem*. **2004**, *271*, 3421–3427.
80. R. M. Weinshilboum, F. A. Raymond, *Biochem. Pharmacol*. **1976**, *25*, 573–579.
81. S. O. Shan, D. Herschlag, *RNA* **2000**, *6*, 795–813.
82. M. Steiner, D. Rueda, R. K. O. Sigel, *Angew. Chem., Int. Ed*. **2009**, *48*, 9739–9742.
83. K. Aoki, *Acta Crystallogr. B* **1976**, *32*, 1454–1459.
84. V. L. Pecoraro, J. D. Hermes, W. W. Cleland, *Biochemistry* **1984**, *23*, 5262–5271.
85. R. K. O. Sigel, B. Song, H. Sigel, *J. Am. Chem. Soc*. **1997**, *119*, 744–755.
86. B. Knobloch, B. Nawrot, A. Okruszek, R. K. O. Sigel, *Chem. Eur. J*. **2008**, 14, 3100–3109 and *14*, 3509.
87. H. Sigel, R. B. Martin, *Chem. Soc. Rev*. **1994**, *23*, 83–91.
88. A. T. Perrotta, M. D. Been, *Biochemistry* **2007**, *46*, 5124–5130.
89. N. Lehman, G. F. Joyce, *Nature* **1993**, *361*, 182–185.
90. N. Lehman, G. F. Joyce, *Curr. Biol*. **1993**, *3*, 723–734.
91. D. N. Frank, N. R. Pace, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 14355–14360.
92. B. Cuenoud, J. W. Szostak, *Nature* **1995**, *375*, 611–614.
93. N. Carmi, L. A. Shultz, R. R. Breaker, *Chem. Biol*. **1996**, *3*, 1039–1046.
94. W. Wang, L. P. Billen, Y. Li, *Chem. Biol*. **2002**, *9*, 507–517.
95. P. J. Bruesehoff, J. Li, A. J. Augustine, Y. Lu, *Comb. Chem. High T. Scr*. **2002**, *5*, 327–335.
96. J. Li, W. C. Zheng, A. H. Kwon, Y. Lu, *Nucleic Acids Res*. **2000**, *28*, 481–488.
97. T. Pan, B. Dichtl, O. C. Uhlenbeck, *Biochemistry* **1994**, *33*, 9561–9565.
98. H. P. Hofmann, S. Limmer, V. Hornung, M. Sprinzl, *RNA* **1997**, *3*, 1289–1300.
99. M. Zivarts, Y. Liu, R. R. Breaker, *Nucleic Acids Res*. **2005**, *33*, 622–631.
100. Y. Lu, J. W. Liu, J. Li, P. J. Bruesehoff, C. M. B. Pavot, A. K. Brown, *Biosens. Bioelectron*. **2003**, *18*, 529–540.
101. T. Li, E. Wang, S. Dong, *Anal. Chem*. **2010**, *82*, 1515–1520.
102. B. Rosenberg, L. Vancamp, J. E. Trosko, V. H. Mansour, *Nature* **1969**, *222*, 385–386.
103. A. M. Pizarro, P. J. Sadler, in *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, Royal Society of Chemistry, Cambridge, UK, 2009, pp. 350–416.
104. A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. Machinami, S. Oda, Y. Miyake, I. Okamoto, Y. Tanaka, *Chem. Commun.* **2008**, 4825–4827.
105. A . Ono, H. Togashi, *Angew. Chem. Int. Ed*. **2004**, *43*, 4300–4302.
106. J. Liu, Y. Lu, *Angew. Chem. Int. Ed*. **2007**, *46*, 7587–7590.
107. X. Liu, Y. Tang, L. Wang, J. Zhang, S. Song, C. Fan, S. Wang, *Adv. Mat*. **2007**, *19*, 1471–1474.
108. R. Freeman, T. Finder, I. Willner, *Angew. Chem. Int. Ed*. **2009**, *48*, 7818–7821.
109. T. Carell, C. Behrens, J. Gierlich, *Org. Biomol. Chem*. **2003**, *1*, 2221–2228.
110. J. S. Lee, L. J. Latimer, R. S. Reid, *Biochem. Cell Biol*. **1993**, *71*, 162–168.
111. P. Aich, S. L. Labiuk, L. W. Tari, L. J. T. Delbaere, W. J. Roesler, K. J. Falk, R. P. Steer, J. S. Lee, *J. Mol. Biol*. **1999**, *294*, 477–485.
112. A. Rakitin, P. Aich, C. Papadopoulos, Y. Kobzar, A. S. Vedeneev, J. S. Lee, J. M. Xu, *Phys. Rev. Lett*. **2001**, *86*, 3670–3673.
113. F. Moreno-Herrero, P. Herrero, F. Moreno, J. Colchero, C. Gómez-Navarro, J. Gómez-Herrero, A. M. Baró, *Nanotechnology* **2003**, *14*, 128–128.
114. M. Fuentes-Cabrera, B. G. Sumpter, J. E. Šponer, J. Šponer, L. Petit, J. C. Wells, *J. Phys. Chem. B* **2007**, *111*, 870–879.
115. S. S. Alexandre, J. M. Soler, L. Seijo, F. Zamora, *Phys. Rev. B* **2006**, *73*, 205112–205112.
116. S. L. Labiuk, L. T. J. Delbaere, J. S. Lee, *J. Biol. Inorg. Chem*. **2003**, *8*, 715–720.
117. Y. Tanaka, S. Oda, H. Yamaguchi, Y. Kondo, C. Kojima, A. Ono, *J. Am. Chem. Soc*. **2007**, *129*, 244–245.
118. S. Johannsen, N. Megger, D. Böhme, R. K. O. Sigel, J. Müller, *Nature Chem*. **2010**, *2*, 229–234.
119. Y. Miyake, H. Togashi, M. Tashiro, H. Yamaguchi, S. Oda, M. Kudo, Y. Tanaka, Y. Kondo, R. Sawa, T. Fujimoto, T. Machinami, A. Ono, *J. Am. Chem. Soc*. **2006**, *128*, 2172–2173.
120. E. Ennifar, P. Walter, P. Dumas, *Nucleic Acids Res*. **2003**, *31*, 2671–2682.
121. D. Böhme, N. Düpre, D. A. Megger, J. Müller, *Inorg. Chem*. **2007**, *46*, 10114–10119.

122. K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shiro, M. Shionoya, *J. Am. Chem. Soc*. **2002**, *124*, 12494–12498.
123. G. H. Clever, C. Kaul, T. Carell, *Angew. Chem. Int. Ed*. **2007**, *46*, 6226–6236.
124. G. H. Clever, K. Polborn, T. Carell, *Angew. Chem. Int. Ed*. **2005**, *44*, 7204–7208.
125. J. Müller, *Eur. J. Inorg. Chem.* **2008**, 3749–3763.
126. G. H. Clever, M. Shionoya, *Coord. Chem. Rev*. **2010**, *254*, 2391–2402.
127. F. H. Westheimer, *Acc. Chem. Res*. **1968**, *1*, 70–78.
128. G. A. Soukup, R. R. Breaker, *RNA* **1999**, *5*, 1308–1325.
129. R. S. Brown, J. C. Dewan, A. Klug, *Biochemistry* **1985**, *24*, 4785–4801.
130. R. Breslow, D.-L. Huang, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 4080–4083.
131. C. Werner, B. Krebs, G. Keith, G. Dirheimer, *Biochim. Biophys. Acta* **1976**, *432*, 161–175.
132. R. S. Brown, B. A. Hingerty, J. C. Dewan, A. Klug, *Nature* **1983**, *303*, 543–546.
133. J. Ciesiolka, W. D. Hardt, J. Schlegl, V. A. Erdmann, R. K. Hartmann, *Eur. J. Biochem*. **1994**, *219*, 49–56.
134. M. Hertweck, M. W. Müller, *Eur. J. Biochem*. **2001**, *268*, 4610–4620.
135. J. Ciesiolka, T. Marciniec, W. J. Krzyzosiak, *Eur. J. Biochem*. **1989**, *182*, 445–450.
136. R. K. O. Sigel, A. Vaidya, A. M. Pyle, *Nat. Struct. Biol*. **2000**, *7*, 1111–1116.
137. R. K. O. Sigel, A. M. Pyle, *Met. Ions Biol. Syst*. **2003**, *40*, 477–512.
138. I. Zagórowska, S. Kuusela, H. Lönnberg, *Nucleic Acids Res*. **1998**, *26*, 3392–3396.
139. C. Berens, B. Streicher, R. Schroeder, W. Hillen, *Chem. Biol*. **1998**, *5*, 163–175.
140. H. J. H. Fenton, *Proc. Chem. Soc*. **1893**, *9*, 113–113.
141. H. J. H. Fenton, *J. Chem. Soc*. **1894**, *65*, 899–910.
142. I. D. Brown, *Acta Crystallogr. B* **1988**, *44*, 545–553.
143. R. Das, K. J. Travers, Y. Bai, D. Herschlag, *J. Am. Chem. Soc*. **2005**, *127*, 8272–8273.
144. I. T. Suydam, S. A. Strobel, *Methods Enzymol*. **2009**, *468*, 3–30.
145. E. L. Christian, M. Yarus, *Biochemistry* **1993**, *32*, 4475–4480.
146. S. Basu, S. A. Strobel, *Methods* **2001**, *23*, 264–275.
147. J. A. Jansen, T. J. McCarthy, G. A. Soukup, J. K. Soukup, *Nature Struct. Mol. Biol*. **2006**, *13*, 517–523.
148. V. D. Sood, T. L. Beattie, R. A. Collins, *J. Mol. Biol*. **1998**, *282*, 741–750.
149. A. D. Griffiths, B. V. Potter, I. C. Eperon, *Nucleic Acids Res*. **1987**, *15*, 4145–4162.
150. C. P. Da Costa, D. Krajewska, A. Okruszek, W. J. Stec, H. Sigel, *J. Biol. Inorg. Chem*. **2002**, *7*, 405–415.
151. C. P. Da Costa, A. Okruszek, H. Sigel, *ChemBioChem* **2003**, *4*, 593–602.
152. J. K. Frederiksen, R. Fong, J. A. Piccirilli, Ed N. V. Hud, in *Nucleic Acid-Metal Ion Interactions*, Royal Society of Chemistry, Cambridge, UK, **2009**, 260–306.
153. S. Basu, S. A. Strobel, *RNA* **1999**, *5*, 1399–1407.
154. S.-O. Shan, A. Yoshida, S. Sun, J. A. Piccirilli, D. Herschlag, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12299–12304.
155. J. K. Frederiksen, J. A. Piccirilli, *Methods* **2009**, *49*, 148–166.
156. M. Forconi, D. Herschlag, *Methods Enzymol*. **2009**, *468*, 311–333.
157. S. Wang, K. Karbstein, A. Peracchi, L. Beigelman, D. Herschlag, *Biochemistry* **1999**, *38*, 14363–14378.
158. W. W. Cleland, *Enzyme Kinetics and Mechanism Part D: Developments in Enzyme Dynamics*, Ed D. L. Purich, Academic Press, **1995**, *249*, 341–373.
159. D. B. Northrop, *Methods* **2001**, *24*, 117–124.
160. A. G. Cassano, V. E. Anderson, M. E. Harris, *Biochemistry* **2004**, *43*, 10547–10559.
161. A. G. Cassano, V. E. Anderson, M. E. Harris, *Biopolymers* **2004**, *73*, 110–129.
162. M. E. Harris, A. G. Cassano, *Curr. Opin. Chem. Biol*. **2008**, *12*, 626–639.
163. Q. Dai, J. K. Frederiksen, V. E. Anderson, M. E. Harris, J. A. Piccirilli, *J. Org. Chem*. **2008**, *73*, 309–311.
164. J. L. Leroy, M. Guéron, *Biopolymers* **1977**, *16*, 2429–2446.
165. T. E. Horton, D. R. Clardy, V. J. DeRose, *Biochemistry* **1998**, *37*, 18094–18101.
166. N. Kisseleva, S. Kraut, A. Jaschke, O. Schiemann, *HFSP J*. **2007**, *1*, 127–136.

167. L. Hunsicker-Wang, M. Vogt, V. J. DeRose, *Methods Enzymol*. **2009**, *468*, 335–367.
168. V. J. DeRose, in *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, Royal Society of Chemistry, Cambridge, UK, **2009**, 154–179.
169. M. Vogt, S. Lahiri, C. G. Hoogstraten, R. D. Britt, V. J. DeRose, *J. Am. Chem. Soc*. **2006**, *128*, 16764–16770.
170. J. E. Penner-Hahn, *Coord. Chem. Rev*. **2005**, *249*, 161–177.
171. I. V. Smirnov, F. W. Kotch, I. J. Pickering, J. T. Davis, R. H. Shafer, *Biochemistry* **2002**, *41*, 12133–12139.
172. K. S. Koutmou, A. Casiano-Negroni, M. M. Getz, S. Pazicni, A. J. Andrews, J. E. Penner-Hahn, H. M. Al-Hashimi, C. A. Fierke, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2479–2484 S2479/2471–S2479/2475.
173. B.-K. Teo, P. Eisenberger, J. Reed, J. K. Barton, S. J. Lippard, *J. Am. Chem. Soc*. **1978**, *100*, 3225–3227.
174. A. P. Hitchcock, C. J. L. Lock, W. M. C. Pratt, *Inorg. Chim. Acta* **1982**, *66*, L45–L47.
175. J. M. Benevides, S. A. Overman, G. J. Thomas, *J. Raman Spec*. **2005**, *36*, 279–299.
176. T. Theophanides, in *Infrared and Raman Spectroscopy of Biological Materials*, Ed H.-U. Gremlich, B. Yan, M. Dekker, New York, **2000**, 205–223.
177. Y. Chen, N. V. Eldho, K. T. Dayie, P. R. Carey, *Biochemistry* **2010**, *49*, 3427–3435.
178. H. Takeuchi, H. Murata, I. Harada, *J. Am. Chem. Soc*. **1988**, *110*, 392–397.
179. H. A. Tajmir-Riahi, T. Theophanides, *J. Biomol. Struct. Dyn*. **1985**, *3*, 537–542.
180. S. Adam, P. Bourtayre, J. Liquier, E. Taillandier, *Nucleic Acids Res*. **1986**, *14*, 3501–3513.
181. M. Langlais, H. A. Tajmir-Riahi, R. Savoie, *Biopolymers* **1990**, *30*, 743–752.
182. J. Duguid, V. A. Bloomfield, J. Benevides, G. J. Thomas, *Biophys. J*. **1993**, *65*, 1916–1928.
183. J. G. Duguid, V. A. Bloomfield, J. M. Benevides, G. J. Thomas, *Biophys. J*. **1995**, *69*, 2623–2641.
184. H. Arakawa, J. F. Neault, H. A. Tajmir-Riahi, *Biophys. J*. **2001**, *81*, 1580–1587.
185. J. Stangret, R. Savoie, *Phys. Chem. Chem. Phys*. **2002**, *4*, 4770–4773.
186. S. Nafisi, Z. Norouzi, *DNA Cell Biol*. **2009**, *28*, 469–477.
187. S. Nafisi, A. Sobhanmanesh, K. Alimoghaddam, A. Ghavamzadeh, H.-A. Tajmir-Riahi, *DNA Cell Biol*. **2005**, *24*, 634–640.
188. E. L. Christian, V. E. Anderson, P. R. Carey, M. E. Harris, *Biochemistry* **2010**, *49*, 2869–2879.
189. J.-H. Chen, B. Gong, P. C. Bevilacqua, P. R. Carey, B. L. Golden, *Biochemistry* **2009**, *48*, 1498–1507.
190. B. Gong, J.-H. Chen, E. Chase, D. M. Chadalavada, R. Yajima, B. L. Golden, P. C. Bevilacqua, P. R. Carey, *J. Am. Chem. Soc*. **2007**, *129*, 13335–13342.
191. B. Gong, Y. Chen, E. L. Christian, J. H. Chen, E. Chase, D. M. Chadalavada, R. Yajima, B. L. Golden, P. C. Bevilacqua, P. R. Carey, *J. Am. Chem. Soc*. **2008**, *130*, 9670–9672.
192. A. Wong, R. Ida, G. Wu, *Biochem. Biophys. Res. Commun*. **2005**, *337*, 363–366.
193. M. L. Gill, S. A. Strobel, J. P. Loria, *J. Am. Chem. Soc*. **2005**, *127*, 16723–16732.
194. R. L. J. Gonzalez, I. Tinoco Jr., *Methods Enzymol*. **2002**, *338*, 421–443.
195. J. Feigon, S. E. Butcher, L. D. Finger, N. V. Hud, *Methods Enzymol*. **2001**, *338*, 400–420.
196. M. C. Erat, O. Zerbe, T. Fox, R. K. O. Sigel, *ChemBioChem* **2007**, *8*, 306–314.
197. S. E. Butcher, F. H.-T. Allain, J. Feigon, *Biochemistry* **2000**, *39*, 2174–2182.
198. P. Legault, C. G. Hoogstraten, E. Metlitzky, A. Pardi, *J. Mol. Biol*. **1998**, *284*, 325–335.
199. M. Seetharaman, N. V. Eldho, R. A. Padgett, K. T. Dayie, *RNA* **2006**, *12*, 235–247.
200. M. Schmitz, I. Tinoco Jr., *RNA* **2000**, *6*, 1212–1225.
201. R. L. Gonzalez, Jr., I. Tinoco Jr., *J. Mol. Biol*. **1999**, *289*, 1267–1282.
202. S. S. Wijmenga, B. N. M. van Buuren, *Prog. Nucl. Mag. Res. Sp*. **1998**, *32*, 287–387.
203. B. A. Connolly, F. Eckstein, *Biochemistry* **1984**, *23*, 5523–5527.
204. M. K. Hansen, J.-P. Simorre, P. Hanson, V. Mokler, L. Bellon, L. Beigelman, A. Pardi, *RNA* **1999**, *5*, 1099–1104.
205. F. Eckstein, *Annu. Rev. Biochem*. **1985**, *54*, 367–402.
206. M. C. Erat, R. K. O. Sigel, *results to be published*.
207. M. Maderia, L. M. Hunsicker, V. J. DeRose, *Biochemistry* **2000**, *39*, 12113–12120.

208. E. M. Osborne, W. L. Ward, M. Z. Ruehle, V. J. DeRose, *Biochemistry* **2009**, *48*, 10654–10664.
209. Y. Tanaka, C. Kojima, E. H. Morita, Y. Kasai, K. Yamasaki, A. Ono, M. Kainosho, K. Taira, *J. Am. Chem. Soc*. **2002**, *124*, 4595–4601.
210. G. Wang, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc*. **2004**, *126*, 8908–8909.
211. G. W. Buchanan, J. B. Stothers, *Can. J. Chem*. **1982**, *60*, 787–791.
212. Y. Tanaka, Y. Kasai, S. Mochizuki, A. Wakisaka, E. H. Morita, C. Kojima, A. Toyozawa, Y. Kondo, M. Taki, Y. Takagi, A. Inoue, K. Yamasaki, K. Taira, *J. Am. Chem. Soc*. **2004**, *126*, 744–752.
213. Y. Tanaka, E. H. Morita, H. Hayashi, Y. Kasai, T. Tanaka, K. Taira, *J. Am. Chem. Soc*. **2000**, *122*, 11303–11310.
214. M. C. Erat, H. Kovacs, R. K. O. Sigel, *J. Inorg. Biochem*. **2010**, *104*, 611–613.
215. V. Sychrovský, J. Šponer, P. Hobza, *J. Am. Chem. Soc*. **2004**, *126*, 663–672.
216. H. Li, R. I. Cukier, Y. Bu, *J. Phys. Chem. B* **2008**, *112*, 9174–9181.
217. I. Bertini, C. Luchinat, *NMR of Paramagnetic Molecules in Biological Systems*, Benjamin/Cummings Pub. Co., Menlo Park, CA, **1986**, p. 319.
218. I. Bertini, C. Luchinat, G. Parigi, *Solution NMR of Paramagnetic Molecules - Applications to Metallobiomolecules and Models*, Elsevier Science B.V., Amsterdam, **2001**, 1–372.
219. J. H. Davis, T. R. Foster, M. Tonelli, S. E. Butcher, *RNA* **2007**, *13*, 76–86.
220. J. Noeske, H. Schwalbe, J. Wöhnert, *Nucleic Acids Res*. **2007**, *35*, 5262–5273.
221. D. O. Campbell, P. Bouchard, G. Desjardins, P. Legault, *Biochemistry* **2006**, *45*, 10591–10605.
222. R. E. Hurd, E. Azhderian, B. R. Reid, *Biochemistry* **1979**, *18*, 4012–4017.
223. J. A. Cowan, *J. Inorg. Biochem*. **1993**, *49*, 171–175.
224. G. Colmenarejo, I. Tinoco Jr., *J. Mol. Biol*. **1999**, *290*, 119–135.
225. J. S. Kieft, I. Tinoco, Jr., *Structure* **1997**, *5*, 713–721.
226. S. Rüdisser, I. Tinoco, Jr., *J. Mol. Biol*. **2000**, *295*, 1211–1223.
227. P. Schultze, N. V. Hud, F. W. Smith, J. Feigon, *Nucleic Acids Res*. **1999**, *27*, 3018–3028.
228. F. W. Smith, J. Feigon, *Biochemistry* **1993**, *32*, 8682–8692.
229. N. V. Hud, P. Schultze, V. Sklenár, J. Feigon, *J. Mol. Biol*. **1999**, *285*, 233–243.
230. D. M. Rose, C. F. Polnaszek, R. G. Bryant, *Biopolymers* **1982**, *21*, 653–664.
231. S. S. Reid, J. A. Cowan, *Biochemistry* **1990**, *29*, 6025–6032.
232. E. Berggren, L. Nordenskiöld, W. H. Braunlin, *Biopolymers* **1992**, *32*, 1339–1350.
233. J. A. Cowan, H. W. Huang, L. Y. Hsu, *J. Inorg. Biochem*. **1993**, *52*, 121–129.
234. L. A. Wright, L. E. Lerner, *Biopolymers* **1994**, *34*, 691–700.
235. L. M. Utschig, J. W. Bryson, T. V. O'Halloran, *Science* **1995**, *268*, 380–385.
236. J. E. Coleman, *Methods Enzymol*. **1993**, *227*, 16–43.
237. R. Ida, G. Wu, *J. Am. Chem. Soc*. **2008**, *130*, 3590–3602.
238. J. P. André, H. R. Mäcke, *J. Inorg. Biochem*. **2003**, *97*, 315–323.
239. S. Basu, A. A. Szewczak, M. Cocco, S. A. Strobel, *J. Am. Chem. Soc*. **2000**, *122*, 3240–3241.
240. J. Reedijk, *Platinum Metals Review* **2008**, *52*, 2–11.
241. L. Helm, A. E. Merbach, *Coord. Chem. Rev*. **1999**, *187*, 151–181.
242. L.-Y. Lian, G. C. K. Roberts, in *NMR of Macromolecules: A Practical Approach*, Ed G. C. K. Roberts, IRL Press at Oxford University Press, Oxford, New York, **1993**, 153–182.
243. H. S. Gutowsky, C. H. Holm, *J. Chem. Phys*. **1956**, *25*, 1228–1228.
244. L. Nordenskjöld, D. K. Chang, C. F. Anderson, M. T. Record, *Biochemistry* **1984**, *23*, 4309–4317.
245. W. H. Braunlin, C. F. Anderson, M. T. Record, *Biochemistry* **1987**, *26*, 7724–7731.
246. W. H. Braunlin, T. Drakenberg, L. Nordenskiöld, *Biopolymers* **1987**, *26*, 1047–1062.
247. B. Halle, V. P. Denisov, *Biopolymers* **1998**, *48*, 210–233.
248. B. Halle, V. P. Denisov, *Methods Enzymol*. **2002**, *338*, 178–201.
249. V. P. Denisov, G. Carlström, K. Venu, B. Halle, *J. Mol. Biol*. **1997**, *268*, 118–136.
250. V. P. Denisov, B. Halle, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 629–633.

251. F. Cesare Marincola, V. P. Denisov, B. Halle, *J. Am. Chem. Soc*. **2004**, *126*, 6739–6750.
252. S. Johannsen, M. M. T. Korth, J. Schnabl, R. K. O. Sigel, *Chimia* **2009**, *63*, 146–152.
253. P. Podbevšek, N. V. Hud, J. Plavec, *Nucleic Acids Res*. **2007**, *35*, 2554–2563.
254. P. Šket, J. Plavec, *J. Am. Chem. Soc*. **2007**, *129*, 8794–8800.
255. D. P. Bancroft, C. A. Lepre, S. J. Lippard, *J. Am. Chem. Soc*. **1990**, *112*, 6860–6871.
256. D. Labuda, D. Poerschke, *Biochemistry* **1982**, *21*, 49–53.
257. S. E. McDowell, N. Špačková, J. Šponer, N. G. Walter, *Biopolymers* **2007**, *85*, 169–184.
258. Y. Hashem, P. Auffinger, *Methods* **2009**, *47*, 187–197.
259. P. Auffinger, L. Bielecki, E. Westhof, *J. Mol. Biol*. **2004**, *335*, 555–571.
260. P. Auffinger, E. Westhof, *J. Mol. Biol*. **2000**, *300*, 1113–1131.
261. P. Auffinger, N. Grover, E. Westhof, *Met. Ions Life Sci*. **2011**, *9*, 1–35.
262. P. Várnai, K. Zakrzewska, *Nucleic Acids Res*. **2004**, *32*, 4269–4280.
263. M. V. Krasovska, J. Sefcikova, K. Reblova, B. Schneider, N. G. Walter, J. Šponer, *Biophys. J*. **2006**, *91*, 626–638.
264. V. K. Misra, D. E. Draper, *J. Mol. Biol*. **2002**, *317*, 507–521.
265. D. E. Draper, *RNA* **2004**, *10*, 335–343.
266. M. J. Serra, J. D. Baird, T. Dale, B. L. Fey, K. Retatagos, E. Westhof, *RNA* **2002**, *8*, 307–323.
267. Y. V. Bukhman, D. E. Draper, *J. Mol. Biol*. **1997**, *273*, 1020–1031.
268. R. Shiman, D. E. Draper, *J. Mol. Biol*. **2000**, *302*, 79–91.
269. L. G. Laing, T. C. Gluick, D. E. Draper, *J. Mol. Biol*. **1994**, *237*, 577–587.
270. C. Hammann, A. Cooper, D. M. J. Lilley, *Biochemistry* **2001**, *40*, 1423–1429.
271. Q. Dong, E. Stellwagen, N. C. Stellwagen, *Biochemistry* **2009**, *48*, 1047–1055.
272. N. C. Stellwagen, S. Magnusdottir, C. Gelfi, P. G. Righetti, *J. Mol. Biol*. **2001**, *305*, 1025–1033.
273. P. M. Gordon, E. J. Sontheimer, J. A. Piccirilli, *Biochemistry* **2000**, *39*, 12939–12952.
274. S.-I. Nakano, D. J. Proctor, P. C. Bevilacqua, *Biochemistry* **2001**, *40*, 12022–12038.
275. A. K. Brown, J. Li, C. M. B. Pavot, Y. Lu, *Biochemistry* **2003**, *42*, 7152–7161.
276. J. M. Wilcox, D. L. Rempel, M. L. Gross, *Anal. Chem*. **2008**, *80*, 2365–2371.
277. R. B. Martin, *J. Chem. Ed*. **1997**, *74*, 1238–1240.
278. D. Grilley, A. M. Soto, D. E. Draper, *Methods Enzymol*. **2009**, *455*, 71–94.
279. H. Krakauer, *Biopolymers* **1971**, *10*, 2459–2490.
280. R. Römer, R. Hach, *Eur. J. Biochem*. **1975**, *55*, 271–284.
281. Y. Bai, M. Greenfeld, K. J. Travers, V. B. Chu, J. Lipfert, S. Doniach, D. Herschlag, *J. Am. Chem. Soc*. **2007**, *129*, 14981–14988.
282. M. Greenfeld, D. Herschlag, *Methods Enzymol*. **2009**, *469*, 375–389.
283. L. M. Hunsicker, V. J. DeRose, *J. Inorg. Biochem*. **2000**, *80*, 271–281.
284. M. L. Bleam, C. F. Anderson, M. T. Record, Jr., *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 3085–3089.
285. A. L. Feig, M. Panek, W. D. Horrocks, Jr., O. C. Uhlenbeck, *Chem. Biol*. **1999**, *6*, 801–810.
286. K. Levenberg, *Quart. Appl. Math*. **1944**, *2*, 164–168.
287. D. W. Marquardt, *SIAM J. Appl. Math*. **1963**, *11*, 431–431.
288. R. Tribolet, H. Sigel, *Eur. J. Biochem*. **1987**, *163*, 353–363.
289. B. Knobloch, W. Linert, H. Sigel, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7459–7464.
290. B. Knobloch, H. Sigel, A. Okruszek, R. K. O. Sigel, *Chem. Eur. J*. **2007**, *13*, 1804–1814.
291. B. Knobloch, D. Suliga, A. Okruszek, R. K. O. Sigel, *Chem. Eur. J*. **2005**, *11*, 4163–4170.
292. B. Knobloch, M. C. Erat, R. K. O. Sigel, *to be submitted for publication*.
293. P. Auffinger, L. Bielecki, E. Westhof, *Structure* **2004**, *12*, 379–388.
294. M. Feig, B. M. Pettitt, *Biophys. J*. **1999**, *77*, 1769–1781.
295. J. S. Kieft, E. Chase, D. A. Costantino, B. L. Golden, *RNA* **2010**, *16*, 1118–1123.
296. T. Hermann, P. Auffinger, W. G. Scott, E. Westhof, *Nucleic Acids Res*. **1997**, *25*, 3421–3427.
297. M. Boero, J. M. Park, Y. Hagiwara, M. Tateno, *J. Phys. Cond. Matter* **2007**, *19*, 365217–365217.
298. C. C. Correll, B. Freeborn, P. B. Moore, J. A. Steitz, *Cell* **1997**, *91*, 705–712.

299. K. Chin, A. M. Pyle, *RNA* **1995**, *1*, 391–406.
300. C. Montigny, P. Champeil, *Anal. Biochem*. **2007**, *366*, 96–98.
301. P. Leverrier, C. Montigny, M. Garrigos, P. Champeil, *Anal. Biochem*. **2007**, *371*, 215–228.
302. D. P. Hanlon, D. S. Watt, E. W. Westhead, *Anal. Biochem*. **1966**, *16*, 225–233.
303. D. E. Allen, D. J. Baker, R. D. Gillard, *Nature* **1967**, *214*, 906–907.
304. R. Nakon, C. R. Krishnamoorthy, *Science* **1983**, *221*, 749–750.
305. P. D. Prenzler, W. D. McFadyen, *J. Inorg. Biochem*. **1997**, *68*, 279–282.
306. Q. Yu, A. Kandegedara, Y. Xu, D. B. Rorabacher, *Anal. Biochem*. **1997**, *253*, 50–56.
307. K. H. Scheller, T. H. J. Abel, P. E. Polanyi, P. K. Wenk, B. E. Fischer, H. Sigel, *Eur. J. Biochem*. **1980**, *107*, 455–466.
308. B. E. Fischer, U. K. Häring, R. Tribolet, H. Sigel, *Eur. J. Biochem*. **1979**, *94*, 523–530.
309. H. Sigel, K. H. Scheller, B. Prijs, *Inorg. Chim. Acta (Bioinorganic Chemistry)* **1982**, *66*, 147–155.
310. S. Gallo, M. Furler, R. K. O. Sigel, *Chimia* **2005**, *59*, 812–816.
311. P. Champeil, T. Menguy, S. Soulie, B. Juul, A. G. de Gracia, F. Rusconi, P. Falson, L. Denoroy, F. Henao, M. le Maire, J. V. Møller, *J. Biol. Chem*. **1998**, *273*, 6619–6631.
312. H. Sigel, *Coord. Chem. Rev*. **1990**, *100*, 453–539.
313. H. Sigel, F. Hofstetter, R. B. Martin, R. M. Milburn, V. Scheller-Krattiger, K. H. Scheller, *J. Am. Chem. Soc*. **1984**, *106*, 7935–7946.
314. H. Sigel, H. Erlenmeyer, *Helv. Chim. Acta* **1966**, *49*, 1266–1274.
315. H. Sigel, P. E. Amsler, *J. Am. Chem. Soc*. **1976**, *98*, 7390–7400.
316. P. E. Amsler, H. Sigel, *Eur. J. Biochem*. **1976**, *63*, 569–581.
317. H. Sigel, *Angew. Chem. Int. Ed*. **1969**, *8*, 167–177.
318. H. Sigel, C. Flierl, R. Griesser, *J. Am. Chem. Soc*. **1969**, *91*, 1061–1064.
319. K. L. Buchmueller, K. M. Weeks, *Nucleic Acids Res*. **2004**, *32*, e184–e184.
320. N. C. Stellwagen, A. Bossi, C. Gelfi, P. G. Righetti, *Anal. Biochem*. **2000**, *287*, 167–175.
321. N. A. Corfù, B. Song, L.-n. Ji, *Inorg. Chim. Acta* **1992**, *192*, 243–251.
322. M. Sokolowska, W. Bal, *J. Inorg. Biochem*. **2005**, *99*, 1653–1660.
323. K. Luby-Phelps, *Int. Rev. Cytol*. **2000**, *192*, 189–221.
324. N. A. Chebotareva, B. I. Kurganov, N. B. Livanova, *Biochemistry (Moscow)* **2004**, *69*, 1239–1251.
325. D. C. Rees, *J. Mol. Biol*. **1980**, *141*, 323–326.
326. N. K. Rogers, G. R. Moore, M. J. E. Sternberg, *J. Mol. Biol*. **1985**, *182*, 613–616.
327. H. Sigel, R. B. Martin, R. Tribolet, U. K. Haring, R. Malini-Balakrishnan, *Eur. J. Biochem*. **1985**, *152*, 187–193.
328. R. Jin, K. J. Breslauer, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8939–8942.
329. M. A. Young, B. Jayaram, D. L. Beveridge, *J. Phys. Chem. B* **1998**, *102*, 7666–7669.
330. A. Okamoto, K. Tainaka, I. Saito, *Bioconjugate Chem*. **2005**, *16*, 1105–1111.
331. M. Furler, B. Knobloch, R. K. O. Sigel, *Inorg. Chim. Acta* **2009**, *362*, 771–776.
332. R. Tribolet, R. Malini-Balakrishnan, H. Sigel, *J. Chem. Soc., Dalton Trans., Inorg. Chem*. **1985**, 2291–2303.
333. S.-I. Nakano, H. T. Karimata, Y. Kitagawa, N. Sugimoto, *J. Am. Chem. Soc*. **2009**, *131*, 16881–16888.
334. S.-I. Nakano, Y. Kitagawa, H. T. Karimata, N. Sugimoto, *Nucleic Acids Symp. Ser*. **2008**, *52*, 519–520.
335. D. Kilburn, J. H. Roh, L. Guo, R. M. Briber, S. A. Woodson, *J. Am. Chem. Soc*. **2010**, *132*, 8690–8696.
336. D. J. Hirsh, G. W. Brudvig, *Nature Protocols* **2007**, *2*, 1770–1781.
337. H.-J. Steinhoff, *Biol. Chem*. **2011**, *385*, 913–920.
338. J. Voss, L. Salwiński, H. R. Kaback, W. L. Hubbell, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12295–12299.
339. J. Voss, W. L. Hubbell, H. R. Kaback, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12300–12303.
340. J. S. Leigh, *J. Chem. Phys*. **1970**, *52*, 2608–2608.

341. N. Toor, K. Rajashankar, K. S. Keating, A. M. Pyle, *Nature Struct. Mol. Biol*. **2008**, *15*, 1221–1222.
342. G. L. Conn, A. G. Gittis, E. E. Lattman, V. K. Misra, D. E. Draper, *J. Mol. Biol*. **2002**, *318*, 963–973.
343. G. W. Collie, S. M. Haider, S. Neidle, G. N. Parkinson, *Nucleic Acids Res*. **2010**, *38*, 5569–5580.
344. A. Gelasco, S. J. Lippard, *Biochemistry* **1998**, *37*, 9230–9239.
345. G. Sposito, *Met. Ions Biol. Syst*. **1986**, *20*, 1–20.

Chapter 2
Nucleic Acid-Metal Ion Interactions in the Solid State

Katsuyuki Aoki and Kazutaka Murayama

Contents

Abstract Metal ions play a key role in nucleic acid structure and activity. Elucidation of the rules that govern the binding of metal ions is therefore an essential step for better understanding of the nucleic acid functions. This review is as an update to a preceding one (*Metal Ions Biol. Syst.*, **1996**, *32*, 91-134), in which we offered a

K. Aoki (✉)
Department of Materials Science, Toyohashi University of Technology,
Tempaku-cho, Toyohashi 441-8580, Japan
e-mail: kaoki@tutms.tut.ac.jp

K. Murayama
Graduate School of Biomedical Engineering, Tohoku University, Aoba,
Sendai 980-8575, Japan
e-mail: kmura@bme.tohoku.ac.jp

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*,
Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_2,

general view of metal ion interactions with mono-, di-, tri-, and oligonucleotides in the solid state, based on their crystal structures reported before 1994. In this chapter, we survey all the crystal structures of metal ion complexes with nucleotides involving oligonucleotides reported after 1994 and we have tried to uncover new characteristic metal bonding patterns for mononucleotides and oligonucleotides with A-RNA and A/B/Z-DNA fragments that form duplexes. We do not cover quadruplexes, duplexes with metal-mediated base-pairs, tRNAs, rRNAs in ribosome, ribozymes, and nucleic acid–drug and –protein complexes. Factors that affect metal binding to mononucleotides and oligonucleotide duplexes are also dealt with.

Keywords A-DNA • A-RNA • B-DNA • crystal structures • metal ions • mononucleotides • oligonucleotides • Z-DNA

1 Introduction

Metal ion interactions with nucleic acids and their constituents have continuously attracted considerable attention over the past five decades because of the involvement of metal ions and their biological significance in nucleic acid processes, resulting in a large body of information [1–7]. In this field X-ray investigations have made an important contribution by providing accurate information on the geometry of metal binding to nucleic acid derivatives. X-ray studies were repeatedly reviewed dealing with transition metal ion complexes of nucleobases [8–13], nucleosides, and nucleotides [8–12,14–16], as well as alkali and alkaline earth metal ion compounds of nucleotides [9,16]. Special reviews on metal complexes of adenosine 5'-triphosphate (ATP) are also available [17,18]. Also, X-ray structural data for all of the metal compounds of *mono*nucleotides reported before 1984 [19] and before 1994 [16] and those of *oligo*nucleotides [1,16] before 1994 were tabulated and/or reviewed. In the previous review [16], we tried to derive general conclusions from the solid state studies of nucleotide–metal complexes and a special effort was made to catalogue characteristic structural patterns observed in the metal complexes of mononucleotides. During three decades, a number of crystal structures of oligonucleotides have been reported, and recent, especially high-resolution X-ray data of oligonucleotides allow a detailed structural analysis of ion binding on DNA and RNA and, in particular, the observation of monovalent cations. Excellent reviews on DNA–cation interactions are available [20–22].

In this chapter we continue to catalogue new characteristic structural patterns, if any, in the crystal structures of metal complexes of *mono*nucleotides reported after 1994. Also, crystal structures of metal complexes of *synthesized oligo*nucleotides reported after 1994 are surveyed to uncover characteristic metal binding patterns. We do not cover quadruplexes (see Chapter 4), duplexes with metal-mediated base-pairs (see Chapters 10 and 11), tRNAs [23], rRNAs in ribosome [24], ribozymes [25], and nucleic acid–drug and –protein complexes. Nomenclature and chemical structure of the common nucleotides are presented in Figure 1. For the conformational terms of the nucleotide structures refer to [26].

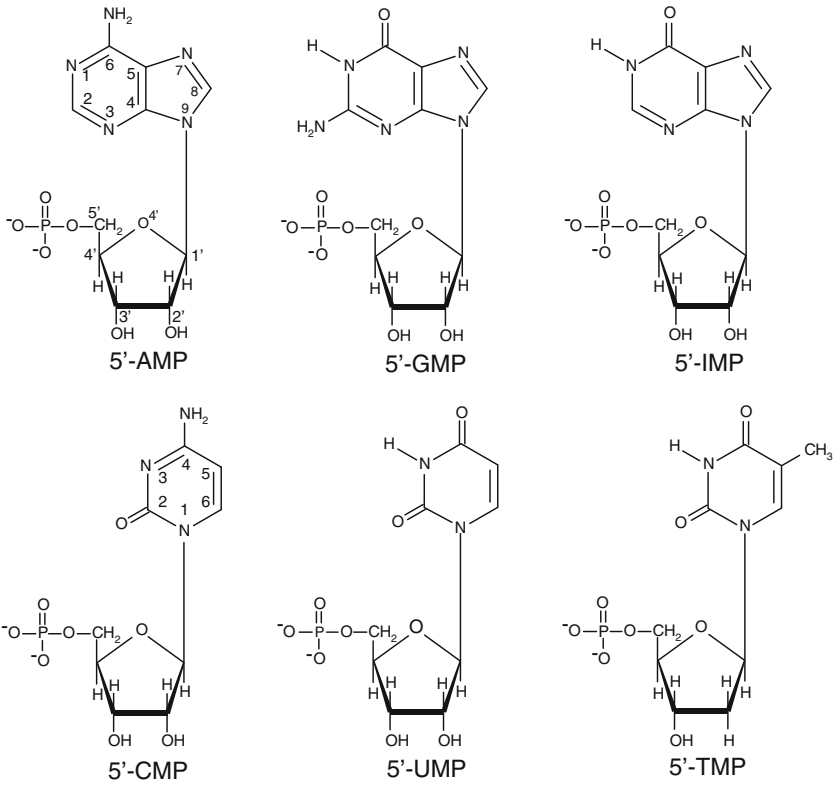


Figure 1 Nomenclature and chemical structure of the common nucleotides. The nucleotides are shown in the *syn* conformation (to save space) even though they mostly occur in the *anti* form.

2 Metal-Mononucleotide Complexes

Only a limited number of crystal structures of metal ion–mononucleotide complexes have been reported after 1994. The total number of the complexes reported after 1994 is summarized in Table 1 for alkali and alkaline earth metal ions and in Table 2 for transition and heavy metal ions, as well as those reported before 1994 in parentheses. Here we deal with only those published after 1994 unless otherwise noted. For those before 1994, refer to Ref. 16.

2.1 Structures of Alkali and Alkaline Earth Metal Ion Complexes

Only a total number of 16 crystal structures are available after 1994, 9 for AMP, one for GMP, 5 for CMP, and one for UMP, and none for hypoxanthine and thymine nucleotides. Their metal coordination data and/or comments are summarized

Table 1 Total number of alkali and alkaline earth metal ion–mononucleotide complexes reported after 1994 and before[a] 1994 in parentheses.

Nucleotide	Sugar[b]	Phosphate[c]	Li^+	Na^+	K^+	Rb^+	Cs^+	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}	Total
AMP	R	5’		4[e]							(1)	5
	R	3’,5’		(2)								2
	R	2’	1[d]	1[f]	1[j]			1[k]	1[l]			5
	D	5’		(1)					(1)			2
ADP	R	5’	(1)		(1)	(1)						3
ATP	R	5’		(2)				(2)	(2)			6
GMP	R	5’		1[g] (2)					(1)		(1)	5
	R	3’,5’		(1)								1
	D	5’		(1)								1
GDP												0
GTP												0
IMP	R	5’		(9)					(2)	(1)	(1)	13
IDP												0
ITP												0
CMP	R	5’		5[h] (1)							(1)	7
	R	3’					(1)					1
	R	2’,3’		(1)								1
	D	5’		(2)								2
CDP	R	5’		(3)								3
CTP												0

UMP	R	5’		1[i] (2)							(1)	4
	R	3’		(1)	(1)							2
	D	5’		(1)								1
UDP	R	5’		(1)	(2)							3
UTP												0
TMP	D	5’							(2)			2
TDP												0
TTP												0
Total			2	42	5	1	1	3	9	1	5	69

[a] For individual complexes reported before 1994, refer to Table 1 in [16]. [b] Abbreviations for sugars: R = ribose; D = deoxyribose.
[c] Abbreviations for phosphates: 5’ = 5’-phosphate; 3’,5’ = 3’,5’-cyclic phosphate; 2’ = 2’-phosphate; 3’ = 3’-phosphate; 2’,3’ = 2’,3’-cyclic phosphate. [d] [31].
[e] [27,28,29,30], containing a $[Mo_5O_{15}]$ cluster [29] or $[UO_2]^{2+}$ ions [30]. [f] [31]. [g] A salt of the Pt^{2+}-GMP complex [32].
[h] [33,34,36]. [i] [36]. [j] [31]. [k] [31]. [l] [31].

Table 2 Total number of transition and heavy metal ion–mononucleotide complexes reported after 1994 and before[a] 1994 in parentheses.

Nucleotide	Sugar[b]	Phosphate[c]	Cu^{2+}	Zn^{2+}	Cd^{2+}	Mn^{2+}	Fe^{2+}	Co^{2+} Co^{3+}	Ni^{2+}	Pd^{2+}	Pt^{2+}	Mo^{4+}	U^{6+}	Total
AMP	R	5’	(2)						1[h] (1)		(2)	1[l]	1[m]	8
	R	3’,5’		(1)										1
ADP	R	5’	1[d] (1)											2
ATP	R	5’	1[e] (1)	(2)	(1)	(1)		(1)						7
GMP	R	5’	(4)	(1)	(1)	(1)	(1)	(2)	(3)	1[i]	6[k] (2)			22
	R	3’	(2)											2
	R	2’	(1)											1
	R	3’,5’	(1)											2
	D	5’	(1)											5
GDP														0
GTP														0
IMP	R	5’	1[f] (9)	(2)	(1)	(1)	(1)	(3)	(2)	1[j]	(4)			25
	R	3’,5’	(1)					(1)						2
	D	5’						(1)						1
IDP														0
ITP														0
CMP	R	5’	(1)	(2)	(2)	(1)		(1)		(1)				8
	R	3’								(2)				2
	D	5’		(1)	(2)									3

CDP														0
CTP														0
UMP	R	5’	1[g] (2)		(2)			(1)						6
	D	5’	(1)	(1)										2
UDP														0
UTP														0
TMP														0
TDP														0
TTP														0
Total			31	9	10	4	2	11	10	5	14	2	1	99

[a] For individual complexes reported before 1994, refer to Table 2 in [16]. [b] Abbreviations for sugars: R = ribose; D = deoxyribose.
[c] Abbreviations for phosphates: 5’ = 5’-phosphate; 3’,5’ = 3’,5’-cyclic phosphate; 2’ = 2’-phosphate; 3’ = 3’-phosphate. [d] [44]. [e] [45,46]. [f] [46]. [g] [46]. [h] [43].
[i] [47]. [j] [47]. [k] [32,47–50], containing a Na^+ salt [32]. [l] [29], containing Na^+ ions.
[m] [30], containing Na^+ ions.

in Table 3. For adenine nucleotides, in both the highly hydrated sodium salts Na_2·(5'-AMP)·12H_2O [27] and Na_2·(N6-methyl-5'-AMP)·12H_2O [28], which are isomorphs to each other, the sodium ion is fully hydrated and does not directly bind to the nucleotide. Interestingly, this is also the case for [Na(5'-dAMPH)]·6H_2O [37], despite the different crystal packings. In the two sodium salts of a pentamolybdate complex of 5'-AMPH [29] and an uranium(VI) complex of 5'-AMP [30], Na^+ ions bind solely to the phosphate groups of nucleotide molecules in the $[(Mo_5O_{15})(5'\text{-AMPH})_2]^{2-}$ fragment [29] or in the trimeric $[(UO_2)_3(O)(5'\text{-AMP})_3]^-$ framework [30], forming an O(P)–Na–O(W)–Na–O(P) linkage (where O(P) and O(W) denote phosphate oxygen and water oxygen, respectively) in the former [29] or a Na–O(P) bond for each of the AMP molecules in the latter [30]. Crystal structures of five (Li^+, Na^+, K^+, Ca^{2+}, and Mg^{2+}) salts of 2'-AMP are reported [31]; results on metal complexes with a 2'-nucleotide are quite scarce, to date only [Cu(2'-GMP)(H_2O)$_3$] is reported [38]. Li^+ ions bind directly to ribose O3' and the phosphate oxygens of the nucleotide. Na^+ and K^+ ions bind directly to phosphate and the ribose O3' and O4' oxygens and indirectly to N7 via a water oxygen, where the involvement of ribose O4' in the metal coordination is sparsely documented for Na^+–5'-ATP [39,40]. Ca^{2+} and Mg^{2+} ions directly bind to phosphate oxygens. In all the 2'-AMP structures, the neutral adenine bases are in the *syn* conformation stabilized by an intramolecular O5'–H···N3 hydrogen bond, except for the Mg^{2+} salt for which the base adopts *anti* conformation stabilized by the indirect O(P)–Mg–O(W)···N3 linkage resulting in macrochelation between the base and the phosphate group.

In guanine nucleotides, such as Na_2[Pt(Me_4dae)(5'-GMP)$_2$]·7H_2O [32], a seven-coordinated $[Na(H_2O)]^+$ ion binds directly to symmetry-related four GMP molecules through O6 of the base, hydroxyl O2' and O3' of the sugar, and three phosphate oxygens. The Na^+–O6 bonding is also observed in [Na_2(5'-dGMP)]·4H_2O [41].

For cytosine nucleotides, five sodium salts of 5'-CMP with different hydration degrees are reported [33–35]. For [Na_8(5'-CMP)$_4$]·37H_2O [34],which crystallized from an aqueous solution, the phase transition proceeds depending on the relative humidity around the crystal to [Na_8(5'-CMP)$_4$]·32.5H_2O [34] followed by [Na_4(5'-CMP)$_2$]·13H_2O (space group $P2_12_12_1$) [34]. The last one also crystallized from a water-methanol solution in the space group $P2_1$ form [33]. In all five structures [33–35], sodium coordination with hydroxyl oxygens O2' and O3' of the ribose is commonly observed, while metal ion-bridging between the base through N3 (and O2) and phosphate oxygen occurs only in less hydrated [Na_4(5'-CMP)$_2$]·13H_2O [33,34] and [Na_4(5'-CMP)$_2$]·13.14H_2O·0.54MeOH [35].

In uracil nucleotide salts, such as [Na_2(5'-UMP)]·13H_2O [36], the metal ion bridges the hydroxyl O2' and O3' oxygens of the ribose. Sodium ion binding to O3' of the deoxyribose is observed in [Na_2(5'-dUMP)]·5H_2O [42], but in this less hydrated salt, metal ions further bind to the base through O4 and the phosphate group.

In summary, alkali and alkaline earth metal ions generally coordinate with water oxygens in well hydrated structures. They bind to the phosphate, sugar and base moieties in less hydrated structures.

Table 3 Coordination data for alkali and alkaline earth metal ion–mononucleotide complexes[a,b].

Complex Geometry about metal (M)	Coordination site(s) (L) on nucleotide	M–L distance (Å)	Sugar puckering[c]	φ_{OO}[c]	χ_{CN}[c]	Ref.
And/or additional remarks						
$Na_2 \cdot (5'\text{-AMP}) \cdot 12H_2O$						
	none		–[d]	–[d]	–[d]	[27]
Na^+ ions form a columnar structure of fully hydrated $[Na(H_2O)_4]_n^{n+}$, where the neighboring octahedra share the edge defined by two water molecules.						
$Na_2 \cdot (N6\text{-methyl-}5'\text{-AMP}) \cdot 12H_2O$						
	none		–	–	–	[28]
Crystals are isomorphous to those of the parent 5'-AMP salt [27].						
$Na_2[(Mo_5O_{15})(5'\text{-AMPH})_2] \cdot 6H_2O$						
Na 6-Coord.	O(P)	2.43(4)	C2'-*endo*	*gg*	*anti*	[29]
One AMPH molecule is independent. Protonation at N1 of the base.						
$Na_{8.5} \cdot [(UO_2)_3(O)(5'\text{-AMP})_3]_2 \cdot 33H_2O$						
Two independent trimeric $[(UO_2)_3(O)(5'\text{-AMP})_3]$ molecules in the asymmetric unit. Each of partially occupied six Na^+ ions is connected to one of the phosphate oxygens in each of six AMP molecules, which adopts *gg* and *anti* conformations.						[30]
$[Li_2(2'\text{-AMPH})_2] \cdot 6H_2O$						
Li1 Tetr.	O3'(Mol.1)	1.988(9)	C2'-*endo*	*gg*	*syn*	[31]
	O(P)(Mol.2)	1.877(9)	C2'-*endo*	*gg*	*syn*	
Li2 Tetr.	O(P)(Mol.2)	1.951(8)				
Two independent molecules in the asymmetric unit. Protonation at the phosphate group.						
$[Na_2(2'\text{-AMPH})_2] \cdot 3H_2O$						
Na1 6-Coord.	O3'(Mol.1)	2.473(3)	C2'-*endo*–C3'-*exo*	*gg*	*syn*	[31]
	O4'(Mol.2)	2.712(3)	C2'-*endo*–C3'-*exo*	*gg*	*syn*	

(continued)

Table 3 (continued)

Complex Geometry about metal (M)	Coordination site(s) (L) on nucleotide	M–L distance (Å)	Sugar puckering[c]	φ_{OO} [c]	χ_{CN} [c]	Ref.
	O(P)(Mol.1*)	2.490(4)				
	O(P)(Mol.2*)	2.514(2)				
Na2 5-Coord.	O3'(Mol.2)	2.301(3)				
	O4'(Mol.1)	2.451(4)				
	O(P)(Mol.1)	2.388(3)				
Two independent molecules in the asymmetric unit. Mol.1* (Mol.2*) is symmetry-related to Mol.1 (Mol.2). Protonation at the phosphate group.						
$[K(2'\text{-AMPH})]\cdot 1.5H_2O$						
6-Coord.	O3'	2.674(6)	C2'-*endo*–C3'-*exo*	*gg*	*syn*	[31]
	O4'	3.040(6)				
	O(P)	2.685(4)				
	O(P)	2.719(7)				
Protonation at the phosphate group. The $[K(H_2O)_2]^+$ ion binds directly to symmetry-related four nucleotide molecules.						
$[Ca_{0.5}(2'\text{-AMPH})]\cdot 4.5H_2O$						
6-Coord.	O(P)	2.184(7)	C2'-*endo*	*gg*	*syn*	[31]
Protonation at the phosphate group. The $[Ca(H_2O)_4]^{2+}$ ion, which locates on a 2-fold symmetry axis, binds directly to two O(P)s of two symmetry-related nucleotides.						
$[Mg(2'\text{-AMP})]\cdot 3H_2O$						
Oct.	O(P)	2.114(2)	C3'-*endo*–C2'-*exo*	*gt*	*anti*	[31]
	O(P)	2.024(2)				
	O(P)	2.055(2)				
The $[Mg(H_2O)_3]^{2+}$ ion binds directly to three O(P)s of three symmetry-related nucleotides.						

$Na_2[Pt(Me_4dae)(5'\text{-GMP})_2]\cdot 7H_2O$							
The Pt atom rides on a two-fold crystallographic axis. The asymmetric unit includes a Na^+ ion and a GMP molecule, which adopts C2'-*endo*, *gg*, *anti* conformations. A seven-coordinated $[Na(H_2O)]^+$ ion binds to four symmetry-related GMP molecules (Mol.1–Mol.4), through O6 for Mol.1, O2' and O3' for Mol. 2, O(P) for Mol. 3, and 2O(P)s for Mol.4.							[32]
$[Na_4(5'\text{-CMP})_2]\cdot 13H_2O$							
Na1	5-Coord.	N3(Mol.1)	2.400(10)	C2'-*endo*	*gg*	*anti*	[31]
		O2(Mol.1)	2.855(10)				
		O(P)(Mol.2)	2.361(9)	C2'-*endo*–C3'-*exo*	*gg*	*anti*	
Na2	6-Coord.	O2'(Mol.2)	2.413(9)				
		O3'(Mol.2)	2.306(9)				
Na3	5-Coord.	N3(Mol.2)	2.521(10)				
		O(P)(Mol.1)	2.334(8)				
Space group $P2_1$. Two independent CMP molecules and four independent Na^+ ions in the asymmetric unit. N3–Na1–O2 chelation. O2'–Na2–O3' chelation. Na^+-bridging between the base and the phosphate group for Na1 and Na3.							
$[Na_4(5'\text{-CMP})_2]\cdot 13H_2O$							
Space group $P2_12_12_1$. Two independent CMP molecules and four independent Na^+ ions in the asymmetric unit, Na1 is fully hydrated. O2'–Na2–O3' chelation. N3–Na–O(P) bridging for Na3 and Na4. Two independent CMP molecules adopt C1'-*exo*–C2'-*endo* or C2'-*endo*, *gg*, and *anti* conformations.							[34]
$[Na_4(5'\text{-CMP})_2]\cdot 13.14H_2O\cdot 0.54MeOH$							
Na1	5-Coord.	N3(Mol.1)	2.3852(9)	C2'-*endo*	*gg*	*anti*	[35]
		O2(Mol.1)	2.8008(9)				
		O(P)(Mol.1')	2.3621(8)				
Na2	6-Coord.	O2'(Mol.2)	2.3995(8)	*C2'-endo*	*gg*	*anti*	
		O3'(Mol.2)	2.3093(8)				
Na3	5-Coord.	N3(Mol.2)	2.5201(8)				
		O(P)(Mol.2')	2.3252(7)				
Space group $P2_12_12_1$. Two independent CMP molecules and four independent Na^+ ions in the asymmetric unit. N3–Na1–O2 chelation for Mol.1 of CMP. O2'–Na2–O3' chelation for Mol.2. Na^+-bridging between the base and the phosphate group for Na1 and Na3.							

(continued)

Table 3 (continued)

Complex Geometry about metal (M)	Coordination site(s) (L) on nucleotide	M–L distance (Å)	Sugar puckering[c]	φ_{OO}[c]	χ_{CN}[c]	Ref.
$[Na_8(5'\text{-CMP})_4]\cdot 37H_2O$						
Out of eight independent Na^+ ions in the asymmetric unit, four are fully hydrated and the other four coordinate with the hydroxyl groups of riboses. The phosphates are fully hydrated and there are no direct interactions between phosphates and Na^+ ions. Four independent CMP molecules adopt C1'-*exo*–C2'-*endo* or C2'-*endo*, *gg*, and *anti* conformations.						[34]
$[Na_8(5'\text{-CMP})_4]\cdot 32.5H_2O$						
Out of eight independent Na^+ ions in the asymmetric unit, four are coordinated with the hydroxyl groups of riboses. The phosphates are fully hydrated and there are no direct interactions between phosphates and Na^+ ions. Four independent CMP molecules adopt C1'-*exo*– C2'-*endo* or C2'-*endo*, *gg*, and *anti* conformations.						[34]
$[Na_2(5'\text{-UMP})]\cdot 13H_2O$						
7-Coord.	O2'	2.55(5)	C2'-*endo*	*gg*	*anti*	[36]
	O3'	2.27(5)				
O2'-Na-O3' chelation.						

[a] Data reported after 1994. [b] Abbreviations: M = metal ion; L = ligand; P = phosphate group; O(P) = phosphate oxygen; Me_4dae = *N*,*N*,*N*',*N*'-tetramethyl-1,2-diaminoethane; Tetr. = tetrahedral; Oct. = octahedral. [c] For the definition of these conformational terms, see [26], pp. 17–24. [d] No data are given.

2.2 Structures of Transition and Heavy Metal Ion Complexes

In the previous review [16], we catalogued the structures of transition and heavy metal ion–mononuclotide *binary* complexes, based on their structural characteristics, into five patterns for *purine* nucleotides and three patterns for *pyrimidine* nucleotides: for *purine* nucleotides, (***1***) a 1:1 M-NMP monomeric structure with the general formula [M(5'-nucleotide)$(H_2O)_5]\cdot nH_2O$, (***2***) a 1:2 mononuclear bis(5'-NMP) structure with the general formula *cis*-$[M(5'\text{-}NMP)_2(amine)_x(H_2O)_y]$, (***3***) a 1:2 mononuclear bis(5'-ATP) structure with the general formula $[M(5'\text{-}ATPH)_2]^{4-}$, (***4***) a 2:2 dimeric structure with the general formula $[M(NMP)]_2$, and (***5***) a polymeric structure; for *pyrimidine* nucleotides, (***i***) a 1:2 mononuclear $M(NMP)_2$, (***ii***) a 2:2 dimeric $[M(NMP)]_2$, and (***iii***) a polymeric M–nucleotide complex. On the other hand, for *aromatic amine*–metal ion–nucleotide *ternary* complexes, the structures are conveniently classified into four categories: (***a***) a dimeric structure with "phosphate-only" metal bonding, (***b***) a structure with "base-only" metal bonding, (***c***) a structure with no direct metal bonding but with intercalation, and (***d***) a tetranuclear bis(NMP) structure with "base and phosphate" metal bonding.

A total number of 15 crystal structures are available after 1994, 5 for adenine nucleotides, 7 for GMP, 2 for IMP, one for UMP, and none for cytosine and thymine nucleotides. Their metal coordination data and/or comments are summarized in Table 4. For adenine nucleotides, $[Ni(5'\text{-}AMP)(H_2O)_5]\cdot H_2O$ [43] belongs to the structure type (***1***) of the metal-nucleotide *binary system* complex. Common features for this type of structure are: the "M–N7 base-only" metal bonding, three intramolecular interligand hydrogen bonds that stabilize the structure (one between a water ligand and O6 or N6 of the base and two between water ligands and phosphate oxygens), *anti*, C3'- or C2'-*endo*, *gauche-gauche* conformations (for the definition of these conformational terms, see Ref. 26). The crystal structure of the same complex was once reported by other authors [52] and was the only example for the structure type (***1***) for 5'-AMP, but, due to the poor quality of the data caused by the crystal twinning problem, the formation of an intramolecular N7–Ni–O(W)···H–N6 hydrogen bond remained to be proven. Such an intramolecular hydrogen bonding is now confirmed by successfully locating all hydrogen atoms [43]. In this structure, the N6···O(W) distance is 2.958(5) Å and considerably longer than the O6···O(W) distance of, for example, 2.838(13) Å in $[Ni(5'\text{-}IMP)(H_2O)_5]\cdot 2H_2O$ [53], a 5'-6-oxo-purine nucleotide. This Ni^{2+}–5'-AMP complex is of particular importance because it provides a rare but clear example showing that the amino group is able to form a hydrogen bond with the ligating water molecule, which eventually acts here as an H-bonding acceptor in contrast to its usual participation as a H-bonding donor, since normally the two hydrogen atoms point outward to lie nearly in the equatorial coordination plane [22]. Two metal cluster complexes, $[(Mo_5O_{15})(5'\text{-}AMPH)_2]\cdot Na_2\cdot 6H_2O$ [29] and $[(UO_2)_3(O)(5'\text{-}AMP)_3]_2\cdot Na_{8.5}\cdot 33H_2O$ [30], provide two novel types of structures for metal–nucleotide *binary system* complexes, that is, "(***6***) a phosphate-only metal bonding" for the former and "(***7***) a trimeric structure with the metal bonding to both the sugar and phosphate" for the latter. $[(Mo_5O_{15})(5'\text{-}AMPH)_2]^{2-}$ is isostructural with $[(Mo_5O_{15})(OPO_3)_2]^{6-}$ [54] and thus, the AMP molecule behaves here as a phosphate derivative.

Table 4 Coordination data for transition and heavy metal ion–mononucleotide complexes[a,b].

Complex Geometry about metal (M)	Coordination site(s) (L) on nucleotide	M–L distance (Å)	Sugar puckering[c]	φ_{OO}[c]	χ_{CN}[c]	Ref.
And/or additional remarks						
$[Ni(5'\text{-AMP})(H_2O)_5]\cdot H_2O$						
Oct.	N7	2.066(3)	C3'-*endo*	*gg*	*anti*	[43]
$Na_2[(Mo_5O_{15})(5'\text{-AMPH})_2]\cdot 6H_2O$						
Mo1 6-Coord.	O(P)	2.40(3)	C2'-*endo*	*gg*	*anti*	[29]
Mo2 6-Coord.	O(P)	2.41(3)				
	O(P)	2.29(2)				
Mo3 6-Coord.	O(P)	2.25(3)				
	O(P)	2.39(3)				
Protonation at N1 of the base. Two-fold crystallographic symmetry in the $[(Mo_5O_{15})(5'\text{-AMPH})_2]^{2-}$ structure, where one AMPH molecule and three Mo atoms are independent.						
$Na_{8.5}\cdot[(UO_2)_3(O)(5'\text{-AMP})_3]_2\cdot 33H_2O$						
Two independent trimeric $[(UO_2)_3(O)(5'\text{-AMP})_3]$ molecules in the asymmetric unit. Six independent AMP molecules adopt *gg* and *anti* conformations.						[30]
$[Cu(terpy)(5'\text{-ADP})]\cdot[Cu(terpy)(H_2O)]\cdot(5'\text{-ADPH}_2)\cdot 16H_2O$						
Sq. pyr.	O(α-P)	1.919(8)	C2'-*endo*	*gg*	*high-anti*	[44]
	O(β-P)	2.244(10)	C2'-*endo*			
Bidentate chelation to α- and β-P of ADP with the Δ handedness[d] in the $[Cu(terpy)(5'\text{-ADP})]^-$ structure.						
$[Cu_4(bpy)_4(5'\text{-ATP})_2]\cdot 20.5H_2O$						
Cu1 Sq. pl.	O(α-P)(Mol.1)	1.943(4)	C2'-*endo*	*gg*	*anti*	[45,46]
	O(γ-P)(Mol.1)	1.951(4)				
Cu2 Sq. pyr.	O(β-P)(Mol.1)	1.982(4)				
	O(γ-P)(Mol.1)	1.934(4)				
	O(β-P)(Mol.2)	2.263(4)	C2'-*endo*	*gg*	*anti*	

Cu3	Sq. pl.	O(α-P)(Mol.2)	1.922(4)				
		O(γ-P)(Mol.2)	1.919(4)				
Cu4	Sq. pyr.	O(β-P)(Mol.2)	1.935(4)				
		O(γ-P)(Mol.2)	1.921(4)				
		O(β-P)(Mol.1)	2.247(4)				
Metal ion-bridged adenine base–bpy π–π stacking.							
$[Pd(en)(5'\text{-}GMPH)_2]\cdot 9H_2O$							
Sq. pl.		N7	2.036(4)	C3'-*endo*	*gg*	*anti*	[47]
Protonation at the phosphate group. The Pd atom rides on a two-fold crystallographic axis. Head-to-tail orientation of the two bases and right-handed chirality (ΔHT conformation). Intramolecular interligand N(en)-H···O(P) hydrogen bonds.							
$[Pt(en)(5'\text{-}GMPH)_2]\cdot 9H_2O$							
Sq. pl.		N7	2.045(4)	C3'-*endo*	*gg*	*anti*	[47]
Isostructural with the Pd^{2+} complex [47].							
$[Pt(MeSCH_2CH_2SMe)(5'\text{-}GMPH)_2]\cdot 6H_2O$							
Sq. pl.		N7	2.054(10)	C3'-*endo*	*gg*	*anti*	[48]
Protonation at the phosphate group. The Pt atom rides on a two-fold axis. ΔHT conformation.							
$[Pt\{(RR)\text{-}Me_4dach\}(5'\text{-}GMPH)_2]\cdot 10D_2O$							
Sq. pl.		N7	2.028(9)	C3'-*endo*	*gg*	*anti*	[49]
Protonation at the phosphate group. The Pt atom rides on a two-fold axis. ΔHT conformation. Intramolecular interligand N–H···O(P) hydrogen bonds.							
$[Pt\{(SS)\text{-}Me_4dach\}(5'\text{-}GMPH)_2]\cdot 14D_2O$							
Sq. pl.		N7	2.028(9)	C3'-*endo*	*gg*	*anti*	[49]
Protonation at the phosphate group. The Pt atom rides on a two-fold axis. ΔHT conformation. Intramolecular interligand N–H···O(P) hydrogen bonds.							
$Na_2[Pt(Me_4dae)(5'\text{-}GMP)_2]\cdot 7H_2O$							
Sq. pl.		N7	2.031(9)	C2'-*endo*	*gg*	*anti*	[32]
The Pt atom rides on a two-fold crystallographic axis. ΛHT conformation. Intramolecular interligand N–H···O(P) hydrogen bonds.							
$[Pt(bpm)(L\text{-arginine})]\cdot(5'\text{-}GMP)\cdot 5H_2O$							
Sq. pl.		None		–[e]	*gg*	*anti*	[50]
Guanine base–bpm π–π stacking.							

(continued)

Table 4 (continued)

Complex Geometry about metal (M)	Coordination site(s) (L) on nucleotide	M–L distance (Å)	Sugar puckering[c]	φ_{OO}[c]	χ_{CN}[c]	Ref.
And/or additional remarks						
$[Cu_4(phen)_4(5'\text{-IMP–H})_2(H_2O)_4](NO_3)_2\cdot 21.5H_2O$						
Cu1 Sq. pyr.	N7(Mol.1)	2.018(4)	C3'-*endo*	*gg*	*anti*	[46]
	O6(Mol.2)	1.980(4)	C3'-*endo*	*gg*	*anti*	
Cu2 Sq. pyr.	N7(Mol.2)	2.018(4)				
	O6(Mol.1)	1.955(4)				
Cu3 Sq. pyr.	N1(Mol.2)	2.021(5)				
	O(P)(Mol.1)	1.943(4)				
Cu4 Sq. pyr.	N1(Mol.1)	2.042(5)				
	O(P)(Mol.2)	1.946(4)				
Deprotonated at N1 of the base.						
$[Pd(en)(5'\text{-IMPH})_2]\cdot 11H_2O$						
Sq. pl.	N7	2.053(8)	C3'-*endo*	*gg*	*anti*	[51]
Protonation at the phosphate group. The Pd atom rides on a 2-fold axis. ΔHT conformation. Intramolecular interligand N(en)-H···O(P) hydrogen bond.						
$[Cu(bpy)(5'\text{-UMP–H})(H_2O)]_2[Cu_4(bpy)_4(\alpha\text{-D-Glc-1P})_2(OH)(H_2O)_2]Cl\cdot 22H_2O$						
Sq. pyr.	N3	2.033(8)	C2'-*endo*	*gg*	*anti*	[46]
	O(P)	1.917(7)				
The $[Cu(bpy)(H_2O)]^{2+}$ unit bridges two UMP molecules, one molecule through the deprotonated N3 of the base and the other through a phosphate oxygen, forming a $[Cu(bpy)(H_2O)(5'\text{-UMP–H})]_n^{2n+}$ polymeric structure.						

[a] Data reported after 1994. [b] Abbreviations: M = metal ion; L = ligand; P = phosphate group; O(P) = phosphate oxygen; terpy = 2,2',2"-terpyridine; bpy = 2,2'-bipyridyl; en = ethylenediamine; *gg* = *gauche-gauche*; Me_4dach = *N*,*N*,*N'*,*N'*-tetramethyl-1,2-diaminocyclohexane; Me_4dae = *N*,*N*,*N'*,*N'*-tetramethyl-1,2-diaminoethane; bpm = 2,2'-bipyrimidine; phen = 1,10-phenanthroline; α-D-Glc-1P = α-D-glucopyranose-1-phosphate; Oct. = octahedral; Sq. pyr. = square pyramidal; Sq. pl. = square planar. [c] For the definition of these conformational terms, see [26] pp. 17–24. [d] For the definition of the Δ handedness, see [26], p. 217. [e] No data are given.

In the trimeric UO_2–AMP complex, each of three UO_2^{2+} units is connected to the O2' and O3' hydroxyls of the sugar unit of an AMP molecule and to both O3' and a phosphate oxygen of the neighboring AMP. This is the first example for O3'–M–O(P) macrochelate formation. The O2'–M–O3' chelation is also quite scarce in heavy metal ion complexes, rare examples being [Cu(5'-UMP)(imidazole)$_2$(H_2O)]·4H_2O [55] and [Cd_2(5'-IMP)(5'-IMPH)$_2$(H_2O)$_6$]·6H_2O [56]. An *aromatic amine*–metal ion–nucleotide *ternary* complex [Cu(terpy)(5'-ADP)]·[Cu(terpy)(H_2O)]·(5'-ADPH$_2$)·16H_2O [44] does not fall into any of the four categories (***a***) – (***d***) noted above, thus adding a novel type of structure: "(***e***) a monomeric structure with phosphate-only metal bonding". The $[Cu(terpy)]^{2+}$ unit binds to the pyrophosphate group of 5'-ADP to form a monomeric $[Cu(terpy)(5'\text{-}ADP)]^-$ molecular species, where the adenine ring moiety is far away from the terpy ligand and thus, any intramolecular ring-ring interaction is impossible. This is in contrast with the type (***a***) *ternary* complexes [M(5'-NMP or 5'-NTP)(*bidentate* aromatic amine)]$_2$ or a tetranuclear bpy-Cu^{2+}-5'-ADPH complex [57], for which the metal ion-bridged face-to-face (stacking) ring-ring interaction occurs, for example, in $[Cu(5'\text{-}AMPH)(bpy)(H_2O)]_2^{2+}$ [58] or $[Cu_4(bpy)_4(5'\text{-}ADPH)_2(H_2O)_2(NO_3)_2]^{2+}$ [57], or the edge-to-face interaction in [Cu(5'-CMP)(dpa)(H_2O)]$_2$ [59]. The coordinative number of the aromatic amine ligand used seems to affect the structure formed, that is, a *monomeric* structure forms for tridentate ligand while a bidentate ligand allows the formation of *dimeric* or *tetranuclear* structures. Another ternary complex, [Cu_4(bpy)$_4$(5'-ATP)$_2$]·20.5H_2O [45,46], resembles [Cu(phen)(5'-ATPH$_2$)]$_2$ [57], [Zn(bpy)(5'-ATPH$_2$)]$_2$ [60], or $[Cu_4(bpy)_4(5'\text{-}ADPH)_2(H_2O)_2(NO_3)_2]^{2+}$ [57] in that it involves the phosphate-only metal bonding and takes the folded structure with the metal ion-bridged intramolecular adenine–aromatic amine ring-ring stacking.

For guanine nucleotides, among the seven crystal structures reported, five are Pd^{2+}– or Pt^{2+}–diamine complexes of 5'-GMP [32,47,49], possible models for bonding of the antitumor drug *cis*-$[Pt(NH_3)_2]^{2+}$ to two adjacent intrastranded guanine bases in DNA. These belong to the structure type (***2***) "a 1:2 mononuclear bis (5'-NMP) structure with the general formula *cis*-[M(5'-NMP)$_2$(amine)$_x$(H_2O)$_y$]". A general rule for this type of structure is: "*cis*-N7–M–N7 base-only" metal bonding, "head-to-tail" relative orientation of the bases, *anti*, C2'-*endo* or C3'-*endo*, *gauche-gauche* conformations. As with all crystal structures known so far of *cis*-[metal(AN$_2$)(nucleotide)$_2$] complexes (AN$_2$ stands for two amines or a diamine), [M(en)(5'-GMPH)$_2$]·9H_2O (M = Pd^{2+}, Pt^{2+}) [47] and [Pt{(*RR* or *SS*)-Me_4dach}(5'-GMPH)$_2$]·*n*D_2O (n = 10 or 14) [49] have the head-to-tail orientation of the two bases in the Δ (right-handed chirality) configuration (ΔHT conformation), where the head-to-tail orientation means that H8's of the two nucleotides are on opposite sides of the coordination plane and Δ is defined [61] as having a negative slope for the line joining the H8's of the two purines when the bis(nucleotide) complex is viewed from the nucleotide coordination side (see Figure 1 in Ref. 32). In sharp contrast, [Pt(Me_4dae)(5'-GMP)$_2$]·Na_2·7H_2O [32] is the only exception adopting the head-to-tail orientation with the Λ configuration (ΛHT conformation). In all these complexes, a pair of intramolecular H-bonds between the anionic phosphate groups and the coordinated en ligands is formed to stabilize the structure. The structure of [Pt($MeSCH_2CH_2SMe$)(5'-GMPH)$_2$]·6H_2O [48] is almost identical to that in [Pt(en)(5'-GMPH)$_2$]·9H_2O

[47] but there forms no intramolecular hydrogen bond between the dithioether ligand and the monoanionic phosphate group, indicating that the orientation of the sugar-phosphate side chain in these bisnucleotide complexes is unrelated to the intramolecular H-bonding. [Pt(bpm)(L-arginine)]·(5'-GMP)·5H_2O [50] involves no direct metal bonding to GMP but a stacking interaction between guanine and bpm rings occurs in the crystal lattice.

For hypoxanthine nucleotides, an *aromatic amine*–metal ion–nucleotide *ternary* complex, [Cu_4(phen)$_4$(5'-IMP–H)$_2$(H_2O)$_4$](NO_3)$_2$·21.5H_2O [46], belongs to the structure type (***d***) and this structure is almost identical to that of [Cu_4(phen)$_4$(5'-IMP–H)$_2$(H_2O)$_4$](NO_3)$_2$·4H_2O [62]. It forms a tetranuclear bis(NMP) structure with "base and phosphate" metal bonding involving unusual M–N1(deprotonated) and M–O6 bonding in addition to the usual M–N7 bonding. The structure of [Pd(en)(5'-IMPH)$_2$]·11H_2O [51] is essentially the same as that of [Pd(en)(5'-GMPH)$_2$]·9H_2O [47] noted above.

For uracil nucleotides, an *aromatic amine*–metal ion–nucleotide *ternary* complex, [Cu(bpy)(5'-UMP–H)(H_2O)]$_2$[Cu_4(bpy)$_4$(α-D-Glc-1P)$_2$(OH)(H_2O)$_2$]Cl·22H_2O [46], provides a novel type of structure, namely "(***f***) a polymeric structure with "base and phosphate" metal bonding". Each [Cu(bpy)(H_2O)]$^{2+}$ fragment binds to two UMP molecules, one molecule through the base (at the deprotonated N3) and the other through the phosphate group to afford a C_2-helical coordination polymer.

In summary, we add the four novel types of structures: (***6***) a "phosphate-only" metal bonding and (***7***) a trimeric structure with "sugar and phosphate" metal bonding for metal–*purine* nucleotide *binary* complexes, and (***e***) a monomeric structure with "phosphate-only" metal bonding and (***f***) a polymeric structure with "base and phosphate" metal bonding for *ternary aromatic amine*–metal ion–nucleotide complexes.

2.3 Preferred Metal Binding Sites or Modes

There are three potential metal binding groups on a nucleotide: phosphate, sugar, and base moieties. Mono-, di-, and triphosphate groups are good ligands not only for alkali and alkaline earth metal ions but transition and heavy metal ions as well.

Sugar hydroxyl groups are good ligands for alkali and alkaline earth metal ions but not for transition and heavy metal ions though exceptions exist [55,56].

Ring nitrogens of bases are good targets for all types of metal ions, especially for transition and heavy metal ions. Coordination to the ring nitrogens is favored over binding to the exocyclic amino and keto groups: the N7 site is the most preferred metal binding site for purine bases and the N3 site for cytosine.

Keto substituents at O2 of cytosine and O2 and O4 of uracil (or thymine) are usual metal binding sites for alkali and alkaline earth metal ions. Metal bonding to guanine O6 is less favored but there are some examples [32,41,63–67] involving the N7–metal–O6 chelation [63–65]. Amino substituents at N6 of adenine, N2 of guanine, and N4 of cytosine have never participated in the metal coordination, though these groups become good ligands when they are deprotonated.

2.4 Factors Affecting Metal Binding Sites

In general, "hard" metal ions such as alkali and alkaline earth metal ions prefer "hard" bases such as phosphate and sugar hydroxyl groups, while transition metal ions prefer phosphate and nucleobase moieties.

Stereospecific interligand interactions, involving hydrogen bonding, electrostatic repulsions, or steric constraints, are important factors that affect nucleobase and site-specific metal bonding [68–70]. For example, in the tripodal tetradentate nta ligand (nta = nitrilotriacetate) system [69a], where nta bears three carboxylate groups that could function as hydrogen bonding acceptors only, the nta-capped octahedral metal ion binds exclusively to adenine at N7 over N9-substituted guanine, cytosine, or uracil; this could be due to the formation of the interligand N6–H···O(nta) hydrogen bonding for adenine.

The hydration degree is an additional factor: when the water content is small in the crystal, the phosphate group is apt to participate in metal coordination. For metal binding effects on nucleotide structures, see [1,8,9].

3 Metal-Oligonucleotide Complexes

All crystal structures of the metal compounds with di- and trinucleotides and those with DNA fragments reported before 1994 were tabulated in Ref. 1 (in Tables 2 and 3, respectively). At that time, X-ray data of oligonucleotides at atomic resolution level were available only for Z-DNA. Now thanks to synchrotron radiation of increasing power, together with better detectors and cryotechniques, high-resolution (better than 1.3 Å [22]) analysis of RNA/DNA structures is possible. However, it is important to keep in mind that the identification of sodium is still problematic because of mixed occupancy of Na^+ with water molecules (sodium ions and water molecules carry the same number of electrons, giving electron density peaks with similar volumes) [21,22,71]. To our knowledge, no crystal structure of metal–*di*- or *tri*nucleotide compounds are available after 1994 to date. Crystal structures of the metal compounds of RNA fragments reported after 1994, excluding those of ribozymes [25], are compiled in Table 5 and those of DNA fragments in Table 6 together with comments in some detail. The ternary drug–metal ion–oligonucleotide complexes are excluded. The Nucleic Acid Data Bank (NDB) [72] compiles all crystal structures of RNA/DNA oligonucleotides so far reported and provides images of any available structure. A data base of metal ion binding sites in RNA structures (MeRNA) [73] catalogs 16122 metal ions corresponding to 23 elements that interact with RNA in 389 PDB entries solved X-ray diffraction, NMR and fiber diffraction, with coordinates released before February 1, 2007 and containing one or more metal cations bound to RNA.

The numbering system of an RNA or DNA fragment employed here is, for example,

GGUGAGG·= 5' – G1-P_2 — G2 – P_3 – U3 – P_4 – G4 – P_5 – A5 – P_6 – G6 – P_7 – G7 — 3'
CCUCACC . . . .
3' – C14-P_{14} - C13-P_{13} -A12-P_{12} - C11-P_{11} -U10 - P_{10} - C9 - P_9 - C8 – 5'.

3.1 RNA Fragments

There is an increasing interest in interactions between metal ions and RNA, because RNA biopolymers can act, in addition to their function as a genetic carrier, as rigid structural scaffolds, conformational switches, and catalysts for chemical reactions and in these cases, metal ions play a crucial role in RNA function [3,74]. It should be noted that RNA often forms complex tertiary structures in which phosphates are often clustered. In those cases, the electrostatic potential would be very different and different metal binding modes would be expected. However, our chapter addresses only the regular double-helical region of RNA (A-RNA) structure.

A-RNA adopts a right-handed double-helical structure with a deep and narrow major groove and a shallow and wide minor groove. In the sodium salt of $[r(GUAUAUA)d(C)]_2$ [75], the sodium ion acts as a glue that connects three duplexes around a 3-fold axis in the crystal lattice by directly binding to O2'-hydroxyl groups of adenine11 (A11) residues from the three duplexes. The high-resolution (1.3 Å) crystal structure of the sodium salt of $[r(GCGUU^{Br}UGAAACGC)]_2$ [76] reveals the base-sequence specific metal bonding. Five Na^+ ions (Na1–Na5) and one Mn^{2+} ion are located in the major groove of the duplex and participate in direct (innersphere) or water-mediated (outersphere) contacts with groups lining the major groove. Na1 and Na2 ions form innersphere contacts with N7 of A9 and N7 of G7, respectively. Na3 and Na4 ions also engage in innersphere contacts with two O4 atoms of U4 and U5 and two O4 atoms of U17 and U18 and thus, the two first bases of UUU triplets constitute a specific site for sodium ions. The GAAA tetranucleotides inserted in a regular RNA duplex, where G forms a non-Watson–Crick G7(*anti*)–G20(*syn*) base pair in the duplex, present an unusual structural plasticity that corresponds to the direct bonding of an octahedral metal ion to N7 of purine bases and O4 of uracil bases. A striking asymmetric pattern of metal ion binding in the two equivalent halves of the palindromic sequences, that is, four (Na1–Na3 + Mn) on the one half while two (Na4 and Na5) on the other half, demonstrates that the sequence and its environment act together to bind metal ions. The highly ionophilic half consisting of six base pairs that binds five metal ions (Na1–Na5) is reminiscent of the loop E of the 5S ribosome RNA [89]. In this structure, a Mn^{2+} (or Mg^{2+}) ion binds to G12 in an outersphere fashion to form an N7···O(W)–Mn–O(W)···O6 linkage.

Five high-resolution crystal structures of magnesium compounds with $[r\{2'\text{-}O\text{-}Me(CGCGCG)\}]_2$ ($P6_122$ form [77] and $P3_212$ form [78]), $[r(CGCG^{F}CG)]_2$ [79], $[r(CGCGAAUUAGCG)]_2$ [80], and r(GGUGAGG)·r(CCUCACC) [81] are

reported. Mg^{2+} ions directly bind to the phosphate groups connecting neighboring duplexes [77] or act as an anchor attached to the 3'-terminal G [79]. Hydrated metal ions are located preferably in the major groove [77–81] compared to the minor groove [81], and when they are located in the major groove, the formation of outersphere N7···O(W)–M–O(W)···O6 bonding at the guanine residue is prominent [78–81].

In the four crystal structures of calcium compounds with $[r(CGCAIGCG)]_2$ [82], r(CUGGGCGG)·r(CCGCCUGG) [83], r(GCCACCCUG)·r(CAGGGUCGGC) [84], and $[r(CCC\mathit{P}GGGG)]_2$ [85], metal–phosphate bonding also occurs [84,85], but no outersphere N7···O(W)–M–O(W)···O6 bonding at G is observed (though a Ca^{2+} ion is located in the major groove [83]). Instead, interestingly, Ca^{2+} ion binding to the *cis*-diol of 3'-terminal sugar moieties of the duplex is frequent [82–85]; the intramolecular O2'–Ca^{2+}–O3' chelation is also observed in $[Ca(5'\text{-IMP})]\cdot 6.5H_2O$ [90].

The manganese complex with $[r(GGCCGAAAGGCC)]_2$ [86], where two G·A and two A·A non-Watson–Crick (W–C) base pairs form an internal loop, is unique in that an octahedral $[Mn(H_2O)_4]^{2+}$ ion directly binds to N7 of G9 in the G9·C4* W–C base pair (C4* is 2-fold symmetry-related to C4) following the internal loop (another Mn^{2+} hydrate is bound to the symmetrically related A8*). Here the formation of the A8·G5* non-W–C base pair in the loop region creates enough space for an octahedral metal ion to bind to the base moiety. In the rhodium (or iridium) complex with $[r(GCUUCGGC)d^{Br}U]_2$ [87], where r(UUCG) is embedded in self-complementary sequences and forms two G·U and two C·U non-W–C base pairs, four $[Rh(NH_3)_6]^{3+}$ (or $[Ir(NH_3)_6]^{3+}$) ions (Rh1–Rh4) per duplex play a role in maintaining the ordering of the molecule in the crystals. Rh1 and Rh2 are located in the major groove and participate in outersphere contacts to O6 atoms of two adjacent G6 and G7 and to O(P) atoms of the preceding C5 and G6.

Finally, a metal complex of a locked nucleic acid (LNA), $r(m^5Cm^5CTm^5CAm^5Cm^5C)^L\cdot r(GGTGAGG)^L$ containing magnesium and cobalt ions, is available [88]. These metal ions play a role in associating two duplexes in the crystal lattice: a $[Mg(H_2O)_4]^{2+}$ ion bridges directly between the two O2' ether oxygens of the sugars from each duplex, while a $[Co(NH_3)_6]^{3+}$ ion is located in the major groove and makes outersphere contacts *via* amine ligands to guanine bases G6 and G7 from each duplex.

In summary, as major binding patterns, the following three are observed for crystalline A-RNA. (i) Metal cations prefer to bind to phosphate groups, especially at tight places where phosphates are clustered in the crystal lattice, through inner- and/or outersphere interactions, to shield the negative phosphate charges. (ii) The *outer*sphere metal coordination of hydrated metal ions to the groups lining up in the *major* groove is frequently observed, especially to both O6 and N7 sites of guanine ($[Mg(H_2O)_6]^{2+}$ [78–81] and $[Mn(H_2O)_6]^{2+}$ [76]), forming a [–N7···O(W)–M–O(W)···O6–] macrochelate. This is consistent with the high negative electrostatic potential in the major groove, predicted for a (dG-dC) model of RNA [91]. (iii) The *inner*sphere coordination of *octahedral* metal ions ($[Na(H_2O)_5]^{2+}$ [76] and $[Mn(H_2O)_4]^{2+}$ [86]) to the *base* moieties in the *major* groove also occurs in a sequence-specific manner, where the direct metal–base bonding becomes possible by sterically inherent structures formed depending on non-complementary base sequences inserted in a regular RNA duplex. Finally, it should be pointed out that the O2' hydroxyl groups, which are RNA specific, are intimately involved in the hydrogen bonding network whereas their participation in innersphere metal coordination is limited [75].

Table 5 Crystal structures of the metal compounds of RNA fragments[a,b].

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Additional remarks							
Right-handed double stranded A-RNA							
Na^+	$[r(GUAUAUA)d(C)]_2$	1	83.66	0.66 Na^+	2.2	15.6	[75]
The duplex carries a 3'-terminal deoxycytidine residue. Two independent Na^+ ions lie on the 3-fold axis. Na1 ($[Na(H_2O)]^+$) binds to O2'-hydroxyl groups of adenine 11 (A11) and its symmetry-related equivalents, mediating the approach of three duplexes. Na2 is fully hydrated.							
Na^+, Mn^{2+} or Mg^{2+}	$[r(GCGUU^{Br}UGAAACGC)]_2$	1	182	5 Na^+, 1 Mn^{2+} (or Mg^{2+})	1.3	15.8	[76]
The duplex involves a G7·G20 Hoogsteen base pair. Three (Na1–Na3) of the five Na^+ ions and one Mn^{2+} ion are in one side of the major groove and the remaining Na4 and Na5 are in the other side. Na1 ($[Na(H_2O)_5]^+$) participates in a direct (innersphere) contact with N7 of A9. Na2 ($[Na(H_2O)_5]^+$) engages in an innersphere contact with N7 of G7. Na3 ($[Na(H_2O)_4]^+$) makes innersphere contacts with two keto O4 oxygens of U4 and U5. Na4 ($[Na(H_2O)_4]^+$) forms innersphere contacts with two keto O4 oxygens of U17 and U18. Na5 ($[Na(H_2O)_5]^+$) forms a disodium cluster with Na4 and participates in water-mediated (outersphere) contacts with O(P)s of G16 and U17. A $[Mn(H_2O)_6]^{2+}$ ion engages in outersphere contacts with N7 and O6 of G12.							
Mg^{2+}	$[r\{2'\text{-}O\text{-}Me(CGCGCG)\}]_2$ ($P6_122$ form)	2	44	2 Mg^{2+}	1.30	17.5	[77]
Mg1 ion lies on a 2-fold axis and bridges between two symmetry-related duplexes through two phosphate oxygens O(P)s of C5. Mg2 ion is coordinated to one O(P) of G2. Both Mg^{2+} ions are located near duplex ends and contribute to the "edge-effects" seen in the helical parameters, rise, and roll, for the terminal base steps.							
Mg^{2+}	$[r\{2'\text{-}O\text{-}Me(CGCGCG)\}]_2$ ($P3_212$ form)	1	65	5 Mg^{2+}	1.19	15.5	[78]
Mg1 ($[Mg(H_2O)_5]^{2+}$) forms an innersphere contact with O(P) of G6 and an outersphere contact with O(P) of C3 of the neighboring molecule. Mg2 ($[Mg(H_2O)_5]^{2+}$) forms an innersphere contact with O(P) of G8 and an outersphere contact with O(P) of C9. Mg3 ($[Mg(H_2O)_4]^{2+}$) bridges two duplexes through innersphere contacts with O(P)s of C5 and C11. The other two fully hydrated $[Mg(H_2O)_6]^{2+}$ ions are in the major groove and form outersphere contacts with N7 and O6 of G6 and N4 of C5 (Mg4) or N7 and O6 of G12 and N4 of C11 (Mg5). The crystals contain also per unit 2 molecules of MPD.							
Mg^{2+}	$[r(CGCG^{F}CG)]_2$	2	132	4 Mg^{2+}	1.45	17.2	[79]
Mg1 and Mg2 locate in the major groove as $[Mg(H_2O)_6]^{2+}$ without disrupting the hydration pattern of the RNA and engage in outersphere contacts with N7 and O6 of guanine residues. Mg3 ($[Mg(H_2O)_6]^{2+}$) is lodged between neighboring duplexes. Mg4 ($[Mg(H_2O)_5]^{2+}$) forms an innersphere contact with O(P) of the terminal G6.							

Mg^{2+}	$[r(CGCGAAUUAGCG)]_2$	0.5	76	2 Mg^{2+}	1.12	14.7	[80]
The duplex is located on a 2-fold axis. Four $[Mg(H_2O)_6]^{2+}$ ions engage in outersphere contacts to the major groove edges of both G and A that form non-W–C base pairs, G4·A9* and its equivalent G4*·A9, forming N7···O(W)-Mg-O(W)···O6(G) or N6(A) macrochelates.							
Mg^{2+}	r(GGUGAGG)·r(CCUCACC)	1	97	2 Mg^{2+}	1.20	19.0	[81]
The sequence corresponds to the *E. coli* $tRNA^{ser}$ acceptor stem microhelix. Mg1 ($[Mg(H_2O)_5]^{2+}$) is located in the major groove and engages in outersphere contacts with N7 and O6 of the subsequent guanine residues G6 and G7. Mg2 ($[Mg(H_2O)_6]^{2+}$) is located in the minor groove and forms outersphere contacts with O2' of G6 and the phosphate backbone of two further symmetry-related RNA molecules.							
Ca^{2+}	$[r(CGCAIGCG)]_2$	0.5	6	1 Ca^{2+}	2.5	24.6	[82]
The duplex is located on a 2-fold axis. A $[Ca(H_2O)_4]^{2+}$ ion forms innersphere contacts with each strand of the duplex through O2' and O3' of the terminal G8 (or its symmetry equivalent G16) and an outersphere contact with O(P) of G6 (or G14) of a neighboring duplex.							
Ca^{2+}	r(CUGGGCGG)·r(CCGCCUGG)	2	~125	6 Ca^{2+}	1.30	23.9	[83]
The sequence of the duplex corresponds to the E region of *Thermus flavus* 5S rRNA. The first crystal structure involving both the D- and L-duplexes (RNA racemate). Ca1 and Ca2 are associated with the outer side of the sugar-phosphate backbone. Ca3 locates in the major groove. Ca4, Ca5, and Ca6 engage in innersphere contacts with O2' and O3' of the 3'-terminal G8 and G16 of the ordered duplex and G8/G16 of the disordered duplex, respectively. The crystals contain also glycerol.							
Ca^{2+}	r(GCCACCCUG)·r(CAGGGUCGGC) ($P4_22_12$ form)	1	23	2 Ca^{2+}	2.2	20.4	[84]
The sequence of the duplex corresponds to the helix II region of *Xenopus laevis* 5S rRNA. Ca1 and Ca2 participate in intermolecular interactions in connecting symmetry-related duplexes. Ca1 sits on a 2-fold axis and is directly associated with four symmetry-related O(P)s of the 3'-terminal G9 from four duplexes. Ca2 also engages in innersphere contacts with O2' and O3' of the 3'-terminal C19 and O(P) of U8.							
Ca^{2+}	$[r(CCCPGGGG)]_2$	1	46	2 Ca^{2+}	1.97	21.7	[85]
RNA duplex containing phenyl (P) ribonucleotides. Ca1 makes innersphere contacts with O2' and O3' of the unpaired 3'-terminal G8 and to O(P) of the residue P4 from a neighboring duplex. Ca2 sits on a 2-fold axis and forms innersphere contacts with two symmetry-related O(P)s of G8.							
Mn^{2+}	$[r(GGCCGAAAGGCC)]_2$	0.5	20	1 Mn^{2+}	2.3	18.6	[86]
Each of two strands of the duplex is related by a 2-fold symmetry. Two G·A and two A·A non-Watson–Crick base pairs form an internal loop. Mn^{2+} ions bind to each strand of the duplex by bridging between N7 of G9 (or G9*) in the W–C region and O(P) of the preceding A8 (or A8*) in the loop. This Mn^{2+} location corresponds to a metal binding site in the hammerhead catalytic RNA.							

(continued)

Table 5 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Rh^{3+}	$[r(GCUUCGGC)d^{Br}U]_2$	1	55	4 $[Rh(NH_3)_6]^{3+}$	1.6	18.0	[87]
The duplex carries a 3'-terminal 5-bromo-2'-deoxyuridine. Among the four $[Rh(NH_3)_6]^{3+}$ ions (Rh1–Rh4), Rh1 is in the major groove and participates in outersphere contacts *via* amine ligands to O(P) atoms of C5 and G6 and to O6 atoms of G6 and G7. Rh2 is in the corresponding positions of strand 2, connected to G16 and G17. Rh3 is in a channel between helices with numerous intermolecular contacts. Rh4 is in a cavity close to $^{Br}U9$ and participates in interactions between the parallel stacks of helices.							
Right-handed double-stranded locked nucleic acid							
Mg^{2+}, Co^{3+}	$r(m^5Cm^5CTm^5CAm^5Cm^5C)^L{\cdot}r(GGTGAGG)^L$	2	100	1 $[Mg(H_2O)_4]^{2+}$, 1 $[Co(NH_3)_6]^{3+}$	1.90	22.4	[88]
A $[Mg(H_2O)_4]^{2+}$ ion is bound to the bridged O2' of two symmetry equivalent 2'-*O*-4'*C*-methylene-β-D-ribofranose sugars. A $[Co(NH_3)_6]^{3+}$ ion acts as an 'anchor' between the two independent duplexes in the asymmetric unit; it participates in outersphere contacts to two adjacent guanine bases G6 and G7 of each duplex. The crystals also contain per unit 1 molecule of cacodylate.							

[a] Abbreviations: $d^{Br}U$ = 5-bromo-2'-deoxyuridine; O(W) = water oxygen; O(P) = phosphate oxygen; Me = methyl; MPD = 2-methyl-2,4-pentenediol; ^FC = 5-fluoro-cytidine; m^5C = 5-methyl-cytidine.

[b] Numbering of, for example, (GGUGAGG)· = 5'-G1—G2—U3—G4—A5—G6—G7-3'
(CCUCACC) · · · · · · ·
3'-C14–C13–A12–C11–U10–C9—C8-5'.

3.2 DNA Fragments

3.2.1 A-DNA Fragments

As with A-RNA, A-DNA adopts a right-handed double-helical structure with a deep and narrow major groove and a shallow and wide minor groove. A review on crystal structures of A-DNA duplexes is available [107]. Reliable detection of light metal ions in X-ray analysis of 3D structures remains a challenge, in particular of sodium ions [93]. Thus, replacement of Na^+ or K^+ by heavier alkali metal ions in the crystallization of a nucleic acid fragment has been attempted for locating sites occupied by Na^+ or K^+. High-resolution quality of crystals and precise measurements of anomalous differences in intensities allowed to locate metal ion binding sites in the salts of $[d(GCGTATACGC)]_2$ with alkali (Na^+, K^+, Rb^+, and Cs^+ [93]) in addition to alkaline earth (Mg^{2+} [93,94,96] and Ba^{2+} [93]) metal ions. There are three binding sites: in two sites metal ions play a role as glues connecting three neighbor duplexes through direct contacts to phosphate-backbone atoms. The third site is located in the major groove and comprises three adjacent base pairs at the G3pT4pA5 step, where metal ions participate in outersphere contacts with N7 and O6 of G3 and O4 of T4 of the same duplex. The former two sites are exclusively occupied by alkali ions, whereas the latter site can accommodate both alkali and alkaline earth metal ions. Crystal structures of Mg^{2+} salts of $[d(ACCGGCCGGT)]_2$, $[r(GCG)d(TATACGC)]_2$, $[r(GC)d(GTATACGC)]_2$, and $[r(G)d(CGTATACGC)]_2$ reveal two characteristic $[Mg(H_2O)_6]^{2+}$ binding modes [97]. One mode has the ion binding in the deep major groove of a GpN step through outersphere contacts with N7 and O6 of guanine bases. The other mode involves binding of the metal ion to phosphates, bridging across the outer mouth of the narrow major groove. Outersphere type interaction of Mg^{2+} ions with guanine bases in the major groove through N7 and/or O6 sites is commonly observed also in other A-DNA duplexes [99,100,101]. On the other hand, *innersphere* interaction with an *inner* base in the *major* groove is observed as still quite rare cases for Ba^{2+} ion in $[d(ACCCGCGGGT)]_2$ [102] and an oxaliplatin complex of d(CCTCTGGTCTCC)·d(GGAGACCAGAGG) [103], where a Ba^{2+} ion binds directly to N7 and O6 of G8 and O6 of adjacent G7 in the former [102] and N7 and O6 of G23 and N7 of adjacent A22 in the latter [103]. It should be noted here that the Ba^{2+} ion is extremely dehydrated in these cases: no water ligand is located in the former and only one water ligand in the latter.

$[Co(NH_3)_6]^{3+}$ ion behaves as $[Mg(H_2O)_6]^{2+}$ ion. Thus, in the crystal structure of $[d(ACCGGCCGGT)]_2$ [102], two of the three independent Co^{3+} ions occupy the same sites as those of Mg^{2+} ions in $[d(ACCGGCCGGT)]_2$ [97]: they bind to two consecutive Gs in one strand, hydrogen bonding to O6 and N7 positions of the bases; another Co^{3+} ion makes outersphere contacts to the duplex via phosphate groups and organizes the crystal packing. On the other hand, the two independent $[Co(NH_3)_6]^{3+}$ ions observed in chimera $[r(GC)d(GTATACGC)]_2$ display a phosphate-only binding mode [102]. Outersphere type interaction of $[Co(NH_3)_6]^{3+}$ ions with guanine bases in the major groove through N7 and/or O6 sites and with phosphate

Table 6 Crystal structures of the metal compounds of DNA fragments[a–c].

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	R (%)	Ref.
Additional remarks							
Right-handed double-stranded A-DNA (involving duplexes with A- and B-segments)							
Na^+	$[d(AGGGGCCCCT)]_2$ ($P2_12_12_1$ form)	1	135	1 Na^+	1.1	19.6	[92]
One possible hydrated Na^+ ion is located in the major groove near the G14 base by displacing one of the water molecules constructing the spine of hydration.							
Na^+	$[d(GCGTATACGC)]_2$ (Na–TOE form)	1	–[d]	2 Na^+	1.30	16.1	[93]
T6-residue is modified at the 2'-position by TOE. Two Na^+ ions occupy the same sites as Site2 and Site3 in the Cs–MEP form [93].							
Na^+, Mg^{2+}	$[d(GCGTATACGC)]_2$ (Na–FET form)	1	–	1 Na^+, 1 Mg^{2+}	1.45	15.9	[93]
T6-residue is modified at the 2'-position by FET. A Na^+ ion occupies the same site as Site3 while a partially ordered Mg^{2+} ion occupies Site2 in the Cs–MEP form [93].							
K^+	$[d(GCGTATACGC)]_2$ (K–MEP form)	1	–	3 K^+	–	20.2	[93]
T6-residue is modified at the 2'- and 3'-positions by MEP. Three K^+ ions occupy the same sites as Site1–Site3 in the Cs–MEP form [93].							
K^+	$[d(GCGTATACGC)]_2$ (K–TOE form)	1	–	3 K^+	1.30	16.1	[93]
T6-residue is modified at the 2'-position by TOE. Three K^+ ions occupy the same sites as Site1–Site3 in the Cs–MEP form [93].							
Rb^+	$[d(GCGTATACGC)]_2$ (Rb–FET form)	1	–	3 Rb^+	1.05	15.4	[93]
T6-residue is modified at the 2'-position by FET. Three Rb^+ ions occupy the same sites as Site1–Site3 in the Cs–MEP form [93].							
Rb^+	$[d(GCGTATACGC)]_2$ (Rb–TOE form)	1	–	3 Rb^+	1.30	16.5	[93]
T6-residue is modified at the 2'-position by TOE. Three Rb^+ ions occupy the same sites as Site1–Site3 in the Cs–MEP form [93].							
Cs^+	$[d(GCGTATACGC)]_2$ (Cs–MEP form)	1	–	3 Cs^+	1.06	15.0	[93]
T6-residue is modified at the 2'- and 3'-positions by MEP. Cs^+ ions occupy three binding sites (1–3). At Site1, a $[Cs(H_2O)_4]^+$ ion makes innersphere contacts with O3' of C9 and O(P) of C10 in one duplex, O3' of T14 andO(P) of A15 in the second duplex, and two O(P)s of C20 in the third duplex. At Site2, a $[Cs(H_2O)_6]^+$ ion is located in the major groove and makes outersphere contacts with N7 and O6 of G3 and O4 of T4 of the same duplex. At Site3, a $[Cs(H_2O)_3]^+$ ion makes innersphere contacts with two O(P)s of C10 in one duplex, O4' of T14 in the second duplex, and O3' of C20 in the third duplex.							
Cs^+	$[d(GCGTATACGC)]_2$ (Cs-TOE form)	1	–	3 Cs^+	1.05	16.1	[93]
T6-residue is modified at the 2'-position by TOE. Three Cs^+ ions occupy the same sites as Site1–Site3 in the Cs–MEP form [93].							

Mg^{2+}	$[d(GCGTATACGC)]_2$ (Mg-TOE form)	1	105	2 Mg^{2+}	1.65	18.9	[94]
T6-residue is modified at the 2'-position by TOE. One Mg^{2+} ion occupies the same site as Site2 in the Cs–MEP form [93]. The other Mg^{2+} ion is located at crown ether-like pockets provided by TOE substituents of adjacent modified T6 residues.							
Mg^{2+}	$[d(GCGTATACGC)]_2$ (Mg-BTL form)	1	31	1 Mg^{2+}	2.30	23.7	[95]
T6-residue is modified at the 2'-position by BTL. No metal ion interaction mentioned.							
Mg^{2+}	$[d(GCGTATACGC)]_2$ (Mg-TFE form)	1	115	1 Mg^{2+}	1.50	17.1	[95]
T6-residue is modified at the 2'-position by TFE. No metal ion interaction mentioned. The crystals also contain spermine.							
Mg^{2+}	$[d(GCGTATACGC)]_2$ (Mg-IME form)	1	114	3 Mg^{2+}	1.60	16.5	[95]
T6-residue is modified at the 2'-position by IME. No metal ion interaction mentioned. The crystals also contain spermine.							
Mg^{2+}	$[d(GCGTA^{M}TACGC)]_2$	1	177	2 $[Mg(H_2O)_6]^{2+}$	0.83	11.7	[96]
^{M}T-residue is modified at the 2'- and 3'-positions by 2'MeO-3'MP. One Mg^{2+} ion is in the major groove and participates in outersphere contacts with the G3-T4 step through O6 of G3 and O4 of T4 and with O4 of T16 of the pairing strand, stabilizing the kink between the base pairs T4·A17 and A5·T16. This binding site is the same as that of a Mg^{2+} ion in the Na–FET form [93]. The other Mg^{2+} ion is attached to the phosphate backbone with outersphere contacts with $O(P_5)$ and $O(P_6)$. The crystals also contain per unit 2 molecules of spermine.							
Mg^{2+}	$[d(ACCGGCCGGT)]_2$	0.5	24	3 $[Mg(H_2O)_6]^{2+}$	1.6	22.9	[97]
The duplex rides on a 2-fold symmetry axis. Four Mg^{2+} ions (Mg1 and Mg2 and symmetry-related Mg1* and Mg2*, where Mg1 and Mg1* are half-weighted) are located in the deep major groove and participate in water-mediated (outersphere) contacts with O6 and N7 sites of guanines: Mg1(or Mg1*) ion contacts with the G4–G5 (or G4*–G5*) step and Mg2 (Mg2*) with the G8–G9 (G8*–G9*) step.							
Mg^{2+}	$[r(GCG)d(TATACGC)]_2$	1	97	4 $[Mg(H_2O)_6]^{2+}$	1.4	19.3	[97]
Three Mg^{2+} ions (Mg1–Mg3) are in the major groove and participate in outersphere contacts with O6 and N7 sites of guanines at GpT steps: Mg1 ion contacts with the G19–C18 step, Mg2 with the G3–T4 step and Mg3 with the G13–T14 step. Mg4 ion is at the exterior of the mouth of the narrow major groove and participates in outersphere contacts with O(P)s.							
Mg^{2+}	$[r(GC)d(GTATACGC)]_2$	1	45	1 $[Mg(H_2O)_6]^{2+}$	1.7	19.6	[97]
A Mg^{2+} ion occupies the same site as that of Mg2 in the $[r(GCG)d(TATACGC)]_2$ complex [97].							
Mg^{2+}	$[r(G)d(CGTATACGC)]_2$	1	78	1 $[Mg(H_2O)_6]^{2+}$	1.7	19.4	[97]
A Mg^{2+} ion is in the major groove and in outersphere contacts with O6 and N7 of guanine at the C2–G3 step, near the site of Mg2 in the $[r(GCG)d(TATACGC)]_2$ complex [97].							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit			Resolution (Å)	*R* (%)	Ref.
		DNA duplex	H_2O	Metal			
Mg^{2+}	$[d(pCCCGCGGG)]_2$	1	73	1 $[Mg(H_2O)_6]^{2+}$	1.6	24.55	[98]
No metal ion interaction mentioned.							
Mg^{2+}	$[d(CCGCCGGCGG)]_2$	0.5	32	1 Mg^{2+}	2.2	19.7	[99]
The duplex rides on a 2-fold axis. Two symmetry-related $[Mg(H_2O)_6]^{2+}$ ions bind indirectly to both strands of the duplex, at the center of the major groove, via water molecules by forming hydrogen bonds to N7 and O6 atoms of G6, G7, and symmetry-related G3* on one side, and G6*, G7*, and G3 on the other side of the 2-fold axis.							
Mg^{2+}	$[d(Cm^5CGCm^5CGGm^5CGG)]_2$ ($P2_12_12_1$ form)	1	79	1 Mg^{2+}	1.7	19.6	[99]
$[Mg(H_2O)_6]^{2+}$ ion is indirectly bound via water molecules to one strand of the duplex through N7 and O6 atoms of G19 and G20 at the duplex end. Methylation causes the displacement of the Mg^{2+}-binding site, from the center of the major groove to the duplex end.							
Mg^{2+}	$[d(Cm^5CGCm^5CGGm^5CGG)]_2$ ($P6_1$ form)	1	80	1 Mg^{2+}	2.15	16.6	[99]
$[Mg(H_2O)_6]^{2+}$ ion is indirectly bound via water molecules to one strand of the duplex through N7 and O6 atoms of G19 and G20 at the duplex end.							
Mg^{2+}	d(GTTTTG(6)-Sd2C(9)CA(11)AAAC)	2	170	5 $[Mg(H_2O)_6]^{2+}$	1.50	20.4	[100]
The asymmetric unit involves two independent hairpins, each of which forms a duplex portion with a stilbenediether (Sd2) linker. One of the five Mg^{2+} ions is the only ion with contacts to a single DNA molecule. It is bound inside of the minor groove and clamps two adjacent base pairs of the A-tract, through H-bonds with N3 of A11, O4' of A12, O2 of T5, and O4' of G6. The other ions mediate interactions between duplexes.							
Mg^{2+}, Pt^{2+}	d(CCTCTGGTCTCC)·d(GGAGACCAGAGG)	2	148	4 Mg^{2+}, 2 *cis*-$[Pt(NH_3)_2]^{2+}$	1.77	17.2	[101]
Two independent duplexes in the asymmetric unit. In each duplex, an A-DNA-like segment at the 5' side of the Pt cross-link and a B-DNA-like segment at the 3'-side. In each duplex, two $[Mg(H_2O)_6]^{2+}$ ions are associated with adjacent purine bases in the major groove on the A-form side, either the terminal G23pG24 or the A20pG21 sequence at the 3'-end of the unplatinated strand, via water ligands as hydrogen bond donors to N7 atoms of guanine and adenine bases and to O6 of guanine. A *cis*-$[Pt(NH_3)_2]^{2+}$ ion *intrastrand* cross-links to the G6pG7 bases through N7 atoms. The duplex is bent by 35–45° at the platinated site toward the major groove without disrupting G·C base pairs.							
Sr^{2+}	$[d(GCGTATACGC)]_2$ (Sr-BOE form)	1	101	1 Sr^{2+}	1.80	17.0	[95]
T6-residue is modified at the 2'-position by BOE. No metal ion interaction mentioned.							

Ba^{2+}	$[d(GCGTATACGC)]_2$ (Ba-AOE form)	1	-	1 Ba^{2+}	1.70	19.0	[93]
T6-residue is modified at the 2'-positions by AOE. A Ba^{2+} ion occupies the same site as Site2 in the Cs–MEP form [93].							
Ba^{2+}	$[d(GCGTATACGC)]_2$ (Ba-MAOE form)	1	115	1 Mg^{2+}	1.50	17.1	[95]
T6-residue is modified at the 2'-position by MAOE. No metal ion interaction mentioned. The crystals also contain spermine.							
Ba^{2+}	$[d(ACCCGCGGGT)]_2$ ($P6_122$ form)	0.5	48	1 Ba^{2+}	2.00	20.6	[102]
The duplex rides on a 2-fold axis. Each of the two Ba^{2+} ions is in the major groove and directly binds to each strand through N7 and O6 of G8 (symmetry-related G8*) and to O6 of the adjacent G7 (G7*) and indirectly via a H-bond to O6 of G9 (G9*).							
Ba^{2+}, Pt^{2+}	d(CCTCTGGTCTCC)·d(GGAGACCAGAGG)	1	12	1 Ba^{2+}, 1 [Pt(*R*,*R*-DACH)(oxalato)]$^{2+}$	2.4	20.9	[103]
The first seven base-pair steps adopt an A-DNA-like conformation and the last two steps take a B-DNA-like one. A Ba^{2+} ion is in the major groove and directly binds to N7 and O6 of G23 and forms a bridge to N7 of the adjacent A22. A Pt^{2+} complex (oxaliplatin) *intrastrand* cross-links two adjacent guanine bases G6 and G7 through their N7 sites. The duplex is bent by ~30° at the platinated site toward the major groove without disrupting G·C base pairs.							
Co^{3+}	$[d(AGGGGCCCCT)]_2$ ($P6_122$ form)	0.5	78	2 $[Co(NH_3)_6]^{3+}$	1.5	23.6	[92]
The duplex rides on a 2-fold axis. The central eight base pairs adopt A-DNA conformation with the two terminal nucleotides A and T flipped out to form base pairs with the neighboring nucleotides. The locations of $[Co(NH_3)_6]^{3+}$ ions are the same as those of $[d(AGGCATGCCT)]_2$ [104].							
Co^{3+}	$[d(GCGTATACGC)]_2$ (Co-MAOE form)	1	124	2 Co^{3+}	1.65	16.3	[95]
T6-residue is modified at the 2'-position by MAOE. No metal ion interaction mentioned. The crystals also contain spermine.							
Co^{3+}	$[d(ACCGGCCGGT)]_2$ ($P6_122$ form)	0.5	65	3 $[Co(NH_3)_6]^{3+}$	1.90	19.4	[102]
The duplex rides on a 2-fold axis. Co1 contacts indirectly via NH_3 ligands with three phosphate groups, two being those of C3 and C12 and one of G9 from the neighboring duplex. Co2 and Co3 are located at the center of the deep major groove near the G4-G5 bases and G8-G9 bases of a duplex, respectively, and form outersphere contacts with N7 and O6 atoms of guanines.							
Co^{3+}	$[r(GC)d(GTATACGC)]_2$	1	72	2 $[Co(NH_3)_6]^{3+}$	1.90	18.2	[102]
Co1 and Co2 ions locate in the interstitial space created by symmetry-related duplexes and indirectly contact with phosphate groups through NH_3 ligands by forming hydrogen bonds.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Co^{3+}	$[d(AGGCATGCCT)]_2$	0.5	76	2 $[Co(NH_3)_6]^{3+}$	1.3	13.9	[104]
The duplex rides on a 2-fold axis. Co1 locates in the major groove and indirectly binds via hydrogen bonds to the 5'-AGG-3' bases at N7 of G2 and G3, O6 of G3 and O(P) of G2. Co2 locates between duplexes and binds via hydrogen bonds to O(P) oxygens of G3, C8, C9, and T10 of four symmetry-related strands.							
Pt^{2+}	d(CCTCTGGTCTCC)·d(GGAGACCAGAGG)	2	31	2 *cis*-$[Pt(NH_3)_2]^{2+}$	2.6	20.3	[105]
In each of two independent duplexes, the duplex consists of the A-type DNA segment at the 5'-side and the B-type segment at the 3'-side with the A–B junction occurring at the base pair T8·A17, the site being adjacent to the platination. In each duplex, a Pt^{2+} ion cross-links two adjacent guanine bases G6 and G7 through their N7 sites. The duplex is bent by 39° for one duplex and 58° for the other at the platinated site toward the major groove without disrupting G·C base pairs.							
Pt^{2+}	d(CCTCTGGTCTCC)·d(GGAGACCAGAGG)	2	48	2 *cis*-$[Pt(NH_3)(NH_2Cy)]^{2+}$	2.4	20.08	[106]
Isomorphous with the cisplatin–DNA structure [101,105]. A Pt^{2+} ion *intrastrand*-cross-links two guanine bases G6 and G7 through their N7 sites, with the NH_2Cy ligand directed toward the 3'-end of the platinated strand. The duplex is bent at the platinated site by ca. 38° toward the major groove without disrupting G·C base pairs.							
Right-handed double-stranded B-DNA							
Na^+	$[d(CCGGTACCGG)]_4$	1	92	0.5 Na^+	2.10	23.0	[108]
Two duplexes ride on a 2-fold axis so as to generate the four-stranded antiparallel stacked-X Holliday junction. A Na^+ ion, which rides on a 2-fold axis, is within the junction and coordinated to two O(P)s of 2-fold symmetry-related two A6 residues.							
Na^+, Mg^{2+}	$[d(CGCGAATTCGCG)]_2$ (Na^+ form; structure X)	1	151	$-^d Na^+$, 1 Mg^{2+}	1.40	19.7	[109]
The DDD duplex. Partially occupied Na^+ ions locate within the minor groove by displacing water molecules constructing the spine of hydration. One Na^+ ion as $[Na(H_2O)_2]^+$ is located at the 5'ApT3' step (A-tract) and directly binds two O2 keto oxygens of T7 and T19 and two O4' sugar oxygens of T8 and T20. A $[Mg(H_2O)_6]^{2+}$ ion locates within the major groove of the G-tract and participates in outersphere contacts with O6 and N7 of G2 and O6 of G22 and with O(P)s of A6 and T7 of a symmetry-related duplex. The crystal also contains per unit 1> spermine.							

Na^+, Mg^{2+}	$[d(GCGAATTCG)]_2$	1	144	2 Na^+, 5 Mg^{2+}, 1 Cl^-	0.89	14.5	[110]
Two Na^+ and two Mg^{2+} ions, and a Cl^- ion are located in the major groove. The Cl^- ion is between two Mg^{2+} ions and interacts with N4 of C2 and with water ligands of Mg^{2+} ions. Mg^{2+} ions are mainly involved in the interaction among different duplexes but none of them are in the minor groove. Two Mg^{2+} ions are at both ends of the minor groove and interact with the ends of the water spine.							
Na^+, Mg^{2+}, Co^{3+}	$[d(GCGAGAGC)]_2$ (8hmt-G_5 Form)	0.5	70	0.33 Na^+, 1 Mg^{2+}, 0.33 Cl^- 0.33 $[Co(NH_3)_6]^{3+}$	2.0	22.0	[111]
The base-intercalated (zipper-like) duplex is formed. The structure is similar to those in $[d(GCGAXAGC)]_2$ (X = A: 8hmt-A_5 form [143]; X = T: 8hmt-T_5 form [111]). A Na^+ and a Co^{3+} ion ride on a 3-fold axis: the Na^+ ion binds to six G5 residues through O6 from three neighboring duplexes related by a 3-fold symmetry; a Co^{3+} ion binds through H-bonds to O6 and N7 of three G3 residues from symmetry-related three duplexes. Mg^{2+} ion not mentioned.							
Na^+, Ca^{2+}	$[d(G^{Br}CGGGAGC)]_2$	1	2	5 Na^+, 2 Ca^{2+}	3.6	29.4	[112]
Two DNA octamers form a base-intercalated duplex. Three duplexes assemble by a crystallographic 3-fold axis to form a hexaplex. Five metal ions locate on a 3-fold axis: one Na^+ ion binds to six G5 residues through O6 from symmetry-related three duplexes; a mixture of Ca^{2+} or Na^+ ions binds to three G4 residues and three G3 residues each through O6.							
Na^+, Ca^{2+}	$[d(CCGGTACm^5CGG)]_4$	1	131	1 Na^+, 1 Ca^{2+}	1.5	21.8	[113]
Two duplexes ride on a 2-fold symmetry axis so as to generate the four-stranded antiparallel stacked-X Holliday junction. A Na^+ ion is moved from the center of the junction in the parent crystal [108] but still coordinates O(P) of A6. A $[Ca(H_2O)_6]^{2+}$ ion is in the minor groove opposite the junction, between C2·G9 and G3·C8 base pairs of the 4-base pair arm.							
Na^+, Co^{3+}	$[d(GCGATAGC)]_2$ (8hmt-T_5 form)	0.5	91	1 Na^+, 0.5 $[Co(NH_3)_6]^{3+}$, 0.33 Cl^-	1.7	22.4	[111]
The base-intercalated (zipper-like) duplex is formed. The structure is similar to those in $[d(GCGAXAGC)]_2$ (X = A: 8hmt-A_5 form [143]; X = G: 8hmt-G_5 form [111]). Na^+ ion not mentioned. Two $[Co(NH_3)_6]^{3+}$ ions, one of which rides on a 3-fold axis and the other rides on the cross of 3-fold and 2-fold axes, are bound through H-bonds to O6 and N7 of three symmetry-related G3 residues for the former Co^{3+} ion and six symmetry-related T5 residues for the latter.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit			Resolution (Å)	*R* (%)	Ref.
		DNA duplex	H_2O	Metal			
Na^+, Pt^{2+}	$[d(CGCGAATTCGCG)]_2$	1	120	1 Na^+, 2 TriplatinNC	1.2	18.18	[114]
The DDD duplex. A $[Na(H_2O)_5]^+$ ion is in the major groove and engaged in outersphere contacts with O6 and N7 of G2 and N7 of G22 and N4 of C21. This position is the same as that of a $[Mg(H_2O)_6]^{2+}$ ion in many DDD crystal structures [109,118,144]. In addition, this Na^+ ion participates in an innersphere contact with O6 of G22. One trinuclear Pt^{2+} compound TriplatinNC (the occupancy is 0.75) forms an outersphere coordination complex through amine ligands with O(P)s of the duplex backbone as phosphate clamps. The other TriplatinNC (the occupancy factor is 0.60) is composed of only one Pt^{2+} unit with a fragment of a hexanediamine chain.								
Na^+, Pt^{2+}	$[d(CGCGAATTCGCG)]_2$	1	83	1 $[Na(H_2O)_4]^+$, 1 TriplatinNC-A	1.70	19.64	[115]
The DDD duplex. A Na^+ ion is in the major groove and contacts outersphere to O6 and N7 of G14 and O6 of G10. The binding mode of a TriplatinNC-A is very similar to that of a TriplatinNC in the complex with DDD [114].							
K^+	$[d(CGCGAAT^NUCGCG)]_2$ (DD1b form)	1	236	1 K^+, 1 Mg^{2+}	1.5	19.3	[116]
The DDD-like sequence with the 8-position replaced by ^{N}U. No metal ion interaction (DD1b form) mentioned.							
K^+, Mg^{2+}	$[d(CGCGAAT^NUCGCG)]_2$ (DD1a form)	1	210	1 K^+, 1 Mg^{2+}	1.6	17.9	[116]
The DDD-like sequence with the 8-position replaced by ^{N}U. No metal ion interaction mentioned.							
K^+, Mg^{2+}	$[d(CGCGAATTCGCG)]_2$ (K^+ form)	1	132	5 K^+, 1 Mg^{2+}	1.50	21.8	[117]
The DDD duplex. Partially occupied by five K^+ ions located within the minor groove, three in the primary layer of the spine of hydration and two in the second layer; K1 bridges between O2 of C9 and N3 of A17. K2 bridges between O2 of T8 and N3 of A18. K3 bridges between O2 of T7 and O2 of T19. A $[Mg(H_2O)_6]^{2+}$ ion is located at the same site within the major groove as that in the sodium form [109a]. The crystals also contain per unit >1 spermine.							
K^+, Mg^{2+}	$[d(CGCGAATTCGCG)]_2$ (K^+/spermine form)	1	132	5 K^+, 1 Mg^{2+}	1.2	13.9	[118]
The DDD duplex. A possible $[K(H_2O)_2]^+$ ion is located in the minor groove of the central 5'-AT-3' step (A-tract) and directly coordinated to O2 atoms of T7 and T19 and to O4' atoms of T8 and T20. A $[Mg(H_2O)_6]^{2+}$ ion is located on the floor of the major groove of one G-tract and engaged in outersphere contacts with O6 and N7 of G2 and O6 of G22 (a partial spermine molecule is on the floor of the major groove of the other G-tract). This Mg^{2+} ion is located within the lip of the minor groove of an adjacent duplex and engaged in outersphere contacts with O(P)s of A6 and T7. The crystals also contain per unit 1> spermine.							

K^+, Mg^{2+}	$[d(CGCGAAT^{f5}UCGCG)]_2$ (A form)	1	150	1 K^+, 1 Mg^{2+}	1.57	21.0	[119]
The DDD-type duplex involving modified ^{f5}U8 and ^{f5}U20 residues. A K^+ ion is located in the minor groove and bound to O2 atoms of T7 and T19. A $[Mg(H_2O)_6]^{2+}$ ion occupies the same position as that in the Dmo^{6}A·C form [136].							
Rb^+, Mg^{2+}	$[d(CGCGAA^{F}TTCGCG)]_2$ (type 1 Rb^+ form)	1	189	1 Rb^+, 3 Mg^{2+}	1.2	17.8	[120]
The DDD duplex involving modified ^{F}T7 and ^{F}T19. A $[Rb(H_2O)_3]^+$ ion (occupancy ca. 0.5) is in the minor groove at the A-tract and directly coordinated to O2 atoms of T7 and T19 and O4' atoms of T8 and T20. Two Mg^{2+} ions bridge phosphates across the minor groove to cause a narrowing of the groove.							
Rb^+, Mg^{2+}	$[d(CGCGAA^{F}TTCGCG)]_2$ (type 2 Rb^+ form)	1	164	1 Rb^+, 2 Mg^{2+}	1.5	17.8	[120]
The DDD duplex involving modified ^{F}T7 and ^{F}T19. Crystals, which were grown at lower Mg^{2+} concentrations than the type 1 Rb^+ form, are isomorphous with the type 1 form [120].							
Cs^+, Mg^{2+}	$[d(CGCGAATTCGCG)]_2$	1	141	–*d* Cs^+, 1 Mg^{2+}	1.8	20.93	[121]
The DDD duplex. Cs^+ ions are in the floor of the A-tract minor groove (occupancy estimated to be 10–40% at each site). A $[Mg(H_2O)_6]^{2+}$ ion is in the major groove and engaged in outersphere contacts with N7 and O6 of G2 and G22 in one duplex and with O(P)s of A6 and T7 of an adjacent one.							
Tl^+	$[d(CGCGAAXXCGCG)]_2$	1	114	8 Tl^+	1.49	19.80	[122]
The DDD duplex (DDD^{4+} form). X is a thymine residue linked at the 5 position to a *n*-propyl-amine. One of eight partially occupied Tl^+ ions (the occupancies of 0.15–0.22) is located in the minor groove at the A-tract and conserved in a series of DDD structures (K^+, Cs^+, and Tl^+). Two other Tl^+ ions are in the major groove and four are in the minor groove. The crystals also contain per unit 15 spermines.							
Tl^+, Mg^{2+}	$[d(CGCGAATTCGCG)]_2$ (Tl^+ form)	1	134	13 Tl^+, 1 Mg^{2+}	1.14	16.9	[123]
The DDD duplex. Thirteen Tl^+ ions are partially occupied (occupancies of 0.10–0.34). Seven Tl^+ ions are in the G-tract major groove and contact with O6 and N7 sites of guanine residues (G2, G4, and G10). Four are in the A-tract or A-tract/G-tract junction minor groove. Two are involved in cation–π interaction. A partially occupied $[Mg(H_2O)_6]^{2+}$ ion is in the G-tract major groove and contacts with G2 and G22 by competing with Tl^+ ions.							
Mg^{2+}	$[d(CGCGAA^{F}TTCGCG)]_2$	1	250	4 Mg^{2+}	0.95	16.0	[96]
The DDD duplex involving modified ^{F}T7 and ^{F}T19. A $[Mg(H_2O)_6]^{2+}$ ion is attached to the phosphate-sugar backbone through outersphere contacts with $O(P_6)$ and $O(P_7)$. The other three Mg^{2+} ions are not mentioned.							
Mg^{2+}	$[d(CGCGAA)T^*d(TCGCG)]_2$	1	163	1 Mg^{2+}	1.20	16.0	[124]
The DDD duplex involving modified T*7 and T*19. No metal ion interaction mentioned.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Mg^{2+}	$[d(CGCGAAT^{F}TCGCG)]_2$ (CGMg form)	1	–	5 Mg^{2+}	1.10	16.3	[125]
The DDD duplex involving modified ^{F}T8 and ^{F}T20. Mg1 ($[Mg(H_2O)_6]^{2+}$) is in the major groove, close to one end of the duplex, and engaged in outersphere contacts with N7 and O6 sites of G2 and G22 and with O(P)s of A6 and T7 of an adjacent duplex. This site is assumed to be the first one to be occupied and causes the DDD duplex to asymmetrically kink into the major groove. The other four Mg^{2+} ions, two of which are located at the periphery of the minor groove, are bridged between the backbones of two duplexes to bring short interduplex phosphate contacts in the lattice.							
Mg^{2+}	$[d(CGCGAA^{RS}TTCGCG)]_2$	1	133	3 Mg^{2+}	1.43	20.8	[126]
The DDD duplex containing interstrand dithiobis(propane) cross-links. The structure is on a par with the sodium form [109a] but no monovalent ion is in the minor groove. Mg1 ($[Mg(H_2O)_6]^{2+}$) is in the major groove between base-pairs 2 and 3 and engaged in outersphere contacts with O6 and N7 of G2 and O6 of G22 and O(P) of G22 and O(P)s of A6 and A7 of a symmetry-related duplex. Mg2 ($[Mg(H_2O)_5]^{2+}$) and Mg3 ($[Mg(H_2O)_5]^{2+}$) occupy interstitial spaces between laterally packed helices, and assist lattice formation via H-bonded bridges between neighboring phosphate oxygens.							
Mg^{2+}	$[d(CGCGAA^{F}T^{F}TCGCG)]_2$	1	177	1 Mg^{2+}	1.55	20.8	[127]
The DDD duplex involving modified ^{F}T7^{F}T8 and ^{F}T19^{F}T20. No metal ion interaction mentioned.							
Mg^{2+}	$[d(CGCGAA^{F}TTCGCG)]_2$	1	186	1 Mg^{2+}	1.55	21.8	[127]
The DDD duplex involving modified ^{F}T7 and ^{F}T19. No metal ion interaction mentioned.							
Mg^{2+}	$[d(CGCGAAt_{OH}t_{OH}CGCG)]_2$	1	66	1 Mg^{2+}	2.14	20.7	[128]
The DDD duplex containing four modified thymidines (t_{OH}) with the sugar O4' replaced by a carbon atom. A $[Mg(H_2O)_6]^{2+}$ ion is located between two adjacent symmetry-related duplexes and engaged in outersphere contacts with three O(P)s of T7 and A6 in one duplex and with N7 and O6 of G2 and O6 of G22 in the other duplex.							
Mg^{2+}	$[d\{CGCGAATT(\varepsilon dA)GCG\}]_2$	1	127	1 Mg^{2+}	2.25	18.1	[129]
No metal ion interaction mentioned.							
Mg^{2+}	$[d(ACCGGCGCCACA)]_2$ (tet form)	1	63	8 Mg^{2+}	2.6	14.5	[130]
Dodecamer duplexes form self-fitting B-DNA molecules by groove–backbone interaction. A $[Mg(H_2O)_6]^{2+}$ ion participates in the groove–backbone interaction, bridging between the flanking guanine residues (G5 and G7) of the central anchoring unit and the invading O(P)s of G19 and C21.							

Mg^{2+}	$[d(ACCGCCGGCGCC)]_2$ (ras form)	1	47	5 Mg^{2+}	2.6	17.3	[130]
Isomorphous to the tet form [130].							
Mg^{2+}	$[d(GCGAATTCG)]_2$	1	88	2 Mg^{2+}	2.05	16.8	[131]
Nonamer duplex consists of an eight-base pair B-form DNA helix with unpaired G bases at its 5'-end, which form (C·G)*G base triples with the adjacent duplex. Two Mg^{2+} ions are in the region of the triplets linking the two together. Mg1 ($[Mg(H_2O)_6]^{2+}$) ion is located in the vicinity of the C2·G18 base pair. Mg2 ($[Mg(H_2O)_4]^{2+}$) ion participates in innersphere contacts with N7 of G1 and O(P) of G12 of a symmetry-related duplex.							
Mg^{2+}	$[d(GCGAATTCG)]_2$	1	151	7 Mg^{2+}, 1 Cl^-	0.89	14.0	[132]
The structure is practically identical to that by Kennard et al. [131b]. One Cl^- and 6 Mg^{2+} ions maintain an ionic network of the mini-triplex, where 4 Mg^{2+} ions interact with two or three duplexes either directly or through hydration waters, while two fully hydrated Mg^{2+} ions (Mg1 and Mg2), which occupy equivalent positions at both ends of the duplex in the major groove, interact only with one duplex through water. The highly ordered Mg1 forms H-bonds with O6 and N7 of G3 and O4 of T16, whereas Mg2 is partially ordered. One Mg^{2+} ion corresponding to Mg1 is also present in Kennard's structure [131b]. The seventh Mg^{2+} ion lies at the edge of the phosphate backbone in the major groove and interacts through hydration waters with $O(P_4)$, $O(P_{14})$ and $O(P_{15})$, each from a different duplex. The Cl^- ion is between two Mg^{2+} ions and interacts with N4 of C2, and O6 of G10 of a neighboring duplex.							
Mg^{2+}	$[d(GGCCAATTGG)]_2$	1	116	2 $[Mg(H_2O)_6]^{2+}$	1.15	16.7	[133]
The decamer duplex consists of an eight-base-pair B-form DNA helix with unpaired GG bases at its 5'-end, which form (C·G)·G triple helices with the adjacent duplex. Two Mg^{2+} ions bridge three adjacent symmetry-related backbone strands, stabilizing the crystal lattice.							
Mg^{2+}	$[d(CCAACGTTGG)]_2$ (CGMg form)	0.5	93	4 Mg^{2+}	0.99	14.09	[134]
The duplex rides on a 2-fold axis. Mg1 and Mg2 are in the minor groove. Mg3 and Mg4 are in the major groove. Mg3 binds to G19 and G20 and Mg4 binds to G16 and T17 via their water ligands.							
Mg^{2+}	$[d(CCAGCGCTGG)]_2$ (AGMg form)	0.5	99	6 Mg^{2+}	0.99	14.05	[134]
Isomorphous with the CGMg form [134]. Metal binding positions of Mg1–Mg3 are the same as those in the CGMg form. Mg4 binds to A3 and G4.							
Mg^{2+}	d(GTTTTG)-Sd2d(CAAAAC)	2	170	5 $[Mg(H_2O)_6]^{2+}$	1.77	17.2	[135]
The asymmetric unit involves two independent Sd2-linked hairpins (Hp1and Hp2), each of which forms a duplex portion. Mg1or Mg2 ions are located in the A-tract minor grooves near the A11-A12 step of Hp1 or the A10-A11 step of Hp2, respectively. Mg3–Mg5 ions are located near the phosphate backbones of Hp1 and Hp2.							
Mg^{2+}	$[d(CGCGmo^6AATCCGCG)]_2$ (Dmo^6A·C form)	1	173	1 $[Mg(H_2O)_6]^{2+}$	1.6	19.6	[136]
The DDD-type duplex involving N6-modified mo^6A5 and mo^6A17 residues. A Mg^{2+} ion is located in the major groove and has outersphere contacts with O6 and N7 of G2 and O6 of G22 in a duplex and O(P)s of A6 and T7 from the neighboring duplex.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Mg^{2+}	[d(CGCGmo^{6}AATTCGCG)]$_2$ (Dmo^{6}A·T form)	1	126	1 $[Mg(H_2O)_6]^{2+}$	1.6	21.1	[137]
The DDD duplex involving N6-modified mo^{6}A5 and mo^{6}A17 residues. One Mg^{2+} ion occupies the same position as that in the Dmo^{6}A·C form [136]. The crystals contain per unit 1 spermine							
Mg^{2+}	[d(CGCGAATTmo^{4}CGCG)]$_2$ (Dmo^{4}C·G-1 form)	1	110	1 Mg^{2+}	1.6	22.3	[138]
The DDD duplex involving N4-modified mo^{4}C9 and mo^{4}C21 residues. No metal ion interaction mentioned.							
Mg^{2+}	[d(CGCGAATTmo^{4}CGCG)]$_2$ (Dmo^{4}C·G-2 form)	1	80	1 $[Mg(H_2O)_6]^{2+}$	2.1	21.1	[139]
The DDD duplex involving N4-modified mo^{4}C9 and mo^{4}C21 residues. One Mg^{2+} ion occupies the same position as that in the Dmo^{6}A·C form [136].							
Mg^{2+}	[d(CGCAAATTmo^{4}CGCG)]$_2$ (Dmo^{4}C·A form)	1	111	1 $[Mg(H_2O)_6]^{2+}$	1.6	21.8	[140]
The DDD-type duplex involving N4-modified mo^{4}C9 and mo^{4}C21 residues. One Mg^{2+} ion occupies the same position as that in the Dmo^{6}A·C form [136].							
Mg^{2+}	[d(CGCGAAT^{f5}UCGCG)]$_2$ (B Form)	1	185	1 $[Mg(H_2O)_6]^{2+}$	1.5	20.7	[119]
The DDD-type duplex involving modified ^{f5}U8 and ^{f5}U20 residues. One Mg^{2+} ion occupies the same position as that in the Dmo^{6}A·C Form [136].							
Mg^{2+}	[d(GCGAAAGC)]$_2$ (*I*422 Form)	1	72	1.25 $[Mg(H_2O)_6]^{2+}$	1.60	19.0	[141]
Two DNA strands form a base-intercalated duplex, and four duplexes are assembled by a crystallographic 222 symmetry to form an octaplex. One Mg^{2+} ion binds through H-bonds to O6 and N7 of each G7. The other one rides on a 4-fold axis but does not contact with any residue.							
Mg^{2+}	[d(CCGAA^{br}UGAGG)]$_2$	1	77	3 $[Mg(H_2O)_5]^{2+}$	1.60	19.2	[200]
Two terminal C1 and G20 bases are swung out from the duplex. One Mg^{2+} ion is directly bound to N7 of G9 next to A8 which is involved in a sheared G–A base pair and indirectly through H-bonds to adjacent O(P)s. The second Mg^{2+} ion, which forms a 2-fold symmetry-related dinuclear cluster, interacts with several phosphate groups, either directly or through H-bonds. The third Mg^{2+} ion binds to O(P) of A10 and interacts through H-bonds to a sheared G7–A14 base pair in the minor groove of a neighboring duplex. The crystals also contain per unit 0.5 spermine.							
Mg^{2+}, Co^{2+}	[d(CGCGAAT^{N}U$_m$CGCG)]$_2$ (DD2 form)	1	236	1 Mg^{2+}, 3 Co^{2+}	1.6	21.5	[116]
The DDD-like sequence with the 8-position replaced by ^{N}U$_m$. No metal ion interaction mentioned.							

Mg^{2+}, Co^{3+}	$[d(CCGAATGAGG)]_2$	1	128	2 $[Mg(H_2O)_5]^{2+}$, 1 $[Co(NH_3)_6]^{3+}$	1.20	18.2	[200]
The two Mg^{2+} ions only interact with phosphate groups. A Co^{3+} ion is bound through H-bonds to O6 and N7 of G7, which is involved in a sheared G–A base-pair, and O(P) of T6 of the same strand.							
Mg^{2+}, Co^{3+}	$[d(GCGAAAGCT)]_2$ ($I4_122$ form)	0.5	16	1.5 $[Mg(H_2O)_6]^{2+}$, 0.5 $[Co(NH_3)_6]^{3+}$	2.5	21.5	[142]
The duplex rides on a 2-fold axis. The sequence C2–A5 forms a parallel duplex,whereas the A6–T9 sequence forms an anti-parallel duplex. The 5'-terminal Gs are flipped out of the duplex. One Mg^{2+} ion is bound through outersphere contacts to O6 and N7 of G3 and O(P) of C2, and N2 of G1 from the neighboring duplex. The other Mg^{2+} ion, which rides on a 2-fold axis, is bound through outersphere contacts to O2 atoms of T9 residues located at the junction between the parallel and anti-parallel duplexes. A Co^{3+} ion, which rides on a 2-fold axis, is bound to the bent corner joining the parallel and anti-parallel duplexes, through outersphere contacts with O6 and N7 of G7 and O(P) of A5.							
Mg^{2+}, Co^{3+}	$[d(G^{I}CGAAAGCT)]_2$ ($P6_322$ form)	0.5	48	0.33 Mg^{2+}, 0.33 $[Co(NH_3)_6]^{3+}$, 0.33 Cl^-	1.4	20.9	[143]
Isomorphous with the $[Co(NH_3)_6]^{3+}$ complex of $[d(G^{Br}CGAAAGCT)]_2$ ($P6_322$ form) [154] having a zipper-like motif.							
Mg^{2+}, Co^{3+}	$[d(GCGAAAGC)]_2$ ($P6_322$ form; 8hmt-A_5 form)	0.5	28	1.33 Mg^{2+}, 0.33 $[Co(NH_3)_6]^{3+}$, 0.33 Cl^-	2.15	22.7	[143]
Isomorphous with the $[d(G^{I}CGAAAGCT)]_2$ ($P6_322$ form) [143]. A hydrated Mg^{2+} ion is bound to the major groove site of G7.							
Ca^{2+}	$[d(CGCGAATTCGCG)]_2$	1	144	3.33 Ca^{2+}	1.45	21.0	[110]
The DDD duplex. A $[Ca(H_2O)_3]^{2+}$ ion, which lies on a 3-fold axis, binds to three $O(P_5)$s from the three duplexes, stabilizing the packing arrangement. Another Ca^{2+} ion stabilizes the interaction between $O(P_{12})$ and $O(P_{24})$ at the ends of neighboring duplexes and also binds to O6 of the 3'-terminal G from the third duplex.							
Ca^{2+}	$[d(CCAACGTTGG)]_2$ (CGCa form)	0.5	109	4 Ca^{2+}	0.99	11.83	[134]
Isomorphous with the CGMg form [125]. Metal binding positions of Ca1–Ca3 are the same as those of Mg1–Mg3 in the CGMg form. Ca3 binds directly to N7 of G19 and O6 of G20.							
Ca^{2+}	$[d(CCAGCGCTGG)]_2$ (AGCa form)	0.5	126	4 Ca^{2+}	0.99	12.14	[134]
Isomorphous with the CGMg form [125]. Metal binding positions of Ca1–Ca3 are the same as those of Mg1–Mg3 in the CGMg form. Ca3 binds directly to N7 of G19 and O6 of G20. Ca4 binds to G4.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Ca^{2+}	$[d(GCGAATTCGCG)]_2$ (GCCa form)	1	125	6 Ca^{2+}	1.30	18.2	[144]
The DDD duplex lacking the 5'-terminal C. The 11-mer forms a decamer duplex with G overhangs at both 3'-termini. Four Ca^{2+} ions (Ca1–Ca3, Ca6) participate in the stabilization of the end-to-end interaction between duplexes, through either innersphere or outersphere contacts with phosphate groups. The other two Ca^{2+} ions (Ca4 and Ca5) are located at the duplex interface. Ca1 ($[Ca(H_2O)_2]^{2+}$) is engaged in four innersphere contacts, three to O(P)s and one to O6 of G12. Ca3 is the only ion that is located in a groove (a position corresponding to Mg1 in the CGMg form [125]). It is coordinated at the boundary between the A-tract and the GCG section at one end of the duplex. In addition to contacting three O4' oxygen atoms and a terminal OH group from a third strand, Ca3 also forms H-bonds to N2 and O2 of the C21·G4 base pair and to N3 of A5 via coordinated water molecules.							
Ca^{2+}	$[d(GGCGAATTCGCG)]_2$ (GGCa form)	1	151	4 Ca^{2+}	1.70	19.4	[144]
The DDD duplex with the 5'-terminal C replaced by G. Three Ca^{2+} ions are located at the positions corresponding to those of Ca1, Ca3 and Ca4 in the GCCa form [144]. Another one is in the minor groove and engaged in outersphere contacts with N2 of G16, O2 of C9, and N3 of A17.							
Ca^{2+}	$[d(CGCGAA^{F}T^{F}TCGCG)]_2$ (CGCa form)	1	177	3 Ca^{2+}	2.20	19.7	[144]
The DDD duplex involving modified $^{F}T7^{F}T8$ and $^{F}T19^{F}T20$. Three Ca^{2+} ions are located at the positions corresponding to those of Ca1, Ca3 and Ca4 in the GCCa form [144].							
Ca^{2+}	$[d(CTCTCGAGAG)]_2$	1	114	6 Ca^{2+}	1.7	18.3	[145]
Three hydrated Ca^{2+} ions are located in the minor groove and mediate the close approach of negatively charged phosphates, arising from the crossed arrangement of duplexes.							
Ca^{2+}	$[d(CCAGTACTGG)]_2$	0.5	131	4 Ca^{2+}	0.74	10.5	[146]
The duplex rides on a 2-fold axis. Each of the four Ca^{2+} ions interacts with symmetry-related duplexes to stabilize the crystal lattice. Two Ca^{2+} ions (Ca2 and Ca3) are in the major groove and the other two (Ca1 and Ca4) are in the minor groove and engaged in outersphere contacts with the edges of the bases (for Ca1–Ca4), with additional contacts to symmetry-related phosphate and sugar oxygen atoms (Ca1 and Ca4). Ca2 additionally binds *directly* to O6 of the *terminal* G10 and N7 of G9. Ca4 is bound in the minor groove of the G-C base-pairs in the center of the duplex.							
Ca^{2+}	$[d(CGCGAATTCGCG)]_2$	1	15	1 Ca^{2+}	2.5	19.3	[147]
The DDD duplex. No Ca^{2+} ion was detected in the minor groove. One Ca^{2+} ion plays a role in stabilizing the packing arrangement.							
Ca^{2+}	$[d(CGCTCTAGAGCG)]_2$	2	47	1 Ca^{2+}	2.25	19.7	[148]
A tetrahedral Ca^{2+} ion interacts with four phosphates at the ends of four different duplexes.							

Sr^{2+}	d(G^{Br}UTTTG)-SE2d(CAAAAC)	4	11	4 Sr^{2+}	3.2	25.7	[149]
The asymmetric unit involves four independent SE2-linked hairpins, each of which forms a duplex portion. In each hairpin duplex, a Sr^{2+} ion is located near the SE-end of the hairpin, in the center of the minor groove, and forms outersphere contacts to O2 of pyrimidine and N3 of purine.							
Mn^{2+}	$[d(AAATATTT)]_2$	5.5	16	6 Mn^{2+}	3.1	27.2	[150]
One of six independent duplexes rides on a 2-fold axis. Mn^{2+} ions are tentatively assigned. The crystals also contain per unit 6 cacodylate molecules.							
Mn^{2+}	$[d(CGTTAATTAACG)]_2$	3	34	7 Mn^{2+}	2.6	23.5	[151]
Six independent strands (A–F) (three duplexes A-B, C-D, and E-F) in the asymmetric unit. The 3'-terminal guanines, which are flipped out of the duplex, enter the minor groove of the next duplex and interact with another guanine through N2···N3 hydrogen bonds. Seven Mn^{2+} ions maintain contacts between neighboring duplexes as follows: N7(G12/B)-Mn-N7(G12/D), N7(G12/C)-Mn-O(P)(T7/F), N7(G12/A)-Mn-N7(G12/F), O(P)(T7/D)-Mn-N7(G12/F), O(P)(G12/A)-Mn-O(P)(G12/B), O(P)(G12/E)-Mn-O(P)(G12/F), and O(P)(G12/C)-Mn-O(P)(G12/D).							
Co^{2+}	$[d(GGCGCC)]_2$	2.5	–	5 Co^{2+}	3.0	22.7	[152]
Isomorphous to the Ni^{2+} and Zn^{2+} complexes [152]. One of three independent duplexes rides on a 2-fold axis. In each strand, a $[Co(H_2O)_5]^{2+}$ ion is engaged in an innersphere contact with N7 of each 5'-terminal G and outersphere contacts with O6 of the same G and O(P)s from the neighboring duplex.							
Co^{2+}	$[d(CGCAATTGCG)]_2$	0.5	46	0.5 Co^{2+}	1.6	21.4	[153]
The duplex rides on a 2-fold axis. The terminal C1 and G10 are flipped out and C1 forms a $C1^{+}$·G8·C3 triplet with the neighboring duplex. A $[Co(H_2O)_4]^{2+}$ ion, which rides on a 2-fold axis, binds to N7 atoms of two flipped out G10 guanines in two symmetry-related duplexes.							
Co^{3+}	$[d(G^{Br}CGAAAGCT)]_2$ ($P6_322$ form)	0.5	18	0.33 $[Co(NH_3)_6]^{3+}$, 1 Cl^-	2.1	18.9	[154]
The duplex rides on a 2-fold axis. The 3'-teminal T residues project out of the duplex. The core adenines are unpaired and intercalated to form a four-step zipper-like motif. A $[Co(NH_3)_6]^{3+}$ion, which rides on a 3-fold axis, links three symmetry-related duplexes through outersphere contacts with N7 and O6 atoms of three G3 residues from three symmetry-related duplexes.							
Co^{3+}	$[d(GCGAAGC)]_2$ ($P2_12_12_1$ form)	1.5	304	9 $[Co(NH_3)_6]^{3+}$, 54 NH_4^+	1.6	19.4	[155]
Three independent duplexes in the asymmetric unit, each of which rides on a 2-fold axis. Two G3·A4* and A4·G3* pairs are crossed and stacked on each other. Three $[Co(NH_3)_6]^{3+}$ ions are in the major groove of each duplex. One is bound through H-bonds to O6 and N7 of G3 and O6 of G3* in the duplex and O(P) of A5 of a neighboring duplex. The other two are bound through H-bonds to O6 and N7 of G6 (and G6*) of one duplex and O(P) of C7 (and C7*) of a neighboring duplex.							
Co^{3+}	$[d(GCAGACGTCTGC)]_2$	3	50	3 $[Co(NH_3)_6]^{3+}$	2.5	23.5	[156]
Three independent duplexes in the asymmetric unit. One Co^{3+} ion is located in the major groove of each duplex and is additionally bridged to the neighboring duplex.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit			Resolution (Å)	*R* (%)	Ref.
		DNA duplex	H_2O	Metal			
Co^{3+}	$[d(CCGAA^{Br}UGAGG)]_2$	1	133	0.5 $[Co(NH_3)_6]^{3+}$	1.50	23.1	[200]
Two terminal C1 and G20 bases are swung out from the duplex. A Co^{3+} ion, which rides on a 2-fold axis, is bound through H-bonds to O6 and N7 atoms of G3 and G17, which are involved in sheared G–A base-pairs.							
Ni^{2+}	$[d(GGCGCC)]_2$	1	–	5 Ni^{2+}	2.9	22.3	[152]
Isomorphous to the Co^{2+} and Zn^{2+} complexes [152].							
Ni^{2+}	$[d(CGTATATACG)]_2$	2	60	9 Ni^{2+}	2.99	20.4	[157]
Two independent duplexes in the asymmetric unit. 3'-Terminal G10 and G20 (G30 and G40) are flipped out and located in the minor groove of a neighboring decamer to form G·C·G triplets. Seven Ni^{2+} ions bind directly to all the guanines through N7, one of which is shared between two guanines of symmetry-related molecules.							
Ni^{2+}	$[d(CGTGTACACG)]_2$	2	12 (with 40 modeled H_2O)	8 Ni^{2+}	3.1	20.9	[158]
Isomorphous with the $[d(CGTATATACG)]_2$ structure [157]. Two independent duplexes in the asymmetric unit. Seven of the eight Ni^{2+} ions bind directly to the N7 sites of all guanines in either extra-helical position (G10, G20, G30, or G40) or terminal position (G2, G12, G22, or G32) of the duplex, where one Ni^{2+} ion is shared between two guanines G10 and symmetry-related G30. The eighth Ni^{2+} ion is between two phosphate groups of neighboring duplexes. No Ni^{2+} ion binds to guanines (G4, G14, G24, or G34) that lie within a *standard B-DNA* tract.							
Ni^{2+}	$[d(GAATTCG)]_2$	1	20	4 $[Ni(H_2O)_5]^{2+}$	2.81	28.2	[159]
The 3'-terminal G is flipped out of the duplex and sticks with the minor groove of the next duplex. Four Ni^{2+} ions bind directly to N7 of all guanines: two bind to two flipped out G7 and G14, accompanying water-mediated contacts with neighboring phosphates; the other two Ni^{2+} ions bind to two 5'-terminal G1 and G8, and one of the two ions acts as a branching point by further connecting the phosphate of A3 of the 2-fold symmetry-related duplex through hydration waters.							
Zn^{2+}	$[d(GGCGCC)]_2$	1	–	5 Zn^{2+}	2.9	21.0	[152]
Isomorphous to the Co^{2+} and Ni^{2+} complexes [152]. The crystals also contain per unit 1 spermine.							

Zn^{2+}	$[d(CGCAATTGCG)]_2$	0.5	29	2.5 Zn^{2+}	2.9	20.8	[160]
The 5'-terminal C is modified by a CH_3CO-$(Arg)_4$-NH-$(CH_2)_6$-NH-p-fragment. The duplex rides on a 2-fold axis. The terminal C1 and G10 are in extra-helical positions. A $[Zn(H_2O)_2]^{2+}$ ion, which rides on a 2-fold axis, binds to four O(P)s of C1 and G10, each from four symmetry-related duplexes. Two other $[Zn(H_2O)_5]^{2+}$ ions bind to N7 of G2 and G10 but not to the internal G8.							
Zn^{2+}	$[d(CGCG\text{-}tcA\text{-}ATTCGCG)]_2$	1	95	1 Zn^{2+}	1.75	17.6	[161]
The DDD duplex modified with the tricycle sugar of A5 (tcA5). 5'-Terminal Cs and 3'-terminal Gs are unpaired. A tetrahedral Zn^{2+} ion binds directly to O(P)s of 3'-terminal G12 and G24 and N3 of 5'-terminal C1 and C13 from four symmetry-related duplexes.							
Rh^{3+}	$[d\{G(dIU)TGCAAC\}]_2$	5	374	8 Δ-α-[Rh{(*R*,*R*)-Me_2trien}phi]$^{3+}$	1.2	17.0	[162]
Five independent duplexes in the asymmetric unit. In each duplex, a Rh^{3+} complex is intercalated into the major groove of the 5'-TG|CA-3' sequence (where | indicates phi insertion), where sequence-specific contacts are formed with the edges of the bases.							
Rh^{3+}	$[d(CGGAAATTCCCG)]_2$	1	75	2 Δ-[Rh$(bpy)_2$(chrysi)]$^{3+}$	1.14	15.1	[163]
One Rh^{3+} complex is intercalated into the major groove of the 5'-AA|TT-3' sequence (where | indicates chrysi insertion). The other Rh^{3+} complex is inserted into the minor groove at a mismatched site by replacing the A4·C21 mismatched base pair.							
Rh^{3+}	$[d(CGGAAATTACCG)]_2$ ($P3_121$ form)	1	89	2 Δ-[Rh$(bpy)_2$(chrysi)]$^{3+}$	1.60	18.4	[164]
Two Rh^{3+} complexes are inserted into the minor groove at each A·A mismatch by ejecting the A4·A21 and A9·A16 mismatched base pairs outward into the major groove.							
Rh^{3+}	$[d(CGGAAATTACCG)]_2$ ($P4_32_12$ form)	0.5	63	1.5 Δ-[Rh$(bpy)_2$(chrysi)]$^{3+}$	1.80	18.3	[164]
The duplex and one of the two Rh^{3+} complexes ride on a 2-fold axis. Thus each duplex contains three Rh^{3+} complexes, one inserted into the minor groove at each A·A mismatched site (A4·A9* and A9·A4*) and a third intercalated into the major groove of the 5'-AA|TT-3' sequence (where | indicates chrysi insertion).							
Pt^{2+}	d(CCTCGCTCTC)·d(GAGAGCGAGG)	1	92	1 *cis*- $[Pt(NH_3)_2]^{2+}$	1.63	16.92	[165]
Decamer duplex forms a B-like DNA structure, where two cytosines C6 and C16 are extruded from the duplex and C6 makes a triplet with a terminal G·C pair of a different molecule. A *cis*-$[Pt(NH_3)_2]^{2+}$ ion *interstrand* cross-links two guanine bases G5 and G15 through their N7 sites. The duplex is bent by 47° at the platinated site toward the minor groove without disrupting G·C base pairs.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Pt^{2+}	d(CCTCTCGTCTCC)·d(GGAGACGAGAGG)	1	–	2 *cis*-$[Pt(NH_3)_2(py)]^{2+}$	2.17	22.5	[166]
Dodecamer duplex forms a B-like DNA structure with platinum mono-functionally coordinated to N7 of G7 in the major groove. The duplex is unwound by ~8° in the vicinity of the platination site.							
Left-handed double-stranded Z-DNA							
Na^+	$[d(CCGCGG)]_2$	1	38	1 Na^+	1.92	18.5	[173]
No metal ion interaction mentioned.							
Na^+	$[d(CGCGCG)]_2$ (spermine form)	1	62	2 Na^+	1.0	18.0	[174]
No metal ion interaction mentioned. The crystals also contain per unit 28 spermine molecules.							
Na^+, Mg^{2+}	$[d(CGCGCG)]_2$	1	117	1 Na^+, 3 Mg^{2+}	1.0	19.1	[175]
A $[Na(H_2O)_6]^+$ ion forms an outersphere contact with O(P) of C5. A $[(H_2O)_3Mg(\mu\text{-}H_2O)_2Mg(H_2O)_4]^{4+}$ cluster ion is formed and engaged in outersphere contacts with O6 of G2 and N4 of C3 in the same strand and N4 of C9 and O6 of G10 in the neighboring duplex. Another $[Mg(H_2O)_6]^{2+}$ ion forms outersphere contacts with O6 of G4 and O6 and N7 of G8 in the same duplex and O(P) of G6 in the neighboring duplex. The crystals also contain per unit 1 spermidine molecule.							
Na^+, Mg^{2+}	$[d(CGCGCG)]_2$	1	–	2 Na^+, 3 Mg^{2+}	1.0	23.2	[176]
Metal ions occupy the same positions as those observed in the polyamine-free complex [178]: a $[Na(H_2O)_6]^+$ ion occupies Site2; two Mg^{2+} ions form a $[(H_2O)_3Mg(\mu\text{-}H_2O)_2Mg(H_2O)_4]^{4+}$ cluster and occupy Site1 and Site1'; another $[Mg(H_2O)_6]^{2+}$ ion occupies Site3. The crystals also contain per unit 1 spermidine molecule.							
Rb^+	$[d(CCGCGG)]_2$	1	91	3(full) 13(partial) Rb^+, 1 Cl^-	1.76	17.5	[177]
Rb^+ ions stabilize the DNA structure by cross-linking two strands or two duplexes of the DNA: each of ten Rb^+ ions directly coordinates to at least one O(P), whereas only one coordinates to N7 of G2; five Rb^+ ions bind directly to O6 atoms of G2, G4, G6, G8, and G12, N2 of G8, O2 atoms of C1 and C7, O3' atoms of C1, G2, C7, and C9, and O5' of G6.							

Mg^{2+}	$[d(CGCGCG)]_2$ (polyamine-free form)	1	–	4 Mg^{2+}	1.0	13.9	[178]
A $[(H_2O)_3Mg(\mu\text{-}H_2O)_2Mg(H_2O)_4]^{4+}$ cluster ion (Mg1 and Mg1') is formed and engaged in an innersphere contact with N7 of G6 and outersphere contacts with O6 and N3 of G6 and O(P) of C5 in the same duplex and O(P)s of G10 and G12 in the neighboring duplex (Site1 and Site1'). Mg2 ion ($[Mg(H_2O)_6]^{2+}$) forms outersphere contacts with O(P) and O4' of G6 and O(P) of C5 of the same strand and O(P)s of C9 and G10 of the neighboring duplex (Site2). Mg3 ion ($[Mg(H_2O)_6]^{2+}$) forms outersphere contacts with O6 of G4, O6 and N7 of G8, and O(P) of G10 in the same duplex and O(P) of G6 in the neighboring duplex (Site3). Mg4 ion ($[Mg(H_2O)_6]^{2+}$) forms outersphere contacts with N7 of G4 and O(P)s of C3 and G6 of the same strand and O(P) of G8 in the neighboring duplex (Site4).							
Mg^{2+}	$[d(CGCGCG)]_2$	1	–	1 Mg^{2+}	1.5	15	[179]
Crystallized at 20°C. A Mg^{2+} ion binds directly to O(P) of G2. The crystals also contain per unit 2 thermine molecules.							
Mg^{2+}	$[d(CGCGCG)]_2$	1	–	3 Mg^{2+}	1.17	23.9	[179]
Crystallized at 10°C under high salt conditions. Three Mg^{2+} ions occupy Site1, Site1', and Site3 in the polyamine-free complex [178]. The crystals also contain per unit 1 thermine molecule.							
Mg^{2+}	$[d(CGCGCG)]_2$	1	–	1 Mg^{2+}	1.0	18	[179]
Crystallized at 4°C. A Mg^{2+} ion is located near Site4 in the polyamine-free complex [178]. The crystals also contain per unit 2 thermine molecules.							
Mg^{2+}	$[d(CGCGCG)]_2$ (spermine/Mg^{2+} form)	1	74	1 Mg^{2+}	0.9	14	[174]
No metal ion interaction mentioned. The crystals also contain per unit 28 spermine molecules.							
Mg^{2+}	$[d(CGCGCG)]_2$ (Mg^{2+} form)	1	85	4 Mg^{2+}	1.0	17.5	[174]
One hydrated Mg^{2+} ion forms outersphere contacts with O6 of G6 in one duplex and O6 of G12 in the next duplex. The other two hydrated Mg^{2+} ions are at C3 and C9 on the convex surface. The forth Mg^{2+} ion forms outersphere contacts with N2 and O(P) of G8 in the minor groove.							
Mg^{2+}	$[d(CGCGCG)]_2$	1	72	1 Mg^{2+}	1.0	19	[180]
A $[Mg(H_2O)_6]^{2+}$ ion forms outersphere contacts with O6 and N7 atoms of G2 and G10 and N4 of C3 in the same duplex and O(P)s of neighboring duplexes. The crystals also contain per unit 2 thermospermine molecules.							
Mg^{2+}	$[d(CGCGCG)]_2$	1	–	1 Mg^{2+}	1.0	13.8	[181]
A $[Mg(H_2O)_4]^{2+}$ ion binds directly to O(P) of C3. The crystals also contain per unit 2 PA(222) molecules.							
Mg^{2+}	$[d(TGCGCA)]_2$	1	58	2 Mg^{2+}	1.64	21.2	[182]
The structure is comparable with that of the $[Co(NH_3)_6]^{3+}$ complex with the identical sequence [185]. Both Mg^{2+} ions make outersphere contacts with atoms of the DNA helix.							

(continued)

Table 6 (continued)

Metal	Sequence	Number of molecules in the asymmetric unit: DNA duplex	H_2O	Metal	Resolution (Å)	*R* (%)	Ref.
Mg^{2+}, Co^{3+}	$[d(Gm^5CGCGCG)]_2$	1	49	1 $[Mg(H_2O)_6]^{2+}$, 1$[Co(NH_3)_6]^{3+}$	1.68	20.7	[183]
The 5'-terminal G1 is overhanged from the duplex. A Co^{2+} ion forms outersphere contacts with O6 and N7 of G5 and O(P) of C4. A Mg^{2+} ion links three duplexes, forming outersphere contacts with O(P) of C4 in one duplex and O(P)s of two neighboring duplexes and with N7 of G12 on one of these duplexes.							
Co^{3+}	$[d(GCGCGCG)]_2$	1	57	2 $[Co(NH_3)_6]^{3+}$	1.65	18.4	[184]
The 5'-terminal G1 is overhanged from the duplex. Co1 forms outersphere contacts with O6 and N7 of G5 and O(P) of C4 in DuplexI and O(P)s of G5 and G14 in DuplexII. Co2 forms outersphere contacts with O(P) of C4 in DuplexI and O(P)s of C13 and G14 in DuplexII and N7 of G12 and O(P) of C11 in DuplexIII.							
Co^{3+}	$[d(TGCGCA)]_2$	1	78	2 $[Co(NH_3)_6]^{3+}$	1.3	12.0	[185]
Co1 forms outersphere contacts with O6 and N7 of G4 and O(P) of C3. Co2 forms outersphere contacts with O(P)s of A6 and A12 in the same duplex.							
Co^{3+}	$[d(CACG(dIU)G)]_2$	1	43	2 $[Co(NH_3)_6]^{3+}$	1.3	16.7	[186]
Co1 forms outersphere contacts with O6 and N7 of G6 in one duplex and O(P)s of G4 and G12 in a neighboring duplex. Co2 forms outersphere contacts with O(P)s of G6 and G10 in the same duplex, O(P) of dIU5 and O(3') of G4 in the second duplex, and N7 of G4 in the third duplex.							
Co^{3+}	d(CGCGCA)·d(TGCGCG) ($P2_12_12_1$ form)	1	45	1 $[Co(NH_3)_6]^{3+}$	1.71	20.8	[187]
A $[Co(NH_3)_6]^{3+}$ ion forms outersphere contcts with N7 and deprotonated N6 of A6 in one duplex, the G12·C1 base pair of the second duplex, and O6 and N7 of G4 in the third duplex.							
Co^{3+}	d(CGCGCA)·d(TGCGCG) ($P6_5$ form)	1.33	–	1 $[Co(NH_3)_6]^{3+}$	2.0	27.3	[187]
One of the two duplexes rides on a 6-fold screw axis. A disordered $[Co(NH_3)_6]^{3+}$ is assumed to bind through H-bonds to O6 and N7 of G12 in one duplex, deprotonated N6 of A6 in the second duplex, O6 and N7 of G8 in the third duplex, and O6 of G8 in the fourth duplex.							
Ru^{3+}	d(CGCGCA)·d(TGCGCG) ($P2_12_12_1$ form)	1	51	1 $[Ru(NH_3)_6]^{3+}$	1.54	18.9	[188]
The metal binding mode is similar to that in the $[Co(NH_3)_6]^{3+}$ complex ($P2_12_12_1$ form) [187].							
Pt^{2+}	$[d(CGTA'CG)]_2$	0.5	24	1 $[Pt(L)_3]^{2+}$	1.60	17.0	[189]
The duplex rides on a 2-fold axis. A square planar $[Pt(NH_3)(L_1)(L_2)]^{2+}$ (L_1, L_2 = NH_3 or H_2O) ion binds to N7 of G6.							

[a] Table presents crystal structures reported after 1994. For those before 1994, see Table 3 in [1]. [b] Abbreviations: O(P) = phosphate oxygen; O(W) = water oxygen; TOE = 2'-*O*-methyl-[tri(oxyethyl)]; FET = 2'-*O*-fluoromethyl; MEP = 2'-*O*-methyl-3'-methylenephosphonate; BTL = 2'-*O*-butyl; TFE = 2'-*O*-[2-(trifluoro)-ethyl; IME = 2'-*O*-[2-(imidazolyl)ethyl]; ^{M}T = 2'-methoxy-3'-methylenephosphonate (2'MeO-3'MP)-thymidine; m^{5}C = 5-methyl-cytidine; Sd2 = bis(2-hydroxyethyl)stilbene-4,4'-diether; BOE = 2'-*O*-[2-(benzyloxy)ethyl; AOE = 2'-*O*-aminooxyethyl; MAOE = 2'-*O*-[2-{(methyleneamino)oxy}ethyl]; DACH = diaminocyclohexane; NH_2cy = cyclohexylamine; DDD = Dickerson–Drew dodecamer; 8hmt-X = octamer adopting the hairpin structure with the sequence d(GCGAXAGC) which is mutated at the 5th residue by X; *I* = *I*-centering lattice; ^{Br}C = 5-bromo-cytidine; TriplatinNC = [{*trans*-Pt(NH_3)$_2$(NH_2(CH_2)$_6$$NH_3^+$)}$_2$-μ-{*trans*-Pt($NH_3$)$_2$($NH_2$($CH_2$)$_6$$NH_2$)$_2$}]$^{8+}$; TriplatinNC-A = [{*trans*-Pt(NH_3)$_3$}$_2$-μ-{*trans*-Pt(NH_3)$_2$(NH_2(CH_2)$_6$$NH_2$)$_2$}]$^{6+}$; ^{N}U = 2'-deoxy-5-(*N*-aminohexyl)carbamoyl-uridine; ^{f5}U = 2'-deoxy-5-formyluridine; ^{F}T = 2'-deoxy-2'-fluoro-arabino-furanosyl thymidine; T* = (L)-α-threofuranosyl thymidine; A^{RS} = A6 and A18 are linked by a –(CH_2)$_2$-S-S-(CH_2)$_2$- tether; t_{OH} = 6'-α-hydroxy carbocyclic thymidine; εdA = 1,N6-ethenoadenosine; mo^{6}A = 2'-deoxy-*N*^{6}-methoxyadenosine; mo^{4}C = 2'-deoxy-*N*^{4}-methoxycytidine; ^{N}U$_m$ = 5-(*N*-aminohexyl)carbamoyl-2'-*O*-methyl-uridine; ^{I}C = 5-iodo-cytidine; ^{Br}U = 5-bromo-uridine; SE2 = bis(2-hydroxyethyl)stilbene-4,4'-diether; tcA = tricycle-sugar modified 2'-deoxy adenosine; dIU = 5-iodo-2'-deoxyuridine; Me_2trien = dimethyltriethylenetetraamine; phi = 9,10-phenanthrenequinone diimine; chrysi = 5,6-chrysenequinone diimine; bpy = 2,2'-bipyridine; py = pyridine; PA(222) = *N*^{1}-[2-(2-amino-ethylamino)-ethyl]-ethane-1,2-diamine; A' = 2-amino adenosine.

[c] Numbering of, for example,

(GGTGAGG)· = 5'-G1—G2—T3—G4—A5—G6—G7-3'
(CCTCACC) · · · · · · ·
3'-C14–C13–A12–C11–T10–C9—C8-5'.

[d] No data are given.

oxygens is also observed in other A-DNA duplexes [92,104]. In all cases, $[Co(NH_3)_6]^{3+}$ ions play important roles in bridging duplexes to facilitate the crystal packing. Four crystal structures of platinum antitumor agents of the dodecamer d(CCTCTGGTCTCC)·d(GGAGACCAGAGG) duplex are of particular interest as models for *intra*strand cross-links between adjacent purine bases: *cis*-$[Pt(NH_3)_2]^{2+}$ (cisplatin) complexes at 2.6 Å resolution [105] and at 1.77 Å resolution [101], $[Pt(R,R\text{-DACH})(oxalato)]^{2+}$ (oxaliplatin) [103], and *cis*-$[Pt(NH_3)(NH_2Cy)]^{2+}$ [106] complexes. In each complex, the duplex consists of the A-type DNA segment at the 5'-side and the B-type segment at the 3'-side with the A-B junction occurring at the base pair T8·A17, the site being adjacent to the platination. Each platinum-diamine compound cross-links the two centrally located guanine bases G6 and G7 of one strand through their N7 sites. The duplex is bent at the platinated site by 30–45° toward the major groove without disrupting G·C base pairs.

In summary, four major metal binding patterns are observed for crystalline A-DNA. One is the same as (i) the *phosphate* bonding noted above in the summary for A-RNA. (ii) *Polyhedral* metal cations, divalent ions in particular, bind to *base* moieties in the *major* groove through *outer*sphere contacts, most preferably to *guanine* bases at N7 and /or O6 sites and, often in addition, to the adjacent phosphate group of the same strand. (iii) When the metal ion is *dehydrated* to an *extreme* extent, though this has been observed only for Ba^{2+}, the *direct* metal bonding to *guanine* bases through N7 and/or O6 sites is observed. (iv) *Square planar* metal ions can *directly* bind to *base* moieties in the *major* groove: *cis*-$[Pt(diamine)]^{2+}$ compounds cross-link two adjacent *guanine* bases by binding to N7 sites.

3.2.2 B-DNA Fragments

B-DNA adopts a right-handed double-helical structure in which the major and minor grooves are wide. The determination [167] of the structure of a $[d(CGCGAATTCGCG)]_2$ duplex, which is a so called Dickerson–Drew dodecamer (DDD), was a landmark in structural biology [168] and DDD and related dodecamers have been extensively studied. Intriguing features of the DDD duplex are the narrowing of the minor groove in the central A-tract and ordered water molecules that line this groove, referred to as a spine of hydration [167] (a ribbon of hydration [120]). The eventual substitution of some of the molecules in the water spine by monovalent cations has been a question of debate [71,109,110,120,125,126,168]. In the high-resolution crystal structures, a single monovalent ion is located with partial occupancy in the minor groove at the 5'ApT3' step (AT-tract) for Na^+ (1.40 Å resolution) [109], K^+ (1.2 Å) [118], Rb^+ (1.2 Å) [120], Cs^+ (1.8 Å) [121], and Tl^+ (1.49 Å) [122], where an octahedral $[M(H_2O)_2]^+$ ion makes innersphere contacts with two O2 keto oxygen atoms of T7 and T19 and two O4' ether oxygen atoms of T8 and T20. On the contrary, monovalent cations are absent in the minor grove in the cross-linked DDD duplex (1.43 Å) [126] and it is argued [126] that occasional binding of monovalent cations within the minor groove of A-tracts may be a consequence of groove narrowing but not the cause. In all the high-resolution DDD structures (1.40 Å [109],

1.50 Å [117], 1.2 Å [118], 1.5 Å [120], 1.10 Å [125], 1.43 Å [126] resolutions), a $[Mg(H_2O)_6]^{2+}$ ion, which is bound in the floor of the major groove at one of the G-tracts of the duplex, has been conserved, where the metal ion is bound through outersphere contacts to O6 and N7 atoms of G2 and O6 of G22 in one duplex and simultaneously to the lip through phosphate oxygens of A6 and T7, but not the interior, of the minor groove of an adjacent duplex. This position is almost the same as that in the original DDD structure [167]. The location and interactions of this Mg^{2+} ion explain the dependence of G-tract bending on divalent cations, that is, the $[Mg(H_2O)_6]^{2+}$ ion pulls cytosine bases C1 and C21 partially out of the helical stack, exposing π-systems to the hydrated Mg^{2+} cation, generating the "dodecamer bend" [109b]. Cation–π interaction between shell water molecules and cytosine π-systems is also observed in a decamer $[d(CGCGAATTCGCG)]_2$ duplex [169].

Outersphere interactions with bases in the major groove are repeatedly observed in the duplexes with other sequences: a fully hydrated Mg^{2+} ion is bound through hydrogen bonds to O6 and N7 of G3 and O4 of T16 in the $[d(GCGAATTCG)]_2$ duplex (2.05 Å [131] and 0.89 Å [132] resolutions), O6 and N7 of G7 in the base-intercalated (zipper-like) $[d(GCGAAAGC)]_2$ duplex (1.60 Å [141]), and O6 and N7 of G3 in the zipper-like $[d(GCGAAAGCT)]_2$ duplex (2.5 Å [142]). Hydrated Mg^{2+} ions also hold neighboring duplexes together and help to neutralize the local high concentration of negative charges from the phosphate oxygens, which otherwise would tend to repel closely packed duplexes.

A $[Ca(H_2O)_6]^{2+}$ ion in the DDD-like or -modified dodecamer duplexes, $[d(GCGAATTCGCG)]_2$ [144], $[d(GGCGAATTCGCG)]_2$ [144], and $[d(CGCGAA\ ^{F}T^{F}TCGCG)]_2$ [144], is located at a position corresponding to the conserved $[Mg(H_2O)_6]^{2+}$ ion in the DDD duplex mentioned above. In addition, in most of the Ca^{2+} compounds of DNA decamers, a hydrated Ca^{2+} ion is bound, indirectly through its water ligands, in the minor groove of the G·C base pairs in the center of the duplex. Examples include $[d(CCGGTACm^5CGG)]_2$ [113], $[d(CTCTCGAGAG)]_2$ [145], $[d(CCAGTACTGG)]_2$ [146], $[d(CGATATATCG)]_2$ [170], $[d(CCAAGCTTGG)]_2$ [171], $[d(CCAACITTGG)]_2$ [172]. It is argued [146] that the calcium ion and its bound water molecules fit well in the wider G·C minor groove, whereas monovalent ions may be more appropriate for AT tracts. The ultrahigh resolution (0.74 Å) of the $[d(CCAGTACTGG)]_2$ duplex [146] further reveals that the other two Ca^{2+} ions bind in the major groove, one Ca^{2+} directly to the terminal G·C pairs through O6 of G10 and N7 of G9 and the other Ca^{2+} via hydrogen bonds to N7 of A3, N7 of G4, and O6 of G4.

In the divalent transition metal ion complexes, the Mn^{2+} complex of $[d(CGTTAATTAACG)]_2$ [151], Co^{2+}, Ni^{2+}, and Zn^{2+} complexes of $[d(GGCGCC)]_2$ [152], Ni^{2+} complexes of $[d(CGTATATACG)]_2$ [157], $[d(CGTGTACACG)]_2$ [158], and $[d(GAATTCG)]_2$ [159], and the Zn^{2+} complex of $[d(CGCAATTGCG)]_2$ [160], the metal ion is bound directly to either *terminal* or *extra-helical* but *not to an interior* guanine base through N7.

As with A/Z-DNA fragments, interactions of $[Co(NH_3)_6]^{3+}$ ions with B-DNA fragments are always of the outersphere type. Cobalt hexammine ions are bound to

guanine bases in the major groove through N7 and/or O6 sites in all the complexes involving $[Co(NH_3)_6]^{3+}$ ions [142,143,154–156,200]. In addition, $[Co(NH_3)_6]^{3+}$ ions support the crystal lattice structure by forming bridges between neighboring duplexes. The Rh^{3+} compounds with planar polycyclic aromatic ligands are unique as sequence-specific metallointercalators or metalloinsertors. The Δ-α-[Rh{(*R*,*R*)-Me_2trien}phi]$^{3+}$ compound is intercalated into the major groove between G4 and C5 of [{G(dIU)TGCAAC}]$_2$ [162]. The Δ-α-[$Rh(bpy)_2$(chrysi)]$^{3+}$ compound is intercalated into the major groove and/or inserted into the minor groove at (a) mismatched site(s) by replacing (a) mismatched base pair(s): in [(CGGAAATTCCCG)]$_2$ [163], one of the two Rh^{3+} molecules is intercalated between A6 and T7 and the other is inserted at the A4·C21 mismatched site; in [(CGGAAATTACCG)]$_2$ ($P3_121$ form [164]), two Rh^{3+} molecules are inserted at the A4·A21 and A9·A16 mismatched sites; and in [(CGGAAATTACCG)]$_2$ ($P4_32_12$ form [164]), one of the three Rh^{3+} molecules is intercalated between A6 and T7 and the other two are inserted at two A·A mismatched sites. Metalloinsertors, unlike metallointercalators which unwind the duplex, perturb the DNA only minimally beyond the local binding site. Trinuclear platinum-polyamine compounds, TriplatinNC [114] and TriplatinNC-A [115], form outersphere coordination complexes with the DDD duplex, through amine ligands with phosphate groups of the duplex backbone as phosphate clamps. The *cis*-$[Pt(NH_3)_2]^{2+}$ complex of the decamer d(CCTCGCTCTC)·d(GAGAGCGAGG) duplex [165] is of special interest as a model for *inter*strand cross-links, where a cisplatin cross-links the two guanine bases G5 and G15 through their N7 sites. On the other hand, in the *cis*-$[Pt(NH_3)_2(py)]^{2+}$ complex of the dodecamer d(CCTCTCGTCTCC)·d(GGAGACGAGAGG) duplex [166], a planar Pt^{2+}-diamine compound with an additional bulky py ligand binds monofunctionally to N7 of guanine G7 in the major groove of one strand.

In summary, eight major metal binding patterns are observed with crystalline B-DNA. Two are the same as noted above in the summary for A-DNA: (i) *phosphate* bonding and (ii) *outersphere* bonding in the *majo*r groove. (iii) Alkali metal ions bind *directly* to *base* moieties in the *minor* groove of the AT-tract by replacing water molecules consisting of the spine of hydration. Divalent alkaline earth ions, involving Mg^{2+} ion [134] and Ca^{2+} ion in particular [113,145,146,170–172], make *outersphere* contacts with *bases* in the *wide minor* groove. When the minor groove is *narrow* as in the DDD structure, Mg^{2+} ion binds through hydrogen bonds to the *lip* of the minor groove [109,117,118,120,125,126]. (iv) Polyhedral metal cations, involving Na^+ [114], Mg^{2+} [200], Ca^{2+} [146], and transition metal ions in particular (Mn^{2+} [151], Co^{2+} [152,153], Ni^{2+} [152,157–159], and Zn^{2+} [160]), bind *directly* to *guanine* bases in the *major* groove through N7 and/or O6 sites. These guanine bases are almost always located in either *terminal*, or *extra-helical*, or *flipped-out* positions or they are involved in the mismatched sheared base pairs. In other words, *polyhedral* metal ions exhibit a difficulty to bind *directly* to the *interior* guanines in the *major* groove in the *standard* B-DNA duplex, except for Na^+ ion that binds to O6 of the third guanine from the 3'-terminal end [114]. (v) *Square planar* metal compounds such as *cis*-$[Pt(NH_3)_2]^{2+}$ and *cis*-$[Pt(NH_3)_2(py)]^{2+}$ bind *directly* to *base* moieties in the *major* groove: the former compound *interstrand* cross-links two

adjacent *guanine* bases through N7 sites [165] and the latter monofunctionally binds to N7 of *guanine* [166]. (vi) Non-covalent polyamine-type bonding occurs in the *minor groove* by polynuclear platinum-amine compounds as phosphate clamps [114,115]. (vii) Non-covalent π–π interaction by metallointercalators and metalloinsertors are observed [162–164]. (viii) Cation–π interaction between DNA base-π systems and hydrated metal cations takes place [109b].

3.2.3 Z-DNA Fragments

Z-DNA adopts a left-handed double-helical structure with a convex surface, instead of the major groove in B-DNA, exposing the outside of base pairs directly to the solvent, and with the minor groove which is very narrow and deep, extending almost to the helical axis. Quite a number of crystal structures of metal salts/complexes of Z-DNA fragments (a total number of 26 [1] before 1994 and 22 after 1994) are reported due to their excellent diffraction properties, among which those of a d(CGCGCG) hexamer amount to half the number reported. In the crystal structures of $[d(CGCGCG)]_2$ duplexes, hydrated Na^+ is bound through outersphere contacts via hydrogen bonds to phosphate oxygens [175,176] and N4 of cytosine [190], and also through innersphere contacts directly to N7 of guanine [190] in addition to phosphate oxygens [190–192]. For hydrated Mg^{2+} ion, outersphere contacts occur usually with N7 and/or O6 of guanine [174-176,178,179,192,193] and rarely with N2 of guanine [174] and N4 of cytosine [175,180] in addition to phosphate oxygens. Direct interactions also occur with N7 of the 3'-terminal guanine [176,178,179,193,194] in addition to phosphate oxygens. On the other hand, hydrated Co^{2+} [195] and Cu^{2+} [196] ions bind exclusively through innersphere contacts to N7 of the *inner* guanines in addition to N7 of the 3'-terminal guanine. For Rb^+ ion [177], extensive direct interactions with base and sugar moieties are observed.

$[Co(NH_3)_6]^{3+}$ [183–185,197,199] and $[Ru(NH_3)_6]^{3+}$ [198] ions form outersphere complexes with $[d(CGCGCG)]_2$ [197,198], $[d(GCGCGCG)]_2$ [184], $[d(Gm^5CGCGCG)]_2$ [183], $[d(TGCGCA)]_2$ [185], and $[d(CGTACGTACG)]_2$ [199] through hydrogen bonds usually with N7 and O6 of guanines in addition to phosphate oxygens. As rare cases, in the complexes of the d(CGCGCA)·d(TGCGCG) duplex, $[Co(NH_3)_6]^{3+}$ [187] and $[Ru(NH_3)_6]^{3+}$ [188] ions are bound through hydrogen bonds to N7 and the *deprotonated N6* of *adenine* in addition to guanines at N7 and O6. A square planar $[Pt(NH_3)(L_1)(L_2)]^{2+}$ (L_1 and L_2 are NH_3 or H_2O) ion binds to $[d(CGTA'CG)]_2$ (A' = 2-amino adenosine) through N7 of G6 [189].

In summary, three major metal binding patterns are observed with crystalline Z-DNA. Two are the same as noted above in the summary for A/B-DNA: (i) *phosphate* bonding and (ii) *outersphere* bonding in the *major* groove, and (iii) hydrated metal ions are much less accessible in the *minor* groove, whereas they do not suffer from any steric constraint on the *convex* surface in attacking the functional groups of *bases*; metal binding sites are always *guanine* bases at N7 (or N7/O6 chelation for Ba^{2+} [199]).

3.3 *Factors Affecting Metal Binding to Oligonucleotide Duplexes*

From X-ray diffraction studies of oligonucleotide duplexes, certain general observations emerge: (i) Metal cations, usually hydrated, preferentially bind to phosphate groups, either directly or mediated by water ligands by forming hydrogen bonds, and serve to partially neutralize the negative charges of phosphates. (ii) In the minor groove in A·T-rich regions in B-DNA, monovalent cations are directly bound to the base moieties by replacing water molecules that constitute a spine of hydration. This is consistent with a high negative electrostatic potential in the minor groove in A·T-rich regions of B-DNA [201]. (iii) In the major groove, metal ions bind preferentially to the guanine base, either directly at the N7 site or through hydrogen bonds with the O6 and N7 sites.

For the *standard* duplexes of A-RNA and A/B-DNA but not for Z-DNA, where usually the metal ion is *polyhedral* with hydrogen-bonding donor ligands like water, the metal ion binds through outersphere contacts (but not innersphere contacts) to the interior guanines in the major groove, as a rule. Only when the guanine base is located at the terminal position or at (or near) extra-helical, flipped-out, or mismatched-base pair positions in *non-standard* duplexes, such a polyhedral metal ion binds directly to the base at the N7 site, often accompanying the formation of an intramolecular interligand hydrogen bond between the water ligand and the O6 substituent. Also as an exception, only when the metal ion is extremely dehydrated, it can bind directly to the interior guanine base in the major groove in the *standard* duplex. What factors correspond to these observations?

First, why does the metal ion bind preferably to the guanine base among the four bases which have some potential metal binding sites at the mononucleotide level? Among the most preferred metal binding sites on the bases, N7 of adenine and guanine and N3 of cytosine, as described in Section 2.3, N3 of cytosine can be ruled out since it takes part in the Watson–Crick base-pair in the RNA/DNA duplexes. The preferred metal bonding to guanine over adenine for the polyhedral metal ion with hydrogen-bonding donor ligands like water can be explained, at least in part, in terms of interligand interaction, as noted in Section 2.4: When such a metal ion binds to N7 of guanine, an intramolecular hydrogen bond is formed between a water ligand as a hydrogen bonding donor and the O6 keto group as an acceptor, stabilizing the structure, whereas when it binds to N7 of adenine, a similar hydrogen bond is formed between the N6-H_2 amino group as a donor and a H_2O ligand as an acceptor but the hydrogen bond is weak compared to the former, as mentioned in Section 2.2. Indeed, a number of crystal structures of [M(5'-purine mononucleotide)$(H_2O)_5$]-type complexes with M–N7 bonding are reported for 6-oxopurine nucleotides but only the Ni^{2+} complex [43] for adenosine nucleotides. Rather it is conceivable that the close contact between the amino NH_2 group and the H_2O ligand in the crowded coordination environment would cause steric hindrance, making it less favorable for the polyhedral metal ion to bind the N7 site of the adenine base.

Second, why does the polyhedral metal ion bind to guanine through outersphere contacts (but no innersphere contacts) to interior guanines in the major groove in A/B-type duplexes but through innersphere contacts in a Z-type duplex? The double-helical ordered structure of oligonucleotides with the distinct shape of the major and minor grooves answers this question. From a structural consideration, Gao et al. originally [195] and others [1,22,69f,157,159] have suggested that for B-DNA with even the widest major groove among right-handed duplexes, there would be steric clashes of the octahedrally hydrated metals bound to N7 of guanine with the previous (the 5'-side) base in the sequence. Square-planar metals like platinum-diamine complexes [105,106,165,166] or highly dehydrated metals, though quite rare only for Ba^{2+} ion [102,103], are able to escape from these clashes. On the other hand, for Z-DNA with a convex surface exposing the edges of base pairs directly to the solvent and a narrow and deep minor groove, polyhedral metal ions have much less access to the minor groove, whereas the lack of steric restriction on the convex surface makes it possible for any metal ion to attack the functional groups of bases, here again exclusively N7 of guanine.

Third, when the polyhedral metal ion makes outersphere contacts through hydrogen bonds with functional groups of bases in the major groove, why does the metal ion bind predominantly to guanine bases? Two factors at least may be responsible for this: a molecular electrostatic potential and a hydrogen bonding potential. (1) Molecular mechanics simulation demonstrated a strong electronegative potential at the G·C region in the major groove for A/B-DNA [201]. (2) Metal cations are bound to nucleic acids usually in a hydrated state and work as hydrogen bonding donors in addition to point charges [22]. The candidates of their targets in the major groove as hydrogen bonding acceptors are N7 of adenine, O6 and N7 of guanine, and O4 of thymine. We suggest that, in order to stabilize the overall structure formed, the *bi*-functional hydrogen bonding to guanine by forming an [–O6(*n*th G)···H_2O–M–OH_2···N7(*n*th G)–] macrochelate would be preferable to the *mono*functional hydrogen bonding to adenine or thymine. In practice, under a given environment, the hydrated metal ion works to form as many hydrogen bonds as possible involving such a hydrogen-bonding macrochelate.

However, despite the above arguments, at present it may be safe to say that whether a polyhedral metal ion could bind directly to the interior-base moiety in the major groove of right-handed duplexes is still an open question from a consideration of the plasticity [20] of the duplex structures. Indeed, some hydrated Mn^{2+} ions bind to internal guanines at N7 sites in the DNA duplex which winds around the nucleosome core particle [202], as a result of a combination of favorable facts [151]: the local conformation of DNA, the longer Mn^{2+}–N7 distance and its variable hydration sphere. In any case, even if it should occur in the standard right-handed DNA duplexes, DNA structures may suffer local deformation to a considerable extent.

In summary, metal binding to oligonucleotide duplexes depends on numerous factors: the shape (the width and the depth) of the major and the minor grooves; electronegative potentials, as modulated by base sequence and conformation; cation size and geometry and dehydration energetics; the concentrations of the competing multivalent cations, polyamines, and others; and the crystal packing environment in the solid state.

4 Concluding Remarks and Future Directions

During the past five decades single-crystal X-ray diffraction studies have provided a wealth of well-defined information on metal interactions with nucleic acid constituents such as bases, nucleosides, and nucleotides, involving oligonucleotides whose crystal structures were analyzed using high-resolution data at the atomic level. Now certain general aspects emerge, as described here. However, known structures of metal complexes of nucleotides (listed in Tables 1 and 2 in Ref. 16 and Tables 1 and 2 in this chapter) and oligonucleotides (in Table 3 in Ref. 1 and Tables 5 and 6 in this chapter) are still limited. In order to refine these general aspects and moreover to add new general observations, further X-ray diffraction studies, particularly those of oligonucleotide fragments with more variable sequences [7a] supported by further technological advances to collect ultra-high resolution data, are needed.

In addition, it should be noted that the localization of metal ions in the crystals does not mean that in solution they are statically bound at specific sites, and rather, those bound even at the most favorable sites on DNA still readily exchange with solution [20], and much remains to be learnt about dynamical behaviors of metal interactions with oligonucleotide duplexes [22].

Finally, since metal ions occur, in practice, in even more complex systems containing proteins and other compounds (though we do not deal with these here), further X-ray studies of such systems are needed to enhance our understanding of the nature of metal complexes and their biological action. It is likely that even more "unusual DNA structures [203,204]" and "unusual RNA folds [204]" and their interactions with metal ions [204] will be discovered. This is a field to be still continuously challenged.

Abbreviations

A	adenine base
A'	2-amino adenosine
AN_2	two amines or a diamine
ADP	adenosine diphosphate
ADPH	monoprotonated adenosine diphosphate
ADPH2	diprotonated adenosine diphosphate
AMP	adenosine monophosphate
AMPH	monoprotonated adenosine monophosphate
ATP	adenosine triphosphate
ATPH	monoprotonated adenosine triphosphate
ATPH2	diprotonated adenosine triphosphate
α-D-Glc-1P	α-D-glucopyranose-1-phosphate
bpy	2,2'-bipyridine
bpm	2,2'-bipyrimidine

^{Br}U	5-bromo-2'-deoxyuridine
C	cytosine base
chrysi	5,6-chrysenequinone diimine
CMP	cytidine monophosphate
Cy	cyclohexyl
d	deoxyribose
D	deuteron
DACH	diaminocyclohexane
dAMP	deoxyadenosine monophosphate
dAMPH	monoprotonated deoxyadenosine monophosphate
DDD	Dickerson–Drew dodecamer
dGMP	deoxyguanosine monophosphate
dIU	5-iodo-2'-deoxyuridine
dpa	2,2'-dipyridylamine
dUMP	deoxyuridine monophosphate
en	ethylenediamine
^{F}C	5-fluoro-cytidine
^{F}T	2'-deoxy-2'-fluoro-arabino-furanosyl thymidine
G	guanine base
gg	*gauche-gauche*
GMP	guanosine monophosphate
GMPH	monoprotonated guanosine monophosphate
I	inosine
IMP	inosine monophosphate
IMPH	monoprotonated inosine monophosphate
IMP–H	mono-deprotonated inosine monophosphate
L	ligand
LNA	locked nucleic acid: an oligonucleotide that contains one or more 2'-*O*,4'-*C*-methylene-β-D-ribofuranosyl monomer(s)
M	metal ion
m^5C	5-methyl-cytidine
Me	methyl
Me_4dach	*N*,*N*,*N'*,*N'*-tetramethyl-1,2-diaminocyclohexane
Me_4dae	*N*,*N*,*N'*,*N'*-tetramethyl-1,2-diaminoethane
MeRNA	data base of metal ion binding sites in RNA
Me_2trien	dimethyltriethylenetetraamine
N	general term for nucleobases
NDB	Nucleic Acid Data Bank
NMP	general term for monophosphate nucleotides
nta	nitrilotriacetate
NTP	general term for triphosphate nucleotides
O(P)	oxygen atom of the phosphate group
O(W)	water oxygen atom
P	phosphate
P	phenyl ribonucleotide

phen	1,10-phenanthroline
phi	9,10-phenanthrenequinone diimine
py	pyridine
r	ribose
T	thymine base
terpy	2,2',2"-terpyridine
TMP	thymidine monophosphate
TriplatinNC	$[\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_3^+)\}_2\text{-}\mu\text{-}\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_2)_2\}]^{8+}$
TriplatinNC-A	$[\{trans\text{-}Pt(NH_3)_3\}_2\text{-}\mu\text{-}\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_2)_2\}]^{6+}$
tRNA	transfer ribonucleic acid
U	uracil base
UMP	uridine monophosphate
UMP–H	mono-deprotonated uridine monophosphate
W	water ligand

References

1. K. Aoki, in *Comprehensive Supramolecuar Chemistry*, Ed J.-M. Lehn, Vol. 5, Pergamon Press, Oxford, 1996, pp. 249–294.
2. *Metal Ions in Biological Systems*, Vol. 32, Eds A. Sigel, H. Sigel, Dekker, New York, 1996, pp. 1–814.
3. *Metal Ions in Life Sciences*, Vol. 9, Eds A. Sigel, H. Sigel, R. K. O. Sigel, Royal Society of Chemistry, Cambridge, UK, 2011, pp. 1–391.
4. (a) B. Lippert, *Prog. Inorg. Chem.* **2005**, *54*, 385–447. (b) E. Freisinger, R. K. O. Sigel, *Coord. Chem. Rev.* **2007**, *251*, 1834–1851. (c) A. Terrón, J. J. Fiol, A. García-Raso, M. Barceló-Oliver, V. Moreno, *Coord. Chem. Rev.* **2007**, *251*, 1973–1986. (d) F. R. Keene, J. A. Smith, J. G. Collins, *Coord. Chem. Rev.* **2009**, *253*, 2021–2035.
5. *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, Royal Society of Chemistry Publishing, Cambridge, UK, 2009.
6. *Metal Complex-DNA Interactions*, Eds N. Hadjiliadis, E. Sletten, John Wiley & Sons, Chichester, UK, 2009.
7. (a) M. Egli, P. S. Pallan, *Curr. Opin. Struct. Biol.* **2010**, *20*, 262–275. (b) J. Müller, *Metallomics* **2010**, *2*, 318–327.
8. D. J. Hodgson, *Prog. Inorg. Chem*. **1977**, *23*, 211–254.
9. V. Swaminathan, M. Sundaralingam, *CRC Crit. Rev. Biochem*. **1979**, *6*, 245–336.
10. L. G. Marzilli, T. J. Kistenmacher, G. L. Eichhorn, in *Metal Ions in Biology*, Ed T. G. Spiro, Vol. 1, John Wiley and Sons, New York, 1980, pp. 180–250.
11. K. Aoki, *J. Cryst. Soc. Jpn*. **1981**, *23*, 309–327.
12. W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984, pp. 201–219.
13. B. Lippert, *Coord. Chem. Rev*. **2000**, *200–202*, 487–516.
14. R. W. Gellert, R. Bau, *Met. Ions Biol. Syst*. **1979**, *8*, 1–55.
15. A. Terrón, *Comm. Inorg. Chem*. **1993**, *14*, 63–88.
16. K. Aoki, *Met. Ions Biol. Syst*. **1996**, *32*, 91–134.
17. R. Cini, *Comm. Inorg. Chem*. **1992**, *13*, 1–34.
18. K. Aoki, I. Fujisawa, in *Nucleoside Triphosphates and Their Analogs*, Ed M. Vaghefi, Taylor and Francis, Boca Raton, London, New York, 2005, pp. 115–132.

19. K. Aoki, in *Landolt-Börnstein: Nukleinsäuren; Teilband b*; *Kristallographische und strukturelle Daten II*, Ed W. Saenger, Springer-Verlag, Berlin, 1989, pp.171–246.
20. N. V. Hud, M. Polak, *Curr. Opin. Struct. Biol*. **2001**, *11*, 293–301.
21. M. Egli, *Chem. Biol*. **2002**, *9*, 277–286.
22. J. A. Subirana, M. Soler-López, *Ann. Rev. Biophys. Biomol. Struct*. **2003**, *32*, 27–45.
23. M. M. Teeter, G. J. Quigley, A. Rich, in *Metal Ions in Genetic Information Transfer*, Eds G. L. Eichhorn, L. G. Marzilli, Elsevier, New York, 1981, pp. 233–272.
24. (a) C. C. Correll, B. Freeborn, P. B. Moore, T. A. Steitz, *Cell* **1997**, *91*, 705–712. (b) M. Selmer, C. M. Dunham, F. V. Murphy IV, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, V. Ramakrishnan, *Science* **2006**, *313*, 1935–1942.
25. R. K. O. Sigel, A. M. Pyle, *Chem. Rev*. **2007**, *107*, 97–113.
26. W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984, pp. 17–24.
27. Y. Sugawara, H. Urabe, K. Kobayashi, Y. Iimura, H. Iwasaki, *Mol. Cryst. Liq. Cryst*. **1996**, *277*, 255–258.
28. U. P. Singh, *J. Mol. Struct*. **2010**, *983*, 73–75.
29. M. Inoue, T. Yamase, *Bull. Chem. Soc. Jpn*. **1996**, *69*, 2863–2868.
30. Z. Szabó, I. Furó, I. Csöregh, *J. Am. Chem. Soc*. **2005**, *127*, 15236–15247.
31. G. S. Padiyar, T. P. Seshadri, J. *J. Biomol. Struct. Dyn*. **1998**, *15*, 803–821.
32. M. Benedetti, G. Tamasi, R. Cini, L. G. Marzilli, G. Natile, *Chem. Eur. J*. **2007**, *13*, 3131–3142.
33. G. Borodi, A. Hernanz, I. Bratu, M. Pop, R. Navarro, *Acta Crystallogr.*, *Sect. E* **2001**, *57*, m514–m516.
34. Y. Sugawara, A. Nakamura, Y. Iimura, K. Kobayashi, H. Urabe, *J. Phys. Chem. B* **2002**, *106*, 10363–10368.
35. S. V. Gonzalez, K. Larsen, W. H. Nelson, E. Sagstuen, and C. H. Görbitz, *Acta Crystallogr.*, *Sect. E* **2005**, *61*, m554–m556.
36. R. Chitra, R. Ranjan-Choudhury, M. Ramanadham, *Appl. Phys. A* **2002**, *74*, S1576–S1578.
37. B. S. Reddy, M. A. Viswamitra, *Acta Crystallogr.*, *Sect. B* **1975**, *31*, 19–26.
38. W. S. Sheldrick, *Acta Crystallogr.*, *Sect. B* **1981**, *37*, 1820–1824.
39. O. Kennard, N. W. Isaacs, W. D. S. Motherwell, J. C. Coppola, D. L. Wampler, A. C. Larson, D. G. Watson, *Proc. Roy. Soc. London*, Ser. A **1971**, *A325*, 401–436.
40. Y. Sugawara, N. Kamiya, H. Iwasaki, T. Ito, Y. Satow, *J. Am. Chem. Soc*. **1991**, *113*, 5440–5445.
41. D. W. Young, P. Tollin, H. R. Wilson, *Acta Crystallogr.*, *Sect. B* **1974**, *30*, 2012–2018.
42. M. A. Viswamitra, T. P. Seshadri, M. L. Post, *Acta Crystallogr.*, *Sect. B* **1980**, *36*, 2019–2024
43. I. Císarová, J. Zachová, *Z. Kristallogr*. **1996**, *211*, 487–488.
44. R. Cini, C. Pifferi, *J. Chem. Soc., Dalton Trans.***1999**, 699–710.
45. M. Kato, T. Tanase, *Inorg. Chem*. **2005**, *44*, 8–10.
46. M. Kato, A. K. Sah, T. Tanase, M. Mikuriya, *Eur. J. Inorg. Chem.* **2006**, 2504–2513.
47. K. J. Barnham, C. J. Bauer, M. I. Djuran, M. A. Mazid, T. Rau, P. J. Sadler, *Inorg. Chem*. **1995**, *34*, 2826–2832.
48. M. I. Djuran, S. U. Milinkovic, A. Habtemariam, S. Parsons, P. J. Sadler, *J. Inorg. Biochem*. **2002**, *88*, 268–273.
49. M. Benedetti, G. Tamasi, R. Cini, G. Natile, *Chem. Eur. J*. **2003**, *9*, 6122–6132.
50. T. Yajima, G. Maccarrone, M. Takani, A. Contino, G. Arena, R. Takamido, M. Hanaki, Y. Funahashi, A. Odani, O. Yamauchi, *Chem. Eur. J*. **2003**, *9*, 3341–3352.
51. T. Rau, R. van Eldik, *Chem. Ber./Recueil* **1977**, *130*, 1551–1555.
52. A. D. Collins, P. D. Meester, D. M. L. Goodgame, A. C. Skapski, *Biochim. Biophys. Acta* **1975**, *402*, 1–6.
53. K. Aoki, *Bull. Chem. Soc. Jpn*. **1975**, *48*, 1260–1271.
54. B. Hedmann, *Acta Crystallogr.*, *Sect. B* **1977**, *33*, 3083–3090.
55. N. S. Begum, H. Manohar, *Polyhedron* **1992**, *11*, 2823–2824.
56. D. M. L. Goodgame, I. Jeeves, C. D. Reynolds, A. C. Skapski, *Nucl. Acids Res*. **1975**, *2*, 1375–1379.

57. W. S. Sheldrick, *Z. Naturforsch. B* **1982**, *37*, 863–871.
58. K. Aoki, *J. Am. Chem. Soc*. **1978**, *100*, 7106–7108.
59. K. Aoki, *J. Chem. Soc., Chem. Commun.*, **1979**, 589–591.
60. P. Orioli, R. Cini, D. Donati, S. Mangani, *J. Am. Chem. Soc*. **1981**, *103*, 4446–4452.
61. R. E. Cramer, P. L. Dahlstrom, *Inorg. Chem*. **1985**, *24*, 3420–3424.
62. R. W. Gellert, B. E. Fisher, R. Bau, *J. Am. Chem. Soc*. **1980**, *102*, 7812–7815.
63. D. Cozak, A. Mardhy, M. J. Olivier, A. L. Beauchamp, *Inorg. Chem*. **1986**, *25*, 2600–2606.
64. A. L. Beauchamp, F. Belenger-Gariepy, A. Mardhy, D. Cozak, *Inorg. Chim. Acta* **1986**, *124*, L23–L24.
65. J. Lorberth, W. Massa, M. E. Essawi, L. Labib, *Angew. Chem. Int. Ed. Engl*. **1988**, *27*, 1160–1161.
66. K. R. Dunbar, J. H. Matonic, V. P. Saharan, C. A. Crawford, G. Christou *J. Am. Chem. Soc*. **1994**, *116*, 2201–2202.
67. M. Coll, X. Solans, M. Font-Altaba, J. A. Subirana, *J. Biomol. Struct. Dyn*. **1987**, *4*, 797–811.
68. L. G. Marzilli, T. J. Kistenmacher, *Acc. Chem. Res*. **1977**, *10*, 146–152.
69. (a) M. A. Salam, K. Aoki, *Inorg. Chim. Acta* **2000**, *311*, 15–24. (b) K. Aoki, M. A. Salam, *Inorg. Chim. Acta* **2002**, *339*, 427–437. (c) M. A. Salam, K. Aoki, *Inorg. Chim. Acta* **2001**, *314*, 71–82. (d) M. S. Rahman, H. Q. Yuan, T. Kikuchi, I. Fujisawa, K. Aoki, *J. Mol. Struct.* **2010**, *966*, 92–101.
70. D. Choquesillo-Lazarte, M. D. P. Brandi-Blanco, I. García-Santos, J. M. González-Pérez, A. Castiñeiras, J. Niclós-Gutiérrez, *Coord. Chem. Rev*. **2008**, *252*, 1241–1256.
71. L. McFail-Isom, C. C. Sines, L. D. Williams, *Curr. Opin. Struct. Biol*. **1999**, *9*, 298–304.
72. H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.-H. Hsieh, A. R. Srinivasan, B. Schneider, *Biophys. J*. **1992**, *63*, 751–759.
73. L. R. Stefan, R. Zhang, A. G. Levitan, D. K. Hendrix, S. E. Brenner, S. R. Holbrook, *Nucl. Acids Res*. **2006**, *34*, D131–D134.
74. (a) A. M. Pyle, *Science* **1993**, *261*, 709–714. (b) D. M. J. Lilley, *FEBS Lett*. **1999**, *452*, 26–30. (c) A. R. Ferré-D'Amaré, J. A. Doudna, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 57–73. (d) R. Hanna, J. A. Doudna, *Curr. Opin. Chem. Biol.* **2000**, *4*, 166–170. (e) A. M. Pyle, *J. Biol. Inorg. Chem.***2002**, *7*, 679–690. (f) V. J. DeRose, *Curr. Opin. Struct. Biol.* **2003**, *13*, 317–324. (g) D. M. J. Lilley, *Curr. Opin. Struct. Biol.* **2005**, *15*, 313–323. (h) D. E. Draper, D. Grilley, A. M. Soto, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 221–243. (i) M. R. Stahley, S. A. Strobel, *Curr. Opin. Struct. Biol.* **2006**, *16*, 319–326. (j) S. A. Strobel, J. C. Cochrane, *Curr. Opin. Chem. Biol.* **2007**, *11*, 636–643. (k) N. Toor, K. S. Keating, A. M. Pyle, *Curr. Opin. Struct. Biol.* **2009**, *19*, 260–266. (l) A. M. Pyle, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 215–232.
75. M. C. Wahl, C. Ban, C. Sekharudu, B. Ramakrishnan, M. Sundaralingam, *Acta Crystallogr., Sect. D* **1996**, *52*, 655–667.
76. Y. Timsit, S. Bombard, *RNA* **2007**, *13*, 2098–2107.
77. D. A. Adamiak, J. Milecki, M. Popenda, R. W. Adamiak, Z. Dauter, W. R. Rypniewski, *Nucl. Acids Res*. **1997**, *25*, 4599–4607.
78. D. A. Adamiak, W. R. Rypniewski, J. Milecki, R. W. Adamiak, *Nucl. Acids Res*. **2001**, *29*, 4144–4153.
79. D. A. Adamiak, J. Milecki, R. W. Adamiak, W. Rypniewski, *New J. Chem*. **2010**, *34*, 903–909.
80. F. Li, P. S. Pallan, M. A. Maier, K. G. Rajeev, S. L. Mathieu, C. Kreutz, Y. Fan, J. Sanghvi, R. Micura, E. Rozners, M. Manoharan, M. Egli, *Nucl. Acids Res*. **2007**, *35*, 6424–6438.
81. A. Eichert, J. P. Fürste, A. Schreiber, M. Perbandt, C. Betzel, V. A. Erdmann, C. Förster, *Biochem. Biophys. Res. Commun*. **2009**, *386*, 368–373.
82. R. J. Carter, K. J. Baeyens, J. SantaLucia, D. H. Turner, S. R. Holbrook, *Nucl. Acids Res*. **1997**, *25*, 4117–4122.
83. W. Rypniewski, M. Vallazza, M. Perbandt, S. Klussmann, L. J. DeLucas, C. Betzel, V. A. Erdmann, *Acta Crystallogr., Sect. D* **2006**, *62*, 659–664.
84. Y. Xiong, M. Sundaralingam, *RNA* **2000**, *6*, 1316–1324.
85. G. Minasov, J. Matulic-Adamic, C. J. Wilds, P. Haeberli, N. Usman, L. Beigelman, M. Egli, *RNA* **2000**, *6*, 1516–1528.

86. K. J. Baeyens, H. L. D. Bondt, A. Pardi, S. R. Holbrook, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12851–12855.
87. W. B. T. Cruse, P. Saludjian, E. Biala, P. Strazewski, T. Prangé, O. Kennard, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4160–4164.
88. A. Eichert, K. Behling, C. Betzel, V. A. Erdmann, J. P. Fürste, C. Förster, *Nucl. Acids Res*. **2010**, *38*, 6729–6736.
89. C. C. Correll, B. Freeborn, P. B. Moore, T. A. Steitz, *Cell* **1997**, *91*, 705–712.
90. E. A. Brown, C. E. Bugg, *Acta Crystallogr., Sect. B* **1980**, *36*, 2597–2604.
91. S. Corbin, R. Lavery, B. Pullman, *Biochim. Bioiphys. Acta* **1982**, *698*, 86–92.
92. Y.-G. Gao, H. Robinson, A. H.-J. Wang, *Eur. J. Biochem*. **1999**, *261*, 413–420.
93. V. Tereshko, C. J. Wilds, G. Minasov, T. P. Prakash, M. A. Maier, A. Howard, Z. Wawrzak, M. Manoharan, M. Egli, *Nucl. Acids Res*. **2001**, *29*, 1208–1215.
94. V. Tereshko, S. Portmann, E. C. Tay, P. Martin, F. Natt, K.-H. Altmann, M. Egli, *Biochemistry* **1998**, *37*, 10626–19634.
95. M. Egli, G. Minasov, V. Tereshko, P. S. Pallan, M. Teplova, G. B. Inamati, E. A. Lesnik, S. R. Owens, B. S. Ross, T. P. Prakash, M. Manoharan, *Biochemistry* **2005**, *44*, 9045–9057.
96. M. Egli, V. Tereshko, M. Teplova, G. Minasov, A. Joachimiak, R. Sanishvili, C. M. Weeks, R. Miller, M. A. Maier, H. An, P. D. Cook, M. Manoharan, *Biopolymers* **1998**, *48*, 234–252.
97. H. Robinson, Y.-G. Gao, R. Sanishvili, A. Joachimiak, A. H.-J. Wang, *Nucl. Acids Res*. **2000**, *28*, 1760–1766.
98. L. G. Fernandez, J. A. Subirana, N. Verdaguer, D. Pyshnyi, L. Campos, L. Malinina, *J. Biomol. Struct. Dyn*. **1997**, *15*, 151–163.
99. C. Mayer-Jung, D. Moras, Y. Timsit, *EMBO J*. **1998**, *9*, 2709–2718.
100. M. Egli, V. Tereshko, G. N. Mushudov, R. Sanishvili, X. Liu, F. D. Lewis, *J. Am. Chem. Soc*. **2003**, *125*, 10842–10849.
101. R. C. Todd, S. J. Lippard, *J. Inorg. Biochem*. **2010**, *104*, 902–908.
102. Y.-G. Gao, H. Robinson, J. H. V. Boom, A. H.-J. Wang, *Biophys. J*. **1995**, *69*, 559–568.
103. B. Spingler, D. A. Whittington, S. J. Lippard, *Inorg. Chem.*, **2001**, *40*, 5596–5602.
104. C. M. Nunn, S. Neidle, *J. Mol. Biol*. **1996**, *256*, 340–351.
105. (a) P. M. Takahara, A.C. Rosenzweig, C. A. Frederick, S.J. Lippard, *Nature* **1995**, *377*, 649–652. (b) P. M. Takahara, C. A. Frederick, S. J. Lippard, *J. Am. Chem. Soc*. **1996**, *118*, 12309–12321.
106. A. P. Silverman, W. Bu, S. M. Cohen, S. J. Lippard, *J. Biol. Chem*. **2002**, *277*, 49743–49749.
107. M. C. Wahl, M. Sundaralingam, *Biopolymers* **1997**, *44*, 45–63.
108. B. F. Eichman, J. M. Vargason, B. H. M. Mooers, P. S. Ho, *Proc. Natl. Sci. USA* **2000**, *97*, 3971–3976.
109. (a) X. Shui, L. McFail-Isom, G. G. Hu, L. D. Williams, *Biochemistry* **1998**, *37*, 8341–8355. (b) L. McFail-Isom, X. Shui, L. D. Williams, *Biochemistry* **1998**, *37*, 17105–17111.
110. (a) M. Soler-López, L. Malinina, J. Liu, T. Huynh-Dinh, J. A. Subirana, *J. Biol. Chem*.**1999**, *274*, 23683–23686. (b) J. Liu, J. A. Subirana, *J. Biol. Chem*. **1999**, *274*, 24749–24752.
111. J. Kondo, S. Umeda, K. Fujita, T. Sunami, A. Takénaka, *J. Synchrotron Rad*. **2004**, *11*, 117–120.
112. J. Kondo, T. Ciengshin, E. C. M. Juan, Y. Sato, K. Mitomi, S. Shimizu, A. Takénaka, *Nucleosides, Nucleotides, Nucleic Acids* **2006**, *25*, 693–704.
113. J. M. Vargason, P. S. Ho, *J. Biol. Chem*. **2002**, *277*, 21041–21049.
114. S. Komeda, T. Moulaei, K. K. Woods, M. Chikuma, N. P. Farrell, L. D. Williams, *J. Am. Chem. Soc*. **2006**, *128*, 16092–16103.
115. S. Komeda, T. Moulaei, M. Chikuma, A. Odani, R. Kipping, N. P. Farrell, L. D. Williams, *Nucl. Acids Res*. **2011**, *39*, 325–336.
116. E. C. M. Juan, J. Kondo, T. Kurihara, T. Ito, Y. Ueno, A. Matsuda, A. Takénaka, *Nucl. Acids Res*. **2007**, *35*, 1969–1977.
117. X. Shui, C. C. Sines, L. McFail-Isom, D. VanDerveer, L. D. Williams, *Biochemistry* **1998**, *37*, 16877–16887.

118. C. C. Sines, L. McFail-Isom, S. B. Howerton, D. VanDerveer, L. D. Williams, *J. Am. Chem. Soc*. **2000**, *122*, 11048–11056.
119. M. Tsunoda, N. Karino, Y. Ueno, A. Matsuda, A. Takénaka, *Acta Crystallogr., Sect. D* **2001**, *57*, 345–348.
120. V. Tereshko, G. Minasov, M. Egli, *J. Am. Chem. Soc*. **1999**, *121*, 3590–3595.
121. K. K. Woods, L. McFail-Isom, C. C. Sines, S. B. Howerton, R. K. Stephens, L. D. Williams, *J. Am. Chem. Soc*. **2000**, *122*, 1546–1547.
122. T. Moulaei, T. Maehigashi, G. T. Lountos, S. Komeda, D. Watkins, M. P. Stone, L. A. Marky, J.-S. Li, B. Gold, L. D. Williams, *Biochemistry* **2005**, *44*, 7458–7468.
123. S. B. Howerton, C. C. Sines, D. VanDerveer, L. D. Williams, *Biochemistry* **2001**, *40*, 10023–10031.
124. C. J. Wilds, Z. Wawrzak, R. Krishnamurthy, A. Eschenmoser, M. Egli, *J. Am. Chem. Soc*. **2002**, *124*, 13716–13721.
125. V. Tereshko, G. Minasov, M. Egli, *J. Am. Chem. Soc*. **1999**, *121*, 470–471.
126. T. K. Chiu, M. Kaczor-Grzeskowiak, R. E. Dickerson, *J. Mol. Biol*. **1999**, *292*, 589–608.
127. I. Berger, V. Tereshko, H. Ikeda, V. E. Marquez, M. Egli, *Nucl. Acids Res*. **1998**, *26*, 2473–2480.
128. S. Portmann, K.-H. Altmann, N. Reynes, M. Egli, *J. Am. Chem. Soc*. **1997**, *119*, 2396–2403.
129. G. A. Leonard, K. E. McAuley-Hecht, N. J. Gibson, T. Brown, W. P. Watson, W. N. Hunter, *Biochemistry* **1994**, *33*, 4755–4761.
130. Y. Timsit, D. Moras, *J. Mol. Biol*. **1995**, *251*, 629–647.
131. (a) L. V. Meervelt, D. Vlieghe, A. Dautant, B. Gallois, G. Précigoux, O. Kennard, *Nature* **1995**, *374*, 742–744. (b) D. Vlieghe, L. V. Meervelt, A. Dautant, B. Gallois, G. Précigoux, O. Kennard, *Acta Crystallogr., Sect. D* **1996**, *52*, 766–775.
132. M. Soler-López, L. Malinina, J. A. Subirana, *J. Biol. Chem*. **2000**, *275*, 23034–23044.
133. D. Vlieghe, J. P. Turkenburg, L. V. Meervelt, *Acta Crystallogr., Sect. D* **1999**, *55*, 1495–1502.
134. T. K. Chiu, R. E. Dickerson, *J. Mol. Biol*. **2000**, *301*, 915–945.
135. M. Egli, V. Tereshko, G. N. Mushudov, R. Sanishvili, X. Liu, F. D. Lewis, *J. Am. Chem. Soc*. **2003**, *125*, 10842–10849.
136. T. Chatake, A. Ono, Y. Ueno, A. Matsuda, A. Takénaka, *J. Mol. Biol*. **1999**, *294*, 1215–1222.
137. T. Chatake, T. Hikima, A. Ono, Y. Ueno, A. Matsuda, A. Takénaka, *J. Mol. Biol*. **1999**, *294*, 1223–1230.
138. M. T. Hossain, T. Chatake, T. Hikima, M. Tsunoda, T. Sunami, Y. Ueno, A. Matsuda, A. Takénaka, *J. Biochem*. **2001**, *130*, 9–12.
139. M. T. Hossain, J. Kondo, Y. Ueno, A. Matsuda, A. Takénaka, *Biophys. Chem*. **2002**, *95*, 69–77.
140. M. T. Hossain, T. Sunami, M. Tsunoda, T. Hikima, T. Chatake, Y. Ueno, A. Matsuda, A. Takénaka, *Nucl. Acids Res*. **2001**, *29*, 3949–3954.
141. Y. Sato, K. Mitomi, T. Sunami, J. Kondo, A. Takénaka, *J. Biochem*. **2006**, *140*, 759–762.
142. T. Sunami, J. Kondo, T. Kobuna, I. Hirao, K. Watanabe, K. Miura, A. Takénaka, *Nucl. Acids Res*. **2002**, *30*, 5253–5260.
143. T. Sunami, J. Kondo, I. Hirao, K. Watanabe, K. Miura, A. Takénaka, *Acta Crystallogr., Sec*t. *D* **2004**, *60*, 90–96.
144. G. Minasov, V. Tereshko, M. Egli, *J. Mol. Biol*. **1999**, *291*, 83–99.
145. D. S. Goodsell, K. Grzeskowiak, R. E. Dickerson, *Biochemistr*y **1995**, *34*, 1022–1029.
146. C. L. Kielkopf, S. Ding, P. Kuhn, D. C. Rees, *J. Mol. Biol*. **2000**, *296*, 787–801.
147. J. Liu, L. Malinina, T. Huynh-Dinh, J. A. Subirana, *FEBS Lett*. **1998**, *438*, 211–214.
148. V. Tereshko, L. Urpí, L. Malinina, T. Huynh-Dinh, J. A. Subirana, *Biochemistry* **1996**, *35*, 11589–11595.
149. F. D. Lewis, X. Liu, Y. Wu, S. E. Miller, M. R. Wasielewski, R. L. Letsinger, R. Sanishvili, A. Joachimiak, V. Tereshko, M. Egli, *J. Am. Chem. Soc*. **1999**, *121*, 9905–9906.

150. N. Valls, M. Richter, J. A. Sbirana, *Acta Crystallogr., Sect. D* **2005**, *61*, 1587–1593.
151. H. Millonig, J. Pous, C. Gouyette, J. A. Subirana, J. L. Campos, *J. Inorg. Biochem*. **2009**, *103*, 876–880.
152. S. L. Labiuk, L. T. J. Delbaere, J. S. Lee, *J. Biol. Inorg. Chem*. **2003**, *8*, 715–720.
153. N. Valls, G. Wright, R. A. Steiner, G. N. Murshudov, J. A. Subirana, *Acta Crystallogr., Sect. D* **2004**, *60*, 680–685.
154. W. Shepard, W. B. T. Cruse, R. Fourme, E. D. L. Fortelle, T. Prangé, *Structure* **1998**, *6*, 849–861.
155. T. Sunami, J. Kondo, I. Hirao, K. Watanabe, K. Miura, A. Takénaka, *Acta Crystallogr., Sect. D* **2004**, *60*, 422–431.
156. J. W. Locasale, A. A. Napoli, S. Chen, H. M. Berman, C. L. Lawson, *J. Mol. Biol*. **2009**, *386*, 1054–1065.
157. N. G. A. Abrescia, L. Malinina, L. G. Fernandez, T. Huynh-Dinh, S. Neidle, J. A. Subirana, *Nucl. Acids Res*. **1999**, *27*, 1593–1599.
158. N. G. A. Abrescia, T. Huynh-Dinh, J. A. Subirana, *J. Biol. Inorg. Chem*. **2002**, *7*, 195–199.
159. N. Valls, I. Usón, C. Gouyette, J. A. Subirana, *J. Am. Chem. Soc*. **2004**, *126*, 7812–7816.
160. M. Soler-López, L. Malinina, V. Tereshko, V. Zarytova, J. A. Subirana, *J. Biol. Inorg. Chem*. **2002**, *7*, 533–538.
161. P. S. Pallan, D. Ittig, A. Héroux, Z. Wawrzak, C. J. Leumann, M. Egli, *Chem. Commun*. **2008**, 883–885.
162. C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton, D. C. Rees, *Nat. Struct. Biol*. **2000**, *7*, 117–121.
163. V. C. Pierre, J. T. Kaiser, J. K. Barton, *Proc. Natl. Sci. USA* **2007**, *104*, 429–434.
164. B. M. Zeglis, V. C. Pierre, J. T. Kaiser, J. K. Barton, *Biochemistry* **2009**, *48*, 4247–4253.
165. F. Coste, J.-M. Malinge, L. Serre, W. Shepard, M. Roth, M. Leng, C. Zelwer, *Nucl. Acids Res*. **1999**, *27*, 1837–1846.
166. K. S. Lovejoy, R. C. Todd, S. Zhang, M. S. McCormick, J. A. D'Aquino, J. T. Reardon, A. Sancar, K. M. Giacomini, S. J. Lippard, *Proc. Natl. Sci. USA* **2008**, *105*, 8902–8907.
167. (a) R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature* **1980**, *287*, 755–758. (b) H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 2179–2183. (c) H. R. Drew, R. E. Dickerson, *J. Mol. Biol*. **1981**, *151*, 535–556.
168. S. Neidle, *Nat. Struct. Biol*. **1998**, *9*, 754–756.
169. K. Grzeskowiak, K. Yanagi, G. G. Privé, R. E. Dickerson, *J. Biol. Chem*. **1991**, *266*, 8861–8883.
170. H. Yuan, J. Quintana, R. E. Dickerson, *Biochemistry* **1992**, *31*, 8009–8021.
171. K. Grzeskowiak, D. S. Goodsell, M. K. Grzeskowiak, D. Cascio, R. E. Dickerson, *Biochemistry* **1993**, *32*, 8923–8931.
172. A. Lipanov, M. L. Kopka, M. K. Grzeskowiak, J. Quintana, R. E. Dickerson, *Biochemistry* **1993**, *32*, 1373–1389.
173. L. Malinina, L. Urpí, X. Salas, T. Huynh-Dinh, J. A. Subirana, *J. Mol. Biol*. **1994**, *243*, 484–493.
174. R. V. Gessner, G. J. Quigley, M. Egli, *J. Mol. Biol*. **1994**, *236*, 1154–1168.
175. H. Ohishi, I. Nakanishi, K. Inubushi, G. V. D. Marel, J. H. V. Boom, A. Rich, A. H.-J. Wang, T. Hakoshima, K. Tomita, *FEBS Lett*. **1996**, *391*, 153–156.
176. H. Ohishi, Y. Tozuka, Z. Da-Yang, T. Ishida, K. Nakatani, *Biochem. Biophys. Res. Commun*. **2007**, *358*, 24–28.
177. X. Dong, *Acta Crystallogr., Sect. D* **2003**, *59*, 1336–1338.
178. H. Ohishi, K. Tsukamoto, Y. Hiyama, N. Maezaki, T. Tanaka, T. Ishida, *Biochem. Biophys. Res. Commun*. **2006**, *348*, 794–798.
179. H. Ohishi, M. Odoko, D.–Y. Zhou, Y. Tozuka, N. Okabe, K. Nakatani, T. Ishida, K. Grzeskowiak, *Biochem. Biophys. Res. Commun*. **2008**, *368*, 382–387.
180. H. Ohishi, N. Terasoma, I. Nakanishi, G. V. D. Marel, J. H. V. Boom, A. Rich, A. H.-J. Wang, T. Hakoshima, K. Tomita, *FEBS Lett*. **1996**, *398*, 291–296.

181. H. Ohishi, K. Suzuki, M. Ohtsuchi, T. Hakoshima, A. Rich, *FEBS Lett*. **2002**, *523*, 29–34.
182. S. Thiyagarajan, P. S. Kumar, S. S. Rajan, N. Gautham, *Acta Crystallogr., Sect. D* **2002**, *58*, 1381–1384.
183. B. H. M. Mooers, B. F. Eichman, P. S. Ho, *J. Mol. Biol*. **1997**, *269*, 796–810.
184. B. Pan, C. Ban, M. C. Wahl, M. Sundaralingam, *Biophys. J*. **1997**, *73*, 1553–1561.
185. A. Harper, J. A. Brannigan, M. Buck, L. Hewitt, R. J. Lewis, M. H. Moore, B. Schneider, *Acta Crystallogr., Sect. D* **1998**, *54*, 1273–1284.
186. G. Schuerman, K. V. Hecke, L. V. Meervelt, *Acta Crystallogr., Sect. D* **2003**, *59*, 1525–1528.
187. S. Thiyagarajan, S. S. Rajan, N. Gautham, *Nucl. Acids Res*. **2004**, *32*, 5945–5953.
188. D. Bharanidharan, S. Thiyagarajan, N. Gautham, *Acta Crystallogr., Sect. F* **2007**, *63*, 1008–1013.
189. G. N. Parkinson, G. M. Arvanitis, L. Lessinger, S. L. Ginell, R. Jones, B. Gaffney, H. M. Berman, *Biochemistry* **1995**, *34*, 15487–15495.
190. B. Chevrier, A. C. Dock, B. Hartmann, M. Leng, D. Moras, M. T. Thuong, E. Westhof, *J. Mol. Biol*. **1986**, *188*, 707–719.
191. D. Bancroft, L. D. Williams, A. Rich, M. Egli, *Biochemistry* **1994**, *33*, 1073–1086.
192. H. Ohishi, S. Kunisawa, G. V. D. Marel, J. H. V. Boom, A. Rich, A. H.-J. Wang, K. Tomita, T. Hakoshima, *FEBS Lett*. **1991**, *284*, 238–244.
193. R.V. Gessner, C. A. Frederick, G. J. Quigley, A. Rich, A. H.-J. Wang, *J. Biol. Chem*. **1989**, *264*, 7921–7935.
194. A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, G. A.V. D. Marel, J. H. V. Boom, A. Rich, *Science* **1981**, *211*, 171–176.
195. Y.-G. Gao, M. Sriram, A. H.-J. Wang, *Nucleic Acids Res*. **1993**, *21*, 4093–4101.
196. T. F. Kagawa, B. H. Geierstanger, A. H.-J. Wang, P. S. Ho, *J. Biol. Chem*. **1991**, *266*, 20175–20184.
197. R.V. Gessner, G. J. Quigley, A. H.-J. Wang, G. A. V. D. Marel, J. H. V. Boom, A. Rich, *Biochemistry* **1985**, *24*, 237–240.
198. P. S. Ho, C. A. Frederick, D. Saal, A. H.-J. Wang, A. Rich, *J. Biomol. Struct. Dyn*. **1987**, *4*, 521–534.
199. R. G. Brennan, E. Westhof, M. Sundralingam, *J. Biomol. Struct. Dyn*. **1986**, *3*, 649–665.
200. Y.-G. Gao, H. Robinson, R. Sanishvili, A. Joachimiak, A. H.-J. Wang, *Biochemistry* 1999, *38*, 16452–16460.
201. R. Lavery, A. Pullman, B. Pullman, *Theoret. Chim. Acta* **1982**, *62*, 93–106.
202. C. A. Davey, T. J. Richmond, *Proc. Natl. Sci. USA* **2002**, *99*, 11169–11174.
203. Unusual DNA Structures, Eds R. D. Wells, S. C. Harvey, Springer, New York, **1988**, 1–306.
204. F. R. Keene, J. A. Smith, J. G. Collins, *Coord. Chem. Rev*. **2009**, *253*, 2021–2035.

Chapter 3
Metal Ion-Promoted Conformational Changes of Oligonucleotides

Bernhard Spingler

Contents

Abstract The review will discuss the influence of metal ions on conformational changes of oligonucleotides. First, a short definition of the torsion angles is given, followed by a concise yet critical overview of the commonly applied experimental techniques. Finally, the possible role of metals upon the following conformational changes of oligonucleotides is discussed: (i) the denaturation of double-strands, (ii) the transition from B- to A-DNA, (iii) the transition from right- to left-handed DNA and RNA, (iv) the condensation, (v) and other conformational changes. We conclude with a summary and outlook.

Keywords circular dichroism • conformational change • crystal structure • DNA • metal • RNA

B. Spingler (✉)
Institute of Inorganic Chemistry, University of Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland
e-mail: spingler@aci.uzh.ch

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*,
Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_3,

1 Introduction

Due to the flexible nature of the oligonucleotides [1], they easily undergo conformational changes. This was already realized in the early days of the structural determination of DNA. DNA fibers equilibrated against different levels of humidity gave raise to different X-ray diffraction patterns [2,3].

This chapter is organized as follows: First, some experimental techniques that are suitable to follow conformational changes of oligonucleotides are shortly discussed. Then, five sections describe the possible role of metals upon oligonucleotides regarding the denaturation of double-strands, the transition from B- to A-DNA, the transition from right- to left-handed helices, the condensation, and other conformational changes. Each section starts with a short general introduction about the process that will be discussed.

In this chapter, we will not discuss the role of metal ions on the formation of G-quadruplexes (see instead Chapter 4) and DNA-protein ternary complexes (see instead Chapter 5), nor the formation of discrete metal assemblies in artificial nucleic acids and of metal mediated base pairs (see instead Chapter 11). Nevertheless, all these processes may lead to conformational changes of oligonucleotides.

Furthermore, we will mainly discuss conformational changes of chemically unmodified oligonucleotides. Given the available space in this review, we had to restrict ourselves to reports where the presence of a metal ion was directly linked to a conformational change of an oligonucleotide composed of naturally occurring nucleotides. Still, the observation of a conformational change within an oligonucleotide by various experimental techniques like circular dichroism [4] or Raman spectroscopy [5,6] frequently means that the actual geometric change within the nucleic acids is unknown. In addition, there have been many studies published that claim a certain conformational transformation based on some spectroscopic change. In our view, some of these claims are not supported by the experimental facts.

Obviously, the clearest method to determine conformational changes is to compare two structures whose geometries have been experimentally determined. Unfortunately, this is only possible for a tiny fraction of all interesting oligonucleotide structures. In addition, when the geometries of different X-ray structures of oligonucleotides with the same sequence in the same general conformation are compared, the helical parameters can vary dramatically [7]. This is also true for the most basic conformational parameter possible, the torsion angle, as shown exemplarily by Lavery and Zakrzewska [8]. A-, B-, and Z-DNA (in case of the purine-pyrimidine step) are all known to exist in two different main conformations [9]. Schneider and coworkers systematically analyzed 447 selected DNA crystal structures with a resolution equal or better than 1.9 Å which included 7739 dinucleotide steps [10]. Their analysis led to the following summary of average torsion values as shown in Figure 1 and Table 1. A similar analysis was done for RNA structures [11].

Furthermore, the interaction of free metal ions with oligonucleotides is sometimes studied in the presence of buffers that are known to form stable metal complexes such as tris(hydroxymethyl)aminomethane (tris), phosphate, or citrate. Preferentially, if a defined pH is wanted, one should use buffers that have weakly coordinating properties to metal ions like cacodylate (see for example the following references:

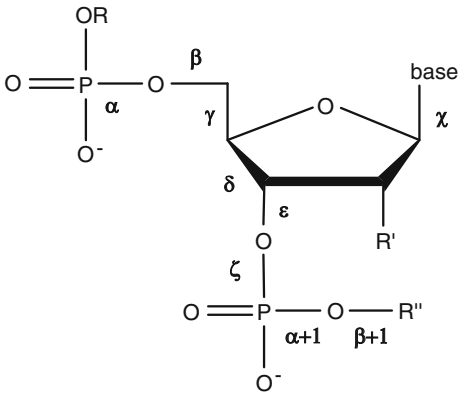


Figure 1 Selected torsion angles of a polynucleotide.

Table 1 Main DNA conformational classes and their selected torsion angles (Figure 1). Data collected from [10].

	α	β	γ	δ	ε	ζ	χ
AI	295	173	54	82	206	285	201
AII	146	192	183	85	197	289	203
BI	299	179	48	133	182	263	250
BII	293	142	46	143	251	168	278
Z, Y-R step	66[a]	186[a]	54	147	264	76	205
ZI, R-Y step	210[a]	233[b]	177	96	242	292	63
ZII, R-Y step	169[a]	162[b]	179	95	187	63	58

[a] α+1 step

[b] β+1 step

[12-15], however, keep the toxicity of cacodylate in mind [16]). The other option is to correct the influence of metal-buffer complexes, if this is possible [17]. A third option is not to use any buffer at all and record the pH instead [18,19].

The possible types of interactions of metal ions with oligonucleotides have been reviewed before [20]. In 1993, Eichhorn has written a short essay about metal-induced conformational changes in general [21].

1.1 Commonly Used Experimental Techniques

UV-vis spectroscopy is a simple, yet powerful as well as sensitive method that allows the researcher to follow the denaturation of oligonucleotides and even the transition of right- to left-handed oligonucleotide helices.

Circular dichroism (CD) is most likely one of the most powerful techniques to detect conformational changes in solutions if the technique is carefully applied. Kypr et al. have written an excellent review how to follow conformational changes of DNA by circular dichroism [22].

Vibrational circular dichroism (VCD) applies the principles of circular dichroism in the infrared region [23]. VCD is a rather insensitive technique. Therefore very high concentrations of oligonucleotides, up to 0.07 M calculated in terms of

phosphates, are needed [24]. Nevertheless, VCD can provide other and more detailed information about the oligonucleotide structure compared with the "normal" infrared spectroscopy.

X-ray and NMR structure determinations are finally crucial to correlate an UV-vis, IR, Raman, CD, or VCD spectrum with a certain geometry. Normally, it is easier to precisely locate metal ions by X-ray structure analysis than with NMR. Nevertheless, even in crystal structures the identity or presence of metal ions can be a fiercely debated topic [25,26].

2 Denaturation of Double-Stranded Oligonucleotides

Double-stranded oligonucleotides can be melted, that is, the two strands are separated in this new state which is also called denatured [27]. This process can be reversed, if the cause of the denaturation is removed. The longer the DNA, the longer it takes to form the duplex state again. In the case of genomic DNA, denaturation can be an irreversible process. The melting can be easily followed by a measurement of the UV absorption at 260 nm, since the molten strands have an UV absorbance that is about 20% higher than those of the original double-strand (corrected for the influence of the temperature) [28]. The hyperchromicity correlates approximately linearly with the content of adenosine-thymidine/uridine base pairs [29]. The determination of melting temperatures (T_m) has been critically discussed [30]. The denaturation of oligonucleotides can be induced by increased temperatures [31], chemical denaturants [32], protonation [33,34], and deprotonation [35].

Owczarzy and coworkers have very carefully studied the factors that influence the melting of short deoxyoligonucleotides (10- to 30mers) in the presence of either 0.069–1.02 M sodium ions [36] or tris, deoxynucleotide-trisphosphates, 1–10 mM magnesium(II), and 50–200 mM potassium [37]. The latter conditions are especially relevant for predicting the optimal temperatures needed for the polymerase chain reaction. In their former work, Owczarzy derived a new relationship to predict the melting temperature. In addition, differential scanning calorimetry (DSC) experiments revealed that the effect of the sodium ion concentration is independent of the oligo concentration but dependent on the guanosine-cytidine content [36]. Furthermore the same group showed that the pH within a range between 6.5 and 8.3 does not have an influence on the melting point. All this work resulted in an online application that can help to predict important properties of oligonucleotides such as T_m, thermodynamics of the denaturation process, and UV spectra [38].

Most transition metals were reported to increase the melting temperature with the exception of copper(II) [13,39]. The influence of various divalent metals (Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd) on the denaturation and aggregation of long genomic DNA and 160 base pairs of nucleosomal fragments has been studied by Raman spectroscopy [40]. The denaturation of calf thymus DNA, which was observed at 90°C, has been followed by IR and VCD spectroscopy [41]. In the IR spectra, bands shifted and became less intense, while the VCD spectrum essentially showed no signals anymore.

The influence of the magnesium ion concentrations on the melting of two RNA strands r[GGUUUGGAGGG] and d[GGCCCUCUAAACC] that are relevant in the context of the P1 duplex of the *Tetrahymena* group I ribozyme were studied by variable temperature ^{1}H NMR [42].

3 B- to A-DNA Transitions

The A-form of pure DNA in aqueous solution can only be accessed in the presence of extremely high concentrations of ethanol (78% percent!) [43]. Additionally, the salt concentrations, e.g., sodium chloride, must not be higher than 1 mM, otherwise precipitation of the DNA occurs. We are not aware of any study that dealt with A-DNA in the presence of transition metals.

4 Right- to Left-Handed Helical Transitions

In the early seventies of the last century, Pohl and Jovin recorded circular dichroic spectra which they interpreted as the "R" and "L" form of poly d(GC) as a function of the used salt concentration (Figure 2) [44].

The transitions from the "R" to the "L" form were determined to have a midpoint at either 2.56 M NaCl, 1.67 M $NaClO_4$, or 0.66 M $MgCl_2$. The enthalpy of the change was almost zero, however the activation energy was determined to be 22 ± 2 kcal/mol [44]. The authors also tested the homopolymer poly d(G)•poly d(C), and poly d(AT) as well as the RNA analogue poly (GC). The RNA poly (GC) did not undergo such a transition, neither did poly d(G)•poly d(C). Whereas the results of the RNA were not withstanding the proof of time, the latter results were afterwards confirmed, since poly d(G) has a strong tendency to form so-called quadruplex DNA [45,46]. On the other hand, poly d(AT) showed minor changes in the CD spectra that were caused by the high salt concentrations. The authors of this publication did speculate about possible biological functions of this new form of DNA, but not about the possible geometrical arrangement.

The questions about a possible geometry were answered when the first crystal structure of a double-stranded DNA was solved in 1979 [47]. It turned out to be a left-handed helix. The repeating unit of the structure was found to be a dinucleotide, rather than a mononucleotide unit as it is the case for most of the right-handed helices. The zigzag of the phosphate backbone led to the term "Z-DNA" which was coined one year later [48]. In the following years, the conditions that favor the formation of Z-DNA or Z-RNA, respectively, have been determined [49]. The most important factor is the presence of an alternating purine-pyrimidine sequence. In fact, this is the ultimate requirement for the formation of Z-DNA or Z-RNA. Within that context, the ease of Z-DNA formation of the different sequences is as follows: $d(GC)_n > d(AC)_n•d(GT)_n >> d(AT)_n$. Furthermore, chemical modifications of the nucleosides modulate the formation of Z-DNA. Increasing the steric demand at the position 8 of guanine by either bromination or methylation

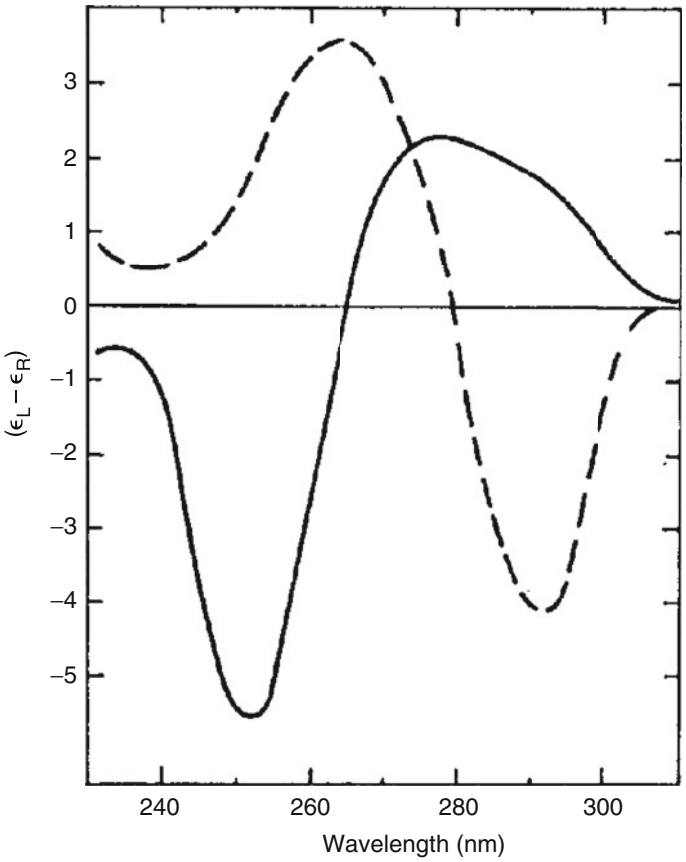


Figure 2 Salt-induced change of the circular dichroism of poly d(GC), solid line: 0.2 M NaCl, pH 7.2 at 25°C, dashed line: same solution after increasing NaCl concentration to 20 %(w/w). Reproduced from [44] by permission of Elsevier; copyright 1972.

[50] helps to form Z-DNA (Figure 3, left: R = Br or CH_3). On the cytosine, bromination and to a lesser extent methylation at the position 5 (in short m5C) facilitate the transition to Z-DNA (Figure 3, right: R = Br, CH_3) [51]. The latter case is very relevant for any physiological role of Z-DNA, since methylation of cytosine at the 5-position has been correlated with gene expression, diverse types of cancer, and schizophrenia [52].

The exploration of left-handed oligonucleotides has been stimulated by multiple findings that Z-DNA has a physiological role. In 1995, it was found that chicken double-stranded RNA adenosine deaminase (later abbreviated as ADAR1) was binding to radioactively labelled Z-DNA [53]. ADAR enzymes hydrolyse adenosine within stretches of double-stranded RNA to give inosine. ADAR1 also contains a Z-DNA binding domain, called Za. Figure 4 shows how the interaction of all these components is thought to be [54]. The polymerase creates a negatively supercoiled

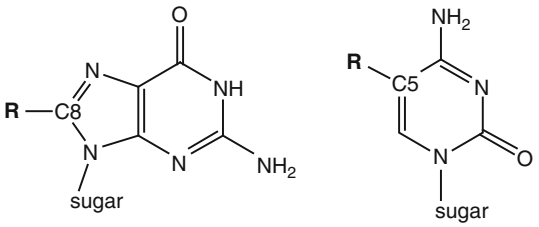


Figure 3 Modified nucleobases, which are relevant for the formation of left-handed DNA.

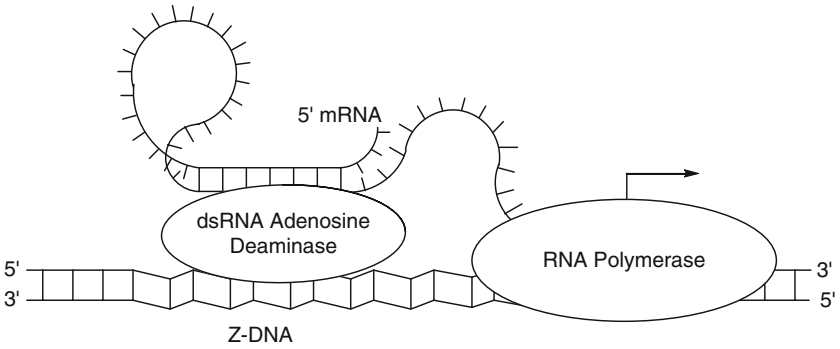


Figure 4 Postulated interaction of ADAR1 within the transcription of DNA. Adapted from [54].

DNA that can release this strain in sections with an alternating purine-pyrimidines sequence by rearranging to the left-handed helix. This temporarily formed Z-DNA is then recognized by domain Za of ADAR1 as a binding site. The actual target of ADAR1, the double-stranded RNA to be edited, is already nearby the Z-DNA. If transcription is stopped, the Z-DNA reverts to its B-form [55]. The possible biochemical role of the Z-DNA has been further clarified in a recent work: Z-DNA forming sequences cause deletions in mammalian cells on a large scale [56]. Spontaneous breaks of the chromosomes can occur in intact DNA and may induce translocation-related diseases. Cellular processes around these Z-DNA forming sequences were very different for bacterial and mammalian cells. In the former, mainly small deletions were found and their frequency was increased by replication. In mammalian cells, however, double-strand breaks were found distributed over a large 400 base pair region. The authors speculated that these large-scale deletions were caused by cellular repair processes.

Recently, an additional role of Z-DNA has been postulated. Dröge and coworkers have cross-linked ADAR1 with the chromatin in A549 cells. Afterwards, they have analyzed the sequences ADAR1 was binding to. They found 186 hotspot regions [57]. Only 2 hotspots were found in promoter regions in contrast to *in silico* predictions [58]. However, 46 of the Z-DNA forming hotspots were located in the centromere regions of the chromosomes. This would imply the involvement of Z-DNA during the processes of mitosis and meiosis.

The group of Tinoco, Jr., showed that Z-RNA could be formed in solutions of high salt concentrations at temperatures above 35°C [59–61]. In the mean time, several

X-ray [62,63] and NMR [64,65] structures of Z-RNA were determined. However, no studies have been published about the induction of Z-RNA by transition metals.

As can be seen in Table 2, the Z-forms of either poly d(GC) or poly d(Gm5C) are inducible by high concentrations of salts, such as sodium chloride or magnesium chloride [66]. Alternatively, Z-DNA can also be formed in the presence of multiple charged cations like inorganic ones such as cobalt(III) hexammine or organic ones such as the tetramine spermine. For example, the midpoint of the B-Z transition of the methylated DNA poly d(Gm5C) occurs at a sub-physiological concentration of 50 mM NaCl in the presence of 2 μM spermine. For a comparison, the spermine concentration in rat liver nuclei was found to be 4 mM [67]. Unfortunately, the authors of reference [66] used varying amounts of DNA and declared its concentrations in only 2 cases. Therefore, the midpoints of these B- to Z-DNA transitions in equivalents of metals respectively polyamines versus poly d(GC) bases cannot be calculated.

Table 2 Required concentrations to induce the midpoint of the right- to left-handed transition of DNA. Ion concentrations are given in mM. Solutions contain 50 mM NaCl, 5 mM tris, pH 8.0. Data collected from [66].

Additive	poly d(GC)	poly d(Gm5C)
NaCl	2500	700
$MgCl_2$	700	0.6
$CaCl_2$	100 (strange CD spectrum)	0.6
$BaCl_2$	40 (strange CD spectrum)	0.6
$[H_3N(CH_2)_2NH_3]Cl_2$	no change seen up to 100 mM	1.0
$[H_3N(CH_2)_3NH_3]Cl_2$	no change seen up to 100 mM	1.0
$[H_3N(CH_2)_4NH_3]Cl_2$	no change seen up to 100 mM	2.0
$[H_3N(CH_2)_5NH_3]Cl_2$	no change seen up to 100 mM	30 (strange CD spectrum)
$[Co(NH_3)_6]Cl_3$	0.02	0.005
Spermidine·3HCl	B-form till 1 mM and then precipitation	0.05
Spermine·4HCl	B-form till 0.05 mM and then precipitation	0.002
ethanol (V/V%)	60%	20%

Not only cobalt(III) hexammine [66] but also simple aqua complexes of transition metals were shown to induce Z-DNA of poly d(GC) (Table 3). It is remarkable that copper(II) does not induce the left-handed form of poly d(GC) [68]. The inhibitory influence of sodium chloride upon the Z-DNA induction promoted by transition metals was described by several authors [69,70]. A systematic study revealed the chloride ion to be the cause for this inhibition [71]. During a titration of poly d(GC) with transition metals, the following processes are thought to happen: First, the metal coordinates to poly d(GC) in the B-form. It was proposed that then the chloride binds to this binary arrangement, thereby inhibiting the B- to Z-DNA transition.

Table 3 Induction of the Z-form of poly d(GC) by divalent transition metals complexes.

Metal salt	Midpoint of transition: equiv. M^{2+} versus DNA bases	Chloride ion concentration [mM]	Ref.
$MnCl_2$	4	10	[69]
$CoCl_2$	9	10	[69]
$NiCl_2$	0.6	0.7	[70]
$Ni(NO_3)_2$	0.5	0	[71]
$NiCl_2$	4	10	[69]
$NiCl_2$	no transition	100	[70]
$ZnCl_2$	2.1	0.1	[72]
$Hg(ClO_4)_2$[a]	0.5	0	[106]

[a]poly d(Gm5C) used

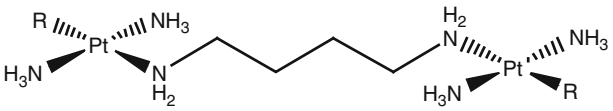
The ability of preformed coordination complexes to induce Z-DNA has been studied by several groups. Table 4 summarizes the reported values for a series of simple nitrogen containing ligands with either copper or zinc [68,72,73]. No clear trend can be observed, neither as a function of metal, nor for the ligands. It remains remarkable that copper polyamine complexes are inducing Z-DNA, while copper(II) or 1,2-ethylenediamine alone do not. A dicobalt complex, which is bridged by imidazole, was shown to be more efficient in partially inducing Z-DNA of a short $d(5mCG)_{15}$ sequence than cobalt hexammine alone, but before the end of the transition was reached, the oligonucleotides started to precipitate [74].

Table 4 Midpoints of the B- to Z-DNA transition of poly d(GC) induced by polyamine metal complexes. Values given as r = metal/nucleobase.

Metal complex	r for M = Cu	r for M = Zn
M(ethylenediamine)Cl_2	0.64 [68]	1.29 [72]
M(dien)Cl_2	0.23 [68]	0.60 [72]
M(trien)Cl_2	not measured	2.99 [68]
M(tren)Cl_2	0.31 [68]	0.07 [68]/0.15 [73]

Early on, it was found that Pt(II)dien was inducing Z-DNA already at a midpoint of about 0.12 (ratio of metal against nucleobase). However cis- and to a lesser extent transplatin were found to cause some changes of the corresponding CD spectra of poly d(GC) without actually inducing the left-handed form [75]. Dinuclear platinum complexes were also shown to induce Z-DNA, such as **1** or even **2** that does not have any open coordination site (Figure 5) [76]. There seems to be a tremendous difference between the dose of platinum complexes needed to generate Z-DNA and the amount of platinum actually found to be attached to DNA. On the one hand, the ratio of added platinum complex to DNA, the so-called r value, ranged between 0.1 and 1. On the other hand, the ratio r_b of platinum complexes like **1** that were found to be *covalently bound* to DNA was only between 0.013 and 0.059. In this work, the conformation of the DNA was not studied after the non-covalently platinum complexes had been removed. Complex **1**, which is able to form a covalent bond to N(7) of guanine, irreversibly induced Z-DNA [77].

This could be shown by adding ethidium bromide, a strong intercalator that normally brings Z-DNA back to the B-form [78,79].



1: R = Cl, **2**: R = NH_3

Figure 5 Dinuclear platinum complexes used for the B- to Z-DNA induction [76].

The ligand 1,4,7,10,13-pentaazacyclohexa-decan-14-16-dione (dione) was introduced by Kimura and coworkers. They used the copper complex of dione derivatives as a model for superoxide dismutase [80,81]. Later, it was shown that nickel complexes of dione like **Ni-F-dione** (Figure 6, left) could induce Z-DNA with a midpoint of the transition of 0.31 (ratio of metal complex against nucleobase) [82]. This work was extended to the corresponding zinc and copper complexes. **Cu-F-dione** and **Cu-F-Me-dione** could promote the B- to Z-DNA transition of poly d(GC), while the **Cu-2F-dione** and **Zn-F-dione** could not [83,84]. This could be explained with the acid-base behavior of the metal complexes (Figure 6, right). Different substituents of the dione ligand influenced the equilibrium between the dicationic form which is able to induce Z-DNA, while the neutral form is not. These studies were continued with pyrazolyl-containing complexes. It was shown that the mono- and the dinuclear nickel complexes of N,N,N',N'-tetrakis-[2-(3,5-dimethylpyrazol-1-yl)-ethyl]-1,2-propylenediamine could induce Z-DNA. In the case of the mononickel complex the authors postulated that a proton could assume the role of the metal [107].

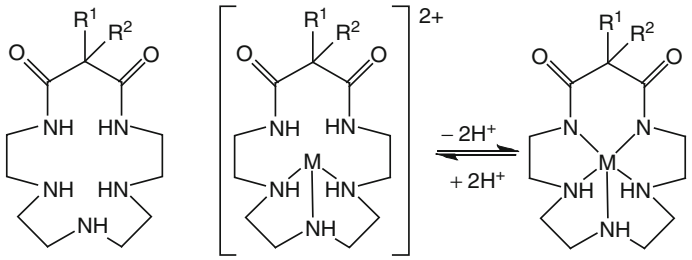


Figure 6 Left: Substituted 1,4,7,10,13-pentaazacyclohexa-decan-14-16-dione (dione) ligands: R^1 = F: **F-dione**, R^1 = F, R^2 = CH_3: **F-Me-dione**, R^1 = R^2 = F: **2F-dione**. Right: Acid-base behavior of metal-dione complexes.

The binding of copper(II) to Z-DNA was explored with the help of crystallography. Kagawa and coworkers studied *preformed* crystals of Z-DNA that were *soaked* with a solution of $CuCl_2$ [85]. The resulting structure was refined at the rather high resolution of 1.2 Å. Copper ions were found to coordinate to the N(7) atoms of *all* guanosine bases of $d(CG)_3$. However, the positions of all metal cations were only partially occupied with reported occupancies ranging from 6.7 to 62.6%.

Furthermore, most of the copper ions in this crystal structure have abnormal high thermal vibrational factors. These so-called B factors indicate that the real occupancies of these copper ions are even lower than reported. In cases of pronounced partial occupancies of these copper atoms, the complete coordination sphere of the metal centers could not always be determined. The reported metal to N(7) distances ranged between 2.17 and 2.57 Å. The latter distance stems from the copper atom that has the lowest occupancy.

The crystal structure of d(GCGCTG) contained a hydrated magnesium cation bridging two base pairs in an interstrand fashion (Figure 7a) [86]. If this crystal was soaked in a solution containing cobalt(II) chloride, the magnesium was replaced by two Co^{2+} ions each coordinating to the N(7) position of a neighboring base pair [87]. The two Co^{2+} atoms had an interatomic distance of 6.1Å and were bridged by water molecules (Figure 7b).

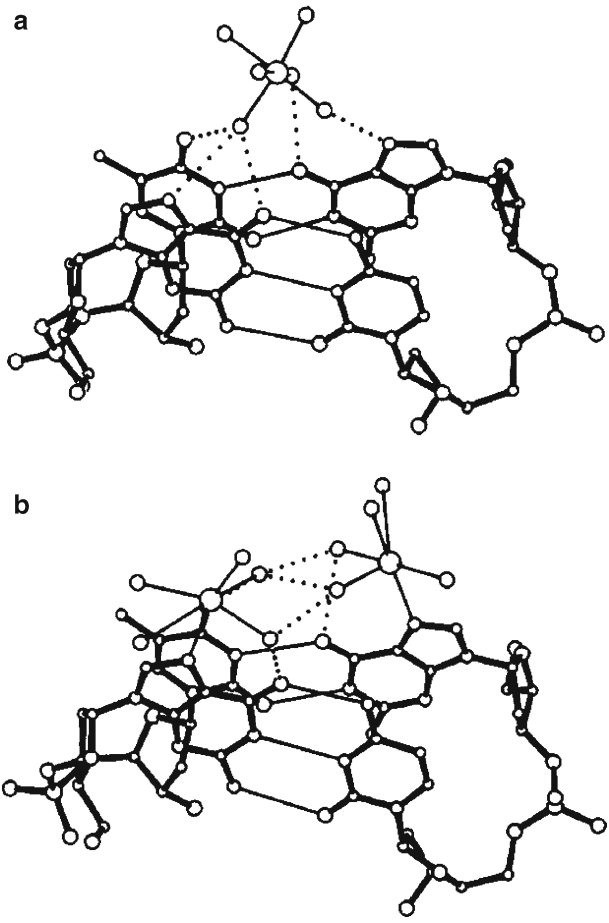


Figure 7 (**a**) Hydrated magnesium ions bridging two neighboring N(7) positions. (**b**) Two Co^{2+} atoms coordinating to the same N(7) positions as in (**a**), connected via several water molecules. Reproduced from [87] by permission from Oxford University Press; copyright 1993.

5 Condensation of Oligonucleotides

Condensation of DNA is the compaction into a regular unit. In the absence of this process, the linear stretched DNA would not fit into any cellular compartment [88]. DNA condensation is also one prerequisite for non-viral gene therapy [89].

Calf thymus DNA was titrated with increasing amounts of chromium(III) (up to 3 equivalents per phosphate) and studied by IR and VCD spectroscopy [41]. The authors observed first a Ψ type condensation in the right-handed form followed by formation of aggregates that they estimate to be in the size of a few micrometer. The behavior of 2.4 equivalents of manganese(II) per phosphate with calf thymus DNA was similar. First, a first Ψ type condensation still in the right-handed form was observed. Upon heating up to 70°C, partial denaturation occurred.

6 Diverse Conformational Changes

The bifunctional platinum complex cisplatin and its clinical successors carboplatin and oxaliplatin induce complex conformational changes, once the bifunctional adduct to the DNA or RNA is formed [90–93]. Since there are almost no protons around the platinated GG cross-link, initial NMR structure determinations without the use of residual dipolar couplings [94] were really challenging and an intensive discourse in the literature about the correct geometry of this unit arose [95–97]. Slightly different is the situation for polynuclear platinum complexes as described by Farell [98].

The interaction of ruthenium complexes with DNA and the following conformational changes of the DNA have been carefully reviewed by Brabec and Novacova in 2006 [99]. Recently, the binding of dinuclear ruthenium complexes to DNA oligomers containing one or two more adenine bulges has been studied by ^{1}H-NMR and molecular modeling [100]. Addition of copper(II) together with dopamine (4-(2-aminoethyl)benzene-1,2-diol) to plasmid DNA pCold caused some significant changes in the CD spectrum of the DNA. Control experiments with Cu(I) indicated that the afore mentioned changes might be copper(I)-dependent [101]. The interaction of a preformed copper(II)-glutathione complex with calf thymus DNA has been studied with the help of UV-vis, viscometry, EPR, CD, thin film IR spectroscopy, and gel electrophoresis. While EPR indicated that the copper complex remained unchanged in the presence of DNA, CD and thin film IR showed some conformational changes of the DNA in the presence of Cu-glutathione [102]. The interaction of calf thymus DNA and copper(II) was studied by IR and VCD spectroscopy [24]. At a copper to DNA phosphate ratio of 0.5–0.7, spectra were observed which were clearly different from B-DNA, but could not be assigned to either A-DNA or a denatured form. Platinum(II) and copper(II) complexes of 4-methyl-2-N-(2-pyridylmethyl) aminophenol have been studied for their antitumor activity and their ability to cleave DNA. The conformational changes upon calf thymus and φ174 supercoiled phage DNA were followed by gel electrophoresis and CD spectroscopy [103]. Also in this

publication, drastic changes were observed in the CD spectra that could not be assigned to a certain DNA geometry.

The folding of RNA in the presence of metal ions is a complex process that is connected with multiple conformational changes. This topic has been reviewed by Draper in 2004 [104]. It seems that the formation and changes between multiple RNA conformations are not yet fully understood. Sections of RNA that change their conformation upon binding of small molecules are called riboswitches. This event is connected with a change of gene expression. Riboswitches can also refold upon metal binding [105].

7 Concluding Remarks and Future Directions

In the preceding sections, we have summarized the influence metal ions can exert upon the conformation of oligonucleotides. Sometimes the metal can be substituted by polyamines, indicating that the effect is mainly coulombic. While many experimental techniques can record conformational changes, the assignment how the geometry actually changed is much more challenging, quite often even impossible to determine. The task can be of course complicated by multiple weak binding sites and overlay of spectra with different geometries. However, the technical progress will allow us to do more studies with less material at a higher resolution in the future.

Abbreviations

ADAR	double-stranded RNA adenosine deaminase
CD	circular dichroism
dien	diethylene triamine
dione	1,4,7,10,13-pentaazacyclohexa-decan-14-16-dione
DSC	differential scanning calorimetry
EPR	electron paramagnetic resonance
IR	infrared
m5C	5-methylcytidine
R	purine
spermidine	N-(3-aminopropyl)butane-1,4-diamine
spermine	N,N'-bis(3-aminopropyl)butane-1,4-diamine
tren	tris(2-aminoethyl)amine
trien	triethylene tetramine
tris	tris(hydroxymethyl)aminomethane
VCD	vibrational circular dichroism
Y	pyrimidine
Zα	Z-DNA binding domain

Acknowledgments The author thanks the University of Zürich for continuous support, all coworkers, past and present, for their input, and Dr. Susmita Gupta for a careful reading of the manuscript.

References

1. In the course of this chapter, we will not differentiate between oligonucleotides and polynucleotides, unless this is necessary.
2. R. E. Franklin, R. G. Gosling, *Acta Cryst*. **1953**, *6*, 673–677.
3. J. Pilet, J. Brahms, *Nature New Biol*. **1972**, *236*, 99–100.
4. W. C. Johnson, in *Circular Dichroism*, 2nd edn., Eds N. Berova, K. Nakanishi, R. W. Woody, Wiley-VCH, New York, 2000, pp 703–718.
5. W. L. Peticolas, E. Evertsz, *Method. Enzymol*. **1992**, *211*, 335–352.
6. W. L. Peticolas, *Method. Enzymol*. **1995**, *246*, 389–416.
7. V. A. Bloomfield, D. M. Crothers, I. Tinoco, Jr., *Nucleic Acids: Structures, Properties and Functions*, 1st edn., University Science Books, Sausalito, 2000, p 800.
8. R. Lavery, K. Zakrzewska, in *Oxford Handbook of Nucleic Acid Structure*, Ed S. Neidle, Oxford University Press, Oxford, 1999, pp 39–76.
9. G. E. Sims, S. H. Kim, *Nucleic Acids Res*. **2003**, *31*, 5607–5616.
10. D. Svozil, J. Kalina, M. Omelka, B. Schneider, *Nucleic Acids Res*. **2008**, *36*, 3690–3706.
11. B. Schneider, Z. Moravek, H. M. Berman, *Nucleic Acids Res*. **2004**, *32*, 1666–1677.
12. S. Devarajan, R. H. Shafer, *Nucleic Acids Res*. **1986**, *14*, 5099–5109.
13. J. G. Duguid, V. A. Bloomfield, J. M. Benevides, G. J. Thomas, *Biophys. J*. **1995**, *69*, 2623–2641.
14. A. M. Polyanichko, V. V. Andrushchenko, E. V. Chikhirzhina, V. I. Vorob'ev, H. Wieser, *Nucleic Acids Res*. **2004**, *32*, 989–996.
15. U. McDonnell, M. R. Hicks, M. J. Hannon, A. Rodger, *J. Inorg. Biochem*. **2008**, *102*, 2052–2059.
16. E. M. Kenyon, M. F. Hughes, *Toxicology* **2001**, *160*, 227–236.
17. N. E. Grossoehme, A. M. Spuches, D. E. Wilcox, *J. Biol. Inorg. Chem*. **2010**, *15*, 1183–1191.
18. M. G. Santangelo, A. Medina-Molner, A. Schweiger, G. Mitrikas, B. Spingler, *J. Biol. Inorg. Chem*. **2007**, *12*, 767–775.
19. M. G. Santangelo, P. M. Antoni, B. Spingler, G. Jeschke, *ChemPhysChem* **2010**, *11*, 599–606.
20. J. Anastassopoulou, *J. Mol. Struct*. **2003**, *651*, 19–26.
21. G. L. Eichhorn, *Coord. Chem. Rev*. **1993**, *128*, 167–173.
22. J. Kypr, I. Kejnovska, D. Renciuk, M. Vorlickova, *Nucleic Acids Res*. **2009**, *37*, 1713–1725.
23. K. J. Jalkanen, V. W. Jurgensen, A. Claussen, A. Rahim, G. M. Jensen, R. C. Wade, F. Nardi, C. Jung, I. M. Degtyarenko, R. M. Nieminen, F. Herrmann, M. Knapp-Mohammady, T. A. Niehaus, K. Frimand, S. Suhai, *Int. J. Quantum Chem*. **2006**, *106*, 1160–1198.
24. V. Andrushchenko, J. H. van de Sande, H. Wieser, *Biopolymers* **2003**, *72*, 374–390.
25. M. Egli, *Chem. Biol*. **2002**, *9*, 277–286.
26. (a) P. Auffinger, L. Bielecki, E. Westhof, *Structure* **2004**, *12*, 379–388. (b) L. D. Williams, in *DNA Binders and Related Subjects*, Springer, Dordrecht, Vol. 253, 2005, pp 77–88.
27. E. P. Geiduschek, A. Holtzer, in *Advances in Biological and Medical Physics*, Eds C. A. Tobias, J. H. Lawrence, Academic Press, New York, 1958, Vol. 6, pp 431–551.
28. R. H. Garrett, C. M. Grisham, *Biochemistry*, 4th edn., Brooks Cole, Boston, 2008, p 1184.
29. A. V. Tataurov, Y. You, R. Owczarzy, *Biophys. Chem*. **2008**, *133*, 66–70.
30. R. Owczarzy, *Biophys. Chem*. **2005**, *117*, 207–215.

31. S. A. Rice, P. Doty, *J. Am. Chem. Soc*. **1957**, *79*, 3937–3947.
32. G. Bonner, A. M. Klibanov, *Biotechnol. Bioeng*. **2000**, *68*, 339–344.
33. J. M. Sturtevant, E. P. Geiduschek, *J. Am. Chem. Soc*. **1958**, *80*, 2911–2911.
34. S. Lewin, D. S. Pepper, *Arch. Biochem. Biophys*. **1965**, *109*, 192–194.
35. M. Ageno, E. Dore, C. Frontali, *Biophys. J*. **1969**, *9*, 1281–1311.
36. R. Owczarzy, Y. You, B. G. Moreira, J. A. Manthey, L. Y. Huang, M. A. Behlke, J. A. Walder, *Biochemistry* **2004**, *43*, 3537–3554.
37. R. Owczarzy, B. G. Moreira, Y. You, M. A. Behlke, J. A. Walder, *Biochemistry* **2008**, *47*, 5336–5353.
38. R. Owczarzy, A. V. Tataurov, Y. Wu, J. A. Manthey, K. A. McQuisten, H. G. Almabrazi, K. F. Pedersen, Y. Lin, J. Garretson, N. O. McEntaggart, C. A. Sailor, R. B. Dawson, A. S. Peek, *Nucleic Acids Res*. **2008**, *36*, W163–W169.
39. G. L. Eichhorn, *Nature* **1962**, *194*, 474–475.
40. J. Duguid, V. A. Bloomfield, J. Benevides, G. J. Thomas, *Biophys. J*. **1993**, *65*, 1916–1928.
41. V. Andrushchenko, D. Tsankov, H. Wieser, *J. Mol. Struct*. **2003**, *661*, 541–560.
42. J. H. Lee, *Bull. Korean Chem. Soc*. **2008**, *29*, 1937–1940.
43. V. I. Ivanov, D. Y. Krylov, *Methods Enzymol*. **1992**, *211*, 111–127.
44. F. M. Pohl, T. M. Jovin, *J. Mol. Biol*. **1972**, *67*, 375–396.
45. T. Brown, W. N. Hunter, *Biopolymers* **1997**, *44*, 91–103.
46. N. W. Luedtke, *Chimia* **2009**, *63*, 134–139.
47. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, A. Rich, *Nature* **1979**, *282*, 680–686.
48. H. Drew, T. Takano, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature* **1980**, *286*, 567–573.
49. B. A. Brown II, A. Rich, *Acta Biochim. Pol*. **2001**, *48*, 295–312.
50. H. Sugiyama, K. Kawai, A. Matsunaga, K. Fujimoto, I. Saito, H. Robinson, A. H.-J. Wang, *Nucleic Acids Res*. **1996**, *24*, 1272–1278.
51. T. M. Jovin, L. P. Mcintosh, D. J. Arndt-Jovin, D. A. Zarling, M. Robert-Nicoud, J. H. van de Sande, K. F. Jorgenson, F. Eckstein, *J. Biomol. Struct. Dyn*. **1983**, *1*, 21–57.
52. S. M. Singh, B. Murphy, R. O'Reilly, *Clin. Genet*. **2003**, *64*, 451–460.
53. A. Herbert, K. Lowenhaupt, J. Spitzner, A. Rich, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7550–7554.
54. A. Herbert, A. Rich, *J. Biol. Chem*. **1996**, *271*, 11595–11598.
55. A. Rich, S. Zhang, *Nat. Rev. Genet*. **2003**, *4*, 566–572.
56. G. Wang, L. A. Christensen, K. M. Vasquez, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2677–2682.
57. H. Li, J. Xiao, J. M. Li, L. Lu, S. Feng, P. Dröge, *Nucleic Acids Res*. **2009**, *37*, 2737–2746.
58. P. Khuu, M. Sandor, J. DeYoung, P. S. Ho, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16528–16533.
59. K. Hall, P. Cruz, I. Tinoco Jr., T. M. Jovin, J. H. van de Sande, *Nature* **1984**, *311*, 584–586.
60. J. H. Riazance, W. A. Baase, W. C. Johnson, K. Hall, P. Cruz, I. Tinoco, *Nucleic Acids Res*. **1985**, *13*, 4983–4989.
61. M. O. Trulson, P. Cruz, J. D. Puglisi, I. Tinoco, R. A. Mathies, *Biochemistry* **1987**, *26*, 8624–8630.
62. M. K. Teng, Y. C. Liaw, G. A. van der Marel, J. H. van Boom, A. H. J. Wang, *Biochemistry* **1989**, *28*, 4923–4928.
63. D. Placido, B. A. Brown II, K. Lowenhaupt, A. Rich, A. Athanasiadis, *Structure* **2007**, *15*, 395–404.
64. P. W. Davis, R. W. Adamiak, I. Tinoco Jr., *Biopolymers* **1990**, *29*, 109–122.
65. M. Popenda, J. Milecki, R. W. Adamiak, *Nucleic Acids Res*. **2004**, *32*, 4044–4054.
66. M. Behe, G. Felsenfeld, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 1619–1623.
67. S. Sarhan, N. Seiler, *Biol. Chem. Hoppe-Seyler* **1989**, *370*, 1279–1284.
68. A. Woisard, G. V. Fazakerley, W. Guschlbauer, *J. Biomol. Struct. Dyn*. **1985**, *2*, 1205–1220.
69. J. H. van de Sande, L. P. McIntosh, T. M. Jovin, *EMBO J*. **1982**, *1*, 777–782.

70. T. Schoenknecht, H. Diebler, *J. Inorg. Biochem*. **1993**, *50*, 283–298.
71. B. Spingler, *Inorg. Chem*. **2005**, *44*, 831–833.
72. G. V. Fazakerley, *Nucleic Acids Res*. **1984**, *12*, 3643–3648.
73. A. Medina-Molner, C. Da Pieve, B. Spingler, unpublished.
74. C. Bauer, A. H.-J. Wang, *J. Inorg. Biochem*. **1997**, *68*, 129–135.
75. B. Malfoy, B. Hartmann, M. Leng, *Nucleic Acids Res*. **1981**, *9*, 5659–5669.
76. A. Johnson, Y. Qu, B. Van Houten, N. Farrell, *Nucleic Acids Res*. **1992**, *20*, 1697–1703.
77. P. K. Wu, M. Kharatishvili, Y. Qu, N. Farrell, *J. Inorg. Biochem*. **1996**, *63*, 9–18.
78. T. D. McGregor, W. Bousfield, Y. Qu, N. Farrell, *J. Inorg. Biochem*. **2002**, *91*, 212–219.
79. Y. Qu, A. Harris, A. Hegmans, A. Petz, P. Kabolizadeh, H. Penazova, N. Farrell, *J. Inorg. Biochem*. **2004**, *98*, 1591–1598.
80. E. Kimura, A. Sakonaka, M. Nakamoto, *Biochim. Biophys. Acta* **1981**, *678*, 172–179.
81. E. Kimura, A. Yatsunami, A. Watanabe, R. Machida, T. Koike, H. Fujioka, Y. Kuramoto, M. Sumomogi, K. Kunimitsu, A. Yamashita, *Biochim. Biophys. Acta* **1983**, *745*, 37–43.
82. H.-C. Shih, N. Tang, C. J. Burrows, S. E. Rokita, *J. Am. Chem. Soc*. **1998**, *120*, 3284–3288.
83. B. Spingler, C. Da Pieve, *Dalton Trans.* **2005**, 1637–1643.
84. B. Spingler, F. Zobi, P. M. Antoni, A. Medina-Molner, R. Alberto, *Chimia* **2005**, *59*, 826–831.
85. T. F. Kagawa, B. H. Geierstanger, A. H.-J. Wang, P. S. Ho, *J. Biol. Chem*. **1991**, *266*, 20175–20184.
86. P. S. Ho, C. A. Frederick, G. J. Quigley, G. A. van der Marel, J. H. van Boom, A. H. J. Wang, A. Rich, *EMBO J*. **1985**, *4*, 3617–3623.
87. Y.-G. Gao, M. Sriram, A. H.-J. Wang, *Nucleic Acids Res*. **1993**, *21*, 4093–4101.
88. V. A. Bloomfield, *Biopolymers* **1998**, *44*, 269–282.
89. P. Saccardo, A. Villaverde, N. Gonzalez-Montalban, *Biotechnol. Adv*. **2009**, *27*, 432–438.
90. A. Gelasco, S. J. Lippard, in *Metallopharmaceuticals I, DNA Interactions*, Eds M. J. Clarke, P. J. Sadler, Springer, Berlin, 1999, Vol. 1, pp 1–43.
91. E. R. Jamieson, S. J. Lippard, *Chem. Rev*. **1999**, *99*, 2467–2498.
92. B. Spingler, D. A. Whittington, S. J. Lippard, *Inorg. Chem*. **2001**, *40*, 5596–5602.
93. V. Brabec, J. Kasparkova, *Drug Resist. Update* **2005**, *8*, 131–146.
94. D. MacDonald, P. Lu, *Curr. Opin. Struc. Biol*. **2002**, *12*, 337–343.
95. J. A. Parkinson, Y. Chen, P. D. S. Murdoch, Z. Guo, S. J. Berners-Price, T. Brown, P. J. Sadler, *Chem. Eur. J*. **2000**, *6*, 3636–3644.
96. L. G. Marzilli, J. S. Saad, Z. Kuklenyik, K. A. Keating, Y. Xu, *J. Am. Chem. Soc*. **2001**, *123*, 2764–2770.
97. M.-A. Elizondo-Riojas, J. Kozelka, *J. Mol. Biol*. **2001**, *314*, 1227–1243.
98. N. Farrell, in *Metal Complexes in Tumor Diagnosis and as Anticancer Agents*, Eds A. Sigel, H. Sigel, Marcel Dekker Inc, 2004, Vol. 42, pp 251–296.
99. V. Brabec, O. Novakova, *Drug Resist. Updates* **2006**, *9*, 111–122.
100. D. P. Buck, J. A. Paul, M. J. Pisani, J. G. Collins, F. R. Keene, *Aust. J. Chem*. **2010**, *63*, 1365–1375.
101. M. Ando, K. Ueda, R. Makino, Y. Nishino, H. Nishida, C. Toda, Y. Okamoto, N. Kojima, *J. Health Sci*. **2009**, *55*, 319–323.
102. K. K. Mukherjea, G. Panda, M. Selim, *Trans. Metal Chem*. **2008**, *33*, 203–210.
103. S. Roy, P. U. Maheswari, M. Lutz, A. L. Spek, H. den Dulk, S. Barends, G. P. van Wezel, F. Hartl, J. Reedijk, *Dalton Trans.* **2009**, 10846–10860.
104. D. E. Draper, *RNA* **2004**, *10*, 335–343.
105. A. R. Ferré-D'Amaré, W. C. Winkler, in *Structural and Catalytic Roles of Metal Ions in RNA*, Eds A. Sigel, H. Sigel, R. K. O. Sigel, Royal Society of Chemistry, Cambridge, 2011, Vol. 9, pp 141–173.
106. D. W. Gruenwedel, *Eur. J. Biochem*. **1994**, *219*, 491–496.
107. B. Spingler, P. M. Antoni, *Chem. Eur. J*. **2007**, 13, 6617–6622.

Chapter 4
G-Quadruplexes and Metal Ions

Nancy H. Campbell and Stephen Neidle

Contents

Abstract Metal ions stabilize quadruplex nucleic acids by coordinating the O6 guanine atoms from G-quartets. These quartets form the basic motif of quadruplex structures. This article systematically surveys the available crystallographic data on native quadruplexes, their ligand complexes and (in one instance) a protein complex. Three categories of quadruplex are examined, tetramolecular, bimolecular, and intramolecular: all are formed by telomeric nucleic acid sequences from human or ciliate organisms.

Keywords crystal structures • ion channel • metal ion coordination • Protein Data Bank • quadruplex

N.H. Campbell • S. Neidle (✉)
CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, UK
e-mail: stephen.neidle@pharmacy.ac.uk

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_4,

1 Introduction – The Basics of Quadruplex Structure

Polynucleotides and oligonucleotides containing long repeats of guanosine nucleotides have long been known to have a marked tendency to self-associate and aggregrate into gel-like materials. Analysis of the fibres formed from gels of guanosine monophosphate [1] led to a structural model with four strands held together by a hydrogen-bonding arrangement of four guanines – the G-quartet (Figure 1). This very stable motif can self-associate and layers of G-quartets are then formed. The structures of polynucleotides containing purely guanosines are thus four-stranded analogues of the Watson-Crick AT and GC base-paired double helices [2].

Short lengths of nucleic acid can also form structures based on the G-quartet motif under certain circumstances. These quadruplex nucleic acids can be formed from one, two or four strands of DNA or RNA sequences that contain a number of short tracts of guanine nucleotides. Each G-tract should contain optimally 3–5 consecutive guanosines, with a typical sequence being:

$$G_mX_nG_mX_oG_mX_pG_m$$

where m is the number of G residues in each short G-tract, which are usually directly involved in G-quartet interactions. X_n, X_o, and X_p can be any combination of residues, including G, forming the linkers between G-quartets. This notation has all the G-tracts of equal length. However, this is not a necessary condition since if one or more of the short G tracts is longer than the others, some G residues will be most likely located in the linker regions. Also, the presence of G residues within connecting sequences can give rise to more complex arrangements in which such individual G residues become inserted into the G-quartet stack and displace the G-tract G residues. The structures and topologies of a number of telomeric quadruplexes DNAs in particular, have been determined by X-ray crystallographic and NMR methods; the reader is referred to the extensive reviews on the subject for further detail [3–5].

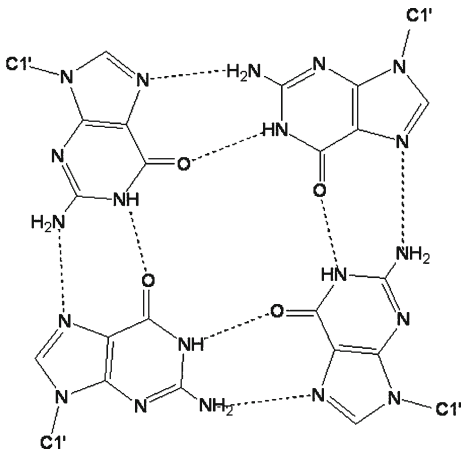


Figure 1 Structure of the G-quartet. Hydrogen bonds are shown as dashed lines.

In principle, tetramolecular (tetrameric) and bimolecular (dimeric) quadruplexes can be formed from the association of non-equal sequences, although almost all bimolecular quadruplexes reported to date are formed by the association of two identical sequences $X_nG_mX_oG_mX_p$, where n and p may or may not be zero. Tetramolecular quadruplexes are formed by four $X_nG_mX_o$ strands associating together, where X_n and X_o may be terminal sequences of zero or non-zero length.

The connecting sequences intervening between successive G-tracts can link stacked G-quartets in a number of distinct ways, so that a wide variety of quadruplex topologies can be formed. These connectors can form loops, of which there are three principal types: diagonal, lateral or propeller (also sometimes termed chain-reversal). The particular type formed is dependent on the number of G-quartets comprising the stem of a quadruplex, on loop length and sequence and sometimes on the nature of the alkali metal ion. Propeller loops necessarily connect two strands in the same parallel orientation, linking the bottom G-quartet with the top G-quartet in a G-stack whereas diagonal and lateral loops connect chains in opposing, anti-parallel orientations. Lateral (sometimes termed edge-wise) loops join adjacent G-strands. Two lateral loops can be located either on the same or opposite faces of a quadruplex, corresponding to head-to-head or head-to-tail, respectively, when in bimolecular quadruplexes. Strand polarities can vary, with one being a head-to-tail lateral loop dimer in which all adjacent strands are anti-parallel, and the other is a head-to-head hairpin quadruplex with one adjacent strand parallel and the other is anti-parallel. The second type of anti-parallel loop is the diagonal loop, which joins opposite G-strands. In this instance the directionalities of adjacent strands must alternate between parallel and anti-parallel, and are arranged around a core of stacked G-quartets. All-parallel quadruplexes have all guanine glycosidic angles in the *anti* conformation, whereas anti-parallel quadruplexes have equal numbers of the guanosine nucleotides in *syn* and *anti* conformations, arranged in a way that is particular for a given topology. All quadruplex structures have four grooves, defined as the cavities bounded by the phosphodiester backbones.

2 The Role of Metal Ions

Early models for four-fold helices and quadruplex oligonucleotide structures based on the G-quartet motif [6,7] suggested the existence of a central channel into which water molecules or cations could be placed. This interpretation was subsequently confirmed and extended by a large number of biophysical and structural studies on G-quartet-containing nucleic acid sequences (reviewed in [8]). This central channel is a universal and unique feature of all quadruplex structures, distinguishing them from other types of nucleic acid arrangements, not least nucleic acid double helices. Of particular significance have been the findings that the stability of these four-stranded structures depends on the presence of physiological concentrations of the alkali metals sodium or potassium with the order of stabilizing ability being $K^+ > Na^+$.

The precise location of an ion in the channel would be expected to depend on its ionic radius. K^+ ions are too large to be accommodated in the plane of a G-quartet, whereas Na^+ ions are small enough to be coordinated in-plane with all four guanine bases in a G-quartet. Several other ions can stabilize quadruplexes, including those of divalent heavier metals such as rubidium and cesium (Rb^+, Cs^+), strontium (Sr^{3+}), thallium (Tl^+), calcium (Ca^{2+}), lead (Pb^{2+}), barium (Ba^{2+}), though all to a lesser extent than K^+. Metal ions in the channel are always located in crystal structures. Metal ion coordination patterns observed in X-ray crystallographic analyses generally concur with findings from several multinuclear NMR studies, which have been able to directly locate Na^+, K^+, Rb^+, and Ca^{2+} ions [9]. The presence of particular ions in the channel can also change quadruplex conformation and fold, an effect that has been most extensively documented for human telomeric quadruplexes, reviewed in [5].

The integral role that metal cations play in G-quadruplex structures was directly observed when the first crystal structure of a quadruplex, the tetramolecular d(TGGGGT), was reported [10]. The crystal structures of bimolecular and intramolecular human telomeric G-quadruplex [11], also showing central channel metal ion arrangements. Here, a novel structure, with an all-parallel topology, was revealed to be the form crystallizing in potassium-containing experimental conditions. Surprisingly, this was fundamentally distinct from an earlier structure reported using NMR techniques [12]. The latter was reported for the identical G-quadruplex-forming sequence in a sodium-containing environment and showed an anti-parallel topology (Figure 2). Earlier studies on the bimolecular quadruplex formed by d(GGGGTTTTGGGG) had suggested that here too, NMR and crystallography were in disagreement [13,14], although a subsequent crystallographic analysis [15] showed that the NMR assignment of topology was correct and that the initial crystal structure [13] was not. G-quadruplex structures are also sensitive to flanking nucleotides, concentration, and molecular crowding conditions [16,17]. In general, metal cations coordinate to the O6 guanine substituent groups lining the central channel of a quadruplex. The positively-charged cations counter the negative electrostatic effect of the carbonyl groups, thus enhancing quadruplex stabilization (Figure 2).

The aim of this review is to survey all the available crystal structure data on the geometry of metal ions in quadruplex crystal structures, as these constitute the most accurate and reliable source of geometric information on ion environments in quadruplexes. We have conducted a systematic search of the Protein Data Bank Database (PDB) [18] using the key search words *quadruplex* or *tetraplex DNA* or *RNA crystal structures*, which resulted in 151 hits (May 2011). The list was further edited by omitting structures that do not represent biologically relevant sequences. The final list contained 36 structures. All 36 structures are from sequences that are of telomeric origin. These have been further categorized into three subgroups in respect of the organism from which the telomeric G-quadruplex-forming sequence originated: *Tetrahymena*, *Oxytricha nova*, and human. Each structure was visualized and metal coordination distances measured and recorded.

A variety of metal ions have been observed in the central quadruplex channel in these crystal structures. For example, mono-cations such as sodium, potassium and thallium and di-cations such as strontium and calcium have all been reported. The observed ions either coordinate to a single G-quartet in (or close to) a square-planar

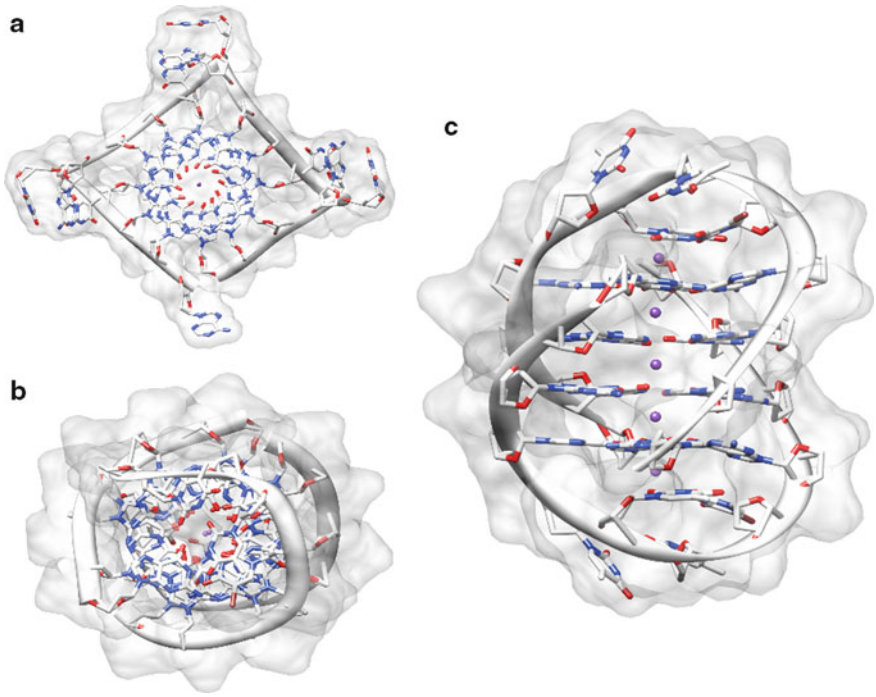


Figure 2 The telomeric quadruplex is parallel in (**a**) human and anti-parallel in (**b**) *Oxytricha nova* quadruplexes; (**c**) the central ion channel. Potassium ions are colored mauve.

arrangement or between two stacking G-quartets in a bipyramidal antiprismatic arrangement. The interplay between specific quadruplex sequence, cation type and size and other factors affecting the coordination arrangement are discussed below.

3 Metal Ions in Quadruplex Structures

3.1 Tetrahymena Telomeric Quadruplexes

The quadruplex crystal structure formed from the telomeric sequence in *Tetrahymena* is tetramolecular. Four strands, each comprising the sequence d(TGGGGT), associate together, adopting the same orientation in the 5' to 3' direction to form a parallel quadruplex (Figure 3). This parallel quadruplex arrangement is maintained in both of the RNA and the DNA forms. The ions that have been observed in various crystal structures of this quadruplex are: sodium, calcium, thallium, strontium or potassium.

Both sodium and thallium ions are able to adopt either the square planar or bipyramidal antiprismatic arrangement. The preference for forming one coordination arrangement compared to another is a consequence of the position of the cation in the ion channel. It is notable that quadruplexes can stack in the crystal structures in a

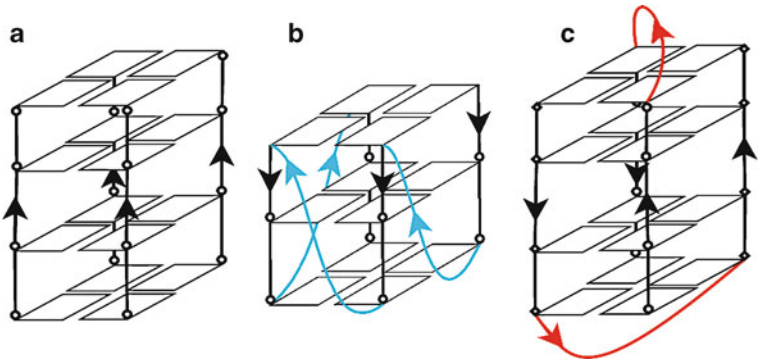


Figure 3 Schematic views of the topology and strand orientations in (**a**) the intermolecular quadruplex d(TGGGGT), (**b**) the human telomeric parallel quadruplex, and (**c**) the *Oxytricha nova* bimolecular anti-parallel quadruplex.

manner such as to allow a continuous one-dimensional array of ions in the ion channel to be formed, proceeding from one quadruplex to the next. At the point where this continuity is broken, for example when just two quadruplexes are stacked together (a typical motif), the ion channel ends. The ion channel hence resembles a linear tube with a center and two ends. The ends are represented by an open G-quartet that does not have a further quadruplex stacking onto it. So, where the ions are coordinating to an "end" G-quartet, sodium and thallium ions adopt a square planar coordination arrangement, and a bipyramidal antiprismatic arrangement where they are positioned away from the "end" G-quartet and closer to the centre of the ion channel (Figure 4).

On the other hand, calcium, strontium and potassium ions show a marked preference for the bipyramidal antiprismatic coordination arrangement. These ions are always positioned between the stacked G-quartets and never within a single G-quartet (Figure 5). The ionic radius for strontium and potassium ions are similar, 1.18 Å and 1.38 Å, respectively. These ions are larger than sodium and thallium (ionic radii of 1.02 Å and 0.89 Å, respectively). It is apparent that the larger size ions simply may not be accommodated except between the stacking G-quartets, for purely steric reasons. Calcium ions, on the other hand, have an ionic radius of 1.00 Å (closer to that of sodium and thallium) but still show a tendency to coordinate between the G-quartets rather than be in-plane or close to them. It may be that the two positive charges of the calcium ion are sufficient to counter the negative electrostatic character of eight O6 carbonyl groups from two consecutive stacked G-quartets and thus calcium ions are observed to be positioned between two G-quartets in all the crystal structures where it is present.

The coordination distances have a smaller distribution of values for the square planar arrangement, ranging between 2.3 to 2.4 Å compared to the range seen with bipyramidal antiprismatic coordination, of 2.3 to 3.6 Å (Table 1). In a square-planar arrangement, coordination distances are limited by the linear dimensions of the relatively small and flat area in the center of the G-quartet where the cation has to fit in order to maintain the cation-G-quartet square planar geometry. In contrast, a

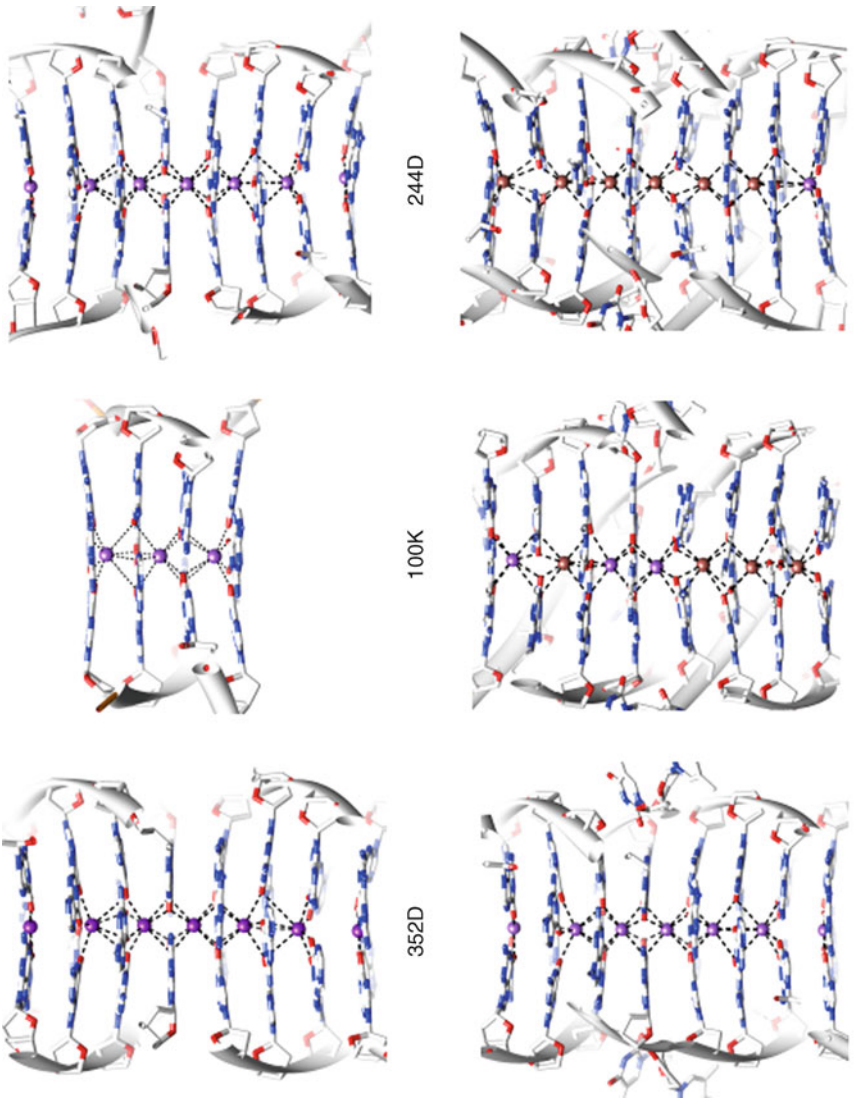


Figure 4 Views of the structures of the tetramolecular quadruplexes listed in Table 1, showing the ion coordination in the channels and the relevant PDB codes. Sodium ions are colored mauve.

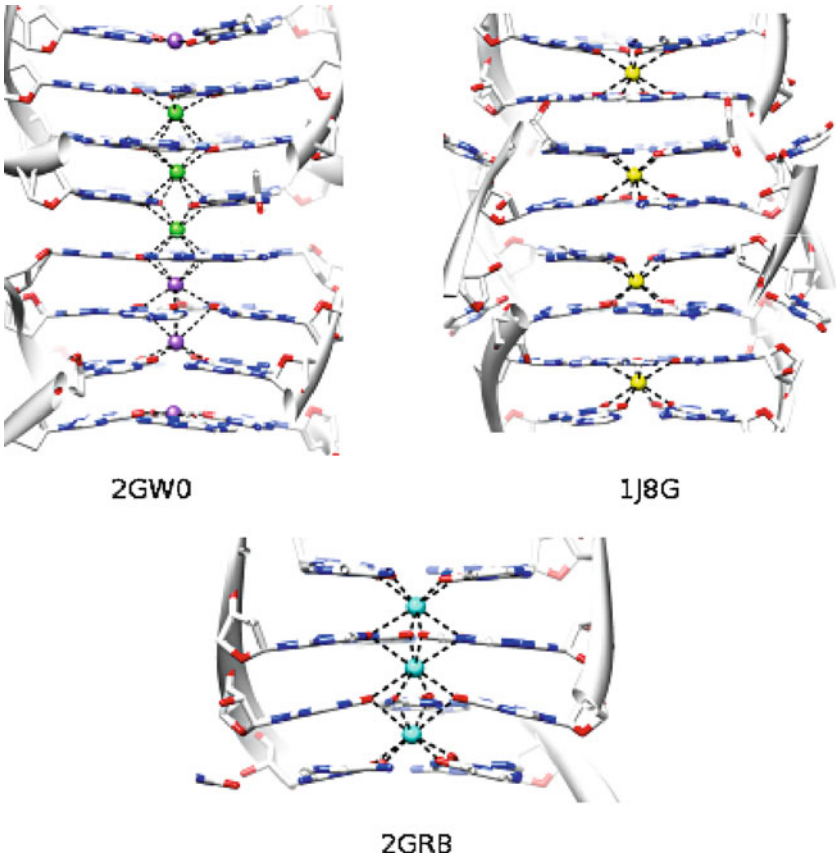


Figure 5 Further tetramolecular quadruplexes, with mixed ion-containing channels. Sodium ions are colored mauve, strontium ions are green and potassium ions are cyan.

larger area is available between two stacked G-quartets for an ion to be fitted in the bipyramidal antiprismatic arrangement.

The range of distances for cations at the interface between two separate G-quadruplexes is 2.6 Å to 3.2 Å (Table 1), with the ions adopting bipyramidal antiprismatic coordination in all instances. This is not surprising since the ion is positioned between two continuous ion channels.

3.2 Oxytricha nova Telomeric Quadruplexes

The quadruplex crystal structures determined for the telomeric sequence in *Oxytricha nova* are all bimolecular. Two strands, each formed of the sequence d(GGGGTTTTGGGG), associate together adopting an anti-parallel quadruplex conformation where the strands run in opposite direction relative to each other in the

Table 1 Ion channel features in *Tetrahymena* telomeric quadruplexes.

		Coordination distances (Å)								
		in channel		between stacking quadruplexes						
		Coordination								
PDB ID	Ion	4	8	8	Additive ion/ ligand	Sequence	Topology	Resolution (Å)	Notes	Refs
352D	Na	2.3	2.4 – 3.6	2.7 – 2.8	Ca	TG_4T	parallel	0.95	DNA	[10]
1O0K	Na	n/a	2.4 – 3.2	n/a	daunomycin	TG_4T	parallel	1.17	DNA	[19]
244D	Na	2.3 – 2.4	2.4 – 3.6	2.7 – 2.8	Ca	TG_4T	parallel	1.20	DNA	[10]
2O4F	Na	2.2 – 2.3	2.4 – 3.2	2.7 – 2.8	Li	TG_4T	parallel	1.50	DNA	[20]
2GW0	Na	2.3 – 2.4	2.4 – 3.2	n/a	Ca	TG_4T	parallel	1.55	DNA	[21]
	Ca	n/a	2.5 – 3.2	2.6 – 2.8						
1S45	Na	n/a	2.5 – 3.5	n/a	Mg	TG_4T	parallel	2.20	DNA	[22]
	Tl	n/a	2.3 – 3.7	2.7 – 2.9						
1S47	Na	n/a	2.7 – 3.6	2.7 – 3.2	Mg	TG_4T	parallel	2.50	DNA	[22]
	Tl	n/a	2.5 – 3.3	n/a						
1J8G	Sr	n/a	2.6	n/a	Na, Ca, spermine	UG_4U	parallel	0.61	RNA	[23]
2GRB	K	n/a	2.7 – 3.1	n/a	Sr	$UGIG_2U$	parallel	1.40	RNA	[24]

5' to 3' direction (Figure 3) and the d(TTTT) sequence forms a diagonal loop. The quadruplex fold is not affected by ion type or presence of bound ligands or protein (Table 2). It is also not affected by the crystal form or space group (i.e. crystal packing); all of the native (ligand-free) structures have this identical topology (PDB IDs 1JPQ, 1JRN, 2GWQ and 2GWE) (Table 2).

In this quadruplex type, both potassium and thallium ions always adopt bipyramidal antiprismatic coordination between two consecutive stacked G-quartets despite the large difference in their ionic radius (Figure 6). The coordination distances to O6 atoms range between 2.4 Å and 3.3 Å. However, sodium ions (even though they are slightly larger than thallium ions) are coordinated in square planar arrangements (Figure 6) within the G-quartets, with coordination distances to O6 atoms ranging between 2.1 Å and 2.5 Å (Table 2).

In contrast to the *Tetrahymena* quadruplex structures, individual *Oxytricha nova* bimolecular quadruplexes cannot stack on one another in the crystal structures due to the presence of the diagonal T_4 loop (Figure 3) capping each terminal G-quartet face and thus preventing the continuation of the ion channel from one quadruplex to the next. No crystal structure of an intramolecular *Oxytricha nova* quadruplex is currently available, but it will be interesting to see what geometries are shown by the channel ions.

3.3 Human Telomeric Quadruplexes

The available database of human telomeric crystal structures consists of eight structures; six DNA quadruplexes, two of which are ligand-free and four ligand-bound, together with two RNA quadruplexes, one ligand-free and one ligand-bound (Table 3).

All the structures reported to date contain potassium ions only in the ion channel, irrespective of the ions present in the experimental conditions used in the crystallization setup – typically a number of different ions are employed simultaneously in crystallization trials. This is a consequence of the higher affinity of potassium ions for quadruplexes relative to any other ion. In all the structures in the database, the potassium ions adopt bipyramidal antiprismatic coordination (Figure 7) with distances ranging between 2.4 Å to 3.3 Å. Similarly, the potassium ion involved in linking two stacking quadruplexes adopts the same bipyramidal antiprismatic coordination with distances in a similar range: 2.6 Å to 3.0 Å (Table 3).

4 Conclusions

All non-human telomeric quadruplex crystal structures reported to date consistently show that the fold of a specific quadruplex sequence is not affected by the ionic conditions used in the crystallization experiment or the coordination of the ions in the ion channel. All these structures are highly hydrated and the crystal lattice environments are crowded with ions, solvent and quadruplex molecules. The relevance of the human telomeric quadruplex crystal structures to biological environments is

Table 2 Ion channel features in *Oxytricha nova* telomeric quadruplexes.

		Coordination distances (Å)						
		Coordination in channel						
PDB ID	Ion	4	6 – 8	Additive ion/ ligand	Sequence	Resolution (Å)	Notes	Refs
3NZ7	K		2.7 – 3.0	S,S fluorosubstituted BSU-6039	$G_4T_4G_4$	1.10	DNA	[25]
3NYP	K		2.9 – 3.3	R,R fluorosubstituted BSU-6039	$G_4T_4G_4$	1.18	DNA	[25]
2HBN	Tl		2.5 – 3.1	n/a	$G_4T_4G_4$	1.55	DNA	[26]
1JPQ	K		2.7 – 3.0	n/a	$G_4T_4G_4$	1.60	DNA	[27]
1L1H	K		2.6 – 3.0	BSU6039	$G_4T_4G_4$	1.75	DNA	[28]
3EUM	K		2.7 – 3.1	BSU-6066	$G_4T_4G_4$	1.78	DNA	[29]
1JB7	Na	2.1 – 2.5	2.6 – 3.1	n/a	$G_4T_4G_4$	1.86	protein complex	[30]
1JRN	K		2.5 – 3.1	n/a	$G_4T_4G_4$	2.00	DNA	[31]
2GWQ	K		2.5 – 3.2	n/a	$G_4T_4G_4$	2.00	DNA	to be published
3ERU	K		2.7 – 3.1	BSU-6045	$G_4T_4G_4$	2.00	DNA	[28]
2GWE	K		2.4 – 3.2	n/a	$G_4T_4G_4$	2.20	DNA	to be published
3EQW	K		2.6 – 3.2	BSU-6042	$G_4T_4G_4$	2.20	DNA	[28]
3ES0	K		2.7 – 3.0	BSU-6048	$G_4T_4G_4$	2.20	DNA	[28]
3EUI	K		2.6 – 3.1	BSU-6042	$G_4T_4G_4$	2.20	DNA	[28]
3EM2	K		2.7 – 3.1	BSU-6038	$G_4T_4G_4$	2.30	DNA	[28]
3ET8	K		2.6 – 3.1	BSU-6054	$G_4T_4G_4$	2.45	DNA	[28]

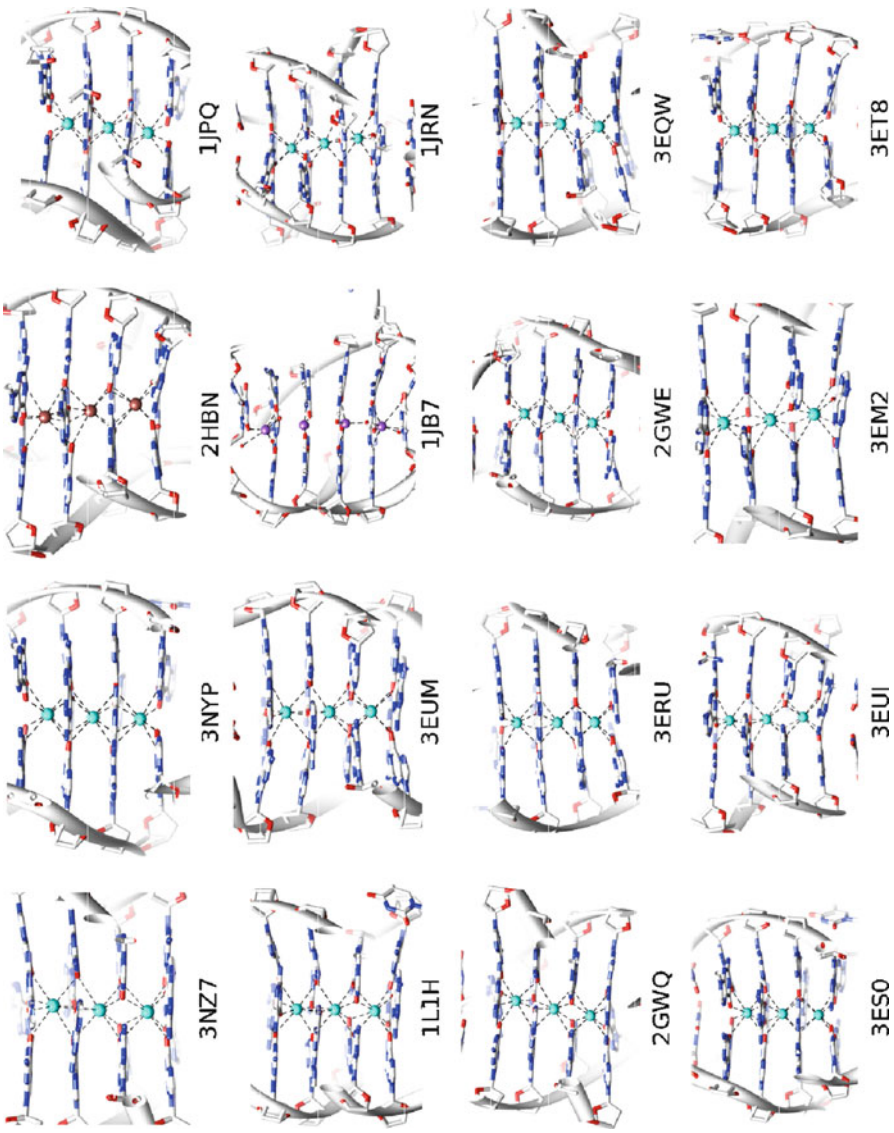
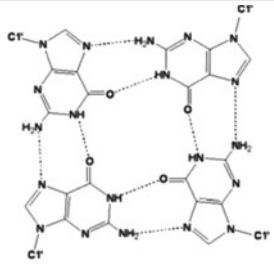


Figure 6 Ion channels in *Oxytricha* nova telomeric quadruplexes. Thallium ions are colored brown.

Table 3 Ion channel features in human telomeric quadruplexes.



		Coordination distances (Å)							
		Coordination = 8							
PDB ID	Ion	in channel	between stacking quadruplexes	Additive ion/ ligand	Stacking	Sequence	Resolution (Å)	Notes	Refs
2HRI	K	2.6 – 3.1	2.8 – 3.0	porphyrin	3' to 3' and 5' to 5'	$AG_3T_2AG_3$	2.09	DNA	[30]
1KF1	K	2.4 – 3.2	2.6 – 2.8	n/a	5' to 5'	$AG_3(T_2AG_3)_3$	2.10	DNA	[11]
3CDM	K	2.6 – 3.1	n/a	naphthalene diimide	5' to 5'	$TAG_3(T_2AG_3)_3$	2.10	DNA	[31]
3CCO	K	2.6 – 3.1	2.9	naphthalene diimide	5' to 5'	$TAG_3T_2AG_3T$	2.20	DNA	[31]
1K8P	K	2.4 – 3.0	n/a	Na	5' to 5'	$^{Br}UAG_3{}^{Br}UTAG_3T$	2.40	DNA	[11]
3CE5	K	2.6 – 3.2	n/a	BRACO19	5' to 3'	$TAG_3T_2AG_3T$	2.50	DNA	[32]
3IBK	K	2.6 – 3.1	n/a	n/a	5' to 5'	$^{Br}UAG_3U_2AG_3U$	2.20	RNA	[33]
3MIJ	K	2.6 – 3.3	n/a	acridine derivative	5' to 5'	$UAG_3U_2AG_3U$	2.60	RNA	[34]

^{Br}U = 5-bromo-2'-deoxyuridine.

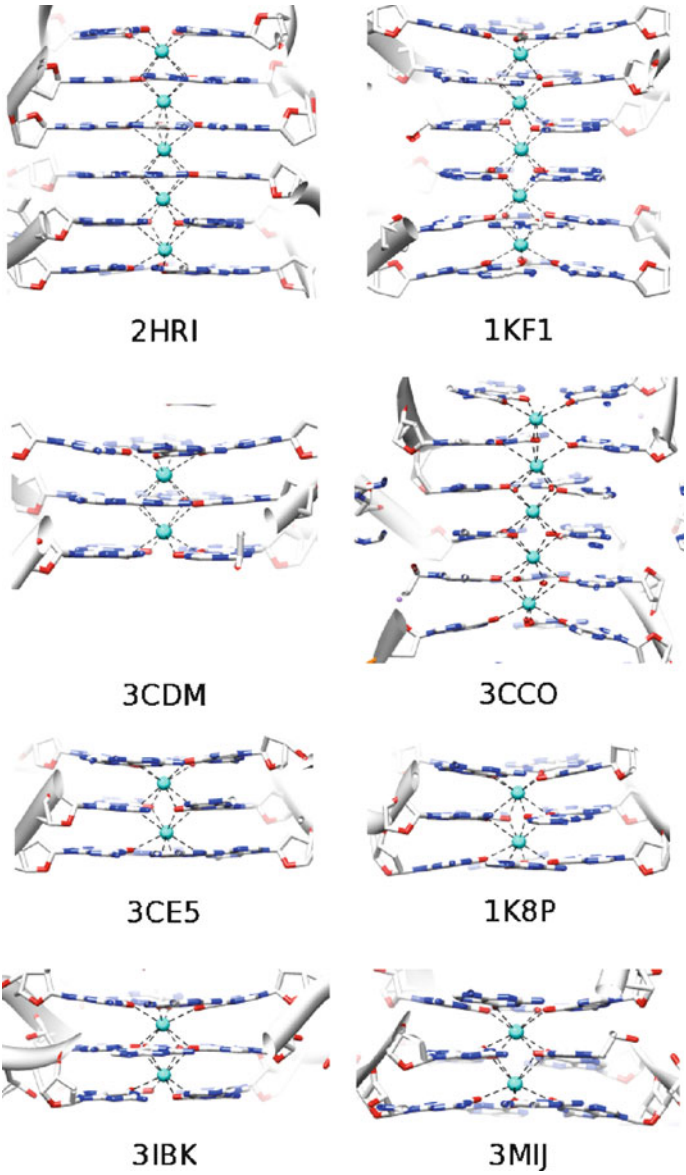


Figure 7 Ion channels in human telomeric quadruplexes.

still controversial [35,36], but whatever the form of these quadruplexes, the ion channel is retained as an essential structural motif.

Smaller ions, such as sodium and thallium, are more inherently likely to exhibit dynamic behavior within the ion channel but they are trapped in a particular position by coordinating within a G-quartet in a square planar arrangement and hence are

observable by X-ray crystallography. It is significant that none of the quadruplex crystal structures show any evidence for disordered sodium ions in the channels, which would indicate high mobility. Similarly sized ions that possess greater net positive charge (such as calcium), may also in principle be able to move along the ion channel, though again they are only observed with bipyramidal antiprismatic coordination between two stacked G-quartets and not within individual channels. This may be due to a preference for their larger net positive charge to simultaneously counter the negative electrostatics of the O6 carbonyl groups in two stacked G-quartets.

Note added in proof

A very recent crystal structure (PDB id 3QXR: D. Wei, G. N. Parkinson, S. Neidle, to be published) has shown that potassium and magnesium ions can be bound in loops and clefts of an intramolecular quadruplex, suggesting that some metal ions have important additional stabilizing roles in these structures over and above their presence in quadruplex ion channels.

Acknowledgments This work was supported by a CRUK Programme Grant to S. Neidle.

References

1. M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013-2018.
2. S. Arnott, R. Chandrasekaran, C. M. Marttila, *Biochem. J*. **1974**, *141*, 537-543.
3. J. T. Davies, *Angew. Chem. Int. Ed*. **2004**, *43*, 668-698.
4. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
5. S. Neidle, *Therapeutic Applications of Quadruplex Nucleic Acids*. Academic Press, San Diego, USA, 2011.
6. F. B. Howard, H. T. Miles, *Biochemistry* **1982**, *21*, 6736-6745.
7. P. Balagurumoorthy, S. K. Brahmachari, *J. Biol. Chem*. **1994**, *269*, 21858-21869
8. N. V. Hud, J. Plavec, in *Quadruplex Nucleic Acids*, Eds S. Neidle, S. Balasubramanian, Royal Society of Chemistry, Cambridge, UK, 2006.
9. R. Ida, G. Wu, *J. Amer. Chem. Soc*. **2008**, *130*, 3590-3602.
10. G. Laughlan, A. I. Murchie, D. G. Norman, M. H. Moore, P. C. Moody, D. M. Lilley, B. Luisi, *Science* **1994**, *265*, 520-524; K. Phillips, Z. Dauter, A. I. Murchie, D. M. Lilley, B. Luisi, *J. Mol. Biol*. **1997**, *273*, 171-182.
11. G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876-880.
12. Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263-282.
13. C. Kang, X. Xhang, R. Ratliff, R. Moysis, A. Rich, *Nature* **1992**, *356*, 126-131.
14. P. Schultze, N. V. Hud, F. W. Smith, J. Feigon, *Nucleic Acids Res.* **1994**, *27*, 3018-3028; P. Schultze, F. W. Smith, J. Feigon, *Structure* **1994**, *2*, 221-233.
15. S. M. Haider, G. N. Parkinson, S. Neidle, *J. Mol. Biol*. **2002**, *320*, 189-200.
16. S. Neidle, *Curr. Opin. Struct. Biol* **2009**, *19*, 239-250
17. S. Neidle, G. N. Parkinson, *Biochimie* **2008**, *90*, 1184-1196.
18. H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westerbrook, C. Zardecki, *Acta Crystallogr*. **2002**, D*58*, 899-907.

19. G. R. Clark, P. D. Pytel, C. J. Squire, S. Neidle, *J. Amer. Chem. Soc*. **2003**, *125*, 4066-4067.
20. C. Creze, B. Rinaldi, R. Haser, P. Bouvet, P. Gouet, *Acta Crystallogr*. **2007**, D*63*, 682-688.
21. M. P. H. Lee, G. N. Parkinson, P. Hazel, S. Neidle, *J. Amer. Chem. Soc*. **2007**, *129*, 10106-10107.
22. C. Cáceres, G. Wright, C. Gouyette, G. Parkinson, J. A. Subirana, *Nucleic Acids Res*. **2004**, *32*, 1097-1102.
23. J. Deng, Y. Xiong, M. Sundaralingam, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13665-13670.
24. B. Pan, K. Shi, M. Sundaralingam, *J. Mol. Biol*. **2006**, *363*, 451-459.
25. N. H. Campbell, D. L. Smith, A. P. Reszka, S. Neidle, D. O'Hagan, *Org. Biomol. Chem*. **2011**, *9*, 1328-1331.
26. M. L. Gill, S. A. Strobel, J. P. Loria, *Nucleic Acids Res*. **2006**, *34*, 4506-4514.
27. S. M. Haider, G. N. Parkinson, S. Neidle, *J. Mol. Biol*. **2003**, *326*, 117-125.
28. N. H. Campbell, M. Patel, A. B. Tofa, R. Ghosh, G. N. Parkinson, S. Neidle, *Biochemistry* **2009**, *48*, 1675-1680.
29. M. P. Horvath, S. C. Schultz, *J. Mol. Biol*. **2001**, *310*, 367-377.
30. G. N. Parkinson, R. Ghosh, S. Neidle, *Biochemistry* **2007**, *46*, 2390-2397.
31. G. N. Parkinson, F. Cuenca, S. Neidle, *J. Mol. Biol*. **2008**, *381*, 1145-1156.
32. N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Amer. Chem. Soc*. **2008**, *130*, 6722-6724.
33. G. W. Collie, S. M. Haider, S. Neidle, G. N. Parkinson, *Nucleic Acids Res*. **2010**, *38*, 5569-5580.
34. G. W. Collie, S. Sparapani, G. N. Parkinson, S. Neidle, *J. Amer. Chem. Soc*. **2011**, *133*, 2721-2728.
35. D. Renčiuk, I. Kejnovská, P. Školáková, K. Bednářová, J. Motlová, M. Vorlíčková, *Nucleic Acids Res*. **2009**, *37*, 6625-6634.
36. Y. Xue, Z. Y. Kan, O. Wang, Y. Yao, J. Liu, Y. H. Hao, Z. Tan, *J. Amer. Chem. Soc*. **2007**, *129*, 11185-11191.

Chapter 5
Metal Ion-Mediated DNA-Protein Interactions

Barbara Zambelli, Francesco Musiani, and Stefano Ciurli

Contents

B. Zambelli • F. Musiani
Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Viale Giuseppe Fanin 40, I-40127 Bologna, Italy
e-mail: barbara.zambelli@unibo.it; francesco.musiani@unibo.it

S. Ciurli
Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Viale Giuseppe Fanin 40, I-40127 Bologna, Italy

CERM (Center of Magnetic Resonance), University of Florence, Via Luigi Sacconi 6, I-50019, Sesto Fiorentino (Florence), Italy
e-mail: stefano.ciurli@unibo.it

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_5,

Abstract The dramatic changes in the environmental conditions that organisms encountered during evolution and adaptation to life in specific niches, have influenced intracellular and extracellular metal ion contents and, as a consequence, the cellular ability to sense and utilize different metal ions. This metal-driven differentiation is reflected in the specific panels of metal-responsive transcriptional regulators found in different organisms, which finely tune the intracellular metal ion content and all metal-dependent processes. In order to understand the processes underlying this complex metal homeostasis network, the study of the molecular processes that determine the protein-metal ion recognition, as well as how this event is transduced into a transcriptional output, is necessary. This chapter describes how metal ion binding to specific proteins influences protein interaction with DNA and how this event can influence the fate of genetic expression, leading to specific transcriptional outputs. The features of representative metal-responsive transcriptional regulators, as well as the molecular basis of metal-protein and protein-DNA interactions, are discussed on the basis of the structural information available. An overview of the recent advances in the understanding of how these proteins choose specific metal ions among the intracellular metal ion pool, as well as how they allosterically respond to their effector binding, is given.

Keywords metal-induced conformational changes • metal ion homeostasis • metal ion selectivity • metal ion sensors • transcriptional regulation

1 Introduction: The Biological Significance of Metal Sensing

Living organisms from all kingdoms of life need transition metal ions as essential micronutrients, involved in a variety of biological reactions and estimated to be required in one third of all proteins. Their essentiality, coupled with their limited environmental availability, has prompted all life forms to develop mechanisms for specific and selective metal ion accumulation and utilization. Accordingly, all organisms possess metal homeostasis networks that ensure the availability and the correct localization of metal ions in metallo-proteins and sub-cellular compartments. On the other hand, the intrinsic toxicity of the majority of metal ions demands a tightly regulated intracellular trafficking that maintains intracellular metal ion concentration under the physiological limits and minimizes the amount of free metal ions [1].

Different transition metal ions have different chemical properties, thus generally each metallo-protein uses a specific metal to carry out a precise function.

The metal-specificity of cellular responses indicates that metal-binding proteins are able to choose, in the mixture of metal ions in solution, the correct cofactor for a specific biological function. Overlaps between the metabolisms of different metal ions were demonstrated, implying that the homeostasis of one single metal ion should not be considered in isolation, but always in the context of the overall intracellular metal ion content [2]. This led to the introduction of the term *metallome* to describe the full content, distribution, and concentration of metal ions found in a certain organism [3].

The study of the metallome has extraordinary importance in the understanding of how life evolved and how cellular metabolisms adapted to the presence of diverse micronutrients in the different environments. Indeed, while changes in the metabolic activity can alter metal demand, variations of environmental metal availability can modify the metabolic activity and enzyme production, in order to minimize the consumption of a limited metal ion or to detoxify the poisoning effect of an abundant metal ion. Systems for resistance to toxic concentrations of heavy metal ions, such as Ag^{+}, AsO_2^{-}, AsO_4^{3-}, Cd^{2+}, Co^{2+}, CrO_4^{2-}, Cu^{2+}, Hg^{2+}, Ni^{2+}, Pb^{2+}, TeO_3^{2-}, Tl^{+}, and Zn^{2+} are present essentially in all microorganisms, including the more ancient life forms [4]. On the other hand, metal hyperaccumulator plants, living on and utilizing heavy metal ions in contaminated soils, often represent an example of recent adaptation to extreme metal ion concentrations caused by pollution of the biosphere in the last few decades [5].

A correct balance of metal acquisition and trafficking is essential to establish relationships between prokaryotic and eukaryotic organisms, such as symbiosis and pathogenesis. For example, leguminous symbionts of the *Rhizobium* family that convert atmospheric nitrogen to ammonium, used by plants as essential nitrogen source [6], rely on metal-dependent enzymes such as Fe-nitrogenase [7] and [Ni,Fe]-hydrogenase [8], thus influencing the metal ion content of the rhizosphere and metal ion availability for their host. Similarly, competition for metal ion acquisition is imperative for host-pathogen interactions. Mammalian hosts often use metal withholding, using metal binding proteins such as the iron-binding lactoferrin [9] or the manganese- and zinc-binding calprotectin [10], to prevent pathogen growth by decreasing metal ion concentration to levels incompatible with bacterial life. To compensate, bacteria synthesize proteins and other molecules, such as bacterioferritins and siderophores, able to chelate metal ions and to subtract them from the host systems. In the case of *Brucella abortus*, an intracellular bacterium living inside the macrophages, mutation of a gene encoding for its iron-binding heme-bacterioferritin causes a deficient growth phenotype in iron starvation conditions [11]. In addition, bacteria sense the presence of low environmental levels of metal ions and respond by inducing the expression of a wide variety of toxins and other virulence determinants [12]. This is the case of the gastric pathogen *Helicobacter pylori*, which needs the expression and activation of two metallo-enzymes, Ni-urease and [Ni,Fe]-hydrogenase, in order to colonize its natural environment.

Accordingly, a significant fraction (ca. 25%) of the global repertoire of bacterial transcription factors is involved in controlling metal homeostasis, particularly for copper, iron, and nickel [13]. Infection with *H. pylori* has been epidemiologically

linked to an impaired iron metabolism in hosts, proving that bacterial infection could influence host metal ion homeostasis [14]. Interestingly, the presence of an alternative urease, requiring iron and not nickel, has been reported for *Helicobacter mustelae*, a bacterium that colonizes the stomach mucosa of carnivore animals [15]. This peculiar metal dependence of urease could have developed as an adaptation to the specific host niche, containing a large amount of iron and a relative lower level of nickel, evidencing the selective pressure that metal ions induce on evolution. This is an indication that the study of bacterial metal trafficking presents a significant, still underestimated, potential for new therapeutic strategies for human and animal health.

The crucial players of metal homeostasis networks are specific metal-responsive transcriptional regulators, generally defined as *metal sensors*, which couple specific metal ion binding with a change in their DNA binding affinity and/or specificity, thus translating the concentration of a certain metal ion into a change in genetic expression [16]. In general, specific interactions between proteins and metal ions, driven by the coordination chemistry and geometry of metal binding sites, are propagated away from the specific metal binding site through changes in protein structure and/or dynamics, along the protein backbone, resulting in a modification of the DNA binding affinity of the protein. This event causes repression, derepression, induction or enhancement of transcription of genes codifying metal ion efflux or uptake pumps, membrane permeases, soluble metallo-chaperones, metal accumulator proteins, metal-dependent enzymes, and metal-dependent transcriptional regulators [17]. This produces a finely tuned metabolic response driven by metal ions, including the coordinated control of the entire machinery of metallo-enzyme synthesis and activation, as well as the systems of homeostasis that involve competitive metal ion uptake, intracellular accumulation, and extrusion. In addition, some metal sensors also control metabolic processes other than metal homeostasis, such as oxidative stress resistance, toxin synthesis, and acid adaptation [18].

In the present chapter, we discuss how metal ion binding to specific proteins influences protein interaction with DNA and how this event can influence the fate of genetic expression, leading to specific transcriptional outputs. The features of representative metal-responsive transcriptional regulators, as well as the molecular basis of metal-protein and protein-DNA interactions are discussed on the basis of the structural information available. An overview of the recent advances in understanding how these proteins choose specific metal ions among the intracellular metal ion pool, as well as how they allosterically respond to their effector binding, is given.

2 Structural Insights into Metal Ion-Driven DNA-Protein Interactions

The dramatic changes in the environmental conditions that organisms encountered during evolution and adaptation to live in specific niches, have influenced intracellular and extracellular metal ion contents and, as a consequence, the cellular ability to sense and utilize different metal ions [2]. This metal-driven differentiation is reflected in the specific panels of metal-responsive transcriptional regulators found in different organisms,

which finely tune the intracellular metal ion content and all metal-dependent processes. In order to understand the processes underlying this complex metal homeostasis network, the study of the molecular processes that determine the protein-metal ion recognition, as well as how this event is transduced into a transcriptional output, is necessary.

2.1 Structural Determinants of Metal-Dependent Transcriptional Regulators

Although the number and variety of metal sensor proteins is extremely large, some functional and structural features are typically conserved among the majority of these proteins, while some other traits are conserved only in a limited number of them, allowing a classification of known metal sensors into different groups on the basis of functional and structural properties. In particular, seven main families of metal-sensing transcriptional regulators have been described in bacteria (Table 1) [16].

In the vast majority of cases, metal sensors negatively control mRNA synthesis, preventing RNA polymerase (RNAP) from initiating the transcription at the promoter. In this context, metal ions can bind metal sensors (i) as co-repressors, increasing the affinity of the protein repressor for the DNA operator sequences: this is the case of proteins that regulate genes for metal ion efflux, storage, trafficking, and tolerance; (ii) as inducers, decreasing the affinity of the repressor for the DNA operator sequences, eventually leading to transcriptional activation of genes for membrane uptake systems. In a limited number of cases, the metal-responsive transcriptional regulators positively regulate gene transcription, functioning as activators. In addition, some single examples are known to act as pleiotropic regulators, exerting a dual control, both as activators and repressors, on different promoters, and playing central roles in general transcriptional networks of specific organisms.

All metal sensors generally function as homo-dimers (or dimers of homo-dimers, as in the case of NikR). In general, each monomer presents a modular architecture, being often divided into distinct N- and C-terminal domains, usually functioning as DNA-binding domain (DBD) and dimerization domain, respectively. The N-terminal DBDs generally protrude on both sites of the protein dimer and contact DNA using either a winged helix-turn-helix (wHTH) motif, containing two wings, three α-helices and three β-strands [19] or ribbon-helix-helix (RHH) motifs, made of a short β-strand followed by two α-helices [20].

Regulatory metal binding sites are positioned in strategic positions along the protein backbone, often located across the interface between two monomers or across the interface between the N- and C-terminal domains.

Metal ion binding to regulatory sites can cause either quaternary structural transitions, propagated along the protein backbone to determine reciprocal DBD orientations, or a change of protein dynamics. In the latter case, an increase of protein mobility promotes protein-DNA binding, likely through a favored induced-fit mechanism, by making the DNA-bound protein conformer energetically more accessible. On the other hand, a decreased mobility of the protein backbone drives protein dissociation from DNA.

Table 1 Structural and functional details of the known families of metal-responsive transcriptional regulators.

Family	Other members of the family		Residues involved in functional metal coordination	Metal ion coordination geometry	Available structures	PDB codes
Notes	Name	Sensed metal ions				
ArsR/SmtB						
Representative: ArsR	*Ec*ArsR	As^{3+}, Sb^{3+}	α3: 3Cys	Trigonal	–	–
Biological mechanism:	*Af*ArsR	As^{3+}	α5C: 3Cys	Trigonal	–	
apo-repression	*Sy*SmtB	Zn^{2+}	α5: 2His, Asp, Glu	Tetrahedral	*S. elongatus* pcc 7942	1R1T, 1R22[a2], 1R23[a1], 1SMT
Role of metal ions:	*Sa*CzrA	Co^{2+}, Zn^{2+}	α5: 3His, Asp	Tetrahedral	*S. aureus*	1R1U, 1R1V[a1], 2KJB[b], 2KJC[a1]
induction	*Sy*ZiaR	Zn^{2+}	α3N: 3Cys, His	Unknown	–	–
DNA binding motifs:			α5: 2His, Asp, Glu			
winged helix-turn-helix	*Sa*CadC	Cd^{2+}, Pb^{2+}, Zn^{2+}, Bi^{3+}	α3N: 4Cys	Tetrahedral (Zn,Cd) Trigonal (Pb)	*S. aureus*	1U2W[a1], 3F72[c]
	*Ca*AztR	Cd^{2+}, Pb^{2+}, Zn^{2+}	α3N: 3Cys, His	Tetrahedral (Zn,Cd) Trigonal (Pb)	–	–
	*Ob*BmxR	Cd^{2+}, Zn^{2+}, Ag^{+}, Cu^{+}	α3N: 4Cys	Tetrahedral (Zn,Cd)	–	–
			α5: 2His, Asp, Glu	Trigonal (Cu)		
	*Mt*NmtR	Ni^{2+}, Co^{2+}	α5C: 5His, Asp	Octahedral	–	
	*Mt*KmtR	Ni^{2+}, Co^{2+}	α5-3: 4His, Glu, Asp	Unknown	–	
	*Mt*CmtR	Cd^{2+}, Pb^{2+}	α4C: 3Cys	Trigonal	*M. tuberculosis*	2JSC[a1]

MerR						
Representative: MerR Biological mechanism:	*Pa*MerR *Sf*MerR	Hg^{2+}	3Cys	Trigonal	–	–
weak apo-repression	*Ec*CueR	Cu^{+}, Ag^{+}, Au^{+}	2Cys	Linear	*E. coli*	1Q05[a1], 1Q06[a1], 1Q07[a1]
Role of metal ions: induction	*Ec*ZntR	Zn^{2+}, Pb^{2+}, Cd^{2+}	I (3Cys, O) II (2Cys, His, O)	Tetrahedral	*E. coli*	1Q08[a1], 1Q09[a1], 1Q0A[a1]
DNA binding motifs:	*St*GolS	Au^{+}	Unknown		–	–
helix-turn-helix-β hairpin	*Pa*CadR	Cd^{2+}	Unknown	Unknown	–	–
	*Ae*PbrR	Pb^{2+}	Unknown	Unknown	–	–
	*Sy*CoaR	Co^{2+}	Unknown	Unknown	–	–
	*Ae*CupR	Au^{+}	Unknown	Unknown	–	–
DtxR						
Representative: DtxR Biological mechanism: holo-repression Role of metal ions: co-repression DNA binding motifs: winged helix-turn-helix	*Cd*DtxR	Fe^{2+}	Met, Cys, Glu, His	Octahedral	*C. diphteriae*	1BI0[a1], 1BI1, 1BI2, 1BI3[a1], 1BYM, 1C0W[a3], 1DDN[a4], 1DPR, 1F5T[a4], 1FWZ[a2], 1G3S[c], 1G3T[c], 1G3W[a2], 1G3Y[a2], 1P92, 1QW1, 1QWP, 1XCV[a2], 2DTR[a1], 2QQ9[a2], 2QQA[a2], 2QQB[a2], 2TDX[a2], 3GLX[a2]
	*Mt*IdeR	Fe^{2+}	Met, Cys, Glu, His	Octahedral	*M. tuberculosis*	1B1B[a1], 1FX7[a1], 1U8R[a3], 2ISY[a1], 2ISZ[a3], 2IT0[a3]
	*Bs*MntR	Mn^{2+}	I (His, Asp, 2Glu, 2O)	1-Octahedral	*B. subtilis*	1ON1[a1], 1ON2[a2], 2EV0[a1],
			II (His, 3Glu, 2O)	2-Octahedral		2EV5[a1], 2EV6[a1], 2F5C[a1], 2F5E[a1], 2F5F[a1], 2HYF, 2HYG
	*Sg*ScaR	Mn^{2+}	Glu, Cys, His, Asp	Incomplete (octahedral?)	*S. gordonii*	3HSR, 3HRT[a1], 3HRU[a1]

(continued)

Table 1 (continued)

Family	Other members of the family		Residues involved in functional metal coordination	Metal ion coordination geometry	Available structures	PDB codes
Notes	Name	Sensed metal ions				
Fur						
Representative: Fur	*Pa*Fur	Fe^{2+}	1 (2His, Asp[e], Glu, O)	Ochtaedral	*P. aeruginosa*	1MZB[a1]
Biological mechanism:	*Ec*Fur	Fe^{2+}	Unknown	Unknown	*E. coli*	2FU4[f]
holo-repression[d1]	*Vc*Fur	Fe^{2+}	2His, 2Glu	Tetrahedral	*V. colerae*	2W57[a1]
Role of metal ions: co-repression	*Hp*Fur	Fe^{2+}	A: 3His, Glu, Glu[e] B: 3His, Glu	Ochtaedral Tetrahedral	*H. pylori*	2XIG[a2]
DNA binding motifs:	*Mt*Zur/ FurB	Zn^{2+}	2His, Asp, Cys	Tetrahedral	*M. tuberculosis*	2O03[a1]
winged helix-turn-helix	*Sc*Zur	Zn^{2+}	2His, Asp, Cys	Tetrahedral	*S. coelicolor*	3MWM[a1]
	*Sc*Nur	Ni^{2+}	4His	Square planar	*S. coelicolor*	3EYY[a1]
	*Rl*Mur	Fe^{2+}, Mn^{2+}	Unknown	Unknown	–	
	*Bs*PerR	Fe^{2+}, Mn^{2+}	3His, 2Asp, Cys	Square pyramidal	*B. subtilis*	3F8N[a1]
NikR						
Representative: NikR	*Ec*NikR	Ni^{2+}	3His, Cys	Square planar	*E. coli*	1Q5V, 1Q5Y[a1,g], 2HZA[a1], 2HZV[a3], 3BKF[a1,g], 3BKT[a1,g],
Biological mechanism: holo-repression[d2]						3BKU[g], 3OD2[a1]
Role of metal ions: co-repression	*Ph*NikR	Ni^{2+}	3His, Cys	Square planar	*P. horikoshii*	2BJ3, 2BJ1[a1], 2BJ7[a1], 2BJ8[a1], 2BJ9[a1]
DNA binding motifs: ribbon-helix-helix	*Hp*NikR	Ni^{2+}	3His, Cys	Square planar	*H. pylori*	2CA9, 2CAD[a1], 2CAJ, 2Y3Y[a1], 2WVB[c], 2WVC[c], 2WVD[c], 2WVE[c], 2WVF[a2], 3LGH[a1,g]

CsoR/RcnR						
Representatives: CsoR/RcnR	*Mt*CsoR	Cu^{+}	His, 2Cys	Trigonal	*M. tuberculosis*	2HH7[a1]
Biological mechanism:	*Tt*CsoR	Cu^{+}	Unknown	Unknown	*T. thermophilus*	3AAI
apo-repression	*Bs*CsoR	Cu^{+}	His, 2Cys	Trigonal	–	–
Role of metal ions:	*Ec*RcnR	Ni^{2+}, Co^{2+}	3His, Cys	Octahedral	–	–
induction						
DNA binding motifs:						
None						
CopY						
Representative: CopY	*Eh*CopY	Cu^{+}	4Cys	Unknown	–	–
Biological mechanism:	*Ll*CopR	Cu^{+}	Unknown	Unknown	*L. lactis*	2K4B[f]
Zn^{2+}-repression						
Role of metal ions:						
induction						
DNA binding motifs:						
winged helix-turn-helix						

[a1] metal-bound; [a2] metal-bound and mutant; [a3] metal- and DNA-bound; [a4] metal- and DNA-bound and mutant; [b] DNA-bound; [c] mutant; [d1] holo- and apo-repression for *Hp*Fur; [d2] holo-repression and holo-activation for *Hp*NikR; [e] bidentate; [f] DNA binding domain only; [g] metal binding domain only.

2.2 Metal Ion Coordination Chemistry in Metal Sensor Proteins

Metal sensor proteins are often evolutionary related orthologs of other transcriptional regulators that control gene expression by binding small organic molecules, from lipophilic agents to intermediary metabolites, or amino acids. The difference between metal sensors and metal-independent transcriptional regulators concerns the evolution of the binding site towards a cavity able to accommodate a precise metal ion with its coordination requirements, thus transducing its coordination chemistry into the appropriate biological response [21]. Indeed, metal sensor proteins usually present very conserved DBDs among members of the same family, while the metal-binding regions are usually divergent among related orthologs, reflecting response to different metal ion effectors. On the other hand, members of different families can often sense the same metal ion.

Different organisms produce characteristic sets of metal sensors, and structurally different metal sensors can be found to perform similar functional roles in different bacteria. The key for understanding how this variety of metal-responsive transcriptional regulators can function *in vivo*, binding the correct metal ion cofactor and coupling it to an allosteric change, ultimately switching on/off genetic expression, resides in the coordination geometry of metal binding sites.

2.3 Structural Basis for Metal Ion Selectivity

Metal-responsive control of gene expression occurs in cellular environment, extremely rich of various types and concentrations of metal ions. During the past few years, while the importance of the cellular metallome has emerged, several studies have addressed the fundamental question aimed to identify the general factors allowing metal sensors, and, more generally, metallo-proteins, to selectively respond to a specific metal ion, avoiding adventitious binding and non-physiological response to wrong metal ion cofactors. A straightforward explanation would be that each protein presents higher affinity for the cognate metal ions as compared to all others. However, metal binding affinity alone is a poor predictor of the metal ion selectivity that proteins exert *in vivo*. This selectivity is rather achieved by combining different factors, such as (i) stereo-electronic factors for preferred metal ion coordination geometry, (ii) metal ion availability in the context of the metallome, (iii) metal-induced conformational changes of the protein structure, and (iv) kinetics.

The type and position of protein residues in the metal binding pocket directly influence the metal ion coordination sphere and, consequently, its affinity and selection: the presence of specific donor ligands (S, O or N from protein residues), the shape and the charge of the metal binding cavity, as well as the position of metal binding residues that allow limited coordination geometries, influence the metal ion preferences for a specific metal binding site.

In every family of transcriptional regulators, homologous folds contain different metal binding sites, able to accommodate diverse metal ion effectors. The residues that form the metal binding pockets have been characterized and in many cases metal ion responsiveness could be rationalized on this basis. In some cases, different metal binding sites, responding to different metal ions, are observed in different positions of the same protein fold. In other cases, a single metal binding site, located in a particular position of the protein structure and containing the same type of metal binding residues, selects differently charged metal ions on the basis of the number and geometry of the metal ligands as well as by fine tuning the electrostatic and/or steric local environment of the metal binding pocket.

The metal ion responsiveness of transcriptional regulators depends on the different types of metal ions that are present inside the intracellular milieu. Different intracellular metal ion pools can dictate which metal ion species are sensed by metal-responsive transcriptional regulators in different cells. Therefore, in principle, the same metal sensor can respond to different metal ions in different organisms, or in different subcellular compartments, differing for the intracellular metal ion content, uptake, trafficking, and sequestration. The competition for metal ions among proteins devoted to these different pathways of metal ion metabolism determines the threshold of the metal concentration inducing sensing and gene regulation. Finally, metal ion selectivity also depends on the redox state of the ion, in turn dictated by cellular metabolism.

Metal ions bind specific protein residues often imposing their preferred coordination geometry, derived from stereo-electronic energy stabilization. The latter, for open-shell transition metal ions, derives from d-d splitting of the ligand field and depends also on the metal ion oxidation state. This event could trigger a local conformational change, which is propagated from the binding pocket along the protein backbone, and results in a change of the protein fold and functionality. Binding of the wrong metal ion in the same position is not able to force the protein into a favorable conformation.

Metal ion binding often does not modulate a direct allosteric structural change in the backbone conformation, but rather a change in protein flexibility, in turn affecting protein-DNA interactions. In this sense, it is possible that the kinetics of metal binding to the protein, and in particular which metal ion is presented to the protein at which stage of its biosynthesis and/or function, coupled with the kinetics of ligand exchange, often mediated by interactions with other proteins, determines which is the metal ion/protein association that eventually leads to metabolic response.

3 Structure-Function Relationships in Prokaryotic Metal Ion-Dependent Transcription Factors

3.1 ArsR/SmtB Family

The ArsR/SmtB family of transcriptional regulators is the most represented and widespread family of known metal sensors. More than 500 sequences of genes encoding ArsR/SmtB members were found among all bacterial taxonomy groups, with the majority

of bacterial genomes possessing at least one sequence [22]. The first members of this family, discovered ca. 20 years ago, are the Zn^{2+}-sensor SmtB from *Synechococcus* PCC7942 [23] and the As^{3+}/Sb^{3+}-sensor ArsR from *Escherichia coli* (*Ec*) [24]. Subsequently, several other metal sensors were found (Table 1), functioning to regulate efflux, detoxification and resistance to excess of cytosolic metal ions [16]. These include the Zn^{2+}-sensor ZiaR [25], the $Cd^{2+}/Bi^{3+}/Pb^{2+}$-sensor CadC [26], the $Co^{2+}/Zn^{2+}/Cd^{2+}/Ni^{2+}/Cu^{2+}$-sensor CzrA [27,28], the $Zn^{2+}/Cd^{2+}/Pb^{2+}$-sensor AztR [29], the Ni^{2+}-sensor NmtR [30], the Co^{2+}-sensor KmtR, [22] the Cd^{2+}/Pb^{2+}-sensor CmtR [31], and the $Cd^{2+}/Zn^{2+}/Ag^{+}/Cu^{+}$-sensor BmxR [32]. In general, these proteins function as repressors in the apo form, recognizing and binding an imperfect 12-2-12 inverted repeat on DNA within the promoter region, and inhibiting the concomitant interaction and activity of RNA polymerase. Metal ions function as inducers, driving protein dissociation from DNA [23].

Four members of this family have been structurally characterized using either X-ray crystallography or high resolution NMR spectroscopy, and in particular the structure of *Synechococcus* (*Sy*) SmtB [27], *Staphylococcus aureus* (*Sa*) CzrA [27,33], *Sa*CadC [34] and *Mycobacterium tuberculosis* (*Mt*) CmtR [35] are known (Figure 1). All these proteins are stable dimers and share a general fold comprising at least five α-helices and a two or three stranded β-sheet. Helices α3 and α4 form two winged helix-turn-helix motifs per dimer, responsible for DNA recognition and binding. The other three helices are involved in hydrophobic interactions within the monomer responsible to build and orient the DNA binding motif. In addition, helices α1 and α5 are involved in dimerization.

Structural information on metal binding in the ArsR/SmtB family has also been obtained through site-directed mutagenesis and X-ray absorption spectroscopy studies. Despite their generally similar folds found within the different members of the family, the position and coordination of metal binding sites are very different in different ArsR/SmtB proteins, suggesting that the protein matrix has been adapted during protein evolution to sense a variety of different metal ions (Table 1). In the vast majority of cases, regulatory metal binding sites are symmetrically placed at the dimer interface, forming inter-subunit sites, with each metal ion bound by residues from opposite protomers.

We can distinguish two different types of metal ion binding sites according to their position on the known or predicted secondary structural elements. One of these is close to the N-terminal part, including the N-terminus and helix α3 (named α3N). This site is generally richer in cysteine residues, being the preferred, but not exclusive, place for "soft" metal ions such as Cd^{2+}, Pb^{2+}, and As^{3+} [36]. In particular, *Sa*CadC [34], *Cyanobacterium anabaena* (*Ca*) AztR [29], and *Ec*ArsR [37] sense metal ion binding using this metal binding site. In *Ec*ArsR, an intra-subunit α3 site is found, where As^{3+} and Sb^{3+} are coordinated by the sole three cysteines in helix α3 [37].

The second type of metal binding sites is located in the C-terminal sequence of the proteins, comprising residues from helix α5. These are usually histidine, aspartate, and glutamate, which form a tetrahedral or distorted tetrahedral coordination site, where Zn^{2+}, Ni^{2+}, and Co^{2+} are preferentially bound (α5 site). *Sy*SmtB [27], *Sa*CzrA [27], NmtR [30], and *Acidithiobacillus ferrooxidans* (*Af*) ArsR [38] respond to metal ions bound in this region of the protein structure. In NmtR, this site is extended to include residues from the C-terminus, giving an octahedral binding site that better coordinates Ni^{2+} and Co^{2+} over Zn^{2+} (α5C site) [30]. *Af*ArsR shows an

atypical trigonal site of the α5C type, where As^{3+} is bound to three cysteines deriving from a single monomer [38].

Every protein of the family normally possesses only one of these metal binding sites with regulatory function. However, some family members, such as *Sy*SmtB [27] and *Sa*CadC [39], contain both α3N and α5 metal binding sites but are regulated by only one of them (α5 for *Sy*SmtB and α3N for *Sa*CadC), while the second site (α3N for *Sy*SmtB and α5 for *Sa*CadC) has a possible structural role. In contrast, the *Synechocystis* (*Sy*) ZiaR apparently senses Zn^{2+} *via* both α3N and α5 binding sites [25]. Analogously, both α3N and α5 binding sites are used by *Oscillatoria brevis* (*Ob*) BxmR, but the α3N site shows less selectivity and binds Cd^{2+}, Ag^{+}, Cu^{+}, and Zn^{2+}, while α5 is specific for Zn^{2+} ion, therefore creating an interesting redundancy in its ability to sense Zn^{2+} [32]. The presence of a hybrid site, with residues from both α5 and α3 helices contributing to metal binding, has been evidenced for *Mt*KmtR (site α5-3) [22]. The *Mt*CmtR metal binding site is found in a different region, involving two cysteines in the α4 helix and a third cysteine in the C-terminus of the other subunit (α4C site), forming a symmetrical pair of metal binding sites (Figure 1) [31].

Binding of metal ions to their specific coordination sites within the ArsR/SmtB family drives an allosteric change that releases the transcriptional regulators from DNA [16]. The structural basis of this conformational transition has been inferred from the available crystal structures, and involves the stabilization of a protein conformer with low affinity for DNA, through a H-bonding network that connects the inducer site to the DNA binding site [27]. In the case of *Sy*SmtB, the crystal structures of apo and metal-bound forms show that Zn^{2+} binding to the regulatory α5 site induces a quaternary structure compaction of the homo-dimer, which alters the relative position of one subunit with respect to the other, therefore changing the positions of the DNA recognition sites [27]. Consistently, the crystal structures of apo and Zn^{2+}-bound *Sa*CzrA, together with the solution structures of the Zn^{2+}- and DNA-bound forms of the protein, indicate that the protein adopts a different conformation upon operator binding, with the peripheral domains moving closer to each other [27,33]. Furthermore, Zn^{2+} binding decreases the internal dynamics of the protein backbone, leading to the unavailability, in energetic terms, of the conformer that features high affinity for DNA [27,33]. On the other hand, DNA binding increases the protein motility around the α5 metal binding site [33]. Similarly, in the apo and metal-bound solution structure of the Cd^{2+}-sensor *Mt*CmtR, binding of the metal ion to the α4C site reduces conformational heterogeneity, locking the two subunits of the dimer in a well-defined reciprocal orientation and decreasing the number of protein conformations available for DNA selective interaction [35].

3.2 MerR Family

Pseudomonas aeruginosa (*Pa*) Hg^{2+}-sensor MerR, regulating the expression of bacterial mercuric ion resistance, was the first protein for which the term "metallo-regulatory" was used [40]. While the majority of metal-dependent transcriptional regulators function as repressors of transcription, proteins belonging to the MerR

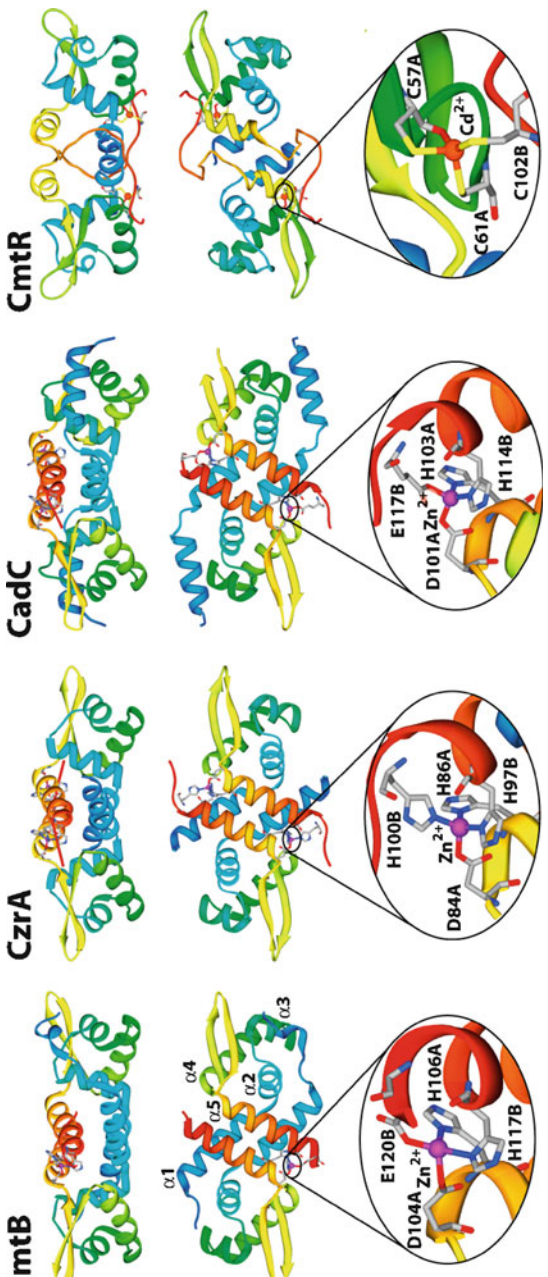


Figure 1 Ribbon diagrams (top and middle panels) and detail of the metal binding sites (bottom panels) of the proteins belonging to the ArsR/SmtB family and for which the structure of the protein is available in the Protein Data Bank (PDB). In all cases, the protein depicted is the metal-bound form of the molecule (PDB code 1R23 for *Sy*SmtB; 1R1V for *Sa*CzrA; 1U2W for *Sa*CadC; and 2JCS for *Mt*CmtR). Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini. In the middle panels the ribbons are rotated by 90° around the horizontal axis with respect to the representation in the top panels. In the bottom panels, the metal ions are depicted as spheres and the residues bound to the metal ions as "sticks". Atoms are colored according to their type.

family mainly function as activators in their holo-forms, inducing the synthesis of proteins for metal efflux and detoxification [41]. These proteins bind their target promoters both in their apo and holo forms, recognizing an unusually long spacer (19–20 bp versus the 16–18 bp found in most prokaryotic promoters [42]) between the –35 and the –10 sequences. This unusual promoter sequence is suboptimal, by itself, for RNAP recognition. Apo-MerR proteins are able to recruit RNAP, but their presence creates a distortion of the DNA structure, thus preventing the formation of a productive complex for transcriptional initiation and contributing to the transcriptional silencing [43]. Metal ion binding functions as inducer of an allosteric change of DNA conformation that realigns the –35 and –10 sequences, so that RNAP can productively contact the promoter sequence and initiate transcription [44].

The MerR family comprises proteins able to sense different metal ions (Table 1), including the Zn^{2+}/Pb^{2+}/Cd^{2+}-sensor *Ec*ZntR [45], the Co^{2+}-sensor *Synechocystis* (*Sy*) PCC 6803 CoaR, the Pb^{2+}-sensor *Alcaligenes eutrophus* (*Ae*) PbrR [46], the Cu^{+}/Ag^{+}/Au^{+}-sensor *Ec*CueR [47], the Cd^{2+}-sensor *Pseudomonas aeruginosa* (*Pa*) CadR [48], the Au^{+}-sensor *Salmonella enterica* (*Se*) GolS [49], and the Au^{+}-sensor *Ae*CupR [50]. Among them, *Ec*ZntR and *Ec*CueR have been structurally characterized by crystallography (Figure 2) [47]. Both these proteins are homo-dimers and present highly homologous folds, with an extended conformation composed by a long dimerization helix from which distinct N-terminal and C-terminal domains protrude. The N-terminal winged helical DNA-binding domain is comprised of a helix-turn-helix β-hairpin structure, while the C-terminal domain (metal binding domain, MDB) is made of a loop containing the metal binding residues.

In both structures, the metal binding sites are positioned in the region close to the dimerization interface. Contrarily to the ArsR/SmtB family regulators, which evolved metal selectivity by positioning metal binding sites in different regions of the protein, the MerR family regulators discriminate among a wide range of different metal ions by finely changing the identity and exact position of the metal binding residues and their environment. In particular, the crystal structure of Zn^{2+}-*Ec*ZntR shows two Zn^{2+} ions bound to two cysteine/histidine residues (Cys^{114} and Cys^{124} for Zn_{I} and Cys^{115} and His^{119} for Zn_{II}) from the MBD of one subunit, and to another cysteine residue (Cys^{79}) from the N-terminus of the dimerization helix in the opposite subunit [47], bridging the two ions. A phosphate ion binds the two Zn^{2+} ions in a bidentate mode. Although the structure of Hg^{2+}-sensor MerR is not available, NMR studies indicate a trigonal planar coordination geometry for Hg^{2+} ions, involving three cysteine residues, corresponding to the cysteine residues that coordinate metal ions in *Ec*ZtnR [51].

The same residues are conserved in *Ae*PbrR and *Pa*CadR [47], suggesting that metal ion binding occurs in similar positions in different proteins responding to different divalent metal ions. Differently, in the Cu^{+}/Au^{+}/Ag^{+}-sensor *Ec*CueR, the Cys^{79} residue is substituted by a serine (Ser^{77}), which abrogates one possible metal ion coordination position and contributes to a buried and hydrophobic environment more adapt for a monovalent ion. In this case, Cu^{+} binds in a linear dithiolate coordination involving the conserved Cys^{112} and Cys^{120} residues [47]. The diverse identity of this residue positioned at the N-terminus of the dimerization helix and close to the MDB in the protein quaternary structure, gives a possible explanation for

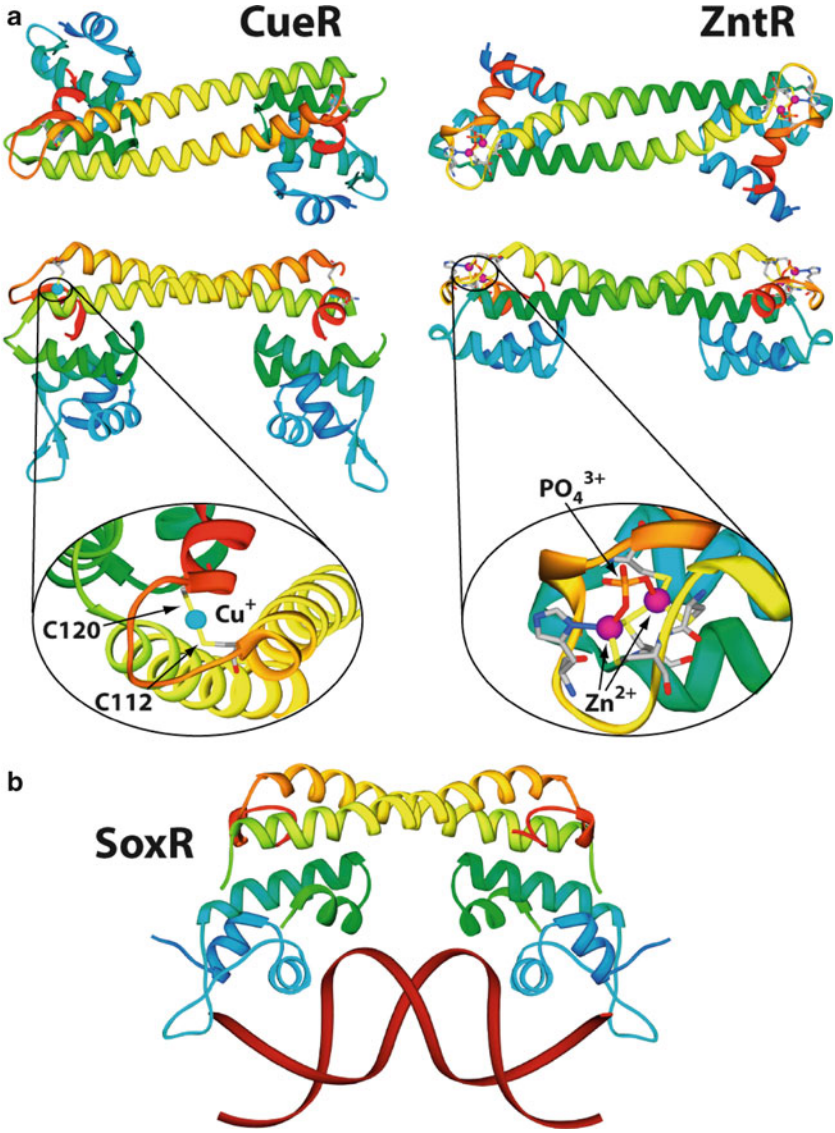


Figure 2 (**a**) Ribbon diagrams (top and middle panels) and details of the metal binding sites (bottom panels) of the proteins belonging to the MerR family and for which the structure of the protein is available in the PDB. In all cases, the protein depicted is the metal-bound form of the molecule (PDB code 1Q05 for *Ec*CueR and 1Q08 for *Ec*ZntR). In the middle panels, the ribbons are rotated by 90° around the horizontal axis with respect to the representation in the top panels. In the bottom panels, the metal ions are depicted as spheres and the residues bound to the metal ions as "sticks". Atoms are colored according to their type. (**b**) Ribbon diagram of *Ec*SoxR (PDB code: 2ZHG) crystallized in complex with DNA. Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini. DNA ribbons are depicted in dark red.

protein selectivity of divalent over monovalent metal ions and *vice versa*. Accordingly, a multiple sequence alignment shows that a serine residue is found in this position in all MerR homologues that recognize +1 charged metal ions, while this place is occupied by cysteine when the protein senses +2 charged metal ions, which often prefer a more covalent-type binding mode [47].

In contrast to *Ec*CueR, which is a poor discriminator between different +1 charged metal ions, *Se*GolS and *Ae*CupR are extremely selective for Au^+. Multiple sequence analysis revealed that *Se*GolS and *Ae*CupR conserve the counterparts of Cys^{112} and Cys^{120} residues in the MBD as well as of Ser^{77} in the N-terminal domain, therefore excluding that first shell metal binding coordination by itself is responsible for the different selectivity of these proteins [50,52]. Mutational analysis, involving the replacement of the metal binding loop of *Se*GolS with the same region of *Ec*CueR, revealed that the precise sequence change of the metal binding loop is sufficient to mimic the Cu^+-sensor in its binding promiscuity to Cu^+, Ag^+, and Au^+ [52]. The involvement of the metal binding loop in determining the metal selective behavior of these proteins is further confirmed by the observation that *Se*GolS and *Ae*CupR share several similarities in this sequence, which are not conserved in *Ec*CueR [50,52]. Although for the other metal-responsive regulators of the MerR family very little biochemical and structural data is available, their sequence alignments show extensive similarities between their primary structure and the sequences of *Ec*ZntR and *Ec*CueR, suggesting a common fold and metal ion coordination.

So far, no structure of DNA-bound protein has been established for the metal sensors of the MerR family. However, this family also includes several non-metal responsive transcriptional regulators, which are involved in the defence against oxidative stress and drug resistance, and bind different effectors in the C-terminal domain. As expected, these proteins show closely related N-terminal DBDs but different C-terminal ligand binding domains, in agreement with the different inducers that are recognized by different proteins [53]. In particular, the structures of *Bacillus subtilis* (*Bs*) BmrR bound to a lipophilic drug (tetraphenylphosphonium) [54], of *Bs*MntA, which lacks any C-terminal domain thus being a constitutive activator [55], and of the oxidative sensor *Ec*SoxR, containing a [Fe-S] cluster bound to four cysteine residues in the MBD [56], have been solved in the presence of their respective DNA operator. When bound to DNA, *Bs*BmrR causes a significant bending of the DNA molecule and a peculiar base pair breaking and sliding of the central base pair step, and the operator bunches up in the middle. This conformation is stabilized by the protein residues that interact with the phosphate backbone of DNA. The results of this interaction twists the DNA molecule, and realigns the –10 and –35 sequences on the same face of the DNA double helix, in a position accessible for RNAP [54]. The structures of *Bs*MntA, obtained both in the absence and in the presence of DNA, shed light on the difference between DNA-free and DNA-bound structures, and revealed a rotation around a hinge region that lies at the N-terminus of the dimerization helix, which changes the relative position of DBDs [55]. Finally, the structure of the DNA-bound *Ec*SoxR presents a network of inter-subunit hydrogen-bonding interactions between the DBDs and the MBDs, proposed to mediate the signal transduction between the two regions of the protein that drives the protein conformational change and consequent transcriptional activation [56].

3.3 DtxR Family

The DtxR family includes the metal-responsive transcriptional regulators that specifically sense Fe^{2+} and/or Mn^{2+}. The prototype of Fe^{2+}-sensors is *Corynebacterium diphteriae* (*Cd*) DtxR, while the paradigm of Mn^{2+}-sensors is *Bs*MntR [16]. These proteins function as metal-dependent repressors of transcription, with metal ions acting as co-repressors by promoting protein interaction with palindromic DNA regions located upstream of the initiation start of specific metal-regulated genes [57,58]. This event represses the synthesis of proteins for metal ion acquisition and storage, as well as of virulence factors. In the case of *Corynebacterium glutamicum* (*Cg*) DtxR, and of the Fe^{2+} responsive *Mt*IdeR, the DtxR ortholog in *M. tuberculosis*, additional roles as holo-activators of genes involved in iron storage and in protection against oxidative stress has been proposed [57,59], suggesting that these proteins are central regulators in the iron-dependent gene transcription network.

Fe^{2+}-responsive *Cd*DtxR and *Mt*IdeR have been extensively characterized from the biochemical, biophysical, and structural point of view. Both these proteins are highly specific for Fe^{2+} *in vivo*, although this selectivity is not maintained *in vitro*, where *Cd*DtxR similarly responds to Fe^{2+}, Ni^{2+} and Co^{2+} [60], while *Mt*IdeR responds to Fe^{2+}, Ni^{2+}, Co^{2+}, Cd^{2+}, Mn^{2+}, and Zn^{2+} [61]. Several crystal structures were initially reported for both proteins in the absence and presence of a variety of metal ions and DNA, even though the structure of the protein in the presence of the physiological cofactor Fe^{2+} has never been obtained [62–70] (Figure 3). These regulators contain a N-terminal winged helix DNA binding domain, followed by a helical dimerization domain, and a C-terminal SH3-like domain, rare in prokaryotes, that was proposed to enhance the DNA-binding activity through stabilization of inter- or intra-subunit contacts [71,72].

The structural studies evidenced the presence, in both *Cd*DtxR and *Mt*IdeR, of two different metal binding sites (see Table 1). Site 1, called "ancillary" site, is proposed to stabilize the dimeric conformation of the protein [73]. In this site, the metal ions are placed between the N- and C-terminal domains, in a distorted trigonal bipyramidal geometry, formed by three residues (two His and one Glu) from the DBD and by two residues of the SH3 domain (one Glu and one Gln). Metal site 2 is the so-called "primary" metal binding site, responsible for protein activation [73], entirely located in the N-terminal domain in an octahedral coordination geometry involving four conserved residues from the DBD (Met, Cys, Glu, His), one of them (Met^{10}) located in the N-terminal helix. In the case of the crystal structure of *Mt*IdeR a third metal binding site has been observed, whose physiological relevance is yet unknown.

In the apo form, *Cd*DtxR and *Mt*IdeR exist in a weakly associated dimeric state, showing a partially disordered region in the DBD that involves the N-terminal helix [74]. This disorder is thought to inhibit protein interaction with DNA, which apparently involves two dimers interacting on opposite sides of the DNA molecule [66]. Metal binding, involving Met10, apparently induces a more folded conformation and stabilizes the dimeric species [60,75]. Metal binding also seems to induce a domain closure of the DBD with respect to the dimerization domain, moving the DNA binding helix in an optimized position for DNA recognition in two consecutive major grooves of DNA [64,68].

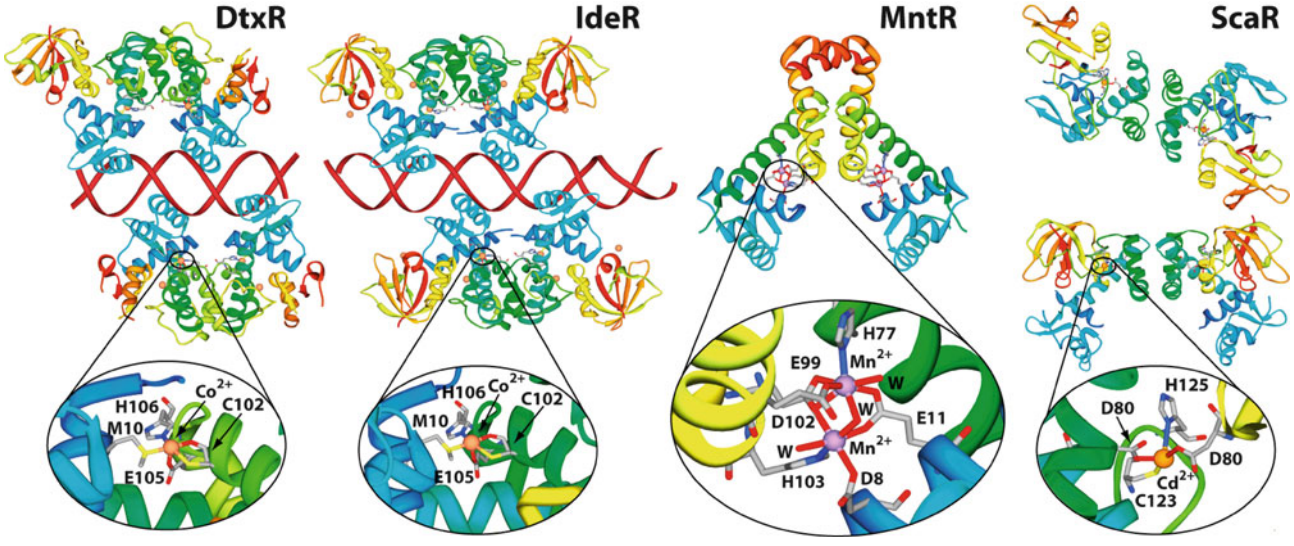


Figure 3 Ribbon diagrams (top panels) and details of the metal binding sites (bottom panels) of the proteins belonging to the DtxR family and for which the structure of the protein is available in the PDB. In all cases, the protein depicted is the metal-bound form of the molecule (PDB code 1C0W for *Cd*DtxR; 1U8R for *Mt*IdeR; 1ON1 for *Bs*Mnt; and 3HRT for *Sg*ScaR). Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini; when present DNA is colored in dark red. In the middle panel of ScaR, the ribbons are rotated by 90° around the horizontal axis with respect to the representation in the top panel. In the bottom panels, the metal ions are depicted as spheres and the residues bound to the metal ions as "sticks". Atoms are colored according to their type.

*Bs*MntR is the prototype of the Mn^{2+}-responsive transcriptional repressor of the DtxR family. Differently from the *Cd*DtxR and *Mt*IdeR repressors, *Bs*MntR exists in solution as a stable dimer, independently of the presence of bound metal ions [76]. This protein is highly specific for Mn^{2+} and Cd^{2+} over other divalent metal ions such as Mg^{2+}, Ca^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, and Zn^{2+} [76]. Structural studies on *Bs*MntR revealed that the overall architecture is conserved between this protein and its Fe^{2+}-responsive homologues, although this protein lacks the C-terminal SH3-like domain [77,78]. Despite this similarity, the location and the coordination of the metal binding sites are considerably different. Although the overall stoichiometry of two Mn^{2+} ions per monomer is maintained, *Bs*MntR does not feature any occupied primary or ancillary sites. On the other hand, the structure of the holo-protein revealed a binuclear metal binding site located at the interface between the N-terminal DBD and the dimerization domain. Two Mn^{2+} ions were found, separated by 4.4 Å, in a distorted octahedral coordination geometry and bridged by a glutamate residue [78].

The comparison of the structures of the metal-bound proteins versus the apo-repressor suggests that binding of metal ions induces a closure of the DBDs with respect to the protein dimer, partially caused by metal binding to specific residues in the inter-domain helix [79]. Moreover, as in the case of DtxR, Mn^{2+} binding to the N-terminal α-helix drives an important conformational change in this region, which is reflected in the protein ability to interact with DNA [79]. Metal ion binding to the protein also results in a global decrease of protein mobility, probably functional to induce a correct recognition of a specific DNA operator [80]. The crystal structure of apo-*Ec*MntR is consistent with these conclusions [81].

Recently, the crystal structure of *Streptococcus gordonii* (*Sg*) ScaR, a Mn^{2+}-responsive protein belonging to the DtxR family, has been reported [82]. The general fold of *Sg*ScaR is more similar to DtxR-like proteins, conserving the three-domain structure and the residues building up the primary site in *Cd*DtxR and *Mt*IdeR. On the other hand, the residues corresponding to the ancillary sites are absent in *Sg*ScaR, in agreement with the stable dimeric behavior shown by this protein in solution. Surprisingly, the holo-structures, obtained in the presence of Cd^{2+} and Zn^{2+}, presented an empty primary site, while revealing a novel metal binding site, located at the interface between the C-terminal domain and the dimerization domain. However, the physiological significance of this metal binding site, possibly caused by an artefact of protein crystallization conditions, remains unclear and should be further investigated in solution [82].

3.4 Fur Family

The Fur family of metal-responsive transcriptional regulators is widespread, being present in the genome of all Gram negative bacteria with the only exception of the plant symbiont *Rhizobium* and some close organisms [83]. The first discovered member and the archetype of the family is the *E. coli* Fe^{2+}-dependent transcriptional repressor [84]. Fur has been also studied from other bacteria, including

P. aeruginosa (*Pa*) [85], *H. pylori* (*Hp*) [86], and *Vibrio colerae* (*Vc*) [87]. The paradigmatic mechanism that explains Fur functioning was initially described for *Ec*Fur [88]: these proteins generally function as holo-repressors, and, in the presence of the cognate metal ions, bind to DNA operators in the target core promoters, thus masking the RNAP binding sequences and preventing transcriptional initiation. Therefore, metal ions act as co-repressor of the gene transcription.

Genetic studies revealed that Fur proteins might also exert positive regulation on some genes that should be over-expressed in the presence of Fe^{2+} ions. However, this apparent activation is not directly mediated by Fur, but it is often mediated by small regulatory RNAs that determine the fate of mRNA molecules at the post-translational level [89]. Distinctively, in addition to the paradigmatic Fe^{2+}-dependent repression occurring on the promoters of holo-Fur repressed genes, *Hp*Fur carries out repression in the absence of bound metal ions, functioning therefore as a repressor on apo-Fur-repressed genes, with Fe^{2+} ions acting as inducers of transcription [90]. By this mechanism, the same input information (high or low concentration of Fe^{2+} ions in the cytoplasm) is translated into two opposite transcriptional outputs in different genes. The wide range and importance of processes directly and indirectly controlled by Fur indicate that this protein is a global transcriptional regulator, involved in several metabolic responses and genetic networks. Accordingly, *Ec*Fur and *Hp*Fur directly or indirectly regulate over 90 and 200 genes respectively, involved not only in iron homeostasis, but also in oxidative stress response, in motility, and several other cellular activities essential for cell growth and environmental adaptation [83,91].

Several Fur homologues that respond to metal ions other than Fe^{2+} have been discovered and characterized. These include the Zn^{2+}-sensor *Mt*Zur [92], the Mn^{2+}/Fe^{2+}-sensor *Rhizobium leguminosarum* (*Rl*) Mur [93], and the Ni^{2+}-sensor *Streptomyces coelicolor* (*Sc*) Nur [94]. The oxygen peroxide-sensor *Bs*PerR [95] can also be included in this list, as it uses metal ions, such as Mn^{2+} or Fe^{2+}, to sense the presence of H_2O_2 and transduce the signal in a transcriptional output. Crystallographic structures are available for *Pa*Fur [96], *Mt*Zur [92], *Bs*PerR [97,98], *Sc*Nur [99], *Vc*Fur [100], and *Hp*Fur [101]. All these proteins share a general dimeric fold, with each monomer containing a N-terminal winged helix-turn-helix DNA binding domain connected to a C-terminal dimerization domain through a short linker (Figure 4).

The structures of different Fur and Fur-like regulators revealed the presence of three different metal binding sites not always conserved in different homologues. The recently reported structure of *Hp*Fur, containing all three metal sites (named S1, S2, and S3), can be used as a reference for the present discussion (Figure 4). S1 corresponds to a tetrahedral structural Zn^{2+}-binding site, important for protein stability and dimerization, found in the C-terminal dimerization domain and usually made of four cysteine thiolates. This site is conserved in the structures of *Mt*Zur [92], *Bs*PerR [97], and *Hp*Fur [101]. In *Ec*Fur, two cysteines and two histidines are the metal coordination residues [102]. S3 in *Hp*Fur has also been observed in the structures of *Pa*Fur and *Vc*Fur: the residues involved in this binding site (two histidines, one glutamate, and an aspartate) are conserved, even though the metal coordination geometry can differ in different protein orthologs. The physiological relevance of this metal site is still controversial [101,103]. Finally, S2 of *Hp*Fur

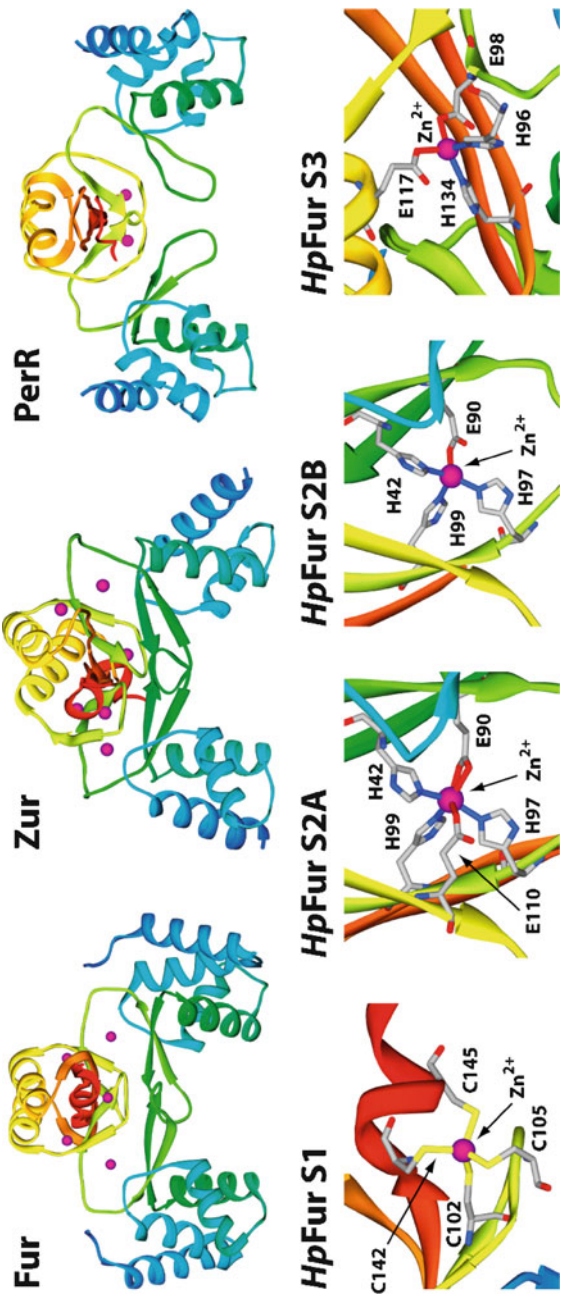


Figure 4 Ribbon diagrams of the proteins belonging to the Fur family and for which the structure of the protein is available in the PDB (top panels) and detail of *Hp*Fur the metal binding sites (bottom panels). In all cases, the protein depicted is the metal bound form of the molecule (PDB code 2XIG for *Hp*Fur, 3MWM for ScZur, and 3F8N for *Bs*PerR). Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini. In the bottom panels, the metal ions are depicted as spheres and the residues bound to the metal ions as "sticks". Atoms are colored according to their type.

contains a "regulatory" metal ion, is located at the interface between the N- and C-terminal domains, and is generally considered responsible for driving protein interaction with its cognate DNA.

The presence of metal ions in the indicated location was reported for all the Fur proteins so far characterized, although with variability of the coordination environment depending on the protein and the metal ion effector. The binding of a metal ion in this region would bring residues from the N- and C-terminal domain in close proximity, thus orienting the relative positions of the two DBDs and stabilizing a conformation with high affinity for DNA [92]. A similar allosteric mechanism has been proposed for *Bs*PerR, for which the presence of a Fe^{2+}/Mn^{2+} ion in the regulatory site would stabilize a conformer with high affinity for DNA, and, at the same time, would create an open coordination site for a H_2O_2 molecule. The latter would oxidize the metal ion, creating a Fe^{3+} species able to catalyze the oxidation reaction of one histidine residue either in the DNA binding domain (His^{37}) or in the dimerization domain (His^{91}) of the protein. This unusual protein modification would subsequently decrease protein affinity for DNA, thus causing dissociating of the repressor and activation of the downstream genes [98].

Another peculiar Fur family homologue, of unknown structure, is the Fe^{2+}-sensor *Bradyrhizobium japonicum* (*Bj*) Irr, which functions as an apo-repressor at low Fe^{2+}-concentrations. When high concentrations of Fe^{2+} ions are available, the heme cofactor binds to a short N-terminal motif and promotes dissociation from DNA as well as protein degradation with an unknown mechanism [104].

3.5 NikR

NikR is a transcription factor that regulates the expression of genes coding for proteins involved in nickel metabolism. It is a highly homologous protein, found in ca. 30 species of bacteria and archea. In particular, the Ni^{2+}-bound NikR from *E. coli* (*Ec*NikR) binds to DNA and represses the transcription of the *nikABCDE* operon, which codes for a specific Ni^{2+} membrane uptake ABC transporter [105,106]. On the other hand, NikR from *H. pylori* (*Hp*NikR) is a pleiotropic regulator of several genes, for example acting as a nickel-dependent repressor of *Hp*NikR itself and of the Ni^{2+} permease NixA, as well as a nickel-dependent promoter of the pathogenic factor urease [107]. In *H. pylori*, the disruption of the Ni^{2+}-chaperone network, involved in metal ion delivery to Ni^{2+}-containing urease, results in a deregulation of the Ni^{2+}-dependent transcriptional regulator NikR [108], suggesting that intracellular metal ion availability influences the activity of this metal sensor, in a mechanism that could be operative in other cases as well.

Several crystal structures of NikR [109–114,159] have established that this protein is a homo-tetramer, made of a dimer of dimers, constituted by three domains. One domain is the central metal-binding domain, made of the C-terminal portion of the protein responsible for tetramerization. This domain hosts four regulatory metal binding sites symmetrically located at the tetramerization interface, where Ni^{2+} ions

bind three fully conserved His and one Cys residues in a square planar coordination geometry stabilized by the ligand field of the open shell d^8 metal ion (Figure 5). An extensive hydrogen-bonding network around the nickel binding sites is thought to couple the metal-specific stereo-electronic preferences of Ni^{2+} with the structural or dynamic properties of the protein, rendering it able to bind DNA. This MBD is flanked by two peripheral DNA-binding domains, separated by unstructured linkers. Each DBD is made of the dimerization of the N-terminal portion of the protein and feature a ribbon-helix-helix motif for DNA binding typical of prokaryotic transcription factors [115] in which two anti-parallel N-terminal β-strands from opposite protomers make a two-stranded anti-parallel β-sheet that contacts the major groove in one-half-site of a two-fold symmetric DNA operator (Figure 5).

Three distinct conformations of NikR have been observed in the solid-state: open, *trans*, and *cis*, characterized by the position of the DBDs with respect to the central MBD (Figure 5). The crystal structure of the *Ec*NikR-DNA complex has shown that the *cis* conformation of the Ni^{2+}-bound protein is able to bind DNA (Figure 5) [111]. The metal specific Ni^{2+}-dependent NikR-DNA interaction is supported by electrophoretic mobility shift assays, fluorescence anisotropy, and DNaseI footprinting studies, as well as calorimetric titrations [106,116,117,160,161].

The effect of Ni^{2+} on the reactivity of NikR toward DNA binding has been recently investigated using atomistic molecular dynamics simulations complemented by solution NMR studies [118]. The results indicate that both apo- and holo-NikR are present in solution as an ensemble of inter-converting structures spanning the conformational space from *cis* to open to *trans* forms. The presence of bound Ni^{2+} ions does not induce, by itself, the stabilization of the *cis* conformation competent for DNA binding, but it appears to unlock the motion of the DBDs with respect to the MDB, supporting the view that the likely mechanism of interaction of the protein with its operator DNA sequence involves a selection of the correct conformation coupled with an induced fit mechanism facilitated by the presence of bound Ni^{2+}.

3.6 CsoR/RcnR Family

The dual nature of essential and toxic element applies to copper. A sensor for Cu^+, the major form of copper in the cell, is CsoR, widely spread in the bacterial world [119]. In general, the apo-form of this transcriptional regulator binds to the promoter region of a copper-sensitive operon, acting as a transcriptional repressor. When copper levels are too high, Cu^+ binds to the protein, inducing its release from DNA and allowing expression of the proteins devoted to decrease intracellular copper concentration.

Two crystal structures are available: one is the Cu^+-bound CsoR from *M. tuberculosis* (*Mt*CsoR) [120], while the other is the apo-form of *Thermus thermophilus* CsoR (*Tt*CsoR) [121] (Figure 6).

*Mt*CsoR is a homo-dimeric protein, containing a N-terminal core with a four-helix bundle fold, and a short C-terminal helix stacked against the basis of the molecule. Apo-*Tt*CsoR is a dimer of dimers in the solid state, with each dimer being similar to

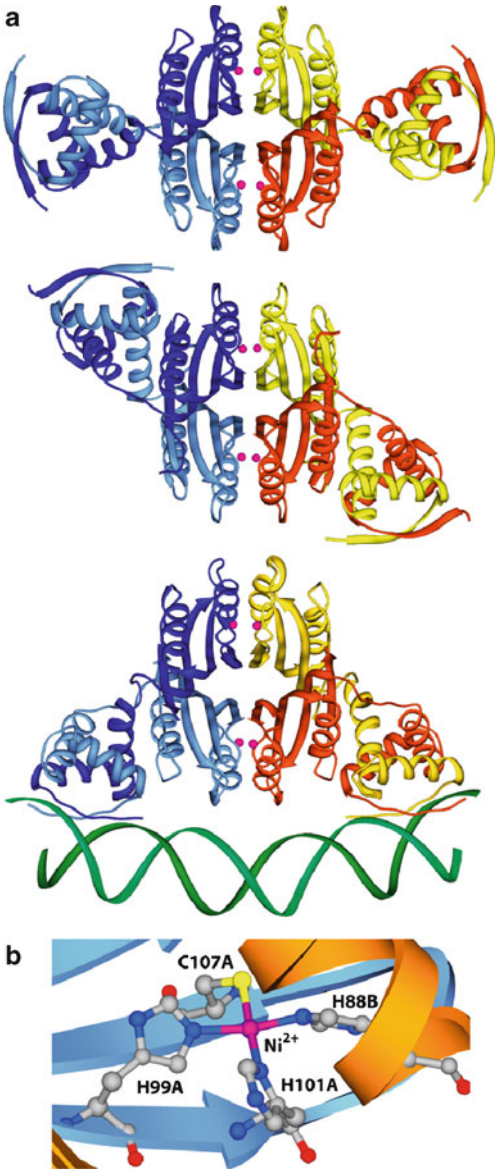


Figure 5 (**a**) Ribbon diagrams of NikR in (top panel) open, (middle panel) *trans*, and (bottom panel) *cis* conformation (PDB: 2HZA, 2CA9, and 2HZV, respectively). (**b**) Structure of the metal binding sites found in *Hp*NikR (2Y3Y). Ribbons are colored according to their monomer chain and atoms are colored according to their type.

the *Mt*CsoR architecture. In *Mt*CsoR, two symmetric metal binding sites are observed, in which each Cu^{+} ion bridges two subunits, with a trigonal planar coordination geometry made of Cys^{36}, $Cys^{65'}$ and $His^{61'}$. In the case of *Tt*CsoR, while residues Cys^{36} and $His^{61'}$ are conserved, $Cys^{65'}$is absent: two histidine residues, His^{5} and His^{70} are found in the close vicinity of the putatively conserved metal binding site, and could be involved

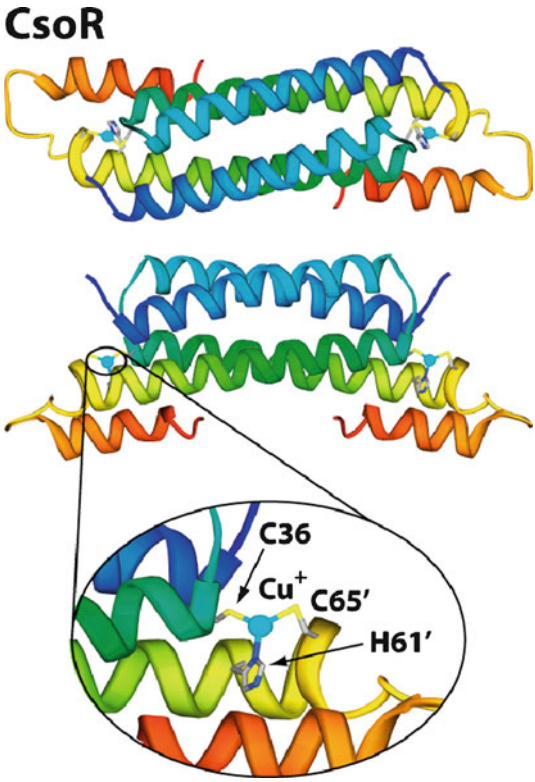


Figure 6 Ribbon diagrams (top and middle panels) and detail of the metal binding site (bottom panel) of *Mt*CsoR (PDB code: 2HH7), representative of the CsoR/RcnR family. In the middle panel, the ribbons are rotated by 90° around the horizontal axis with respect to the representation in the top panel. Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini. In the bottom panel, the metal ion is depicted as a sphere and the residues bound to the metal ion as "sticks". Atoms are colored according to their type.

in metal binding. No clear metal-induced allosteric changes, possibly modulating protein-DNA interactions, can be deduced on the basis of these structural studies.

B. subtilis CsoR, another member of the family, has been shown to bind Cu^{+} with high affinity in a trigonal S_2N coordination site [122]. This protein is also able to bind Ni^{2+}, Zn^{2+}, and Co^{2+} with high affinity, but adopts different metal coordination geometries in each case, with non-physiological responses, suggesting that, in this case, metal-ligand coordination geometry drives physiological metal selectivity rather than metal binding thermodynamics [21].

RcnR is a Ni^{2+} and Co^{2+} sensor that regulates the expression of a nickel and cobalt efflux protein [123,124], RcnA, a membrane permease [125]. No structural information is available for RcnR. This protein shares very low sequence similarity with CsoR, but it has been proposed to be an all α-helical protein with a fold similar to that of CsoR [124]. RcnR proteins possess a conserved His-Cys-His-His metal binding motif. Ni^{2+} and Co^{2+} binding experiments indicate nanomolar affinity for both ions, while data from electronic, electron paramagnetic resonance, and X-ray absorption spectroscopies reveal that RcnR binds Ni^{2+} or Co^{2+} in a six-coordinated octahedral geometry that includes all four conserved metal binding residues as well as, possibly, the N-terminal $–NH_2$ group and a peptide amide group [124].

3.7 CopY Family

CopY is a Cu^+ responsive transcription factor, present in the Gram-positive phylum of Firmicutes [16]. The first characterized protein of the family has been *Enterococcus hirae* (*Eh*) CopY, which regulates the Cu^+ uptake and efflux as well as the synthesis of *Eh*CopZ, a Cu^+ chaperone [126]. The generally accepted mechanism of action involves the Zn^{2+}-bound *Eh*CopY, which binds the cognate DNA operator acting as repressor of transcription [127]. Cu^+-*Eh*CopZ then delivers Cu^+ ion to the repressor, which, in this form, releases the gene transcription from the promoter [128]. The protein has been reported to function as a dimer, and spectroscopic studies revealed that each monomer binds one Cu^+ ion to four conserved cysteine residues located at the C-terminus [128].

The solution structure of the monomeric N-terminal domain of *Lactococcus lactis* CopR, a member of the CopY family, revealed that this protein region contains a winged helix-turn-helix motif, involved in the DNA binding [129]. Structural studies on the integral protein will be necessary to understand how Cu^+ can modulate the DNA binding properties of the protein, leading to a specific transcriptional output.

3.8 Other Bacterial Metal Sensor Systems

In addition to the seven major families of metal-sensors described above, we can distinguish some individual metal sensors belonging to other, structurally distinct, families of regulators not primarily involved in metal-dependent transcriptional regulation. An example is the structurally characterized molybdenum-sensor *Ec*ModE from the LysR family [130]. This family is the largest group of transcriptional regulators in bacteria. LysR-like proteins are usually dimers of dimers with each monomer made of a conserved N-terminal winged helical DNA binding domain and a divergent C-terminal ligand binding domain [131]. These proteins recognize a two-fold symmetric operator sequence with half sites separated by 10–15 base pairs. In contrast with other LysR-like proteins, *Ec*ModE apparently functions as a single dimer, binding the oxyanion form of Mo(VI) and recognizing the operator sequence in the promoter DNA, negatively regulating the transcription of genes encoding MoO_4^{2-} membrane transporters [132]. Crystal structures of *Ec*ModE in the apo [133] and molybdate-bound [134] forms (Figure 7) reveal that the metal ion is bound at the C-terminal domain through hydrogen bonding interactions of the oxyanion oxygens with nearby protein residues. This causes an allosteric change of the relative orientations of the DBDs, stabilizing a conformation competent for high affinity DNA binding.

In order to monitor changes in their extracellular environment, bacteria have evolved multiple two-components histidine kinase transcriptional regulators [135]. Each two-component system consists of a sensor protein-histidine kinase, often associated to the plasma membrane, and a response regulator [136]. Generally, the sensor, in dimeric form, detects an external stimulus using the periplasmic domains and transmits it to the cytoplasmatic domains, containing the ATP binding pockets and

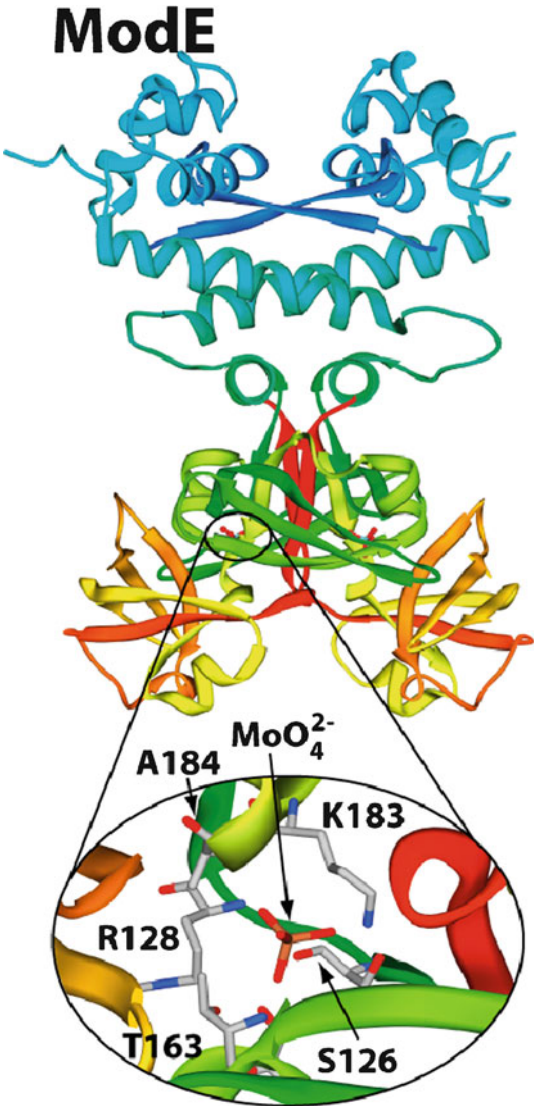


Figure 7 Ribbon diagram (top panel) and detail of the metal binding site (bottom panel) of *Ec*ModE (PDB code: 1O7L). Ribbons are colored from blue in the proximity of the N-termini to red at the C-termini. In the bottom panel, the metal ion and the residues bound to it are depicted as "sticks". Atoms are colored according to their type.

the histidine sequences for self-phosphorylation. Subsequently, each His-bound phosphoryl group is transferred to an aspartate residue in the response regulator component, which binds DNA and activates the transcription of specific genes [136]. Some two-component systems have been identified as responsible for controlling extracellular metal ion sensing: for example, the *Ec*CusR/S detects copper concentration in the periplasm and responds by regulating the expression of genes involved in metal ion export and tolerance [137]. Similarly, the *Synechocystis* 6803 ManR/S controls the expression of genes involved manganese import into the cytoplasm [138].

Sigma factors are dissociable subunits of RNAP, responsible for promoter recognition by the enzyme and DNA melting at the transcriptional initiation site. Competition of different types of sigma factors for RNAP, resulting in a specific transcriptional response, is a general mechanism to finely tune metabolic changes [139,140]. The use of extra-cytoplasmatic sigma factors is an alternative and important mechanism used by bacteria to transduce the external stimuli from the periplasm to the cytoplasm *via* a trans-membrane metal cascade [141]. The archetype of the metal-responsive extra-cytoplasmatic sigma factors is the *Ec*FecI, part of the *Ec*FecIRA signal cascade [142]. *Ec*FecA is a Fe^{3+}-citrate receptor present in the bacterial outer membrane. Ligand binding to *Ec*FecA is transduced through *Ec*FecR, located in the plasma membrane, to the cytosolic sigma factor *Ec*FecI. This regulator then binds the –10 and the –35 sequences, as well as the RNAP, promoting transcriptional initiation of the Fe^{3+}-citrate importer system.

Another example of a three-components signal transduction complex is the *Cupriavidus metallidurans* (*Cm*) CnrYXH, involved in regulation of bacterial tolerance to Co^{2+} and Ni^{2+} ions [143]. *Cm*CnrX is a dimeric membrane-anchored protein with a periplasmic domain acting as a receptor for metal ions. It interacts with the trans-membrane protein *Cm*CnrY, which sequestrates CnrH in its cytoplasmatic side. Binding of metal ions to *Cm*CnrX in the periplasm triggers a signal to *Cm*CnrY, which causes, with a still unknown mechanism, the release of *Cm*CnrH. The latter then can diffuse in the cytoplasm to the DNA molecule and positively regulate the synthesis of proteins involved in Co^{2+} and Ni^{2+} resistance [144]. The recently reported crystal structure of the soluble dimeric domain of *Cm*CnrX revealed that Ni^{2+} and Co^{2+} ions, the inducers of the signalling cascade, bind in an octahedral geometry, while Zn^{2+} ions, leading to an inactive form of the protein, bind in a trigonal bipyramidal environment [145].

4 Metal Ion-Dependent DNA-Protein Interactions in Eukaryotes

4.1 Metal Sensors in Yeast

The study of meta-regulated gene expression in eukaryotes has mainly focused on the yeast *Saccharomyces cerevisiae*. In this organism, metal homeostasis is controlled by (i) metal ion and protein compartmentalization, (ii) metal ion-mediated protein-DNA interactions, and (iii) metal ion-mediated masking of transcriptional activation domains. No metal-responsive signal cascade from the plasma membrane has been discovered yet.

Mechanisms of metal ion tolerance involve their accumulation in vacuoles, diminishing their intracellular availability [146,147]. On the other hand, the localization of proteins in different compartments, such as cytoplasm or nucleus, determines their activity as transcriptional regulators. This is the case of the Fe^{2+}-sensor Aft1/Aft2, which is localized in the nucleus when Fe^{2+} is present at low concentrations, inducing genes that encode Fe^{2+} transporters [148].

Mac1 and Ace1 are metal-dependent activators that bind DNA in response of low and high Cu^+ ions concentrations, respectively. They contain regulatory motifs rich in cysteines that *in vitro* bind Cu^+ ions and form poly-copper structures [149]. The current model indicates that at a threshold Cu^+ concentration, these poly-copper clusters inhibit the DNA binding activity of Mac1, but induce the activity of Ace1, with a consequent repression of Mac1-regulated genes and an activation of the Ace1-regulated genes [150].

Zap1 is a Zn^{2+}-dependent transcriptional regulator that activates and represses gene transcription in a Zn^{2+}-dependent manner. This protein contains a transcriptional activation domain placed between two atypical zinc finger motifs [151]. Zn^{2+} binding to these zinc fingers is more labile than in standard motifs and binding occurs cooperatively. This metal binding event masks the activation domain, thus inactivating gene expression by Zap1 [152].

4.2 Metal Sensors in Higher Eukaryotes

Eukaryotic cells often sense metal ions at the pre-transcriptional level, by changing the stability and translation of mRNA, or at the post-translational level, by determining the protein fate through trafficking and degradation. With few exceptions, metal ion sensors are yet to be discovered in higher organisms. MTF1 is the only known metal responsive transcriptional regulator in animals. This protein is commonly, but not exclusively, acting as a transcriptional activator [153]. In mammalians, MTF1 responds to Zn^{2+} ions that bind Cys_2His_2 zinc fingers motifs, stabilizing a complex with DNA and activating transcription [154]. In addition to the zinc-fingers, *Drosophila* MTF1 contains a copper cluster that confers to the protein the ability to respond to copper ions and to up-regulate the synthesis of a copper-metallothionein [155]. In plants, genetic studies have discovered two metal-responsive transcriptional regulators: FER/FIT1 regulates transcription of genes in response to Fe^{2+} ions deficiency [156,157], while the Cu^+-sensor SPL7 activates expression of genes involved in Cu^+ assimilation [158]. No structural detail on these proteins has been discovered so far.

5 General Conclusions

The study of metal-dependent gene regulation can be considered the new frontier of bio-inorganic chemistry. It requires full grasp of metal ion coordination chemistry, molecular biology, biochemistry, cell biology, structural biology, and several biophysical methods. It therefore represents a novel challenge both for researchers with a chemical background aiming to explore the chemistry of metal ions in biological systems, and for biochemists or molecular biologists eager to widen their knowledge

in the inorganic chemistry of metals that are essential for life as we know it. In most cases, these two perspectives require a closely entangled multi-disciplinary approach that may induce different research groups to undergo a synergistic effort to unravel this new fascinating side of biology.

Abbreviations

Ae	*Alcaligenes eutrophus*
Af	*Acidithiobacillus ferrooxidans*
ATP^{4-}	adenosine 5'-triphosphate
Bj	*Bradyrhizobium japonicum*
bp	base pair
Bs	*Bacillus subtilis*
Ca	*Cyanobacterium anabaena*
Cd	*Corynebacterium diphteriae*
Cg	*Corynebacterium glutamicum*
Cm	*Cupriavidus metallidurans*
DBD	DNA-binding domain
Ec	*Escherichia coli*
Eh	*Enterococcus hirae*
His	histidine
Hp	*Helicobacter pylori*
HTH	helix-turn-helix
MBD	metal-binding domain
mRNA	messenger RNA
Mt	*Mycobacterium tuberculosis*
NMR	nuclear magnetic resonance
Ob	*Oscillatoria brevis*
Pa	*Pseudomonas aeruginosa*
PDB	Protein Data Bank
RHH	ribbon-helix-helix (motif)
Rl	*Rhizobium leguminosarum*
RNAP	RNA polymerase
Sa	*Staphylococcus aureus*
Sc	*Streptomyces coelicolor*
Se	*Salmonella enterica*
Sg	*Streptococcus gordonii*
Sy	*Synechococcus or Synechocystis*
Tt	*Thermus thermophilus*
Vc	*Vibrio colerae*
wHTH	winged helix-turn-helix (motif)

References

1. L. A. Finney, T. V. O'Halloran, *Science* **2003**, *300*, 931–936.
2. K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson, *Nature* **2009**, *460*, 823–830.
3. R. J. P. Williams, *Coord. Chem. Rev*. **2001**, *216*, 583–595.
4. S. Silver, T. L. Phung, *J. Ind. Microbiol. Biotechnol*. **2005**, *32*, 587–605.
5. U. Kramer, *Annu. Rev. Plant. Biol*. **2010**, *61*, 517–534.
6. P. van Rhijn, J. Vanderleyden, *Microbiol. Rev*. **1995**, *59*, 124–142.
7. B. N. Kaiser, S. Moreau, J. Castelli, R. Thomson, A. Lambert, S. Bogliolo, A. Puppo, D. A. Day, *Plant J*. **2003**, *35*, 295–304.
8. J. M. Palacios, H. Manyani, M. Martinez, A. C. Ureta, B. Brito, E. Bascones, L. Rey, J. Imperial, T. Ruiz-Argueso, *Biochem. Soc. Trans*. **2005**, *33*, 94–96.
9. D. Legrand, A. Pierce, E. Elass, M. Carpentier, C. Mariller, J. Mazurier, *Adv. Exp. Med. Biol*. **2008**, *606*, 163–194.
10. B. D. Corbin, E. H. Seeley, A. Raab, J. Feldmann, M. R. Miller, V. J. Torres, K. L. Anderson, B. M. Dattilo, P. M. Dunman, R. Gerads, R. M. Caprioli, W. Nacken, W. J. Chazin, E. P. Skaar, *Science* **2008**, *319*, 962–965.
11. M. A. Almiron, R. A. Ugalde, *J. Microbiol*. **2010**, *48*, 668–673.
12. C. M. Litwin, S. B. Calderwood, *Clin. Microbiol. Rev*. **1993**, *6*, 137–149.
13. A. Danielli, V. Scarlato, *FEMS Microbiol. Rev*. **2010**, *34*, 738–752.
14. K. Muhsen, D. Cohen, *Helicobacter* **2008**, *13*, 323–340.
15. J. Stoof, S. Breijer, R. G. Pot, D. van der Neut, E. J. Kuipers, J. G. Kusters, A. H. van Vliet, *Environ. Microbiol*. **2008**, *10*, 2586–2597.
16. Z. Ma, F. E. Jacobsen, D. P. Giedroc, *Chem. Rev*. **2009**, *109*, 4644–4681.
17. S. Silver, L. T. Phung, *Ann. Rev. Microbiol*. **1996**, *50*, 753–789.
18. D. Osman, J. S. Cavet, *Nat. Prod. Rep*. **2010**, *27*, 668–680.
19. K. S. Gajiwala, S. K. Burley, *Curr. Opin. Struct. Biol*. **2000**, *10*, 110–116.
20. E. R. Schreiter, C. L. Drennan, *Nat. Rev. Microbiol*. **2007**, *5*, 710–720.
21. M. A. Pennella, D. P. Giedroc, *Biometals* **2005**, *18*, 413–428.
22. D. R. Campbell, K. E. Chapman, K. J. Waldron, S. Tottey, S. Kendall, G. Cavallaro, C. Andreini, J. Hinds, N. G. Stoker, N. J. Robinson, J. S. Cavet, *J. Biol. Chem*. **2007**, *282*, 32298–32310.
23. A. P. Morby, J. S. Turner, J. W. Huckle, N. J. Robinson, *Nucleic Acids Res*. **1993**, *21*, 921–925.
24. J. Wu, B. P. Rosen, *Mol. Microbiol*. **1991**, *5*, 1331–1336.
25. C. Thelwell, N. J. Robinson, J. S. Turner-Cavet, *Proc. Natl. Acad. Sci*. **1998**, *95*, 10728–10733.
26. G. Endo, S. Silver, *J. Bacteriol*. **1995**, *177*, 4437–4441.
27. C. Eicken, M. A. Pennella, X. Chen, K. M. Koshlap, M. L. VanZile, J. C. Sacchettini, D. P. Giedroc, *J. Mol. Biol*. **2003**, *333*, 683–695.
28. C. M. Moore, A. Gaballa, M. Hui, R. W. Ye, J. D. Helmann, *Mol. Microbiol*. **2005**, *57*, 27–40.
29. T. Liu, J. W. Golden, D. P. Giedroc, *Biochemistry* **2005**, *44*, 8673–8683.
30. J. S. Cavet, W. Meng, M. A. Pennella, R. J. Appelhoff, D. P. Giedroc, N. J. Robinson, *J. Biol. Chem*. **2002**, *277*, 38441–38448.
31. J. S. Cavet, A. I. Graham, W. Meng, N. J. Robinson, *J. Biol. Chem*. **2003**, *278*, 44560–44566.
32. T. Liu, X. Chen, Z. Ma, J. Shokes, L. Hemmingsen, R. A. Scott, D. P. Giedroc, *Biochemistry* **2008**, *47*, 10564–100575.
33. A. I. Arunkumar, G. C. Campanello, D. P. Giedroc, *Proc. Natl. Acad. Sci*. **2009**, *106*, 18177–18182.
34. J. Ye, A. Kandegedara, P. Martin, B. P. Rosen, *J. Bacteriol*. **2005**, *187*, 4214–4221.
35. L. Banci, I. Bertini, F. Cantini, S. Ciofi-Baffoni, J. S. Cavet, C. Dennison, A. I. Graham, D. R. Harvie, N. J. Robinson, *J. Biol. Chem*. **2007**, *282*, 30181–30188.
36. M. D. Wong, Y. F. Lin, B. P. Rosen, *J. Biol. Chem*. **2002**, *277*, 40930–40936.

37. W. Shi, J. Dong, R. A. Scott, M. Y. Ksenzenko, B. P. Rosen, *J. Biol. Chem*. **1996**, *271*, 9291–9297.
38. J. Qin, H. L. Fu, J. Ye, K. Z. Bencze, T. L. Stemmler, D. E. Rawlings, B. P. Rosen, *J. Biol. Chem*. **2007**, *282*, 34346–34355.
39. L. S. Busenlehner, T. C. Weng, J. E. Penner-Hahn, D. P. Giedroc, *J. Mol. Biol*. **2002**, *319*, 685–701.
40. T. O'Halloran, C. Walsh, *Science* **1987**, *235*, 211–214.
41. N. L. Brown, J. V. Stoyanov, S. P. Kidd, J. L. Hobman, *FEMS Microbiol. Rev*. **2003**, *27*, 145–163.
42. C. B. Harley, R. P. Reynolds, *Nucleic Acids Res*. **1987**, *15*, 2343–2361.
43. D. F. Browning, S. J. Busby, *Nat. Rev. Microbiol*. **2004**, *2*, 57–65.
44. A. Z. Ansari, M. L. Chael, T. V. O'Halloran, *Nature* **1992**, *355*, 87–89.
45. C. E. Outten, F. W. Outten, T. V. O'Halloran, *J. Biol. Chem*. **1999**, *274*, 37517–37524.
46. B. Borremans, J. L. Hobman, A. Provoost, N. L. Brown, D. van Der Lelie, *J. Bacteriol*. **2001**, *183*, 5651–5658.
47. A. Changela, K. Chen, Y. Xue, J. Holschen, C. E. Outten, T. V. O'Halloran, A. Mondragon, *Science* **2003**, *301*, 1383–1387.
48. K. R. Brocklehurst, S. J. Megit, A. P. Morby, *Biochem. Biophys. Res. Commun*. **2003**, *308*, 234–239.
49. L. B. Pontel, M. E. Audero, M. Espariz, S. K. Checa, F. C. Soncini, *Mol. Microbiol*. **2007**, *66*, 814–825.
50. X. Jian, E. C. Wasinger, J. V. Lockard, L. X. Chen, C. He, *J. Am. Chem. Soc*. **2009**, *131*, 10869–10871.
51. L. M. Utschig, J. W. Bryson, T. V. O'Halloran, *Science* **1995**, *268*, 380–385.
52. S. K. Checa, M. Espariz, M. E. Audero, P. E. Botta, S. V. Spinelli, F. C. Soncini, *Mol. Microbiol*. **2007**, *63*, 1307–1318.
53. M. Ahmed, L. Lyass, P. N. Markham, S. S. Taylor, N. Vazquez-Laslop, A. A. Neyfakh, *J. Bacteriol*. **1995**, *177*, 3904–3910.
54. E. E. Heldwein, R. G. Brennan, *Nature* **2001**, *409*, 378–382.
55. K. J. Newberry, R. G. Brennan, *J. Biol. Chem*. **2004**, *279*, 20356–20362.
56. S. Watanabe, A. Kita, K. Kobayashi, K. Miki, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4121–4126.
57. B. Gold, G. M. Rodriguez, S. A. Marras, M. Pentecost, I. Smith, *Mol. Microbiol*. **2001**, *42*, 851–865.
58. I. Brune, H. Werner, A. T. Huser, J. Kalinowski, A. Puhler, A. Tauch, *BMC Genomics* **2006**, *7*, 1–19.
59. I. Brune, H. Werner, A. T. Huser, J. Kalinowski, A. Puhler, A. Tauch, *BMC Genomics* **2006**, *7*, 21.
60. M. M. Spiering, D. Ringe, J. R. Murphy, M. A. Marletta, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3808–3813.
61. M. P. Schmitt, M. Predich, L. Doukhan, I. Smith, R. K. Holmes, *Infect. Immun*. **1995**, *63*, 4284–4289.
62. X. Qiu, C. L. Verlinde, S. Zhang, M. P. Schmitt, R. K. Holmes, W. G. Hol, *Structure* **1995**, *3*, 87–100.
63. X. Qiu, E. Pohl, R. K. Holmes, W. G. Hol, *Biochemistry* **1996**, *35*, 12292–12302.
64. N. Schiering, X. Tao, H. Zeng, J. R. Murphy, G. A. Petsko, D. Ringe, *Proc. Natl. Acad. Sci. U S A* **1995**, *92*, 9843–9850.
65. X. Ding, H. Zeng, N. Schiering, D. Ringe, J. R. Murphy, *Nat. Struct. Biol*. **1996**, *3*, 382–387.
66. A. White, X. Ding, J. C. vanderSpek, J. R. Murphy, D. Ringe, *Nature* **1998**, *394*, 502–506.
67. E. Pohl, R. K. Holmes, W. G. Hol, *J. Mol. Biol*. **1999**, *292*, 653–667.
68. E. Pohl, R. K. Holmes, W. G. Hol, *J. Mol. Biol*. **1999**, *285*, 1145–1156.
69. M. D. Feese, B. P. Ingason, J. Goranson-Siekierke, R. K. Holmes, W. G. Hol, *J. Biol. Chem*. **2001**, *276*, 5959–5966.
70. G. Wisedchaisri, R. K. Holmes, W. G. Hol, *J. Mol. Biol*. **2004**, *342*, 1155–1169.

71. G. P. Wylie, V. Rangachari, E. A. Bienkiewicz, V. Marin, N. Bhattacharya, J. F. Love, J. R. Murphy, T. M. Logan, *Biochemistry* **2005**, *44*, 40–51.
72. C. Liu, K. Mao, M. Zhang, Z. Sun, W. Hong, C. Li, B. Peng, Z. Chang, *J. Biol. Chem*. **2008**, *283*, 2439–2453.
73. C. J. Chou, G. Wisedchaisri, R. R. Monfeli, D. M. Oram, R. K. Holmes, W. G. Hol, C. Beeson, *J. Biol. Chem*. **2004**, *279*, 53554–53561.
74. P. D. Twigg, G. Parthasarathy, L. Guerrero, T. M. Logan, D. L. Caspar, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11259–11264.
75. M. Semavina, D. Beckett, T. M. Logan, *Biochemistry* **2006**, *45*, 12480–12490.
76. S. A. Lieser, T. C. Davis, J. D. Helmann, S. M. Cohen, *Biochemistry* **2003**, *42*, 12634–12642.
77. A. Glasfeld, E. Guedon, J. D. Helmann, R. G. Brennan, *Nat. Struct. Biol*. **2003**, *10*, 652–657.
78. J. I. Kliegman, S. L. Griner, J. D. Helmann, R. G. Brennan, A. Glasfeld, *Biochemistry* **2006**, *45*, 3493–3505.
79. M. A. DeWitt, J. I. Kliegman, J. D. Helmann, R. G. Brennan, D. L. Farrens, A. Glasfeld, *J. Mol. Biol*. **2007**, *365*, 1257–1265.
80. M. Golynskiy, S. Li, V. L. Woods, Jr., S. M. Cohen, *J. Biol. Inorg. Chem*. **2007**, *12*, 699–709.
81. T. Tanaka, A. Shinkai, Y. Bessho, T. Kumarevel, S. Yokoyama, *Proteins* **2009**, *77*, 741–746.
82. K. E. Stoll, W. E. Draper, J. I. Kliegman, M. V. Golynskiy, R. A. Brew-Appiah, R. K. Phillips, H. K. Brown, W. A. Breyer, N. S. Jakubovics, H. F. Jenkinson, R. G. Brennan, S. M. Cohen, A. Glasfeld, *Biochemistry* **2009**, *48*, 10308–10320.
83. S. C. Andrews, A. K. Robinson, F. Rodriguez-Quinones, *FEMS Microbiol. Rev*. **2003**, *27*, 215–237.
84. A. Bagg, J. B. Neilands, *Biochemistry* **1987**, *26*, 5471–5477.
85. R. W. Prince, D. G. Storey, A. I. Vasil, M. L. Vasil, *Mol. Microbiol*. **1991**, *5*, 2823–2831.
86. S. Bereswill, F. Lichte, T. Vey, F. Fassbinder, M. Kist, *FEMS Microbiol. Lett*. **1998**, *159*, 193–200.
87. A. R. Mey, E. E. Wyckoff, V. Kanukurthy, C. R. Fisher, S. M. Payne, *Infect. Immun*. **2005**, *73*, 8167–8178.
88. V. de Lorenzo, S. Wee, M. Herrero, J. B. Neilands, *J. Bacteriol*. **1987**, *169*, 2624–2630.
89. E. Masse, H. Salvail, G. Desnoyers, M. Arguin, *Curr. Opin. Microbiol*. **2007**, *10*, 140–145.
90. I. Delany, G. Spohn, R. Rappuoli, V. Scarlato, *Mol. Microbiol*. **2001**, *42*, 1297–1309.
91. A. Danielli, D. Roncarati, I. Delany, V. Chiarini, R. Rappuoli, V. Scarlato, *J. Bacteriol*. **2006**, *188*, 4654–4662.
92. D. Lucarelli, S. Russo, E. Garman, A. Milano, W. Meyer-Klaucke, E. Pohl, *J. Biol. Chem*. **2007**, *282*, 9914–9922.
93. E. Diaz-Mireles, M. Wexler, G. Sawers, D. Bellini, J. D. Todd, A. W. Johnston, *Microbiology* **2004**, *150*, 1447–1456.
94. B. E. Ahn, J. Cha, E. J. Lee, A. R. Han, C. J. Thompson, J. H. Roe, *Mol. Microbiol*. **2006**, *59*, 1848–1858.
95. J. W. Lee, J. D. Helmann, *Nature* **2006**, *440*, 363–367.
96. E. Pohl, J. C. Haller, A. Mijovilovich, W. Meyer-Klaucke, E. Garman, M. L. Vasil, *Mol. Microbiol*. **2003**, *47*, 903–915.
97. D. A. Traore, A. El Ghazouani, S. Ilango, J. Dupuy, L. Jacquamet, J. L. Ferrer, C. Caux-Thang, V. Duarte, J. M. Latour, *Mol. Microbiol*. **2006**, *61*, 1211–1219.
98. D. A. Traore, A. El Ghazouani, L. Jacquamet, F. Borel, J. L. Ferrer, D. Lascoux, J. L. Ravanat, M. Jaquinod, G. Blondin, C. Caux-Thang, V. Duarte, J. M. Latour, *Nat. Chem. Biol*. **2009**, *5*, 53–59.
99. Y. J. An, B. E. Ahn, A. R. Han, H. M. Kim, K. M. Chung, J. H. Shin, Y. B. Cho, J. H. Roe, S. S. Cha, *Nucleic. Acids Res*. **2009**, *37*, 3442–3451.
100. M. A. Sheikh, G. L. Taylor, *Mol. Microbiol*. **2009**, *72*, 1208–1220.
101. C. Dian, S. Vitale, G. A. Leonard, C. Bahlawane, C. Fauquant, D. Leduc, C. Muller, H. de Reuse, I. Michaud-Soret, L. Terradot, *Mol. Microbiol*. **2011**, *79*, 1260–1275.
102. L. Pecqueur, B. D'Autreaux, J. Dupuy, Y. Nicolet, L. Jacquamet, B. Brutscher, I. Michaud-Soret, B. Bersch, *J. Biol. Chem*. **2006**, *281*, 21286–21295.

103. A. C. Lewin, P. A. Doughty, L. Flegg, G. R. Moore, S. Spiro, *Microbiology* **2002**, *148*, 2449–2456.
104. S. K. Small, S. Puri, M. R. O'Brian, *Biometals* **2009**, *22*, 89–97.
105. K. De Pina, V. Desjardin, M. A. Mandrand-Berthelot, G. Giordano, L. F. Wu, *J. Bacteriol*. **1999**, *181*, 670–674.
106. P. T. Chivers, R. T. Sauer, *J. Biol. Chem*. **2000**, *275*, 19735–19741.
107. A. H. van Vliet, F. D. Ernst, J. G. Kusters, *Trends Microbiol*. **2004**, *12*, 489–494.
108. E. L. Benanti, P. T. Chivers, *J. Bacteriol*. **2009**, *191*, 2405–2408.
109. E. R. Schreiter, M. D. Sintchak, Y. Guo, P. T. Chivers, R. T. Sauer, C. L. Drennan, *Nat. Struct. Biol*. **2003**, *10*, 794–799.
110. P. T. Chivers, T. H. Tahirov, *J. Mol. Biol*. **2005**, *348*, 597–607.
111. E. R. Schreiter, S. C. Wang, D. B. Zamble, C. L. Drennan, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13676–13681.
112. C. Dian, K. Schauer, U. Kapp, S. M. McSweeney, A. Labigne, L. Terradot, *J. Mol. Biol*. **2006**, *361*, 715–730.
113. C. M. Phillips, E. R. Schreiter, Y. Guo, S. C. Wang, D. B. Zamble, C. L. Drennan, *Biochemistry* **2008**, *47*, 1938–1946.
114. A. L. West, F. St John, P. E. Lopes, A. D. MacKerell, Jr., E. Pozharski, S. L. Michel, *J. Am. Chem. Soc*. **2010**, *132*, 14447–14456.
115. P. T. Chivers, R. T. Sauer, *Protein Sci*. **1999**, *8*, 2494–2500.
116. P. T. Chivers, R. T. Sauer, *Chem. Biol*. **2002**, *9*, 1141–1148.
117. N. S. Dosanjh, S. L. Michel, *Curr. Opin. Chem. Biol*. **2006**, *10*, 123–130.
118. F. Musiani, B. Bertosa, A. Magistrato, B. Zambelli, P. Turano, V. Losasso, C. Micheletti, S. Ciurli, P. Carloni, *J. Comput. Theor. Chem*. **2010**, *6*, 3503–3515.
119. T. Liu, A. Ramesh, Z. Ma, S. K. Ward, L. Zhang, G. N. George, A. M. Talaat, J. C. Sacchettini, D. P. Giedroc, *Nat. Chem. Biol*. **2007**, *3*, 60–68.
120. T. Liu, A. Ramesh, Z. Ma, S. K. Ward, L. Zhang, G. N. George, A. M. Talaat, J. C. Sacchettini, D. P. Giedroc, *Nat. Chem. Biol*. **2007**, *3*, 60–68.
121. K. Sakamoto, Y. Agari, K. Agari, S. Kuramitsu, A. Shinkai, *Microbiology* **2010**, *156*, 1993–2005.
122. Z. Ma, D. M. Cowart, R. A. Scott, D. P. Giedroc, *Biochemistry* **2009**, *48*, 3325–3334.
123. J. S. Iwig, J. L. Rowe, P. T. Chivers, *Mol. Microbiol*. **2006**, *62*, 252–262.
124. J. S. Iwig, S. Leitch, R. W. Herbst, M. J. Maroney, P. T. Chivers, *J. Am. Chem. Soc*. **2008**, *130*, 7592–7606.
125. A. Rodrigue, G. Effantin, M. A. Mandrand-Berthelot, *J. Bacteriol*. **2005**, *187*, 2912–2916.
126. D. Strausak, M. Solioz, *J. Biol. Chem*. **1997**, *272*, 8932–8936.
127. P. A. Cobine, C. E. Jones, C. T. Dameron, *J. Inorg. Biochem*. **2002**, *88*, 192–196.
128. P. A. Cobine, G. N. George, C. E. Jones, W. A. Wickramasinghe, M. Solioz, C. T. Dameron, *Biochemistry* **2002**, *41*, 5822–5829.
129. F. Cantini, L. Banci, M. Solioz, *Biochem. J*. **2009**, *417*, 493–499.
130. L. A. Anderson, T. Palmer, N. C. Price, S. Bornemann, D. H. Boxer, R. N. Pau, *Eur. J. Biochem*. **1997**, *246*, 119–126.
131. S. E. Maddocks, P. C. Oyston, *Microbiology* **2008**, *154*, 3609–3623.
132. L. A. Anderson, E. McNairn, T. Lubke, R. N. Pau, D. H. Boxer, *J. Bacteriol*. **2000**, *182*, 7035–7043.
133. D. R. Hall, D. G. Gourley, G. A. Leonard, E. M. Duke, L. A. Anderson, D. H. Boxer, W. N. Hunter, *EMBO J*. **1999**, *18*, 1435–1446.
134. A. W. Schuttelkopf, D. H. Boxer, W. N. Hunter, *J. Mol. Biol*. **2003**, *326*, 761–767.
135. A. Khorchid, M. Ikura, *Int. J. Biochem. Cell. Biol*. **2006**, *38*, 307–312.
136. R. Gao, A. M. Stock, *Ann. Rev. Microbiol*. **2009**, *63*, 133–154.
137. G. P. Munson, D. L. Lam, F. W. Outten, T. V. O'Halloran, *J. Bacteriol*. **2000**, *182*, 5864–5871.
138. T. Ogawa, D. H. Bao, H. Katoh, M. Shibata, H. B. Pakrasi, M. Bhattacharyya-Pakrasi, *J. Biol. Chem*. **2002**, *277*, 28981–28986.
139. K. S. Murakami, S. Masuda, S. A. Darst, *Science* **2002**, *296*, 1280–1284.

140. D. G. Vassylyev, S. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov, S. Yokoyama, *Nature* **2002**, *417*, 712–719.
141. A. Staron, H. J. Sofia, S. Dietrich, L. E. Ulrich, H. Liesegang, T. Mascher, *Mol. Microbiol*. **2009**, *74*, 557–581.
142. V. Braun, S. Mahren, A. Sauter, *Biometals* **2006**, *19*, 103–113.
143. C. Grosse, S. Friedrich, D. H. Nies, *J. Mol. Microbiol. Biotechnol*. **2007**, *12*, 227–240.
144. G. Grass, B. Fricke, D. H. Nies, *Biometals* **2005**, *18*, 437–448.
145. J. Trepreau, E. Girard, A. P. Maillard, E. de Rosny, I. Petit-Haertlein, R. Kahn, J. Coves, *J. Mol. Biol.* **2011**, in press.
146. C. W. MacDiarmid, L. A. Gaither, D. Eide, *EMBO J*. **2000**, *19*, 2845–2855.
147. L. Li, O. S. Chen, D. McVey Ward, J. Kaplan, *J. Biol. Chem*. **2001**, *276*, 29515–29519.
148. R. Ueta, N. Fujiwara, K. Iwai, Y. Yamaguchi-Iwai, *Mol. Biol. Cell*. **2007**, *18*, 2980–2890.
149. K. R. Brown, G. L. Keller, I. J. Pickering, H. H. Harris, G. N. George, D. R. Winge, *Biochemistry* **2002**, *41*, 6469–6476.
150. M. M. Pena, K. A. Koch, D. J. Thiele, *Mol. Cell. Biol*. **1998**, *18*, 2514–2523.
151. A. J. Bird, K. McCall, M. Kramer, E. Blankman, D. R. Winge, D. J. Eide, *EMBO J*. **2003**, *22*, 5137–5146.
152. Z. Wang, L. S. Feng, V. Matskevich, K. Venkataraman, P. Parasuram, J. H. Laity, *J. Mol. Biol*. **2006**, *357*, 1167–1183.
153. U. Wimmer, Y. Wang, O. Georgiev, W. Schaffner, *Nucleic Acids Res*. **2005**, *33*, 5715–5727.
154. J. H. Laity, G. K. Andrews, *Arch. Biochem. Biophys*. **2007**, *463*, 201–210.
155. X. Chen, H. Hua, K. Balamurugan, X. Kong, L. Zhang, G. N. George, O. Georgiev, W. Schaffner, D. P. Giedroc, *Nucleic Acids Res*. **2008**, *36*, 3128–3138.
156. H. Q. Ling, P. Bauer, Z. Bereczky, B. Keller, M. Ganal, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13938–13943.
157. E. P. Colangelo, M. L. Guerinot, *Plant Cell* **2004**, *16*, 3400–3412.
158. H. Yamasaki, M. Hayashi, M. Fukazawa, Y. Kobayashi, T. Shikanai, *Plant Cell* **2009**, *21*, 347–361.
159. S. Benini, M. Cianci, S. Ciurli, *Dalton Trans*. **2011**, *40*, 7831–7833
160. B. Zambelli, M. Bellucci, A. Danielli, V. Scarlato, S. Ciurli, *Chem. Commun*. **2007**, 3649–3651
161. B. Zambelli, A. Danielli, S. Romagnoli, P. Neyroz, S. Ciurli, V. Scarlato, *J. Mol. Biol*. **2008**, *383*, 1129–1143

Chapter 6
Spectroscopic Investigations of Lanthanide Ion Binding to Nucleic Acids

Janet R. Morrow and Christopher M. Andolina

Contents

J.R. Morrow (✉) • C.M. Andolina
Department of Chemistry, University at Buffalo, State University of New York, Buffalo, NY 14260-3000, USA
e-mail: jmorrow@buffalo.edu; christopher.andolina@gmail.com

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_6,

Abstract Luminescent lanthanide (Ln(III)) ions are valuable spectroscopic probes for metal ion binding sites in nucleic acids. In this chapter, we briefly review Ln(III) luminescence and the information available from these experiments. An emphasis is placed on direct excitation Eu(III) spectroscopy as a tool. Eu(III) excitation spectroscopy is used to show that solutions containing micromolar Eu(III), 100 mM NaCl, and 20 mM MES buffer contain predominantly a mononuclear Eu(III) aqua complex and an Eu(III) hydroxide complexes. The binding of these species to various RNA and DNA sequences are monitored by using Eu(III) excitation spectroscopy. Eu(III) luminescence lifetime data shows that the Eu(III) ion typically loses 1–3 water molecules to form innersphere complexes with RNA and DNA that contain tandem base pair mismatches or hairpin loops. In addition, early studies that used nucleobase-sensitized Eu(III) or Tb(III) luminescence within transfer RNA or in the hammerhead ribozyme are presented. Luminescence resonance energy transfer studies are shown to be useful for determining distances between bound Ln(III) ion and organic fluorophores or between two different Ln(III) ions. To supplement luminescence data, the binding sites of paramagnetic Ln(III) ions are determined by monitoring the chemical shifts of nucleotide protons. Binding sites are identified by following the protons that are influenced by the Ln(III) pseudo-contact shift.

Keywords DNAzyme • europium • fluorescence resonance energy transfer • lanthanide ion aqueous complexes • lanthanide ion complexes • lanthanide ion coordination chemistry • lanthanide ion dimerization • lanthanide ion hydrolysis • laser induced luminescence • luminescence lifetimes • lanthanide luminescence • lanthanide ion macrocyclic complexes • lanthanide ion solution chemistry • nucleic acid materials • nucleic acid structure • pseudo-contact shift • ribozyme • RNA NMR structure

1 Introduction

1.1 Lanthanide(III) Ions as Luminescent Probes

Exquisitely detailed information about the Ln(III) coordination sphere can be obtained from luminescence studies of Ln(III) complexes [1–5]. The information obtained from luminescence experiments is especially valuable for the characterization of metal ion binding sites in biopolymers. A primary example of this is the use of Eu(III) or Tb(III) ions as spectroscopically active metal ion surrogates for the study of Ca(II) ion binding sites in proteins and peptides [6,7]. Ln(III) luminescence spectroscopy has enabled the characterization of these metal ion binding sites by giving information on the type of ligand donor group, geometry of the complex, number of bound waters and

Ln(III)-Ln(III) internuclear distances in the case of multiple metal ion binding sites. In contrast, there are relatively few studies that use Ln(III) ions for probing metal ion binding sites in nucleic acids, despite the importance of metal cation interactions in DNA and RNA. In this chapter, studies on Ln(III) binding to RNA, DNA, and to nucleic acids containing non-natural ligands that bind Ln(III) ions are reviewed.

There are two commonly used methods to excite Ln(III) luminescence [3,8–10]. The most frequently used method relies on excitation of a sensitizing ligand, generally an organic dye, as discussed further below. The second method is direct photoexcitation of Ln(III) ion luminescence through electronic transitions within the $4f^n$ manifold. However, the electronic transitions between states with a $4f^n$ configuration are electric dipole (parity) and spin forbidden transitions giving rise to extremely weak luminescence from dilute solutions of Ln(III) ions. Powerful (mJ/pulse) excitation sources are needed to study direct Ln(III) luminescence. The weak luminescence of the Ln(III) ions is exacerbated further in aqueous solution due to effective non-radiative quenching of luminescence by adjacent OH oscillators of water. Thus, dilute solution studies of Ln(III) luminescence under biologically relevant conditions are challenging. Laser-induced direct excitation Ln(III) luminescence was studied in the 1980s and 1990s especially in the groups of Horrocks and Choppin [2,5,11]. These systems used a pumped dye laser for excitation over a narrow wavelength window. More recently in the Morrow laboratory, a novel spectroscopic system was built based on an injection seeded frequency tripled (λ = 355 nm) Nd:YAG pump laser coupled with a master oscillator power oscillator (MOPO) which offers continuously tunable output in the UV (230–400 nm), visible (420–690 nm), and IR (735–1800 nm) spectrum with approximately 50 mJ/pulse at 580 nm [12]. The superior detection limits of this instrument make it possible to study Eu(III) excitation spectroscopy at nanomolar concentrations of lanthanide complexes in water, approximately 1000-fold lower concentrations than previously studied. The lower concentration limits for studying lanthanide ions in water by using the MOPO/laser system enables us to study interactions with nucleic acids under more biologically relevant conditions than used previously. This chapter has an emphasis on the application of direct excitation Ln(III) luminescence spectroscopy as a tool to study Ln(III) coordination sphere and to gain insight on cation interactions with nucleic acids.

Lanthanide luminescence is more typically studied through excitation of a sensitizing ligand, generally an organic dye [3,8,10,13]. The sensitizing ligand acts as an antennae which, upon photoexcitation, transfers energy to the lanthanide ion leading to Ln(III) luminescence. A good sensitizing ligand has a large molar absorptivity for conversion to the excited singlet state, efficient intersystem crossing to the triplet state and a triplet state energy between 2500–3500 cm^{-1} above that of a Ln(III) excited state. Good progress has been made on the synthesis of Ln(III) complexes that have high quantum yields by judicious consideration of photophysics of Ln(III) ions and sensitizers [8,9,14]. The natural nucleobases of DNA and RNA are not optimal for sensitization of most Ln(III); however, Tb(III) luminescence is sensitized in a base sequence and structure dependent manner [15,16]. Early studies showed that Tb(III) ions could be sensitized by both single-stranded and double-stranded nucleic acids with especially efficient sensitization by guanine rich sequences [17].

1.2 Metal Ion Binding to Nucleic Acids

Nucleic acids contain only four different nucleosides including adenosine, guanosine, thymidine and cytidine for DNA and uridine occurs instead of thymidine for RNA. Each nucleoside is connected through a phosphate diester linkage which bears a negative charge. DNA and RNA are polynucleotides with large negative charges that are neutralized by association with cations, which may include polyamines, cationic amino acids in protein complexes or divalent or monovalent metal cations. Most of these metal ion cations are associated loosely with the nucleic acid in order to neutralize charge. Such loosely associated "diffuse" cations are described by a non-linear Poisson-Boltzmann (NLPB) equation [18–20]. According to the NLPB effect, there is a higher fraction of monovalent cations associated with phosphate esters of double-stranded versus single-stranded nucleic acids due to the closer spacing of phosphate esters in the former and the need for the alleviation of electrostatic stress. Interestingly, there is also a correspondingly greater association of cations to A-form RNA compared to B-form DNA, in part due to the closer spacing of phosphate diesters in A-form nucleic acids [21]. In addition, for diffuse metal ions, divalent cations are more effective at relieving electrostatic stress and thus associate more tightly with the polynucleotide than do monovalent cations [20].

In addition to diffuse metal ions, there are site specifically bound metal ions, especially in highly folded RNAs that have multiple ligand donor groups that might simultaneously interact with the metal ion. Site specifically bound metal ions may be outersphere and interact through water ligands or may have waters replaced to give innersphere interactions. In RNA structures that contain either single nucleotide or two nucleotide bulges, there is crystallographic evidence that metal cations such as Ca(II) or Mg(II) bind to specific donor groups such as the N7 of guanosine [22]. Mg(II) aqua ion has high hydration energies and frequently forms outersphere complexes. However, interesting innersphere Mg(II) complexes that have bidentate chelation have been reported [23–26]. Other donor ligands include the carbonyl groups of nucleobases and the 2'-hydroxyl group of RNA. Different types of binding sites have been characterized for monovalent and divalent cations in large nucleic acid structures such as the ribosome as studied by X-ray crystallography [27].

Solution state characterization of these metal ion binding sites is very difficult, especially for spectroscopically silent metal ions such as Mg(II), Na(I), K(I) or Ca(II). Recently developed techniques that enable the study of natural metal ions include methods such as Raman crystallography [28] and solution Raman spectroscopy to monitor binding of Mg(II) to phosphate esters [29]. However, there is a long tradition of employing metal ion surrogates that are spectroscopically active to probe metal ion binding sites in nucleic acids [30]. Briefly, one of the more commonly used spectroscopic probes is Mn(II), a paramagnetic cation, which is used as a Mg(II) surrogate. Mn(II) binding can be probed either by using EPR spectroscopy or by monitoring with NMR the effect of the Mn(II) ion on protons that are in close proximity to the Mn(II) binding site [31,32]. Such protons undergo effective line broadening due to the paramagnetic Mn(II) nucleus [33,34]. Alternatively,

$Co(NH_3)_6^{3+}$ is used as a surrogate for outersphere binding of fully hydrated Mg(II) [33,35]. Nuclear Overhauser experiments are used to monitor through space interactions between the protons on $Co(NH_3)_6^{3+}$ and those of nucleic acids close to the binding site [36,37]. Other methods use phosphorothioate modifications to facilitate the ^{31}P NMR studies of Cd(II) binding [38,39]. NMR spectroscopic studies of ^{15}N-labeled nucleobases have been used to monitor binding of both NMR silent as well as NMR active divalent metal ions such as $^{113}Cd(II)$ [40–42]. Such metal ion surrogates or mimics have provided insight into the general properties of metal ion binding to nucleic acids despite criticism that the properties of surrogate ions are not identical to those of the naturally occurring metal ions. Recent work in this area is summarized in the first chapter of this volume.

1.3 Lanthanide(III) Ions as Probes for Metal Ion Binding Sites in Nucleic Acids

Ln(III) ions have been used as surrogates for alkaline earth ion binding to nucleic acids, especially in proteins [1]. Ln(III) ions are close in size and coordination number (8–9) to Ca(II) ions, not Mg(II) ions. The large positive charge of trivalent Ln(III) ions is expected to result in stronger nucleic acid binding constants than observed for divalent ions. However, the oxophilicity and hard base character of the Ln(III) are properties held in common with Mg(II) and Ca(II). Large hydration energies are another common feature of Ln(III) and Mg(II), one that makes it difficult to form innersphere contacts [43].

In addition to luminescence properties, the paramagnetic properties of the Ln(III) are useful in probing metal ion binding sites in nucleic acids [44]. We show that Ln(III) ion binding sites are mapped through pseudo-contact shifts of ^{1}H and ^{31}P resonances of the nucleic acids for groups that are in close proximity to the bound Ln(III). In this type of experiment, the different magnetic but similar chemical properties of each of the paramagnetic Ln(III) ions are utilized. Metal ion binding sites are mapped through a series of ^{1}H or ^{31}P NMR titrations that capitalize on the different distance range of measurable pseudo-contact shift for each Ln(III) [45].

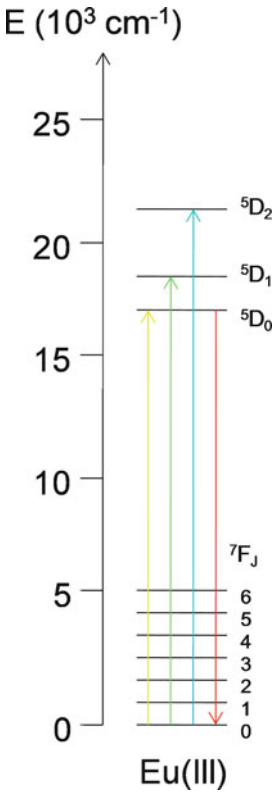
We begin the chapter by reviewing typical luminescence experiments with an emphasis on direct excitation Eu(III) luminescence. In Section 3, the aqueous chemistry of Eu(III) as studied by direct excitation luminescence is presented as background for understanding binding to nucleic acids. In Section 4, Ln(III) binding to RNA is discussed for structurally complex large RNAs including transfer RNA, ribozymes, and splicesomes. This is followed by presentation of studies of simple structural motifs such as duplexes or hairpin loops that can readily be investigated by NMR spectroscopy methods in addition to luminescence methods. The last section will cover Ln(III) complexes of modified nucleic acids which are of interest in the field of supramolecular chemistry for the preparation of organized assemblies of metal ions.

2 Direct Excitation Lanthanide Luminescence

2.1 Luminescence Properties of the Lanthanide Ions

The 4f orbitals of the Ln(III) are isolated from the environment by the filled outer-core 5s and 5p electrons that are of lower energy but extend further into space. Thus the 4f orbitals do not participate to any large extent in bonding. A consequence of f-orbital shielding and minimal interaction with the environment or crystal field is that the selection rules for f-f transitions lead to weak absorption and emission spectra. The atomic line-like spectra of the Ln(III) do not change markedly with environment compared to transition metals, except for crystal field splitting of the f-f transitions [10].

Shown in Scheme 1 is the energy level diagram for Eu(III) showing the ground state ^{7}F manifold and the first three excited states. Energy differences between terms arise primarily from differences in Coulombic or interelectronic repulsion to give separations on the order of 10^4 cm^{-1}. Terms are further split by spin-orbit coupling (10^3 cm^{-1}) to give several J levels. As shown in Scheme 1, the ground state ^{7}F manifold for the free Eu(III) ion has seven J levels. The energy level diagrams for the other



Scheme 1 Energy level diagram of Eu(III).

Ln(III) are available [46]. Finally, the matrix or the ligand donors in complexes (the crystal field) may split each of the J levels into further sublevels. Of note, apart from Eu(III), none of the other Ln(III) ions has a ground state or a lowest excited state that is non-degenerate. This feature makes Eu(III) an especially good probe because excitation and emission bands are relatively simple for certain transitions.

Of special note is the energy difference between the first excited state and the highest level of the ground state manifold for each Ln(III). The smaller this energy difference, the more readily the Ln(III) ion is quenched non-radiatively by vibronic processes including water OH oscillations or by other solvent or ligand NH, OH or CH stretches according to the energy gap law [46]. Highly quenched Ln(III) ions have low quantum yields and luminescence lifetimes that are generally in the sub-microsecond (Er(III), Tm(III), Ho(III)) or microsecond (Yb(III), Sm(III), Dy(III)) range. Conversely, Ln(III) with large energy gaps can have lifetimes in the millisecond range (Eu(III), Tb(III), Gd(III)). Eu(III) and Tb(III) are the most commonly used luminescent Ln(III), in part due to their longer lifetimes, moderate susceptibility to OH quenching, which is useful for water counting experiments, and their emission bands, which are in the visible region of the spectrum. These long-lived excited states facilitate experiments that involve time-gating luminescence to obtain higher signal to noise ratios by applying a temporal delay to allow background fluorescence to decay after an excitation pulse.

2.2 Ln(III) Excitation Spectroscopy

The Eu(III) ion is one of the most useful of the Ln(III) luminescent probes [47]. Europium's utility as a probe is partly attributed to the non-degenerate ground state ($^7F_0$) and first excited state ($^5D_0$). Thus excitation through the $^7F_0 \rightarrow {}^5D_0$ transition gives rise to a single peak for each Eu(III) species because this f-f transition is not split by ligand fields. This transition is extremely useful for monitoring the solution chemistry of Eu(III) complexes. For example, coordination complex isomers that differ by ligand conformation or ligand ionization state can be distinguished. Research in Horrock's and Choppin's laboratories in particular popularized this transition for the study of the solution chemistry of Eu(III) in simple coordination complexes and in proteins [5,48–50]. In Figure 1 are shown several examples of the $^7F_0 \rightarrow {}^5D_0$ transition for different Eu(III) complexes both in solution and in the solid state. These studies show that the detection limit is 10 nM for Eu(III) aqua ion and 1 nM for Eu(EDTA) [12].

The position of the $^7F_0 \rightarrow {}^5D_0$ Eu(III) excitation peak has been related to the type of donor atom and its charge in an empirical correlation [51,52]. Anionic ligands such as carboxylates give rise to red-shifted peaks while neutral donors such as amides do not shift the excitation peak as much as the carboxylates relative to aqua ligands. The position of the excitation peak depends on the Eu(III) coordination number as well, with a reduction in coordination number giving rise to a blue shift in the excitation peak. Unfortunately, the lack of data for donor groups similar

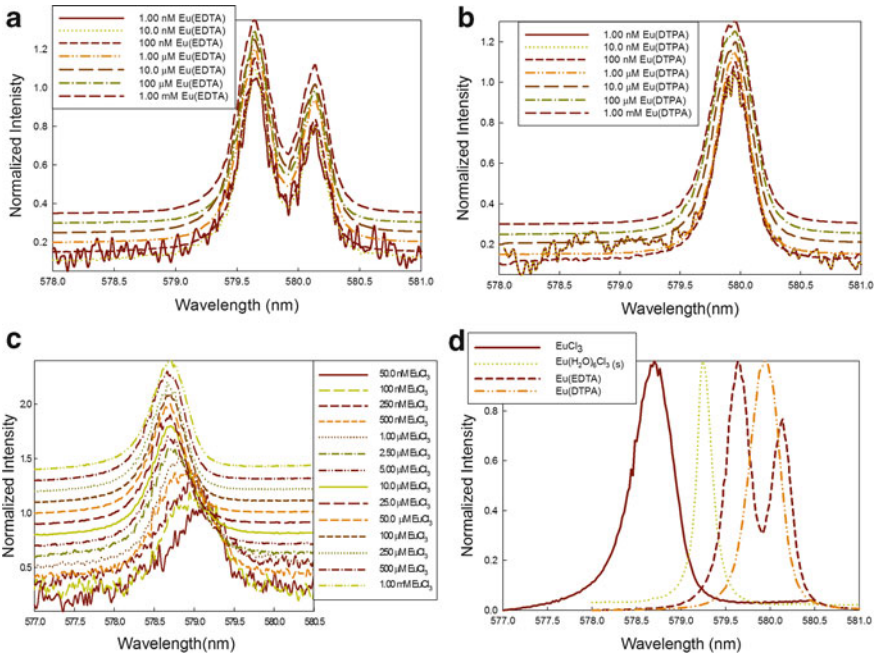


Figure 1 Excitation spectra ($\lambda_{em} = 628 \pm 27$ nm) as a function of concentration for (**a**) Eu(EDTA), (**b**) Eu(DTPA), (**c**) $EuCl_3$(aq), (**d**) comparison of 1.00 mM solutions of Eu(EDTA), Eu(DTPA), $EuCl_3$(aq), and $EuCl_3 \cdot 6H_2O$ solid excitation spectra. All complexes were in prepared in 20 mM MES buffer, pH 6.50, $I = 0.10$ M (NaCl). Spectra are offset for clarity. Reprinted from [12] with permission from the Society of Applied Spectroscopy, copyright 2009.

to those found in nucleic acids makes it difficult to use Eu(III) excitation peak position for interpreting experiments discussed below.

The MOPO/laser system makes it feasible to excite several of the Eu(III) transitions that occur in the visible region of the spectrum (Figure 2). These excitation bands have the advantage that they are more allowed and of correspondingly higher intensity than the $^7F_0 \rightarrow {}^5D_0$ transition which is forbidden also by the $\Delta J \neq 0$ rule. Thus, excitation through these transitions gives rise to more intense luminescence with a corresponding increase in sensitivity for samples. However, the drawback of this approach is that the higher energy excitation bands are split by ligand fields so that the interpretation of speciation is complicated. Excitation spectra are typically collected on the laser/MOPO system by monitoring the $^7F_2$ emission through a band pass filter.

For comparison, the excitation spectra of several different Ln(III) including Tb(III), Dy(III), and Sm(III) are also shown in Figure 2. The excitation peaks of these Ln(III) species are highly split by crystal fields in both ground and excited states. For these Ln(III) species, excitation peak differences can be distinguished for the various complexes, but these differences are not readily interpreted due to the complexity of the peaks.

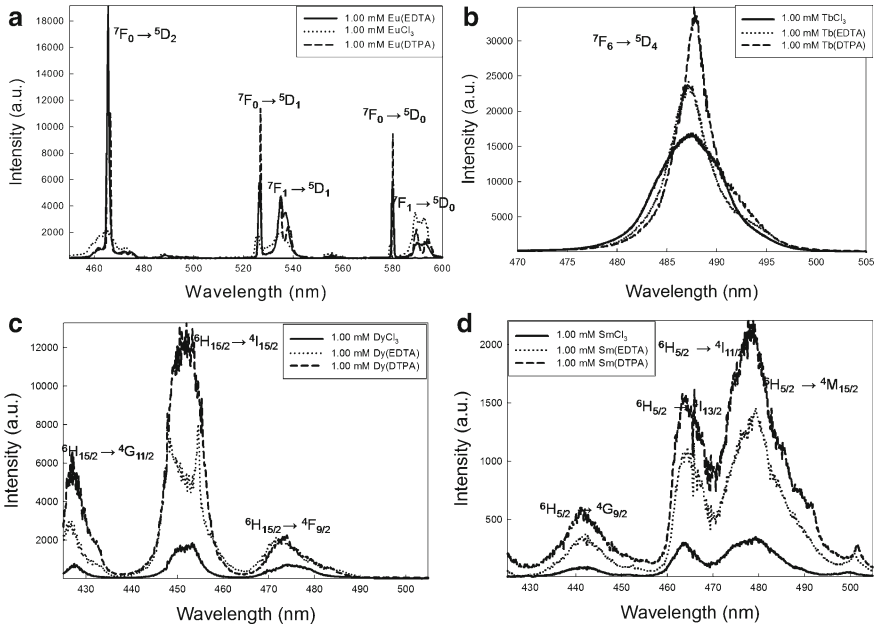


Figure 2 Excitation spectra of 1.00 mM Ln(III) complexes in 20 mM MES buffer, pH 6.50, I = 0.10 M (NaCl) using band pass filters specified in the text: (**a**) Eu(EDTA), Eu(DTPA), and $EuCl_3$(aq), (**b**) Tb(EDTA), Tb(DTPA), and $TbCl_3$(aq), (**c**) Dy(EDTA), Dy(DTPA), and $DyCl_3$ (aq), (**d**) Sm(EDTA), Sm(DTPA), and $SmCl_3$ (aq). Reprinted from [12] with permission from the Society of Applied Spectroscopy, copyright 2009.

2.3 Ln(III) Emission Spectroscopy

In aqueous solution, the higher level Ln(III) excited states are mostly quenched by non-radiative processes such as vibronic coupling to the OH of water leading to population of the lowest excited state [53]. Thus for Eu(III), most luminescence arises from the $^5D_0$ state for complexes in water regardless of the mode of excitation. The non-degenerate nature of the $^5D_0$ state simplifies the splitting of the Eu(III) emission bands. The most intense are the electric dipole induced transitions, especially the $^5D_0 \rightarrow {}^7F_2$ ($\Delta J = 2$) and the $^5D_0 \rightarrow {}^7F_4$ ($\Delta J = 4$) that have hypersensitive character. For hypersensitive transitions, the intensity of these peaks changes more dramatically in response to ligand environment than other transitions. Other emission peaks are useful as well. The $^5D_0 \rightarrow {}^7F_0$ ($\Delta J = 0$) emission peak reflects the number of Eu(III) environments, similar to excitation spectra. Also, the $^5D_0 \rightarrow {}^7F_1$ ($\Delta J = 1$) band (a magnetic dipole transition) is useful because it splits into as many as three components for strong ligand fields and may reflect the absence or presence of axial symmetry. The presence of multiple emission peaks, some of which are not sensitive to the coordination sphere such as the $\Delta J = 1$, facilitates the development of Eu(III) ratiometric sensors [9,54]. Shown in Figure 3 are emission spectra of the

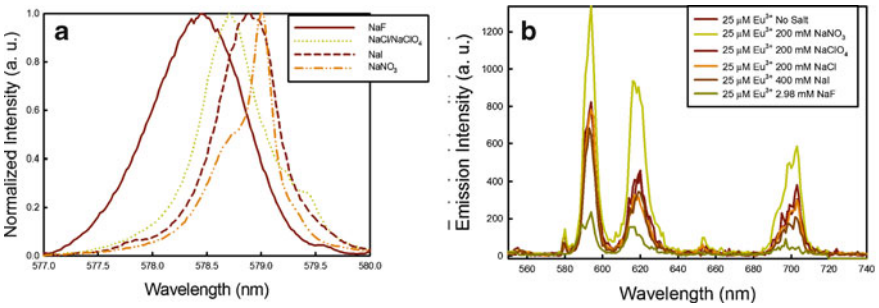


Figure 3 Emission spectra of 25.0 μM Eu(III) samples containing various salts at 100–200 mM concentrations, pH 6.5 and 20.0 mM MES. (**a**) Excitation of $^5D_2 \rightarrow {}^7F_0$ transition (~465 nm), (**b**) excitation spectra of 25.0 μM Eu(III) samples containing various salts at 100–200 mM concentrations, pH 6.5 and 20.0 mM MES. Reprinted from [55] with permission from Wiley-VCH, copyright 2009.

Eu(III) aqua ion in the presence of various simple anions including Cl^-, F^-, NO_3^{2-} [55]. Similarly, direct excitation of Tb(III), Dy(III), and Sm(III) leads to multiple emission peaks in the visible spectrum [12].

2.4 Time-Resolved Ln(III) Luminescence Spectroscopy

Luminescence intensity decays give information about the Ln(III) coordination sphere [56,57]. There is a large non-radiative quenching contribution primarily by OH groups in water as well as by other ligand groups including NH of amides or amines or OH groups of alcohols [46]. Replacement of bound OH with OD or NH with ND leads to less efficient de-excitation as a result of energy transfer to higher lying vibrational overtones of the OD oscillator than OH oscillator. Thus, luminescence lifetime experiments conducted alternately in H_2O and D_2O give information about the number of OH groups coordinated to the Ln(III). The relationship between the number of bound water molecules, q, and the difference in the rate constants for photoluminescence in H_2O compared to D_2O is given (eq. 1) for Eu(III). In general, the efficiency of quenching is proportional to the distance of the quencher from the Ln(III) and the vibrational overtone that matches the energy gap of the Ln(III). For Eu(III), eq. (1) holds with the following contributions: $\alpha = 0.25$ ms^{-1} for outersphere contribution, $\delta = 0.075$ ms^{-1} for amide NH, A = 1.2 ms [57].

$$q = A\left[k_{H_2O} - k_{D_2O} - k_{XH}\right]$$
$$k_{XH} = \alpha + \delta n_{O=CNH} \quad (1)$$

Despite efforts to precisely determine quenching parameters for Ln(III) complexes, it is important to keep in mind that there are uncertainties in these parameters that may introduce substantial errors. For example, outersphere quenching contributions

have been determined only for a few of classes of complexes [57] and there are widely different values for the NH oscillators of amines [58]. Additional outersphere and innersphere quenching contributions of yet undetermined magnitude will likely be introduced in unusual environments such as those of nucleic acids. It is also important to note that the error in the determination of luminescence lifetime decays can be as high as 8–10% for separate samples prepared and measured on different days [12].

2.5 *Luminescence Resonance Energy Transfer*

Luminescence resonance energy transfer between Ln(III) and organic fluorophores [59–61] or a second Ln(III) [6] is a useful way to determine distances between luminescent species. Luminescence resonance energy transfer (LRET) involves radiationless dipole-dipole coupled energy transfer from an excited luminophore through space to another luminophore [62,63]. For LRET, the emission band of the excited luminophore donor must overlap the absorbance band of the acceptor luminophore. The distance between the donor and acceptor molecules can be measured by using eq. (2). In eq. (2), the efficiency of energy transfer is E, τ_{DA} is the time-resolved luminescence lifetime of the donor-acceptor pair, τ_D is the lifetime of the donor, R_0 is the distance for 50% energy transfer to occur or Förster distance, and r is the calculated distance between the donor and acceptor.

$$E = 1 - \frac{\tau_{DA}}{\tau_D} = \frac{R_0^6}{R_0^6 + r^6} \tag{2}$$

Luminescence resonance energy transfer between Ln(III) ions in aqueous solutions has been used to study dimerization of Ln(III) ions through bridging ligands or to determine distances between Ln(III) ions in dinuclear systems [64–66]. LRET has also been used to map metal ion binding sites in biomolecules, especially calcium binding proteins [7,67]. These studies entail using two different lanthanide cations to estimate distances between metal ion binding sites. The Eu(III) and Nd(III) donor acceptor pair is most useful. Nd(III) as an acceptor has one of the largest extinction coefficients (~10 $M^{-1}cm^{-1}$) of all the Ln(III), facilitating LRET over longer distances. The most useful donor is Eu(III) because the $^7F_0 \rightarrow {}^5D_0$ excitation peak is not split by ligand fields and this allows for precise control over which Eu(III) species is excited. One drawback in LRET between two Ln(III) ions is that the small overlap integral leads to a small R_0, the distance at which energy transfer is 50%. A typical R_0 for a Eu(III)/Nd(III) couple in a dinuclear complex is 8 Å [66]. Likewise, for Ln(III) bound to nucleic acids, only closely spaced Ln(III) ions can be detected.

In contrast, LRET between sensitized luminescent Ln(III) donor and an organic fluorophore acceptor typically can be observed over much longer distances (≈100 Å) [59,68]. Such LRET systems have applications in the study of protein-protein interactions, ion channels in cells, and mapping Ln(III) ion binding sites in complex RNA molecules [62,69].

3 Aqueous Solution Chemistry of the Lanthanide(III) Ions

3.1 Speciation as a Function of pH

The aqueous chemistry of Ln(III) ions has been studied by a variety of techniques including potentiometric pH titrations [43,70], luminescence spectroscopy [71–74], and most recently by X-ray scattering [75]. As with aqueous solutions of most metal ions, the speciation is complicated. Under certain conditions, hydroxide complexes are present as well as multinuclear complexes formed from oligomerization. In addition, at high concentrations of salt in buffered solutions, there may be a mixture of outer-sphere and innersphere complexes of the simple anions with the buffer. A recent study in our laboratory showed that detailed information could be gained from Eu(III) excitation spectroscopy of solutions containing nanomolar to micromolar concentrations of Eu(III) under conditions of controlled pH and ionic strength [55]. This work showed that there is a speciation change between pH 6 and 7 that leads to a more highly luminescent Eu(III) species (Figure 4). This pH-dependent equilibrium was attributed to hydrolysis of the Eu(III) aqua species to form a Eu(III) hydroxide complex. Above pH 7, the appearance of new excitation peaks are consistent with the formation of additional hydroxide complexes and perhaps aggregates, but these were not investigated further.

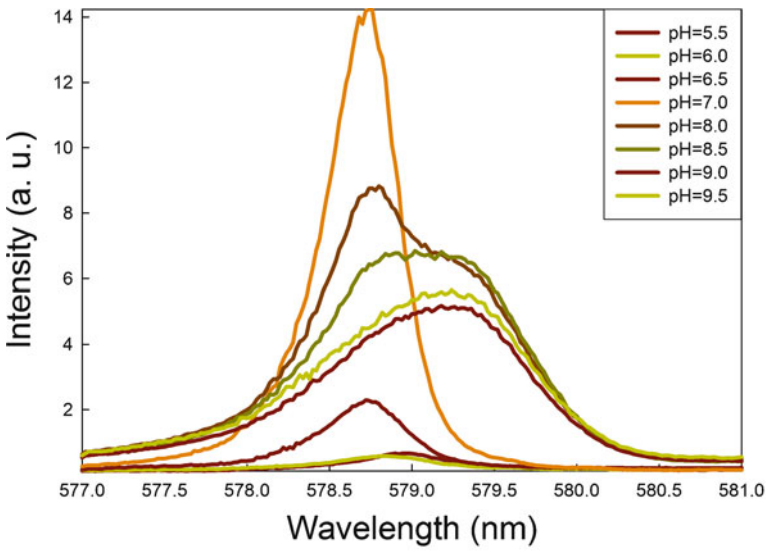


Figure 4 Excitation spectra of the luminescence 25 μM Eu(III) $^7F_0 \rightarrow {}^5D_0$ and dependence on pH I = 0.100 M (NaCl), 20.0 mM buffer. Reprinted from [55] with permission from Wiley-VCH, copyright 2009.

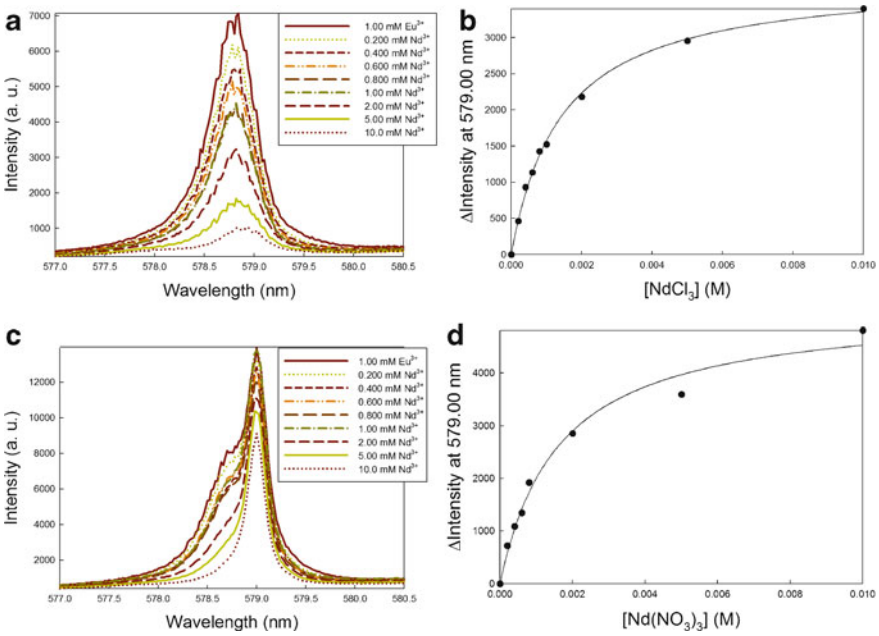


Figure 5 Nd(III) quenching of Eu(III) $^7F_0 \rightarrow {}^5D_0$ excitation spectra in 20.0 mM MES pH 6.5. (**a**) Titration of 1.00 mM Eu(III) with Nd(III), I = 0.100 M (NaCl), (**b**) binding curve of the absolute change of intensity at 579.00 nm as a function of Nd(III) concentration, I = 0.100 M (NaCl), K_d = 1.43 ± 0.07 mM, (**c**) Titration of 1.00 mM Eu(III) with Nd(III), I = 0.100 M ($NaNO_3$), (**d**) binding curve of the absolute change of intensity at 579.00 nm as a function of Nd(III) concentration, I = 0.100 M (NaCl), K_d = 1.6 ± 0.2 mM. Reprinted from [55] with permission from Wiley-VCH, copyright 2009.

3.2 Outer- and Innersphere Anions

Early work used luminescence lifetime spectroscopy to monitor Ln(III) complexes of simple anions including the halides, perchlorate, and nitrate [73,74]. However, reliance of luminescence lifetime data alone to determine inner- versus outersphere binding is problematic because many physical phenomena in addition to water quenching can lead to a change in the luminescence lifetimes, making interpretation difficult. For example, outersphere quenching contributions are not well defined and very high concentrations of anions may give non-ideal solutions. Some anions may quench by back energy transfer, electron transfer or by creation of charge transfer states from the excited state of the Ln(III) [3, 76] making it difficult to determine the actual hydration state.

Eu(III) excitation spectroscopy provides an alternative way to monitor anion binding to Eu(III) aqua species in solution [55,71,77]. These studies showed that nitrate and fluoride formed innersphere complexes as demonstrated by the appearance of new $^7F_0 \rightarrow {}^5D_0$ Eu(III) excitation peaks at pH 6.5 (Figure 5) [55]. Other anions including chloride and perchlorate did not markedly change the excitation peak

frequency at either pH 5.0 or 6.5, consistent with outersphere complex formation. Relative emission band intensities also changed most markedly with the innersphere anions, nitrate and fluoride. MES buffer was shown to bind weakly to Eu(III) in solutions containing 25 μM Eu(III) and 100 mM NaCl.

3.3 Aggregation of Ln(III) in Solution at Neutral pH

Dimerization and higher order aggregation of Ln(III) ions has been studied through LRET [77], typically from plots of quenching of Eu(III) luminescence by Nd(III). In these studies Stern–Volmer constants obtained from plots of luminescence lifetime with added quencher are compared with those determined from quenching of luminescence intensity. Differences between the two are attributed to an additional static quenching component obtained by the formation of a dimeric complex [65,78,79]. In this way it was shown that simple complexes of Ln(III) dimerize at high pH [80]. In recently published work, excitation spectroscopy of Eu(III) was used to study the formation of aggregates. A Eu(III) solution at pH 6.5 was titrated with Nd(III) quencher [55]. Efficient quenching of the excitation peak with leveling off of the luminescence intensity as the dimer completely formed was observed. This result supports the presence of multinuclear Ln(III) ion at near neutral pH. The dimerization constant of 1.46 mM is consistent with an appreciable concentration of dimer at millimolar concentrations but not at the micromolar concentrations used in the nucleic acid studies described below. Importantly, this shows that under conditions of 100 mM NaCl and MES buffer, the Eu(III) species is predominantly mononuclear. The increase in the intensity of the excitation peak from pH 5.0 to 6.5 suggests that there is some Eu(III)-hydroxide formation, but of unknown extent (Figure 4). Thus, under the conditions of our experiments that are used for studying binding of Eu(III) to nucleic acids (25 μM $EuCl_3$ in MES buffer, at pH 6.5, 100 mM NaCl) the predominant species are the mononuclear Eu(III) aqua or hydroxide complexes.

4 Lanthanide(III) Binding to RNA

The characterization of Ln(III) ion nucleic acid binding sites by using luminescence spectroscopy is quite difficult even when the full battery of experiments are used. In the studies described below, Ln(III) binding sites are characterized using Eu(III) excitation spectroscopy, nucleobase or modified nucleobase sensitized Ln(III) luminescence, Ln(III) luminescence lifetimes, and LRET between two different Ln(III) or a Ln(III) and an organic fluorophore. A common assumption is that there are a few strong binding sites that predominate under the conditions of the experiment. This assumption is supported by the presence of a single $^7F_0 \rightarrow {}^5D_0$ excitation peak or by a reasonably good statistical fit of the luminescence decay trace to a single exponential. These assumptions generally work well in simple Ln(III)

coordination complexes or in Ln(III) complexes with proteins because there are a variety of ligand donors with distinct coordination properties. However, nucleic acid ligand donor groups are not as diverse as those of proteins or coordination complexes. Thus, luminescence data alone may not fully describe the details of Ln(III) binding to nucleic acids. Toward the end of Sections 4 and 5 on nucleic acid binding, we show that additional spectroscopic techniques including mapping of binding sites by Ln(III)-induced paramagnetic shifting of DNA or RNA proton resonances may enable a more complete description of Ln(III) binding sites.

4.1 Transfer RNA

Transfer RNA was the first structurally complex RNA molecule that showed distinct Mg(II) ion binding sites by X-ray crystallography; thus, it is fitting that some of the first important studies describe binding of Ln(III) to different transfer RNA molecules. The structure of a 15-year old crystal shows one Mg(II) in the D-loop, one in the D-stem, and one at the junction of the D and TΨC loops [81]. Studies that used Ln(III) ions as mimics for Mg(II) in transfer RNA date back nearly 35 years [69,82]. Luminescence studies with tRNA are facilitated by the presence of modified bases such as 4-thiouridine in *E. coli* tRNA or wybutine in yeast tRNA that are good sensitizers for Ln(III) luminescence. These studies show that certain transfer RNAs such as *E. coli* $tRNA^{fMet,Glu}$ bind 3–4 Eu(III) ions, on average 600-fold more strongly than Mg(II) ions. Direct excitation studies undertaken in the Horrocks' laboratory showed 2–3 strongly bound Ln(III) ions with a 1.0 μM average dissociation constant in 0.5 M NaCl, pH 6 [83]. A broad $^7F_0 \rightarrow {}^5D_0$ Eu(III) excitation peak with a slight asymmetry was observed and fit to two peaks for two different Eu(III) environments. Luminescence lifetime data taken at wavelengths across the excitation peak were consistent with approximately six bound water molecules for the Eu(III) ions in these two environments. Assuming a nine-coordinate Eu(III) ion, there are three direct innersphere contacts with donor groups of the RNA. LRET studies between the wybutine base and Eu(III) were combined with LRET studies using Nd(III) quenching of Eu(III) sites to obtain further information about the number of Ln(III) binding sites. These energy transfer data were consistent with at least three independent Eu(III) sites. The discrepancy between the number of binding sites determined by LRET and the $^7F_0 \rightarrow {}^5D_0$ excitation spectra highlights a common problem: that the different RNA sites do not always produce sufficiently distinct environments to give resolved Eu(III) excitation peaks.

Direct excitation Eu(III) spectroscopy was used to study binding of Eu(III) to the isolated anticodon loop of $tRNA^{phe}$ from *E. coli* [84]. Addition of the loop to Eu(III) ion at 30 μM, 100 mM NaCl and pH 5.0, gave rise to an increase in the Eu(III) $^7F_0 \rightarrow {}^5D_0$ excitation peak. Data was most consistent with a single strong binding site with a K_d of 1 μM, four bound waters, and presumably five innersphere RNA contacts. A DNA analog bound Eu(III) 8-fold more weakly than the RNA analog and had a hydration number of five.

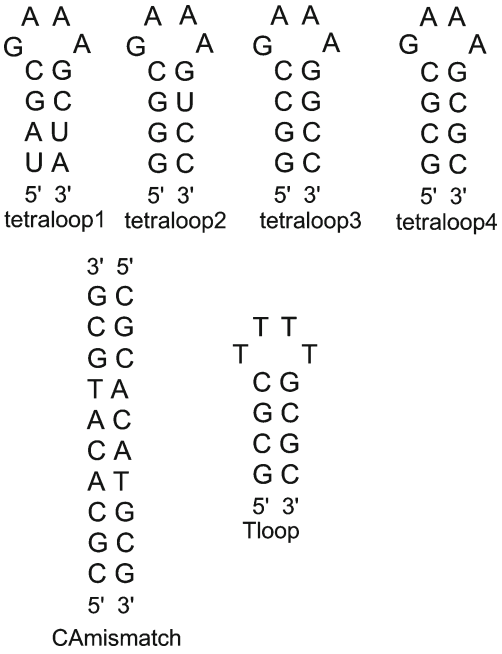
4.2 Ribozymes and Spliceosomes

The hammerhead ribozyme is an example of a small catalytic self-cleaving RNA that requires divalent metal ions for optimal activity. There are numerous elegant studies with a focus on metal ion dependence of hammerhead ribozyme folding and cleavage [85–89]. Studies in Uhlenbeck's laboratory in collaboration with Horrocks laboratory used Ln(III) luminescence as a means to study Ln(III) binding to hammerhead ribozyme [90,91]. Luminescence data from both sensitized Tb(III) luminescence and Eu(III) excitation spectroscopy were consistent with a single strong Ln(III) binding site. For example, a single symmetrical $^{7}F_{0} \rightarrow {}^{5}D_{0}$ excitation peak was observed for Eu(III) bound to hammerhead ribozyme, although the peak width at half-height was 1 nm, much broader than a typical excitation peak for a single Eu(III) species, suggesting the possibility of overlapping excitation peaks from more than one type of site. The photoluminescence decay trace of this Eu(III) species was fit to a single lifetime and was consistent with the replacement of three bound waters to give a hydration number of 6. Crystallographic studies showed a single Tb(III) ion bound close to the catalytic core of the ribozyme. Interestingly, Tb(III) inhibited cleavage of the hammerhead ribozyme, presumably by binding to a distinct site that blocked cleavage or rearrangement to the active species.

An interesting approach in Greenbaum's laboratory features LRET for monitoring binding of Tb(III) to RNA [69,92]. In this approach, irradiation at 280 nm gives both direct excitation and nucleobase sensitized Tb(III) luminescence. LRET to an appended cyanine dye gives distances between the two luminophores. In an application of this approach, Tb(III) was bound to two small nuclear RNAs that pair to form a conserved complex at the catalytic cores of the spliceosome. The authors studied the human U2-U6 complex, which contains the RNA components that are required in both steps of splicing. The binding of Tb(III) was monitored by LRET to Cy3, a member of the cyanine dye family, that was incorporated into either the 5'- or the 3'-end of the U2-U6 complex. Excitation of the Tb(III) at 280 nm led to fluorescence of the Cy3 dye through energy transfer. Analysis of LRET data was consistent with three distinct Tb(III) sites that were sufficiently close to the Cy3 dye. The presence of three Tb(III) sites was supported by luminescence intensity relationships as a function of Tb(III) ion. Also, deconvolution of the observed photoluminescence decay trace into multiple Tb(III) lifetimes reflected the identification of multiple Tb(III) environments in the absence of the energy acceptor, Cy3. The Tb(III) sites were mapped to two RNA internal loops and a four-way junction.

4.3 GAAA RNA Tetraloops

The 5'-GAAA-3' tetraloop is a member of the 5'-GNRA-3' family of loops that form unusually stable hairpins [93,94] and are found in numerous RNAs including



Scheme 2 DNA and RNA sequences that bind lanthanide(III) ions.

group I and II ribozymes [95–97]. These loops have been a popular motif for metal ion binding studies. Spectroscopic data show that Mg(II), Mn(II), $Co(NH_3)_6^{3+}$ bind to this loop [37,38,98]. DeRose's group showed that addition of Mg(II) to the tetraloop1 (Scheme 2) produces a shift in the ^{31}P NMR resonances of phosphate diesters 5' to A6, A7 and G5 [38]. Incorporation of phosphorothioate substitutions into each of the three adenosine residues of the tetraloop followed by addition of the thiophilic metal ion, Cd(II), shows that binding occurs 5' to the A7 in the loop. Ln(III) ions also bind to GAAA tetraloops as shown by luminescence spectroscopy.

Excitation spectroscopy of Eu(III) bound to a GAAA tetraloop (tetraloop2) at pH 5 to 5.5 in Greenbaum's laboratory showed an asymmetric excitation peak that could be deconvoluted into two peaks [99]. One peak was slightly blue-shifted in comparison to Eu(III) aqua ion but the second peak was highly blue-shifted. Luminescence lifetime decays were consistent with a highly dehydrated Eu(III) center with only 1–2 remaining bound waters. These are highly unusual properties for Eu(III) bound to RNA in comparison to other work and deserve comment (see below). Interestingly, a tetraloop containing all four loop nucleotides replaced with DNA and retaining the RNA stem had similar excitation peak frequencies and luminescence lifetimes. The homogeneous DNA loop, however, bound fully hydrated Eu(III) much more weakly than the RNA analog. ITC measurements were consistent with entropic driven binding for the highly dehydrated Eu(III) binding to RNA and enthalpic binding for the fully hydrated Eu(III) binding to the DNA sequence [100].

Initial studies on Eu(III) binding to RNA in our laboratory were hampered by the presence of acetate and fluoride impurities from RNA deprotection which bound

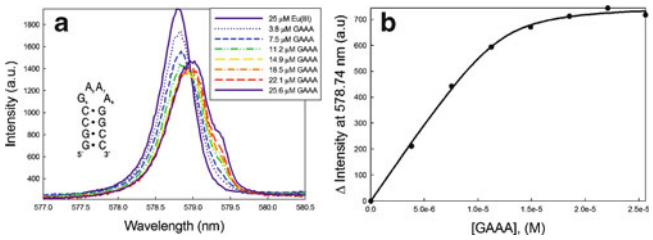


Figure 6 Excitation spectra of the $^7F_0 \rightarrow {}^5D_0$ transition of 25 μM Eu(III), 20 mM MES, pH of 6.5, I = 0.100 M (NaCl), titrated with (**a**) GAAA, (**b**) binding isotherm fit to eq. (3) from a change in the luminescence intensity at 578.74 nm of Eu(III) in the presence of the GAAA tetraloop (K_d of 0.97 ± 0.51 μM (n = 2.1 ± 0.1)).

Eu(III) as shown by unique new excitation peaks [55,101]. Extra steps were necessary to scrupulously separate the RNA from these ions prior to studies of Eu(III) binding to various tetraloop sequences. Fluoride in particular gives distinct highly blue-shifted excitation peaks, binds with micromolar dissociation constants and quenches Eu(III) luminescence in aqueous solutions to give apparent hydration states that are quite low [55]. Such quenching contributions cannot be readily distinguished from those of water ligands and may actually shorten the lifetime to less than that of fully hydrated Eu(III). Such highly blue shifted Eu(III) excitation peaks as observed by Greenbaum (578.40 nm) in solutions containing RNA deprotected by using tetraethyl ammonium fluoride are likely to arise from the presence of Eu(III) fluoride complexes [99].

Our work on binding of Eu(III) to the GAAA tetraloop (GGCCGAAAGGCC, tetraloop3) used a slightly different stem sequence than that of Greenbaum's group and was carried out at pH 6.5, 100 mM NaCl, and 20 mM MES buffer with 25 μM Eu(III) (Scheme 2) [101,102]. Addition of tetraloop3 to Eu(III) gave excitation peak frequencies that had decreased intensities and were only slightly red-shifted from Eu(III) aqua ion in NaCl (Figure 6). There are two distinct excitation peaks representing at least two different Eu(III) ion environments. The Eu(III) species bound to GAAA tetraloop3 are only slightly dehydrated. A hydration number close to 8 was determined upon excitation both at the major excitation peak maximum and at the more red-shifted shoulder. Binding curves were fit to a two site model (eq. (3)) with both Eu(III) having an apparent K_d of 1 μM for tetraloop3 [103]. Here LS is the concentration of the Eu(III)-nucleic acid complex, K_d is the dissociation constant, M is the total concentration of the Eu(III), A is the nucleic acid concentration, n_s is the number of identical binding sites, x_M is the mole fraction of M, x_{LS} is the mole fraction of LS, I_M is the luminescence intensity of free Eu(III), and I_{LS} is the intensity of LS.

$$[LS] = 0.5\left(K_d + M + n_s A - \sqrt{(K_d + M + n_s A)^2 - 4 n_S MA} \right)$$

$$I_{obs} = x_M I_M + x_{LS} I_{LS} \qquad (3)$$

Consistent with the two metal ion site model, LRET studies with Eu(III) donor and Nd(III) acceptor bound to tetraloop3 were consistent with two metal ions

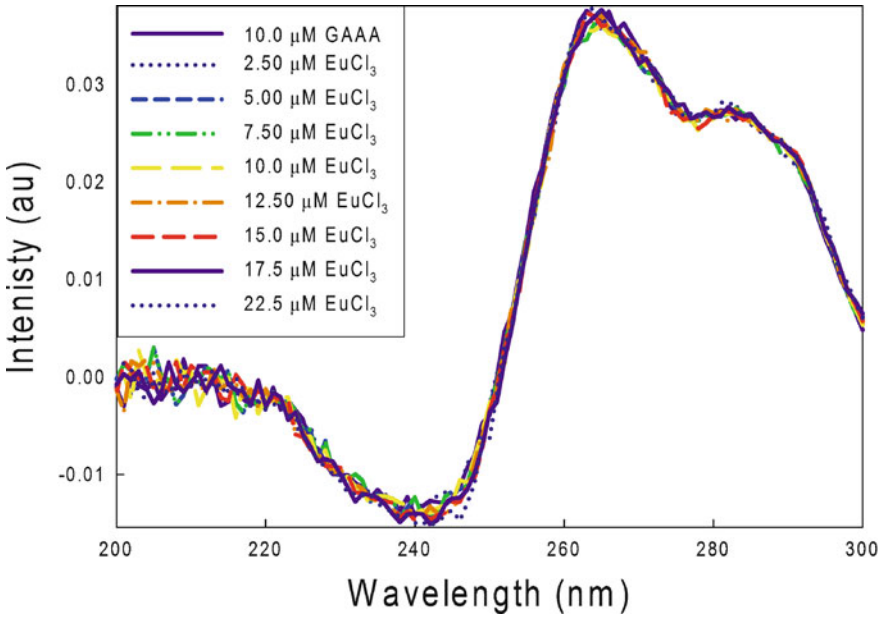


Figure 7 Circular dichroism spectra of 10 µM GAAA tetraloop RNA with 0 to 23 µM Eu(III). All solutions contain 20.0 mM MES, pH of 6.5, and $I = 0.100$ M (NaCl).

approximately 6.3 Å apart. Small sequence changes in the stem region (tetraloop4) gave a similar Eu(III) excitation spectrum upon addition of the tetraloop and similar metal ion binding properties. Binding of Eu(III) to tetraloop3 did not markedly change the CD spectrum (Figure 7). This shows that the Eu(III) ion does not denature the RNA hairpin loop upon binding.

Paramagnetic Ln(III) ions induce long range pseudo-contact shifts (PCS) that are mediated by through-space interactions of the paramagnetic Ln(III) with nuclear spins such as protons (^{1}H) or phosphorous (^{31}P) nuclei as well as other nuclei. The work of the Bertini and Otting laboratories have successfully utilized paramagnetic Ln(III) ions for the determination of NMR solution state structures of proteins [45,104–106]. The pseudocontact shift varies for each Ln(III) with certain Ln(III) ions influencing nuclear spins up to 40 Å away [44]. Ln(III)-induced line broadening through relaxation of protons adjacent to the paramagnetic center also occurs, but generally over a shorter distance than do PCS. Thus, Ln(III) ions are excellent paramagnetic probes for mapping binding sites through inducing chemical shift changes of proximal nuclei.

The NMR structure of tetraloop3 was solved in order to make proton NMR assignments. Three different paramagnetic Ln(III) ions were chosen for titration into tetraloop3 [101,102]. Ce(III) and Eu(III) were chosen for their moderate PCS and Yb(III) was chosen for its longer range PCS. A titration with Eu(III) ion is shown in Figure 8. A control titration using the diamagnetic cation La(III) was also carried out

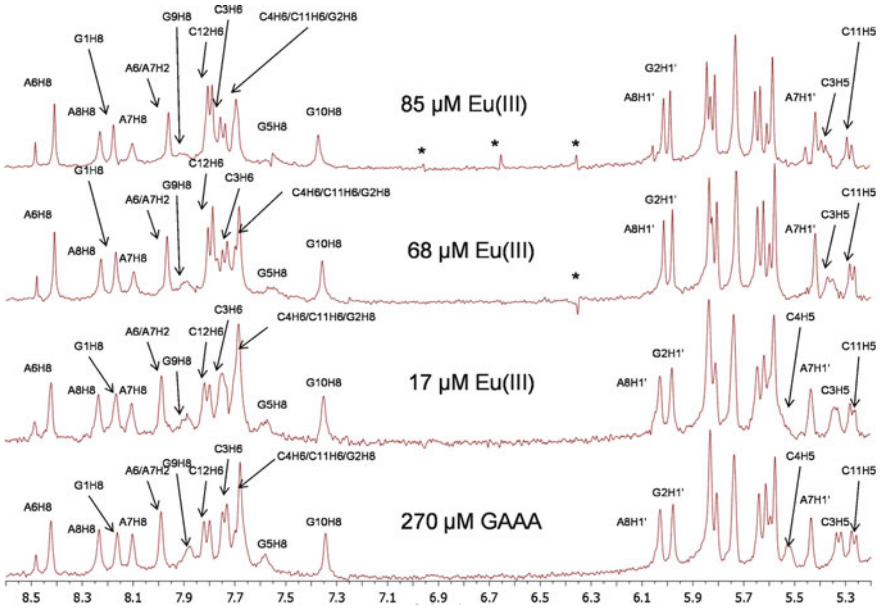


Figure 8 ^{1}H NMR spectra PCS shifts from titrations of 270 μM GAAA tetraloop titrated with Eu(III) 2.0 mM MES, pH of 6.5, $I = 0.100$ M (NaCl) in D_2O at 20.5°C. Artifacts are marked with *. Each spectrum has been normalized to the peak of greatest intensity for clarity.

to enable us to extract information on the distance between the paramagnetic Ln(III) and the shifted protons. This is to account for the possible change in secondary or tertiary structure of the nucleic acid upon binding Ln(III) cations. The NMR titration data was consistent with two Ln(III) sites including one in the GAAA tetraloop in close contact with the O6 of G5 of the sheared GA base pair. The second was in the stem on the other side of the closing CG base pair interacting with the C3 and C4 bases in the minor groove.

4.4 Double-Stranded RNA

Binding of Eu(III) to the fully self-complementary sequences GGCCGGCC and GCGCGCGC was studied for comparison to the tetraloops [101,102]. Upon binding to the double-helical RNA, the Eu(III) excitation peak red-shifted with a decrease in luminescence intensity. The excitation peaks for both sequences could be fit to two peaks, consistent with at least two types of Eu(III) environments. Luminescence lifetimes taken on the left of the peak gave hydration numbers of 8–9 while lifetimes on the red side of the peak gave hydration numbers close to 7, reflecting two different types of sites.

5 Lanthanide(III) Binding to DNA

5.1 Single-Stranded and Double-Stranded DNA

Early studies suggested that single-stranded nucleic acids were more effective at sensitizing Tb(III) than double-stranded nucleic acids, presumably due to wrapping of the flexible single-stranded DNA around Tb(III) [16,17]. This wrapping interaction may position Tb(III) closer to the nucleobase sensitizer. All bases were shown to sensitize Tb(III) luminescence although guanosine containing sequences appear to be the most efficient sensitizers. Early Eu(III) excitation studies showed that oligodeoxynucleotides including homopolymers of G, A, C, T bound to Eu(III) at pH 6.5 to give broad $^7F_0 \rightarrow {}^5D_0$ excitation peaks [107,108]. Different classes of sites were proposed based on the distinct lifetimes obtained as a function of DNA. The apparent number of bound waters varied from two for the strongly bound sites to six or seven for the weakly bound sites. However, the ill-defined structure of these single-stranded sequences makes it difficult to determine the nature of the binding site.

More recent studies in the Morrow group show that Eu(III) binds to self-complementary double helical DNA including the 24-mers $d(GCGC)_6$ and $d(GGCC)_6$ [109]. As with the RNA GC rich sequences, the Eu(III) excitation peak decreases in intensity and red-shifts (Figure 9). Also similar to the RNA analog, the luminescence lifetimes are relatively short, even slightly shorter than full hydrated Eu(III). Lifetimes of 105 µs are observed for bound Eu(III) compared to 115 µs for the aqua ion. This reproducibly shorter lifetime is attributed to an environment that gives rise to unusually strong outersphere or other quenching contributions. Apparent q numbers are 9 and 10 for Eu(III) bound to the self-complementary sequences $(GCGC)_6$ and $(GGCC)_6$, respectively.

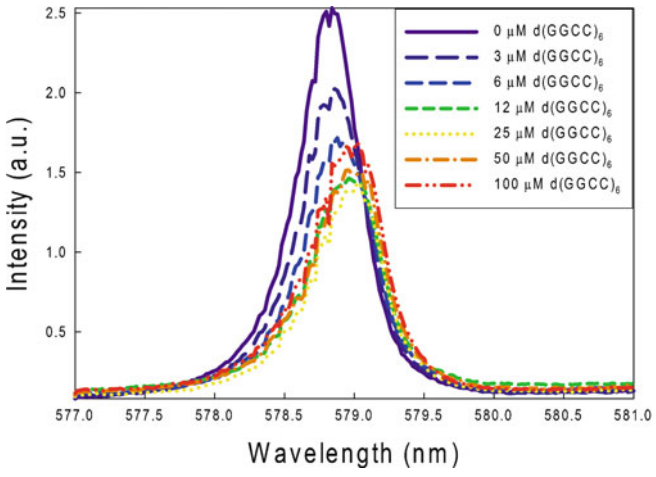


Figure 9 The steady-state excitation spectra of Eu(III) with $d(GGCC)_6$ DNA at pH = 6.5 and I = 0.10 M (NaCl) and [Eu(III)] = 25 µM.

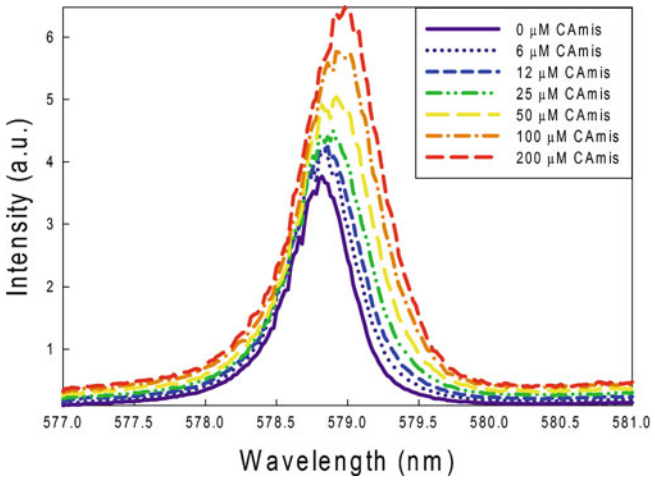


Figure 10 The steady-state excitation spectra of Eu(III) with the CA mismatch DNA (CAmis) sequence at pH = 6.5, I = 0.10 M (NaCl), and [Eu(III)] = 25 μM.

5.2 *Base Mismatches*

The importance of forming a Ln(III) ion DNA binding site in order to obtain efficient sensitization was verified by recent studies in the laboratories of Turro and Lu [15,110]. Fully complementary DNA helices did not sensitize luminescence, but guanine bases in mismatches that may produce binding sites, including GG and GA tandem mismatches, effectively sensitized Tb(III) ions. This Tb(III) luminescence sensitization requirement may provide a means to identify mismatched sequences. Recent studies in Yi Lu's lab [110] showed that DNAzyme sensitized Tb(III) luminescence. Sensitization of Tb(III) luminescence was dependent on the presence of a GT wobble base pair in the DNAzyme substrate complex. Replacement of the GT wobble with a GC Watson-Crick base pair led to a decrease in sensitized Tb(III) luminescence. Similarly, replacement of the strand containing the DNAzyme with a fully complementary DNA strand reduced Tb(III) luminescence. Luminescence lifetime experiments of Tb(III) bound to the DNAzyme were consistent with a q of close to 7, consistent with two innersphere contacts with RNA. Similar to the self-cleaving hammerhead ribozyme, Tb(III) strongly inhibited cleavage of the substrate by the DNAzyme with a K_i of 3.3 μM in the presence of 100 μM Zn(II).

Consistent with these results, work in the Morrow laboratory showed that DNA structures containing non-cannonical Watson-Crick base pairs such as tandem base mismatches produce an increase in the Eu(III) excitation peak intensity and dehydration of the Eu(III) ion [109,111]. For example, the $^7F_0 \rightarrow {}^5D_0$ excitation peak of Eu(III) is shown as a function of concentration of DNA containing a tandem CA mismatch (Figure 10, Scheme 2). This shows that disruption of the normal double-helical

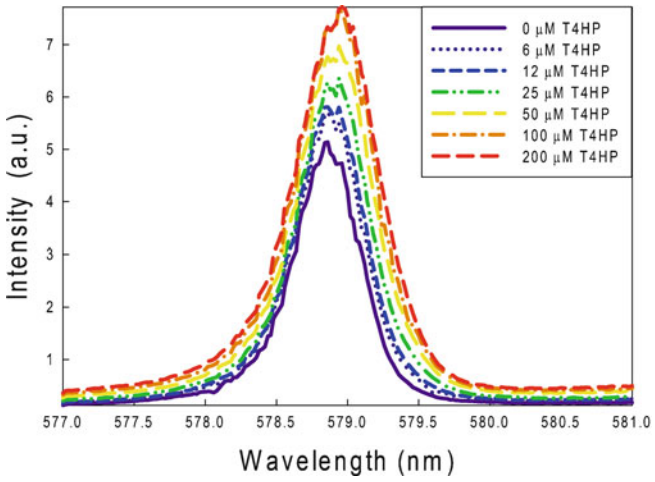


Figure 11 The steady-state excitation spectra of Eu(III) and the T4 hairpin (T4HP) DNA at pH = 6.5, I = 0.10 M (NaCl), and [Eu(III)] = 25 μM.

structure by a mismatch leads to Eu(III) binding sites with an apparent q of 6 and three direct DNA contacts. Fitting the data to a 1:1 binding isotherm [112] gives a dissociation constant of 44 μM. Optical thermal melting experiments show that the Eu(III) ion does not destabilize the duplex upon binding.

5.3 Hairpin Loops and Quadruplexes

Both Tb(III) and Eu(III) bind to guanine rich oligonucleotides that fold into four-stranded structures known as G-quadruplexes [66]. Binding of Tb(III) to the G-quadruplexes sensitizes Tb(III) luminescence. It is proposed that Tb(III) binds tightly to the cavity in the quadruplex which serves to place Tb(III) close to the guanine sensitizers. Tb(III) is also proposed to bind, albeit more weakly, to the TTA loops.

Hairpins in DNA that contain relatively flexible loops are sites for Ln(III) ion binding. The hairpin loop (Scheme 2) containing four thymines in the loop with a GC containing stem binds to Eu(III) and increases the Eu(III) $^7F_0 \rightarrow {}^5D_0$ excitation peak intensity (Figure 11) [109]. Luminescence lifetime studies are consistent with a slightly dehydrated Eu(III) with a q of 7. A binding constant of 33 μM was obtained assuming a 1:1 binding isotherm. The increase of luminescence intensity for the T4 loop, compared to the decrease in luminescence intensity for the stem sequence alone suggests that the loop has unique Eu(III) ion binding sites. This postulate is supported by ^{1}H NMR titrations of the T4 loop with Eu(III) which shows a pronounced chemical shift change of the thymidine protons (0.3 ppm for T6 and 0.1 ppm for T5) in the presence of the paramagnetic Eu(III) center.

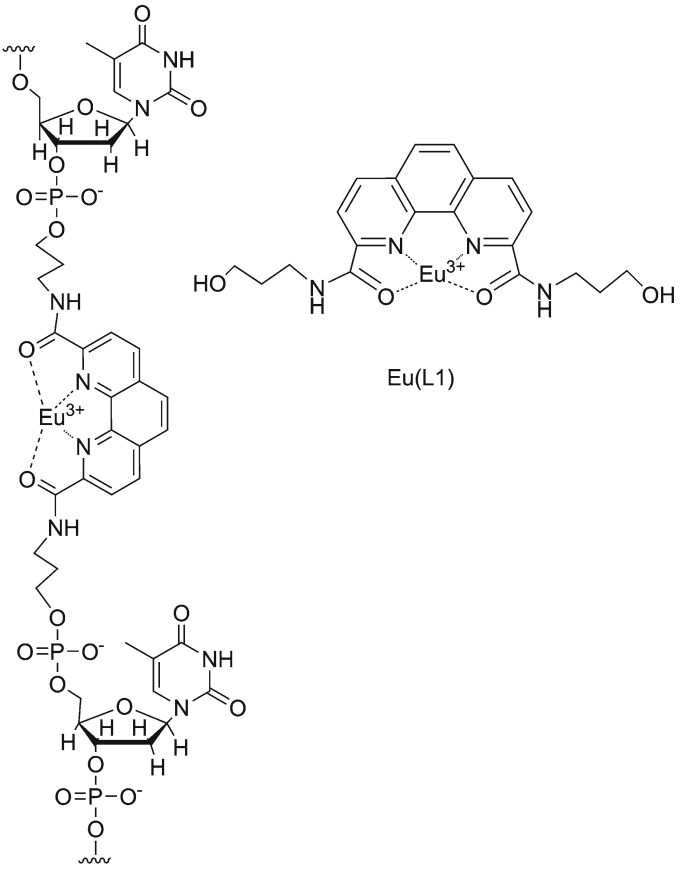
6 Lanthanide Ion Binding to Modified Nucleic Acids Containing Additional Ligand Donor Groups

6.1 Macrocyclic Ligands Conjugated to DNA

Macrocyclic ligands and their Ln(III) complexes have been conjugated to DNA to study artificial nuclease-catalyzed cleavage of complementary RNA [113–117]. These macrocycles bind the Ln(III) ion tightly under physiological conditions. Thus, most coordination sites of the Ln(III) are filled with macrocycle donor groups, but a few sites contain water ligands that can be displaced upon binding of the phosphate diesters of RNA. Coordination of the phosphate ester to the Ln(III) complex is an important first step in the catalytic cleavage of RNA [118]. Most studies have focused on RNA cleavage activity and not the Ln(III) complex as a spectroscopic probe. However, recent studies on the characterization of a Eu(III) macrocyclic complex conjugated to DNA or to a DNA/RNA chimera show that Eu(III) is bound both to the macrocycle and to additional groups on the DNA or RNA. Binding to groups other than the phosphate diester may inhibit cleavage by blocking available coordination sites for catalysis [102].

6.2 DNA Non-nucleosidic Linkers for Ln(III) Binding

An area of increasing research interest is the incorporation into DNA of nucleobase-like ligands or non-nucleosidic linkers that bind metal ions [119,120]. This facilitates the preparation of new molecules that use the self-assembly properties of DNA to form large metal ion containing structures and materials. Most reports to date involve the incorporation of ligands that bind transition metal ions or Zn(II). Recently, we reported an example of a non-nucleosidic linker that bound Eu(III) [111]. In this study, a phenanthroline derivative is incorporated into the DNA backbone in place of a nucleotide (Scheme 3). The Eu(III) excitation peak frequency, q numbers, and phenanthroline-sensitized luminescence are all consistent with Eu(III) binding to the non-nucleosidic linker rather than the natural DNA donor groups. A series of double-stranded derivatives were prepared that contained the phenanthroline linker in different bulge like positions. Eu(III) binding constants vary with the structure and position of the linker in DNA. Binding to the free ligand, L1, is relatively weak at 46 μM, while binding to the linker in a bulge-like structure is 2 μM. The origin of this remarkable change in binding properties with the placement of the linker is not known, but shows promise for the design of new types of self-assembling Ln(III) materials with unique properties.



Scheme 3 A Eu(III) complex and its incorporation as a non-nucleosidic linker into the DNA backbone.

7 Concluding Remarks

Lanthanide luminescence spectroscopy adds a battery of useful experiments for the study of metal ion binding sites in nucleic acids. A wide range of experiments are available that give information on the type of binding sites and their distance from other fluorophores including other Ln(III) ions. Sensitized luminescence through either naturally occurring nucleobases or modified bases are useful for studying Ln(III) sites. Alternatively, direct excitation Eu(III) luminescence, especially of the $^7F_0 \rightarrow {}^5D_0$ excitation peak gives information on the number of different Ln(III) environments and produces luminescence of Ln(III) ion in sites that do not have sensitizing antennae. However, it is clear that the different Eu(III) environments produced by the ligand donor groups of nucleic acids do not change the excitation peak position sufficiently to entirely resolve different species. Further information

is available by combining structural studies from NMR spectroscopy with paramagnetic shift properties of the Ln(III) ions to delineate the nucleic acid binding sites. Development of these additional techniques may lead to a more complete picture of metal ion binding to nucleic acids.

Abbreviations

CD	circular dichroism
Cy3	a cyanine dye
DTPA	diethylenetriamine pentaacetate
EDTA	ethylenediamine N,N,N',N'-tetraacetate
ITC	isothermal titration calorimetry
Ln(III)	trivalent lanthanide ion
LRET	luminescence resonance energy transfer
MES	2-morpholinoethanesulfonic acid buffer
MOPO	master oscillator power oscillator
NLBP	non-linear Poisson-Boltzmann equation
PCS	pseudo contact shift
q	number of bound water molecules
tRNA	transfer RNA

Acknowledgment We acknowledge Dr. Ryan Matthews of the University at Buffalo and Dr. Matthew Fountain of the State University of New York, College at Fredonia, for contributing to this manuscript. We gratefully acknowledge the National Science Foundation (CHE0911375) for support of this work and for a major instrumentation award (CHE-0321058) to build the MOPO laser system.

References

1. W. D. J. Horrocks, Jr., D. R. Sudnick, *Acc. Chem. Res*. **1981**, *14*, 384–392.
2. W. D. Horrocks, Jr., *Adv. Inorg. Biochem*. **1982**, *4*, 201–261.
3. J.-C. G. Bünzli, *Acc. Chem. Res*. **2006**, *39*, 53–61.
4. S. V. Eliseeva, J.-C. G. Bünzli, *Chem. Soc. Rev*. **2010**, *39*, 189–227.
5. G. R. Choppin, D. R. Peterman, *Coord. Chem. Rev*. **1998**, *174*, 283–299.
6. W. D. Horrocks, Jr., W. E. Collier, *J. Am. Chem. Soc*. **1981**, *103*, 2856–2862.
7. W. D. Horrocks, Jr., M.-J. Rhee, A. P. Snyder, D. R. Sudnick, *J. Am. Chem. Soc*. **1980**, *102*, 3650–3652.
8. E. G. Moore, A. P. S. Samuel, K. N. Raymond, *Acc. Chem. Res*. **2009**, *42*, 542–552.
9. C. P. Montgomery, B. S. Murray, E. J. New, R. Pal, D. Parker, *Acc. Chem. Res*. **2009**, 42, 925–937.
10. V. S. Sastri, J.-C. G. Bünzli, R. Rao, G. V. S. Rayudu, J. R. Perumareddi, *Modern Aspects of Rare Earths and Their Complexes*, 1st edn., Elsevier B. V., Amsterdam, The Netherlands, 2003.
11. W. D. Horrocks, Jr., *J. Am. Chem. Soc*. **1977**, *99*, 2378.
12. C. M. Andolina, W. G. Holthoff, P. M. Page, R. A. Mathews, J. R. Morrow, F. V. Bright, *Appl. Spectroscopy* **2009**, *63*, 483–493.

13. A. P. S. Samuel, J. Xu, K. N. Raymond, *Inorg. Chem*. **2009**, *48*, 687–698.
14. J.-C. G. Bünzli, A.-S. Chauvin, C. D. B. Vandevyver, S. Bo, S. Comby, *Ann. New York Acad. Sci*. **2008**, *1130*, 97–105.
15. P. L. Fu, C. Turro, *J. Am. Chem. Soc*. **1999**, *121*, 1–7.
16. D. P. Ringer, S. Burchett, D. E. Kizer, *Biochemistry* **1978**, *17*, 4818–4825.
17. M. D. Topal, J.R. Fresco, *Biochem*. **1980**, *19*, 5531–5537.
18. D. E. Draper, *RNA* **2004**, *10*, 335–343.
19. D. E. Draper, D. Grilley, A. M. Soto, *Ann. Rev. Biophys. Biomol. Struct*. **2005**, *34*, 221–243.
20. V. K. Mirsa, D. E. Draper, *Proc. Nat. Acad. Sci. USA* **2001**, *98*, 12456–12461.
21. S. A. Pabit, X. Qiu, J. S. Lamb, L. Li, S. P. Meisburger, L. Pollack, *Nucl. Acid. Res*. **2009**, ***37***, 3887–3896.
22. E. Ennifar, P. Walter, P. Dumas, *Nucl. Acid. Res*. **2003**, *31*, 2671–2682.
23. E. Freisinger, R. K. O. Sigel, *Coord. Chem. Rev*. **2007**, *251* 1834–1851.
24. R. K. O. Sigel, S. Gallo, *Chimia* **2010**, *64*, 126–131.
25. R. K. O. Sigel, H. Sigel, *Acc. Chem. Res*. **2010**, *43*, 974–984.
26. A. S. Petrov, J. C. Bowman, S. C. Harvey, L. D. Williams, *RNA* **2011**, *17*, 291–297.
27. D. J. Klein, P. B. Moore, T. A. Steitz, *RNA* **2006**, *10*, 1366–1379.
28. B. Gong, Y. Chen, E. L. Christian, J.-H. Chen, E. Chase, D. M. Chadalavada, R. Yajima, B. L. Golden, P. C. Bevilacqua, P. R. Carey, *J. Am. Chem. Soc*. **2008**, *130*, 9670–9672.
29. E. L. Christian, V. E. Anderson, P. R. Carey, M. E. Harris, *Biochemistry* **2010**, *49*, 2869–2879.
30. V. J. DeRose, in *Nucleic Acid-Metal Ion Interactions*, Ed N. V. Hud, RSC Publishing, Cambridge, 2009, pp 154–179.
31. S. R. Morrissey, T. E. Horton, V. J. DeRose, *J. Am. Chem. Soc*. **2000**, *122*, 3473–3481.
32. S. R. Morrissey, T. E. Horton, C. V. Grant, C. G. Hoogstraten, R. D. Britt, V. J. DeRose, *J. Am. Chem. Soc*. **1999**, *121*, 9215–9218.
33. S. E. Butcher, F. H.-T. Allain, J. Feigon, *Biochemistry* **2000**, *39*, 2174–2182.
34. N. V. Hud, J. Feigon, *Biochemistry* **2002**, *41*, 9900–9910.
35. Z. Gdaniec, H. Sierzputowska-Gracz, E. C. Theil, *Biochemistry* **1998**, *37*, 1505–1512.
36. J. H. Davis, T. R. Foster, M. Tonelli, S. E. Butcher, *RNA* **2007**, *13*,76–86.
37. S. Rüdisser, I. Tinoco, Jr., *J. Molec. Biol*. **2000**, *295*, 1211–1223.
38. M. Maderia, T.E. Horton, V. J. DeRose, *Biochemistry* **2000**, *39*, 8193–8200.
39. M. Maderia, L.M. Hunsicker, V. J. DeRose, *Biochemistry* **2000**, *39*, 12113–12120.
40. Y. Tanaka, C. Kojima, E. H. Morita, Y. Kasai, K. Yamasaki, A. Ono, M. Kainosho, K. Taira, *J. Am. Chem. Soc*. **2002**, *124*, 4595–4601.
41. Y. Tanaka, K. Taira, *Recent Research Developments in Organic Chemistry* **2005**, *9*, 93–118.
42. G. Wang, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc*. **2004**, *126*, 8908–8909.
43. J. Burgess, *Ions in Solution: Basic Principles of Chemical Interactions*, Horwood, West Sussex, 1999.
44. C. F. G. C. Geraldes, C. Luchinat, in *The Lanthanides and Their Interrelations with Biosystems*, Vol. 40 of *Metal Ions in Biological Systems*, Eds A. Sigel, H. Sigel, Dekker, NY, 2003, pp. 513–588.
45. G. Pintacuda, M. John, X.-C. Su, G. Otting, *Acc. Chem. Res*. **2007**, *40*, 206–212.
46. G. Stein, E. Wurzberg, *J. Chem. Phys*. **1975**, *62*, 208–213.
47. W. D. Horrocks, D. R. Sudnick, *Science* **1979**, *206*, 1194–1196.
48. S. Amin, D. A. Voss, Jr; W. D. Horrocks, Jr., C. H. Lake, M. R. Churchill, J. R. Morrow, *Inorg. Chem*. **1995**, ***34***, 3294–3300.
49. S. T. Frey, C. A. Chang, J. F. Carvalho, A. Varadarajan, L. M. Schultze, K. L. Pounds, W. D. Horrocks, *Inorg. Chem*. **1994**, *33*, 2882–2889.
50. D. M. Epstein, L. L. Chappell, H. Khalili, R. M. Supkowski, W. D. Horrocks, Jr., J. R. Morrow, *Inorg. Chem*. **2000**, *39*, 2130–2134.
51. S. T. Frey, W. D. Horrocks, Jr., *Inorg. Chim. Acta* **1995**, ***229***, 383–390.
52. G. R. Choppin, Z. M. Wang, *Inorg. Chem*. **1997**, *36*, 249–252.
53. Y. Haas, G. Stein, E. Wurzberg, *J. Chem. Phys*. **1974**, *60*, 258–265.
54. R. Pal, D. Parker, *Org. Biomol. Chem*. **2008**, *6*,1020–1033.

55. C. M. Andolina, R. A. Mathews, J. R. Morrow, *Helv. Chim. Acta* **2009**, *92*, 2330–2348.
56. W. D. Horrocks, D. R. Sudnick, *J. Am. Chem. Soc*. **1979**, *101*, 334–340.
57. A. Beeby, I. M. Clarkson, R. S. Dickins, S. Faulkner, D. Parker, L. Royle, A. S. de Sousa, J. A. G. Williams, M. Woods, *J. Chem. Soc., Perkin Trans. 2* **1999**, 493–504.
58. K. Nwe, J. P. Richard, J. R. Morrow, *J. Chem. Soc., Dalton Trans.* **2007**, 5171–5178.
59. J. G. Reifernberger, P. R. Selvin, *Rev. Fluorescence* **2005**, *2*, 399–431.
60. P. R. Selvin, J. E. Hearst, *Proc. Nat. Acad. Sci. USA* **1994**, *91*, 10024–10027.
61. F. Yuan, N. L. Greenbaum, *Spectrum* **2007**, *20*, 14–17.
62. P. R. Selvin, *Ann. Rev. Biophys. Biomolec. Struc*. **2002**, *31*, 275–302.
63. T. Förster, *Ann. Phys*. **1948**, *2*, 55.
64. H. G. Brittain, *Inorg. Chem*. **1978**, *18*, 1740–1745.
65. L. Spaulding, H. G. Brittain, *Inorg. Chem*. **1983**, *22*, 3486–3488.
66. C. M. Andolina, J. R. Morrow, *Eur. J. Inorg. Chem.* **2011**, 154–164.
67. M.-J. Rhee, D. R. Sudnick, V. K. Arkle, W. D. J. Horrocks, Jr., *Biochemistry* **1981**, *20*, 3328–3334.
68. P. R. Selvin, *IEEE Journal of Selected Topics in Quantum Electronics* **1996**, *2*, 1077–1087.
69. F. Yuan, N. L. Greenbaum, *Methods* **2010**, *52*, 173–179.
70. C. F. Baes, R. E. Mesmer, *The Hydrolysis of Cations*, 2nd edn., Robert E. Krieger Publishing Company, Inc., Malabar, Florida, 1976.
71. P. J. Breen, W. D. Horrocks, Jr., *Inorg. Chem*. **1983**, *22*, 536–540.
72. K. Cernochova, J. N. Mathur, G. R. Choppin, *Radiochim. Acta* **2005**, *93*, 733–739.
73. I. Billard, in *Handbook on the Physics and Chemistry of Rare Earths*, Vol. 33, Eds K. A. Gscheidner, Jr., J.-C. G. Bünzli, V. K. Pecharsky, Elsevier, Amsterdam, The Netherlands, 2003, 465–514.
74. A. Nehlig, M. Elhabiri, I. Billard, A.-M. Albrecht-Gary, K. Lutzenkirchen, *Radiochim. Acta* **2003**, *91*, 37–43.
75. C. H. Huang, J. Hammell, S. J. Ratnakar, A. D. Sherry, J. R. Morrow, *Inorg. Chem*. **2010**, *49*, 5963–5970.
76. F. Kielar, C. P. Montgomery, E. J. New, D. Parker, R. A. Poole, S. L. Richardson, P. A. Stenson, *Org. Biomol. Chem*. **2007**, *5*, 2975–2982.
77. I. Sanchez-Lombardo, C. M. Andolina, J. R. Morrow, A. K. Yatsimirsky, *Dalton Trans*. **2010**, *39*, 864–873.
78. L. Spaulding, H. G. Brittain, *Inorg. Chem*. **1984**, *23*, 2165–2170.
79. S. Amin, D.A. Voss, Jr., W. D. Horrocks., Jr., J. R. Morrow, *Inorg. Chem.* **1996**, *35*, 7466–7467.
80. G. Hernandez, H. G. Brittain, M. F. Tweedle, R. G. Bryant, *Inorg. Chem*. **2010**, *29*, 864–873.
81. L. Jovine, S. Djordjevic, D. Rhodes, *J. Molec. Biol*. **2000**, *301*, 401–414.
82. J. M. Wolfson, D. R. Kearns, *Biochemistry* **1975**, *14*, 1436–1444.
83. E. H. Cornwall, *Ph. D. Thesis*, Pennsylvania State University, University Park, PA,1993.
84. N.L. Greenbaum, C. Mundoma, D. R. Peterman, *Biochemistry* **2001**, *40*, 1124–1134.
85. J. Schnabl, R. K. O. Sigel, *Curr. Opin. Chem. Biol*. **2010**, *14*, 269–275.
86. R. K. O. Sigel, A. M. Pyle, *Chem. Rev*. **2007**, *107*, 97–113.
87. J. L. O'Rear, S. Wang, A. L. Feig, L. Beigelman, O. C. Uhlenbeck, D. Herschlag, *RNA* **2001**, *7*, 537–545.
88. T. Humphry, S. Iyer, O. Iranzo, J. R. Morrow, J. P. Richard, P. Paneth, A. C. Hengge, *J. Am. Chem. Soc*. **2008**, *130*, 17858–17866.
89. J. R. Morrow, *Comments Inorg. Chem*. **2008**, *29*, 169–188.
90. A. L. Feig, M. Panek, W. D. Horrocks, Jr, O. C. Uhlenbeck, *Chem. Biol*. **1999**, *6*, 801–810.
91. A. L. Feig, W. G. Scott, O. C. Uhlenbeck, *Science* **1998**, *279*, 81–84.
92. F. Yuan, L. Griffin, L. Phelps, V. Buschmann, K. Weston, N. L. Greenbaum, *Nucl. Acid. Res*. **2007**, *35*, 2833–2845.
93. K. Nwe, C. M. Andolina, J. R. Morrow, *J. Am. Chem. Soc*. **2008**, *130*, 14861–14871.
94. H. A. Heus, A. Pardi, *Science* **1991**, *253*, 191–194.
95. M. Costa, F. Michel, *EMBO J* **1995**, ***14***, 1276–1285.
96. F. L. Murphy, T. R. Cech, *J. Molec. Biol*. **1994**, *263*, 49–63.

97. M.-Y. Yang, J. R. Morrow, J. P. Richard, *Bioorg. Chem*. **2007**, *35*, 366–374.
98. C. D. Downey, J. L. Fiore, C. D. Stoddard, J. H. Hodak, D. J. Nesbitt, A. Pardi, *Biochemistry* **2006**, *45*, 3664–3673.
99. C. Mundoma, N. L. Greenbaum, *J. Am. Chem. Soc*. **2001**, *124*, 3525–3532.
100. C. Mundoma, N. L. Greenbaum, *Biopolymers* **2003**, *69*, 100–109.
101. R. A. Mathews, C. S. Rossiter, J. R. Morrow, J. P. Richard, *Dalton Trans.* **2007**, 3804–3811.
102. J.-M. Escudier, C. Dupouy, M. A. Fountain, I.-M. A. del Mundo, E. M. Jacklin, J. R. Morrow, *Org. Biomolec. Chem*. **2009**, *7*, 3251–3257.
103. C. Henoumont, L. Vander Elst, S. Laurent, R. Muller, *J. Biol. Inorg. Chem*. **2009**, *14*, 683–691.
104. I. Bertini, C. Luchinat, G. Parigi, *Concepts Magn. Reson*. **2002**, *14*, 259–286.
105. L. Banci, I. Bertini, K. L. Bren, M. A. Cremonini, H. B. Gray, C. Luchinat, P. Turano, *J. Biol. Inorg. Chem*. **1996**, *1*, 117–126.
106. G. Otting, *J. Biomolec. NMR* **2008**, *42*, 1–9.
107. S. L. Klakamp, W. D. Horrocks, Jr., *J. Inorg. Biochem*. **1992**, *46*, 193–205.
108. S. L. Klakamp, W. D. Horrocks, Jr., *J. Inorg. Biochem*. **1991**, *46*, 175–192.
109. R. A. Mathews, *Ph.D. Thesis*, University of Buffalo, State University of New York, 2008.
110. H.-K. Kim, J. Li, N. Nagraj, Y. Lu, *Eur. J. Chem*. **2008**, *14*, 8696–8703.
111. C.-H. Huang, A. Parish, F. Samain, F. Garo, R. Haner, J. R. Morrow, *Bioconjugate Chem*. **2010**, *21*, 476–482.
112. Y. Ye, H.-W. Lee, W. Yang, S. Shealy, J. J. Yang, *J. Am. Chem. Soc*. **2005**, *127*, 3743–3750.
113. L. Huang, L. L. Chappell, O. Iranzo, B. F. Baker, J. R. Morrow, *J. Biol. Inorg. Chem*. **2000**, *5*, 85–92.
114. B. F. Baker, S. S. Lot, J. Kringel, S. Cheng-Flournoy, P. Villiet, H. M. Sasmor, A.M. Siwkowski, L. L. Chappell, J. R. Morrow, *Nucl. Acid Res*. **1999**, *27*, 1547–1551.
115. D. Magda, R. A. Miller, J. L. Sessler, B. L. Iverson, *J. Am. Chem. Soc*. **1994**, *116*, 7439–7440.
116. D. Magda, M. Wright, S. Crofts, A. Lin, J. L. Sessler, *J. Am. Chem. Soc*. **1997**, *119*, 6947–6948.
117. J. Hall, D. Husken, U. Pieles, H. E. Moser, R. Haner, *Chem Biol.* **1994**, *1*, 185–190.
118. J. R. Morrow, T. L. Amyes, J. P. Richard, *Acc. Chem. Res*. **2008**, *41*, 539–548
119. K. Tanaka, M. Shionoya, *Coord. Chem. Rev*. **2007**, *251*, 2731–2742.
120. H. Guido, C. K. Clever, T. Carell, *Angew. Chem. Int. Ed*. **2007**, *46*, 6226–6236.

Chapter 7
Oxidative DNA Damage Mediated by Transition Metal Ions and Their Complexes

Geneviève Pratviel

Contents

Abstract DNA damage by redox-active metal complexes depends on the interaction of the metal complex with DNA together with the mechanism of oxygen activation. Weak interaction, tight binding, and direct involvement of DNA in the coordination sphere of the metal are described. Metal complexes acting through the production of diffusing radicals and metal complexes oxidizing DNA by metal-centered active species are compared. Metal complexes able to form high-valent metal-oxo species in close contact with DNA and perform DNA oxidation in a way reminiscent of enzymatic chemistry are the most elegant systems.

G. Pratviel (✉)
Laboratoire de Chimie de Coordination, CNRS, 205, Route de Narbonne,
F-31077 Toulouse-Cedex 04, France
e-mail: genevieve.pratviel@lcc-toulouse.fr

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*,
Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_7,

Keywords DNA-metal interaction • hydroxyl radical • iron bleomycin • manganese porphyrin • nickel complexes • platinum complexes

1 Introduction

Metal ion complexes are well known for their redox chemistry and their properties of oxygen activation [1–12]. Since DNA is sensitive to oxidation, DNA damage/cleavage by metal complexes has attracted much attention because (i) metal complexes may be useful tools for studies of molecular DNA lesions due to oxidative stress, which is implicated in many disorders including aging, cancer, and neurological diseases, (ii) they may also be used as probes of nucleic acids structure (DNA, RNA) or probes of DNA binding sites for proteins and drugs, (iii) versatile and specific artificial restriction nucleases based on metal complexes would supplement the tools for molecular biology and finally, (iv) metal complexes may still be viewed as potential cytotoxic anticancer drugs.

Unfortunately, literature often misses detailed molecular studies on the mechanisms of DNA damage. Therefore, this review will not be exhaustive but will focus on some examples illustrating fundamental mechanisms in the chemistry of DNA oxidation by metal complexes.

Three different mechanisms are considered. The *first* one is DNA degradation by metal complexes located at a certain distance from DNA. It includes the generation (by a redox-active metal complex) of reactive diffusible hydroxyl radicals by the so-called Fenton chemistry. This classical chemistry of DNA damage does not require a very special interaction between the metal complex and the biological target and is easy to achieve. Outersphere electron transfer can also be observed with metal complexes not interacting closely with DNA. The *second* one deals with the reaction of high valent, metal-centered reactive species, for which a strong and precise interaction of the metal complex and DNA occurs. Thus, the activated species comes in close contact with DNA. This chemistry is reminiscent of the highly selective oxidation chemistry mediated by heme or non-heme enzymes with a substrate entering a binding pocket containing an activated metal oxo (or its hydroperoxo precursor) species. For DNA damage, the metal complex has to adapt to and fit into the structure of DNA to bring the activated metal ion within bonding distance with the biological target. The *third* one is the chemistry of DNA oxidation through direct coordination of a metal complex to DNA. Here, oxidation of DNA is facilitated by the fact that heteroatoms from DNA belong to the inner coordination sphere of the metal ion.

2 Metal Complexes with Weak Interaction with DNA

2.1 Oxidative Chemistry at a Distance

We include in this section DNA oxidation mediated by diffusible species (principally hydroxyl radicals) and by electron transfer (outersphere electron transfer).

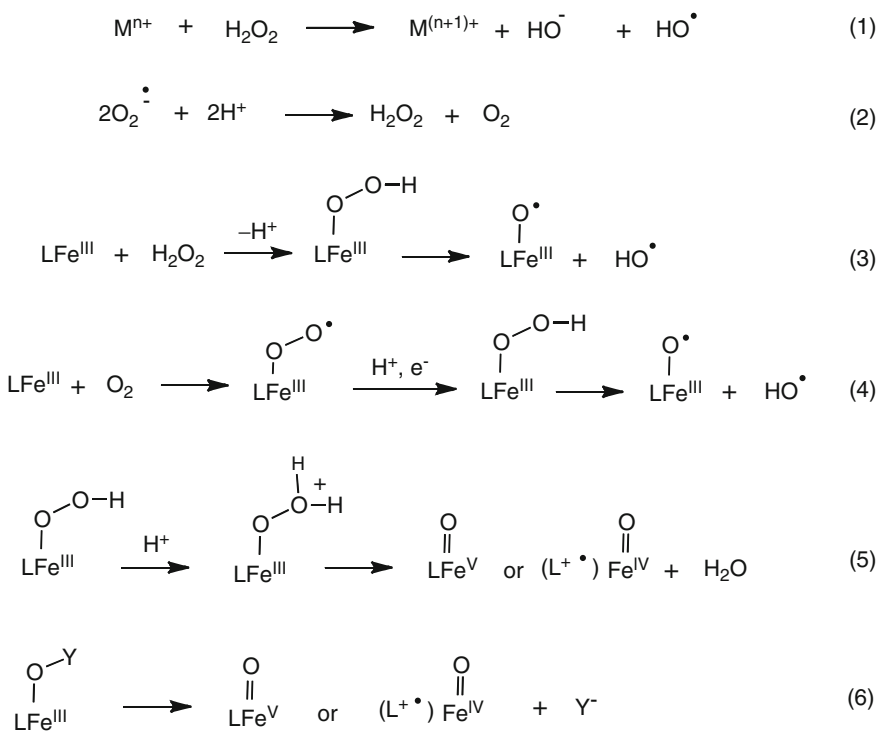


Figure 1 Mechanisms of oxygen activation by metal complexes. L stands for a ligand.Y stands for a good leaving group for the heterolytic cleavage of the O–O bond.

Redox-active metal complexes, even when unable to make contact with DNA, are able to mediate DNA damage through the production of hydroxyl radicals (HO•). This reaction is easy and does not require a special ligand to stabilize high valent states. One-electron reduction of H_2O_2 by Fe^{II} (Fenton reaction, Figure 1, reaction 1) is a classical way of production of HO•. Other metals are capable of serving in the Fenton reaction (copper, manganese, cobalt, etc.). The redox potential of H_2O_2 is 460 mV (NHE). The reaction requires a reductant in order to (re)generate the reduced form of the metal ion. Molecular oxygen may be a source of H_2O_2 through disproportionation of the superoxide radical anion (Figure 1, reaction 2). Hydroxyl radicals may also be generated by homolytic cleavage of the O–O bond of coordinated H_2O_2 or O_2 as illustrated with iron (Figure 1, reactions 3 and 4).

Hydroxyl radicals are small, neutral, and consequently diffusible. HO• are capable of reaction with any nucleobase as well as with 2-deoxyribose entities. HO• produced at a distance from the target, attack DNA at any accessible DNA region. HO• produced in the vicinity of DNA react with different sites yet in a more restricted area.

One way to distinguish complexes producing hydroxyl radicals (Figure 1, reactions 1, 3, and 4) from complexes acting by a metal-centered oxidant (Figure 1, reactions 5 or 6) is the use of radical quenchers such as ethanol, mannitol, DMSO. However, these radical traps are not (or less) active when the radicals are generated in

the vicinity of DNA. Therefore, in the field of DNA damage, it is often difficult to ascertain the mechanism of action of positively charged metal complexes that interact via electrostatic interactions with the phosphates or metal complexes bearing aromatic ligands intercalating between DNA base pairs. When the radical trap fails to inhibit the reaction, no conclusion can be drawn in the absence of further studies such as the characterization of active intermediates and careful analysis of DNA damage.

The second type of oxidative chemistry that can be mediated without a close specific interaction between the metal complex and DNA is electron abstraction from guanine (G) bases. Some positively charged, redox-active metal complexes may interact with DNA by electrostatic interactions or by intercalation and thus, within this relatively short distance, may be able to perform outersphere electron transfer reaction with DNA, provided the redox potential of the oxidized form of the metal complex is high enough. Electron abstraction from the 2-deoxyribose units does not occur. Electron abstraction takes place at the level of nucleic acid bases. Guanine has the lowest redox potential, 1.29 V (NHE) [13]. Electron abstraction occurs only at the G bases, while $HO^{\bullet}$ radicals potentially react with all nucleic bases (and sugars).

Simple metal complexes with no special device to target DNA can only oxidize DNA by these so-called remote chemistries and are in general poorly efficient. A quick evaluation test, commonly used in the literature, consists of cleavage of circular supercoiled plasmid DNA analyzed by agarose gel electrophoresis. This test is extremely sensitive and allows detection of nanomolar concentrations of cleavage events. Typically, micromolar concentrations of these metal complexes are necessary to transform the initial closed supercoiled circular DNA material (referred to as form I) into nicked circular DNA (form II). Thus, the yield of cleavage with respect to the cleaving agent is extremely low. Unfortunately, this method underestimates the real amount of damage events since it quantifies direct DNA breaks while alkali-labile lesions that are important in DNA oxidation are missed.

2.2 Fenton Chemistry

A hydroxyl radical is capable of two types of chemical reactions: (i) homolytic cleavage of a C–H bond (2-deoxyribose) and (ii) addition on a double bond (aromatic heterocycles of nucleobases). These two reactions generate a transient radical species on DNA that further transforms to a DNA lesion principally after reaction with O_2, which has a diradical character. Only some sugar radicals lead to direct strand breaks (C5', C3', and C4' radicals). Radicals generated on the other sugar positions (C1', C2', sometimes C4' positions) and base oxidations lead to alkali-labile lesions that are revealed after an alkaline treatment. For reviews on the various DNA oxidation products due to $HO^{\bullet}$ on the sugars and the nucleic bases see [14–18].

The fact that complexes devoid of any affinity for DNA are nevertheless able to mediate DNA damage by a $HO^{\bullet}$ mechanism is illustrated by the well known iron complex of EDTA. Fe-EDTA is negatively charged and thus cannot approach DNA. Thus, radicals must diffuse some distance before encountering the DNA. The damaged sites on DNA reflect their relative accessibility. The sugar C–H bonds are not equally accessible. The

C1'–H1' bond is the most deeply buried in the minor groove while C4'–H4' and C5'–H5' are more accessible from the bulk solvent. The two C5'–H5' bonds directly connected to the phosphate groups are located on the outside of the minor (and major groove) of the double helix [18]. C3'–H3' and C2'–H2' bonds lie in the major groove.

Hydroxyl radical generators are often used in footprinting experiments and folding structure determinations of nucleic acids [19,20]. Moreover, in the case of covalent linking of the metal complex on a DNA recognition entity (protein, nucleic acid, etc.) the pattern of scission has been used to provide information on the binding of the DNA recognition element [21–23]. Furthermore, the DNA cleavage pattern is different when the reagent attacks DNA from the major or the minor groove. A minor groove location leads to a 3'-shift in the cleavage events on a DNA sequence whereas a major groove location leads to a 5'-shift. Even though it proves useful, Fe-EDTA is not a good DNA cleaving agent. Excess of this reagent is necessary (mM concentration) to achieve DNA damage. This is probably due to the self-degradation of the ligand.

2.3 Electron Transfer

Electron transfer is restricted to damage at G residues. In general, guanine oxidation by electron transfer is not sequence-specific due to the phenomenon of long-range hole transfer (or long-range charge transport) that takes place within the π-stacked base pairs of the DNA double helix. An initial guanine radical cation (hole) is unstable and is able to abstract one electron from a neighboring G, thus creating a new $G^{+\bullet}$ (hole migration) [24]. Consequently, DNA damage may be observed far away from the site where it was initiated. The eventual trapping of the radical cation by water or by superoxide radical anion leads to irreversible DNA damage at guanines that are revealed by alkaline treatment. For a review on guanine oxidation products derived from guanine oxidation by electron transfer see [25]. Long-range charge transport was discovered with light activated rhodium complexes with aromatic ligands [26,27].

On the other hand, if the active species of the metal complex is not a one-electron oxidant but a two-electron oxidant, guanine oxidation by electron transfer leads to a guanine cation. This two-step process includes the deprotonation of the first formed $G^{+\bullet}$ to a neutral radical $[G(-H)]^{\bullet}$ followed by a second one-electron oxidation leading to a guanine cation $[G(-H)]^{+}$. The guanine cation entity is no more prone to long-range charge transport. It reacts with nucleophiles (such as H_2O, peroxides, amine groups, etc.) and transforms to guanine lesions. Thus, DNA damage by electron transfer through a two-electron mechanism will be observed at the site where the metal complex interacts with DNA. A high-valent manganese-oxo porphyrin, porphyrin-Mn^{V}=O or $(porphyrin)^{+\bullet}$-Mn^{IV}=O (Figure 1, Reaction 6), is able to mediate this type of guanine oxidation [28–30]. This chemistry was observed when this highly reactive reagent, endowed with a two redox equivalent capacity, approaches GC-rich DNA through non-specific and weak electrostatic interactions between the positively charged macrocyclic ligand and DNA. The exact nature of the interaction responsible for guanine oxidation is not known in the present case. We will see in the next section that the same porphyrin reagent, when in tight interaction with the minor groove of DNA, performs sugar oxidation at the C5'-carbon.

3 Metal Complexes with a Specific and Tight Non-covalent Interaction with DNA

3.1 Oxidative Chemistry Mediated in Close Contact with DNA

This section deals with selective 2-deoxyribose oxidation by metal-centered active species (Figure 1, Reactions 5 and 6). In this section the metal complexes are endowed with a high affinity for DNA. For a ligand to be capable of positioning and maintaining an activated metal species in a precise site within DNA, a tight binding to DNA is absolutely necessary. Additionally, this binding mode must be associated with a proper positioning of the metal-centered oxidant with respect to a given C–H bond of sugar. Moreover, the total absence of generation of diffusing radicals by the metal center is characterized by the selective oxidation of one (and only one) C–H bond of 2-deoxyribose at the site of interaction. The chemistry is exclusively directed toward sugars, nucleic bases are not touched. The chosen examples of this section have an activated metal species located in the minor groove. Three possible oxidation mechanisms may occur in the minor groove, namely, C1'-, C4'-, and C5'-carbon oxidation. Iron bleomycin directs its reactivity to the C4'–H4' bond of 2-deoxyribose at the 5'-pyrimidine nucleoside of a 5'-GPy-3' site (5'-GC-3' or 5'-GT-3' sites) while manganese tetrapyridiniumylporphyrin oxidizes only the C5'–H5' bond of the nucleoside located on the 3'-side after a three consecutive AT base pairs site. The consequence of this type of chemistry is the ability of these metal complexes to undergo plasmid cleavage at submicromolar concentrations [31,32].

3.2 Iron Bleomycin

The natural glycopeptide bleomycin (BLM) appears as the paradigm for DNA cleaving agents [33–36]. The mechanism of action of this antibiotic and antitumoral drug is due to DNA cleavage by the iron bleomycin complex, which takes place in the presence of O_2 and a reductant or in the presence of H_2O_2. The mechanism of DNA damage involves a metal-centered active species able to perform a two-electron oxidation reaction. A BLM–Fe^{III}–OOH species was characterized, yet it is probably a precursor of a reactive unstable ($BLM^{+\bullet}$)–Fe^{IV}=O (Figure 1, Reaction 5), which was not characterized [4]. The mechanism of DNA damage by iron bleomycin is due to selective abstraction of $H^{\bullet}$ from the C4'–H4' bond of 2-deoxyribose by the activated form of bleomycin and to subsequent formation of a carbon-centered radical at C4'. This radical may either react with O_2 or may be further oxidized to a cation at C4' that reacts with a molecule of water (Figure 2). Both pathways lead to DNA cleavage. This elegant selective reaction of iron bleomycin can now be explained by the structure of an inactive analogue of activated bleomycin, namely BLM–Co^{III}–OOH, in complex with DNA at the specific cleavage site [37–39]. The binding constant of BLM–Co^{III}–OOH with a 5'GPy-3' site is $K_a \sim 10^6$–10^7 M^{-1}. As shown in Figure 3, the metal binding

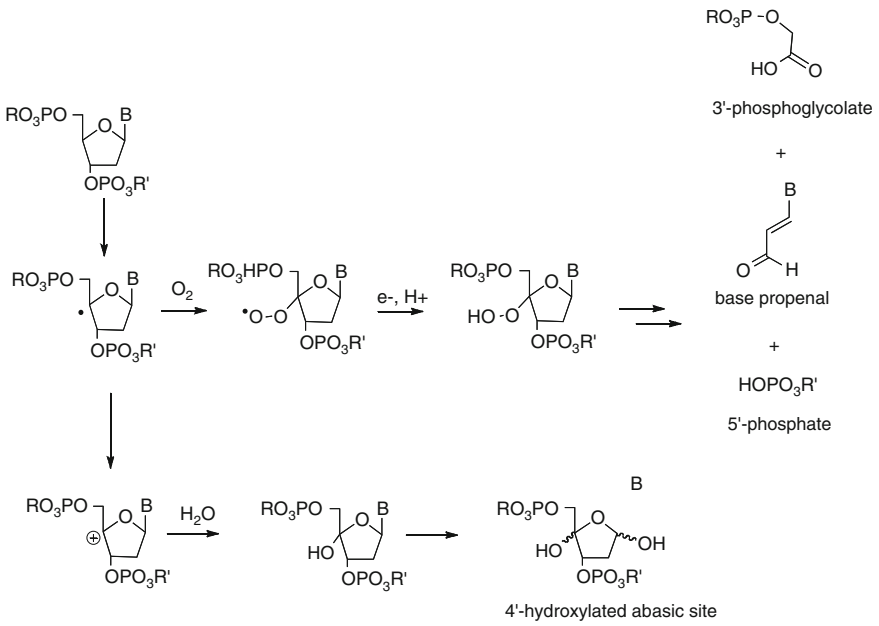


Figure 2 DNA damage by iron bleomycin.

domain of BLM is held in the minor groove of DNA by the establishment of two hydrogen bonds between the guanine base of a 5‘-GT-3’ site and the pyrimidine moiety of BLM. The bithiazole of BLM intercalates “above” the T residue of the 5’-GT-3’ site. The sugars of BLM lie in the minor groove and help the fitting of the metal binding domain toward the targeted sugar. This precise interaction of iron bleomycin with a 5’-GPy-3’ site and the capacity of BLM to undergo oxygen activation at the metal center are the keys of the efficiency of this reagent for single-strand cleavage of DNA depicted here. However, iron bleomycin is also able to perform double-strand breaks of DNA. A double-strand break may involve sequential oxidation of 2-deoxyriboses on opposite strands (probably by the same molecule of BLM) starting with the 3’-phosphoglycolate (Figure 2) pathway on the first strand, which leaves one redox equivalent on the metal center of BLM, followed by the attack of the second strand. Double-strand breaks of BLM are still not fully understood [36,40–42].

3.3 Manganese Porphyrin

As emphasized in the case of iron bleomycin, a strong and precise interaction of the manganese porphyrin, *meso*-tetrakis(4-*N*-methylpyridiniumyl)-porphyrinato) manganese(III) (Mn-TMPyP4), associated with the capacity of the ligand to stabilize high valent metal-oxo species, are also the keys of the efficiency of this

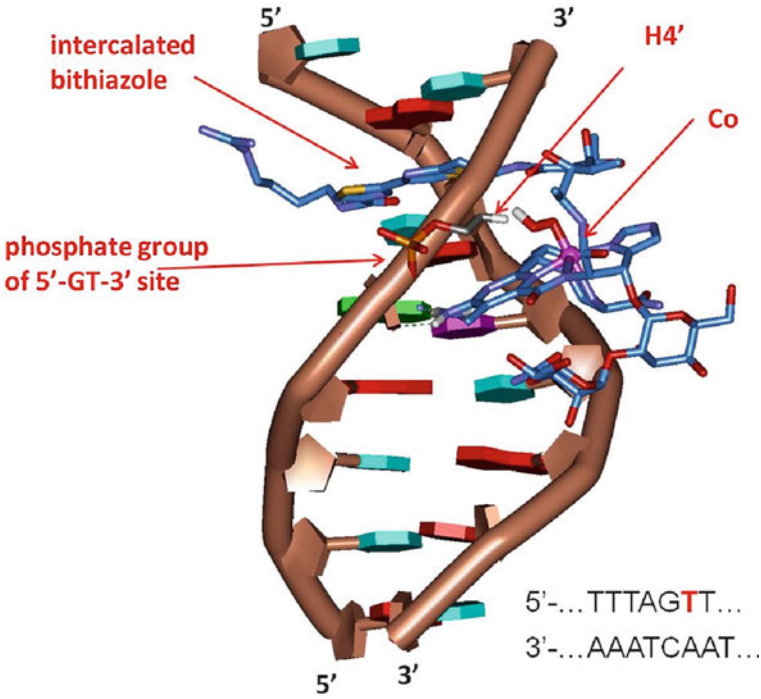


Figure 3 Structure of BLM–Co^{III}–OOH in interaction with a 5'-GT-3' site in double-stranded DNA (PDB file 2R2S) [39]. BLM is shown in CPK with a modeled peroxide. Hydrogen atoms are omitted for clarity. Simplified DNA is shown with color coded bases (red, adenine; cyan, thymine; green, guanine; purple, cytosine). The targeted C4'–H4' bond of the T residue of 5'GT-3', which is the site of attack of bleomycin, is indicated together with the neighboring phosphate group between G and T. Hydrogen bonds between guanine and the pyrimidine moiety of bleomycin are shown by dashed lines. The DNA sequence used in the crystal with a red oxidized T-residue is shown at the bottom of the figure.

reagent in DNA cleavage. The manganese complex is activated into a manganese-oxo species, Mn^{V}=O (equivalent to a (porphyrin$^{+\bullet}$)Mn^{IV}=O entity) with $KHSO_5$ [43]. In this case, the use of the dissymmetric peroxide $KHSO_5$ as an oxygen atom donor ensures a better efficiency in the formation of high-valent metal-oxo species due to the good leaving group KSO_4^- (see Reaction 6, Figure 1). The highest affinity binding site of the cationic Mn-TMPyP4 is in the minor groove of sequences consisting of at least three consecutive AT base pairs ($K_a \sim 10^6$–$10^7\ M^{-1}$). This interaction site is referred to as an $(AT)_3$-box. In AT-rich sequences a negative potential attracts many minor groove binding molecules [21]. As can be seen in the modeling structure [44] of Figure 4, it turns out that, within the minor groove of the $(AT)_3$-box, the metal-oxo entity of the porphyrin locates at suitable distance with respect to the C5'–H5' bonds of the 2-deoxyribose unit of the 3'-neighboring nucleoside of the $(AT)_3$-box. It was demonstrated that Mn-TMPyP/$KHSO_5$ cleaves DNA through typical cytochrome P450 chemistry [4,6,45]: H$^\bullet$ abstraction from the C5'–H5' bond leads to a carbon-centered radical at C5'. Oxygen rebound mechanism

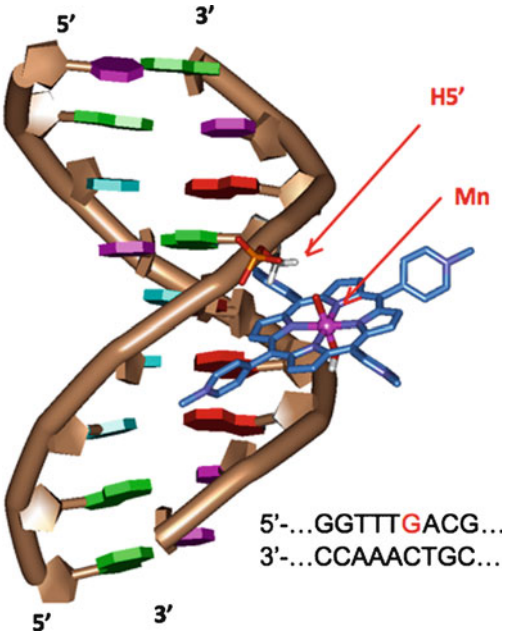


Figure 4 Modeling of the interaction of a high-valent manganese-oxo porphyrin, (TMPyP4–Mn^{V}=O) (with a *trans* hydroxo ligand) with the minor groove of the three consecutive AT base pair sites in double-stranded DNA [44]. Manganese porphyrin is shown in CPK. Hydrogen atoms are omitted for clarity as well as the counter ions. Simplified DNA is shown with color coded bases (red, adenine; cyan, thymine; green, guanine; purple, cytosine). The targeted C5'–H5' bond of the G residue on the 3'-side of the TTT, which is the site of attack of the manganese porphyrin, is indicated together with the neighboring 5'-phosphate of G. The DNA sequence used in the modeling study with a red oxidized G-residue is shown at the bottom of the figure.

gives rise to 5'-aldehyde and 3'-phosphate ending fragments and to direct strand cleavage (Figure 5). After two successive β-eliminations an oxidized product of sugar, furfural (FUR) is released as a marker of C5'-oxidation [46].

The active metal center of the manganese porphyrin is not so deeply engaged in the minor groove compared to iron bleomycin. This is why the former activates the C5'–H5' bond near the phosphate groups at the edge of the DNA backbone while the later reaches the deeper situated C4'–H4' bond in the minor groove.

In contrast to iron bleomycin for which only one mechanism of DNA damage was reported (C4' oxidation), the Mn-TMPyP4/$KHSO_5$ system showed several ways of damaging DNA depending on a different interaction mode of the metal complex with DNA. In DNA sequences devoid of AT-rich regions, Mn-TMPyP4 cannot find a way to interact in the minor groove of DNA. The overall binding constant of the metalloporphyrin for GC-rich DNA is one to two orders of magnitude lower compared to AT-rich DNA. In GC-rich sequences electron abstraction from the G bases [29,47] together with hydroxylation of the C1'–H1' bond by an oxygen transfer mechanism were observed [29,45]. Reaction of the high-valent metal-oxo entity with the most difficult to reach C–H target at the bottom of the minor groove is tentatively attributed

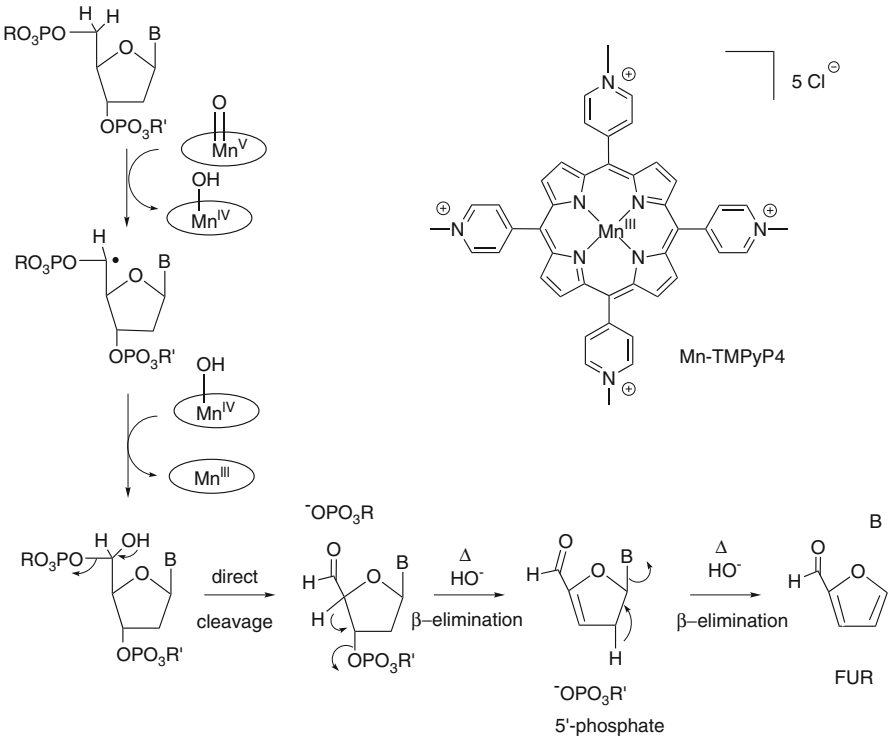


Figure 5 Structure of manganese-porphyrin (Mn-TMPyP4) and mechanism of DNA damage by hydroxylation of the C5'–H5' bond.

to partial denaturation of damaged DNA or to a partial intercalation of manganese porphyrin between DNA base pairs. Guanine oxidation by a two-electron transfer mechanism is tentatively attributed to an outersphere electron transfer from an outside binding porphyrin perhaps interacting with the phosphate groups. Thus, even if the Mn-TMPyP4/$KHSO_5$ system is clearly a reagent that does not produce any diffusing species, depending on the DNA sequence it interacts with, several mechanisms of DNA degradation are possible. This example illustrates the importance of the interaction of a metal complex with DNA for the mechanism of DNA damage.

4 DNA Adducts of Metal Complexes

4.1 Oxidative Chemistry at Accessible Guanine Residues

Metal ions may coordinate to nucleobases (especially to the most basic site, namely N7 in guanine). Simple coordination of a metal ion at G does not necessarily lead to DNA oxidation. However, some high valent metal complexes proved capable of selective oxidation of coordinated guanine bases. Two examples are selected for

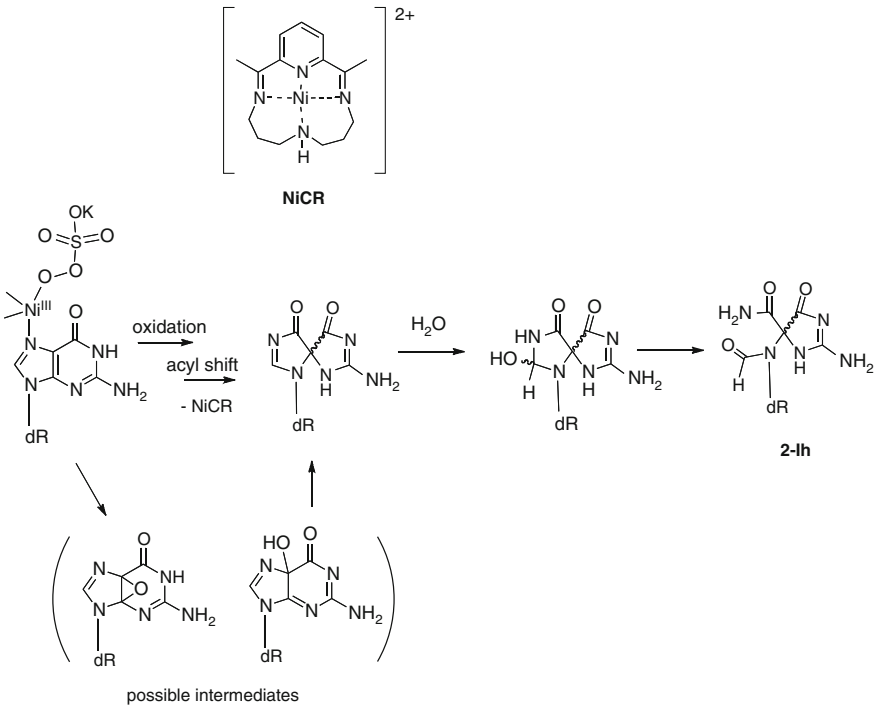


Figure 6 Structure of the NiCR dication and proposed mechanism of accessible guanine oxidation by the NiCR/$KHSO_5$ system; dR = deoxyribose.

which the mechanism was studied in some detail: a nickel(II) complex associated to $KHSO_5$ and a platinum(IV) complex. In double-stranded DNA, the N7 atom of G is located in the major groove but substitution of one coordination bond in a metal complex with guanine N7 is precluded due to steric hindrance. Thus, in this section, DNA damage occurs at Gs that are readily accessible (unpaired, bulged, looped or terminal guanine residues). DNA damage at G bases does not lead to direct strand breaks but to alkali-labile lesions that transform to strand breaks upon alkaline treatment.

4.2 Nickel Complexes

The nickel(II) tetraazamacrocyclic complex $[Ni(Me_2[14]py\text{-}diene\text{-}N_4)]^{2+}$ (NiCR) (Figure 6) induces oxidation of accessible guanine residues in the presence of potassium monopersulfate, $KHSO_5$ [48-51]. NiCR exists as a four-coordinate square planar complex and exhibits little tendency to acquire axial ligands in its initial Ni^{II} state. Selective association with accessible guanine is established only after axial ligation is induced by oxidizing the nickel to the Ni^{III} state, which in turn adopts octahedral, six-coordinate geometry [49,52,53]. $KHSO_5$ is a strong oxidant with a potential sufficiently high to cause oxidation of the Ni^{II} complex.

From systematic studies and correlations between the guanine oxidation reaction and the redox potential of various complexes the dependence of the DNA reaction on redox potential was consistent with the notion that Ni^{III} species with higher potential are the most efficient in guanine oxidation. Furthermore, ligand conformational flexibility was found important to adapt to DNA binding. Additionally, ligand flexibility may allow formation of a *cis*-coordinated complex bringing guanine and a bound HSO_5^- into adjacent positions around the metal ion. Indeed, the reactive nickel species responsible for the oxidation of DNA proved not simply a Ni^{III} complex. Simple electron transfer oxidants were ineffective at promoting the oxidation of DNA even though the formation of the Ni^{III} complex was detected by UV-visible spectroscopy [50]. Among all the tetraazamacrocyclic ligands tested, the CR ligand affords a good balance between high in-plane donor strength such that a stable, square planar complex with a relatively high Ni^{III}/Ni^{II} redox potential is obtained and a certain flexibility. Interestingly, the NiCR/$KHSO_5$ system leads to only one type of guanine lesion, 5-carbamido-5-formamido-2-iminohydantoin (2-Ih) (Figure 6) [54], which was characterized previously from an intermediate epoxide at the C4–C5 bond of G [55,56]. The molecular mechanism for the formation of this product remains to be established.

The reaction of the NiCR/$KHSO_5$ system with accessible Gs is relatively efficient. The reaction of the single-stranded 5'-T_7GT_7-3' oligonucleotide with stoichiometric amount of NiCR and excess amount of $KHSO_5$ (added in intervals) leads to 60% yield in 2-Ih-modified oligonucleotide within 30 min at pH = 7.4 and 37°C. Thus, NiCR appears as a probe for accessible guanine residues found in a variety of DNA and RNA structures in which there is a strong correlation between the extent of guanine modification and the exposure of N7.

Interestingly, direct coordination of the metal ion on guanine N7 was also postulated in the case of a di-copper complex which reacts with unpaired Gs resulting in the formation of the same guanine oxidation product (2-Ih) under aerobic conditions and in the presence of 3-mercaptopropionic acid as reductant [57].

4.3 Platinum Complexes

Most Pt^{IV} anticancer drugs are considered to be Pt^{II} prodrugs. Because Pt^{IV} complexes are kinetically inert, reaction with DNA is not expected until they have been reduced to Pt^{II} in the cellular medium. However, direct Pt^{IV} binding to DNA has been observed *in vitro* and [Pt^{IV}(dach)Cl_4], which has a high redox potential, oxidizes guanine without UV irradiation or additional oxidant. *In vitro* substitution reaction of [Pt^{IV}(dach)Cl_4], with an oxidizable ligand, such as a guanine, leads to an innersphere two-electron transfer from guanine to the platinum ion [58–61]. Considering the redox potential of G (1.29 V versus NHE) [13], an outersphere electron transfer mechanism is not possible between G and the Pt^{IV} complex. Instead, an innersphere mechanism induced by the coordination of Pt^{IV} on G was considered. In this example, only guanine residues at 5'-end of a DNA strand were shown to be reactive [59].

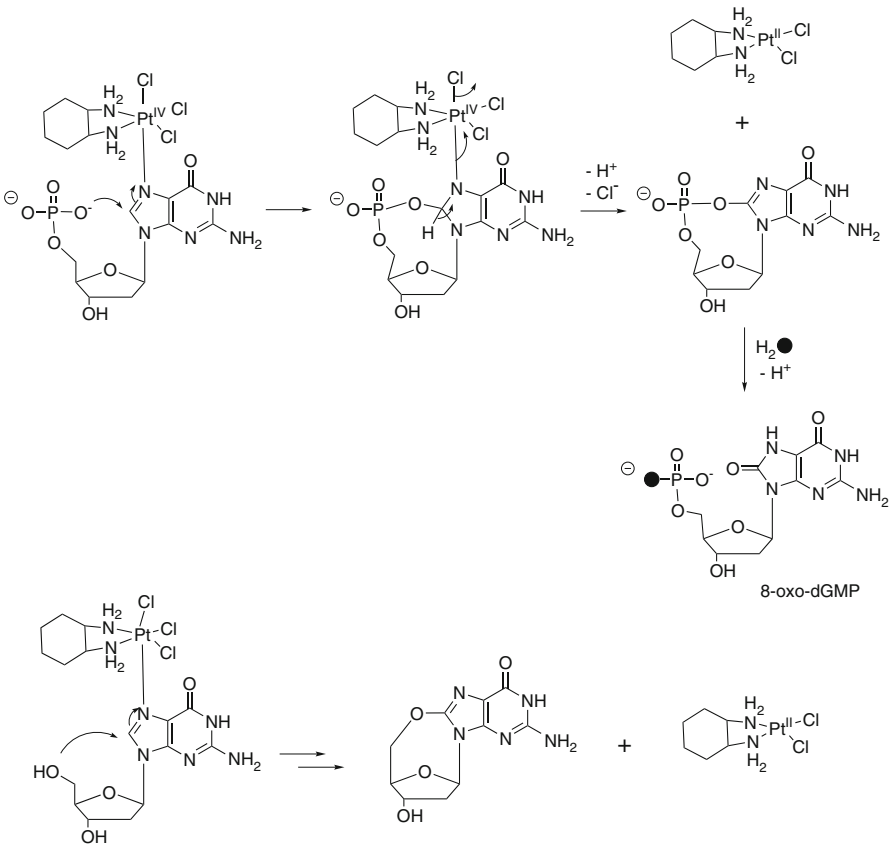


Figure 7 Mechanism of terminal guanine oxidation by $Pt^{IV}(dach)Cl_4$ through the formation of a Pt^{IV} adduct at N7 of guanine.

Binding of Pt^{IV} to N7 of G followed by nucleophilic attack of a 5'-phosphate or a 5'-hydroxyl oxygen to C8 of the G moiety and an innersphere, two-electron transfer produce cyclic (5'-O–C8)-G and a Pt^{II} complex (Figure 7). The identity of the final oxidized G residue depends on the hydrolysis rate of the cyclic intermediate. The cyclic phosphodiester slowly hydrolyzes to the 5'-phosphate group and an oxidized G-residue (8-oxo-G) whereas the cyclic ether does not hydrolyze, and this cyclic form is the final oxidation product. Simple binding of a Pt^{IV} complex to N7 of intra-strand G moieties was also observed experimentally but without further oxidation reactions at G. Substitution reactions of Pt^{IV} complexes with G are catalyzed by trace amounts of Pt^{II} [60–62]. Therefore, since the reaction induces the release of a free Pt^{II} complex, the overall reaction is autocatalytic after an induction time for the reaction to start.

What is the yield of this reaction? After 10h, a reaction mixture of initial 10 mM $[Pt^{IV}(dach)Cl_4]$ and 20 mM 5'-dGMP, afforded 5 mM of 8-oxo-dGMP at pH = 8.6

and 37°C [58,59]. Kinetic studies confirmed that substitution is accelerated by Pt^{II} [61]. Addition of Pt^{II} (~20% with respect to initial to Pt^{IV}) reduced the reaction time by half.

Although such a reaction may not be biologically relevant due to competition with the intracellular reduction of the Pt^{IV} complexes, it illustrates the potential oxidative behavior of Pt^{IV} complexes toward DNA.

These two last mentioned examples of high-valent metal complexes capable of direct coordination to N7 of G, illustrate oxidation of a G base through an inner-sphere electron transfer process or through the formation of a *cis*-coordinatd complex with a strong oxidation power [63]. However, further studies are needed to better understand the mechanisms underlying guanine oxidation mediated by DNA adducts of metal complexes. Particularly, labeling studies will clarify the origin of the oxygen atoms incorporated in oxidized guanine residues and tell whether oxidation proceeds by an electron transfer or an oxygen atom transfer process.

5 Concluding Remarks and Future Directions

The chemistry of DNA damage depends on the location of the metal ion with respect to DNA, on the mechanism of oxygen activation by the metal complex, and on the stability of the association between the metal complex and DNA. For efficient DNA damage the reactive metal ion center must be positioned in close and appropriate location with respect to DNA to avoid self-degradation of the metal complex. This requisite is achieved only in a few cases.

Up to now, metal complexes able to damage DNA by oxidation have proved remarkable tools for the study of molecular DNA damage under oxidative stress and for studies of nucleic acid structures. However, the field is still open for the design of versatile sequence-specific DNA cleavers and of double-strand cleaving agents that may have a future as artificial nucleases and cytotoxic drugs.

Abbreviations

A	adenine
BLM	bleomycin
C	cytosine
CR	2,12-dimethyl-3,7,11,17-tetraazabicyclo[11.3.1]heptadeca-1(17),2,11,13,15-pentaene
dach	diaminocyclohexane
5'-dGMP	2-deoxyguanosine 5'-monophosphate
DMSO	dimethylsulfoxide
EDTA	ethylenediamine-*N*,*N*,*N'*,*N'*-tetracetate
FUR	furfural

G	guanine
5'-GPy-3'	5'-guanine pyrimidine-3' sequence
2-IH	5-carbimido-5-formamido-2-iminohydantoin
K_a	affinity constant
NHE	normal hydrogen electrode
8-oxo-dGMP	8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-monophosphate
8-oxo-G	8-oxo-7,8-dihydro-guanine
T	thymine
TMPyP4	*meso*-tetrakis(4-*N*-methylpyridiniumyl)-porphyrin dianion

Acknowledgments The contribution of coauthors whose names appear in the references is acknowledged. This article is dedicated to Dr Marguerite Pitié.

References

1. J. T. Groves, *J. Inorg. Biochem*. **2006**, *100*, 434–447.
2. L. Que, Jr., *Acc. Chem. Res*. **2007**, *40*, 493–500.
3. L. Que, Jr., W. B. Tolman, *Nature* **2008**, *455*, 333–340.
4. M. Pitié, G. Pratviel, *Chem. Rev*. **2010**, *110*, 1018–1059.
5. A. Gunay, K. H. Theopold, *Chem. Rev*. **2010**, *110*, 1060–1081.
6. B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev*. **2004**,*104*, 3947–3980.
7. S. Shaik, H. Hirao, D. Kumar, *Acc. Chem. Res*. **2007**, *40*, 532–542.
8. T. Punniyamurthy, S. Velusamy, J. Iqbal, *Chem. Rev*. **2005**, *105*, 2329–2363.
9. S. V. Kryatov, E. V. Rybak-Akimova, S. Schindler, *Chem. Rev*. **2005**, *105*, 2175–2226.
10. M. M. Abu-Omar, A. Loaiza, N. Hontzeas, *Chem. Rev*. **2005**, *105*, 2227–2252.
11. E. I. Solomon, P. Chen, M. Metz, S. K. Lee, A. E. Palmer, *Angew. Chem. Int. Ed*. **2001**, *40*, 4570–4590.
12. R. A. Himes, K. D. Karlin, *Curr. Opin. Chem. Biol*. **2009**, *13*, 119–131.
13. S. Steenken, S. V. Jovanovic, *J. Am. Chem. Soc*. **1997**, *119*, 617–618.
14. C. von Sonntag, *Free-Radical-Induced DNA Damage and Its Repair, A Chemical Perspective*, Springer-Verlag, Berlin, Heidelberg, New York, 2006.
15. J. R. Wagner, J. Cadet, *Acc. Chem. Res*. **2010**, *43*, 564–571.
16. J. Cadet, T. Douki, J. L. Ravanat, *Free Radical Biol. Med*. **2010**, *49*, 9–21.
17. C. J. Burrows, J. G. Muller, *Chem. Rev*. **1998**, *98*, 1109–1152.
18. W. K. Pogozelski, T. D. Tullius, *Chem. Rev*. **1998**, *98*, 1089–1108.
19. T. D. Tullius, J. A. Greenbaum, *Curr. Opin. Chem. Biol*. **2005**, *9*, 127–134.
20. S. S. Jain, T. D. Tullius, *Nat. Protoc*. **2008**, *3*,1092–1100.
21. P. B. Dervan, *Science* **1986**, *232*, 464–471.
22. D. P. Mack, B. L. Iverson, P. B. Dervan, *J. Am. Chem. Soc*. **1988**, *110*, 7572–7574.
23. R. Baliga, J. W. Singleton, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10393–10397.
24. J. C. Genereux, J. K. Barton, *Chem. Rev*. **2010**, *110*, 1642–1662.
25. G. Pratviel, B. Meunier, *Chem. Eur. J*. **2006**, *12*, 6018–6030.
26. C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. H. Bossmann, N. J. Turro, J. K. Barton, *Science* **1993**, *262*, 1025–1029.
27. D. B. Hall, R. E. Holmlin, J. K. Barton, *Nature* **1996**, *382*, 731–735.
28. C. Vialas, G. Pratviel, C. Claparols, B. Meunier, *J. Am. Chem. Soc*. **1998**, *120*, 11548–11553.
29. M. Pitié, C. Boldron, G. Pratviel, *Adv. Inorg. Chem*. **2006**, *58*, 77–130.
30. B. Mestre, A. Jakobs, G. Pratviel, B. Meunier, *Biochemistry* **1996**, *35*, 9140–9149.

31. R. E. Kilkuskie, H. Suguna, B. Yellin, N. Murugesan, S.M. Hecht, *J. Am. Chem. Soc*. **1985**, *107*, 260–261.
32. J. Bernadou, G. Pratviel, F. Bennis, M. Girardet, B. Meunier, *Biochemistry* **1989**, *28*, 7268–7275.
33. D. H. Petering, Q. Mao, W. Li, E. DeRose, W. E. Antholine, *Met. Ions Biol. Syst*. **1996**, *33*, 619–648.
34. R. M. Burger, *Chem. Rev*. **1998**, *98*, 1153–1170.
35. C. A. Claussen, E. C. Long, *Chem. Rev*. **1999**, *99*, 2797–2816.
36. J. Chen, J. Stubbe, *Curr. Opin. Chem. Biol*. **2004**, *8*, 175–181.
37. S. T. Hoehn, H. D. Junker, R. C. Bunt, C. J. Turner, J. Stubbe, *Biochemistry* **2001**, *40*, 5894–5905.
38. C. Zhao, C. Xia, Q. Mao, H. Forsterling, E. DeRose, W.E. Antholine, W.K. Subczynski, D. H. Petering, *J. Inorg. Biochem*. **2002**, *9*1, 259–268.
39. K. D. Goodwin, M. A. Lewis, E. C. Long, M. M. Georgiadis, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5052–5056.
40. J. Chen, J. Stubbe, *Nat. Rev. Cancer* **2005**, *5*, 102–112.
41. A. T. Abraham, X. Zhou, S.M. Hecht, *J. Am. Chem. Soc*. **2001**, *123*, 5167–5175.
42. J. Chen, M. K. Ghorai, G. Kenney, J. Stubbe, *Nucleic Acids Res*. **2008**, *36*, 3781–3790.
43. J. T. Groves, J. B. Lee, S. S. Marla, *J. Am. Chem. Soc*. **1997**, *119*, 6269–6273.
44. P. Arnaud, K. Zakrzewska, B. Meunier, *J. Comput. Chem*. **2003**, *24*, 797–805.
45. M. Pitié, J. Bernadou, B. Meunier, *J. Am. Chem. Soc*. **1995**, *117*, 2935–2336.
46. G. Pratviel, M. Pitié, J. Bernadou, B. Meunier, *Angew. Chem. Int. Ed*. **1991**, *30*, 702–704.
47. C. Vialas, G. Pratviel, B. Meunier, *Biochemistry* **2000**, *39*, 9514–9522.
48. X. Chen, S. E. Rokita, C. J. Burrows, *J. Am. Chem. Soc*. **1991**, *113*, 5884–5886.
49. J. G. Muller, X. Chen, A. C. Dadiz, S. E. Rokita, C. J. Burrows, *J. Am. Chem. Soc*. **1992**, *114*, 6407–6411.
50. J. G. Muller, X. Chen, A. C. Dadiz, S. E. Rokita, C. J. Burrows, *Pure Appl. Chem*. **1993**, *65*, 545–550.
51. C. J. Burrows, S. E. Rokita, *Acc. Chem. Res*. **1994**, *27*, 295–301.
52. H.-C. Shih, N. Tang, C. J. Burrows, S. E. Rokita, *J. Am. Chem. Soc*. **1998**, *120*, 3284–3288.
53. H.-C. Shih, H. Kassahun, C. J. Burrows, S. E. Rokita, *Biochemistry* **1999**, *38*, 15034–15042.
54. P. Ghude, M. A. Schallenberger, A. M. Fleming, J. G. Muller, C. J. Burrows, *Inorg. Chim. Acta* **2011**, doi:10.1016/j.ica.2010.12.063.
55. W. Ye, R. Sangaiah, D. E. Degen, A. Gold, K. Jayaraj, K. M. Koshlap, G. Boysen, J. Williams, K. B. Tomer, L. M. Ball, *Chem. Res. Toxicol*. **2006**,*19*, 506–510.
56. W. Ye, R. Sangaiah, D. E. Degen, A. Gold, K. Jayaraj, K. M. Koshlap, G. Boysen, J. Williams, K. B. Tomer, V. Mocanu, N. Dicheva, C. E. Parker, R. M. Schaaper, L. M. Ball, *J. Am. Chem. Soc*. **2009**, *131*, 6114–6123.
57. L. Li, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc*. **2005**, *127*, 520–521.
58. S. Choi, R. B. Cooley, A. S. Hakemian, Y. C. Larrabee, R. C. Bunt, S. D. Maupas, J. G. Muller, C. J. Burrows, *J. Am. Chem. Soc*. **2004**, *126*, 591–598.
59. S. Choi, R. B. Cooley, A. Voutchkova, C. H. Leung, L. Vastag, D. E. Knowles, *J. Am. Chem. Soc*. **2005**, *127*, 1773–1781.
60. S. Choi, L. Vastag, C. H. Leung, A. M. Beard, D. E. Knowles, J. A. Larrabee, *Inorg. Chem*. **2006**, *45*, 10108–10114.
61. S. Choi, L. Vastag, Y. C. Larrabee, M. L. Personick, K. B. Schaberg, B. J. Fowler, R. K. Sandwick, G. Rawji, *Inorg. Chem*. **2008**, *47*, 1352–1360.
62. R. M. Roat, M. J. Jerardi, C. B. Kopay, V. Heath, J. A. Clark, J. A. DeMars, J. M. Weaver, E. Bezemer, J. Reedijk, *J. Chem. Soc., Dalton Trans.* **1997**, 3115–3621.
63. Note added in proof: Coordination of copper at the N7 of guanine was also proposed to be at the origin of guanine oxidation, see A. M. Fleming, J. G. Muller, I. Ji, C. J. Burrows. *Org. Biomol. Chem*. **2011**, *9*, 3338–3348.

Chapter 8
Metal Ion-Dependent DNAzymes and Their Applications as Biosensors

Tian Lan and Yi Lu

Contents

T. Lan
Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Y. Lu (✉)
Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, and

Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
e-mail: yi-lu@illinois.edu

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_8,

Abstract Long considered to serve solely as the genetic information carrier, DNA has been shown in 1994 to be able to act as DNA catalysts capable of catalyzing a trans-esterification reaction similar to the action of ribozymes and protein enzymes. Although not yet found in nature, numerous DNAzymes have been isolated through *in vitro* selection for catalyzing many different types of reactions in the presence of different metal ions and thus become a new class of metalloenzymes. What remains unclear is how DNA can carry out catalysis with simpler building blocks and fewer functional groups than ribozymes and protein enzymes and how DNA can bind metal ions specifically to perform these functions. In the past two decades, many biochemical and biophysical studies have been carried out on DNAzymes, especially RNA-cleaving DNAzymes. Important insights have been gained regarding their metal-dependent activity, global folding, metal binding sites, and catalytic mechanisms for these DNAzymes. Because of their high metal ion selectivity, one of the most important practical applications for DNAzymes is metal ion detection, resulting in highly sensitive and selective fluorescent, colorimetric, and electrochemical sensors for a wide range of metal ions such as Pb^{2+}, UO_2^{2+}, including paramagnetic metal ions such as Cu^{2+}. This chapter summarizes recent progresses in *in vitro* selection of metal ion-selective DNAzymes, their biochemical and biophysical studies and sensing applications.

Keywords biosensor • catalysis • DNA • DNAzyme • metalloenzyme • metal sensing

1 Introduction

For a long time, DNA's only function was perceived as being the genetic material for all organisms. The genetic information is encoded by certain combinations of the four nucleobases: adenine (A), thymine (T), guanine (G), and cytosine (C). Because of the limitation of functional group diversity in the building blocks of DNA in comparison with other biomolecules such as proteins, and because of the mostly invariant double helical structure, DNA was considered to be incapable of catalyzing chemical reactions.

The discovery of ribozymes [1,2], which are RNAs that can catalyze chemical or biological reactions, in the early 1980s was the first hint that DNA might also be

Table 1 Different types of reactions catalyzed by DNAzymes.

Reaction type	Metal cofactor	Refs
RNA cleavage	Pb^{2+}	[3–10]
	Mg^{2+}	
	Zn^{2+}	
	none	
	Ca^{2+}	
	Mn^{2+} or Mg^{2+}	
	Mn^{2+} or Mg^{2+}	
	UO_2^{2+}	
DNA cleavage (oxidative)	Cu^{2+}	[11–13]
DNA cleavage (hydrolytic)	Zn^{2+} or Mn^{2+}	[14–16]
Phosphoramidate bond cleavage	Mg^{2+}	[17]
DNA ligation	Cu^{2+}, Zn^{2+} or Mn^{2+}	[18,19]
RNA ligation (3′-5′)	Zn^{2+} or Mg^{2+}	[20,21]
RNA ligation (branch formation)	Mn^{2+} or Mg^{2+}	[22–25]
RNA ligation (lariat formation)	Mn^{2+}	[26,27]
Nucleopeptide linkage formation	Mn^{2+} or Mg^{2+}	[28]
DNA phosphorylation	Mn^{2+}	[29]
DNA capping	Mg^{2+}, Cu^{2+}	[30]
DNA depurination (IO_4^--dependent)	none	[31]
DNA depurination	Ca^{2+}	[32]
Diels-Alder reaction	Ca^{2+}	[33]
Porphyrin metallation	Cu^{2+} or Zn^{2+}	[34,35]

capable of catalysis, because of the structural similarity between DNA and RNA. Such reasoning was not unfounded: in 1994, Ronald Breaker and Gerald Joyce isolated the first deoxyribozyme (DNAzyme) via *in vitro* selection [3]. This DNAzyme can catalyze a phosphodiester bond cleavage in the presence of Pb^{2+} and had a rate enhancement of 10^5 over the uncatalyzed reaction. After this initial discovery, many more DNAzymes have been successfully isolated which are able to catalyze various types of reactions. Table 1 lists representative DNAzymes that have been obtained, including DNAzymes that catalyze cleavage of RNA [3–10], DNA [11–16] or the phosphoramidate bond [17], the ligation of DNA [18,19] or RNA [20,21], the formation of an RNA branch [22–25], an RNA lariat [26,27], or a nucleopeptide bond [28], phosphorylation [29], adenylation [30], and depurination [31,32] of DNA, the Diels-Alder reaction [33], and porphyrin metallation [34,35]. One of the RNA-cleaving DNAzymes is capable of catalyzing a phosphodiester transfer reaction with a $k_{cat}/K_m \approx 10^9\ M^{-1}\ min^{-1}$ that rivals that of protein enzymes [4].

Interestingly, most DNAzymes carry out their catalysis with the aid of metal cofactors and some are very selective in this respect. For example, the first reported DNAzyme is about 400,000-fold selective for Pb^{2+} [3,36] over other competing metal ions and the UO_2^{2+}-dependent DNAzyme (39E) is >1 million-fold selective over other competing metal ions [10,37]. This kind of metal selectivity can be found in many other DNAzymes. Because of their generally high metal ion selectivity, DNAzymes have been converted into fluorometric [10,36,38–46], colorimetric [47–53], and electrochemical [54,55] sensors for metal ions. In addition, they have also been used to

detect and degrade specific RNA sequences because of the intrinsic sequence requirement for the DNAzyme activity [56–61].

Even though DNAzymes have proved their usefulness in many applications, the detailed structure and reaction mechanisms of the DNAzyme catalysis are still largely unclear. A large number of biochemical and biophysical studies have already been carried out for several DNAzymes, but the absence of any high-resolution, three-dimensional structure for the active conformations of DNAzymes is currently a major obstacle for understanding them [62,63]. In the meantime, biochemical and biophysical studies have offered insights into the conserved sequences and structural features responsible for DNAzymes' selective binding of metal ions and efficient catalysis.

In this chapter, we will first introduce *in vitro* selection (a method for isolating DNAzymes), then present the results of detailed biochemical and biophysical studies on RNA-cleaving DNAzymes. These studies have shed some light on how metal ions interact with DNAzymes. Finally, we will discuss the biosensing applications of DNAzymes.

2 *In vitro* Selection of DNAzymes

2.1 In vitro Selection

Although not yet found in nature, DNAzymes have been isolated via the combinatorial method of *in vitro* selection. This method was first developed in the 1990s to isolate ribozymes and RNA aptamers, single-stranded oligonucleotides with high affinities for different target molecules [64–66]. In 1994, Ronald Breaker and Gerald Joyce used this strategy to isolate the first DNA catalyst [3]. The first DNAzyme used Pb^{2+} as a cofactor to catalyze the cleavage of a phosphodiester bond.

In vitro selection starts with a "random pool" containing typically $10^{14–15}$ different DNA sequences. The pool is designed by flanking a random sequence with two constant regions that are primer binding sites (Figure 1). The length of the sequence generally varies from 40 to 80 nucleotides. After the "random pool" is then synthesized, it is incubated with the substrate to carry out the desired reaction in the presence of a cofactor. The active DNA sequences are partitioned and collected. The active sequences can be separated from the inactive ones by various methods, such as polyacrylamide gel electrophoresis (PAGE) or affinity resin. Polyacrylamide gel electrophoresis separates the reaction products according to their size-dependent mobility on a polymer matrix. Affinity separation normally employs biotinylated DNA and a streptavidin resin. After the separation step, polymerase chain reaction (PCR) is used to amplify the desired DNA sequences. Since *in vitro* selection normally only utilizes a single-stranded DNA (ssDNA), such ssDNA is then generated from the PCR products and they are incubated under the desired condition with the substrate to start the next round of selection. After iterative rounds of selection, separation, and amplification, the large initial pool of random sequences can be reduced into a small population of sequences that are enriched with active DNAzymes. Typically, it takes 5–15 rounds

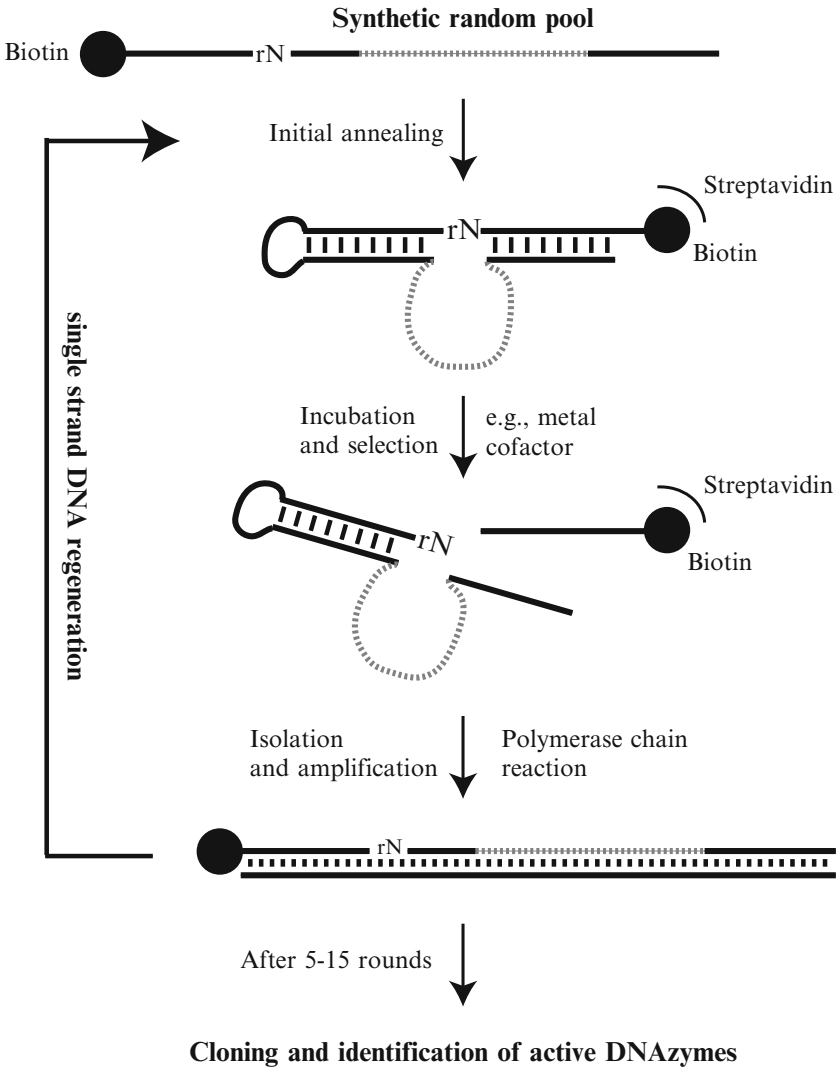


Figure 1 The schematic view of an *in vitro* selection process. rN denotes a ribonucleotide (as the cleavage site). Grey dotted line denotes the random region.

of selection to obtain an active DNAzyme. At this point, the population is cloned and sequenced. The resulting sequences are grouped into different families based on sequence similarity and each family is tested for activity. The DNAzyme selection is an application of the idea of "survival of the fittest", allowing researchers to find new DNAzymes with desirable activities in the presence of intended targets.

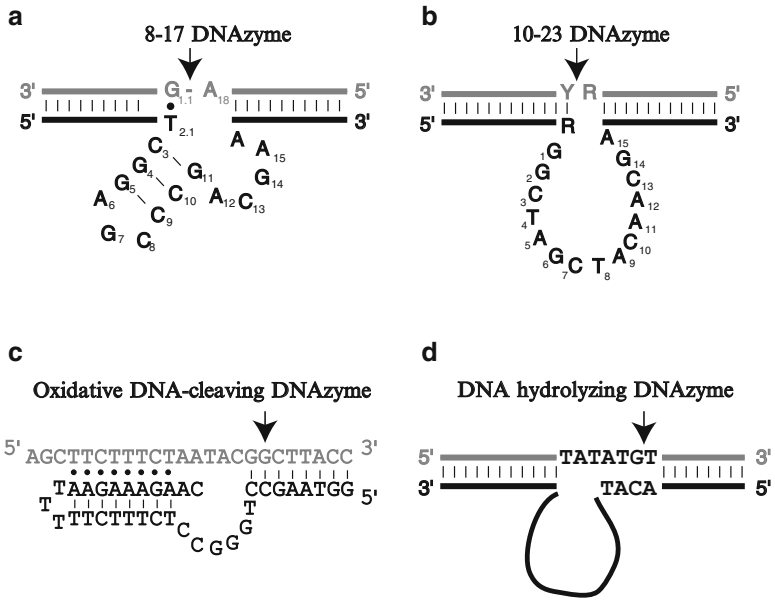


Figure 2 The predicted secondary structures for various RNA- and DNA-cleaving DNAzymes. Enzyme strands are in black and substrates are in grey. Black arrows indicate the cleavage site. Black lines between the enzyme and substrate strand indicate Watson-Crick base pairing. (**a**) The 8-17 DNAzyme; black dot indicates a wobble base pair. (**b**) The 10-23 DNAzyme, Y = pyrimidine, R = purine. (**c**) Ascorbate-independent DNA-cleaving DNAzyme, black dots indicate formation of triple helix; oxidative cleavage occurs on the sugar ring. (**d**) DNA-hydrolyzing DNAzyme.

2.2 Examples of Selected DNAzymes

Many DNAzymes have been isolated by *in vitro* selection. While most of the DNAzymes selected to date involve either cleaving or ligating the phosphodiester bond, DNAzymes with other functionalities have also been isolated, such as catalyzing porphyrin metallation and the Diels-Alder reaction.

2.2.1 RNA-Cleaving DNAzymes

The RNA-cleaving DNAzyme is the most frequently isolated and studied DNAzyme. They catalyze the phosphodiester bond cleavage of their RNA substrate. The 10-23 and 8-17 DNAzymes, first reported by Santoro and Joyce in 1997 [4], are the best-studied and most widely used RNA-cleaving DNAzymes (Figure 2a and 2b). They are named by the round and clone number in their respective selections. The computer algorithm mfold [67] has predicted that both DNAzymes have small catalytic cores of only 13 to 15 nucleotides. The 10-23 catalytic core is composed of an unstructured loop while the 8-17 catalytic core is composed of a three-base-pair stem, an

AGC trinucleotide loop, and an ACGA(A) loop. The cleavage site of 10-23 is a phosphodiester bond at the 3′-end of a single nucleotide flanked by two completely complementary substrate binding arms. The 8-17 cleavage site is similar to that of the 10-23 DNAzyme, but has an unpaired nucleotide with a G•T wobble pair immediately downstream from it. The initial selection of the 8-17 DNAzyme is only reported to cleave a 5′-AG-3′ junction, while the 10-23 DNAzyme is capable of cleaving any purine-pyrimidine (5′-RY-3′) junction [4]. Later biochemical characterizations indicated that both DNAzymes carry out the catalysis in a similar manner: metal assisted deprotonation of the 2′-hydroxyl, and the resulting oxyanion attacks the phosphate group nearby leading to cleavage [5,56,68–71]. This strategy is also used by naturally occurring ribozymes and endonucleases [72,73].

One surprising aspect of the 8-17 DNAzyme is that this DNAzyme has been isolated numerous times by different groups under very different conditions [5,7,9,70,74,75]. Several later selections were also carried out and isolated many more 8-17 variants. All these studies suggest that the 8-17 DNAzyme is a small, efficient sequence that occurs very frequently during *in vitro* selections. The large number of 8-17 variants aids later biochemical studies substantially [76].

The 10-23 DNAzyme has been widely used as a therapeutic agent for suppressing RNA levels in various systems [56,57,60] and for degrading viral RNA [58,59,61]. The 8-17 DNAzyme has numerous applications in nucleic acid detection [77,78], metal ion sensing [79–81] and DNA computing [82–87].

2.2.2 DNA-Cleaving DNAzymes

Besides catalyzing RNA cleavage, DNAzymes can also catalyze DNA cleavage via different types of mechanisms [11–16]. Because DNA is intrinsically more stable than RNA, DNAzymes that are capable of catalyzing DNA cleavage demand a higher degree of complexity than their RNA-cleaving siblings. The first DNA-cleaving DNAzymes were isolated in 1996 by Carmi, Schultz, and Breaker [11]. Both of the DNAzyme classes identified in this experiment catalyze DNA cleavage via an oxidative mechanism; what distinguishes them is that they have different requirements for cofactors: the Class I catalyst requires both Cu^{2+} and ascorbate while the Class II catalyst requires only Cu^{2+}. However, these DNAzymes are not site-selective and their oxidative cleavage mechanism limits their application to *ex vivo* purposes [11–13]. Other DNA-cleaving DNAzymes were isolated in 2009 by Chandra, Sachdeva, and Silverman [14]. Unlike the Cu^{2+}/ascorbate-dependent DNAzymes, these new DNAzymes are capable of cleaving DNA by a hydrolytic mechanism using Zn^{2+} and Mn^{2+} as cofactors [14–16].

2.2.3 DNAzymes that Catalyze Ligation Reactions

The reverse reaction of RNA or DNA cleavage, RNA or DNA ligation, can also be catalyzed by DNAzymes. The first RNA-ligating DNAzyme was isolated in 2003

by Flynn-Charlebois et al. [88,89]. This DNAzyme is capable of ligating a RNA fragment with a 2′,3′-cyclic phosphate to the 5′-hydroxyl of a second RNA fragment. The phosphodiester bond formed by this DNAzyme is a 2′-3′ linkage, unlike the naturally occurring 3′-5′ linkage in DNA [88,89]. On the other hand, an RNA-ligating DNAzyme that formed a 3′-5′ linkage was isolated in 2005. A 5′-triphosphate RNA and a DNAzyme that cleaves the 3′-5′ linkage were introduced into the selection process to favor the formation of the natural linkage [21,90]. Detailed information on the RNA-ligating DNAzymes can be found elsewhere [56,91]. Besides RNA ligation, DNA-ligating DNAzymes have also been reported [18,19].

2.2.4 Other DNAzymes

Other types of reactions that DNAzyme catalyze are listed in Table 1. While most DNAzymes selected so far catalyze reactions that modify a nucleic acid substrate, two DNAzymes in particular catalyze reactions on non-nucleotide substrates: the PS2.M DNAzyme and the DAB22 DNAzyme [33,34].

The PS2.M DNAzyme was originally isolated as a DNA sequence that specifically binds to *N*-methylmesoporphyrin (NMM), a transition-state analogue of a porphyrin metallation reaction [34,35,92,93]. Later studies indicate that hemin [Fe(III)-protoporphyrin IX], an inhibitor of porphyrin metallation, can also bind to the PS2.M DNAzyme, and the DNAzyme/hemin complex can catalyze the peroxidation of 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) or luminol [35,93,94]. The rate of the DNAzyme catalyzed peroxidation is approximately 250 fold higher than the background rate measured in the presence of a control oligo. Instead of forming a helical structure, the 18-nucleotide PS2.M DNAzyme is proposed to have a G-quadruplex structure due to its high guanine content. This G-quadruplex structure of the PS2.M DNAzyme is also crucial for its metallation and peroxidase activity [93–96].

The DAB22 DNAzyme was isolated to catalyze a Diels-Alder reaction. It catalyzes the carbon-carbon bond formation between an anthracene and a maleimide with a rate of 0.7 M^{-1} s^{-1} [33]. This rate is comparable to the ribozyme counterpart which carries out the same reaction [97].

3 Biochemical Mass Spectrometric Studies of DNAzymes

Among the various types of DNAzymes, the RNA-cleaving DNAzymes have been best characterized because of their wide practical applications. These DNAzymes have been studied biochemically via activity assays, mass spectrometry, and UV-induced cross-linking assays. These biochemical assays have provided information about the conserved motifs and catalytic mechanisms of these DNAzymes.

Table 2 Activity of the 8-17 and 10-23 DNAzymes in the presence of different metal ions.

	Metal cofactor	k_{obs}, min^{-1}	k_{cat}/K_m	pK_a	Refs
10-23 DNAzyme	50 mM Mg^{2+}	3.4	4.5×10^9	11.4	[4,71,98]
	25 mM Mn^{2+}	1.19	7.0×10^7	10.6	
	25 mM Ca^{2+}	0.863	1.4×10^7	12.9	
	25 mM Mg^{2+}	0.961	2.2×10^7	11.4	
	25 mM Ba^{2+}	0.101	0.26×10^7	13.5	
	10 mM Mn^{2+}	>4	n.r.	10.6	
	10 mM Mg^{2+}	0.28	n.r.	11.4	
	10 mM Ca^{2+}	0.12	n.r.	12.9	
	10 mM Sr^{2+}	0.026	n.r.	13.2	
	10 mM Ba^{2+}	0.015	n.r.	13.5	
8-17 DNAzyme	2 mM Mg^{2+}	≈0.01	n.r.	11.4	[4,68,70]
	0.2 mM Pb^{2+}	0.47	n.r.	7.8	
	5 mM Zn^{2+}	12	n.r.	8.96	
	3 mM Mn^{2+}	≈0.1	n.r.	10.6	
	3 mM Mg^{2+}	≈0.002	n.r.	11.4	
	3 mM Ca^{2+}	≈0.02	n.r	12.9	
17E DNAzyme (8-17 variant)	0.1 mM Pb^{2+}	5.75	n.r.	7.8	
	10 mM Zn^{2+}	1.35	n.r.	8.96	[68]
	10 mM Mn^{2+}	0.24	n.r.	10.6	
	10 mM Mg^{2+}	0.017	n.r.	11.4	
	10 mM Ca^{2+}	0.015	n.r.	12.9	
Mg5 DNAzyme (8-17 variant)	0.2 mM Pb^{2+}	2.1	n.r.	7.8	
	5 mM Zn^{2+}	0.74	n.r.	8.96	[68,70]
	3 mM Mn^{2+}	> 3	n.r.	10.6	
	3 mM Mg^{2+}	0.06	n.r.	11.4	
	3 mM Ca^{2+}	1	n.r.	12.9	

n.r. = not reported

3.1 Activity Assays

When these DNAzymes cleave their RNA-containing substrates, they generate products with different lengths. Thus, their activity can be easily monitored by PAGE. Vital information on these DNAzymes can be obtained by combining the activity assays with mutational studies.

The 10-23 and 8-17 DNAzymes have been studied most extensively by these methods [4,68–70,98]. Basic biochemical parameters and their metal-dependent changes have been extracted from these studies and are shown in Table 2 (different variants of the 8-17 DNAzymes have also been shown for comparison). Both 10-23 and 8-17 DNAzymes can catalyze reactions very efficiently: their catalytic rates approach 10 min^{-1}, which is very close to that of naturally existing ribozymes [4]. Studies on the metal dependence of the 8-17 DNAzyme and its variants are much more extensive than that of the 10-23 DNAzyme. A study by the Lu group and

others in 2003 demonstrated that the 17E, as well as other 8-17 DNAzymes, has a higher reaction rate in the presence of Pb^{2+} [68]. Interestingly, although both the 10-23 and 8-17 DNAzymes were selected in the presence of Mg^{2+}, Mg^{2+} is not the metal cofactor that can induce the highest catalytic activity. Instead, Mn^{2+} promotes the highest activity for the 10-23 DNAzyme and Pb^{2+} has the same effect for the 8-17 DNAzyme [68,71,98].

In addition to the rate constants and metal dependence, important information about the conserved sequences required for the activities of these enzymes was obtained through mutational studies [98–100]. In these studies, nucleotides in different positions in the DNAzyme were replaced with natural or unnatural bases, revealing the importance of the nucleotide or functional group at particular positions (Figure 2a and 2b). A deletion study carried out in 2004 [100], agrees with the original hypothesis that the 10-23 core is almost intolerant to mutations: only C_7 and T_8 deletion remain highly active. A C_7/T_8 double deletion resulted in a 10-fold decrease in the catalytic rate in the presence of Mg^{2+} [4,101]. Interestingly, a four-nucleotide deletion ($A_5G_6C_7T_8$) from the 10-23 DNAzyme still resulted in an active enzyme and the resulting DNAzyme was highly selective for Ca^{2+} [98]. Systematic mutagenesis of the 10-23 DNAzyme's catalytic core revealed that the nucleotides closer to the borders (e.g., G_1, G_2, T_4, G_6, and G_{14}) are the least tolerant to mutation, while nucleotides in the center are much more tolerant [99]. Only a few of the mutations in the region from positions 7-12 reduce the activity of the DNAzyme, which is not consistent with the original report in 1997 [4]. In the same study, the authors have also tried to replace certain nucleobases with inosine and other nucleobase analogs such as 2-aminopurine and purine. The nucleotide analog substitutions indicate that the functional groups on certain bases are crucial for the 10-23 DNAzyme's activity, such as the 2-amino and the 6-keto groups on G_{14}, as well as the exocyclic amino group on A_5.

Similar investigations on the 8-17 DNAzyme and its variants [68,102] suggested a sequence requirement that is very close to the original proposal [4], including the importance of a G•T wobble pair next to the cleavage site, a AGC loop, a 3-base pair stem and a single-stranded region with the sequence WCGA(A) (W = A or T). An unnatural nucleotide substitution study has also been performed on the 8-17 catalytic core, stressing the critical role of N7 on A_6, O6 on G_7, and N2 on G_{14} [102]. Although the 8-17 DNAzyme was initially reported to cleave only a 5′-AG-3′ junction, later studies on other 8-17 variants demonstrated their capability to cleave all sixteen possible dinucleotide junctions with different rates [5,9,68,75]. Mass spectrometric analysis of the substrate cleaved by the 8-17 DNAzyme suggested that different products were formed in the presence of different metal ions [5,68]. When the cleavage reaction is catalyzed by Zn^{2+} or Mg^{2+}, the formation of 2′,3′-cyclic phosphate (at the 3′-terminus) was observed by MALDI-TOF mass spectrometry; when the reaction was catalyzed by Pb^{2+}, 3′ (or 2′) monophosphate formation is observed at the 3′-terminus [68].

In a recent mutational study, an abasic nucleotide or C3 acrylic spacer was substituted for the key residues within the catalytic core of the 8-17 and 10-23 DNAzymes [103]. Besides reconfirming that certain nucleotides are absolutely

required for the activities of these DNAzymes, the authors proposed that many other nucleotides within the catalytic core perform a chaperone-like function and they are only important for the optimal activity of the DNAzyme. Four nucleotides are absolutely required for both the 8-17 and 10-23 DNAzymes and coincidentally, the positions and neighboring nucleotides of these four nucleotides are very similar.

3.2 Cross-linking Studies

The activity assays have provided vital information on the sequence and metal cofactor requirements of the DNAzymes, but very little information about the structure of the DNAzymes. Several cross-linking studies carried out by the Sen group have provided some insight into the spatial arrangement of the DNAzyme [104–106]. In these studies, halo- and thiosubstituted nucleotides were used to replace some of the nucleotides in the catalytic core of the 8-17 DNAzyme. Upon irradiation with UV light, these halo- and thiosubstituted nucleotides formed cross-linked adducts with nucleotides that are contacting. These adducts can be treated with piperidine, which causes the linked nucleotides to undergo deglycosylation and β-elimination at the 3′ phosphate, resulting in strand cleavage [107]. The site of the crosslink can then be identified via PAGE [104].

The cross-linking study on the 8-17 DNAzyme in the presence of Mg^{2+} revealed strong interaction between nucleotides in the $A_6G_7C_8$ loop, and particularly $C_{13}G_{14}$ with $T_{2.1}$ in the wobble pair (Figure 2a). This result is in close agreement with previous mutational studies [68,102], stressing the importance of these nucleotides for the catalytic activity of the 8-17 DNAzyme. The authors provided a more interactive picture on the active site of the 8-17 DNAzyme based on available data: the AGC loop and the WCGAA region bend toward the cleavage site and the G•T wobble pair, forming close contacts between $T_{2.1}$, A_6, G_7, C_{13} and G_{14} [104]. A follow-up study addressing the cross-links formed in the presence of Pb^{2+} showed only local contacts. No global rearrangements, e.g., contacts between the AGC loop and $T_{2.1}$, were observed [105]. This Pb^{2+}-dependent study, together with the mass spectrometric analysis of the cleavage products [68] and FRET studies [108,109], suggest that Pb^{2+} is a unique metal cofactor for the 8-17 DNAzyme. A stereo-chemical study of the active site residues of the 8-17 DNAzyme have also been reported using an iodine-mediated cross-link approach. The study revealed that C_{13} occupies an asymmetrical position which is critical for catalysis [106].

Another study by Leung and Sen used an electron hole flow pattern, and suggested that C_{13} and G_{14} in the single-stranded loop region are highly solvent exposed. This observation agrees with what was previously known: that these two nucleotides are critical for catalysis. G_7, in contrast, is not solvent-exposed, but is similar to a guanine within a double helix. This observation suggests that G_7 helps organizing the active site [110]. G_7's lesser extent of solvent exposure seems to be in agreement with a previous study that indicated that G_7 is sensitive to steric hindrance [102].

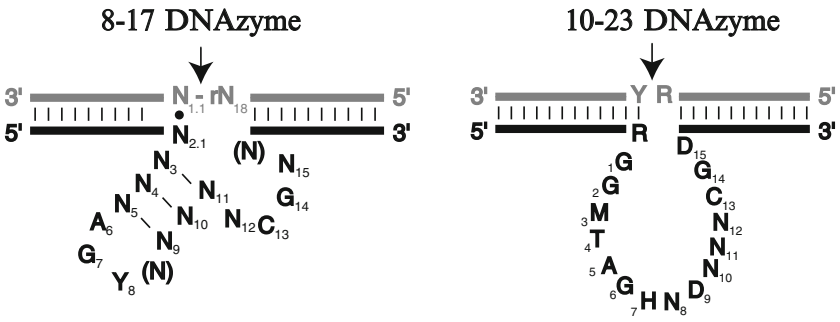


Figure 3 The conserved nucleotides in the 8-17 and 10-23 DNAzymes. N = A, T, C or G; r = ribonucleotide; Y = T or C; R = A or G; M = A or C; H = A, C or T; D = A, G or T; the arrow denotes the cleavage site.

3.3 Conserved Sequences and Motifs of DNAzymes

Among all the DNAzymes, the conserved sequence for the 8-17 DNAzyme is the most extensively studied. The members of this sequence family isolated from *in vitro* selections by several groups [4,5,7,70,111] provide a good starting point for analyzing its sequence conservation. The first published study of this DNAzyme suggested that its conserved sequence included the G•T wobble base pair, a 3-base pair stem, an AGC trinucleotide loop, and a single-stranded region with the sequence WCGR(A) (W = A or T; R = A or G) [68,102]. Later studies have shown that the conserved sequence of the 8-17 DNAzyme is smaller than previously thought, especially when the cleavage site is taken into account [9,75,112]. Based on these studies, only four nucleotides are strictly conserved: A_6, G_7 in the tri-nucleotide loop, and C_{13}, G_{14} in the single-stranded region (see Figure 3).

Although not as extensively studied as the 8-17 DNAzyme, the motif for the 10-23 DNAzyme has also been defined (Figure 3) [4,100,101,103]. Other DNAzyme motifs have been identified through *in vitro* selection as well, but the sequence conservation of these DNAzyme motifs has not been studied yet in detail.

3.4 Reaction Mechanisms

Based on the biochemical studies carried out to date, the reaction mechanism of most RNA-cleaving DNAzymes is believed to be consistent with metal-assisted general acid-base catalysis. However, without the three-dimensional structure of any active DNAzyme, a detailed mechanism is still out-of-reach. In the case of the 8-17 DNAzyme, cleavage product analysis shows that (with the exception of the Pb^{2+}-catalyzed reaction) the products contain a 2′,3′-cyclic phosphate at the 3′-terminus and a 5′-hydroxyl at the 5′-terminus [5,68]. These cleavage products indicate

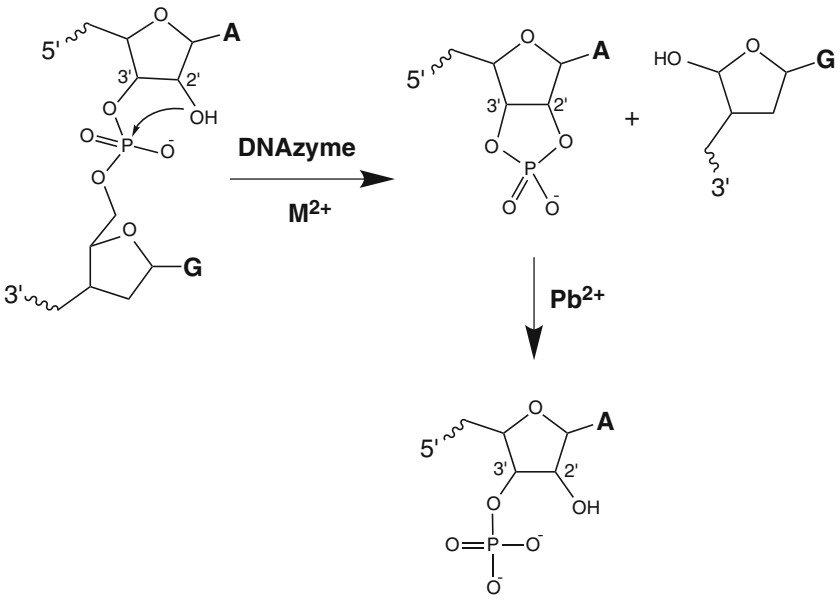


Figure 4 The proposed mechanism for the 8-17 catalysis based on biochemical studies on the 8-17 DNAzyme. The hydrolysis of the cyclic phosphate to monophosphate only occurs in the presence of Pb^{2+}.

a nucleophilic attack from the 2′-hydroxyl on the scissile phosphodiester bond. The attack results in a 5-coordinated phosphate intermediate and the cleavage of the P-O bond leading to the formation of the products. In the presence of Pb^{2+}, the cyclic phosphate is further hydrolyzed to a monophosphate (Figure 4) [68]. This second step hydrolysis by Pb^{2+} is observed in leadzyme (ribozyme) and protein ribonucleases. From other studies, the rate of 8-17 catalyzed reactions plotted versus pH on a logarithmic scale is linear with a slope of one [5,68,113,114]. This observation suggests a single deprotonation event is the rate-limiting step, and this can be assigned to the nucleophilic attack of the 2′-hydroxyl.

A linear relationship between the rates of the reaction and the pK_a values of the active metal cofactor hydrates has also been observed. This observation supports the idea that the deprotonation of the 2′-hydroxyl is assisted by the corresponding metal cofactors [72,73]. Direct metal coordination is probably required for the 8-17 DNAzyme catalysis, since the cobalt complex $[Co(NH_3)_6]^{3+}$ (a structural analogue of hexahydrate Mg^{2+}, but incapable of coordinating to the ribose, phosphate and water) does not support the activity of the DNAzyme [115]. That no significant activity can be observed for the DNAzyme in the presence of high concentration of monovalent ions, may also suggest that the metal cofactor plays a direct role in catalysis [116]; in addition, metal cofactors can also assist the folding of this DNAzyme. The 10-23 DNAzyme also employs a very similar mechanism, except for the hydrolysis of the cyclic phosphate to the monophosphate [69,71,103,117].

4 Biophysical Studies of DNAzymes

To date, no three-dimensional X-ray or NMR structure of a DNAzyme in an active conformation has been published. The only crystal structure available is a 10-23 DNAzyme that resembles a 2:2 enzyme:substrate complex in a four-way junction. This stoichiometry and conformation do not reflect the active form of the 10-23 DNAzyme. As work to obtain the three-dimensional structure of an active DNAzyme continues, the structure and folding of DNAzymes are being studied by various spectroscopic methods, such as Förster resonance energy transfer (FRET), circular dichroism (CD), Tb^{3+} luminescence, and NMR.

4.1 FRET and Single Molecule FRET Studies

FRET is a widely accepted technique for studying the structural dynamics of biomolecules in solution [118–121]. Various types of macromolecule structures, such as double-stranded DNA, three- and four-way DNA junctions, hammerhead and hairpin ribozymes; as well as dynamic process, such as DNA/RNA folding, ribozyme catalysis, DNA/protein interactions, have been studied by FRET [122–134]. In a FRET experiment, the molecule is labelled with donor and acceptor fluorophores that exhibit spectral overlap between donor emission and acceptor excitation. Upon the excitation of the donor, fluorescent energy can transfer to the acceptor depending on the distance between the two. The efficiency of the energy transfer is highly sensitive to inter-fluorophore distance. From the change of the FRET efficiency, the global conformation change can be deduced [135]. Although FRET studies are normally carried out with dual fluorophores, multi-fluorophore FRET has also been applied in protein studies [136-138].

The first FRET study on the 8-17 DNAzyme involved a tri-fluorophore labelled DNAzyme [135]. The 8-17 variant used in the study, 17E, was an inactive construct where the ribose at the cleavage site has been replaced with a deoxyribose. The three arms of the DNAzyme are labelled with three fluorophores: 6-carboxyfluorescein (FAM), Cy5, and 6-carboxytetramethylrhodamine (TMR) (as shown in Figure 5a). The FRET efficiency changes between FAM-TMR, FAM-Cy5, and TMR-Cy5 were monitored and the global conformation changes of the DNAzyme were monitored in the presence of Zn^{2+}. The result indicates that the 17E DNAzyme undergoes a two-step folding process in the presence of Zn^{2+}: at low Zn^{2+} concentrations, the FAM-labelled arm folds towards the Cy5-labelled arm with a K_d of 19 μM; at higher concentration of Zn^{2+} (K_d = 260 μM), the TMR-labelled arm folds towards the other two arms, forming a more compact structure.

Another tri-fluorophore labelled FRET study emphasizing the folding of the 8-17 DNAzyme with different cleavage junctions has also been carried out [139]. The study also revealed that the DNAzyme undergoes two-step folding in the presence of Mn^{2+}, where the first folding step occurs at 1–10 mM Mn^{2+} and is

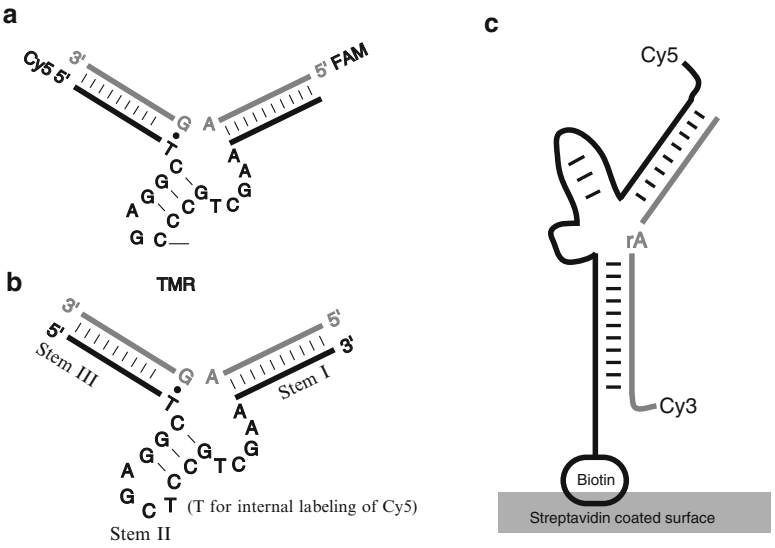


Figure 5 The 8-17 constructs used in various FRET studies. (**a**). The construct used for the tri-color FRET. (**b**). The construct for the bulk FRET study. Pairs of Cy5/Cy3 are separately labelled on Stem I/II, Stem II/III and Stem I/III. A thymine on Stem II is used for internal Cy5 labelling. FRET efficiencies between different stems were measured. (**c**). The construct for a smFRET study contained a riboA at the cleavage site. The Cy5 and Cy3 labelled DNAzyme:substrate complex is attached on the glass slide via a biotin-streptavidin interaction. The time-dependent FRET efficiency was monitored in the presence of different divalent metal ions.

independent of the cleavage junction, and the second folding step occurs at 20–200 mM Mn^{2+} and is dependent on the cleavage junction. The 8-17 DNAzymes with purine-purine junctions folded at lower Mn^{2+} concentration than the ones with pyrimidine-pyrimidine junctions [139].

In 2007, two more FRET studies provided a more complete view on the metal-dependent global folding and folding dynamics of the 8-17 DNAzyme (Figure 5b) [108,109]. Global folding of the DNAzyme was observed for both Zn^{2+} and Mg^{2+} and lower concentration of Zn^{2+} than Mg^{2+} is required for the process. This observation is consistent with the fact that the 8-17 DNAzyme was active at lower concentrations of Zn^{2+}. However, no conformational change is observed in the presence of Pb^{2+}, the metal ion cofactor that supports the highest level of activity [108]. The second single molecule FRET (smFRET) study addressed the problem that the 8-17 DNAzyme construct used in the ensemble FRET was inactive (Figure 5c). The reaction catalyzed by the 8-17 DNAzyme is followed in real-time in the smFRET study. Fluorescence time trace for each individual molecule was recorded and the FRET efficiency was monitored. Changes in the FRET efficiency corresponded to an individual state in the dynamic process. In the presence of Zn^{2+} and Mg^{2+}, a folded state is observed before cleavage. In contrast, in the presence of Pb^{2+}, cleavage is observed

without a folding step. These results reveal that the 8-17 DNAzyme uses an "induced fit" catalytic mechanism in the presence of Zn^{2+} and Mg^{2+} while using a "lock-and-key" mechanism in the presence of Pb^{2+}. These results may help to explain why Pb^{2+} can support the highest rate of reaction [109].

A comprehensive investigation on the folding and activity of the 8-17 DNAzyme in the presence of monovalent ions has been conducted by Mazumdar et al. [116]. Folding is observed by ensemble FRET for all monovalent ions examined in the study at high concentrations (Li^+, Na^+, NH_4^+, Rb^+, and Cs^+), while only Li^+, Na^+, and NH_4^+ confer activity.

A newly developed, three-color, alternating-laser excitation FRET method has been reported for studying the 8-17 DNAzyme with similar results. This new method may provide more detailed insights into the folding and reaction dynamics of DNAzymes in general [140,141].

4.2 NMR Studies

Although NMR spectroscopy can provide vital information on the specific interactions between the nucleotides in the DNAzyme, only a limited number of NMR studies have been carried out with DNAzymes. The ^1H NMR spectra for the lead-dependent DNAzyme selected in 1994 was first obtained. No significant chemical shift was observed when 0–19 mM Mg^{2+} was titrated into the DNAzyme in solution. Unfortunately, no Pb^{2+} titration was performed, since Pb^{2+} is the active metal cofactor. NMR spectroscopy has also been attempted on the 10-23 DNAzyme, but only the base pairing features between the substrate binding arms can be resolved; no information on the catalytic core was provided [101]. Further NMR analysis on DNAzymes could assist the understanding of the specific interactions, such as the DNA/DNA and DNA/metal interactions, and the catalytic mechanism [142].

4.3 Other Spectroscopic Studies

Lanthanide ions (Lu^{3+}, Tb^{3+}, Eu^{3+}) have often been used as spectroscopic probes for nucleic acids, since they show sensitized luminescence by energy transfer from the nucleic acids [143–149]. They are used in many ribozyme studies to probe the metal binding sites by competing with other cations. Luminescence lifetime measurements can be used to determine the number of water molecules coordinated to the lanthanide ions.

Tb^{3+} luminescence spectroscopy was recently used to study the metal binding in the 8-17 DNAzyme [150]. Tb^{3+} can competitively and reversibly inhibit the Zn^{2+}- and Pb^{2+}-dependent enzyme activity. This property of Tb^{3+} allowed a further study of the metal ion binding of the DNAzyme. Luminescence lifetime measurements indicated that 6–7 water molecules were coordinated to Tb^{3+}. Since nine water

molecules are normally coordinated to Tb^{3+}, 2–3 water molecules were replaced by functional groups in the DNAzyme. A competition study indicated that Tb^{3+} competes differently for the Pb^{2+} binding site from binding sites of other metal ions such as Zn^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, and Mg^{2+}. This study provides additional evidence for a unique interaction between Pb^{2+} and the 8-17 DNAzyme, which has been indicated by all previous biochemical and biophysical studies.

In contrast to the inhibitory effect observed for the 8-17 DNAzyme, lanthanide ions are able to support the RNA-cleaving activity of the lead-dependent DNAzyme selected in 1994 [151]. However, the mode of catalysis appears to be different from that of the lead driven catalysis. No pH-dependent activity was observed for any of the lanthanide ions tested, while a pH dependent activity is generally observed for Pb^{2+}. The luminescence measurements indicate that the substrate alone is sufficient for binding the Tb^{3+} ion. Absence of the 2′-hydroxyl group on the substrate eliminated Tb^{3+} binding.

CD spectroscopy has also been used to study structural changes of the 8-17 DNAzyme [116]. The transition between the right-handed B-form DNA to the left-handed Z-form DNA can be monitored by CD spectroscopy. Formation of Z-DNA, as indicated by a negative peak at 294 nm, can be induced in the presence of molar concentrations of monovalent ions, such as Li^+, Na^+, Rb^+, Cs^+ or NH_4^+ for DNA sequences with alternating GC bases [152–157]. More importantly, the Z-DNA formation is also observed in the presence of micromolar concentrations of Zn^{2+} and millimolar concentrations of Mg^{2+}. The Z-DNA formation is attributed to the three-base-pair stem in the catalytic core, since the mutation of the middle GC to AT eliminates the Z-DNA formation. No Z-DNA formation is observed in the presence of Pb^{2+}. The CD study correlates very well with previous FRET studies showing that Zn^{2+} and Mg^{2+} have a different mode of interaction with the DNAzyme from Pb^{2+} [108,109].

4.4 Structural Features Learned from the Biophysical Studies

Without any high-resolution crystal or NMR structure of an active DNAzyme, the precise structure of the DNAzyme responsible for the catalytic activity is largely unclear. However, from the biochemical and biophysical studies, certain important structural features of DNAzymes, especially those of the 8-17 DNAzyme, can be deduced and are summaried below.

First, global folding of the 8-17 DNAzyme in the presence of Zn^{2+} and Mg^{2+} (as well as high concentrations of monovalent ions) can be observed and this folding is required for activity. Second, a local B-to-Z DNA transition occurs in the stem region inside the catalytic core. Third, the cross-linking study points out that several key nucleotides make contact in the catalytic core: the AGC trinucleotide loop folds back and contacts the dinucleotide cleavage junction; C_{13} and G_{14} are also in close contact with the cleavage junction. Fourth, the metal binding site appears to be in the WCGAA region as mutagenesis studies indicated. Last, Pb^{2+} is the exception to

all four points mentioned above as no global folding or B-to-Z DNA transition is observed in the presence of Pb^{2+}. In addition, limited nucleotide contacts are detected via the photo cross-linking study and the site of Pb^{2+} binding, as indicated from all the studies, is different from that of Zn^{2+} and Mg^{2+} binding.

Some of these structural features can be also observed for the 10-23 DNAzyme, such as the structural rearrangement to a more compact structure upon Mg^{2+} binding [158]. The nucleotide responsible for metal binding is also deduced [117]. However, these studies on the 10-23 or other DNAzymes are much less comprehensive in comparison with those of the 8-17 DNAzyme.

5 Biosensing Applications of DNAzymes

Heavy metal ion contamination can often pose significant hazards to the health of the general public and the environment. Developing analytical tools for their detection is an important way to monitor contamination. Currently, heavy metal ions are analyzed by various instrumental techniques, such as atomic absorption spectrometry, inductively coupled plasma mass spectroscopy, anodic stripping voltammetry, and X-ray fluorescence spectrometry. Although these instrumental techniques are very sensitive to heavy metals (often lower than ppb level) [159], they require advanced equipment, skilled operators, and pretreatment of samples, and therefore, on-site and real-time monitoring is difficult to achieve [160–162].

The rational design of small-molecule based chemical sensors for metal ions are challenging because it often involves try-and-error processes. The combinatorial search for metal ion sensors presents a new opportunity. Isolated from combinatorial *in vitro* selection experiments, DNAzymes often possess high metal selectivity, such as the 8-17, 39E, and the classic lead-dependent DNAzyme [3,5,10,68]. These RNA-cleaving DNAzymes are specific for a wide range of metal ions, have small catalytic domains, and have a fast reaction rate [163–167]. Combined with the high stability and the low cost of synthesis for DNA, DNAzymes represent good candidates for biosensing applications.

5.1 Fluorescence Sensors

Fluorescence-based sensors can provide high sensitivity with common benchtop or even portable fluorometers. Fluorophores and quenchers are normally covalently attached to DNA in fluorescence-based DNAzyme sensors. Recently, fluorescence sensors without chemically linked fluorophores have also been developed. These labelled free sensors without the chemically linked fluorophores have benefits due to their lower cost.

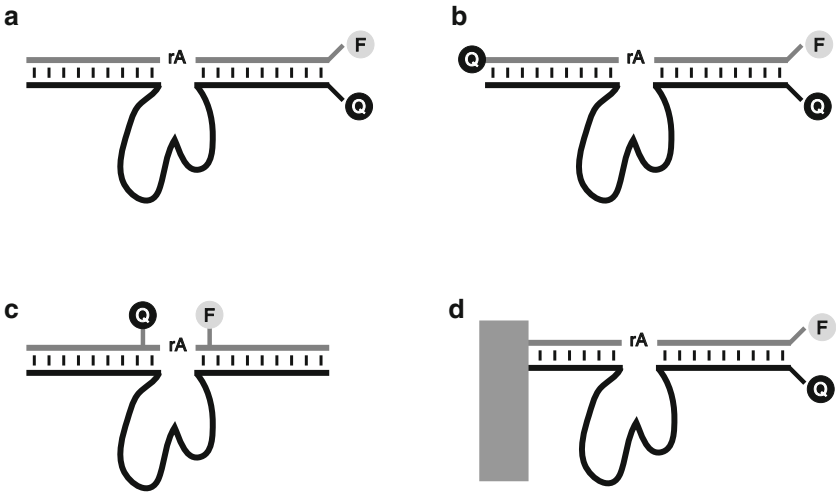


Figure 6 Designs for labeled fluorescence sensors. F denotes a fluorophore and Q denotes a quencher in general. rA denotes the cleavage site. Different fluorophores and quenchers could be used in different studies.

5.1.1 Labelled Fluorescence Sensors

The first fluorescence DNAzyme sensor was reported in 2000 by Li and Lu based on the 8-17 DNAzyme for the detection of Pb^{2+} (Figure 6a) [38]. In this sensor, the fluorophore 6-carboxytetramethylrhodamine (TAMRA) was covalently attached to the 5′-end of the substrate strand and a quencher (4-(4′-dimethylaminophenylazo)benzoic acid, or Dabcyl) was attached to the 3′-end of the enzyme strand. As a result, in the absence of Pb^{2+}, the enzyme and substrate were held together by Watson-Crick base-pairing because the melting temperature was above room temperature, resulting in a low fluorescence quenched state. In the presence of Pb^{2+}, the 8-17 DNAzyme catalyzed the cleavage of the substrate. The cleaved substrate was unstable at room temperature due to its lowered melting temperature with its complementary strand, and therefore it was released from the enzyme strand. The end result was that the fluorophore was separated from the quencher and an increase in the fluorescence could be detected with a fluorometer. By this design, the DNAzyme based biosensor is able to detect Pb^{2+} at concentrations as low as 10 nM, which is lower than 72 nM, the maximum contamination level for drinking water as defined by the U.S. EPA.

The sensor also has 80–1000 fold selectivity against other metal ions because of the intrinsic selectivity of the 8-17E DNAzyme [68]. One drawback of this design is the high background fluorescence even when the assay is carried out at 4°C. As a result, only a low fluorescence enhancement is observed (60%). In a new design, a second intramolecular quencher was introduced into the system and an over six-fold

increase in fluorescence was obtained at room temperature (Figure 6b) [39]. Because of the robust performance, many of the newer fluorescent sensors have been based on this dual-quencher design [10,36,40–42,168].

Some of these sensors possess extremely high sensitivity and selectivity, in some cases they can surpass many instrumental methods. For example, a fluorescent sensor based on the UO_2^{2+}-dependent 39E DNAzyme with incorporated asymmetric substrate binding arms and a dual-quencher can detect as low as 45 pM of UO_2^{2+} in solution, which is lower than the detection limit of ICP-MS for uranium (420 pM) [10]. The toxic level in water for uranium is 130 nM defined by the U.S. EPA. Besides its high sensitivity, this sensor also has an over 1-million fold selectivity against the next best competing ion Th(IV) and hundreds of millions fold selectivity over other metal ions [169]. The improved Pb^{2+} sensor based on the lead-dependent DNAzyme can also offer a 40,000-fold selectivity against other metal ions [36], which is a significant improvement over the 8-17 DNAzyme-based sensors.

Besides modifying the termini of the DNAzyme and the substrate, internal labelling of the fluorophore and quencher can also be used (Figure 6c). Both rational design and combinatorial selection methods have been used for such a sensor design. In the rational design study, different fluorophore and quencher pairs were placed on various positions across the cleavage site of the 8-17 DNAzyme. In general, higher fluorescence enhancement was obtained from the internally labelled fluorophore and quencher, with the strongest signal producing an 85-fold enhancement [170]. Combinatorial selection of metal sensors can be accomplished by introducing a fluorophore and a quencher into the *in vitro* selection of metal-dependent DNAzymes. This type of selection has been carried out by Li and coworkers and they have isolated DNAzymes with high fluorescent enhancements and high catalytic rates [171,172].

Replacing organic dyes with quantum dots (QDs) has become popular. Quantum dots have advantages over organic dyes in terms of their size-dependent emission, single wavelength excitation, higher quantum yield and higher photo stability [173,174]. Therefore, using QD can potentially give higher sensitivity and allow multiplex detection. Using QD-linked DNAzymes, Wu and coworkers demonstrated the simultaneous detection of Pb^{2+} and Cu^{2+} with very high sensitivity (0.2 nM for Pb^{2+} and 0.5 nM for Cu^{2+}). In this study, QDs with different emissions were attached to the substrates of the 8-17 DNAzyme and a Cu^{2+}-dependent DNAzyme; the substrate was also covalently attached to a quencher at the 3′-terminus. Both DNAzymes also had a quencher attached at the 5′-terminus. Before cleavage, the fluorescence of the QDs were quenched; after addition of metal ions, cleavage of the substrates released the quenchers and fluorescence enhancement could be observed when the QDs were excited with a single excitation wavelength [175]. Replacing the quencher with gold nanoparticles can also achieve a lower detection limit because of the better quenching by gold nanoparticles compared to organic dyes [176,177]. A recent study has demonstrated that use of a gold nanoparticle, a fluorophore dual-labelled substrate, and the 8-17 DNAzyme can achieve a detection limit as low as 5 nM [178].

5.1.2 Label-Free Fluorescence Sensors

Labelling the DNAzyme with fluorescent tags costs extra time and materials and they may potentially interfere with enzymatic activities. Therefore, fluorescence sensors without these external tags will have an advantage in these aspects. Recently, several label-free fluorescence sensors have been developed by various groups.

One of the methods for designing label-free sensors involved DNA intercalating dyes. The extrinsic fluorophore 2-amino-5,6,7-trimethyl-1,8-naphthyridine (ATMND) was demonstrated to be a general route to design label-free fluorescence sensors. In two different studies, an abasic site or vacant site was introduced into the double-stranded region of the 8-17 DNAzyme-substrate complex. The fluorophore ATMND can bind to the opposite nucleotide of the abasic or vacant site via hydrogen bonding and base stacking. When ATMND is bound, its fluorescence is quenched. Upon cleavage, the release of the substrate causes the release of ATMND into solution, resulting in a fluorescence enhancement. This design was applied to both the 8-17 DNAzyme and 39E DNAzyme to develop sensors for Pb^{2+} and UO_2^{2+}. Both sensors retained high sensitivity and selectivity [179,180].

Another label-free fluorescence sensor was reported by Zhang et al. utilizing the double-stranded DNA chelating dye Picogreen and the 17E DNAzyme for the detection of Pb^{2+}. The dye binds to double-stranded DNA in the absence of Pb^{2+}, producing a high fluorescence state; after addition of Pb^{2+}, catalytic cleavage of the substrate causes dissociation of the enzyme-substrate complex, releasing the chelated dye and reducing the overall fluorescence. Because of the "turn-off" nature of this system, a higher detection limit of 10 nM is reported when compared to the ATMND system [181].

5.2 Colorimetric Sensors

Although fluorescence sensors can offer high sensitivity and selectivity for detection, they still require instruments for signal output. Even a small portable fluorometer can be a limitation for on-site and real-time detection because of cost issues. Colorimetry is ideal for its convenience because the presence of the analyte can be seen directly by a color change. Metallic nanoparticles, especially gold nanoparticles (AuNPs), display strong distance-dependent optical properties and very high extinction coefficients [182]. The change from a disperse state to an aggregate state of 13 nm AuNPs causes a color change from red to blue. Due to the polyanionic nature of DNA, aggregation of DNA-functionalized AuNPs can be prevented even in molar concentration of NaCl while bare AuNPs are less stable because of the salt-induced screening effect [183]. In addition, DNA has also been demonstrated to control the assembly or disassembly of AuNPs [184–186]. These properties have been utilized to design colorimetric sensors [182–185,187,188].

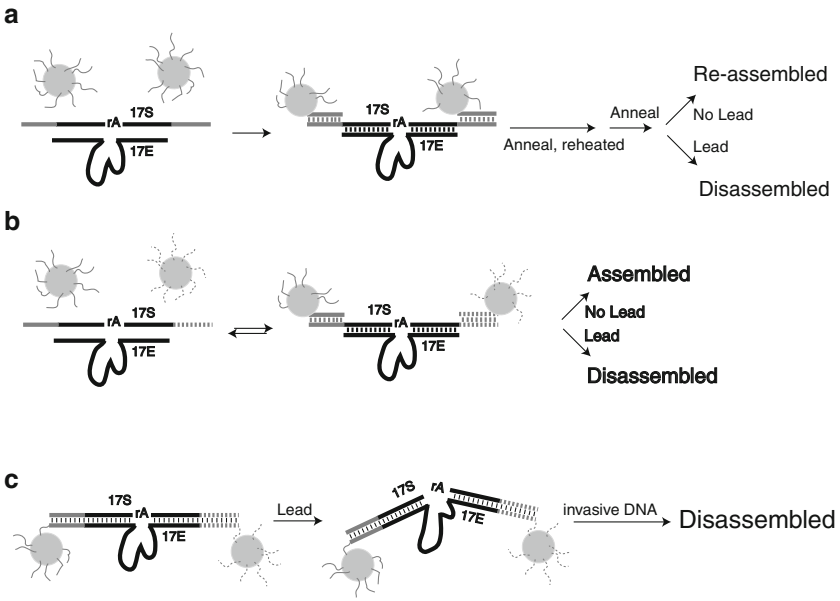


Figure 7 Designs for labelled colorimetric sensors. Grey circles denote gold nanoparticles (AuNPs). (**a**) The first colorimetric DNAzyme sensor based on the assembly and disassembly of AuNP-DNA aggregates. In this design, the AuNPs were arranged in a head-to-tail maner. (**b**) Improved design based on a head-to-head AnNP organization. (**c**) Colorimetric sensor design that utilizes an invasive DNA.

5.2.1 Labelled Colorimetric Sensors

The first colorimetric sensor for Pb^{2+} detection was based on the assembly and disassembly of gold-nanoparticles in the presence of the 8-17 DNAzyme. In this design, both substrate arms were extended and the two extended regions were designed to hybridize to their complementary strands on DNA-functionalized AuNPs (Figure 7a). In the presence of the enzyme strand, aggregate of AuNPs linked by the enzyme-substrate complex can form after a heating-and-cooling hybridization process. This DNA-AuNPs aggregate has a blue color. To perform sensing, the aggregate is heated to 50°C and then followed by a cooling process controlled by the presence of Pb^{2+}. If Pb^{2+} is present, cleavage of the substrate can prevent the formation of the blue aggregate again and a red color can be observed by spotting the sensing solution onto a TLC plate [47]. The initial colorimetric sensor required multi-step operation and long detection time; additionally, the detection limit (~100 nM) was about 50 times higher than the fluorescence sensor. Since this design was originally published, several improvements to the original sensor have been made. The heating-and-cooling process was eliminated by designing a tail-to-tail arrangement of AuNPs and the detection time was decreased to 5 minutes by increasing the size of the AuNPs used from 13 nm to 42 nm. These improvements

to the kinetics of the sensor came at the price of changing the original "turn-on" sensor to a "turn-off" sensor, since no color change can be observed in the presence of Pb^{2+} (Figure 7b) [52].

To achieve fast and "turn-on" sensing with DNAzymes, small fragments of invasive DNA have been used to assist the disassembly of the DNA-AuNPs aggregate (Figure 7c) [189,190]. Later, by using asymmetric substrate binding arms, the disassembly of the DNA-AuNPs aggregate can occur without the presence of invasive DNA at room temperature [191]. A similar design has been applied to the 39E DNAzyme to develop a UO_2^{2+} sensor [48].

Other than cleaving DNAzymes, a Cu^{2+}-dependent ligation DNAzyme has also been used for designing colorimetric sensors by similar methods. The use of a ligation DNAzyme has the advantage of an extremely low background and therefore a high sensitivity [192].

5.2.2 Label-Free Colorimetric Sensors

The colorimetric, label-free detection of an analyte is probably the most favorable method from the end user's perspective. Such a detection method is possible with DNAzymes and AuNPs. A label-free sensor employs the different absorption properties of ss- and dsDNA on citrate-coated AuNPs. Single-stranded DNA can adsorb onto AuNPs via its nucleobases and enhance the salt stability of AuNPs. Without exposed nucleobases, the negatively charged phosphate backbone of dsDNA repels the negatively charged AuNPs. Therefore, no enhanced salt stability can be observed when dsDNA is present to AuNPs [193].

Both Pb^{2+} sensors (8-17 DNAzyme) and a UO_2^{2+} sensor (39E DNAzyme) have been developed [48,49]. Since no modification is required for detection, after the addition of metal ions to the solution containing the DNAzyme and the substrate, the reaction is quenched by addition of EDTA or adjusting the pH after a period of time. Detection is performed by addition of 13 nm AuNPs to the reaction mixtures. The presence of cleaved products then can protect AuNPs from aggregating due to salts present in the reaction system and as a result, no color change is observed. Although being a "turn-off" sensor, this system can achieve very low detection limits (1 nM for UO_2^{2+} and 3 nM for Pb^{2+}). A similar label-free Pb^{2+} sensor has also been reported using bare gold nanoparticles, but with a higher detection limit of 500 nM [50].

All of the sensing methods described above require careful and accurate aliquots of small microliter amounts of solution, making it difficult for the general public with minimal scientific background or training to carry out the experiment successfully. Toward this goal, a dipstick test has been developed based on the 8-17 DNAzyme and AuNP conjugates for detection of Pb^{2+} with practical application for detecting Pb^{2+} in paints (Figure 8) [53]. In this study, the substrate carried a 3′-biotin modification and the 5′-end was extended with a thiol group. The substrate was conjugated to the AuNPs via thiol-gold chemistry. The dipstick is made of four overlapping pads: the wicking pad, the conjugation pad, the lateral flow

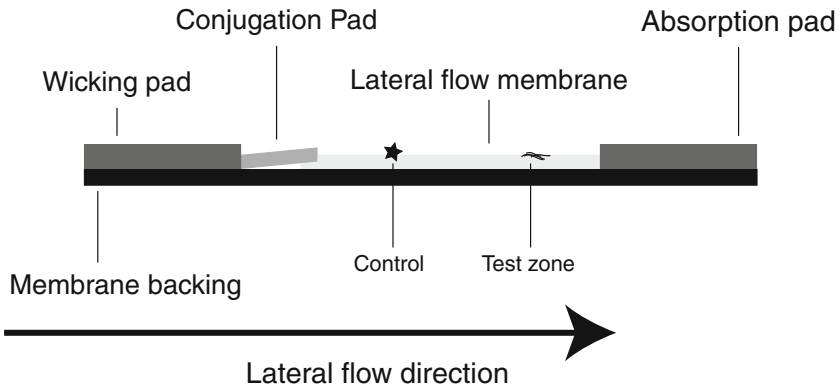


Figure 8 Design of the dipstick for lead detection. The star denotes a control line with streptavidin and the black lines at the test zone denote the capture DNA. The dipstick involved has four overlapping pads on a membrane backing. Buffer flows from the wicking pad towards the absorption pad. Uncleaved DNAzyme-substrate-AuNP conjugates are captured at the control line by streptavidin while the cleaved substrate-AuNP conjugates are captured at the test line by the capture DNA.

membrane and the absorption pad, held in place on a backing sheet. The enzyme-substrate was annealed and spotted onto the conjugation pad of the dipstick. Streptavidin was spotted on the membrane at the control zone. Capture DNA which is complimentary to the releasing part of the cleaved substrate was spotted as the test zone. After all the components dried on the membrane, the dipstick can be used for sensing. If no Pb^{2+} is present, all the DNA-AuNPs conjugates were captured by the streptavidin at the control zone, and one red line results; if Pb^{2+} is present, the cleaved substrate with the conjugated AuNPs can travel up the membrane with the buffer flow and be captured by the capture DNA. Since the cleavage reaction is generally not 100%, two lines can often be observed. With this procedure, a 5 μM detection limit is obtained. With a modified procedure, in which the reaction is carried out in solution and detection is performed on the dipstick, a 500 nM detection limit is achieved.

5.3 Electrochemical Sensors

Besides sensors with optical detections, electrochemical DNAzyme sensors have also been demonstrated. Electrochemical sensors have the advantages of high sensitivity, low cost of the electronic device detector, and convenience due to its miniaturization [79]. The first electrochemical DNAzyme sensor was reported by Xiao et al. [54]. In this study, the 8-17 DNAzyme was modified with a 3′-methylene blue (MB) group and it was 5′-thiolated. The DNAzyme was then conjugated to a gold electrode. When the substrate hybridized to the enzyme strand, the relatively rigid complex prevented

the MB group from approaching the electrode, thus preventing electron transfer. In contrast, the cleavage and release of the substrate allowed the MB to transfer electrons to the electrode. The detection limit for Pb^{2+} with this sensor was 500 nM.

Another electrochemical sensor based on the 8-17 DNAzyme was also reported but using a $[Ru(NH_3)_6]^{3+}$ and DNA-AuNPs for signal amplification [55]. The ruthenium complex can bind to the phosphate backbone of the DNA and the DNA-AuNPs hybridized to an extended region on the releasing arm of the substrate. In the absence of Pb^{2+}, electron transfer from the ruthenium complex to the electrode was maximized and amplified by the AuNPs; in the presence of Pb^{2+}, the cleaved substrate was released from the DNA-AuNPs conjugate, resulting in a reduced number of ruthenium complexes close to the electrode surface and loss of the amplification from the AuNPs. Although being a "turn-off" sensor, this sensor was still highly sensitive, demonstrating a 1 nM detection limit because of signal amplification.

5.4 Other DNAzyme Sensors

Instead of using other molecules to generate a fluorescence signal, a hemin DNAzyme (PS2.M) has also been used. This DNAzyme was first isolated as a cofactor to bind *N*-methylmesoporphyrin [92]. Later, it was found that this sequence is able to catalyze porphyrin metallation [34,94]. The hemin-DNAzyme complex has peroxidase activity, and as a result, its activity can be monitored by using common peroxidase substrates, such as luminol and ABTS [93,95,96]. The Willner group was the first to apply this DNAzyme in biosensing applications [194,195]. The strategy for using the hemin/DNAzyme complex as the signal module generally involves splitting the DNAzyme into two portions and each portion has an extension that can bind to the target [196]. In the case of a "turn-off" sensor, the presence of the trigger DNA will sterically hinder the formation of the hemin DNAzyme, and thus result in a "turn-off" state; when the trigger DNA is absent, the hemin DNAzyme is in the active state and can produce signals when luminol or ABTS is present [196]. A "turn-on" sensor was developed by first locking a portion of the hemin DNAzyme using its complimentary sequence; in the presence of the target, the fully functional DNAzyme is released and produces a signal [197–200]. More detailed reviews of this system are provided elsewhere [86].

Other than solution phase sensors, DNAzyme-based sensors can also be prepared by surface immobilization. Surface immobilization allows lower background and sensor regeneration since unhybridized and cleaved substrate can be washed away. By covalently attaching the 8-17 DNAzyme with a fluorophore-labelled substrate to a gold surface, a detection limit of 1 nM was achieved [201]. This detection limit is about ten times lower than most of the solution phase fluorescent sensors using the same DNAzyme. Besides a lower detection limit, immobilization of the DNAzyme on a gold-coated nanocapillary membrane can retain the activity of the sensor for a 30-day period at room temperature in a dried state [202,203]. Other than gold surfaces, carbon nanotubes and silica gel have also been used to immobilize and entrap DNAzyme sensors [204,205]. One step further for immobilized DNAzyme sensors

is the application in micro- and nanofluidic devices. Immobilizing DNAzyme sensors in these devices allow detection with a very small amount of materials. Generally, less than 1 nL of DNA solution is needed for sensing [44,206].

6 Concluding Remarks and Future Directions

In summary, a large number of studies have been carried out to understand the basic biochemistry of DNAzymes using the 8-17 and 10-23 DNAzymes as model systems. Many sensing applications based on these DNAzymes have been demonstrated and some of these sensors are now available commercially [207].

However, many aspects of these DNAzymes still remain to be understood, such as the nature of their metal binding site, their exact catalytic mechanisms, and most importantly, their active structures. Many more studies on these aspects will be needed to obtain a more complete picture of these DNAzymes.

The RNA-cleaving DNAzymes only represent one member of the DNAzyme family. Fundamental studies on the other DNAzymes are important to be able to understand this new class of enzymes.

Besides the basic studies on existing DNAzymes, we look forward to see more DNA catalysts, which can catalyze new types of reactions. New reactions, where the substrates are no longer nucleotides, are highly exciting, since they expand the practical application of DNAzymes into new territories.

For sensing applications, a major challenge is to apply the success of *in vitro* tests to sense and image metal ions in cells or even the human body. The study of the structure and mechanism of DNAzymes and their applications in sensing is a primary example of scientific endeavors that expand both our fundamental understanding and exploration of the frontier of practical applications. Further advances in this field will have great impacts in many areas of science, engineering, and society.

Abbreviations

ABTS	2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid)
ATMND	2-amino-5,6,7-trimethyl-1,8-naphthyridine
AuNP	gold nanoparticle
CD	circular dichroism
Cy5	a cyanine dye
Dabcyl	4-(4′-dimethylaminophenylazo)benzoic acid
DNAzyme	deoxyribozyme
EDTA	ethylenediamine-N,N,N′,N′-tetraacetate
EPA	Environmental Protection Agency
FAM	6-carboxyfluorescein
FRET	Förster resonance energy transfer

hemin	Fe(III)-protoporphyrin IX
ICP-MS	inductively coupled plasma-mass spectrometry
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
MB	methylene blue
NMR	nuclear magnetic resonance
NMM	*N*-methylmesoporphyrin
PCR	polymerase chain reaction
PAGE	polyacrylamide gel electrophoresis
QD	quantum dot
smFRET	single molecule Förster resonance energy transfer
ssDNA	single-stranded DNA
TLC	thin layer chromatography
TMR = TAMRA	6-carboxytetramethylrhodamine

Acknowledgments The research of the Lu group described in this chapter has been generously supported by the U.S. Department of Energy, the National Institutes of Health, the Department of Defense, the Department of Housing and Urban Development, the Environmental Protection Agency, the National Science Foundation, and the Illinois Sustainable Technology Center.

References

1. K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147–157.
2. C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace, S. Altman, *Cell* **1983**, *35*, 849–857.
3. R. R. Breaker, G. F. Joyce, *Chem. Biol*. **1994**, *1*, 223–229.
4. S. W. Santoro, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4262–4266.
5. J. Li, W. Zheng, A. H. Kwon, Y. Lu, *Nucleic Acids Res*. **2000**, *28*, 481–488.
6. C. R. Geyer, D. Sen, *Chem. Biol*. **1997**, *4*, 579–593.
7. D. Faulhammer, M. Famulok, *Angew. Chem., Int. Ed. Engl*. **1996**, *35*, 2837–2841.
8. A. R. Feldman, D. Sen, *J. Mol. Biol*. **2001**, *313*, 283–294.
9. R. P. G. Cruz, J. B. Withers, Y. Li, *Chem. Biol*. **2004**, *11*, 57–67.
10. J. Liu, A. K. Brown, X. Meng, D. M. Cropek, J. D. Istok, D. B. Watson, Y. Lu, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2056–2061.
11. N. Carmi, L. A. Shultz, R. R. Breaker, *Chem. Biol*. **1996**, *3*, 1039–1046.
12. N. Carmi, H. R. Balkhi, R. R. Breaker, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2233–2237.
13. N. Carmi, R. R. Breaker, *Bioorg. Med. Chem*. **2001**, *9*, 2589–2600.
14. M. Chandra, A. Sachdeva, S. K. Silverman, *Nat. Chem. Biol*. **2009**, *5*, 718–720.
15. Y. Xiao, M. Chandra, S. K. Silverman, *Biochemistry* **2010**.
16. Y. Xiao, E. C. Allen, S. K. Silverman, *Chem. Commun*. **2011**, *47*, 1749–1751.
17. J. Burmeister, G. von Kiedrowski, A. D. Ellington, *Angew. Chem., Int. Ed. Engl*. **1997**, *36*, 1321–1324.
18. B. Cuenoud, J. W. Szostak, *Nature* **1995**, *375*, 611–614.
19. A. Sreedhara, Y. Li, R. R. Breaker, *J. Am. Chem. Soc*. **2004**, *126*, 3454–3460.
20. R. L. Coppins, W. E. Purtha, Y. Wang, S. K. Silverman "Synthesis of native 3'-5' RNA linkages by deoxyribozymes", *229th ACS National Meeting*, San Diego, CA, USA, 2005, ORGN–653.
21. W. E. Purtha, R. L. Coppins, M. K. Smalley, S. K. Silverman, *J. Am. Chem. Soc*. **2005**, *127*, 13124–13125.
22. Y. Wang, S. K. Silverman, *Biochemistry* **2003**, *42*, 15252–15263.

23. Y. Wang, S. K. Silverman, *J. Am. Chem. Soc*. **2003**, *125*, 6880–6881.
24. R. L. Coppins, S. K. Silverman, *Nat. Struct. Mol. Biol*. **2004**, *11*, 270–274.
25. R. L. Coppins, S. K. Silverman, *J. Am. Chem. Soc*. **2005**, *127*, 2900–2907.
26. E. D. Pratico, Y. Wang, S. K. Silverman, *Nucleic Acids Res*. **2005**, *33*, 3503–3512.
27. Y. Wang, S. K. Silverman, *Angew. Chem., Int. Ed. Engl*. **2005**, *44*, 5863–5866.
28. P. I. Pradeepkumar, C. Hoebartner, D. A. Baum, S. K. Silverman, *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 1753–1757.
29. W. Wang, L. P. Billen, Y. Li, *Chem. Biol*. **2002**, *9*, 507–517.
30. Y. Li, Y. Liu, R. R. Breaker, *Biochemistry* **2000**, *39*, 3106–3114.
31. C. Höbartner, P. I. Pradeepkumar, S. K. Silverman, *Chem. Commun.* **2007**, 2255–2257.
32. T. L. Sheppard, P. Ordoukhanian, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7802–7807.
33. M. Chandra, S. K. Silverman, *J. Am. Chem. Soc*. **2008**, *130*, 2936–2937.
34. Y. Li, D. Sen, *Nat. Struct. Biol*. **1996**, *3*, 743–747.
35. Y. Li, D. Sen, *Biochemistry* **1997**, *36*, 5589–5599.
36. T. Lan, K. Furuya, Y. Lu, *Chem. Commun*. **2010**, *46*, 3896–3898.
37. A. K. Brown, J. Liu, Y. He, Y. Lu, *ChemBioChem* **2009**, *10*, 486–492.
38. J. Li, Y. Lu, *J. Am. Chem. Soc*. **2000**, *122*, 10466–10467.
39. J. Liu, Y. Lu, *Anal. Chem*. **2003**, *75*, 6666–6672.
40. N. Nagraj, J. Liu, S. Sterling, J. Wu, Y. Lu, *Chem. Commun.* **2009**, 4103–4105.
41. J. Liu, Y. Lu, *Angew. Chem., Int. Ed. Engl*. **2007**, *46*, 7587–7590.
42. J. Liu, Y. Lu, *J. Am. Chem. Soc*. **2007**, *129*, 9838–9839.
43. T. S. Dalavoy, D. P. Wernette, M. Gong, J. V. Sweedler, Y. Lu, B. R. Flachsbart, M. A. Shannon, P. W. Bohn, D. M. Cropek, *Lab Chip* **2008**, *8*, 786–793.
44. I.-H. Chang, J. J. Tulock, J. Liu, W.-S. Kim, D. M. Cannon, Jr., Y. Lu, P. W. Bohn, J. V. Sweedler, D. M. Cropek, *Environ. Sci. Technol*. **2005**, *39*, 3756–3761.
45. T. Li, S. Dong, E. Wang, *J. Am. Chem. Soc*. **2010**, *132*, 13156–13157.
46. X. B. Zhang, Z. Wang, H. Xing, Y. Xiang, Y. Lu, *Anal. Chem*. **2010**, *82*, 5005–5011.
47. J. Liu, Y. Lu, *J. Am. Chem. Soc*. **2003**, *125*, 6642–6643.
48. J. H. Lee, Z. Wang, J. Liu, Y. Lu, *J. Am. Chem. Soc*. **2008**, *130*, 14217–14226.
49. Z. Wang, J. H. Lee, Y. Lu, *Adv. Mater*. **2008**, *20*, 3263–3267.
50. H. Wei, B. Li, J. Li, S. Dong, E. Wang, *Nanotechnology* **2008**, *19*, 095501–095505.
51. J. W. Liu, Y. Lu, *J. Fluoresc*. **2004**, *14*, 343–354.
52. J. Liu, Y. Lu, *Chem. Mater*. **2004**, *16*, 3231–3238.
53. D. Mazumdar, J. Liu, G. Lu, J. Zhou, Y. Lu, *Chem. Commun*. **2009**, *46*, 1416–1418.
54. Y. Xiao, A. A. Rowe, K. W. Plaxco, *J. Am. Chem. Soc*. **2007**, *129*, 262–263.
55. L. Shen, Z. Chen, Y. Li, S. He, S. Xie, X. Xu, Z. Liang, X. Meng, Q. Li, Z. Zhu, M. Li, X. C. Le, Y. Shao, *Anal. Chem*. **2008**, *80*, 6323–6328.
56. D. A. Baum, S. K. Silverman, *Cell. Mol. Life Sci*. **2008**, *65*, 2156–2174.
57. C. R. Dass, P. F. Choong, L. M. Khachigian, *Mol. Cancer. Ther*. **2008**, *7*, 243–251.
58. S. Chakraborti, A. C. Banerjea, *Mol. Ther*. **2003**, *7*, 817–826.
59. R. Goila, A. C. Banerjea, *Biochem. J*. **2001**, *353*, 701–708.
60. S. W. Santoro, in *Synthetic Nucleic Acids as Inhibitors of Gene Expression*, Ed L. M. Khachigian, CRC Press, Boca Raton, 2005, pp. 53–68.
61. T. Toyoda, Y. Imamura, H. Takaku, T. Kashiwagi, K. Hara, J. Iwahashi, Y. Ohtsu, N. Tsumura, H. Kato, N. Hamada, *FEBS Lett*. **2000**, *481*, 113–116.
62. J. Nowakowski, P. J. Shim, G. F. Joyce, C. D. Stout, *Acta Crystallogr. D. Biol. Crystallogr*. **1999**, *D55*, 1885–1892.
63. J. Nowakowski, P. J. Shim, G. S. Prasad, C. D. Stout, G. F. Joyce, *Nat. Struct. Biol*. **1999**, *6*, 151–156.
64. C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.
65. A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.
66. A. A. Beaudry, G. F. Joyce, *Science* **1992**, *257*, 635–641.
67. M. Zuker, *Nucleic Acids Res*. **2003**, *31*, 3406–3415.

68. A. K. Brown, J. Li, C. M. B. Pavot, Y. Lu, *Biochemistry* **2003**, *42*, 7152–7161.
69. S. W. Santoro, G. F. Joyce, *Biochemistry* **1998**, *37*, 13330–13342.
70. A. Peracchi, *J. Biol. Chem*. **2000**, *275*, 11693–11697.
71. Q.-C. He, J.-M. Zhou, D.-M. Zhou, Y. Nakamatsu, T. Baba, K. Taira, *Biomacromolecules* **2002**, *3*, 69–83.
72. R. R. Breaker, G. M. Emilsson, D. Lazarev, S. Nakamura, I. J. Puskarz, A. Roth, N. Sudarsan, *RNA* **2003**, *9*, 949–957.
73. G. M. Emilsson, S. Nakamura, A. Roth, R. R. Breaker, *RNA* **2003**, *9*, 907–918.
74. K. Schlosser, Y. Li, *Biochemistry* **2004**, *43*, 9695–9707.
75. K. Schlosser, J. Gu, L. Sule, Y. Li, *Nucleic Acids Res*. **2008**, *36*, 1472–1481.
76. K. Schlosser, Y. Li, *ChemBioChem* **2010**, *11*, 866–879.
77. S. Sando, T. Sasaki, K. Kanatani, Y. Aoyama, *J. Am. Chem. Soc*. **2003**, *125*, 15720–15721.
78. S. Sando, A. Narita, T. Sasaki, Y. Aoyama, *Org. Biomol. Chem*. **2005**, *3*, 1002–1007.
79. J. Liu, Z. Cao, Y. Lu, *Chem. Rev*. **2009**, *109*, 1948–1998.
80. X. Zhang, R. Kong, Y. Lu, *Annu. Rev. Anal. Chem.* **2011**, *4*, DOI: 10.1146/annurev.anchem.111808.073617.
81. Y. Lu, J. Liu, J. Li, P. J. Bruesehoff, C. M. B. Pavot, A. K. Brown, *Biosens. Bioelectron*. **2003**, *18*, 529–540.
82. M. N. Stojanovic, T. E. Mitchell, D. Stefanovic, *J. Am. Chem. Soc*. **2002**, *124*, 3555–3561.
83. M. N. Stojanovic, D. Stefanovic, *J. Am. Chem. Soc*. **2003**, *125*, 6673–6676.
84. M. N. Stojanovic, S. Semova, D. Kolpashchikov, J. Macdonald, C. Morgan, D. Stefanovic, *J. Am. Chem. Soc*. **2005**, *127*, 6914–6915.
85. H. Lederman, J. Macdonald, D. Stefanovic, M. N. Stojanovic, *Biochemistry* **2006**, *45*, 1194–1199.
86. I. Willner, B. Shlyahovsky, M. Zayats, B. Willner, *Chem. Soc. Rev*. **2008**, *37*, 1153–1165.
87. J. Elbaz, O. Lioubashevski, F. Wang, F. Remacle, R. D. Levine, I. Willner, *Nature Nanotech*. **2010**, *5*, 417–422.
88. A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf, S. K. Silverman, *J. Am. Chem. Soc*. **2003**, *125*, 2444–2454.
89. A. Flynn-Charlebois, T. K. Prior, K. A. Hoadley, S. K. Silverman, *J. Am. Chem. Soc*. **2003**, *125*, 5346–5350.
90. Y. Wang, S. K. Silverman, *Biochemistry* **2005**, *44*, 3017–3023.
91. C. Hobartner, S. K. Silverman, *Biopolymers* **2007**, *87*, 279–292.
92. Y. Li, C. R. Geyer, D. Sen, *Biochemistry* **1996**, *35*, 6911–6922.
93. P. Travascio, Y. Li, D. Sen, *Chem. Biol*. **1998**, *5*, 505–517.
94. P. Travascio, A. J. Bennet, D. Y. Wang, D. Sen, *Chem. Biol*. **1999**, *6*, 779–787.
95. P. Travascio, D. Sen, A. J. Bennet, *Can. J. Chem*. **2006**, *84*, 613–619.
96. H.-W. Lee, D. J. F. Chinnapen, D. Sen, *Pure Appl. Chem*. **2004**, *76*, 1537–1545.
97. B. Seelig, A. Jaschke, *Chem.Biol*. **1999**, *6*, 167–176.
98. N. Sugimoto, Y. Okumoto, T. Ohmichi, *J. Chem. Soc., Perkin Trans. 2* **1999**, 1381–1386.
99. Z. Zaborowska, J. P. Fuerste, V. A. Erdmann, J. Kurreck, *J. Biol. Chem*. **2002**, *277*, 40617–40622.
100. Z. Zaborowska, S. Schubert, J. Kurreck, V. A. Erdmann, *FEBS Lett*. **2004**, *579*, 554–558.
101. G. F. Joyce, *Methods Enzymol*. **2001**, *341*, 503–517.
102. A. Peracchi, M. Bonaccio, M. Clerici, *J. Mol. Biol*. **2005**, *352*, 783–794.
103. B. Wang, L. Cao, W. Chiuman, Y. Li, Z. Xi, *Biochemistry* **2010**, *49*, 7553–7562.
104. Y. Liu, D. Sen, *J. Mol. Biol*. **2008**, *381*, 845–859.
105. Y. Liu, D. Sen, *J. Mol. Biol*. **2010**, *395*, 234.
106. G. S. Sekhon, D. Sen, *Biochemistry* **2010**, *49*, 9072–9077.
107. C. J. Burrows, J. G. Muller, *Chem. Rev*. **1998**, *98*, 1109–1151.
108. H. K. Kim, J. Liu, J. Li, N. Nagraj, M. Li, C. M. B. Pavot, Y. Lu, *J. Am. Chem. Soc*. **2007**, *129*, 6896–6902.
109. H. K. Kim, I. Rasnik, J. Liu, T. Ha, Y. Lu, *Nat. Chem. Biol*. **2007**, *3*, 763–768.

110. E. K. Y. Leung, D. Sen, *Chem. Biol*. **2007**, *14*, 41–51.
111. D. Faulhammer, M. Famulok, *J. Mol. Biol*. **1997**, *269*, 188–202.
112. K. Schlosser, J. Gu, J. C. Lam, Y. Li, *Nucleic Acids Res*. **2008**, *36*, 4768–4777.
113. M. Bonaccio, A. Credali, A. Peracchi, *Nucleic Acids Res*. **2004**, *32*, 916–925.
114. K. Schlosser, Y. Li, *Nucleic Acids Res*. **2009**, *37*, 413–420.
115. J. A. Cowan, *J. Inorg. Biochem*. **1993**, *49*, 171–175.
116. D. Mazumdar, N. Nagraj, H. K. Kim, X. Meng, A. K. Brown, Q. Sun, W. Li, Y. Lu, *J. Am. Chem. Soc*. **2009**, *131*, 5506–5515.
117. B. Nawrot, K. Widera, M. Wojcik, B. Rebowska, G. Nowak, W. J. Stec, *FEBS Lett*. **2007**, *274*, 1062–1072.
118. R. M. Clegg, *Methods Enzymol*. **1992**, *211*, 353–388.
119. M. Lorenz, A. Hillisch, S. Diekmann, *Rev. Mol. Biotechnol*. **2002**, *82*, 197–209.
120. D. M. J. Lilley, T. J. Wilson, *Curr. Opin. Chem. Biol*. **2000**, *4*, 507–517.
121. N. G. Walter, *Methods* **2001**, *25*, 19–30.
122. C. Gohlke, A. I. H. Murchie, D. M. J. Lilley, R. M. Clegg, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11660–11664.
123. F. Stuehmeier, J. B. Welch, A. I. H. Murchie, D. M. J. Lilley, R. M. Clegg, *Biochemistry* **1997**, *36*, 13530–13538.
124. R. M. Clegg, A. I. H. Murchie, D. M. J. Lilley, *Biophys. J*. **1994**, *66*, 99–109.
125. G. S. Bassi, A. I. H. Murchie, F. Walter, R. M. Clegg, D. M. J. Lilley, *EMBO J*. **1997**, *16*, 7481–7489.
126. J. B. Murray, A. A. Seyhan, N. G. Walter, J. M. Burke, W. G. Scott, *Chem. Biol*. **1998**, *5*, 587–595.
127. N. G. Walter, J. M. Burke, D. P. Millar, *Nat. Struct. Biol*. **1999**, *6*, 544–549.
128. K. J. Hampel, J. M. Burke, *Biochemistry* **2001**, *40*, 3723–3729.
129. M. J. B. Pereira, D. A. Harris, D. Rueda, N. G. Walter, *Biochemistry* **2002**, *41*, 730–740.
130. A. Jenne, W. Gmelin, N. Raffler, M. Famulok, *Angew.Chem., Int. Ed*. **1999**, *38*, 1300–1303.
131. X.-w. Fang, T. Pan, T. R. Sosnick, *Nat. Struct. Biol*. **1999**, *6*, 1091–1095.
132. M. I. Wallace, L. Ying, S. Balasubramanian, D. Klenerman, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5584–5589.
133. K. M. Parkhurst, M. Brenowitz, L. J. Parkhurst, *Biochemistry* **1996**, *35*, 7459–7465.
134. V. V. Didenko, *BioTechniques* **2001**, *31*, 1106–1121.
135. J. Liu, Y. Lu, *J. Am. Chem. Soc*. **2002**, *124*, 15208–15216.
136. W. H. Sawyer, R. Y. S. Chan, J. F. Eccleston, B. E. Davidson, S. A. Samat, Y. Yan, *Biochemistry* **2000**, *39*, 5653–5661.
137. A. Bhattacharyya, B. Bhattacharyya, S. Roy, *Eur. J. Biochem*. **1993**, *216*, 757–761.
138. A. K. Tong, S. Jockusch, Z. Li, H.–R. Zhu, D. L. Akins, N. J. Turro, J. Ju, *J. Am. Chem. Soc*. **2001**, *123*, 12923–12924.
139. J. C. F. Lam, Y. Li, *ChemBioChem* **2010**, *11*, 1710–1719.
140. N. K. Lee, H. R. Koh, K. Y. Han, S. K. Kim, *J. Am. Chem. Soc*. **2007**, *129*, 15526–15534.
141. N. K. Lee, H. R. Koh, K. Y. Han, J. Lee, S. K. Kim, *Chem. Commun*. **2010**, *46*, 4683–4685.
142. Y.-J. Choi, H.-J. Han, J.-H. Lee, S.-W. Suh, B.-S. Choi, *Bull. Korean Chem. Soc*. **2000**, *21*, 955–956.
143. D. E. Draper, *Biophys.Chem*. **1985**, *21*, 91–101.
144. A. L. Feig, W. G. Scott, O. C. Uhlenbeck, *Science* **1998**, *279*, 81–84.
145. A. L. Feig, M. Panek, W. D. Horrocks, Jr., O. C. Uhlenbeck, *Chem. Biol*. **1999**, *6*, 801–810.
146. D. S. Gross, H. Simpkins, *J. Biol. Chem*. **1981**, *256*, 9593–9598.
147. M. D. Topal, J. R. Fresco, *Biochemistry* **1980**, *19*, 5531–5537.
148. R. K. O. Sigel, A. M. Pyle, *Met. Ions Biol. Syst*. **2003**, *40*, 477–512.
149. N. L. Greenbaum, C. Mundoma, D. R. Peterman, *Biochemistry* **2001**, *40*, 1124–1134.
150. H. K. Kim, J. Li, N. Nagraj, Y. Lu, *Chem. Eur. J*. **2008**, *232*, 8696.
151. C. R. Geyer, D. Sen, *J. Mol. Biol*. **1998**, *275*, 483–489.
152. F. M. Pohl, T. M. Jovin, *J. Mol. Biol*. **1972**, *67*, 375–396.

153. T. J. Thamann, R. C. Lord, A. H. Wang, A. Rich, *Nucleic Acids Res*. **1981**, *9*, 5443–5457.
154. M. W. Germann, K. H. Schoenwaelder, J. H. Van de Sande, *Biochemistry* **1985**, *24*, 5698–5702.
155. Y. Wang, G. A. Thomas, W. L. Peticolas, *Biochemistry* **1987**, *26*, 5178–5186.
156. L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. Van der Marel, J. H. Van Boom, *Biochemistry* **1988**, *27*, 6327–6331.
157. B. Hernandez, V. Baumruk, C. Gouyette, M. Ghomi, *Biopolymers* **2005**, *78*, 21–34.
158. M. Cieslak, J. Szymanski, R. W. Adamiak, C. S. Cierniewski, *J. Biol. Chem*. **2003**, *278*, 47987–47996.
159. G. W. Ewing, *Analytical Instrumentation Handbook*, M. Dekker, New York, 1997.
160. R. Y. Tsien, A. W. Czarnik, in *Fluorescenct Chemosensors for Ion and Molecule Recognization*, Vol. 538 of *ACS Symposium Series*, Ed A. W. Czarnik, American Chemical Society, Washington, DC, 1993, pp. 130–146.
161. A. W. Czarnik, *Chem. Biol*. **1995**, *2*, 423–428.
162. A. W. Czarnik, *Acc. Chem. Res*. **1994**, *27*, 302–308.
163. K. Schlosser, S. A. McManus, Y. Li, in *The Aptamer Handbook: Functional Oligonucleotides and Their Applications*, Ed S. Klussmann, Wiley-VCH, Weinheim, 2006, pp. 228–261.
164. *Functional Nucleic Acids for Sensing and Other Analytical Applications*, Vol. 8 of *Integrated Analytical Systems*, Eds Y. Li, Y. Lu, Springer, New York, 2009.
165. S. K. Silverman, *Chem. Commun*. **2008**, *30*, 3467–3485.
166. S. K. Silverman, *Acc. Chem. Res*. **2009**, *42*, 1521–1531.
167. Y. Lu, *Chem. Eur. J*. **2002**, *8*, 4589–4596.
168. H. Wang, Y. Kim, H. Liu, Z. Zhu, S. Bamrungsap, W. Tan, *J. Am. Chem. Soc*. **2009**, *131*, 8221–8226.
169. D. W. Boomer, M. J. Powell, *Anal. Chem*. **1987**, *59*, 2810–2813.
170. W. Chiuman, Y. Li, *Nucleic Acids Res*. **2007**, *35*, 401–405.
171. S. H. J. Mei, Z. Liu, J. D. Brennan, Y. Li, *J. Am. Chem. Soc*. **2003**, *125*, 412–420.
172. Z. Liu, S. H. J. Mei, J. D. Brennan, Y. Li, *J. Am. Chem. Soc*. **2003**, *125*, 7539–7545.
173. S. Hohng, T. Ha, *ChemPhysChem* **2005**, *6*, 956–960.
174. D. M. Willard, A. Van Orden, *Nature Mater*. **2003**, *2*, 575–576.
175. C. S. Wu, M. K. Khaing Oo, X. Fan, *ACS Nano* **2010**, *4*, 5897–5904.
176. R. A. Reynolds, III, C. A. Mirkin, R. L. Letsinger, *J. Am. Chem. Soc*. **2000**, *122*, 3795–3796.
177. J. Yguerabide, E. E. Yguerabide, *Anal. Biochem*. **1998**, *262*, 137–156.
178. J. H. Kim, S. H. Han, B. H. Chung, *Biosens. Bioelectron*. **2011**, *26*, 2125–2129.
179. Y. Xiang, A. Tong, Y. Lu, *J. Am. Chem. Soc*. **2009**, *131*, 15352–15357.
180. Y. Xiang, Z. Wang, H. Xing, N. Y. Wong, Y. Lu, *Anal. Chem*. **2010**, *82*, 4122–4129.
181. L. Zhang, B. Han, T. Li, E. Wang, *Chem. Commun*. **2011**, *47*, 3099–3101.
182. J. Liu, Y. Lu, *Nat. Protoc*. **2006**, *1*, 246–252.
183. I. I. Lim, W. Ip, E. Crew, P. N. Njoki, D. Mott, C. J. Zhong, Y. Pan, S. Zhou, *Langmuir* **2007**, *23*, 826–833.
184. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff, *Nature* **1996**, *382*, 607–609.
185. J. J. Storhoff, A. A. Lazarides, R. C. Mucic, C. A. Mirkin, R. L. Letsinger, G. C. Schatz, *J. Am. Chem. Soc*. **2000**, *122*, 4640–4650.
186. H. Li, L. Rothberg, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 14036–14039.
187. W. Zhao, F. Gonzaga, Y. Li, M. A. Brook, *Adv. Mater*. **2007**, *19*, 1766–1771.
188. K. Sato, K. Hosokawa, M. Maeda, *J. Am. Chem. Soc*. **2003**, *125*, 8102–8103.
189. J. Liu, Y. Lu, *J. Am. Chem. Soc*. **2004**, *126*, 12298–12305.
190. J. Liu, D. P. Wernette, Y. Lu, *Angew. Chem., Int. Ed. Engl*. **2005**, *44*, 7290–7293.
191. J. Liu, Y. Lu, *Org. Biomol. Chem*. **2006**, *4*, 3435–3441.
192. J. Liu, Y. Lu, *Chem. Commun.* **2007**, 4872–4874.
193. H. Li, L. J. Rothberg, *J. Am. Chem. Soc*. **2004**, *126*, 10958–10961.
194. I. Willner, R. Baron, B. Willner, *Biosens. Bioelectron*. **2007**, *22*, 1841–1852.
195. I. Willner, B. Willner, E. Katz, *Bioelectrochemistry* **2007**, *70*, 2–11.
196. Y. Xiao, V. Pavlov, R. Gill, T. Bourenko, I. Willner, *ChemBioChem* **2004**, *5*, 374–379.

197. T. Niazov, V. Pavlov, Y. Xiao, R. Gill, I. Willner, *Nano Letters* **2004**, *4*, 1683–1687.
198. V. Pavlov, Y. Xiao, R. Gill, A. Dishon, M. Kotler, I. Willner, *Anal. Chem*. **2004**, *76*, 2152–2156.
199. Y. Xiao, V. Pavlov, T. Niazov, A. Dishon, M. Kotler, I. Willner, *J. Am. Chem. Soc*. **2004**, *126*, 7430–7431.
200. J. Elbaz, M. Moshe, B. Shlyahovsky, I. Willner, *Chem. Eur. J*. **2009**, *15*, 3411–3418.
201. C. B. Swearingen, D. P. Wernette, D. M. Cropek, Y. Lu, J. V. Sweedler, P. W. Bohn, *Anal. Chem*. **2005**, *77*, 442–448.
202. D. P. Wernette, C. B. Swearingen, D. M. Cropek, Y. Lu, J. V. Sweedler, P. W. Bohn, *Analyst* **2006**, *131*, 41–47.
203. D. P. Wernette, C. Mead, P. W. Bohn, Y. Lu, *Langmuir* **2007**, *23*, 9513–9521.
204. T.-J. Yim, J. Liu, Y. Lu, R. S. Kane, J. S. Dordick, *J. Am. Chem. Soc*. **2005**, *127*, 12200–12201.
205. Y. Shen, G. Mackey, N. Rupcich, D. Gloster, W. Chiuman, Y. Li, J. D. Brennan, *Anal. Chem*. **2007**, *79*, 3494–3503.
206. A. K. Shaikh, K. S. Ryu, E. D. Goluch, J.-M. Nam, J. Liu, C. S. Thaxton, N. T. Chiesl, A. E. Barron, Y. Lu, C. A. Mirkin, C. Liu, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9745–9750.
207. ANDalyze, Inc., 2010, http://www.andalyze.com. (accessed on April 8, 2011).

Chapter 9
Enantioselective Catalysis at the DNA Scaffold

Almudena García-Fernández and Gerard Roelfes

Contents

Abstract The unique chiral structure and the highly specific Watson-Crick base-pairing interactions that characterize natural double-stranded DNA, make this natural biopolymer an attractive ligand for asymmetric catalytic processes. In this chapter the applications of DNA as scaffold and chiral ligand in enantioselective transition metal catalysis are presented. An overview of the state of the art for the different approaches to metal-DNA based catalysts is given, followed by an overview of the mechanistic studies that have been performed to date.

A. García-Fernández • G. Roelfes (✉)
Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4,
NL-9747 AG Groningen, The Netherlands
e-mail: j.g.roelfes@rug.nl

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*,
Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_9,

Keywords asymmetric catalysis • chirality • DNA • DNA-based catalysis • reactions in water

1 Introduction

"*This structure has novel features which are of considerable biological interest*". This is the sentence that Watson and Crick used at the beginning of the publication in which they presented the three-dimensional structure of the DNA helix [1]. Almost sixty years later, the understanding of DNA and the development of new applications, both biological and chemical, continue to be an exciting and challenging topic for the 21st century scientific community. Especially in the past two decades, DNA has emerged as a highly versatile molecule for applications beyond its natural function, including nanotechnology [2–4], DNA-templated synthesis [5], and hybrid catalytic systems [6,7].

The main attractive features of DNA for catalysis lie in the highly specific Watson-Crick base-pairing hydrogen-bonding interactions and the well-defined helical structure that characterize natural double-stranded DNA. These two properties can be used for the specific assembly of catalyst structures, for substrate binding and enantiodiscrimination in asymmetric catalysis. In addition, DNA is a cheap and commercially readily available biopolymer. Furthermore, its high solubility in water makes DNA a suitable starting point for the development of new transformations in aqueous medium. Catalysis in water is one of the important research areas with respect to green chemistry [8,9].

This chapter deals with the application of DNA as scaffold and source of chirality in enantioselective transition metal catalysis. An overview of the developments to date will be followed by a discussion of the role of DNA in catalysis and finally an outlook for this novel approach to enantioselective catalysis will be given.

2 DNA in Metal Catalysis

Three general approaches towards DNA-based metal catalysis can be discerned: (i) metal-dependent DNAzymes, (ii) DNA-directed and templated catalysis, and (iii) DNA-based asymmetric catalysis.

DNAzymes are catalytically active DNA's that fold into well-defined tertiary structures analogous to enzymes. Some of these rely on metal ions, although it is often not clear what the role of the metal ion is, i.e., structural or catalytic. As discussed in Chapter 8 of this book, metalloDNAzymes have found many applications [10]; sensing being a prime example. However, to date no (metallo)DNAzymes for enantioselective catalysis have been reported.

In DNA-directed and templated catalysis, DNA hybridization is used to bring catalysts and substrates together or to promote the formation of an active catalyst on a DNA template [11]. Since this does not involve enantioselective synthesis, this topic is not further discussed here.

Finally, in DNA-based asymmetric catalysis, DNA is used as the chiral scaffold for the formation of a hybrid catalyst that can catalyze reactions in an enantioselective fashion. This approach will be detailed below.

3 Concept of DNA-Based Asymmetric Catalysis

The unique chirality of the DNA double helix and its capacity to bind a wide variety of small molecules via intercalation and/or groove binding makes this natural biopolymer an attractive ligand for asymmetric catalytic reactions.

3.1 General Concept

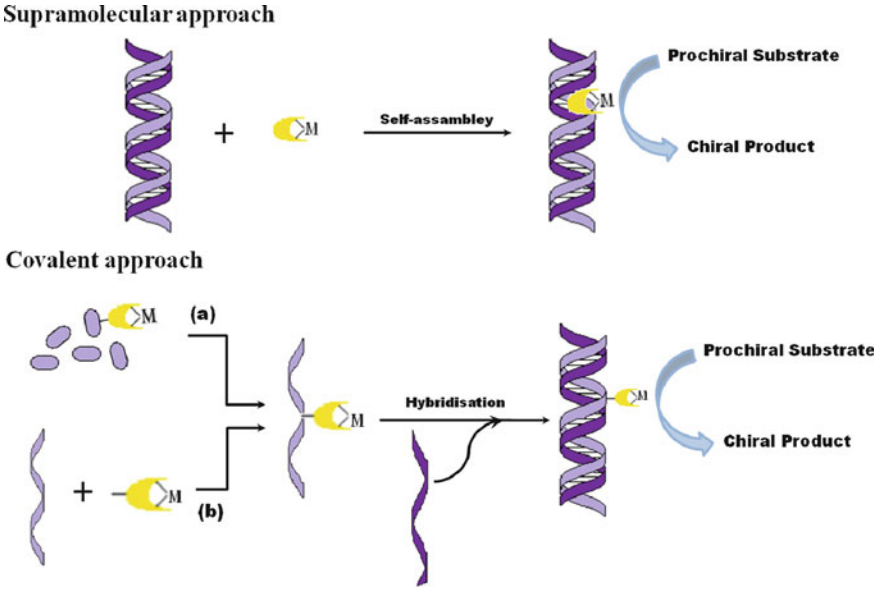
DNA-based asymmetric catalysis involves the assembly of a hybrid catalyst from duplex DNA and a catalytically active transition metal ion, which is coordinated to a non-chiral ligand. Thus, the transition metal complex is brought into close proximity of the DNA and the catalyzed reaction takes place in, or very close to, the chiral DNA helix structure, resulting in the preferential formation of one mirror image form of the reaction product.

3.2 Approaches to DNA-Based Asymmetric Catalysis

Two general approaches can be followed in the design of DNA-based asymmetric catalysis: (i) the supramolecular approach, in which the catalyst is bound using non-covalent interactions and (ii) the covalent approach, in which the catalyst is attached to the DNA via a covalent linkage (Scheme 1).

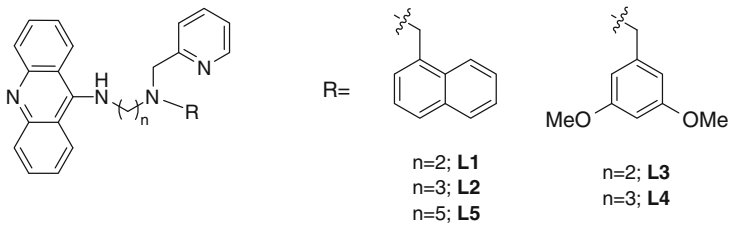
The covalent approach guarantees the well-defined positioning of the catalyst in the DNA, which thus gives rise to excellent control over the catalyst structure. However, the synthetic complexity and time-consuming nature of preparing covalently modified DNAs is complicating the design and optimization process. In contrast, the supramolecular approach allows an easy assembly of the catalytic system, the use of inexpensive DNA from natural sources, and the rapid screening and optimization of the catalysts. However, a potential complication of the supramolecular approach lies in the fact that the catalyst may not be very well defined due to the possibility of the transition metal complex binding at multiple positions to the DNA. As a consequence, the DNA-based catalyst may actually be a heterogeneous mixture of different catalysts that reside in a different micro-environment.

To date, two generations of supramolecular DNA-based catalysts have been developed, which differ in the design of the metal chelator: (a) The first generation ligands, which contain a 9-amino acridine intercalating moiety that is connected to a 2-(aminomethyl)pyridine metal binding moiety through a spacer, and (b) the second generation DNA-based catalysts employ bipyridine-type ligands. Compared to the first generation of ligands, the catalytically active metal center can be brought much closer to the DNA due to the absence of a spacer moiety (Figure 1).



Scheme 1 Schematic representation of DNA-based asymmetric catalysis using the supramolecular and covalent anchoring strategies: (**a**) nucleotide synthesis, (**b**) postsynthetic modifications.

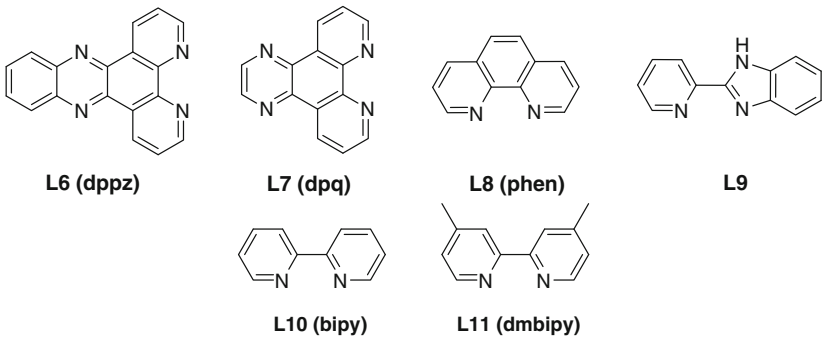
First generation ligands



n=2; **L1**
n=3; **L2**
n=5; **L5**

n=2; **L3**
n=3; **L4**

Second generation ligands



L6 (dppz) **L7 (dpq)** **L8 (phen)** **L9**

L10 (bipy) **L11 (dmbipy)**

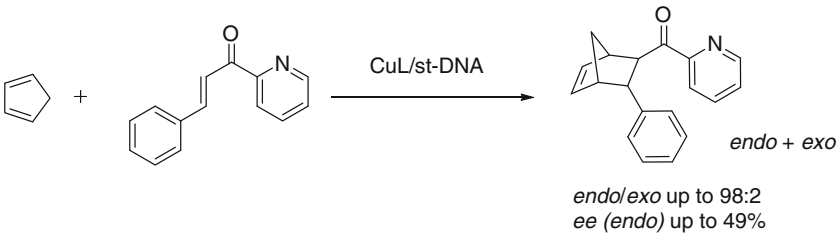
Figure 1 First and second generation ligands used in supramolecular DNA-based catalysis.

4 Reaction Scope of DNA-Based Asymmetric Catalysis

The concept of DNA-based asymmetric catalysis has been applied successfully to various asymmetric reactions, including Lewis acid-catalyzed C-C and C-X bond-forming reactions and organometallic reactions.

4.1 Lewis Acid-Catalyzed C-C Bond-Forming Reactions

In the first application of a DNA-metal complex hybrid catalyst, Roelfes and Feringa reported the use of salmon testes DNA (st-DNA) as source of chirality in the Cu(II)-catalyzed enantioselective Diels-Alder reaction of azachalcone with cyclopentadiene [12] (Scheme 2).



Scheme 2 DNA-based asymmetric Diels-Alder reaction. Typical conditions: 1.3 mg/mL st-DNA; [catalyst] = 0.3 mM; [azachalcone] = 1 mM; [cyclopentadiene] = 5 mM; [3-(N-morpholino)-propanesulfonic acid (= MOPS buffer)] = 20 mM; pH = 6.5; 5°C; 3 days.

The reasons for selecting the Diels-Alder reaction were: (i) The Diels-Alder reaction greatly benefits from water as reaction medium [13,14]. (ii) Kobayashi and Manabe [15] and Engberts et al. [16] had demonstrated the feasibility of Lewis acid catalysis in water in this and related reactions. (iii) This reaction does not involve changes in the oxidation state of the metal center which may give rise to damage of the DNA [17], and (iv) previous work on catalytic antibodies [18,19] and RNAzymes [20] suggested the potential of the Diels-Alder reaction in DNA-based asymmetric catalysis.

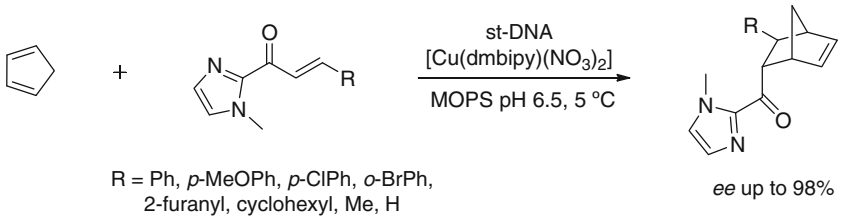
The hybrid catalyst was self-assembled by combination of the copper complex $[CuL(NO_3)_2]$, where L is a first generation ligand, and salmon testes DNA. The structure of the ligand L determines the enantiomeric excess observed in each case. Thus, the highest *ee*'s were obtained when the substituent R in the first generation ligands is an aromatic group such as 1-naphthylmethyl or 3,5-dimethoxybenzyl. This suggests the involvement of π-π stacking interactions between the substituent and the dienophile, as was previously put forward in the case of Cu(II) catalysts based on amino acids [15,16]. Alternatively, cation-π interactions between the arylmethyl group and Cu(II) have also been proposed [21].

When **L2**, with R = 1-naphtylmethyl, was used as ligand (see Figure 1), almost complete *endo* selectivity was achieved (*endo*/*exo* 98:2) and 49% *ee* was observed for the (−) enantiomer of the *endo* product. Interestingly, changing the length of the spacer from three to two carbon atoms (**L1**) gave rise to the formation of (+) enantiomer with 48% *ee*. Elongation of the spacer length up to n = 5 (**L5**) led to strong decrease in enantioselectivity, indicating that close contact between the DNA and the catalyst is required for efficient transfer of chirality.

In the case of R = 3,5-dimethoxybenzyl (**L3, L4**) a dependence between the length of the spacer and the enantiomeric preference of the reaction was not observed; the (+) enantiomer of the Diels-Alder product was obtained in excess in all evaluated cases, with *ee* values up to 37%.

The use of the second generation ligands, which are based on 2,2'-bipyridine-type structures, in the metal-DNA based asymmetric Diels-Alder reaction led to a dramatic improvement of the stereo- and enantioselectivity in this reaction [22]. Nearly complete *endo* selectivity and up to 99% *ee* was observed in the reaction when 4,4'-dimethyl-2,2'-bipyridine (dmbipy) was used as ligand for copper.

An important step towards increasing the synthetic applicability of the Diels-Alder reaction catalyzed by hybrid metal-DNA was the introduction of the α, β-unsaturated 2-acyl(1-methyl)imidazole substrates as dienophiles [23] (Scheme 3). This kind of α,β-unsaturated ketones have proven to be excellent substrates in a variety of enantioselective Lewis acid catalyzed reactions [24–27].

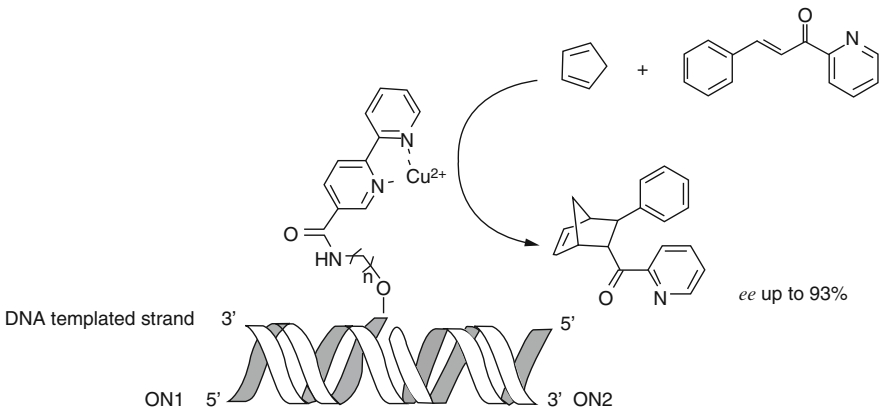


Scheme 3 DNA-based catalytic Diels-Alder reaction with α,β-unsaturated 2-acyl(1-methyl)-imidazole substrates.

Also in this case 4,4'-dimethyl-2,2'-bipyridine was found to be the best ligand, providing excellent diastereoselectivities (up to 99:1) and good to excellent enantioselectivities (80–98% *ee*) with all the substrates evaluated. Moreover, the imidazole moiety could subsequently be removed to generate the corresponding methyl ester [28].

Two more examples of DNA-based catalysis have been reported for the Diels-Alder reaction using Cu(II) as catalytic metal ion. These other examples use the covalent approach to catalyst assembly, attaching the metal ion to the DNA via a covalent linkage. In one report, polyaza crown ethers were incorporated into DNA-conjugates as catalytic Cu(II) binding sites. The catalytic activity of these systems was evaluated in the Diels-Alder reaction of cyclopentadiene and azachalcone [29]. However, only a low enantioselectivity of 10% was observed.

In the other approach, a modular approach towards catalyst assembly was used that involves three oligonucleotide components: an oligonucleotide that is functionalized with a ligand, i.e., 2,2'-bipyridine (bipy) at the 5' or 3' terminal phosphate moieties (ON1), an unfunctionalized oligonucleotide (ON2), and a template oligonucleotide strand that is complementary to both ON1 and ON2. Hybridization in the presence of a metal ion gives rise to a duplex DNA structure that has the catalytically active metal complex in an internal position at the interface between ON1 and ON2 [30] (Scheme 4).

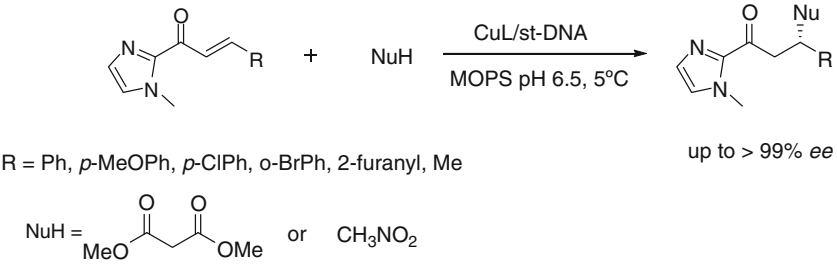


Scheme 4 Diels-Alder cycloaddition using covalent approach to asymmetric DNA-based catalysis.

The conversion and enantiomeric excess of the product proved to be dependent on the design of the catalyst, in particular the length of the spacer and the DNA sequence around the interface. The DNA sequence in the region of the metal center was readily optimized by variation of the unfunctionalized (ON2), resulting in a maximum *ee* of up to 93% for the major (*endo*) enantiomer.

The α,β-unsaturated 2-acyl(1-methyl)imidazole substrates have also been applied successfully in DNA-based catalytic enantioselective Michael addition reactions [31] (Scheme 5).

The hybrid catalyst composed of st-DNA and the complex [Cu(dmbipy)$(NO_3)_2$] catalyzed the addition of dimethyl malonate to (*E*)-1-(1-methyl-1*H*-imidazol-2-yl)-3-phenylprop-2-en-1-one in water to give full conversion and 91% *ee*.

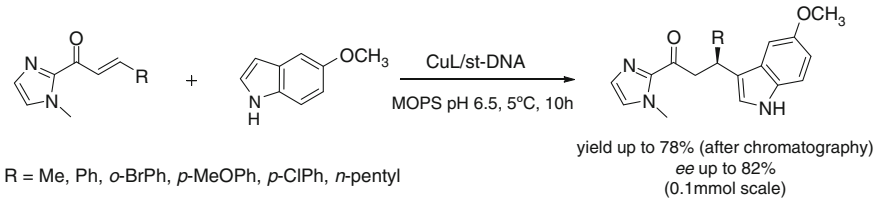


Scheme 5 Asymmetric Michael addition with enolate nucleophiles catalyzed by the copper-ligand complex in the presence of st-DNA.

The substrate scope of this reaction proved to be quite broad: good to excellent enantioselectivities, ranging from 58 to 99% *ee*, were obtained in all cases. When nitromethane was used as Michael donor, a significantly increased concentration of the nucleophile was needed to achieve full conversion. Again, with multiple substrates good to excellent enantioselectivities were observed, albeit slightly lower than obtained when using dimethyl malonate.

The aqueous solution containing the DNA-based catalyst could be recycled several times without loss of catalytic activity or enantioselectivity. Moreover, the use of organic co-solvents such as acetonitrile (10% v/v), allowed, in the case of the Michael addition of dimethyl malonate to (*E*)-1-(1-methyl-1*H*-imidazol-2-yl)-3-phenylprop-2-en-1-one, to carry out the reaction on gram scale, giving rise to 85% isolated yield and 95% *ee* [32]. Furthermore, using the protocol of Evans et al. [24], the imidazole group of the Michael addition products was converted to a good leaving group upon *N*-alkylation, followed by treatment with methanol to yield the corresponding methyl ester derivates.

Using the same α,β-unsaturated 2-acyl(1-methyl)imidazole substrates, the first catalytic enantioselective vinologous Friedel-Crafts alkylation of indoles and pyrroles in aqueous medium was developed [33]. Again, the highest enantioselectivities were obtained using a DNA-based hybrid catalyst based on [Cu(dmbipy)$(NO_3)_2$] (Scheme 6).



Scheme 6 Asymmetric Friedel-Crafts alkylation reaction catalyzed by Cu-L11 (for L11 see Figure 1) in the presence of DNA.

A variety of indole π nucleophiles reacted with α, β-unsaturated 2-acyl imidazoles (Scheme 6) having aryl or alkyl substituents (R) on the β position. Full conversion was found after 10 h in most cases. Only when R was an aryl group, the reaction was slower and needed 20 equivalents of indole for completion. The *ee* varied from 69 to 83% using st-DNA and could be increased to 93% by using the self-complementary oligonucleotide d(TCAGGGCCCTGA)$_2$ (*vide infra*).

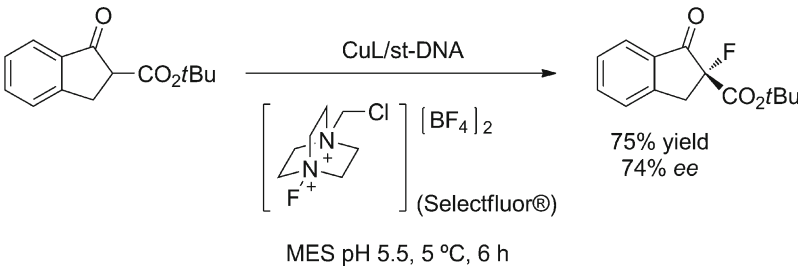
An interesting observation is that in the reaction of (*E*)-1-(1-methyl-1*H*-imidazol-2-yl)but-2-en-1-one (R = CH_3) with 5-methoxy indole, the copper complex loading could be lowered to 0.3 mol%. Based on the binding affinity $K_b = (1.12 \pm 0.02) \times 10^4\ M^{-1}$, determined for the binding of Cu(II)-dmbipy, it can be calculated that under the reaction conditions only 16% of the copper complexes are actually bound to DNA, which corresponds to an effective catalyst loading of only 0.05 mol%. Yet, compared to the highest catalyst loadings, when 95% is bound to DNA, the enantioselectivity did not decrease. This suggests a significant acceleration of the reaction by the presence of DNA (*vide infra*).

Finally, the presence of an organic co-solvent, e.g., methanol, ethanol, dimethylformamide, dimethylsulfoxide or 1,4-dioxane, in the reaction medium is tolerated and, in some cases, caused the reaction to proceed significantly faster [32]. The reaction of 5-methoxyindole with (*E*)-1-(1-methyl-1*H*-imidazol-2-yl)but-2-en-1-one in the presence of 30% v/v methanol made it possible to perform the reaction on gram scale and at –18ºC. Under these conditions the product could be isolated in 85% yield with an increase in the enantioselectivity from 81 to 93%.

4.2 Lewis Acid-Catalyzed C-X Bond-Forming Reactions

In the light of the excellent results obtained with the catalytic carbon-carbon bond-forming reactions, the enantioselective formation of carbon-heteroatom bonds using metal-DNA based catalysts was investigated.

Shibata, Toru, and coworkers reported the DNA-based catalytic enantioselective C-F bond-forming reaction. The Cu(II)-dmbipy/st-DNA hybrid catalyzed the fluorination reaction of indanone-β-ketoesters in aqueous medium using 1-(chloromethyl)-4-fluoro-1,4-diazonibicyclo[2.2.2]octane ditetrafluoroborate (Selectfluor®) as fluorination reagent [34] (Scheme 7).

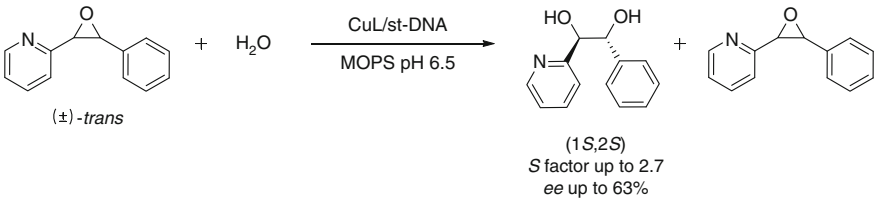


Scheme 7 Asymmetric fluorination reaction of β-keto esters.

In this case, the indanone-2-carboxylate esters used as substrates bind to the Cu(II) center in a bidentate fashion, resulting in formation of the corresponding enolates. These then react with an electrophilic fluorine source, giving rise to the corresponding fluorinated product. The catalytic reaction proved to be very sensitive to the substitutions on the β-keto ester substrates; the best enantioselectivities (up to 74%) where obtained when the indanone carboxylates presented bulky esters as substituents.

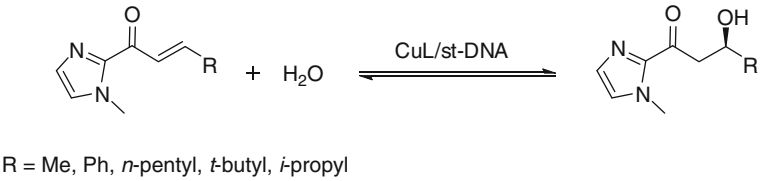
Also the hydrolytic kinetic resolution of 2-pyridyloxiranes in water catalyzed by [Cu(II)-(dmbipy)/st-DNA] has been reported [35]. The selectivity of the reaction showed a strong dependence of the substituents on the oxirane. The best results were obtained with trans-2-(3-phenyloxiranyl)pyridine as substrate. In this case, the selectivity factor observed was 2.7 and the remaining epoxide was recovered

with 63% *ee* (Scheme 8). However, the effectiveness of these hybrid catalysts has to be improved significantly for the reaction to become of interest for synthetic applications.



Scheme 8 Hydrolytic kinetic resolution with a DNA-based catalyst.

Recently, the DNA-based catalytic enantioselective conjugate addition of water to α,β-unsaturated ketones in aqueous medium has been reported [36]. The use of water as nucleophile in conjugate addition reactions represents a challenging topic in asymmetric catalysis and is of considerable synthetic interest, since it allows direct incorporation of hydroxy groups without requiring additional deprotection steps. The DNA-based reaction represents the first non-enzymatic catalytic enantioselective *syn*-hydration of α,β-unsaturated ketones in water, giving rise to chiral β-hydroxy ketones with up to 79% *ee*, which could be further improved to 82% by performing the reaction in D_2O. Furthermore, the catalytic solution could be recovered and recycled at least four times without observing changes in the conversion and in the enantioselectivity of the reaction (Scheme 9).



Scheme 9 Asymmetric conjugate addition of water to enones catalyzed by Cu-L in the presence of DNA.

Surprisingly, the best results in this transformation were obtained using ligands of the first generation. While in absence of the copper complex no conversion was observed, in absence of the ligand, the reaction rate was slower and the opposite enantiomer of the β-hydroxy ketone product was obtained. These results demonstrate that the copper complex is responsible of the activation of the enone for the hydration step, but the combination of the ligand coordinated to the metal center and the DNA controls the stereocontrol outcome of the reaction.

The catalytic activity of the complex [Cu(**L3**)$(NO_3)_2$/st-DNA] (for **L3** see Figure 1) was evaluated for several 2-acyl-(1-methyl)imidazole derivatives. In general, a dependence of the enantioselectivities, which were in the range from 28 to 72%, on

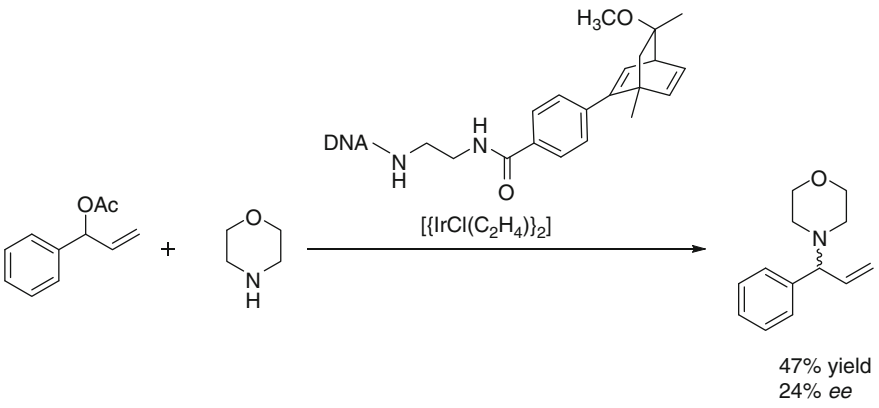
the steric demand of the R substituents was observed: bulky alkyl substituents gave rise to higher enantioselectivities. Only, when R was an aryl substituent (phenyl), no hydration reaction was observed, due to the unfavorable thermodynamics for this reaction.

4.3 Organometallic Catalysis

Compared to Lewis acid catalysis, the development of DNA-based systems for organometallic transformations is less well explored to date.

Kamer et al. reported the synthesis of a uridine nucleobase substituted with a diphosphine group [37]. This modified uridine was used as ligand in the allylic amination catalyzed by Pd(II) complexes, inducing *ee*'s up to 82% when THF was employed as solvent. Unfortunately, when this nucleobase was included as part of a trimer a lower reactivity and the loss of the enantioselectivity was observed. Other approaches to include phosphine ligands into DNA structures have been explored [38,39], but to date no catalysis was reported.

An alternative approach to achieving DNA-based organometallic catalysis involved the covalent attachment of a diene ligand that was used to bind Ir(I) [40]. The DNA-based catalysts were synthesized by coupling of the bicyclooctadiene diene ligands with a 19-*mer* oligodeoxynucleotide and addition of an iridium(I) source. These iridium(I) catalysts have been evaluated in the allylic amination of 1-phenylallyl acetate with morpholine in a water/1,4-dioxane mixture (7:3) (Scheme 10). Although the *ee* values and stereoselectivity factors obtained were low, a chirality transfer from the nucleic acid to the iridium(I) complex could be observed. Using different DNA or RNA oligonucleotides allowed modulating the chiral environment surrounding the metal center. As a result, the enantiomeric outcome of the reaction could be changed, e.g. from an *ee* of +23 to -27% in one case.



Scheme 10 DNA-based iridium(I)-catalyzed allylic amination.

5 Role of DNA in Catalysis

An intriguing question concerns the role of DNA in these DNA-based catalytic reactions is. Is it "merely" the source of chirality or does the DNA affect the catalysis in more ways? For this reason the non-covalent approach to DNA-based catalysis has been subjected to a detailed mechanistic study, particularly focusing on the C-C bond-forming reactions. Most of the information to date has resulted from a study of the Diels-Alder reaction of azachalcone with cyclopentadiene, which proved to be a very convenient model reaction. However, similar trends were found for the other DNA-based catalytic enantioselective C-C bond-forming reactions (*vide infra*). The *syn* hydration of enones has quite different characteristics. For the other reactions and for the covalently linked catalysts, mechanistic information is not yet available.

5.1 Effect on Reactivity

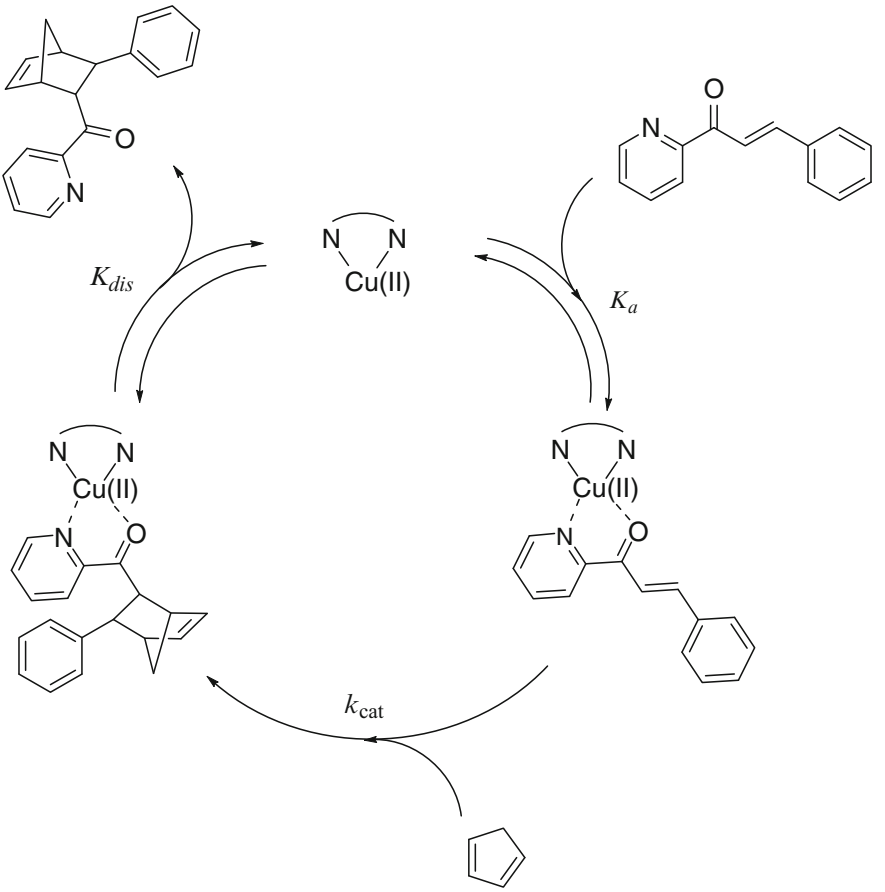
The effect of the DNA on the rates of the catalyzed reactions has been investigated in a kinetic study [7]. In the case of the Diels-Alder cycloaddition of azachalcone with cyclopentadiene catalyzed by Cu-L/st-DNA it was demonstrated that the DNA sequence and the DNA structure have a significant effect on the reaction rate.

In case of catalysts from the first generation, a slight deceleration of the reaction was observed in the presence of DNA compared to the reaction in the absence of DNA [41]. In contrast, using the second generation catalysts, i.e., those based on 2,2'-bipyridine-type ligands, an acceleration of the reaction was found in the presence of DNA. This ranged from a modest 2.4-fold acceleration in case of Cu-bipy to a surprising 58-fold increase in the apparent second order rate constant (k_{app}) with the Cu-dmbipy complex.

The proposed mechanism for this reaction involves reversible bidentate binding of the Cu(II) center to the azachalcone, followed by the irreversible Diels-Alder reaction, and finally dissociation of the product from the catalysts, thus making it available for the next cycle (Scheme 11) [42]. The binding constant K_a for the binding of azachalcone to the Cu(II) complex was found not to be significantly affected by the presence of DNA [43]. Using these values and the k_{app}, the k_{cat} for the actual could be determined. Indeed, it was found that in the presence of DNA the k_{cat} of the reaction increases by two orders of magnitude.

5.2 DNA Sequence Dependence

Salmon testes DNA is a natural DNA which contains, from the perspective of DNA-based catalysis, a random sequence of nucleotides. Since no apparent sequence selectivity for binding of Cu-dmbipy was observed [43], which means that the



Scheme 11 Proposed catalytic cycle of the Diels-Alder reaction catalyzed by Cu(II)-L.

Table 1 Sequence dependence of first and second generation catalysts.

Entry	DNA	first generation (**L3**) *ee* (%)	second generation (**L11**) *ee* (%)
1	st-DNA	37(+)	98.5(+)
2	poly (dA-dT)	6(+)	15(−)
3	poly (dG-dC)	62(+)	78(+)
4	$d(GCGCGCGCGCGC)_2$	54(+)	95(+)
5	$d(GCGCGCGC)_2$	27(+)	86(+)
6	$d(GACTGACTAGTCAGTC)_2$	34(+)	78(+)
7	$d(TCGGGTACCCGA)_2$	16(+)	98.6(+)
8	$d(TCAGGGCCCTGA)_2$	10(+)	99.4(+)

Conditions: all experiments were carried out with the following reagents: 0.3 mM $[Cu(L)(NO_3)_2]$, 1.3 mg mL^{-1} st-DNA, 1 mM azachalcone, 16 mM cyclopentadiene, 20 mM MOPS pH 6.5, 5ºC, 3 days.

Cu-dmbipy binds everywhere on the DNA, the sequence dependence of both the enantioselectivity and rate acceleration was studied. For this purpose a series of self-complementary oligonucleotides were investigated as scaffold in the catalyzed reactions.

The enantioselectivity observed for the Diels-Alder reaction catalyzed by the complex of the first generation [Cu(**L3**)$(NO_3)_2$] was increased from 37%, when st-DNA was used as chiral scaffold, to 62%, when the reaction took place in presence of poly(dG-dC)(dG-dC) which presents alternating GC sequences. In contrast, AT- rich sequences gave rise to a strong decrease of the enantioselectivity values.

With the second generation catalyst based on Cu-dmbipy, a different behavior was observed. Sequences containing alternating GC and AT base pairs, or sequences containing alternations of G and C generally gave lower *ee*'s than obtained with salmon testes DNA. However, sequences containing tracts of three G or three C lead to higher *ee* values. The highest enantioselectivity, 99.4% *ee*, was obtained with $d(TCAGGGCCCTGA)_2$.

The kinetic studies also demonstrate that the sequences have a significant influence over the catalytic Diels-Alder reactions. Surprisingly, it was found that the DNA sequences that give rise to the highest enantioselectivities are also those that cause the largest rate accelerations (Figure 2) [43].

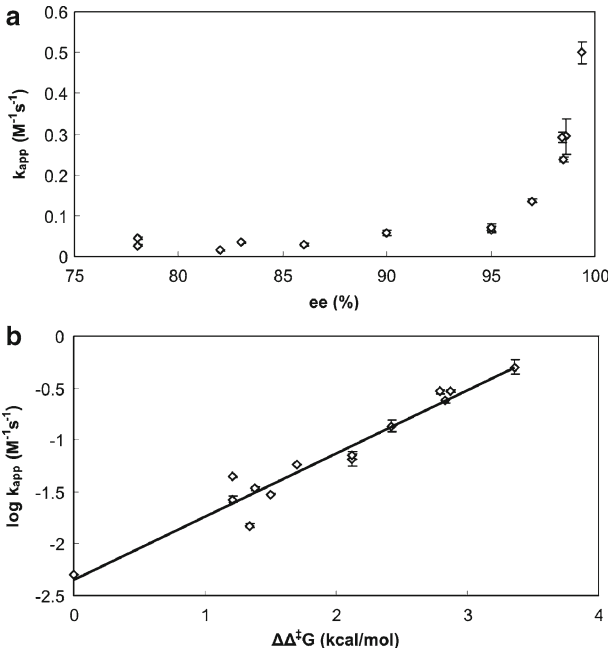


Figure 2 (**a**) Relationship between k_{app} and *ee* obtained with different double-stranded oligonucleotides and st-DNA. (**b**) Relationship between log k_{app} and $\Delta\Delta G^{\ddagger}$. Reprinted with permission from [43]; copyright 2008 American Chemical Society.

A linear relation was found between the *ee* (expressed in $\Delta\Delta G^{\ddagger}$) and the *log* of the apparent 2nd order rate constant (k_{app}).

A combination of these results explains why DNA-based asymmetric catalysis with Cu-dmbipy/st-DNA works, even though the catalyst is actually a heterogeneous mixture of copper complexes that all are located in a different DNA sequence and, hence, provide a different microenvironment for the reaction: the catalysts that are bound to the DNA sequences that give the highest *ee* dominate the overall outcome of the reaction because they also accelerate the reaction the most.

Recently, Moses et al. have reported on alternative DNA structures as scaffold for the Diels-Alder reaction. Using parallel or antiparallel G-quadruplex structures, moderate to good enantioselectivities were obtained [44].

Analogous to the observations made for the Diels-Alder reaction, the Cu-dmbipy/st-DNA catalyzed Friedel-Crafts reactions also show an increase of the apparent rate constant (k_{app}) of up to 30-fold in the presence of the DNA [33]. In this case, the enantioselectivity is also dependent on the DNA sequences following a similar pattern of sequence dependence found for the Cu-dmbipy/DNA catalyzed Diels-Alder reaction.

Also the reaction rate of the Cu-dmbipy/st-DNA catalyzed Michael addition was affected by the presence of DNA. With dimethyl malonate as the Michael donor, a modest 3- to 6-fold rate acceleration was observed. Most likely, the observed rate acceleration is lower than in the Diels-Alder and Friedel-Crafts reaction because of electrostatic repulsion: the Michael addition requires the approach of the negatively charged enolate of dimethyl malonate to the Cu(II)-bound enone, that is located in the highly negatively charged DNA. In case of nitromethane as the Michael donor even a deceleration of the reaction in the presence of DNA was observed.

Finally, also the enantioselectivities in the *syn* hydration of α,β-unsaturated ketones, catalyzed by first generation catalysts, i.e., Cu-**L3**/DNA, proved to be sequence-dependent. Interestingly, the sequence dependence in the hydration reaction proved to be very different from that observed in the Diels-Alder reaction with the same catalyst [36]. The highest enantioselectivites in the hydration reaction were obtained with DNA sequences that containing central AT base pairs. In particular, the use of the self-complementary oligonucleotides $d(CAAAAATTTTTG)_2$ and $d(GCGCTATAGCGC)_2$ provided the highest *ee* values, i.e., up to 82% in D_2O.

5.3 Origin of Enantioselectivity

A complete stereochemical model that explains the transfer of the chirality of the DNA to the catalyzed reaction is currently not available, since the DNA binding mode of Cu-dmbipy and, hence, the catalyst structure, is still not known. However, some interesting observations can be made with regard to the observed stereochemical course of the catalyzed reactions. The absolute configurations of the benchmark products from the DNA-based catalytic C-C bond-forming reactions have been determined either directly, by comparison of the optical rotations to those reported in the literature, or via conversion to a known analogue by displacement of the imidazole moiety with a methoxy group. From the comparison of the absolute configurations it could be

concluded that the diene or nucleophile in these different reactions, always approach the enone from the same π face, that is, the *Si* face in case of R = aryl, and the *Re* face in case of R = alkyl (Figure 3). Therefore, it is likely that the enantiodiscrimination in these reactions is the result of the same mechanism. This could suggest that one face of the enone moiety is shielded by the DNA, resulting in preferred attack from the other face. However, this disagrees with the observed increase in the reaction rate, which would be unlikely in the case of shielding. An alternative hypothesis is that the DNA scaffold actively directs the incoming nucleophile or diene to one π face of the Cu(II)-bound enone, analogous to mechanisms commonly found with enzymes.

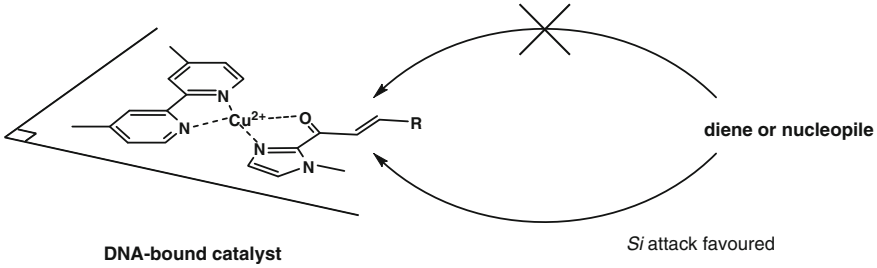


Figure 3 Stereochemistry of the approach of the diene or nucleophile.

Attack from the same π face was observed in the Diels-Alder reaction catalyzed by the first generation catalyst Cu-**L3**/st-DNA. A different behavior was observed in the conjugate addition of water to α,β-unsaturated ketones with the same catalyst [36]: the water nucleophile attacks from the *Re* face of the enone. It is proposed that the induction of enantioselectivity by the DNA in this case might be related to the hydrogen bonding capabilities of the nucleobases in the groove; a spine of hydration is formed in the groove, with highly localized water molecules [45], which could assist and direct the attack of the water nucleophile on the enone.

6 Lessons from DNA-Based Catalysis

Despite being introduced only five years ago DNA-based asymmetric catalysis has already demonstrated its potential in a variety of important catalytic reactions. Based on the results reported to date, some general lessons can be learned.

6.1 Comparison of the Different Approaches

Two different approaches for binding a catalytically active transition metal complex have been developed, involving either supramolecular or covalent anchoring. Covalent anchoring in principle allows the greatest level of control over the catalyst structure

and thus allows the most precise design of the catalyst. However, a significant drawback is the laborious and time-consuming optimization of the catalyst, since it involves non-trivial multistep synthesis. This is reflected in the fact that, to date, only very few examples of successful DNA-based catalysts that involve covalent anchoring of the catalytically active metal complex have been reported.

In contrast, the supramolecular approach has proven quite successful to date, even though the catalyst itself is structurally not well-defined. To a large part this is due to the ease of assembly of the DNA-based catalyst: it just involves mixing of the metal complex with the DNA. As a result, the optimization process is straightforward as has been demonstrated for several important reaction classes. Moreover, this approach allows the use of natural DNA, such as salmon testes DNA, which is inexpensive, as the scaffold. This makes this approach to DNA-based asymmetric catalysis also attractive for applications in organic synthesis. Indeed, multiple catalytic reactions have been performed on gram scale, giving rise to good yields and high enantioselectivities of the isolated products. The fact that the aqueous solutions containing the catalyst are readily recycled further increases the appeal of this concept.

6.2 Implications for Catalysis

The DNA-based catalysis concept has two unique characteristics that distinguish it from the conventional approach to homogeneous catalysis. First of all, whereas conventionally a transition metal catalyst is designed as a well defined and homogeneous species, "the DNA-based catalyst" is actually a heterogeneous mixture of catalysts that all reside in a different microenvironment. Yet, excellent *ee*'s can be obtained in multiple reaction classes. This is attributed to the fact that the fraction of complexes that are in the optimal environment to achieve high enantioselectivities dominate the reaction because they cause the largest rate accelerations. This relates to a second important characteristic: whereas often enantioselectivity is achieved at the expense of catalytic activity, here they go hand in hand. This is most likely due to the fact the DNA-based catalyst does not rely on shielding of one prochiral face of the substrate, which is often the basis under design of ligands for transition metal catalysis. Instead, it is proposed that the DNA actively participates in the reaction, e.g., by directing reagents to the preferred face of the substrate, making use of the so-called second coordination sphere interactions presented by the DNA scaffold [46,47]. As such, the DNA-based catalyst is a true hybrid catalyst, combining some of the attractive features of synthetic and enzymatic catalysis.

7 Concluding Remarks and Future Directions

This novel catalytic concept successfully combines the well-known catalytic activity of the transition metal ions with the unique chirality of the DNA. This has resulted in novel catalysts that not only give rise to high enantioselectivity in a variety of

reaction classes, but also display some very special mechanistic characteristics that can be related to the second coordination sphere interactions provided by the DNA. This has even resulted in the discovery of a new catalytic asymmetric reaction.

Future research in this field will focus on two aspects. First of all, the catalytic scope of the DNA-based catalysis will be expanded. For example, the reactions reported to date almost exclusively involve Lewis acid catalysis. Especially DNA-based organometallic reactions will be highly desirable, since this reaction class is very important in organic synthesis. Moreover, the recent catalytic enantioselective *syn* hydration reaction is an attractive starting point for the further development of novel catalytic chemistry. It is envisioned that especially in reactions such as these, involving small and "difficult" nucleophiles such as water, the DNA-based catalysis concept will reveal its full potential.

The second important point is to achieve an understanding of these second coordination sphere interactions provided by the DNA. This will rely on extensive mechanistic and spectroscopic studies, which are complicated by the heterogeneous nature of the DNA-based catalytic system. However, this gives rise to valuable new insights that will contribute to our knowledge about catalyst design.

In conclusion, although the field of DNA-based asymmetric catalysis is still in its infancy, already extraordinary advances have been made. The continuous progress in this research field will contribute to the development of new enantioselective catalytic reactions of interest for synthetic applications.

Abbreviations

A	adenine
AT	adenine-thymine
BF_4^-	tetrafluoroborate
bipy	2,2'-bipyridine
Bu	butyl
dmbipy	4,4'-dimethyl-2,2'-bipyridine
dppz	dipyrido[3,2-a:2',3'-c]phenazine
dpq	pyrazino[2,3-*f*][1,10]phenanthroline
$\Delta\Delta G^{\ddagger}$	increase of Gibbs free energy of the transition state
ee	enantiomeric excess
G	guanine
GC	guanine-cytosine
h	hour
K_a	equilibrium constant for binding of the substrate to the metal ion
K_b	equilibrium constant for binding of the metal complex to DNA
K_d	dissociation constant
k_{app}	apparent rate constant
k_{cat}	rate constant for the catalyzed reaction
L	ligand

Me	methyl
MES	2-(*N*-morpholino)ethanesulfonic acid
MOPS	3-(*N*-morpholino)propanesulfonic acid
Nu	nucleophile
ON	oligonucleotide
OTf	triflate
Ph	phenyl
phen	1,10-phenanthroline
s	second
st-DNA	salmon testes deoxyribonucleic acid
T	template
t	tertiary
THF	tetrahydrofuran

References

1. J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737–738.
2. N. C. Seeman, *Angew. Chem. Int. Ed*. **1998**, *37*, 3220–3238.
3. P. W. K. Rothemund, *Nature* **2006**, *440*, 297–302.
4. J. J. Storhoff, C. A. Mirkin, *Chem. Rev.*, **1999**, *99*, 1849–1862.
5. X. Li, D. R. Liu, *Angew. Chem. Int. Ed*. **2004**, *43*, 4848–4870.
6. S. K. Silverman, *Angew. Chem. Int. Ed*. **2010**, *49*, 7180–7201.
7. A. J. Boersma, R. P. Megens, B. L. Feringa, G. Roelfes, *Chem. Soc. Rev*. **2010**, *39*, 2083–2092.
8. *Organic Reactions in Water: Principles, Strategies and Applications*, Ed U. M. Lindström, 1st ed., Wiley-Blackwell, Oxford, 2007, pp. 1–424.
9. U. M. Lindström, *Chem. Rev*. **2002**, *102*, 2751–2772.
10. J. Liu, Z. Cao, Y. Lu, *Chem. Rev*. **2009**, *109*, 1948–1958.
11. E. W. Dijk, B. L. Feringa, G. Roelfes, in *Topics in Organometallic Chemistry*, Vol. 25, Ed. T. R. Ward, Springer, Berlin, 2009, pp. 1–24.
12. G. Roelfes, B. L. Feringa, *Angew. Chem. Int. Ed*. **2005**, *44*, 3230–3232.
13. D. C. Rideout, R. Breslow, *J. Am. Chem. Soc*. **1980**, *102*, 7816–7817.
14. S. Otto, J. B. F. N. Engberts, *J. Am. Chem. Soc*. **1999**, *121*, 6798–6806.
15. S. Kobayashi and K. Manabe, *Acc. Chem. Res*. **2002**, *35*, 209–217.
16. S. Otto, G. Boccaletti, and J. B. F. N. Engberts, *J. Am. Chem. Soc*. **1998**, *120*, 4238–4239.
17. M. Pitié, G. Pratviel, *Chem. Rev*. **2010**, *110*, 1018–1059.
18. A. C. Braisted, P. G. Schultz, *J. Am. Chem. Soc*. **1990**, 112, 7430–7431.
19. D. Hilvert, K. W. Hill, K. D. Nared, M. T. M. Auditor, *J. Am. Chem. Soc*. **1989**, *111*, 9261–9262.
20. M. Helm, M. Petermeier, B. X. Ge, R. Fiammengo, A. Jäschke, *J. Am. Chem. Soc*. **2005**, *127*, 10492–10493.
21. K. Ishihara, M. Fushimi, *Org. Lett*. **2006**, *8*, 1921–1924.
22. G. Roelfes, A. J. Boersma, B. L. Feringa, *Chem. Commun*. **2006**, 635–637.
23. A. J. Boersma, B. L. Feringa, G. Roelfes, *Org. Lett*. **2007**, *9*, 3647–3650.
24. D. A. Evans, K. R. Fandrick, H. J. Song, *J. Am. Chem. Soc*. **2005**, *127*, 8942–8943.
25. M. C. Myers, A. R. Bharadwaj, B. C. Milgram, K. A. Scheidt, *J. Am. Chem. Soc*. **2005**, *127*, 14675–14680.
26. D. A. Evans, K. R. Fandrick, *Org. Lett*. **2006**, *8*, 2249–2252.

27. D. A. Evans, H. J. Song, K. R. Fandrick, *Org. Lett*. **2006**, *8*, 3351–3354.
28. D. A. Evans, K. R. Fandrick, H.–J. Song, K. A. Scheidt, R. Xu, *J. Am. Chem. Soc*. **2007**, *129*, 10029–10041.
29. U. Jakobsen, K. Rohr, S. Vogel, *Nucleosides, Nucleotides, and Nucleic Acids* **2007**, *26*, 1419–1422.
30. N. Sancho Oltra, G. Roelfes, *Chem. Commun*. **2008**, 6039–6041.
31. D. Coquière, B. L. Feringa, G. Roelfes, *Angew. Chem. Int. Ed*. **2007**, *46*, 9308–9311.
32. R. P. Megens, G. Roelfes, *Org. Biomol. Chem*. **2010**, *8*, 1387–1393.
33. A. J. Boersma, B. L. Feringa, G. Roelfes, *Angew. Chem. Int. Ed*. **2009**, *48*, 3346–3348.
34. N. Shibata, H. Yasui, S. Nakamura, T. Toru, *Synlett* **2007**, 1153–1157.
35. E. W. Dijk, B. L. Feringa, G. Roelfes, *Tetrahedron: Asymmetry* **2008**, *19*, 2374–2377.
36. A. J. Boersma, D. Coquière, D. Geerdink, F. Rosati, B. L. Feringa, G. Roelfes, *Nature Chem*. **2010**, *2*, 991–995.
37. L. Ropartz, N. J. Meeuwenoord, G. A. van der Marel, P. W. N. M. van Leeuwen, A. M. Z. Slawin, P. C. J. Kamer, *Chem. Commun*. **2007**, 1556–1558.
38. M. Caprioara, R. Fiammengo, M. Engeser, A. Jäschke, *Chem.–Eur. J.* **2007**, *13*, 2089–2095.
39. M. Nuzzolo, A. Grabulosa, A. M. Z. Slawin, N. J. Meeuwenoord, G. A. van der Marel, P. C. J. Kämer, *Eur. J. Org. Chem.* **2010**, 3229–3236.
40. P. Fournier, R. Fiammengo, A. Jäschke, *Angew. Chem. Int. Ed*. **2009**, *48*, 4426–4429.
41. F. Rosati, A. J. Boersma, J. E. Klijn, A. Meetsma, B. L. Feringa, G. Roelfes, *Chem.–Eur. J*. **2009**, *15*, 9596–9605.
42. S. Otto, F. Bertoncin, J. B. F. N. Engberts, *J. Am. Chem. Soc*. **1996**, *118*, 7702–7707.
43. A. J. Boersma, J. E. Klijn, B. L. Feringa, G. Roelfes, *J. Am. Chem. Soc*. **2008**, *130*, 11783–11790.
44. S. Roe, D. J. Ritson, T. Garner, M. Searle, J. E. Moses, *Chem. Commun*. **2010**, *46*, 4309–4311.
45. H. M. Berman, B. Schneider, in *Oxford Handbook of Nucleic Acid Structure*, Ed S. Neidle, Oxford University Press, Oxford, 1999, p. 295.
46. F. Rosati, G. Roelfes, *ChemCatChem* **2010**, *2*, 916–927.
47. C. M. Thomas, T. R. Ward, *Chem. Soc. Rev*. **2005**, *34*, 337–346.

Chapter 10
Alternative DNA Base Pairing through Metal Coordination

Guido H. Clever and Mitsuhiko Shionoya

Contents

Abstract Base-pairing in the naturally occurring DNA and RNA oligonucleotide duplexes is based on π-stacking, hydrogen bonding, and shape complementarity between the nucleobases adenine, thymine, guanine, and cytosine as well as on the hydrophobic-hydrophilic balance in aqueous media. This complex system of multiple supramolecular interactions is the product of a long-term evolutionary process and thus highly optimized to serve its biological functions such as information

G.H. Clever (✉)
Institute for Inorganic Chemistry, Georg-August University Göttingen, Tammannstr. 4, D-37077 Göttingen, Germany
e-mail: gclever@gwdg.de

M. Shionoya
Department of Chemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
e-mail: shionoya@chem.s.u-tokyo.ac.jp

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_10,

storage and processing. After the successful implementation of automated DNA synthesis, chemists have begun to introduce artificial modifications inside the core of the DNA double helix in order to study various aspects of base pairing, generate new base pairs orthogonal to the natural ones, and equip the biopolymer with entirely new functions. The idea to replace the hydrogen bonding interactions with metal coordination between ligand-like nucleosides and suitable transition metal ions culminated in the development of a plethora of artificial base-pairing systems termed "metal base-pairs" which were shown to strongly enhance the DNA duplex stability. Furthermore, they show great potential for the use of DNA as a molecular wire in nanoscale electronic architectures. Although single electrons have proven to be transmitted by natural DNA over a distance of several base pairs, the high ohmic resistance of unmodified oligonucleotides was identified as a serious obstacle. By exchanging some or all of the Watson-Crick base pairs in DNA with metal complexes, this problem may be solved. In the future, these research efforts are supposed to lead to DNA-like materials with superior conductivity for nano-electronic applications. Other fields of potential application such as DNA-based supramolecular architecture and catalysis may be strongly influenced by these developments as well. This text is meant to illustrate the basic concepts of metal-base pairing and give an outline over recent developments in this field.

Keywords coordination chemistry • DNA • metal-base pairing • nanotechnology

1 Introduction

1.1 Use of DNA as Intelligent Material

The natural role of DNA is the conservation and propagation of genetic information in each organism. As the result of million years of evolution, it features a maximal density of functionality embedded in its double-helical structure [1].

The interior of the duplex is formed by a parallel stack of pairwise-bonded aromatic nucleobases. Hydrogen bonding, π-stacking, and shape complementarity all play an important role as factors governing the strength and structural integrity of the double strand. Depending on the base-pair sequence and environmental factors such as humidity, salt content, and binding partners, DNA can adopt a number of secondary structures of which the right-handed B-type helix is the predominant form. The latter is characterized by a base-pair distance of 3.4 Å and a helical pitch of 36 degrees per base pair resulting in a complete helix turn approximately every ten base pairs. The structure features two grooves, called the major and minor groove, through which the nucleobases can be accessed and recognized by DNA-binding proteins such as transcription factors in a sequence-dependent manner. The outer surface of the double strand is lined by the negatively-charged sugar-phosphate backbone providing superior water solubility and shielding for the internal hydrophobic π-stack. DNA is a remarkably stable molecule that allows handling in a

relatively wide pH range and does not undergo substantial decomposition when heated to a temperature of 100°C for a limited time.

Short DNA strands are routinely synthesized by an automated solid-phase process with full control over the desired sequence. The fully synthetic nature of this process involving no enzymatic reagents allows the introduction of modified building blocks such as artificial nucleosides, unnatural backbones or end-groups regardless of the biological function of DNA. However, the choice of a suitable protecting group strategy complying with the solid-phase synthesis and the subsequent transfer of the products to the aqueous media is required.

Since automated, solid-phase oligonucleotide synthesis is only feasible for the production of relatively short oligonucleotides (in practice up to lengths of about 100 nucleotides), longer DNA strands of hundreds or thousands of base pairs containing artificial building blocks is achievable through the combination of solid-phase DNA synthesis with subsequent enzymatic techniques. Short modified DNA strands can be stitched together and ligated to form longer constructs using proteins called ligases. A new development of recent years is the use of artificial nucleoside triphosphates in the automated polymerase chain reaction (PCR) employing specifically designed or evolved polymerases that are able to process non-natural dNTPs [2].

In the quickly developing field of bottom-up nanotechnology, DNA has become an important building block owing to its superior properties [3]. Several lines of research may benefit from the use of modified DNA structures. In particular, the molecular electronics approach is currently seen as a promising new technology because the ongoing miniaturization of electronic circuits is facing a limit in structure size using the classic silicon-based photolithographic processes [4]. Furthermore, electrofunctional organic building blocks such as light-harvesting and charge separating units of future photovoltaic devices might be wired up with such molecular cables.

A great variety of synthetic DNA modifications have been realized in recent years such as fluorescent dyes or surface anchors covalently bound to the 3' or 5' ends, outside pointing functionalities protruding from the sides of the double helix and artificial backbone structures dramatically changing the solubility of the DNA constructs (Figure 1) [5]. Most interesting, however, turned out to be the modifications that were introduced right into the middle of the DNA double helix, thus replacing the natural base pairs. Such artificial nucleobases were developed for a range of reasons such as studying the hydrogen bonding and stacking forces that hold together the double helix [6]. Other artificial bases were incorporated into DNA to create new sets of base pairs that are orthogonal to the natural Watson-Crick system and therefore can be seen as an extension of the genetic code [7]. Owing to their wide application in medicinal diagnostics, many artificial bases were developed for the analytical recognition of single nucleotide polymorphisms (SNPs) [8].

Before we restrict our focus on the internal modification of DNA with metal ions coordinated to ligand-like nucleobases, another approach to DNA nanotechnology based on treating DNA as a programmable building material is briefly introduced. The basis for this field, often called "DNA origami", is that the sequence-dependent hybridization of DNA single strands to double strands and the possibility to generate junctions of several duplexes was recognized as a method for the creation of complex

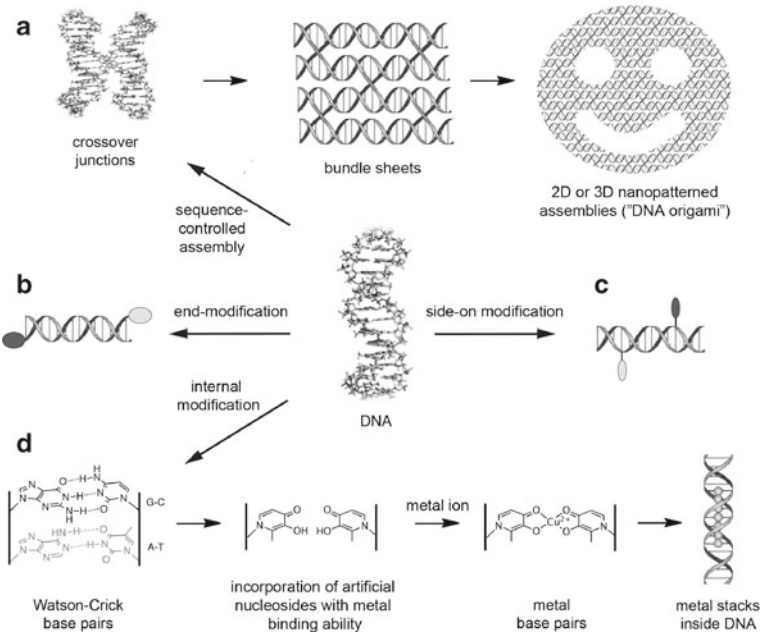


Figure 1 Developments in the area of DNA nanotechnology: (**a**) The "DNA origami" approach makes use of the sequence-specific self-assembly features of DNA and the rational design of inter-duplex connections via cross-links based on the naturally occurring Holliday junction, (**b**) the covalent attachment of functionalities to the 3'- or the 5'-end(s) of DNA is achievable with contemporary standard methods of automated DNA synthesis, (**c**) several strategies have been developed for the covalent attachment of functionalities to internal nucleosides protruding out of the DNA duplex, (**d**) the incorporation of modifications inside the core of the DNA double helix requires the synthesis of artificial nucleobase phosphoramidite building blocks but allows to take direct influence on the process of double strand formation.

self-assembling 2- and 3-dimensional molecular architectures. This branch of DNA nanotechnology has been pioneered by Seeman [9], followed by reports of complex 2D structures by Winfree et al. [10], Rothemund [11], Yan [12], and Somoza [13]. By going into the 3rd dimension, the topic has recently been brought to another level of sophistication by Shih et al. [14], Gothelf, Kjems, and coworkers [15] (Figure 1a).

1.2 Can DNA Be Used as a Molecular Wire?

While exploring the suitability of DNA as a molecular wire, it turned out that the electron conducting abilities of natural DNA are not sufficient for the use of unmodified double strands in molecular electronic circuits [68]. The literature values for the conductivity actually spread over a great range from 1 to 1×10^7 MΩ. This inconsistency was attributed to the differences in experimental techniques such as the method for contacting the DNA (AFM-based, silicon break junctions, carbon nanotube contacts,

etc.), environmental factors (humidity, salt content, buffer), and surface effects [16]. It can also be assumed that the specific sequences of the various DNA strands under investigation play a role for the measured conductivity values.

Astonishingly, the conductance of single excess electrons and positive charges through short patches of DNA is part of biological processes involved in DNA repair [17]. The lossless conductance of many charges, however, over distances of several hundred base pairs does not seem to be a function of DNA that was worth to be optimized by natural evolution. Besides a low ohmic resistance, the transport of a steady current of electrical charges along a molecular wire would also require a high robustness of the oligonucleotide for the successful implementation into a bio-artificial hybrid electronic device [18].

1.3 Metal Coordination of DNA Containing Only Natural Nucleotides

The idea of doping the interior of oligonucleotide duplexes with metal ions in order to yield materials with enhanced conductivity or other interesting electronic effects is not so new [19]. Early experiments to introduce metal ions into the core of a double helix aiming at enhanced electrical conductivity made use of unmodified DNA strands at high pH and metal ions such as Zn(II), Co(II), and Ni(II) [20]. Based on this idea, Lee et al. proposed a model system for a field-effect transistor based on M-DNA [21]. The structure and conductive properties of this so-called M-DNA were, however, controversially discussed and despite some proposed structural models, the exact position of the metal ions inside or around the DNA strands remains rather unclear [19].

More reliable evidence for the binding of metal ions inside DNA duplexes exclusively consisting of natural base pairs was obtained for thymine-rich sequences upon addition of Hg(II) salts. As early as 1952 the effect of Hg(II) on DNA samples was studied. In 1963, the formation and structure of a **T**-Hg-**T** metal-base pair, in which each thymine base is deprotonated at the N(3) position, was proposed by Katz (**5** in Figure 3, $X = CH_3$; see further below) [22]. Later, this observation was picked up by Buncel et al. and Kuklenyik and Marzilli, who supported the picture of Hg(II) binding to pairs of thymine bases by NMR experiments [23]. Recently, Ono and Togashi picked up this principle again and contributed further studies on the **T**-Hg-**T** metal-base pairing [24].

A well documented method for elucidating the interstrand binding of metal ions to oppositely arranged nucleosides is the observation of a duplex-stabilizing effect in terms of an increase in the melting temperature T_M (for a more detailed explanation, see below). Whereas **TT** mismatches are known to strongly destabilize the DNA double helix, the addition of Hg(II) to sequences containing such **TT** mismatches leads to a significant rise of the thermal duplex stability as observed in the DNA melting profiles. Likewise, Müller, Sigel, and coworkers found that the analogous **U**-Hg-**U** metal-base pairs (**5** in Figure 3, X = H) can be formed in an RNA context [25]. Recently, Ono et al. reported the **C**-Ag-**C** metal-base pair **6** which forms upon addition of Ag(I) to double strands containing **CC** mismatches [26]. Megger and Müller subsequently found evidence for the formation of parallel-stranded helices containing

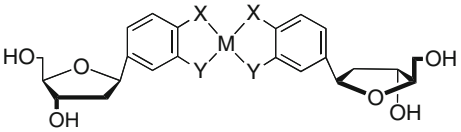


Figure 2 The phenylenediamine, aminophenol, and catechol metal-base pairs. Reprinted from [19a] with permission from Wiley-VCH; copyright 2007.

eight consecutive **C**-Ag-**C** base pairs by CD and UV measurements [27]. The fact that **TT** mismatches bind Hg(II) but no Ag(I) and **CC** mismatches show an opposite behavior was exploited recently by Willner et al. to create logic AND as well as OR gates based on metal-binding oligonucleotides which are attached to quantum dots [28].

Noteworthy is furthermore the observation of an Au(III) ion bound inside a **GC** base pair (with guanosine deprotonated at the N(1) position) that was found in a crystallographic study aimed at screening the interaction of an RNA duplex with various metal ions [29]. Some data on the potential for electron conductance through metal-containing oligonucleotides based on natural nucleobases have already been gathered in a small number of experiments. Joseph and Schuster studied the effect of a **T**-Hg-**T** base pair on the long-distance radical cation ("hole") hopping properties but found no significant effect of this metal-base pair on the charge transport [30]. Voityuk however, proposes that a stack of **T**-Hg-**T** base pairs might be beneficial for excess electron transfer [31]. Electrical transport through cation-stabilized G-quadruplex DNA was recently reported by Erbe et al. [32].

In an alternative approach, DNA strands bound to surfaces were used as templates for the spatially controlled deposition of metal ions, and subsequent reductive formation of nanoscale wires by thermal soldering or photographic development [33]. Concerning the various roles of metal ions in naturally occurring nucleic acids, the reader may refer to [34] (see also Chapter12 of this book).

Since binding of metal ions to the natural nucleobases was found to be limited to only a small number of metals and Watson-Crick base pairing might interfere with metal binding, a new approach of metal-base pairing starting from designed, artificial nucleobases with metal binding potential was developed.

In these so-called "metal-base pairs", the natural nucleobase attached to the sugar is replaced by a ligand capable of coordinating a (transition) metal ion in a linear or square-planar fashion [19]. Such a metal-base pair is ideally represented by a flat coordination complex that is integrated into DNA as part of the base pair stack without distorting the shape of the duplex.

1.4 Artificial Metal-Base Pairs

The field of metal-base pairing based on artificial nucleosides was pioneered by Tanaka and Shionoya who studied the coordination chemistry of monomeric ligand-modified nucleosides (**1** – **4**, Figure 2) in aqueous solutions. Although only serving

as model compounds, this first work established the foundation of the concept of metal-base pairing [35].

However, it proved difficult to incorporate the oxidation-sensitive *ortho*-phenylenediamine, aminophenol, and catechol compounds **1** - **4** into oligonucleotides. Consequently, metal-base pairing inside DNA double strands could not yet be realized in these early experiments. Shortly following this first work, the groups of Schultz and Shionoya both succeeded in using the known pyridine nucleoside **P** for metal-mediated base pairing inside the DNA double helix. The Shionoya group showed that Ag(I) ions can be coordinated between two (or even three) pyridine nucleosides inside a DNA duplex resulting in the formation of not only double- but also triple-stranded systems (**8** in Figure 3, see also Section 2.1) [36]. Schultz et al. developed an unsymmetrical metal-base pair consisting of one pyridine base **P** and one pyridine ligand equipped with two extra coordinating groups resulting in a family of metal-base pairs denoted as **10** in Figure 3 [37]. Two of the **Dipic**-Cu-**P** base pairs (**10** with X = O) were successfully incorporated into the dodecamer sequence [d(5'-CGCG**Dipic**AT**P**CGCG)$_2$]. The structure was elucidated by X-ray crystallography showing the Jahn-Teller distorted octahedral coordination of the Cu(II) ions between the **P** and **Dipic** ligands forming the metal-base pair and two oxygen atoms of the neighboring nucleotides [38].

Since then, a variety of further metal-base pairs was developed (Figure 3). Only selected examples will be discussed here and the reader may refer to preceding review articles [19,39,40] and the respective original publications mentioned in the caption to Figure 3.

A schematic overview of synthesis and incorporation of ligand-modified nucleosides is given in Figure 4 [46]. Following automated synthesis, deprotection and purification, double strands are obtained by mixing equimolar amounts of matching single strands in an aqueous solution containing an electrolyte such as NaCl of high ionic strength and a buffer to adjust the required pH. Experience has shown that the choice of an appropriate buffer system is crucial for successful formation of the DNA metal-base pairs since a number of buffer components (e.g., tris(hydroxymethyl)-aminomethane, TRIS, and ethylenediamine tetraactetate, EDTA) is known to interfere with metal coordination by forming strong complexes with transition metal ions themselves. The metal-base pairs are finally assembled by adding the required amount of metal ions, which can be monitored by thermal denaturation experiments, UV- or CD-based titrations and mass spectrometry.

In a thermal denaturation experiment, also termed a "melting curve" experiment, the extinction at the absorption maximum of the nucleobases at 260 nm (which is higher for single strands than for the double strand, an effect called hyperchromicity) is plotted against the sample temperature (see Figure 6 below for an example). The transition point of the curve is defined as the melting point T_M of the DNA duplex. It is dependent on the length, sequence, and concentration of the DNA as well as the nature and concentration of the additives (salt, buffer). Typical parameters used are: DNA concentration of 1–10 μM, buffer concentration 10 mM (pH around 7, but may depend on the requirements imposed by the kinds of ligand and metal ion), NaCl concentration 50–150 mM ($NaClO_4$ was successfully employed instead in studies using Ag(I) as metal ion under investigation).

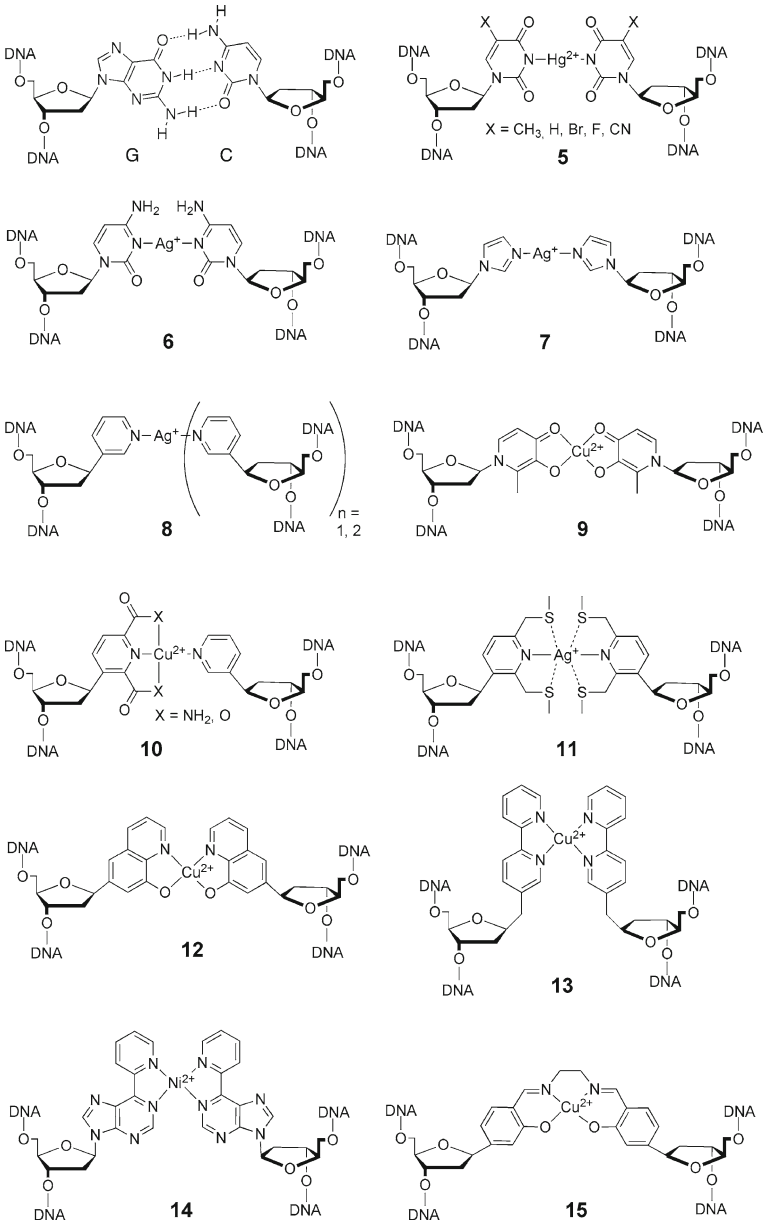


Figure 3 The structure of a Watson-Crick GC base pair compared with a selection of the reported metal-base pairs. Only one kind of metal ion is shown in each case and some derivatives are omitted. Since in most cases no or only limited structural data were reported, the real coordination geometries might differ from the structures drawn here. References not mentioned in the text: **5** (X = Br, F, CN) [41], **11** [42], **12** (also reported with a simplified propanediol backbone) [43], **13** [44], and **14** [45]. Reprinted from [19b] with permission from Elsevier; copyright 2010.

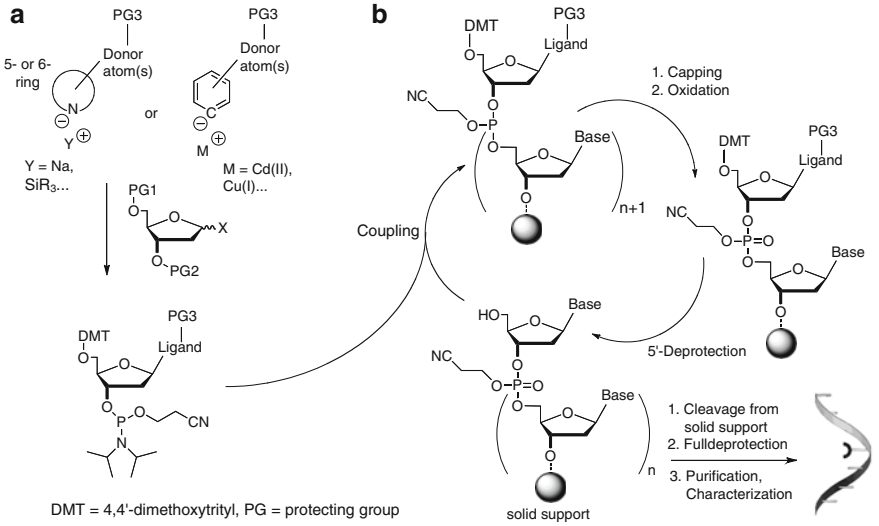


Figure 4 Synthesis of ligand-modified nucleosides and incorporation of artificial nucleosides into oligonucleotides by automated solid-phase synthesis using phosphoramidite chemistry. The key step in the synthesis of artificial nucleosides is the coupling of the nucleobase (or the artificial ligand) with the deoxyribose sugar moiety. The examples in Figure 3 show that either a C-N bond or a C-C bond has to be formed and several methods for either case have been reported in the respective publications. The choice of a suitable protecting group strategy for the ligand which is compatible with DNA synthesis but allows for a smooth deprotection of the artificial nucleobases after DNA synthesis in aqueous solution is mandatory. After DNA synthesis, the single strands are cleaved from the solid support, fully deprotected and dissolved in an aqueous buffer. The oligonucleotides may be purified by HPLC, gel electrophoresis, or other techniques, characterized by mass spectrometry and their concentration is estimated by UV spectroscopy (requiring knowledge of the molar extinction coefficient of the modified nucleosides). Reprinted from [19b] with permission from Elsevier; copyright 2010.

For the ease of interpretation, the stabilizing effect of the used transition metal ions on the DNA strands containing the ligands in oppositely arranged positions can be expressed as the difference ΔT_M between the melting points of the duplex before and after addition of the respective metal ions. When duplex stabilizing effects in terms of ΔT_M are compared for the different metal-base pairing systems described in the literature, care has to be taken because the melting temperature T_M is not a thermodynamic value. Nevertheless, the fact that a T_M value is relatively easy to determine and its descriptive character have made it a key parameter employed for discussing thermal duplex stability.

Deduction of the thermodynamic parameters ΔH and ΔS for describing the base pairing situation upon addition of metal ions requires a thoroughly conducted series of melting point experiments carried out with a number of samples of different DNA concentration and subsequent van't Hoff analysis. Isothermal titration calorimetry (ITC) [47] may be used as a direct method to determine the thermodynamic parameters associated with metal complexation, however, to the best of our knowledge this method has never been used by the respective community.

2 Selected Examples

2.1 The Hydroxypyridone Metal-Base Pair

3-Hydroxy-4-pyridone is a heterocyclic, bidentate chelate ligand capable of forming square-planar 2:1 complexes with metal ions such as Cu(II) and octahedral 3:1 complexes with metal ions such as Fe(III). A hydroxypyridone nucleoside **H** (**9** in Figure 3; carrying a methyl group at the 2-position) was synthesized by attachment to the C1' position of deoxyribose via its nitrogen atom. The synthesis and incorporation into oligonucleotides were reported by Tanaka and Shionoya et al. [48]. Subsequently, addition of Cu(II) to a duplex containing two oppositely arranged hydroxypyridone bases was shown to result in the formation of an **H**-Cu-**H** base pair inside the double helix (Figure 5) [68].

It was found that the hydroxypyridones lead to a destabilization compared to **AT** or **GC** base pairs at the same position in the absence of transition metal ions. Upon addition of Cu(II) ions, however, the duplex underwent a stabilization of $\Delta T_M = 13$ K (in the given sequence context, see Figure 6) which was interpreted as formation of one metal base pair inside the DNA double strand [48]. Addition of the same

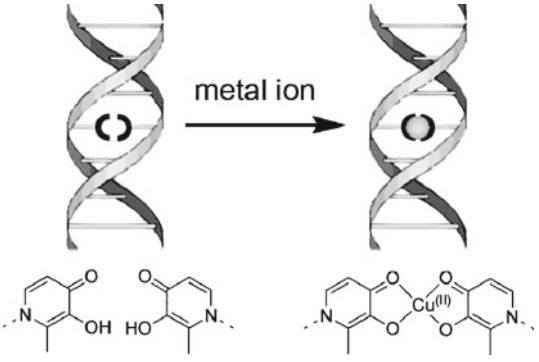


Figure 5 The hydroxypyridone metal base pair **H**-Cu-**H.** Reprinted from [19b] with permission from Elsevier; copyright 2010.

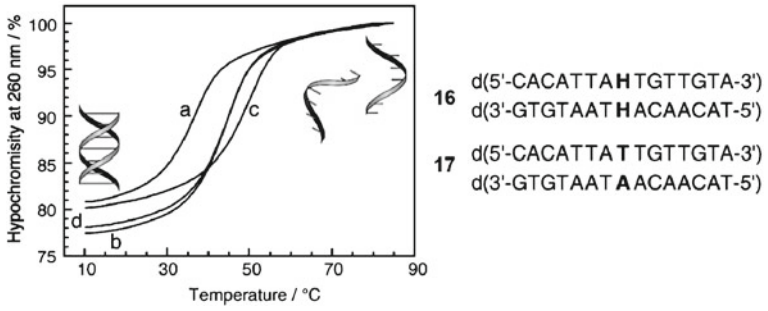


Figure 6 Melting curves of the duplexes **16** (a and c) and **17** (b and d). [**16**] = [**17**] = 2.0 μM in 10 mM sodium phosphate buffer, 50 mM NaCl (pH 7.0). (a) and (b): no Cu^{2+}; (c) and (d): [Cu^{2+}] = 2.0 μM. Reprinted from Ref. [48] with permission from ACS Publications; copyright 2002.

amount of Cu(II) to a similar sequence containing an **AT** base pair instead of the hydroxypyridones did not result in any change of T_M.

The same system was further characterized by UV-based titration studies, CD spectroscopy, and ESI mass spectrometry, all indicating the complexation of one Cu(II) ion inside the double strand. An X-ray structure of the monomeric **H**-Cu-**H** base pair **9** confirmed the square-planar coordination of the Cu(II) ion by two hydroxypyridone ligands via their four oxygen atoms [48]. Since this model complex is not embedded within the surrounding DNA structure, the ligands, however, were found to be coordinated in an *anti* configuration as opposed to the *syn* configuration that was anticipated for the **H**-Cu-**H** metal-base pair inside the DNA double helix by molecular modeling. Recently, the expected *syn* configuration of the hydroxypyridone ligands around the Cu(II) ion was confirmed by a crystal structure containing two **H**-Cu-**H** metal-base pairs **9** inside a modified duplex based on an acyclic three-carbon propylene glycol phosphodiester backbone [49].

In a related study by the Shionoya group aiming at the coordination of soft metal ions such as Pd(II) by ligand-modified nucleobases, two sulfur containing ligands based on hydroxypyridone **H** in which the two coordinating oxygen atoms were exchanged one by one for a sulfur atom were synthesized and tested for their metal coordinating abilities [50]. Although the formation of Pd(II) and Pt(II) metal-base pairs was successfully achieved using the monomeric nucleosides, the incorporation of these sulfur-containing nucleosides into DNA duplexes has not been reported yet. Section 3 will focus on the incorporation of up to five **H**-Cu-**H** metal-base pairs **9** in a way that Cu(II) ions are stacked on top of each other inside the double strands.

Depending on the offered kind of metal ion, the hydroxypyridone ligand is also able to form 3:1 octahedral complexes. The use of Fe(III) was recently shown to result in the formation of triple helical constructs (Figure 7) [51]. Mixing three equivalents of a tetrameric $\mathbf{H}_4$ sequence with four equivalents of Fe(III) ions resulted in the formation of a triple helical complex $(\mathbf{H}_4)_3Fe_4$ after two days at 85°C as shown by UV-based titration studies, CD measurements, and ESI mass spectrometric data.

Noteworthy in the context of triple helix formation is also the earlier work of the Shionoya group using a pyridine ligand **P** inside the duplex [d(3'-T_{10}**P**T_{10}-5')•d(3'-A_{10}**P**A_{10}-5')] which was shown to bind another single strand d(3'-T_{10}**P**T_{10}-5') upon addition of Ag(I) in fashion of a heterotriplex (**P** = pyridine in base pair **8**, Figures 3 and 7b) [36]. Recently, an observation of parallel triplex stabilization by Ag(I) ions was reported by Jyo et al. [52].

2.2 The Salen Metal-Base Pair

The assembly of most metal-base pairs can be interpreted as a two-component process comprising the DNA duplex containing the preorganized ligands as one reaction partner and the metal ion as the second component.

The situation is different for the salen metal-base pair **15** (**S**) introduced by Clever, Colborn, and Carell [53]. In contrast to the hydroxypyridone system, the assembly of the salen metal-base pair inside the DNA double helix requires the addition of

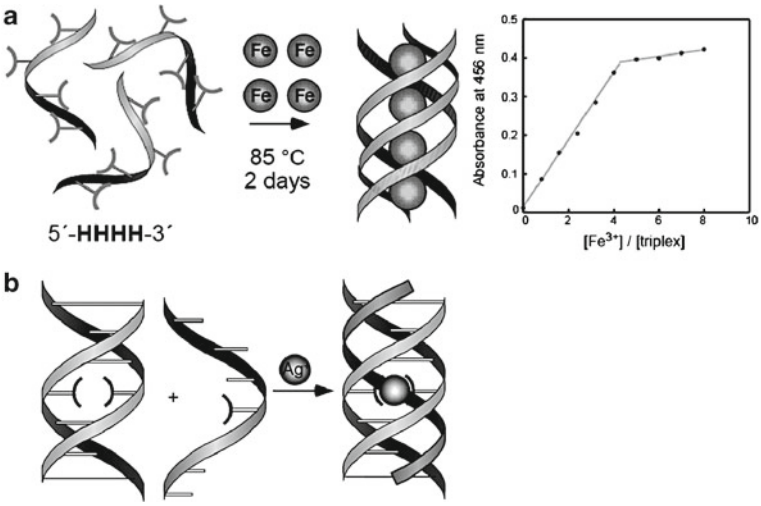


Figure 7 Formation of triple helices. (**a**) From three $\mathbf{H}_4$ strands (consisting of hydroxypridone nucleotides exclusively) and four Fe(III) ions; (**b**) A metal-containing triple helix can also be formed upon addition of Ag(I) ions to a mixture of duplex [d(3'-T_{10}**P**T_{10}-5')•d(3'-A_{10}**P**A_{10}-5')] and single strand d(3'-T_{10}**P**T_{10}-5'). Reprinted from [19b] with permission from Elsevier; copyright 2010.

two components to the duplex containing a pair of salicylic aldehydes at positions facing each other. The first additive is ethylenediamine which reacts with both aldehyde groups of the ligand precursors (salicylic aldehyde) under elimination of two molecules of water to the well known salen ligand (Figure 8). Since the formation of such aromatic imines is reversible in water, an excess of ethylenediamine is added in order to drive the equilibrium to the side of the salen ligand. Subsequently a transition metal ion such as Cu(II), Mn(II) (oxidized to Mn(III) upon complexation), Fe(III), or vanadyl (VO^{2+}) is added and the metal-base pair **15** is formed [53b].

The structure of the metal-base pairing system was elucidated by X-ray analysis of a monomeric Cu(II) complex of the salen ligand that was formed from the free salicylic acid nucleoside and ethylenediamine [53b,69]. The structure confirms the square-planar coordination of the metal ion (going along with a slight propeller twist between the planes of the two aromatic rings). A superposition of the molecular structure of a natural, hydrogen-bonded Watson–Crick base pair with the salen Cu(II)-base pair is depicted in Figure 8b. The excellent structural agreement between the natural **AT** base pair and the salen metal-base pair indicated that the chosen molecular design should result in a very good fit of the salen metal-base pair inside the DNA double helix. This assumption could be supported by further experimental results such as CD measurements of salen metal-containing DNA strands [53a].

A comparative study was conducted comparing the ligand depicted in Figure 8 (glycosidic bond *para* to the carbonyl group) with an isomeric salicylic aldehyde precursor (glycosidic bond *para* to the hydroxyl group) inside the same sequence context [53b]. In full accordance with the results obtained from the X-ray crystallographic investigations, it was found that the former isomer indeed fits much better inside the

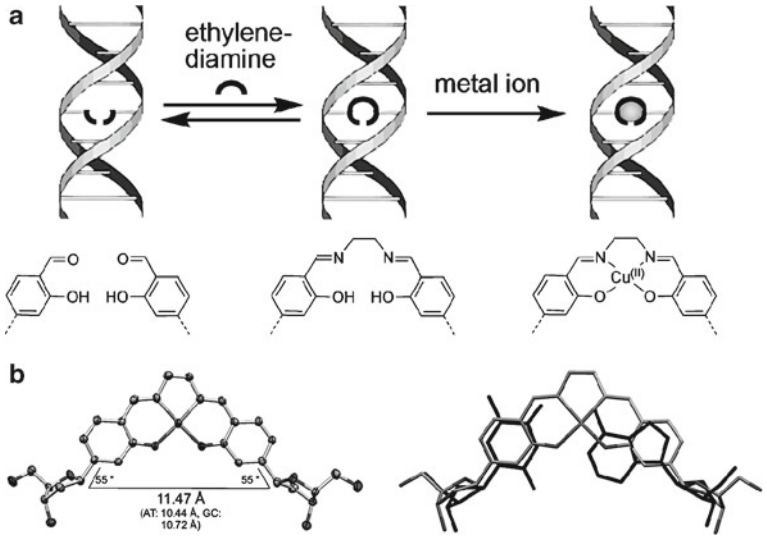


Figure 8 (**a**) Formation of the salen ligand from perfectly preorganized salicylic aldehyde precursors inside the DNA double helix and subsequent coordination of the metal ion giving the salen-metal-base pair; (**b**) X-ray structure of the monomeric salen-Cu(II)-base pair and superposition with a Watson-Crick AT base pair. Reprinted from [19b] with permission from Elsevier; copyright 2010.

DNA duplex compared with the metal-base pair based on the latter isomer. Metal-base pair formation using ethylenediamine and Cu(II) is such a favorable process, however, that even the second isomer is undergoing the reaction inside the DNA, presumably under distortion of the duplex structure.

The mechanistic difference of the metal-base pair formation with the salen system compared to the hydroxypyridone system (and other comparable ligands) could be deduced from the results of thermal denaturation experiments. The same sequence as in the work of Shionoya et al. [48] was used: [d(5'-CACATTA**S**TGTTGTA-3')• d(3'-GTGTAAT**S**ACAACAT-5')]). Whereas the duplex containing the two salicylic aldehydes **S** has a melting temperature T_M of 40°C (about 10 K destabilization compared to AT), the addition of ethylenediamine and Cu(II) resulted in the formation of the salen Cu(II)-base pair (**15** in Figure 3) accompanied by an increase in T_M to 82°C. This very large value for ΔT_M of +42 K is the highest ever reported increase in melting temperature for a metal-base pairing system (Figure 9) [53a].

The reason for this large duplex stabilization value was elucidated by a series of melting point experiments using different combinations of the single components added to the DNA duplex. If only ethylenediamine was added, the melting point T_M was raised by only 5 K, attributed to the formation of the imine cross-link which is reversible in the aqueous environment. If only Cu(II) was added to the duplex, T_M was raised by 15 K, a value comparable to the examples of duplex stabilization reported for the other metal-base pairs. The high stabilization of ΔT_M = +42 K was achieved only in case, both, ethylenediamine and Cu(II) ions, were added to the duplex [53a]. A reasonable explana-

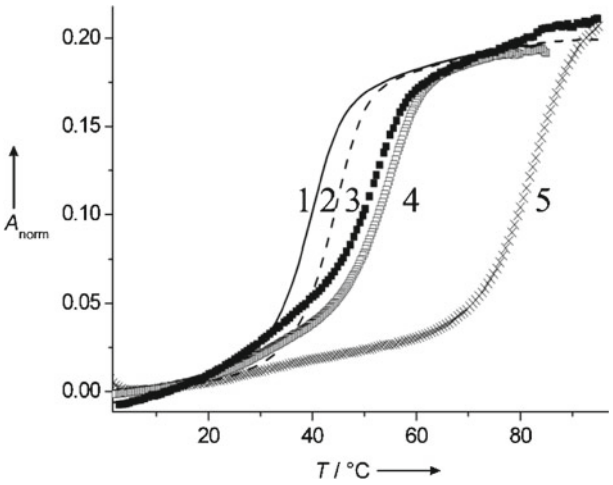


Figure 9 Comparison of the melting curves of the sequence d(5'-CACATTA**S**TGTTGTA-3')· d(3'-GTGTAAT**S**ACAACAT-5') (1) without any additives (39.9°C, solid line); (2) with only ethylenediamine (45.5°C, dashed line); (3) with methylamine and Cu(II) (52.3°C, filled boxes); (4) with only Cu(II) (54.9°C, open boxes); and (5) with ethylenediamine and Cu(II) (82.4°C, crosses) (3 μM DNA, 150 mM NaCl, 10 mM CHES buffer). Reprinted from [19b] with permission from Elsevier; copyright 2010.

tion is that the formation of the salen system from the salicylic aldehydes and ethylenediamine provides an excellent coordination environment for the metal and the binding of the Cu(II) ion to the tetradentate chelate ligand results in a stabilization of the imine bonds towards hydrolytic cleavage. Both the action of the cross-linking agent ethylenediamine and the coordinated metal ion are responsible for the remarkable duplex stabilization because a metal-stabilized cross-link is formed that is of far higher strength than metal coordination alone (Figure 10). Supporting this hypothesis is further the observation that the addition of Cu(II) and methylamine (instead of ethylenediamine) resulted in a stabilization ΔT_M of only 12 K. This value is comparable to the melting temperature obtained for the DNA sample containing Cu(II) alone [53a].

In the case of the Cu(II)-containing system, the thermal de- and renaturing profiles (heating, cooling curves) are superimposable. The measurements of the same sequence containing ethylenediamine and Mn(III), however, reproducibly showed a strong hysteresis between the de- and renaturing profiles.

This behavior was explained with the thermal instability of the salen-Mn(III) complex when exposed to temperatures above T_M for elongated times. In these experiments, the transition in the heating curve can be assigned to the metal-containing and thus higher melting duplex, whereas the cooling curve shows that the duplex re-hybridizes before reincorporation of the metal (expressed by a lower T_M).

After the sample is allowed to spend some time at a temperature below T_M, however, the metal is again fully incorporated giving rise to a denaturing profile superimposable with the preceding denaturing curve. This observation allowed the conclusion that the salen Cu(II)-base pair **15** is more stable than the corresponding salen Mn(III)-base pair even at high temperatures. This fact was supported also

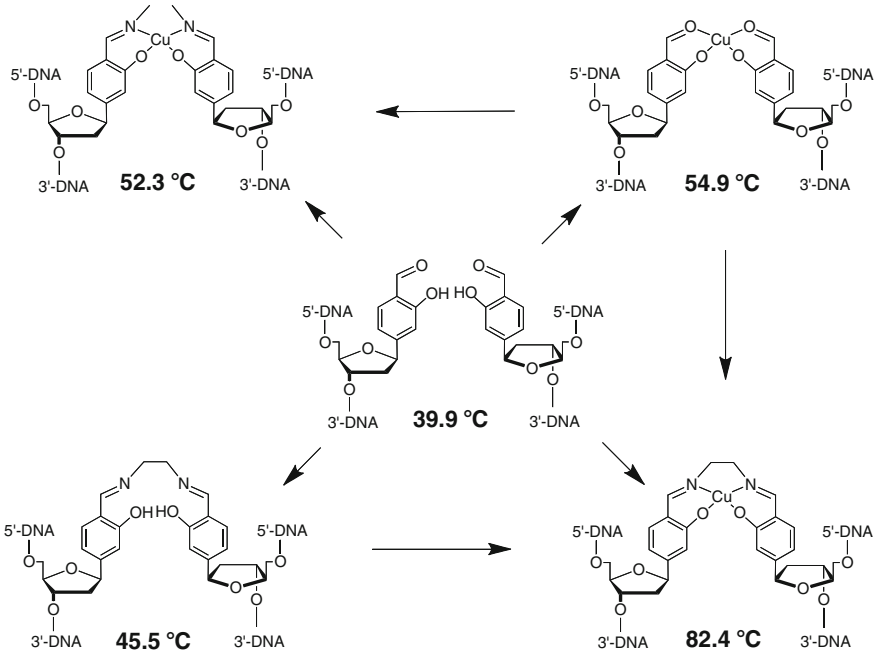


Figure 10 Cooperative assembly of the salen-metal-base pairs in DNA. Reprinted from [19b] with permission from Elsevier; copyright 2010.

by the higher ΔT_M caused by the addition of Cu(II) (+42 K) in comparison to the addition of Mn(III) (+28 K).

Another form of cooperative binding of Cu(II) ions was observed in the case of duplex d(5'-CACATT**SS**TGTT GTA-3')•d(3'-GTGTAA**SS**ACAACAT-5') containing two neighboring salen ligands [53c]. The addition of 0.5 equivalents of Cu(II) results in a 1:1 mixture of metal-free duplexes and duplexes containing two Cu(II) ions as observed by a thermal denaturation study (Figure 11). The binding of the first Cu(II) ion to the duplex was shown to enhance the affinity for binding of the second Cu(II) ion so that all available Cu(II) ions in the solution end up being incorporated pairwise. At any given ratio of Cu(II) ions to DNA duplexes (<2) the reaction mixture thus contains only two species: metal-free duplexes and duplexes containing two metal ions. This model also explains the observation of isosbestic points in the titration curves of all examined double strands containing two or more adjacent metal binding sites such as **HH** or **SS** (see Section 3).

2.3 The Triazole Metal-Base Pair

The use of 5-membered nitrogen-containing heterocycles as artificial nucleobases with metal coordinating abilities such as imidazole, triazole, and tetrazole was systematically studied by Müller and coworkers [40,54]. In their systematic studies,

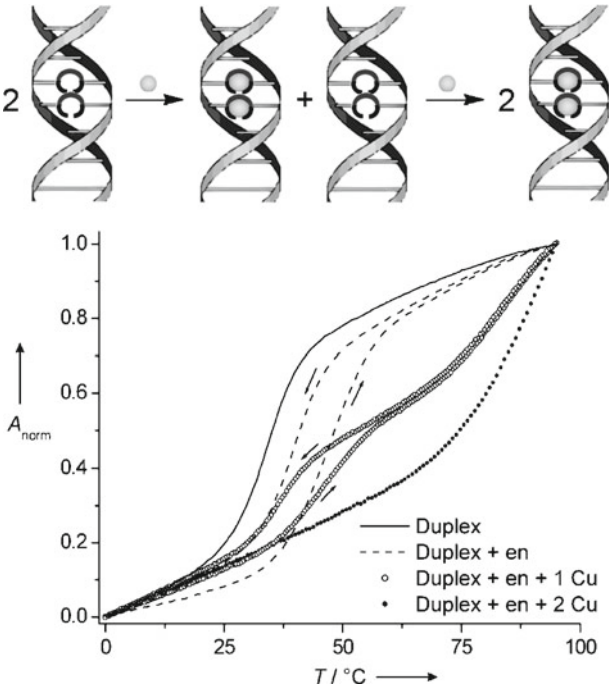


Figure 11 Cooperative uptake of metal ions inside duplexes containing two neighboring salen ligands. Addition of 0.5 equivalents of Cu(II) results in a 1:1 mixture of metal-free duplexes and duplexes containing two Cu(II) ions. The observed hysteresis effect is characteristic for the melting of the duplex containing ethylenediamine but no Cu(II) ions [53c]. Reprinted from [19b] with permission from Elsevier; copyright 2010.

the ability to coordinate metal ions such as Ag(I) and Hg(II) (in terms of stability constant and ligand-to-metal ratio) was examined prior to the incorporation of ligands such as **7** into DNA sequences and the results were taken as indicators for predicting the behavior of the same ligand-metal systems inside the DNA context [55]. In addition, DFT calculation and experimental determination of the pK_a values of the monomeric nucleosides was applied to estimate the proper pH range that is suitable for the formation of metal-base pairs inside the oligonucleotide environment. Out of the examined 5-membered heterocycles, the incorporation of 1,2,4-triazole **Z** into DNA was chosen because of its ability to differentiate between Ag(I) and Hg(II) (Figure 12).

A highly interesting observation was made when the hybridization behavior of the palindromic sequence d(5'-$A_7\mathbf{Z}_3T_7$-3') containing three neighboring 1,2,4-triazole ligands **Z** was examined in thermal de-/renaturation experiments (Figure 13) [54]. In the absence of metal ions, the strand forms a hairpin due to the destabilizing effect of the **Z** bases lacking any ability to engage in mutual hydrogen bonding. The loop region of this hairpin is formed by the central stretch of three **Z** ligands. Upon addition of Ag(I) ions, however, the oligonucleotides undergo a structural transition and metal binding between pairs of 1,2,4-triazole ligands **Z** leads to dimerization of the

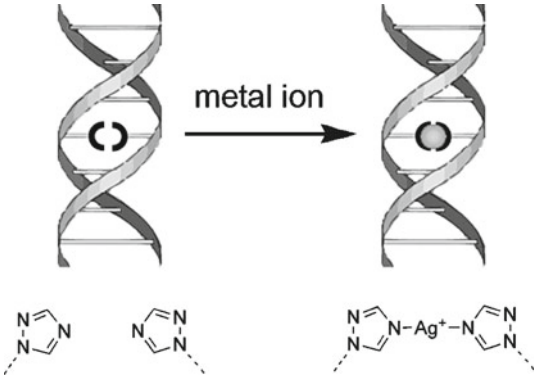


Figure 12 Assembly of the Ag(I)-mediated base pair of Müller et al. [54]. Reprinted from [19b] with permission from Elsevier; copyright 2010.

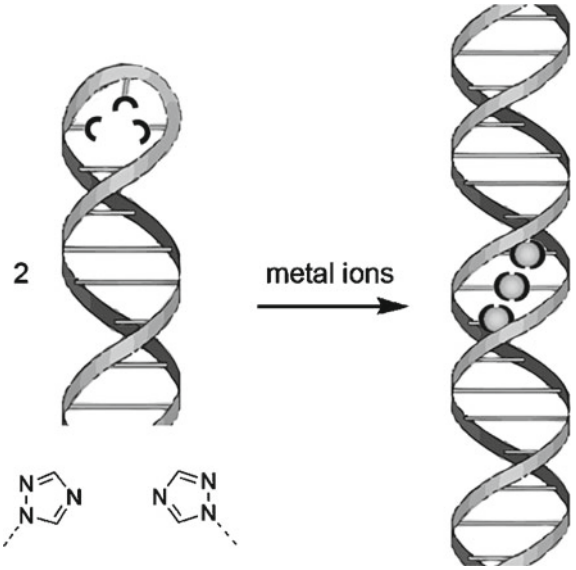


Figure 13 Control of hairpin to duplex transition by metal addition [54]. Reprinted from [19b] with permission from Elsevier; copyright 2010.

palindromic sequences resulting in a double strand containing three stacked Ag(I) ions. The evidence was brought by UV spectroscopic and MALDI-TOF mass spectrometric experiments. Furthermore it was found that the hairpin exhibits a concentration-independent melting temperature owing to its unimolecular melting behavior. In contrast, the melting curves of the sample containing Ag(I) ions showed a dependence on the oligonucleotide concentration which verifies the existence of a double strand. Subsequently, the proposed structural switching process was supported by fluorescence resonance energy transfer (FRET) experiments with dye-labeled strands of the same sequence and by the observation of an increase in molecular size in dynamic light scattering (DLS) experiments [54]. Recently, the groups around Müller and Sigel have determined the structure of a related palindromic sequence containing three consecutive imidazole-Ag(I)-base pairs **7** by NMR spectroscopy [63]. The results are discussed in the next Section.

3 Stacking and Mixing of Metals Inside DNA

After the fundamentals for the formation of metal-base pairs inside the DNA double helix using the discussed set of ligands and metal ions were elucidated, the focus in the field of metal-base pairing has shifted towards the realization of longer metal stacks inside DNA double strands [56]. Whereas the thorough examination of oligonucleotides containing only one metal-base pair is of great importance for determining basic parameters like binding constant and thermal duplex stabilization (as expressed by a ΔT_M value), the study of duplexes containing a number of directly stacked metal ions is supposed to contribute to the goal of realizing new DNA-based materials with single molecule electrical conductivity. The first report on consecutive metal stacking inside DNA double strands was given by Kuklenyik and Marzilli who examined the stacking of three Hg(II) ions between consecutive **TT** mismatches inside the middle of a short oligonucleotide $[d(5'\text{-GCGC}\mathbf{TTT}\text{GCGC-}3')]_2$ (Figure 14) [23b].

The extension to longer stretches of stacked metal ions, however, was not successful any more when using the higher homologous sequence $[d(5'\text{-GCGC-}\mathbf{TTTT}\text{GCGC-}3')]_2$. Instead of duplex formation, the addition of Hg(II) ions to this

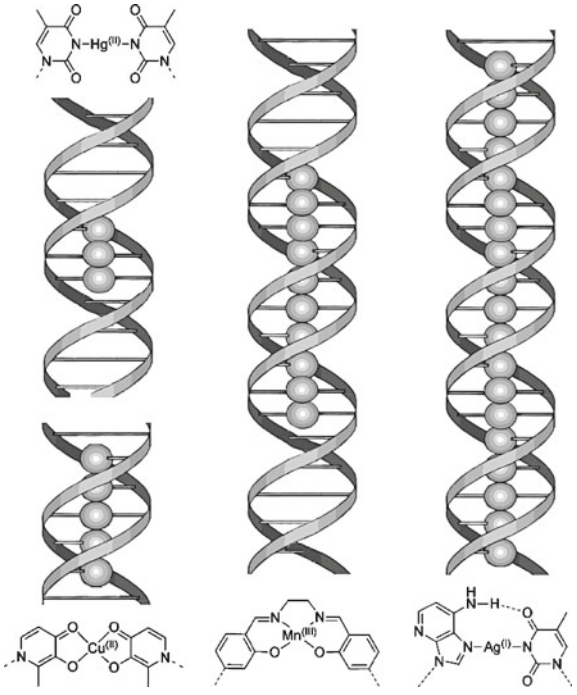


Figure 14 Examples of reported metal stacks inside DNA double helices. The sequences of stretches of natural Watson-Crick base pairs adjacent to the metal-base pair regions have been omitted for clarity. Reprinted from [19b] with permission from Elsevier; copyright 2010.

strand resulted in the arrangement of a hairpin due to metal-mediated intrastrand crosslinking of the terminal **T** residues of the $\mathbf{T}_4$ stretch. In contrast to this, Müller and Sigel et al. recently presented conclusive evidence for the formation of RNA duplexes containing up to six consecutive **U**-Hg-**U** base pairs (**5**, X = H) inside RNA duplexes by performing NOESY-NMR experiments with ^{15}N and ^{13}C labeled oligonucleotides [57].

Using the hydroxypyridone base pair **H**, Tanaka, Shionoya et al. achieved the stacking of up to five Cu(II) ions inside a duplex formed from the artificial oligonucleotides $[d(5'\text{-}G\mathbf{H}_nC\text{-}3')]_2$ ($n = 1 - 5$) [58]. The exclusive formation of the double strand $[d(5'\text{-}G\mathbf{H}_5C\text{-}3')]_2$ containing five Cu(II) ions was indicated by UV spectroscopic titration experiments and mass spectrometry. Most interestingly, EPR experiments showed a ferromagnetic interaction between the stacked paramagnetic Cu(II) ions. The EPR experiment also allowed a rough determination of the distance between the metal-base pairs to be 3.7 ± 0.1 Å which is comparable to the distance between the base pairs in natural DNA [58]. A subsequent theoretical DFT study by Mallajosyula and Pati supported the ferromagnetic coupling in this system of stacked **H**-Cu-**H** units (albeit assuming a Cu-Cu distance of 3.2 Å with the neighboring metals bridged by two oxygen atoms of the ligands in a $\{Cu_2O_2\}$ convex quadrangle structure) [59]. The same theoretical work proposed an antiferromagnetic coupling for two neighboring Cu(II)-salen-base pairs **S**-Cu-**S**. Indeed this was found to be true by a recent EPR study conducted by Schiemann et al. using a DNA double strand containing two adjacent Cu(II)-salen-metal-base pairs [60]. Nakanishi et al. also contributed theoretical studies for both above-mentioned base pairing systems [61].

Use of the salen ligand **S** introduced in Section 2.2 allowed the incorporation of ten metal ions in a row inside a double strand [53c]. The sequences were designed to consist of non-palindromic **GC** stretches at their ends in order to achieve preorganization of the ten adjacent pairs of salicylic aldehydes [(5'-CGGCC$\mathbf{S}_{10}$CGCGC -3')·(3'-GCCGG$\mathbf{S}_{10}$GCGCG-5')]. The aforementioned differences in duplex stabilization of Cu(II) versus Mn(III) were found to play an important role in the selective and well-controlled formation of double strands containing more than one metal-base pair. Whereas the higher stability of the salen-Cu(II)-base pair with respect to the salen-Mn(III)-base pair is beneficial in systems containing only a single salen ligand, the situation in case of the longer stacks inside DNA follows an opposite trend. Due to the high stability of the salen-Cu(II)-base pair, misfolded, undefined cross-linked or overlapping sequences are anticipated to be formed as kinetic products short after addition of the metal salt to the salen-containing DNA strands. The result was the observation of a large distribution of signals in the ESI mass spectra around the expected value for the duplex containing ten neatly stacked salen-Cu(II) complexes. In contrast, the addition of Mn(III) resulted in the formation of a much cleaner metalated DNA product after the reaction mixture was allowed to equilibrate at room temperature for several days as observed from the high resolution ESI-FTICR mass spectra [53c]. This result was ascribed to the "self-healing" capabilities of the salen-Mn(III)-containing duplexes allowing initially formed kinetic products to transform into one thermodynamic product after some time.

Besides mass spectrometry, UV-based titration experiments and CD spectroscopy were employed as analytical methods for the characterization of the multinuclear metal arrangements inside the DNA systems.

By combining both Ag(I) coordination and hydrogen bonding within one base pair consisting of one thymine **T** and one artificial 1-deazaadenine **D**, Polonius and Müller succeeded in the stacking of 19 Ag(I) ions inside a 20-mer sequence [62]. The **D** and **A** bases are thought to form a doubly hydrogen-bonded Hoogsteen base pair and the addition of Ag(I) ions leads to substitution of one of these hydrogen bonds as illustrated in Figure 14. Likewise to the considerations concerning the effect of the relative stabilities of the salen-Cu(II)- and Mn(III)-base pairs on the formation of neatly metalated duplexes, the rather low duplex stabilization observed for the **D**-Ag-**T** base pair was seen as an advantage here as well. Again, the error-free formation of long metal-containing double strands was thought to require structural equilibration to reach the thermodynamically most favored structure. In a joint work, Müller and Sigel et al. recently reported the solution structure determination of a DNA double strand containing three consecutive imidazole-Ag(I)-base pairs **7** by a detailed NMR spectroscopic study involving the direct measurement of the $^{1}J(^{15}N,^{107/109}Ag)$ coupling constants between the ligands and the metal ions. According to these measurements, the duplex adopts a B-type conformation with only minor deviations in the region of the metal-base pairs [63].

Since stacking of several metal ions of the same kind inside the DNA duplex has now been realized in practice, next steps in the direction of using artificial DNA strands as molecular wires such as single molecule conductivity measurements have to be undertaken. From the viewpoint of information storage and processing with DNA-based systems, the expansion of the genetic code with artificial base pairs is another highly interesting field of research. Also here, metal-base pairing has contributed new ideas since most metal-base pairs offer a binding mode orthogonal to the natural Watson–Crick binding motifs.

Going further in this direction, even the mixing of two different metal ions inside DNA was shown to be possible. Therefore, two ligand-modified nucleosides with orthogonal coordination capabilities for two different kinds of metal ions were chosen and incorporated into artificial oligonucleotides in the form of mixed arrays [64].

The groups around Shionoya and Carell succeeded in combining two different types of ligands which were shown to be able to bind Cu(II) and Hg(II) ions sequence-specifically in the same duplex [65]. The system presented by Shionoya et al. comprised the hydroxypyridone ligand **H** known for its ability to coordinate Cu(II) and the pyridine ligand **P** capable of coordinating Hg(II) (Figure 15). The formation of a duplex containing one Hg(II) and two Cu(II) ions was shown by UV- and CD-spectroscopic titration experiments with consecutive addition of both metal ions to the oligonucleotide d(5'-G**HPH**C-3') and subsequent characterization of the reaction products by ESI mass spectrometry [65].

Carell et al. synthesized strands containing a predetermined sequence of five salen ligands **S** with the propensity to bind Cu(II) and five **TT** mismatches capable of binding Hg(II) (Figure 16). Also in this case, the incorporation of the anticipated number of Cu(II) and Hg(II) ions according to the programmed sequence was

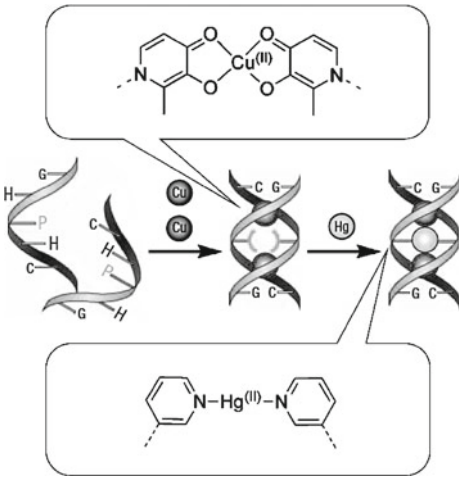


Figure 15 A mixed metal array based on the hydroxypyridone **H** and pyridine **P** ligands. Reprinted from [19b] with permission from Elsevier; copyright 2010.

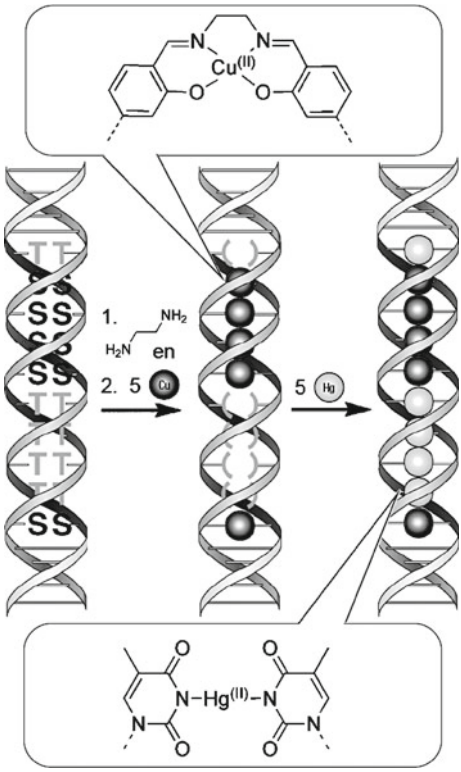


Figure 16 A mixed metal array based on the salen **S** and thymine **T** ligands. Reprinted from [19b] with permission from Elsevier; copyright 2010.

indicated by CD-spectroscopic titration studies and the products were indentified by ESI mass spectrometry [65].

Several different sequences of Cu(II) and Hg(II) ions could be aligned inside the DNA strands following a rational design and thus the binary pattern that was encoded into the oligonucleotide by the automated solid-phase synthesis could be translated into a determined array of metal ions [65]. In a similar manner, Ono et al. were recently able to incorporate **T**-Hg-**T** and **C**-Ag-**C** base pairs into the same double strand [66].

4 Conclusion and Future Prospects

The research field of metal-base pairing inside DNA has developed within 10 years from the first attempts to coordinate single metal ions by simple ligand-modified nucleosides to complex oligonucleotide systems such as metal-mediated triple helices and programmable multi-metal stacks [19]. From a structural point of view, the achieved duplex stabilization compared to unmodified oligonucleotides might find interesting applications in the field of DNA nanotechnology, for example by introducing metal-base pairs into DNA origami constructs. Furthermore, the switching of hybridization in the absence or presence of the required metal ions (or a change in their redox states) and the presented possibility to control structural interconversions (i.e., hairpin to duplex transitions) might add switchable and dynamic functions to DNA nanoscale architectures.

From another perspective, the precise positioning of transition metal ions in a three-dimensional space made up by a surrounding oligonucleotide framework could also be a new approach to mimic metalloenzymes with functions such as catalysis, charge separation, or selective guest recognition. The spin-spin interactions between neighboring paramagnetic metal ions such as Cu(II) inside DNA duplexes have already been shown to result in interesting magnetic phenomena. Combined with the potential charge transport capabilities, the newly discovered magnetic properties of the metal-containing oligonucleotides may become of great importance for the emerging field of spintronics. Also an application in future quantum computing devices is imaginable. Furthermore, the substitution of some or all natural base pairs by metal-base pairs may present great potential for the overall enhancement of charge conductivity through oligonucleotide-based materials and so contribute to the area of DNA-based molecular electronics. Further synthetic modification of the ligands and a broader choice of metal ions might be necessary to fine-tune the prospect electronic functions.

Upcoming work has to unambiguously establish the conductive properties of metal-containing DNA strands, e.g., by single molecule measurements in break junctions [16a], single electron transfer studies with charge injectors and traps attached to model duplexes [17] or spectroscopic methods that yield information about charge carrier availability [67].

Keeping in mind the various switching processes, even more sophisticated electronic functions beyond pure charge transport such as processing of logical inputs may be achieved using metal-modified DNA. The first examples going into this direction were introduced by Willner et al. by the application of orthogonal metal-base pairing (**T**-Hg-**T** vs. **C**-Ag-**C**) in combination with the optical properties of quantum dots as sensors for Hg(II) and Ag(I) ions and as logical AND as well as OR gates [28]. The obtained mixed-metal arrays serve as striking examples for the possibility to incorporate new base pairing principles orthogonal to and alongside with the natural Watson-Crick base pairs into DNA strands. All the discussed and many other works have by now generated a good insight into the factors leading to stable metal-base pairing (complex stability, geometrical prerequisites, compatibility with the natural base pairs) thereby paving the way for future application of artificial metal-base pairs in even more complex nanoscopic constructs.

Notes added in proof

Recently, Shionoya et al. [68] were able to prove the direct conductance of electrons through metal-containing DNA double strands in single-molecule experiments based on the attachment of DNA strands between the ends of single-walled carbon nanotubes. It was found that the conductivity along hydroxypyridone ligand-containing DNA duplexes can be switched on and off in the presence or absence of metal ions such as Cu(II), respectively. Longer stretches of stacked metal base pairs resulted in enhanced conductivities.

Carell et al. [69] demonstrated recently that metal-salen base pairs can be formed inside DNA double strands by enzymatic DNA polymerization using salicylic aldehyde triphosphates, ethylenediamine as reversible cross-linker. A single-crystal X-ray structure of the salen-Cu(II)-containing DNA strand in complex with a fragment of DNA polymerase I from *Bacillus stearothermophilus* showed that the metal base pair fits perfectly inside the B-DNA-like helical structure.

Abbreviations

AFM	atomic force microscopy
CD	circular dichroism
CHES	2-(cyclohexylamino)ethanesulfonic acid
DFT	density functional theory
Dipic	pyridine-2,6-dicarboxylate
DLS	dynamic light scattering
dNTP	2’-deoxyribonucleoside 5’-triphosphate
EDTA	ethylenediamine-N,N,N’,N’-tetraacetate

EPR	electron paramagnetic resonance
ESI	electro spray ionization
FRET	fluorescence resonance energy transfer
FTICR	Fourier-transform ion cyclotron resonance
HPLC	high performance liquid chromatography
ITC	isothermal titration calorimetry
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
NOESY	nuclear Overhauser effect spectroscopy
PCR	polymerase chain reaction
salen	N,N'-ethylenebis(salicylimine)
SNPs	single nucleotide polymorphisms
T_M	melting temperature
TRIS	tris(hydroxymethyl)-aminomethane

Acknowledgments G.H.C. thanks the "Fonds der Chemischen Industrie" and the "Deutsche Forschungsgemeinschaft" (IRTG 1422 – Metal Sites in Biomolecules) for generous support. This work was supported by grants-in-Aids from MEXT of Japan and the Global COE Program for Chemistry Innovation.

References

1. H. Lodish, A. Berk, S. L. Zipursky, P. Matzudaira, D. Baltimore, J. Darnell, *Molecular Cell Biology*, 6th edn, WH Freeman and Co, New York, **2007**.
2. (a) S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum, M. Famulok, *J. Am. Chem. Soc.* **2005**, *127*, 15071; (b) P. M. E. Gramlich, C. T. Wirges, J. Gierlich, T. Carell, *Org. Lett.* **2008**, *10*, 249.
3. (a) F. A. Aldaye, A. Palmer, H. F. Sleiman, *Science* **2008**, *321*, 1795; (b) U. Feldkamp, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2006**, *45*, 1856; (c) M. Endo, H. Sugiyama, *ChemBioChem* **2009**, *10*, 2420; (d) F. C. Simmel, *Angew. Chem. Int. Ed.* **2008**, *47*, 5884; (e) F. A. Aldaye, P. K. Lo, P. Karam, C. K. Mclaughlin, G. Cosa, H. F. Sleiman, *Nat. Nanotechnol.* **2009**, *4*, 349; (f) A. Heckel, M. Famulok, *Biochimie* **2008**, *90*, 1096.
4. G. E. Moore, *Electronics* **1965**, *38*, 114.
5. (a) C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2001**, *40*, 4128; (b) K. V. Gothelf, T. H. LaBean, *Org. Biomol. Chem.* **2005**, *3*, 4023; (c) J. Wengel, *Org. Biomol. Chem.* **2004**, *2*, 277; (d) J. J. Storhoff, C. A. Mirkin, *Chem. Rev.* **1999**, *99*, 1849; (e) T. Nguyen, A. Brewer, E. Stulz, *Angew. Chem. Int. Ed.* **2009**, *48*, 1974; (f) M. D. Sorensen, M. Petersen, J. Wengel, *Chem. Comm.* **2003**, 2130; (g) D. J. Hurley, Y. Tor, *J. Am. Chem. Soc.* **1998**, *120*, 2194.
6. (a) F. Seela, Y. He, *J. Org. Chem.* **2003**, *68*, 367; (b) S. A. Benner, *Acc. Chem. Res.* **2004**, *37*, 784.
7. A. T. Krueger, E. T. Kool, *Chem. Biol*. **2009**, *16*, 242.
8. A. Okamoto, K. Kanatani, I. Saito, *J. Am. Chem. Soc*. **2004**, *126*, 4820.
9. (a) N. C. Seeman, *Nature* **2003**, *421*, 427; (b) N. C. Seeman, *Int. J. Nanotechnol.* **2005**, *2*, 348.
10. H. T. Maune, S. Han, R. D. Barish, M. Bockrath, W. A. Goddard III, P. W. K. Rothemund, E. Winfree, *Nat. Nanotechnol*. **2010**, *5*, 61.
11. P. W. K. Rothemund, *Nature* **2006**, *440*, 297.
12. H. Yan, *Science* **2004**, *306*, 2048.
13. A. Somoza, *Angew. Chem. Int. Ed*. **2009**, *48*, 9406.
14. S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, W. M. Shih, *Nature* **2009**, *459*, 414.

15. E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. D. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, J. Kjems, *Nature* **2009**, *459*, 73.
16. (a) X. Guo, A. A. Gorodetsky, J. Hone, J. K. Barton, C. Nuckolls, *Nat. Nanotechnol.* **2008**, *3*, 163; (b) R. Mas-Balleste, O. Castillo, P. J. Sanz Miguel, D. Olea, J. Gomez-Herrero, F. Zamora, *Eur. J. Inorg. Chem.* **2009**, 2885.
17. (a) C. J. Murphy, M. R. Arkin, T. Jenkins, N. D. Ghatlia, S. H. Bossmann, N. J. Turro, J. K. Barton, *Science* **1993**, *262*, 1025; (b) S. Breeger, M. von Meltzer, U. Hennecke, T. Carell, *Chem. Eur. J.* **2006**, *12*, 6469; (c) B. Elias, J. C. Genereux, J. K. Barton, *Angew. Chem. Int. Ed.* **2008**, *47*, 9067.
18. C. M. Niemeyer, M. Adler, *Angew. Chem. Int. Ed*. **2002**, *41*, 3779.
19. (a) G. H. Clever, C. Kaul, T. Carell, *Angew. Chem. Int. Ed.* **2007**, *46*, 6226; (b) G. H. Clever, M. Shionoya, *Coord. Chem. Rev.* **2010**, *254*, 2391.
20. (a) A. Rakitin, P. Aich, C. Papadopoulos, Y. Kobzar, A. S. Vedeneev, J. S. Lee, J. M. Xu, *Phys. Rev. Lett.* **2001**, *86*, 3670; (b) P. Aich, S. L. Labiuk, L. W. Tari, L. J. T. Delbaere, W. J. Roesler, K. J. Falk, R. P. Steer, J. S. Lee, *J. Mol. Biol.* **1999**, *294*, 477.
21. S. Nokhrin, M. Baru, J. S. Lee, *Nanotechnology* **2007**, *18*, 095205.
22. (a) S. Katz, *J. Am. Chem. Soc.* **1952**, *74*, 2238; (b) S. Katz, *Biochim. Biophys. Acta* **1963**, *68*, 240.
23. (a) E. Buncel, C. Boone, H. Joly, R. Kumar, A. R. J. Norris, *Inorg. Biochem.* **1985**, *25*, 61; (b) Z. Kuklenyik, L. G. Marzilli, *Inorg. Chem.* **1996**, *35*, 5654.
24. (a) A. Ono, H. Togashi, *Angew. Chem. Int. Ed.* **2004**, *43*, 4300; (b) Y. Miyake, H. Togashi, M. Tashiro, H. Yamaguchi, S. Oda, M. Kudo, Y. Tanaka,Y. Kondo, R. Sawa, T. Fujimoto, T. Machinami, A. Ono, *J. Am. Chem. Soc*. **2006**, *128*, 2172; (c) Y. Tanaka, S. Oda, H. Yamaguchi, Y. Kondo, C. Kojima, A. Ono, *J. Am. Chem. Soc.* **2007**, *129*, 244.
25. S. Johannsen, S. Paulus, N. Düpre, J. Müller, R. K. O. Sigel, *J. Inorg. Biochem*. **2008**, *102*, 1141.
26. A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. Machinami, S. Oda, Y. Miyake, I. Okamoto, Y. Tanaka, *Chem. Comm.* **2008**, 4825.
27. D. A. Megger, J. Müller, Nucleosides, *Nucleotides and Nucleic Acids* **2010**, *29*, 27.
28. R. Freeman, T. Finder, I. Willner, *Angew. Chem. Int. Ed*. **2009**, *42*, 7818.
29. E. Ennifar, P. Walter, P. Dumas, *Nucleic Acids Res*. **2003**, *31*, 2671.
30. J. Joseph, G. B. Schuster, *Org. Lett*. **2007**, *9*, 1843.
31. A. A. Voityuk, *J. Phys. Chem. B* **2006**, *110*, 21010.
32. S.-P. Liu, S.-H. Weisbrod, Z. Tang, A. Marx, E. Scheer, A. Erbe, *Angew. Chem. Int. Ed*. **2010**, *49*, 3313.
33. (a) A. Houlton, A. R. Pike, M. A. Galindo, B. R. Horrocks, *Chem. Comm.* **2009**, 1797; (b) G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen, T. Carell, *J. Am. Chem. Soc.* **2006**, *128*, 1398.
34. J. Müller, *Metallomics* **2010**, *2*, 318.
35. K. Tanaka, M. Shionoya, *J. Org. Chem*. **1999**, *64*, 5002.
36. K. Tanaka, Y. Yamada, M. Shionoya, *J. Am. Chem. Soc*. **2002**, *124*, 8802.
37. (a) E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg, P. G. Schultz, *J. Am. Chem. Soc.* **2000**, *122*, 10714; (b) N. Zimmermann, E. Meggers, P. G. Schultz, *Bioorg. Chem.* **2004**, *32*, 13.
38. S. Atwell, E. Meggers, G. Spraggon, P. G. Schultz, *J. Am. Chem. Soc*. **2001**, *123*, 364.
39. K. Tanaka, M. Shionoya, *Coord. Chem. Rev*. **2007**, *251*, 2732.
40. J. Müller, *Eur. J. Inorg. Chem.* **2008**, 3749.
41. I. Okamoto, K. Iwamoto, Y. Watanabe, Y. Miyake, A. Ono, *Angew. Chem. Int. Ed*. **2009**, *48*, 1648.
42. N. Zimmermann, E. Meggers, P. G. Schultz, *J. Am. Chem. Soc*. **2002**, *124*, 13684.
43. L. Zhang, E. Meggers, *J. Am. Chem. Soc*. **2005**, *127*, 74.
44. H. Weizman, Y. Tor, *J. Am. Chem. Soc*. **2001**, *123*, 3375.
45. (a) C. Switzer, S. Sinha, P. H. Kim, B. D. Heuberger, *Angew. Chem. Int. Ed.* **2005**, *44*, 1529; (b) C. Switzer, D. Shin, *Chem. Commun.* **2005**, 1342.
46. M. J. Gait (Ed.), *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, New York, **1990**.

47. C. A. Schalley (Ed.), *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH, Weinheim, **2007**.
48. K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shiro, M. Shionoya, *J. Am. Chem. Soc*. **2002**, *124*, 12494.
49. (a) M. K. Schlegel, L.-O. Essen, E. Meggers, *J. Am. Chem. Soc.* **2008**, *130*, 8158; (b) M. K. Schlegel, L. Zhang, N. Pagano, E. Meggers, *Org. Biomol. Chem.* **2009**, *7*, 476.
50. Y. Takezawa, K. Tanaka, M. Yori, S. Tashiro, M. Shiro, M. Shionoya, *J. Org. Chem*. **2008**, *73*, 6092.
51. Y. Takezawa, W. Maeda, K. Tanaka, M. Shionoya, *Angew. Chem. Int. Ed*. **2009**, *48*, 1081.
52. T. Ihara, T. Ishii, N. Araki, A. W. Wilson, A. Jyo, *J. Am. Chem. Soc*. **2009**, *131*, 3826.
53. (a) G. H. Clever, K. Polborn, T. Carell, *Angew. Chem. Int. Ed.* **2005**, *44*, 7204; (b) G. H. Clever, Y. Söltl, H. Burks, W. Spahl, T. Carell, *Chem. Eur. J.* **2006**, *12*, 8708; (c) G. H. Clever, T. Carell, *Angew. Chem. Int. Ed.* **2007**, *46*, 250.
54. D. Böhme, N. Düpre, D. A. Megger, J. Müller, *Inorg. Chem*. **2007**, *46*, 10144.
55. (a) J. Müller, D. Böhme, N. Düpre, M. Mehring, F.-A. Polonius, *J. Inorg. Biochem.* **2007**, *101*, 470; (b) J. Müller, D. Böhme, P. Lax, M. Morell Cerda, M. Roitzsch, *Chem. Eur. J.* **2005**, *11*, 6246.
56. J. K. Klosterman, Y. Yamauchi, M. Fujita, *Chem. Soc. Rev*. **2009**, *38*, 1714.
57. S. Johannsen, S. Paulus, N. Düpre, J. Müller, R. K. O. Sigel, *J. Inorg. Biochem*. **2008**, *102*, 1141.
58. K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shionoya, *Science* **2003**, *299*, 1212.
59. S. S. Mallajosyula, S. K. Pati, *Angew. Chem. Int. Ed*. **2009**, *48*, 4977.
60. G. H. Clever, S. J. Reitmeier, T. Carell, O. Schiemann, *Angew. Chem. Int. Ed*. **2010**, *49*, 4927.
61. (a) Y. Nakanishi, Y. Kitagawa, Y. Shigeta, T. Saito, T. Matsui, H. Miyachi, T. Kawakami, M. Okumura, K. Yamaguchi, *Polyhedron* **2009**, *28*, 1945; (b) T. Matsui, H. Miyachi, Y. Nakanishi, Y. Shigeta, T. Sato, Y. Kitagawa, M. Okumura, K. Hirao, *J. Phys. Chem. B*. **2009**, *113*, 12790.
62. F.-A. Polonius, J. Müller, *Angew. Chem. Int. Ed*. **2007**, *46*, 5602.
63. S. Johannsen, N. Megger, D. Böhme, R. K. O. Sigel, J. Müller, *Nat. Chem*. **2010**, *2*, 229.
64. J. Müller, *Nature* **2006**, *444*, 698.
65. K. Tanaka, G. H. Clever, Y. Takezawa, Y. Yamada, C. Kaul, M. Shionoya, T. Carell, *Nat. Nanotechnol*. **2006**, *1*, 190.
66. K. Yanagida, N. Hamochi, K. Sasano, I. Okamoto, A. Ono, *Nucleic Acids Symp. Ser*. **2007**, *51*, 179.
67. (a) K. Kawai, Y. Osakada, M. Fujitsuka, T. Majima, *J. Phys. Chem. B* **2008**, *112*, 2144; (b) R. Yamagami, K. Kobayashi, A. Saeki, S. Seki, S. Tagawa, *J. Am. Chem. Soc*. **2006**, *128*, 2212.
68. S. Liu, G. H. Clever, Y. Takezawa, M. Kaneko, K. Tanaka, X. Guo, M. Shionoya, *Angew. Chem. Int. Ed*. **2011**, *50*, 8886.
69. C. Kaul, M. Müller, M. Wagner, S. Schneider, T. Carell, *Nat. Chem*. **2011**, in press, DOI: 10.1038/nchem.1117.

Chapter 11
Metal-Mediated Base Pairs in Nucleic Acids with Purine- and Pyrimidine-Derived Nucleosides

Dominik A. Megger, Nicole Megger, and Jens Müller

Contents

D.A. Megger • N. Megger • J. Müller (✉)
Institute for Inorganic and Analytical Chemistry, University of Münster, Corrensstr. 28/30, D-48149 Münster, Germany
e-mail: mueller.j@uni-muenster.de

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_11,

Abstract Metal-mediated base pairs are transition metal complexes formed from complementary nucleosides within nucleic acid double helices. Instead of relying on hydrogen bonds, they are stabilized by coordinative bonds. The nucleosides acting as ligands do not necessarily have to be artificial. In fact, several examples are known of naturally occurring nucleobases (e.g., thymine, cytosine) capable of forming stable metal-mediated base pairs that are highly selective towards certain metal ions. This chapter provides a comprehensive overview of metal-mediated base pairs formed from natural nucleosides or from closely related artificial nucleosides that are pyrimidine or purine derivatives. It addresses the different strategies that lead to the development of these base pairs. The article focuses on structural models for metal-mediated base pairs, their experimental characterization within a nucleic acid, and on their possible applications.

Keywords cytosine • deazaadenine • M-DNA • metal-mediated base pairs • polymerase • sensor • thymine

1 Introduction to Metal-Mediated Base Pairs

The interaction of metal salts with nucleic acids has long been investigated (see refs in [1]). Early experiments focused on the reaction of DNA from natural sources or of polynucleotides with transition metal ions such as Co^{2+}, Ni^{2+}, Mn^{2+}, Zn^{2+}, Cd^{2+}, Cu^{2+}, Hg^{2+}, or Ag^{+} [2]. It became clear that the metal ions are extremely diverse in their reaction with nucleic acids [3,4], and that in particular Hg^{2+} and Ag^{+} were able to bind strongly to the nucleobases without unwinding the double helices and without binding to the phosphate groups [4]. Replacement of protons from the nucleobases was suggested as one possible binding mode, termed "type II binding" in the case of Ag^{+} [5]. The fact that one Ag^{+} was included per two nucleotide residues, accompanied by proton displacement, was attributed to the conversion of a N–H···N hydrogen bond within a complementary base pair to two coordinative bonds in N-Ag-N [5]. In the case of Hg^{2+}, experimental data pointed towards the preferential incorporation of the metal ion between the deprotonated N3 atoms of two (mismatched) thymidine residues [6].

The advent of automated solid-phase synthesis of oligonucleotides enabled a much more detailed picture of the metal-binding properties to be obtained by allowing sequence-specific studies. Numerous metal-mediated base pairs, i.e., base pairs in which hydrogen bonds are formally replaced by coordinative bonds to metal ions, have been generated site-specifically within nucleic acid double helices [7–9]. Hence, it became possible to investigate in more detail the above mentioned metal-mediated base pairs comprising natural nucleotides. In addition, artificial nucleotides have been developed with a larger affinity towards metal ions, increasing significantly the scope of this type of functionalized nucleic acid. Using artificial nucleosides the structures of metal-modified DNA double helices could be determined [10,11]. In this chapter, metal-mediated base pairs with natural nucleosides as well as artificial purine- or pyrimidine-derived nucleosides will be discussed. Selected examples

for metal-mediated base pairs formed from artificial nucleosides without close structural resemblance to their natural counterparts are dealt with in Chapter 10 of this book [12].

2 M-DNA: A Metal Ion Complex of DNA

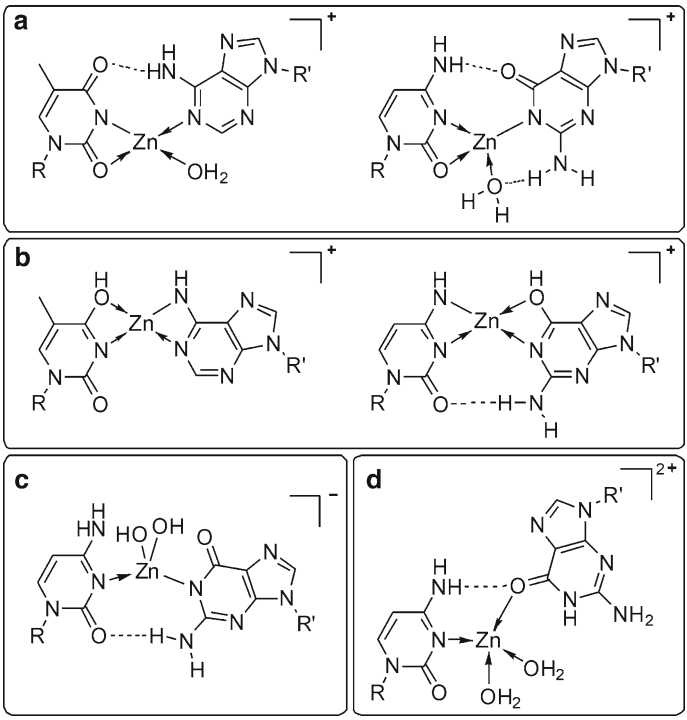
2.1 Formation of M-DNA

So-called M-DNA has been reported to form from regular B-DNA, both bacterial and synthetic, upon treatment with divalent cations such as Zn^{2+}, Ni^{2+}, or Co^{2+} at elevated pH values (pH > 8) [13,14]. Reaction with EDTA converts M-DNA back to B-DNA. The properties of this metal complex of DNA are quite distinct from those of B-DNA. For example, the well-established intercalator ethidium does not intercalate into M-DNA [13]. In fact, this observation has led to the development of a convenient assay for M-DNA formation based on the decrease in fluorescence that accompanies M-DNA formation. NMR studies suggest that the imino protons of thymine and guanine are absent in M-DNA [13], even though according to their pK_a values of 10.5 and 10.0 (values given for the respective 5'-phosphates) [15] they should still be protonated under slightly alkaline conditions. A detailed study revealed that each metal ion displaces one proton from the original B-DNA [14]. Interestingly, the UV and CD spectra of M-DNA almost resemble those of B-DNA, suggesting similar structures [13]. Based on these observations, base pairing schemes were suggested for M-DNA, in which one imino (or amino) proton of a base pair is replaced by a metal ion, leading to metal-mediated base pairs.

2.2 Structural Models

In addition to the originally proposed models for the metal-mediated base pairs (Scheme 1a, b) [14], combined computational and experimental modelling studies have identified other potential geometries [16]. Two of these are shown in Schemes (1c) and (1d).

According to the base pairing scheme shown in Scheme (1a), the metal ion replaces an imino proton of either thymine (A:T base pair) or guanine (G:C base pair). It is additionally coordinated to one aqua ligand residing in the minor groove [14]. Alternatively, the metal ion was suggested to displace an amino proton of either adenine (A:T base pair) or cytosine (G:C base pair) as shown in Scheme (1b). No aqua ligand is necessary to achieve a four-fold coordination of the metal ion in this model. However, the complementary nucleobase is required to be present in its rare enol form [14]. Taking into account the slightly alkaline conditions, a different model has been put forward for a metal-mediated G:C base pair in which the Zn^{2+} ion is coordinated by hydroxido ligands rather than an aqua ligand (Scheme 1c) [17]. Based



Scheme 1 Base pairing schemes suggested for M-DNA (R, R' = DNA backbone). (**a**) Replacement of the thymine and guanine imino protons by Zn^{2+} [14]. (**b**) Replacement of an adenine and cytosine amino proton by Zn^{2+} [14]. (**c**) Replacement of the guanine imino proton by Zn^{2+} [17]. (**d**) Loss of two hydrogen bonds without proton displacement by Zn^{2+} [18].

on the single-crystal X-ray diffraction analysis of a Zn^{2+} complex of 1-methylcytosine, a model for metal cross-links occurring during DNA melting in the presence of Zn^{2+} ions had been suggested [18] that was later proposed as a feasible base pairing scheme for M-DNA [16]. However, in this model, no proton displacement by Zn^{2+} takes place (Scheme 1d). All of these models contain a divalent cation coordinated by four donor atoms. In the case of Zn^{2+}, this typically translates into a tetrahedral coordination environment. According to a computational study, only the base pair shown in Scheme (1d) can adopt a planar geometry after full geometrical optimization [16], which is usually considered important for the formation of a stable DNA double helix with stacking neighboring base pairs. However, this base pair still contains the imino proton, and the resulting duplex would need to be parallel-stranded [16]. Both of these requirements are not in agreement with the experimental results.

Interestingly, the energy required to "flatten" the Zn^{2+}-mediated G:C base pair shown in Scheme 1(a) is relatively low (<8 kcal/mol), suggesting that it indeed represents a plausible model for building an antiparallel double helix [16]. Unfortunately, no structural information has yet been obtained for M-DNA. Single-crystal X-ray diffraction

analyses of oligonucleotides crystallized in the presence of Zn^{2+}, Co^{2+}, or Ni^{2+} under slightly alkaline conditions show a coordination of the metal ions to the N7 positions of the terminal guanine residues rather than an incorporation along the helical axis [19].

2.3 Physical Properties

M-DNA has gained a lot of attention because of reports that it supports metallic-like conduction [20]. Based on such electronic properties, the creation of DNA-based devices for nanoelectronic applications was suggested to be feasible. Evidence to support molecular wire behavior of M-DNA was collected using a variety of methods. For example, current-voltage curves of bundles of M-DNA duplexes placed between two electrodes separated by a gap of at least 1 μm have been measured directly [20]. Moreover, electrochemical investigations of self-assembled DNA monolayers on gold electrodes showed an increased electron transfer rate between different solute redox probes and the gold surface for M-DNA in comparison with B-DNA [21,22]. In particular, a differential behavior towards base mismatches was observed for B-DNA and M-DNA that was exploited for mismatch detection [23]. Finally, the attachment of donor and acceptor fluorophores at opposite ends of an M-DNA duplex enabled the investigation of charge transfer through the double helix. The observed fluorescence quenching was interpreted in terms of an electron hopping mechanism for electron transfer [24].

However, it needs to be stated that the enhanced electrical conductivity of M-DNA is heavily disputed. A magnetic study of the electronic states showed no substantial differences between B-DNA and M-DNA except for Mn^{2+}-doped DNA. Nonetheless, even in the latter case a charge transport through the metalated DNA was not observed [25]. Instead, it was suggested that the divalent cations replace the Na^+ ions shielding the negative charge of the phosphate backbone. A topographic characterization by atomic force microscopy showed that M-DNA is about five times shorter and ten times higher than regular B-DNA [26]. Electrostatic experiments performed with M-DNA adsorbed on mica surface showed an insulating behavior. The results of the electrochemical characterization have also been interpreted differently: The apparent enhancement of electrical conductivity of M-DNA could be attributed to the mere presence of divalent cations. Increased charge neutralization by these cations facilitates a penetration of the DNA monolayer by the negatively charged redox mediator, hence leading to the observed increase in conductivity [27]. The release of ethidium from B-DNA upon formation of M-DNA might be a result of the relatively low solubility of M-DNA: Upon precipitation under the conditions of M-DNA formation, ethidium could actually just dissociate from the precipitating B-DNA [28]. The displacement of the protons usually involved in hydrogen bonding could at least in part be explained by a (transient) coordination of the divalent cations to the nucleobases. From model chemistry it is well-known that the attachment of a metal ion to a nucleobase results in an acidification of both endocyclic imino and exocyclic amino groups [15]. Finally, the fluorescence

quenching observed upon M-DNA formation could be attributed to its more compact size, resulting in either intermolecular FRET between aggregated double helices or intramolecular FRET as a result of metal-mediated fraying of the helix ends [28].

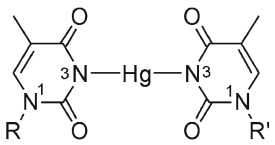
The prospect of self-assembling nanoscale electronic architectures based on DNA is certainly a fascinating one [29]. At present, it is not clear whether this goal can be achieved with M-DNA. Nonetheless, the studies of M-DNA have aroused interest in an investigation of metal-mediated base pairs with both natural and artificial nucleobases.

3 Metal-Mediated Base Pairs with Pyrimidine-Derived Nucleobases

3.1 Hg^{2+}-Mediated T:T Base Pairs

3.1.1 History of the T–Hg–T Base Pair

In the early 1950s, a decrease in DNA viscosity was observed upon the addition of $HgCl_2$ [30]. First attempts of an explanation involved a decrease in the overall size of the nucleic acid caused by a diminishing electrostatic repulsion between the phosphate groups due to the positively charged Hg^{2+}, potentially accompanied by intramolecular cross-links of the type P–O–Hg–O–P [30]. After it became clear that the nucleobases act as the binding sites for Hg^{2+} rather than the phosphate groups [31], the formation of a 2:1 complex was proposed, containing two deprotonated thymine residues and a central mercuric ion [6]. It was suggested that these base pairs (Scheme 2) were formed between two proximate thyminate residues of opposite strands by means of a chain shift mechanism [32].



Scheme 2 Chemical structure of a T–Hg–T base pair (R, R' = DNA backbone).

The molecular structure of the complex between deprotonated 1-methylthymine and Hg^{2+} strongly supports the proposed metal-mediated base pair (Figure 1). However, the structure obtained by single crystal X-ray diffraction analysis contains the nucleobases in a transoid orientation whereas a cisoid arrangement of the glycosidic bonds is required for an incorporation into a B-type duplex [33].

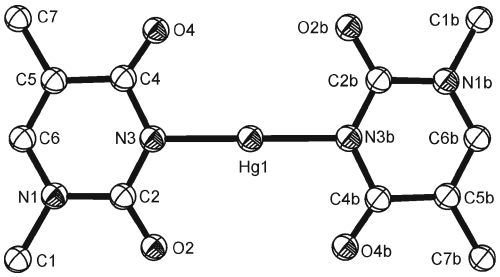


Figure 1 Molecular structure of [Hg(1-methylthyminate)$_2$] [33].

3.1.2 Characterization of the T–Hg–T Base Pair within a DNA Duplex

As mentioned above, the advent of solid phase oligonucleotide synthesis enabled sequence-specific studies of the metal-binding behavior of the various nucleobases. Accordingly, the selective formation of T–Hg–T base pairs inside a DNA duplex could be investigated. Using hairpin structures comprising a loop of two or three neighboring thymine residues, a rearrangement towards a regular duplex was observed upon the addition of Hg^{2+}. This rearrangement was accompanied by the formation of T–Hg–T base pairs, as could be shown by ^{1}H NMR spectroscopy [34]. Interestingly, a hairpin containing four thymine residues in the loop did not rearrange towards a regular duplex upon the addition of Hg^{2+}. Instead, one intramolecular T–Hg–T cross-link between the first and the last thymine residues was observed [34], indicating a complex relationship between oligonucleotide sequence and conformation.

A quantitative study regarding the thermal stabilization of a T–Hg–T base pair indicated that the addition of Hg^{2+} to a DNA duplex containing one T:T mismatch results in an increase of the melting temperature T_m of 10°C (from 37 to 47°C) [35]. Control experiments using a regular A:T base pair instead of the mismatch showed a T_m of 44°C, indicating that the metal-mediated base pair is more stable than the naturally occurring A:T pair [35]. An isothermal titration calorimetry study revealed that the binding constant for specific binding of Hg^{2+} to a T:T mismatch amounts to almost 10^6 M^{-1}[36].

^{1}H NMR spectroscopy was applied to investigate in more detail the formation of this metal-mediated base pair. By using a duplex containing two T:T mismatches, the formation of T–Hg–T base pairs could be followed by observing the resonances of the thymine imino protons. In the absence of Hg^{2+}, four imino proton resonances were detected in the corresponding chemical shift region (Figure 2a, 10.4–11.1 ppm) [35]. Addition of Hg^{2+} led to a decrease in signal intensity, indicating a substitution of the imino protons by Hg^{2+} and hence the formation of a metal-mediated base pair (Figure 2c). In experiments using approximately half the amount of metal ions required for base pair formation, three independent species were observed in the ^{1}H NMR spectrum, i.e., the Hg^{2+}-free state and two partially Hg^{2+}-bound states (Figure 2b) [35]. This clearly indicates that the metal ions are tightly bound and that the dissociation and association processes occur slowly on the NMR timescale. Hence, T–Hg–T base pairs are formed in an equilibrium process (Figure 2d) [35].

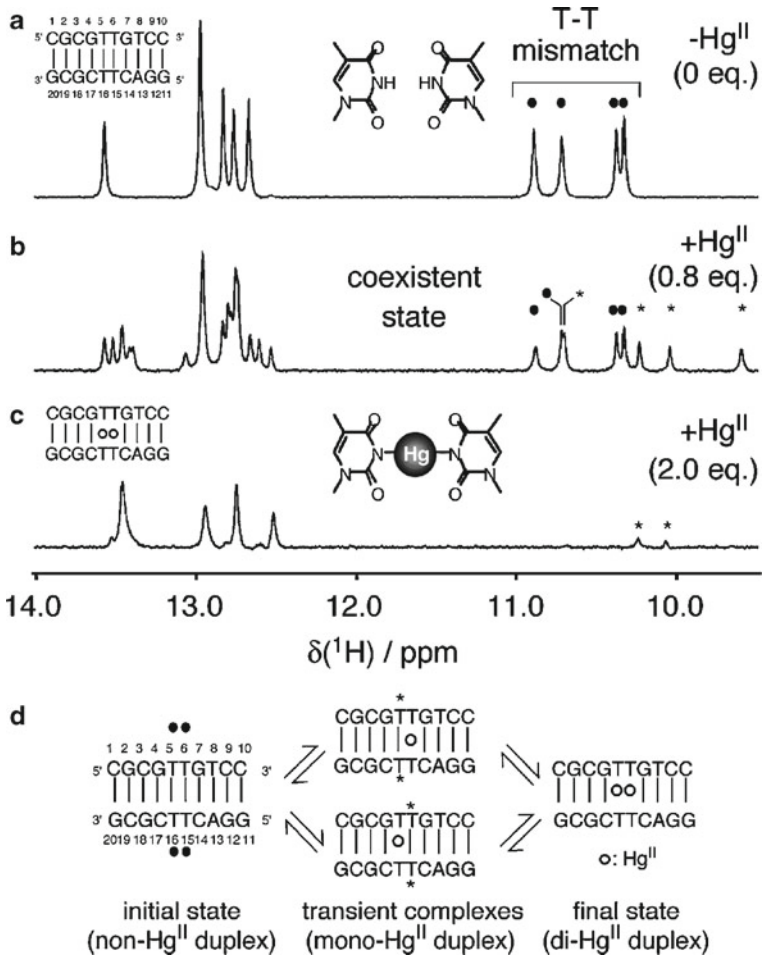


Figure 2 Imino proton region of the ^{1}H NMR spectra of an oligonucleotide double helix. (**a**) In the absence of Hg^{2+}; (**b**) in the presence of Hg^{2+} (0.8 equivalents with respect to the duplex); (**c**) in the presence of Hg^{2+} (2.0 equivalents with respect to the duplex); (**d**) equilibrium system between the duplex and Hg^{2+}. Reprinted with permission from [35]; copyright 2006 American Chemical Society.

Moreover, the two metal-mediated base pairs within the duplex are not formed cooperatively but independently from each other.

The first direct proof for the formation of a T–Hg–T base pair within a double helix was obtained by ^{15}N NMR spectroscopy. By incorporating ^{15}N-labelled thymine residues at different positions within oligonucleotide strands, metal ion binding to these residues could be unequivocally confirmed [37]: In the presence of Hg^{2+}, a significant downfield shift of the ^{15}N resonances of the endocyclic N3 atoms of thymine by ~30 ppm was observed. Such a significant shift can only be explained by the substitution of the imino proton by Hg^{2+}. Most interestingly, when both thymine residues within the metal-mediated base pair were ^{15}N-labelled a splitting of the ^{15}N resonance (2.4 Hz) was

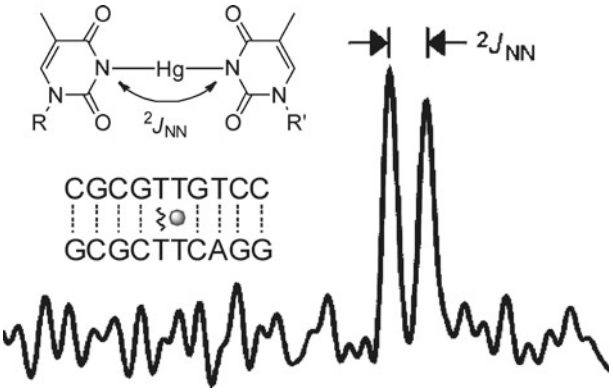


Figure 3 Detail of the ^{15}N NMR spectrum of the partially ^{15}N-labelled oligonucleotide duplex (R, R' = DNA backbone). The $^2J_{^{15}N,^{15}N}$ coupling of 2.4 Hz is highlighted by arrows. Adapted with permission from [37]; copyright 2007 American Chemical Society.

observed (Figure 3). This splitting was attributed to a $^2J_{^{15}N,^{15}N}$ coupling across the central metal ion, providing direct proof for the formation of the T–Hg–T base pair [37].

3.1.3 Application in Hg^{2+} Sensors

With Hg^{2+} being a toxic environmental pollutant, its reliable detection from aqueous solution is highly desirable. Accordingly, the formation of T–Hg–T base pairs has been exploited for the development of DNA-based Hg^{2+}-sensor systems. In the first example of this type, a thymine-rich oligonucleotide functionalized with a fluorophore at its 3'-end and a quencher at its 5'-end was applied [38]. In the absence of Hg^{2+} the oligonucleotide adopts a randomly coiled single-stranded conformation. However, in the presence of Hg^{2+} a hairpin structure comprising several T–Hg–T base pairs is adopted. In this conformation, the fluorescence is quenched with increasing concentration of Hg^{2+} because fluorophore and quencher are located close to each other [38]. To ensure a high specificity of the sensor, control experiments with other metal ions were performed, such as Ca^{2+}, Mg^{2+}, Cu^{2+}, Fe^{2+}, Cd^{2+}, Pb^{2+}, Zn^{2+}, Ni^{2+}, Mn^{2+}, Co^{2+}. None of these influence significantly the fluorescence spectra. Hence, this Hg^{2+} sensor is sensitive enough to detect small amounts of Hg^{2+} in solutions containing an excess of other metal ions [38]. By combining the T–Hg–T base pairs with Au nanoparticle-functionalized reporter DNA [39] or electrochemiluminescent tags [40], the detection limit can be lowered down to the picomolar concentration range.

As mercury cations can exist in two oxidation states, the T–Hg–T base pairs have also been applied as sensors for the redox environment [41]. Here, an oligonucleotide containing only thymine residues was labelled with fluorophore and quencher as described above. In the presence of Hg^{2+} a hairpin structure with metal-mediated base pairs is adopted, leading to a quenched fluorescence. Upon introduction of a

reductive environment (via the addition of Na_2SO_3), Hg^{2+} is reduced to Hg_2^{2+}. As the formation of Hg^{2+}-mediated base pairs is no longer possible, the folded structure is converted back to the randomly coiled conformation, leading again to a recovering fluorescence [41].

3.1.4 U–Hg–U, an RNA Analogue

Apart from a gold-mediated G:C base pair accidentally formed during the crystallization of a nucleic acid [42], only one metal-mediated base pair has been reported as yet for RNA [43]: In analogy to the well-established T–Hg–T base pair, RNA duplexes containing uracilate–Hg^{2+}–uracilate (U–Hg–U) base pairs have been synthesized. As the presence or absence of a methyl group at the C5 position represents the only difference between thymine and uracil, a similar complexation behavior towards Hg^{2+} was expected.

So far, five RNA duplexes have been described, comprising stretches of up to 20 consecutive uracil residues, thereby being able to potentially form up to 20 Hg^{2+}-mediated base pairs [43]. A combination of different spectroscopic techniques and dynamic light scattering (DLS) were applied to demonstrate that the N3 imino proton of uracil is substituted by Hg^{2+} [43], in analogy to the situation described above for thymine. The RNA oligonucleotides under investigation were either regular duplexes with U:U mismatches (most likely in the form of a wobble pair) or hairpins with the uracil residues located in the loop. The latter case resembles the situation described in Section 3.1.2, where the formation of T–Hg–T base pairs leads to a hairpin-to-duplex transition [34]. For the RNA hairpins, this conformational change was followed by DLS and DOSY NMR spectroscopy, as double helix and hairpin show significantly different hydrodynamic radii r_H: Upon the addition of Hg^{2+}, r_H increases by about 30%, indicating that a conformational change takes place from hairpin to a regular duplex with U–Hg–U base pairs [43].

This conformational change was further monitored by NMR spectroscopic techniques other than DOSY. In particular, the $^1H,^1H$-NOESY spectrum of the hairpin r(GGAGCGCGUUUUUUCGCGCUCC) comprising six uracil residues in the loop displays well-resolved and dispersed peaks in the absence of Hg^{2+} [43]. The fact that a sequential walk could be followed through the entire sequence indicates the formation of a well-structured loop. The A-type conformation of the stem region was confirmed by characteristic cross-peaks [43]. Upon the addition of Hg^{2+}, the ^1H NMR spectra changed in the same manner as previously observed for DNA duplexes comprising T–Hg–T base pairs: The signal intensity of the N3 imino proton resonances of the uracil residues involved in metal-mediated base pairing decreased [43]. Moreover, a study applying a ^{13}C-labelled oligonucleotide confirmed that only those uracil residues not involved in Watson-Crick base pairing (i.e., all uracil residues except of U20) bind to Hg^{2+}. As can be seen from the $^1H,^{13}C$-HSQC spectrum (Figure 4), the resonances of the central uracil residues shift significantly after the addition of Hg^{2+}, indicating the formation of U–Hg–U base pairs, whereas the resonance of U20 is least affected [43].

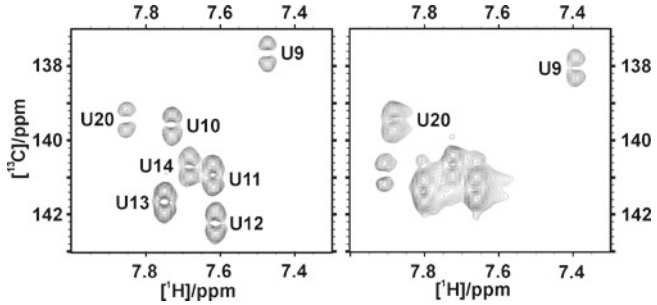


Figure 4 Section of the $^{1}H,^{13}C$-HSQC spectra of r(GGAGCGCGUUUUUUCGCGCUCC) in the absence (left) and presence (right) of Hg^{2+}, showing the ^{1}J correlation between H6 and C6 of all uracil residues. Reprinted with permission from [43], copyright 2008 Elsevier.

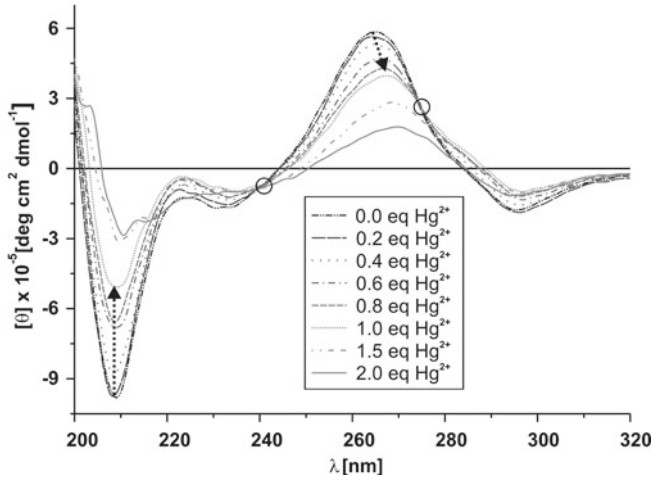


Figure 5 CD spectra of r(GGAGCGCGUUUUUUCGCGCUCC), confirming an A-type duplex conformation in the absence and presence of up to one equivalent of Hg^{2+}. Here, one equivalent is defined as the amount of metal ion required to saturate every possible metal-mediated base pair. The isosbestic points observed up to the addition of 1 equivalent of Hg^{2+} are highlighted by circles. Reprinted with permission from [43], copyright 2008 Elsevier.

In analogy to the T–Hg–T base pairs, the formation of U–Hg–U base pairs is accompanied by an increase in thermal stability [43]. CD spectroscopy was used to confirm the A-type helical conformation (Figure 5). This study showed that U–Hg–U base pairs are formed in an equilibrium process, as indicated by the appearance of isosbestic points upon the stepwise addition of Hg^{2+}. Once all U–Hg–U base pairs are formed, excess Hg^{2+} binds unspecifically to other coordination sites of the duplex [43].

Interestingly, Hg^{2+} can be removed from the U–Hg–U base pairs by treating the RNA duplexes with a chelating resin, restoring the previous nucleic acid

conformation [43]. This behavior is unlike that found for T–Hg–T base pairs [37] (and also for the artificial imidazole–Ag^+–imidazole base pairs) [11], where the metal ions cannot be removed by using chelating agents. Hence, this different reactivity might be attributed to the A-type conformation of the RNA duplex, in which the metal ions are not aligned along the helical axis but twisted around it.

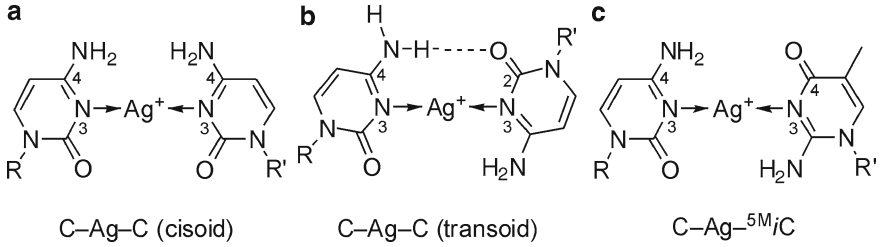
The incorporation of U–Hg–U base pairs into RNA duplexes demonstrates that this nucleic acid should not be disregarded in the context of metal-mediated base pairs. As a result of the different helical conformation compared with DNA, the scope of this type of nucleic acid functionalization can be significantly broadened by including RNA.

3.2 Ag^+-Mediated C:C Base Pairs

Apart from thymine and uracil, the third naturally occurring pyrimidine nucleobase cytosine has also been utilized in the construction of metal-mediated base pairs. In contrast to the previously discussed examples, metal-mediated self-pairs of cytosine are specifically formed in the presence of Ag^+.

The reversible binding of Ag^+ to oligonucleotides containing cytosine-rich regions has been known for about 50 years [44]. In the course of some of the first studies, the interaction of Ag^+ with poly-d(C) was investigated by potentiometric and photometric methods [3,5]. Even though the metal-binding sites were not clear at that point, the formation of Ag^+-mediated base pairs in the so-called "type II binding" was suggested as a result of the substitution of a proton involved in hydrogen bonding by a metal ion. In contrast to the molecular structure of [Hg(1-methylthyminate)$_2$] (Figure 1), the complex [Ag(1-MeC)](NO_3) synthesized from 1-methylcytosine (1-MeC) and $AgNO_3$ shows a 1:1 stoichiometry between ligand and metal ion [45,46]. In this dinuclear complex, Ag^+ is bound by two cytosine residues in an N3–Ag–O2 fashion and, due to the coordination of the nitrate counter ions, displays a distorted trigonal coordination sphere. Based on this structure, models were suggested for Ag^+-mediated base pairs, e.g., in nucleic acid regions of high G:C content [45].

The Ag^+-mediated cytosine self-pair C–Ag–C was recently discovered as another example for a metal-meditated self-pair of naturally occurring nucleobases [47]. By using an oligonucleotide double helix comprising a single C:C mismatch within otherwise regular Watson-Crick base pairs, Ag^+ was shown to be the only metal ion significantly stabilizing the duplex, with an increase of the melting temperature T_m of 8°C [47]. Other metal ions such as Hg^{2+}, Cu^{2+}, Ni^{2+}, Pd^{2+}, Co^{2+}, Mn^{2+}, Zn^{2+}, Pb^{2+}, Cd^{2+}, Mg^{2+}, Ca^{2+}, Fe^{2+}, Fe^{3+}, or Ru^{3+} had no notable effects on T_m, making the C:C mismatch highly selective for the binding of Ag^+. An investigation by means of ^{1}H NMR spectroscopy confirmed the formation of C–Ag–C base pairs, as is obvious by significant changes in the imino proton region of the spectra. In analogy to what had been observed for T–Hg–T base pairs [35], a spectrum recorded at a lower Ag^+ concentration revealed the existence of two independent species, namely the metalated and non-metalated duplex [47]. The binding constant for specific binding of Ag^+ to a C:C mismatch within a regular DNA duplex lies in the range of $10^5\,M^{-1}$ [48].



Scheme 3 Postulated structures of the Ag^+-mediated base pairs C–Ag–C (cisoid and transoid) and C–Ag–^{5M}*i*C (R, R' = DNA backbone) [47,49,50]. The latter base pair probably also contains one hydrogen bond. Coordinative bonds are drawn as arrows and hydrogen bonds as broken lines.

In these examples, the formation of a Watson-Crick duplex containing a single C:C mismatch enforces a cisoid orientation of the nucleobases inside the C–Ag–C base pair (Scheme 3a). However, the corresponding transoid base pair is expected to be of higher stability, due to diminished electrostatic repulsion and the possibility to form an additional hydrogen bond (Scheme 3b). In fact, CD and UV spectroscopic investigations of $d(C_8)$ containing only cytosine moieties suggest the preferred formation of transoid C–Ag–C base pairs if the strand orientation is not predetermined by surrounding nucleobases [49]. Importantly, a comparison of the CD spectra in the presence of Ag^+ and under slightly acidic conditions, respectively, rules out that the addition of Ag^+ leads to an i-motif-like structure.

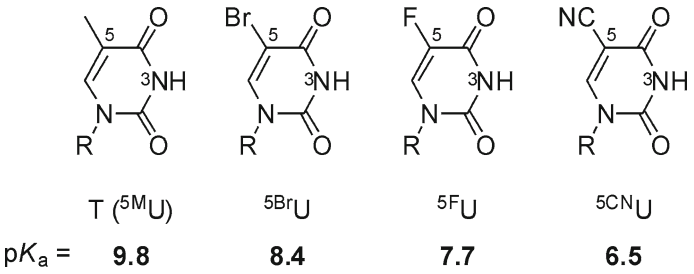
The preferred orientation of a C–Ag–C base pair was probed by a series of experimental and theoretical investigations. For instance, various Watson-Crick duplexes differing only in a single base mismatch comprising cytosine and/or 5-methylisocytosine (^{5M}*i*C) (Scheme 3c) were synthesized and their melting temperature determined [50]. The formation of the C–Ag–^{5M}*i*C base pair (ΔT_m = +18.9°C) results in a significantly larger increase in T_m than that of the cisoid C–Ag–C base pair (ΔT_m = +4.8°C). As the C–Ag–^{5M}*i*C base pair geometry closely resembles that of the transoid C–Ag–C one, these data strongly suggest that the latter is the thermodynamically more stable one. Moreover, experiments with DNA duplexes covalently locked in either a parallel or an antiparallel orientation proved the stabilization of parallel-stranded duplexes by transoid C–Ag–C base pairs [51]. Very recently, dispersion-corrected DFT calculations have been carried out revealing that the transoid C–Ag–C is indeed stabilized by one hydrogen bond and therefore about 7 kcal mol^{-1} more stable than the corresponding cisoid one [52].

Applications of the Ag^+- mediated cytosine self-pair include the oligonucleotide-based detection of this metal ion. In analogy to the Hg^{2+} sensor mentioned in Section 3.1.3, a turn-off fluorescent sensor utilizing the quenching of fluorescence as a result of a conformational change of an oligonucleotide upon the formation of C–Ag–C base pairs has been reported [47]. A similar turn-on sensor based on a fluorescein-labelled single-stranded oligonucleotide has also been reported, taking advantage of the fact that graphene oxide selectively adsorbs and quenches the labelled single-stranded DNA probe but not the duplex comprising C–Ag–C base pairs [53]. Both sensors are highly selective towards Ag^+ with detection limits in the nanomolar range.

3.3 Chemically Modified Pyrimidines

3.3.1 5-Substituted Uracil Derivatives

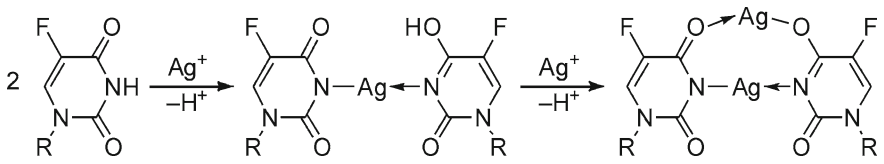
As the naturally occurring pyrimidine nucleobases have been found to be excellent ligands for the formation of metal-mediated base pairs, several chemically modified pyrimidine derivatives have been investigated in an attempt to control the stability and selectivity towards different metal ions. A promising approach for such a fine-tuning is based on the use of 5-substituted uracil derivatives [54]. The pK_a value of the amide proton of these pyrimidine derivatives varies from 6.5 to 9.8, depending on the electron-withdrawing effect of the substituent in the 5-position (Scheme 4). As the formation of a metal-mediated base pair requires the deprotonation of this site, a change in pK_a should directly influence the stability of such a base pair. Moreover, the affinity towards metal ions depends on the nucleophilic character of the ligand which often goes along with its basicity.



Scheme 4 5-Substituted uracil derivatives used for metal-mediated base pairs (^{5M}U: 5-methyluracil = thymine; ^{5Br}U: 5-bromouracil; ^{5F}U: 5-fluorouracil; ^{5CN}U: 5-cyanouracil; R = DNA backbone) [54]. The pK_a values of the amide protons are shown as well (for R = CH_3).

The investigation of oligonucleotide duplexes containing one base mismatch each of the 5-substituted uracil derivatives revealed significant differences in the reactivity towards Hg^{2+} and Ag^+ in the pH range of 5.5–9. For example, in the case of a T:T mismatch the addition of Ag^+ does not lead to a change of T_m (within error limits), whereas the addition of Hg^{2+} results in a duplex stabilization of $\Delta T_m \approx$ 7°C almost irrespective of the pH [54]. In contrast, 5-fluorouracil (^{5F}U) shows a completely different reactivity. Under slightly acidic conditions, the presence of Hg^{2+} results in a significant stabilization of the duplex (ΔT_m = 7°C) whereas the addition of Ag^+ does not affect T_m at all. By changing the pH to slightly basic conditions, the reactivity is reversed: Now Ag^+ stabilizes the duplex more significantly than Hg^{2+} (ΔT_m = 14°C vs. 6°C) [54].

The differential binding of these metal ions was corroborated by a mass spectrometric study from a solution containing both metal ions, directly proving the specific binding of Hg^{2+} at pH 7 and of Ag^+ at pH 9, respectively [54]. Interestingly, a UV titration suggested that the base pair stabilized by Ag^+ contains two metal ions (Scheme 5). This finding did not come as a total surprise, as it had previously been predicted based



Scheme 5 Proposed formation of $^{5F}U–Ag_2–^{5F}U$ (R = DNA backbone) [54].

on the single crystal X-ray diffraction analysis of the complex [Ag(1-methyluracilate)] [55]. In summary, a carefully designed manipulation of a nucleobase at a position that is not involved in metal-ion binding can affect the stability of the resulting metal-mediated base pairs and concomitantly the affinity towards different metal ions.

3.3.2 4-Substituted Pyrimidinone Derivatives

All base pairs discussed so far in Section 3 contain a metal ion with linear coordination geometry (Hg^{2+} or Ag^{+}). Even though metal complexes are known that reveal the ability of thymine and cytosine to act as chelating ligands [56,57], no proof exists of comparable metal-mediated self-pairs of pyrimidines within a nucleic acid duplex. Hence, to extend the concept of pyrimidine self-pairs to metal ions preferring a more sophisticated coordination geometry, a functionalization of the metal-binding site of the nucleobase is necessary. In this context, the synthesis and metal-binding properties of oligonucleotides containing the artificial nucleobase 4-(2'-pyridyl)-pyrimidinone (Pyr^{Py}) have been reported [58]. This nucleobase is a cytosine derivative containing a pyridyl substituent attached to C4 instead of the amino group. Oligonucleotide duplexes containing the Pyr^{Py}:Pyr^{Py} mismatch as well as combinations of Pyr^{Py} with naturally occurring nucleobases have been investigated. In the absence of metal ions, Pyr^{Py} has a destabilizing effect on the duplex ($\Delta T_m = -15.1$°C compared to a regular G:C base pair). However, in the presence of Ni^{2+} the formation of a Pyr^{Py}–Ni–Pyr^{Py} base pair (Figure 6a) was delineated from a significant increase in T_m. The artificial Pyr^{Py}–Ni–Pyr^{Py} base pair is formed in a highly selective manner, as neither the presence of other divalent metal ions (Co^{2+}, Cu^{2+}, Zn^{2+}, Fe^{2+}, Mn^{2+}) nor a replacement of one of the Pyr^{Py} residues by a natural nucleobase results in a significant increase of T_m. As a result, the artificial Pyr^{Py}–Ni–Pyr^{Py} base pair has a stability and a mismatch discrimination comparable to those of natural Watson-Crick pairs [58].

Following the strategy of increasing the number of metal-binding sites on a cytosine derivative to modify its metal-binding properties, a 2,2'-bipyridine residue was attached at C4 instead of a pyridine ring, leading to a nucleobase with a terpyridine-derived binding motif (Pyr^{Bipy}) (Figure 6b) [59]. As Pyr^{Bipy} is a tridentate ligand, the formation of a planar base pair requires a monodentate ligand in the complementary position, e.g., a pyridyl moiety glycosylated at C3 (3Py) or C4 (4Py), respectively. Interestingly, only the combination of Pyr^{Bipy} with 4Py leads to the formation of a stable Ag^{+}- mediated base pair ($\Delta T_m = +12.9$°C) [59]. The addition of other metal ions or the choice of a different nucleobase complementary to Pyr^{Bipy} (e.g., Pyr^{Bipy}, 3Py, G, A, C, T) does not result in a remarkable duplex stabilization [59].

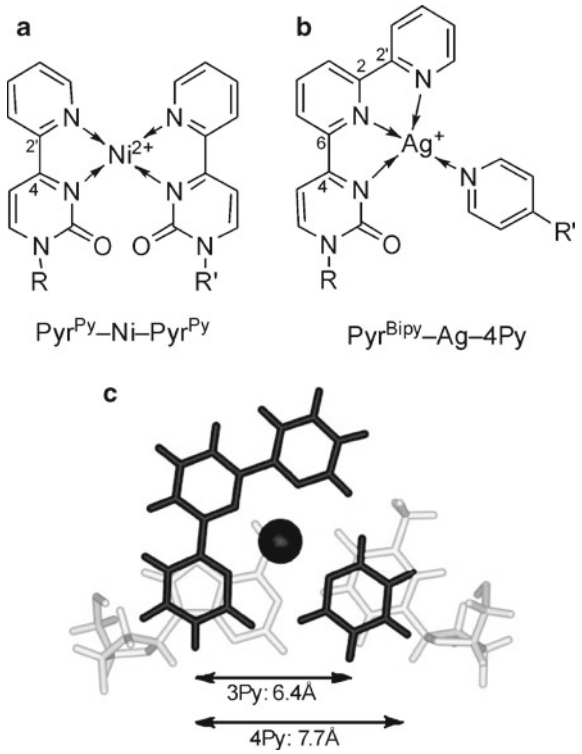


Figure 6 Schematic representations of (**a**) the Pyr^{Py}–Ni–Pyr^{Py} base pair [58] and (**b**) the Pyr^{Bipy}–Ag–4Py base pair [59]; (**c**) superposition of the geometry-optimized structure of Pyr^{Bipy}–Ag–Py on a natural A:T base pair (R, R' = DNA backbone). The distances between the glycosidic bonds are indicated by arrows. This figure has been adapted from [59] with permission of The Royal Society of Chemistry; copyright 2005.

The different stability of 3Py and 4Py is a prime example for the importance of base pair dimension and geometry. As shown by the geometry-optimized structure of a Pyr^{Bipy}–Ag–Py base pair, the distance between the glycosidic bonds is larger for Pyr^{Bipy}–Ag–4Py and therefore resembles more closely the distance found in natural DNA double helices (Figure 6c) [59].

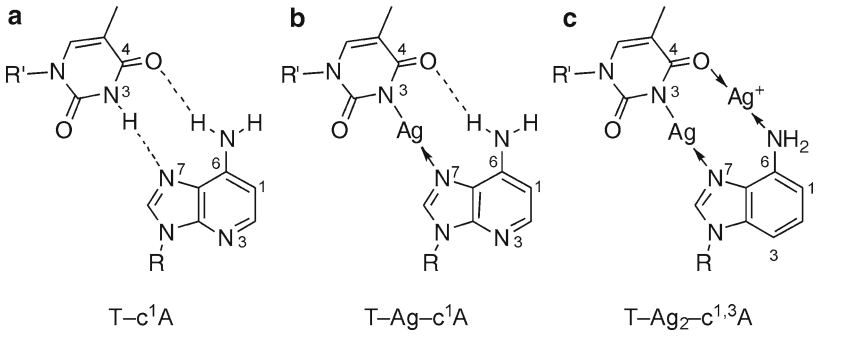
4 Metal-Mediated Base Pairs with Purine-Derived Nucleobases

4.1 1-Deazaadenine and 1,3-Dideazaadenine

Metal-mediated base pairs containing 1-deazaadenine (c^1A) or 1,3-dideazaadenine ($c^{1,3}A$) are special inasmuch as they adopt a Hoogsteen geometry. As the purine bases are defunctionalized at their Watson-Crick edge, the base pairs cannot be

incorporated into a Watson-Crick duplex. What makes them special is that they have a thermal stability comparable to that of natural base pairs [60], contrary to most other metal-mediated base pairs. This provides the opportunity to create DNA duplexes containing a large number of metal-mediated base pairs without losing the reversibility of the hybridization process. In principle, metal-mediated Hoogsteen base pairs can also be formed from unmodified purine bases, but at current examples for such base pairs are limited to triple-helical systems [61] or unusual left-handed DNA duplexes [62]. Hence, 1-deaza- and 1,3-dideazapurine derivatives are required for the exclusive formation of a metal-modified Hoogsteen double helix.

Hoogsteen duplexes from 1-deazaadenine and thymine adopt an established structural conformation in the absence of transition metal ions (Scheme 6a) [63]. In agreement with other metal-mediated base pairs formally derived by the substitution of an imino proton by a metal ion, Ag^+- mediated T–Ag–c^1A base pairs could be devised (Scheme 6b) [60]. For this, different oligonucleotides containing 1-deazaadenine were synthesized and investigated in the presence of complementary thymine-rich sequences and Ag^+. For an equimolar mixture of $d(T_{20})$ and $d(c^1A_{19}A)$ the specific binding of one metal ion per artificial base pair was observed, accompanied by an increase in melting temperature of $\Delta T_m > 36°C$ (approx. 2°C per base pair) [60]. As a result, T_m is still low enough to ensure a completely reversible duplex formation in aqueous solution. CD spectroscopic experiments revealed only marginal differences in the absence and presence of Ag^+, indicating that the insertion of a metal ion into the Hoogsteen base pair does not affect the overall duplex structure [60].



Scheme 6 Artificial Hoogsteen base pairs formed from (**a**) thymine and 1-deazaadenine (c^1A) [63]; (**b**) thyminate, Ag^+, and 1-deazaadenine [60]; (**c**) thyminate, Ag^+, and 1,3-dideazaadenine ($c^{1,3}$A) (R, R' = DNA backbone) [64].

The use of 1,3-dideazaadenine as an artificial nucleobase was investigated as well [64]. At first glance a reactivity similar to that of 1-deazaadenine could be expected, as both nucleobases comprise an identical Hoogsteen edge. However, the nucleobases show different base-pairing properties in the absence of transition metal ions, as 1,3-dideazaadenine does not form stable Hoogsteen base pairs with thymine [65]. A differential reactivity was also established in the presence of Ag^+. The self-complementary oligonucleotide sequences $d(c^{1,3}AT)_9$ and $d(c^{1,3}A_9T_9)$

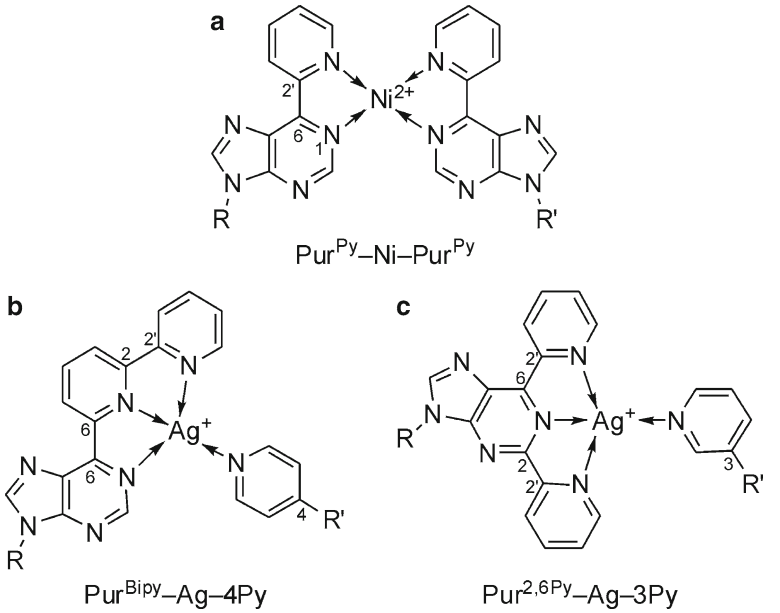
comprising 1,3-dideazaadenine ($c^{1,3}A$) form duplexes consisting of doubly Ag^+-mediated T–Ag_2–$c^{1,3}A$ base pairs (Scheme 6c). This can be explained by the more basic exocyclic amino group of 1,3-dideazaadenine as compared to 1-deazaadenine [66], enabling it to bind an additional metal ion. The sequence $d(c^{1,3}AT)_9$ with alternating purine and pyrimidine nucleosides (T_m = 82°C) was found to adopt a more stable duplex structure than $d(c^{1,3}A_9T_9)$ (T_m = 62°C) in the presence of Ag^+ [64]. The formation of stable metalated duplexes was further proven by dynamic light scattering revealing an increase in the hydrodynamic radius after the addition of Ag^+ and by mass spectrometric experiments. The doubly metalated base pair was investigated by dispersion-corrected DFT calculations, confirming that it is energetically favored over the corresponding non- and singly-metalated base pairs [64]. In the geometry-optimized structure of T–Ag_2–$c^{1,3}A$, the two Ag^+ ions are in close vicinity (2.88 Å), suggesting an argentophilic interaction [64]. This interaction contributes some 16 kcal mol^{-1} to the overall stability of the base pair. Its presence is supported by the fact that, upon excitation at 270 nm, the oligonucleotide duplexes exhibit fluorescence at ~465 nm in the presence of Ag^+ [64].

The base pairs formed from 1-deazaadenine and 1,3-dideazaadenine, respectively, represent another example for the potential that lies in the modification of the base pairing properties by subtly changing the nucleobases.

4.2 6-Substituted Purine Derivatives

In analogy to the pyrimidinone-derived nucleobases (Section 3.3.2), the affinity of purine bases towards metal ions can be enhanced by the functionalization with a heterocyclic aryl moiety, too. For example, a modified adenine derivative Pur^{Py} (6-(2'-pyridyl)-purine) containing a 2-pyridyl residue instead of an amino group at the C6 position has been reported (Scheme 7a) [67]. In oligonucleotides containing one self-pair of this artificial nucleoside, only Ni^{2+} was shown to form a stable metal-mediated base pair (Pur^{Py}–Ni–Pur^{Py}), whereas the addition of other metal ions (Co^{2+}, Cu^{2+}, Zn^{2+}, Ag^+, Fe^{2+}, Mn^{2+}, Eu^{3+}, Pd^{2+}) does not result in a stabilization of the duplex. Moreover, the choice of a different nucleobase complementary to Pur^{Py} (e.g., G, A, C, T) does not result in a remarkable duplex stabilization [67]. Quantum chemical calculations revealed an almost perfect overlap of the glycosidic bond positions of the artificial Pur^{Py}–Ni–Pur^{Py} and the natural A:T base pair within an antiparallel Watson-Crick duplex [67].

The attachment of one bipyridyl residue or two pyridyl moieties to a purine ring leads to terpyridine-like binding motifs. Accordingly, oligonucleotides containing the purine-derived nucleobases $Pur^{2,6Py}$ and Pur^{Bipy} (Scheme 7b, c) have been investigated [68]. In analogy to the bipyridine-functionalized pyrimidinone Pyr^{Bipy} (Figure 6b), these nucleobases form stable Ag^+-mediated base pairs with pyridine. Because of the different geometrical orientation of the terpyridine-like motif, $Pur^{2,6Py}$ preferentially forms metal-mediated base pairs with 3Py, whereas Pur^{Bipy} prefers 4Py [68].



Scheme 7 Schematic representation of (**a**) the Pur^{Py}–Ni–Pur^{Py} base pair [67], (**b**) the Pur^{Bipy}–Ag–4Py base pair [68], and (**c**) the $Pur^{2,6Py}$–Ag–3Py base pair [68] (R, R' = DNA backbone).

5 Processing of Metal-Mediated Base Pairs by Polymerases

As it is known that polymerases can recognize modified substrates and catalyze a replication reaction involving the formation of artificial base pairs, attempts have been made to directly incorporate metal-mediated base pairs into DNA duplexes by *in vitro* transcription in the presence of the appropriate metal ion. In the case of the T–Hg–T base pair, a primer-extension reaction using various DNA polymerases has been reported to take place, incorporating thymidine 5'-triphosphate (TTP) opposite to a thymine residue in the template strand in the presence of Hg^{2+} [69]. It has been proposed that this finding suggests a potential mechanism for the mutagenicity of Hg^{2+} [69]. As a result of the highly specific incorporation of Hg^{2+} into the T–Hg–T base pair, the addition of other metal ions like Mn^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}, Cu^{2+}, Zn^{2+}, Pb^{2+}, Ni^{2+}, or Au^{+} does not lead to the incorporation of TTP [69]. In another example, Taq polymerase has been shown to be capable of bypassing T–Hg–T and C–Ag–C base pairs in the primer region [70]. Hence, the amplification products are formed only in the presence of the appropriate metal ion. By applying this strategy, molecular logic gates have been devised that are triggered by Ag^{+} and Hg^{2+}, respectively [70]. Both studies indicate that metal-mediated base pairs with a geometry similar to that of the natural base pairs can be processed by polymerases, suggesting that an expansion of the genetic alphabet might be possible via a metal-mediated enzymatic incorporation of artificial nucleobases into nucleic acids [69].

6 Concluding Remarks and Outlook

Metal-mediated base pairs from natural nucleosides (or close derivatives thereof) have come a long way. Originally proposed to explain changes in the physical properties of natural DNA that was observed in the presence of certain transition metal ions, they are nowadays discussed and utilized in the context of sophisticated DNA-based applications such as metal ion sensors [71,72], sensors for the redox environment [41], molecular logic gates [70], the detection of single nucleotide polymorphisms [73], long-distance charge transfer [74], and DNA-based machines like bipedal walkers and steppers [75]. This wealth of possible applications will undoubtedly increase even further with the development of new metal-mediated base pairs including entirely artificial nucleosides [12].

Several studies have underlined the potential that lies in fine-tuning the metal-binding properties of nucleosides by replacing single atoms or entire groups of atoms not directly involved in the binding of the metal ions [54,64,76]. This approach will certainly play a major role in diversifying the range of nucleosides available for metal binding. Factors contributing to the stabilization of nucleic acid duplexes with metal-mediated base pairs, many of which display an increased positive charge along the helical axis, are only about to emerge. In this context, a non-covalent metallophilic attraction might play a significant role in stabilizing consecutive metal-mediated base pairs comprising metal ions with a closed-shell electronic structure [77]. The argentophilic interaction has already been shown to be a main contributor to the stability of doubly Ag^+-mediated base pairs [64].

Finally, the formation of metal-mediated base pairs is not limited to DNA but has also been found for RNA [43] and the nucleic acid analogues GNA (glycol nucleic acid) [78] and PNA (peptide nucleic acid) [79]. Hence, the possibility to combine particular properties of a nucleic acid with the functionality introduced by a metal-mediated base pair will enable the access to supramolecules that are structurally similar yet display chemically and physically distinct properties.

Abbreviations

C	cytosine, (deoxy)cytidine
c^1A	1-deazaadenine
$c^{1,3}A$	1,3-dideazaadenine
C–Ag–C	cytosine–Ag^+–cytosine
C–Ag–^{5M}iC	cytosine–Ag^+–5-methylisocytosine
CD	circular dichroism
DLS	dynamic light scattering
DOSY	diffusion-ordered spectroscopy
DFT	density functional theory
EDTA	ethylenediamine-*N,N,N',N'*-tetraacetate

FRET	fluorescence resonance energy transfer
G	guanine, (deoxy)guanosine
GNA	glycol nucleic acid
HSQC	heteronuclear single quantum coherence
^{5M}iC	5-methylisocytosine
1-MeC	1-methylcytosine
NOESY	nuclear Overhauser enhancement spectroscopy
PNA	peptide nucleic acid
$Pur^{2,6Py}$	2,6-di-(2'-pyridyl)-purine
Pur^{Bipy}	6-[6-(2,2'-bipyridyl)]-purine
Pur^{Py}	6-(2'-pyridyl)-purine
3Py	3-pyridyl
4Py	4-pyridyl
Pyr^{Bipy}	4-[6-(2,2'-bipyridyl)]-pyrimidinone
Pyr^{Py}	4-(2'-pyridyl)-pyrimidinone
r_H	hydrodynamic radius
T	thymine, thymidine
T–Hg–T	thyminate–Hg^{2+}–thyminate
TTP	thymidine 5'-triphosphate
U	uracil, uridine
^{5Br}U	5-bromouracil
^{5CN}U	5-cyanouracil
^{5F}U	5-fluorouracil
U–Hg–U	uracilate–Hg^{2+}–uracilate
UV	ultraviolet

Acknowledgment Financial support of our research by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

References

1. J. Müller, *Metallomics* **2010**, *2*, 318–327.
2. G. L. Eichhorn, Y. A. Shin, *J. Am. Chem. Soc*. **1968**, *90*, 7323–7328.
3. G. L. Eichhorn, J. J. Butzow, P. Clark, E. Tarien, *Biopolymers* **1967**, *5*, 283–296.
4. R. M. Izatt, J. J. Christensen, J. H. Rytting, *Chem. Rev*. **1971**, *71*, 439–481.
5. R. H. Jensen, N. Davidson, *Biopolymers* **1966**, *4*, 17–32.
6. S. Katz, *Nature* **1962**, *194*, 569.
7. J. Müller, *Eur. J. Inorg. Chem.* **2008**, 3749–3763.
8. G. H. Clever, C. Kaul, T. Carell, *Angew. Chem. Int. Ed*. **2007**, *46*, 6226–6236.
9. K. Tanaka, M. Shionoya, *Coord. Chem. Rev*. **2007**, *251*, 2732–2742.
10. S. Atwell, E. Meggers, G. Spraggon, P. G. Schultz, *J. Am. Chem. Soc*. **2001**, *123*, 12364–12367.
11. S. Johannsen, N. Megger, D. Böhme, R. K. O. Sigel, J. Müller, *Nat. Chem*. **2010**, *2*, 229–234.
12. G. H. Clever, M. Shionoya, Chapter 10 of this book.
13. J. S. Lee, L. J. P. Latimer, R. S. Reid, *Biochem. Cell Biol*. **1993**, *71*, 162–168.
14. P. Aich, S. L. Labiuk, L. W. Tari, L. J. T. Delbaere, W. J. Roesler, K. J. Falk, R. P. Steer, J. S. Lee, *J. Mol. Biol*. **1999**, *294*, 477–485.

15. B. Lippert, *Prog. Inorg. Chem*. **2005**, *54*, 385–447.
16. M. Fuentes-Cabrera, B. G. Sumpter, J. E. Šponer, J. Šponer, L. Petit, J. C. Wells, *J. Phys. Chem. B* **2007**, *111*, 870–879.
17. S. S. Alexandre, J. M. Soler, L. Seijo, F. Zamora, *Phys. Rev. B* **2006**, *73*, 205112.
18. E. C. Fusch, B. Lippert, *J. Am. Chem. Soc*. **1994**, *116*, 7204–7209.
19. S. L. Labiuk, L. T. J. Delbaere, J. S. Lee, *J. Biol. Inorg. Chem*. **2003**, *8*, 715–720.
20. A. Rakitin, P. Aich, C. Papadopoulos, Y. Kobzar, A. S. Vedeneev, J. S. Lee, J. M. Xu, *Phys. Rev. Lett*. **2001**, *86*, 3670–3673.
21. C.-Z. Li, Y.-T. Long, H.-B. Kraatz, J. S. Lee, *J. Phys. Chem. B* **2003**, *107*, 2291–2296.
22. M. J. Dinsmore, J. S. Lee, *J. Electroanal. Chem*. **2008**, *617*, 71–77.
23. X. Li, Y. Zhou, T. C. Sutherland, B. Baker, J. S. Lee, H.-B. Kraatz, *Anal. Chem*. **2005**, *77*, 5766–5769.
24. S. D. Wettig, D. O. Wood, P. Aich, J. S. Lee, *J. Inorg. Biochem*. **2005**, *99*, 2093–2101.
25. K. Mizoguchi, S. Tanaka, T. Ogawa, N. Shiobara, H. Sakamoto, *Phys. Rev. B* **2005**, *72*, 0033106.
26. F. Moreno-Herrero, P. Herrero, F. Moreno, J. Colchero, C. Gómez-Navarro, J. Gómez-Herrero, A. M. Baró, *Nanotechnology* **2003**, *14*, 128–133.
27. B. Liu, A. J. Bard, C.-Z. Li, H.-B. Kraatz, *J. Phys. Chem. B* **2005**, *109*, 5193–5198.
28. B. Q. Spring, R. M. Clegg, *J. Phys. Chem. B* **2007**, *111*, 10040–10052.
29. G. H. Clever, M. Shionoya, *Coord. Chem. Rev*. **2010**, *254*, 2391–2402.
30. S. Katz, *J. Am. Chem. Soc*. **1952**, *74*, 2238–2245.
31. C. A. Thomas, *J. Am. Chem. Soc*. **1954**, *76*, 6032–6034.
32. S. Katz, *Biochim. Biophys. Acta* **1963**, *68*, 240–253.
33. L. D. Kosturko, C. Folzer, R. F. Stewart, *Biochemistry* **1974**, *13*, 3949–3952.
34. Z. Kuklenyik, L. G. Marzilli, *Inorg. Chem*. **1996**, *35*, 5654–5662.
35. Y. Miyake, H. Togashi, M. Tashiro, H. Yamaguchi, S. Oda, M. Kudo, Y. Tanaka, Y. Kondo, R. Sawa, T. Fujimoto, T. Machinami, A. Ono, *J. Am. Chem. Soc*. **2006**, *128*, 2172–2173.
36. H. Torigoe, A. Ono, T. Kozasa, *Chem. Eur. J*. **2010**, *16*, 13218–13225.
37. Y. Tanaka, S. Oda, H. Yamaguchi, Y. Kondo, C. Kojima, A. Ono, *J. Am. Chem. Soc*. **2007**, *129*, 244–245.
38. A. Ono, H. Togashi, *Angew. Chem. Int. Ed*. **2004**, *43*, 4300–4302.
39. R.-M. Kong, X.-B. Zhang, L.-L. Zhang, X.-Y. Jin, S.-Y. Huan, G.-L. Shen, R.-Q. Yu, *Chem. Commun.* **2009**, 5633–5635.
40. C.-X. Tang, Y. Zhao, X.-W. He, X.-B. Yin, *Chem. Commun*. **2010**, *46*, 9022–9024.
41. Y. Miyake, A. Ono, *Tetrahedron Lett*. **2005**, *46*, 2441–2443.
42. E. Ennifar, P. Walter, P. Dumas, *Nucleic Acids Res*. **2003**, *31*, 2671–2682.
43. S. Johannsen, S. Paulus, N. Düpre, J. Müller, R. K. O. Sigel, *J. Inorg. Biochem*. **2008**, *102*, 1141–1151.
44. T. Yamane, N. Davidson, *Biochim. Biophys. Acta* **1962**, *55*, 609–621.
45. L. G. Marzilli, T. J. Kistenmacher, M. Rossi, *J. Am. Chem. Soc*. **1977**, *99*, 2797–2798.
46. T. J. Kistenmacher, M. Rossi, L. G. Marzilli, *Inorg. Chem*. **1979**, *18*, 240–244.
47. A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. Machinami, S. Oda, Y. Miyake, I. Okamoto, Y. Tanaka, *Chem. Commun.* **2008**, 4825–4827.
48. H. Torigoe, Y. Miyakawa, A. Ono, T. Kozasa, *Nucleosides Nucleotides Nucleic Acids* **2011**, *30*, 149–167.
49. D. A. Megger, J. Müller, *Nucleosides Nucleotides Nucleic Acids* **2010**, *29*, 27–38.
50. H. Urata, E. Yamaguchi, Y. Nakamura, S.-i. Wada, *Chem. Commun*. **2011**, *47*, 941–943.
51. T. Ono, K. Yoshida, Y. Saotome, R. Sakabe, I. Okamoto, A. Ono, *Chem. Commun*. **2011**, *47*, 1542–1544.
52. D. A. Megger, C. Fonseca Guerra, F. M. Bickelhaupt, J. Müller, *J. Inorg. Biochem*. **2011**, *105*, 1398–1404.
53. Y. Wen, F. Xing, S. He, S. Song, L. Wang, Y. Long, D. Li, C. Fan, *Chem. Commun*. **2010**, *46*, 2596–2598.

54. I. Okamoto, K. Iwamoto, Y. Watanabe, Y. Miyake, A. Ono, *Angew. Chem. Int. Ed*. **2009**, *48*, 1648–1651.
55. K. Aoki, W. Saenger, *Acta Cryst*. **1984**, *C40*, 775–778.
56. B. Lippert, *Coord. Chem. Rev*. **2000**, *200–202*, 487–516.
57. J. Ruiz, M. D. Villa, V. Rodríguez, N. Cutillas, C. Vicente, G. López, D. Bautista, *Inorg. Chem*. **2007**, *46*, 5448–5449.
58. C. Switzer, D. Shin, *Chem. Commun.* **2005**, 1342–1344.
59. D. Shin, C. Switzer, *Chem. Commun.* **2007**, 4401–4403.
60. F.-A. Polonius, J. Müller, *Angew. Chem. Int. Ed*. **2007**, *46*, 5602–5604.
61. T. Ihara, T. Ishii, N. Araki, A. W. Wilson, A. Jyo, *J. Am. Chem. Soc*. **2009**, *131*, 3826–3827.
62. M. G. Santangelo, P. M. Antoni, B. Spingler, G. Jeschke, *ChemPhysChem* **2010**, *11*, 599–606.
63. F. Seela, T. Wenzel, *Helv. Chim. Acta* **1994**, *77*, 1485–1499.
64. D. A. Megger, C. Fonseca Guerra, J. Hoffmann, B. Brutschy, F. M. Bickelhaupt, J. Müller, *Chem. Eur. J.* **2011**, *17*, 6533–6544.
65. F. Seela, T. Wenzel, *Helv. Chim. Acta* **1995**, *78*, 833–846.
66. L. E. Kapinos, A. Holý, J. Günter, H. Sigel, *Inorg. Chem*. **2001**, *40*, 2500–2508.
67. C. Switzer, S. Sinha, P. H. Kim, B. D. Heuberger, *Angew. Chem. Int. Ed*. **2005**, *44*, 1529–1532.
68. B. D. Heuberger, D. Shin, C. Switzer, *Org. Lett*. **2008**, *10*, 1091–1094.
69. H. Urata, E. Yamaguchi, T. Funai, Y. Matsumura, S.-i. Wada, *Angew. Chem. Int. Ed*. **2010**, *49*, 6516–6519.
70. K. S. Park, C. Jung, H. G. Park, *Angew. Chem. Int. Ed*. **2010**, *49*, 9757–9760.
71. R. Freeman, T. Finder, I. Willner, *Angew. Chem. Int. Ed*. **2009**, *48*, 7818–7821.
72. N. Kanayama, T. Takarada, M. Maeda, *Chem. Commun*. **2011**, *47*, 2077–2079.
73. H. Torigoe, T. Kozasa, A. Ono, *Nucleic Acids Symp. Ser*. **2006**, *50*, 89–90.
74. J. Joseph, G. B. Schuster, *Org. Lett*. **2007**, *9*, 1843–1846.
75. Z.-G. Wang, J. Elbaz, I. Willner, *Nano Lett*. **2011**, *11*, 304–309.
76. J. Müller, D. Böhme, P. Lax, M. Morell Cerdà, M. Roitzsch, *Chem. Eur. J*. **2005**, *11*, 6246–6253.
77. L. Benda, M. Straka, Y. Tanaka, V. Sychrovský, *Phys. Chem. Chem. Phys*. **2011**, *13*, 100–103.
78. M. K. Schlegel, L. Zhang, N. Pagano, E. Meggers, *Org. Biomol. Chem*. **2009**, *7*, 476–482.
79. R. M. Franzini, R. M. Watson, G. K. Patra, R. M. Breece, D. L. Tierney, M. P. Hendrich, C. Achim, *Inorg. Chem*. **2006**, *45*, 9798–9811.

Chapter 12
Metal Complex Derivatives of Peptide Nucleic Acids (PNA)

Roland Krämer and Andriy Mokhir

Contents

Abstract Peptide nucleic acid (PNA) is a non-cyclic pseudopeptide-nucleic acid structural mimic with promising applications within diagnostics and drug discovery. This review focuses on metal complex derivatives of PNA. Metal ions and their complexes display unique physical and chemical properties and offer the opportunity to introduce new labels and probes for bioanalytical and diagnostic applications of PNA, but also to modulate or to introduce new (for example catalytic) functions and biological activities.

R. Krämer (✉) • A. Mokhir
Inorganic Chemistry Institute, University of Heidelberg, Im Neuenheimer Feld 270, D-69120 Heidelberg, Germany
e-mail: roland.kraemer@urz.uni-heidelberg.de; Andriy.Mokhir@urz.uni-heidelberg.de

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2_12,

Keywords cellular uptake • labels • metal complexes • nucleic acid detection • peptide nucleic acids

1 Introduction

Since their introduction about 20 years ago [1], peptide nucleic acids (PNAs) have caught the interest in many fields of science from pure chemistry, over (molecular) biology, drug discovery, and (genetic) diagnostics, to nanotechnology and prebiotic chemistry [2]. A very wide variety of PNA derivatives and PNA-type molecules have been synthesized, but most applications focus on the original aminoethylglycine PNA (Figure 1).

The latter is a non-cyclic pseudopeptide-nucleic acid structural mimic (Figure 1). PNA binds stronger to sequence complementary oligonucleotides than does natural DNA or RNA because there is no electrostatic repulsion between strands. Consequently, PNA effectively targets nucleic acids at variable ionic strength [3].

Beside the formation of duplexes with complementary single-stranded nucleic acids, an invasive binding mode with double stranded nucleic acids is observed for PNA. In particular, homopyrimidine PNA oligomers bind to sequence complementary homopurine targets in duplex DNA such that one PNA oligomer hybridizes the complementary DNA strand by Watson–Crick base pairing, while a second PNA strand associates in a Hoogsteen triplex mode to this newly formed PNA–DNA duplex, thus leaving the homopyrimidine DNA strand of the original DNA duplex as a single-stranded loop [4,5] (Figure 2). In addition, other PNA binding modes have been discovered and exploited for various applications.

In living cells, PNA displays antisense activity since it can bind to complementary sequences of mRNA and can modulate its function. PNA can also act as antigene agent because it can break up a DNA duplex. In contrast to unmodified oligo-RNA

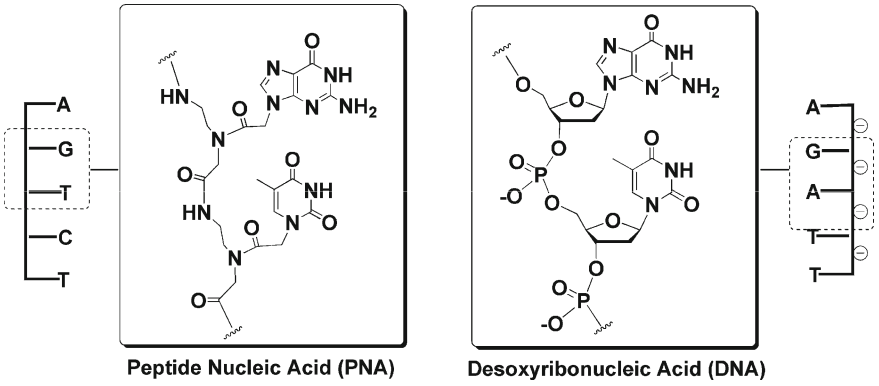


Figure 1 Structures of aminoethylglycine peptide nucleic acid (PNA) and desoxyribonucleic acid (DNA).

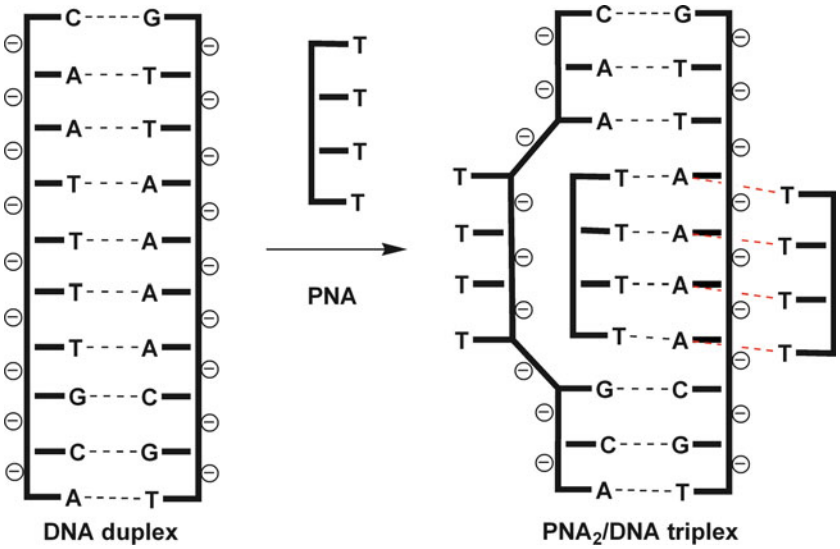


Figure 2 Duplex invasion of homopurine ds-DNA sequence by homopyrimidine PNA.

(and DNA), which are degraded by nucleases in living cells, PNA is not a substrate of nucleases nor peptidases and is thus resistant to hydrolytic degradation and rather stable within cells and organisms. Alternative splicing correction is a particular attractive RNA-interference approach of PNA, an exciting demonstration being the treatment of Duchenne muscular dystrophy by mRNA splicing correction in a mouse model [6].

There is little doubt that one of the biggest hurdles still to be solved in terms of PNA drug discovery is delivery and thus, *in vivo* bioavailability. PNA oligomers are not taken up by eukaryotic or bacterial cells in culture and show only very limited if any antisense-derived activity in animal experiments [7]. A large variety of methods have been introduced and applied to improve cellular bioavailability of PNAs, ranging from chemical modification, conjugation to cell-penetrating peptides, or complexation to physical or electrical temporal disruption of the cell membrane. However, the design of effective, robust and simple delivery methods remains a challenge [8].

PNA technology is still a very active area of research with promising applications within diagnostics and drug discovery. This review will focus on the metal complex derivatives of peptide nucleic acids. Metal ions and their complexes display a variety of unique physical and chemical properties. This offers the opportunity to introduce new labels and probes for bioanalytical and diagnostic applications of PNA, but also to modulate or to introduce new (for example catalytic) functions and biological activities. Exchange-inert metal complexes are typically coupled to PNA directly via a covalent linker. Alternatively, for more labile complexes, a ligand might be coupled to PNA and the metal ion added *in situ* to form the metal complex derivative [9,10].

2 Metal Ion Binding to Unmodified Peptide Nucleic Acid

It is well-known that the stability of DNA duplexes is critically dependent on high concentrations of Na^+ or K^+ ions which contribute to a shielding of the negative charges of the two strands from each other. In sharp contrast to the behavior of negatively charged DNA and RNA, PNA complexes show almost unaltered stability even at very low ionic strength; in fact, PNA-DNA duplexes are slightly stabilized with decreasing ionic strength [3].

Since *in situ* prepared PNA-metal complexes may contain excess free metal ions, we have examined the effect of selected transition metal ions Zn^{2+}, Cu^{2+}, and Ni^{2+} on unmodified PNA-DNA duplexes. At low micromolar concentration, these metal ions have little influence on the stability of PNA-DNA duplexes: At best a slight, up to 2°C, decrease of the melting point has been observed at 4 μM metal ions.

Kraatz and coworkers [11,12] have investigated the interaction of PNA and of PNA-DNA hybrids with selected metal ions at higher concentrations by electrochemical impedance spectroscopy. Interaction of Co, Ni, Zn or Mg at 400 μM concentration was observed with thin nucleic acid films on gold electrodes. Data for the PNA-DNA duplex are interpreted in terms of metal ion binding to both the polyanionic backbone and – in particular for Ni^{2+}– to the nucleobases [11]. Surprisingly, there is also evidence for interactions of the metal ions with the neutral polyamide backbone of PNA [12].

3 Effect of Metal Ions on Hybridization of Modified DNA

Hybridization of carefully designed chemically modified peptide nucleic acids can be both facilitated and inhibited by metal ions. Such metal-dependent PNAs can potentially find applications in analysis of, e.g., metal ions, nucleic acids, and proteins, in nanotechnology and medicine.

3.1 Terminally Modified Peptide Nucleic Acids

Duplexes of nucleic acids with peptide nucleic acids, which carry organic ligands on one of the termini, are stabilized in the presence of metal ions in accordance with three possible mechanisms (Figure 3). In all cases the role of the metal ion is to link the PNA strand to the nucleic acid via either coordinative or electrostatic interactions.

In particular, the groups of Spiccia and Graham have attempted to make use of a strong interaction of the Zn-cyclen complex (cyclen is 1,4,7,10-tetraazacyclododecane) with thymidine (dT) or uracil (U) to achieve Zn^{2+}-dependent binding of PNAs to ribonucleic acids via mode I (Figures 3 and 4) [13]. The major goal of the authors in this project was to develop PNAs, which bind more strongly to their complementary

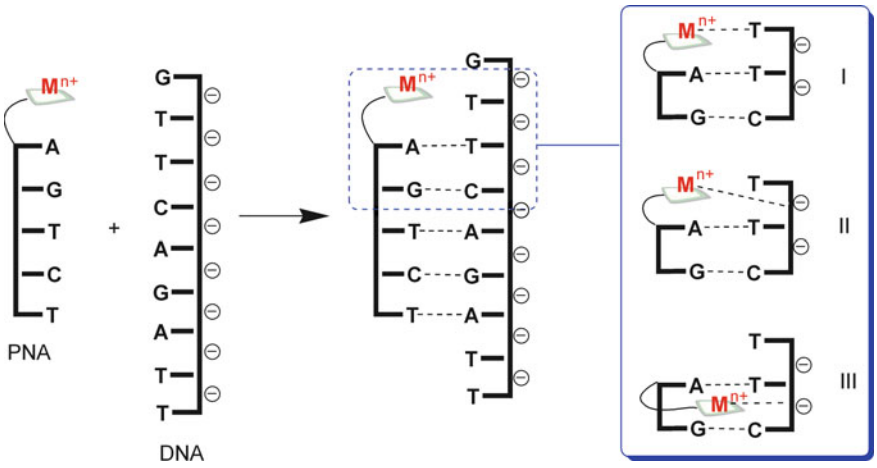


Figure 3 Possible modes of interaction of (metal complex)~PNA conjugates with nucleic acids (DNAs or RNAs).

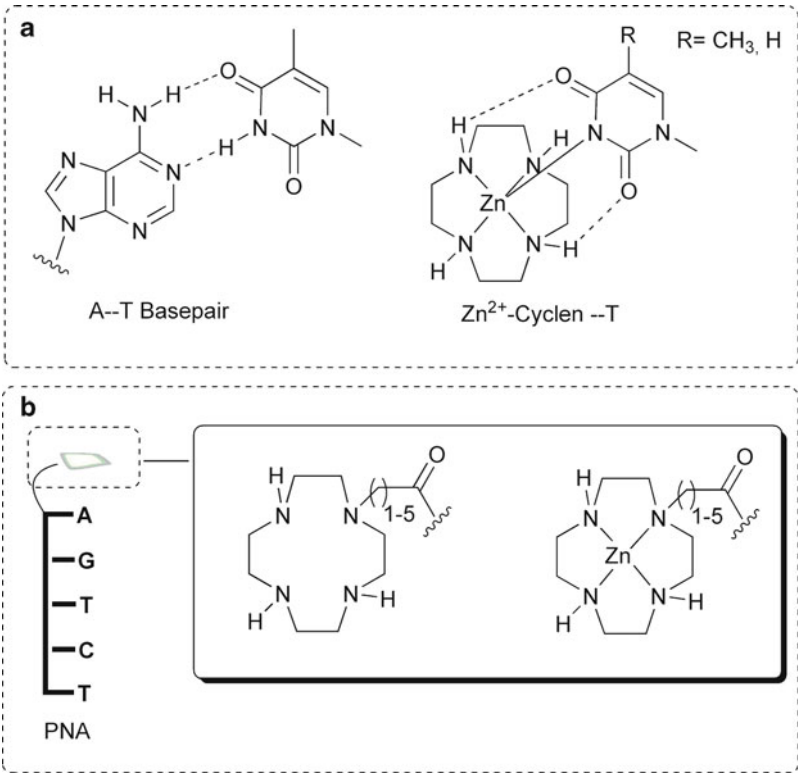


Figure 4 (**a**): Natural A-T base pair and its artificial counterpart, a cyclen-Zn–T complex; (**b**): conjugates of peptide nucleic acids with cyclen and a cyclen-Zn^{2+} complex, which were applied as binders of HIV-1 TAR mRNA.

RNA targets. Both dT–Zn-cyclen and U–Zn-cyclen complexes are strongly stabilized by the coordinative interaction of Zn^{2+} with the deprotonated amide of dT(U) and additionally by two hydrogen bonds between the secondary amino groups of the cyclen ligand and the carbonyl groups of dT (U) (Figure 4A). Therefore, these complexes are stable in aqueous solution and at physiological pH: log K_a = 5.6 (dT complex) and 5.2 (U complex) [14,15].

Spiccia, Graham, and their coworkers have prepared a series of N-cyclen~PNA conjugates having flexible linkers between the ligand and the PNA fragment and studied their binding to the stem-loop part of the HIV-1 TAR messenger RNA both in the absence and presence of Zn^{2+} by using UV melting temperature measurements and native gel-shift assays [13]. They have observed that, though cyclen~PNA conjugates were efficient binders of the mRNA, Zn^{2+} did not exhibit any significant effect on the stability of the corresponding cyclen~PNA/mRNA structures. Based on these data the authors have concluded that either the cyclen-Zn–U(RNA) complex was not formed due to the unfavorable orientation of the cyclen-Zn unit and the opposite U-nucleobase in the PNA/RNA structure or that the cyclen-Zn–U complex was formed, but could not be well accommodated within the PNA/RNA structure. The latter destabilizing effect would cancel out the stabilization due to the linking of the PNA and RNA via the coordinative cyclen-Zn–U interaction.

Mokhir and Krämer prepared a series of N-modified ligand~PNA conjugates and studied their hybridization to complementary DNA targets in the presence of kinetically labile metal ions: Ni^{2+}, Cu^{2+} and Zn^{2+} (Figure 5) [9]. They observed that binding of some of these PNA conjugates to their DNA targets was enhanced in the presence of Zn^{2+} and Ni^{2+}, whereas the effect of Cu^{2+} was substantially smaller. Cu^{2+} is generally a stronger binder than Ni^{2+} and Zn^{2+}. For example, it can bind to nucleobases of the PNA. Therefore, the weak duplex stabilizing effect of Cu^{2+} could be explained by formation of alternative structures, e.g., complexes in which the PNA itself saturates all coordination sites at the copper ion via coordination of the ligand and donor atoms of the nucleobases. These interactions would stabilize the PNA in its single-stranded form and therefore, destabilize the duplex.

It was observed that the change of the PNA charge upon complexation of the metal ion is an important factor, which determines the degree of the metal-induced stabilization of the PNA/DNA duplexes. In particular, hybridization of the ligand~PNA conjugates containing neutral ligands (e.g., bis(picolyl)amine) was found to be especially strongly affected by Ni^{2+} and Zn^{2+}. In this case the change of the charge was maximal: +2 [9].

Mokhir and Krämer have found that the metal-induced stabilization of the PNA/DNA duplexes is independent from the overhang sequence of the DNA target. Therefore, they have concluded that the metal ions do not interact directly with the nucleobases of the DNA as in mode I, but rather as in mode II, they bind to the phosphodiester group of the DNA backbone via electrostatic interactions (see Figures 3 and 6). This binding mode was confirmed by studies of the duplex stabilities at different salt concentrations. In particular, at high salt conditions, which are expected to reduce the electrostatic interactions, the metal-induced stabilization was substantially reduced.

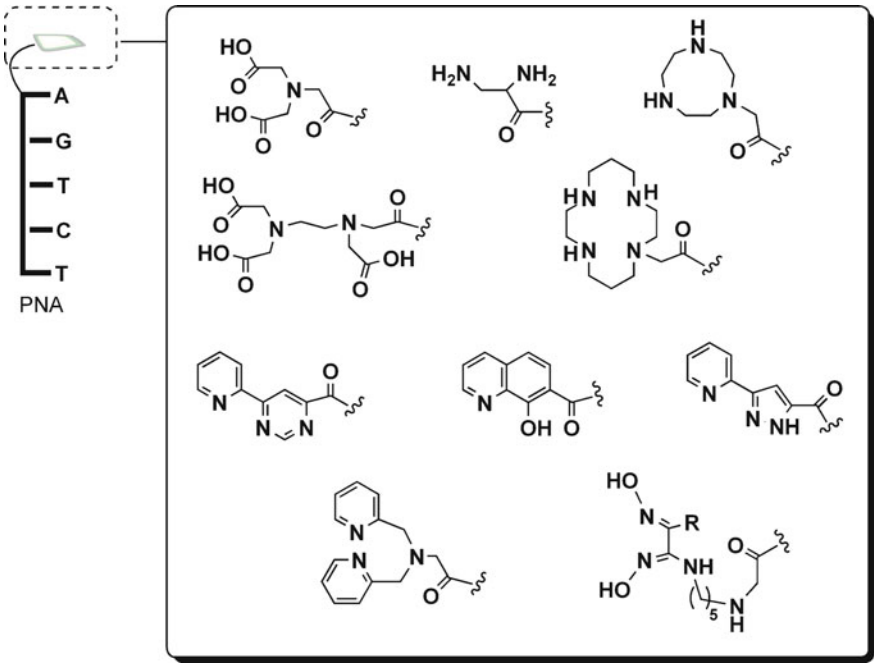


Figure 5 Peptide nucleic acids N-terminally modified with a series of ligands for kinetically labile metal ions: Ni^{2+}, Cu^{2+} and Zn^{2+} [9].

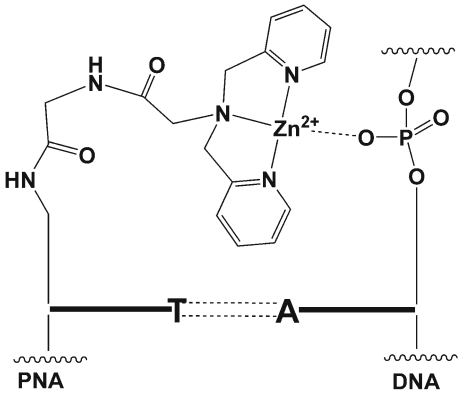


Figure 6 Zn^{2+}-mediated stabilization of PNA/DNA duplexes via electrostatic interaction between the metal ion and the phosphodiester group of the DNA target: mode II of Figure 3.

The same authors have also prepared a conjugate containing a bis(picolyl)amine ligand attached to the N-terminus of the PNA fragment via an aromatic linker naphthalenediimide derivative (NADI), which can intercalate within PNA/DNA duplexes (L~NADI~PNA) [16,17]. In the absence of Zn^{2+} the conjugate L~NADI~PNA is a poor binder of its DNA targets, whereas in the presence of the metal ions the stable PNA/DNA duplex is formed. This duplex is stabilized due to the intercalation of

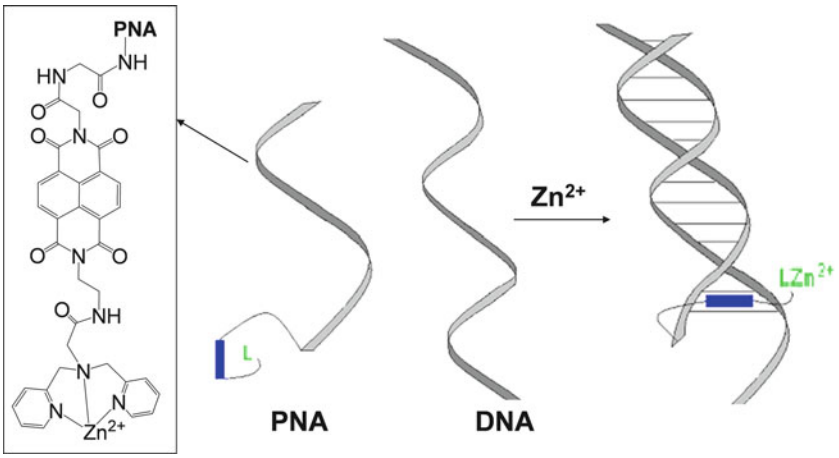


Figure 7 Zn^{2+}-mediated stabilization of PNA/DNA duplexes via combined electrostatic interaction between the metal ion and the phosphodiester group of the DNA target and intercalation of the naphthalenediimide fragment within the PNA/DNA duplex [16,17]. Reproduced from [10] with permission of Elsevier; copyright 2006.

NADI and the additional electrostatic interaction of the L-Zn^{2+} complex with the phosphodiester group of the DNA: mode III (see Figures 3 and 7).

The Zn^{2+}-dependent peptide nucleic acids described above can be potentially used for the design of antisense agents specific for the specialized cells, which contain elevated amounts of chelatable Zn^{2+} [10,17]. For example, $Zn^{2+}$is present at high concentrations in the brain [42], pancreas [43], and spermatozoa [44]. In breast cancer tissues zinc levels are increased by 72% [45]. It has been also found that the malaria parasite (*Plasmodium falciparum*), infecting red blood cells, accumulates Zn^{2+} [46]. In particular, concentration of free (chelatable) Zn^{2+} in these parasites is increased up to 50-fold in comparison with the host erythrocyte.

Hybridization of DNAs with PNAs, which contain metal-binding ligands on both termini, is expected to be inhibited in the presence of metal ions. This has been already demonstrated by Krämer and coworkers by the example of DNA/DNA hybridization (Figure 8) [18].

We can also envision that hybridization of PNAs containing poly-T stretches on their termini will be inhibited in the presence of Hg^{2+} ions. In this case, hairpin structures will be formed, which are stabilized by artificial T—Hg—T base pairs. An analogous mechanism was demonstrated for DNAs (Figure 9) [19].

3.2 Internally Modified Peptide Nucleic Acids

Internally modified peptide nucleic acids are mostly applied for the preparation of metal-containing DNA-based nanostructures and in the analysis of nucleic acids and metal ions.

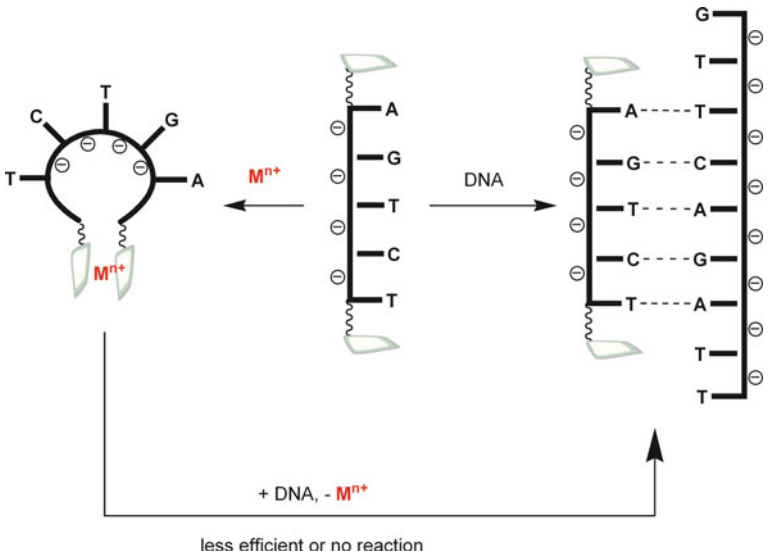


Figure 8 Metal ion-induced inhibition of DNA hybridization; analogous mechanism can potentially be applicable also for PNA hybridization [18].

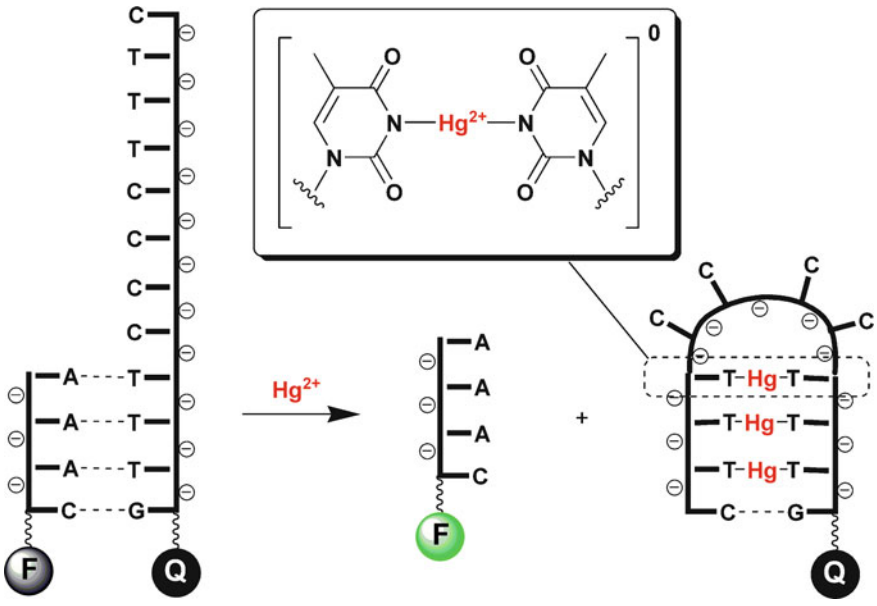


Figure 9 Hg^{2+}-induced decomposition of DNA/DNA duplexes (F = fluorophore; Q = quencher). In the initial duplex the fluorescence of the fluorophore is quenched; upon duplex decomposition F becomes fluorescent, which allows monitoring the Hg^{2+}-induced decomposition of the duplex by using fluorescence spectroscopy [19].

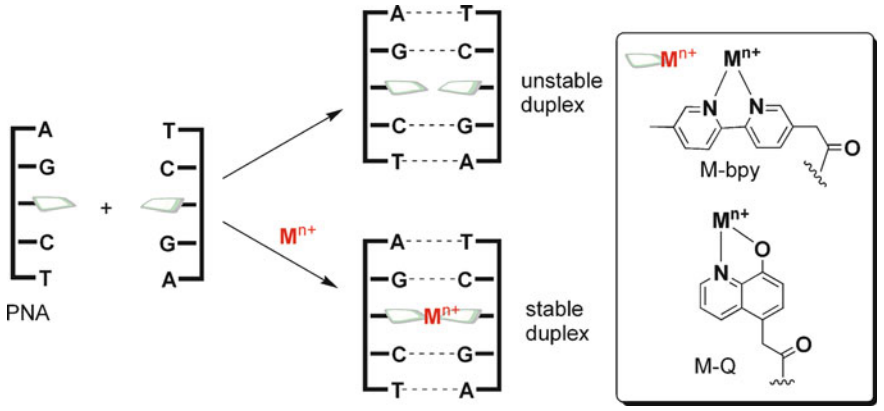


Figure 10 Formation of metal-containing PNA duplexes from PNA strands, which are internally modified with bidentate ligands.

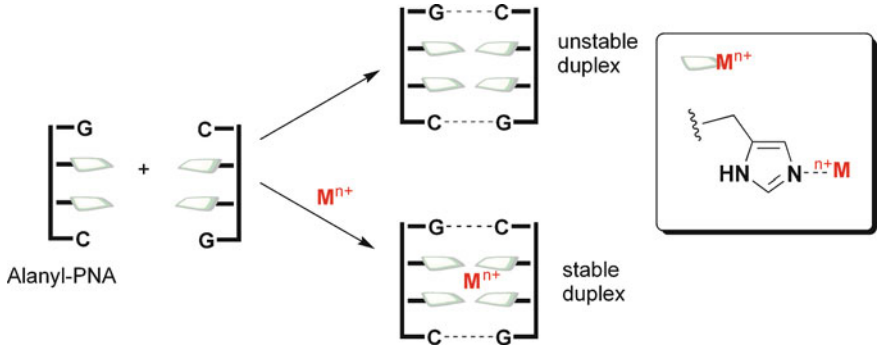


Figure 11 Self-assembly of histidine-modified, self-complementary alanyl-PNAs in the presence of Zn^{2+} ions; Cu^{2+} ions do not induce such a process [22].

Achim and coworkers have prepared PNAs containing either 2,2'-bipyridine (Bpy) or 8-hydroxyquinoline (HQ) ligands in their interior (Figure 10) [20,21].

Binding of complementary Bpy-containing PNAs to each other is strongly facilitated in the presence of equimolar concentrations of Ni^{2+}: the melting point (Tm) of the duplex is 48°C in the absence of the metal ion and 59°C in its presence. The formation of the metal-containing duplex has been also confirmed by UV-visible and CD spectroscopy.

In contrast, HQ~PNAs have been found to form preferably homoduplexes in the presence of Cu^{2+} ions. These structures exhibited high degree of mismatch toleration, i.e., even the duplexes with many mismatches were obtained due to formation of highly stable CuQ_2 complexes.

Diederichsen et al. have prepared self-complementary alanyl-PNAs with two internal histidine residues (Figure 11) [22]. These conjugates form stable duplexes

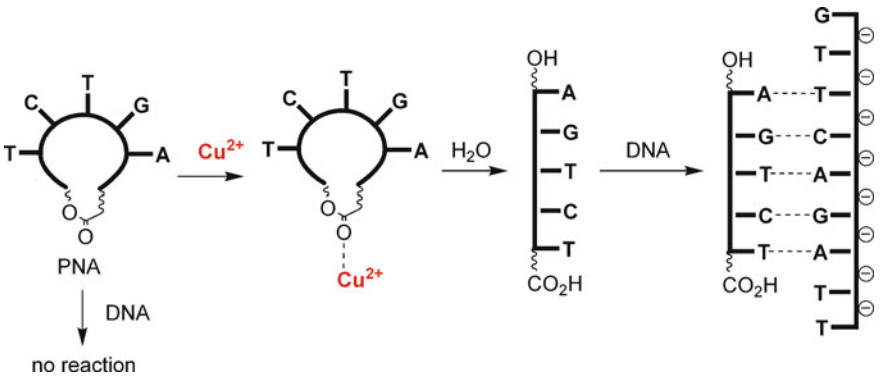


Figure 12 Cu^{2+}-controlled cyclic peptide nucleic acids: c-PNAs [23–25].

only in the presence of Zn^{2+} ions. It has been suggested that the duplex in this case is stabilized by formation of a ZnN4(His) complex in the interior of alanyl-PNA duplexes. In contrast, Cu^{2+} ions do not support the duplex formation. In accord with EXAFS data copper is coordinated via 2 N-containing ligands and 3 O-containing ligands in the alanyl-PNA/Cu^{2+} complex rather then via N-ligands from His. Thus, the 1:1 Cu/PNA complex is stabilized over the desired 1:2 Cu/PNA complex.

Mokhir and coworkers have reported on another type of metal-dependent, internally modified peptide nucleic acids: cyclic PNAs, c-PNA (Figure 12) [23–25]. Such c-PNAs contain an ester group in their backbone. This group is cleaved in the presence of Cu^{2+} ions. The cycle size in these PNAs has been optimized to obtain compounds, which are not able to hybridize to complementary nucleic acid targets. In the presence of Cu^{2+}, the ester group is cleaved and linear PNA is released. This compound binds nucleic acids efficiently. Thus, Cu^{2+} controls hybridization of c-PNAs. It has been demonstrated that c-PNAs can be applied for detection of Cu^{2+} ions [23] and nucleic acids [25].

3.3 Quadruplex Structures

G-quadruplex structures are self-assembled in solutions of G-rich DNAs in the presence of K^+ or Na^+ ions due to formation of hydrogen bonded, metal ion stabilized G-tetrads (Figure 13). Armitage and co-workers have applied CD spectroscopy and fluorescence resonance energy transfer (FRET) to demonstrate that hybrid PNA/DNA G-quandruplexes can be also obtained by simply mixing and annealing DNA and PNA strands with the sequence of *Oxytricha nova*telomeres: G4T4G4 [26]. They have observed that Na^+ ions were required for the formation of these poly-stranded structures, whereas Li^+ did not support their self-assembly.

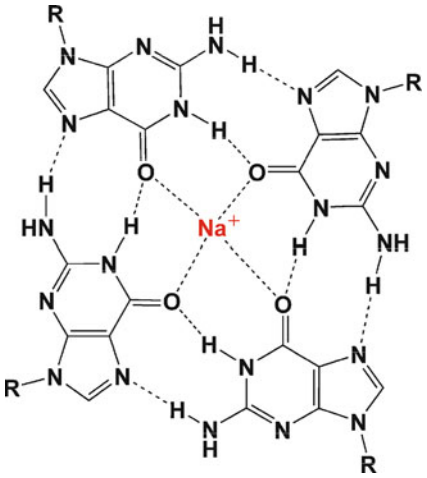


Figure 13 Structure of a G-tetrad, which is found in G-quadruplex structures made of DNA or hybrid structures made of DNA and PNA.

4 Labeling of Peptide Nucleic Acids with Metal Complexes

Labeling of peptide nucleic acids with metal complexes allows introducing new functionalities within PNA. These modified PNAs can be applied for the detection of nucleic acids by using electrochemical approaches, optical or vibrational microscopy and radioactivity-based imaging.

4.1 Organometallic Labels

A number of PNA monomers containing organometallic labels have been reported. These compounds will not be described here in detail. We will concentrate only on the conjugates containing metal complexes bound to PNA oligomers (longer than 5-mers).

The groups of Metzler-Nolte and Gasser have attached a variety of ferrocene residues to termini and interior of PNAs by using either amide bond-forming [27, 28] or Cu^+-catalyzed "click" reactions [29,30]. These mono- and poly-modified PNAs can be potentially applied for the electrochemical detection of nucleic acids. The same group has described the approach for the conjugation of four different ferrocene derivatives to PNAs [31]. Each of these complexes exhibits a unique, well distinguishable electrochemical potential that makes them potentially suitable for the simultaneous detection of four different DNA sequences by using electrochemical methods. Two of these electrochemical probes have recently been applied to develop an assay for the parallel electrochemical detection of two DNA sequences (Figure 14) [32].

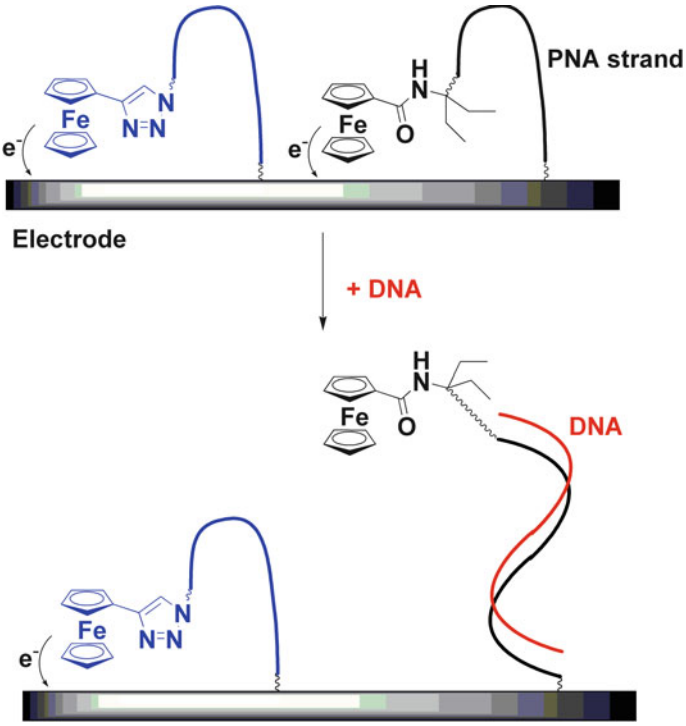


Figure 14 A principle of the electrochemical assay for the parallel detection of two different DNA sequences [32].

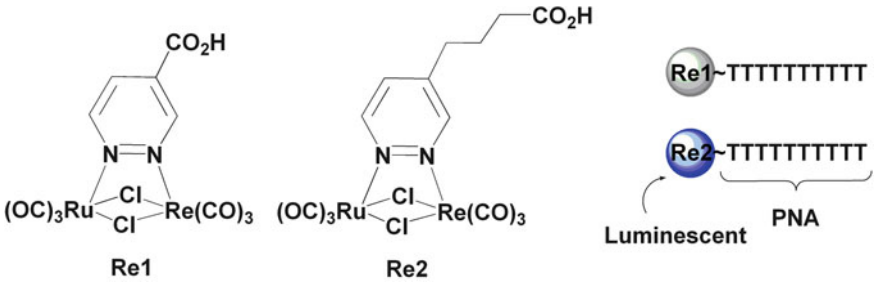


Figure 15 Structures of two luminescent Re^{+} complexes: Re1 and Re2, which have been conjugated to T10-PNA sequences; the Re2~T10 conjugate was successfully used for labeling the cytoplasm and nucleus of HEK293 cells [33].

The groups of D'Alfonso and Licandro have prepared conjugates of PNAs with luminescent, metallorganic, dinuclear complexes of Re^{+} containing 1,2-diazine, chloride and carbon monoxide ligands (Figure 15) [33]. Interestingly, conjugation of the complex Re1 to the T10-PNA fragment leads to almost complete quenching

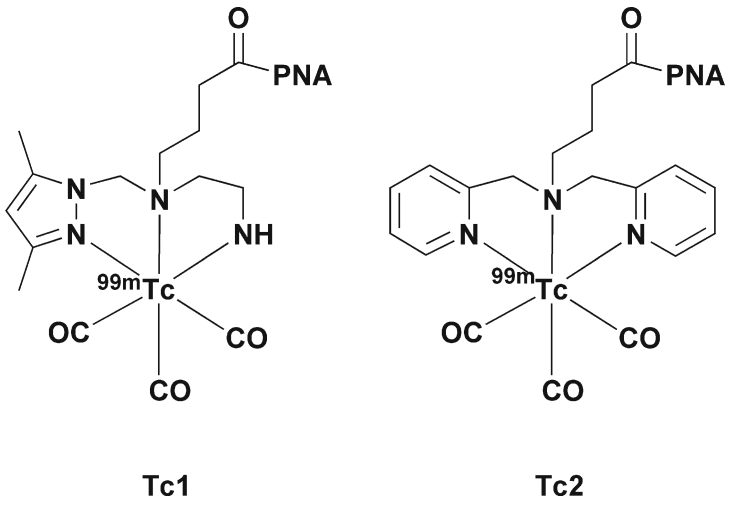


Figure 16 Structures of ^{99m}Tc-complexes, which are suitable for labeling peptide nucleic acids.

of its luminescence, whereas the luminescence of the closely related complex Re2 is practically not affected by the conjugation. The Re2~T10 conjugate is taken up by HEK293 cells and is accumulated both in the nucleus and in the cytoplasm. The emission of the dye in the nucleus is blue-shifted with respect to that one in the cytoplasm. Thus, both cellular compartments can be imaged simultaneously with a single probe.

4.2 Radioactive Labels

Peptide nucleic acids labeled with radioactive probes, e.g., metal complexes, are ideally suitable for non-invasive imaging of PNA distribution and localization *in vivo*. The information obtained from these data can be used to conclude about expression of specific genes. Mostly popular labels are complexes of ^{99m}Tc due to their favorable properties and availability. A number of ^{99m}Tc-labeled PNAs have been prepared and tested.

Synthetic work on the development of new radioactive complexes is usually conducted on Re analogues, since the chemical properties of Re and Tc are similar. After all optimizations are conducted, the ^{99m}Tc-complex is prepared in the minimal number of steps (ideally in 1 step). The complex formation is confirmed by using HPLC equipped with a highly sensitive radioactivity detector on the basis of comparison of retention factors of the Tc complex and the analogous Re complex. Structures of two recently developed ^{99m}Tc-complexes suitable for PNA labeling are presented in Figure 16 [34,35].

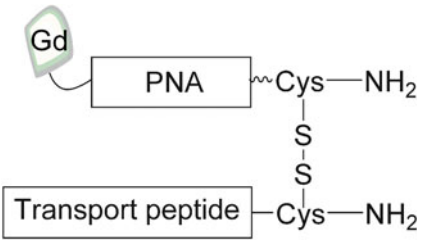


Figure 17 A conjugate of a peptide nucleic acid with a paramagnetic contrast agent, which is potentially suitable for the study of the PNA distribution *in vivo* by magnetic resonance imaging (MRI) [36].

4.3 Paramagnetic Labels

Magnetic resonance imaging (MRI) is another highly efficient method for studying the distribution of biomolecules *in vivo*. To visualize peptide nucleic acids in experimental animals and in humans by MRI, PNAs have to be linked to metal complexes, which act as paramagnetic contrast agents.

One report on the preparation and characterization of such a conjugate has been published by the group of Lehmann [36]. In particular, they have synthesized a multifunctional probe containing a PNA fragment as a recognition element of the c-*myc*gene, a gadolinium^{3+} complex with diethylenetriamine pentacetic acid (DTPA) as a paramagnetic contrast agent, and a membrane transport protein as a transporting unit (Figure 17). The resulting construct has been thoroughly characterized by using electrospray ionization (ESI) mass spectrometry and size exclusion chromatography (SEC) coupled to high resolution inductively coupled plasma mass spectrometry (ICP-MS). Application of this conjugate for MRI of PNAs *in vivo* has not been demonstrated yet.

5 Catalyst-Modified Peptide Nucleic Acids for Amplified Detection of Nucleic Acids

Sequence-specific detection of nucleic acids is a crucial research aspect in biological sciences and it is becoming increasingly important in diagnostics and genomics. Since the amount of nucleic acids in biological samples is limited, a hybridization event is usually used as a trigger of some catalytic reaction to amplify a nucleic acid target.

Mokhir, Krämer, and coworkers have developed a DNA-templated catalytic reaction in which a metal complex~peptide nucleic acid conjugate is the catalytic component (PNA-L-Cu^{2+}) (Figure 18) [37,38]. Another component of the reaction is an ester substrate attached to a short PNA (S-linker-PNA). S-linker-PNA and

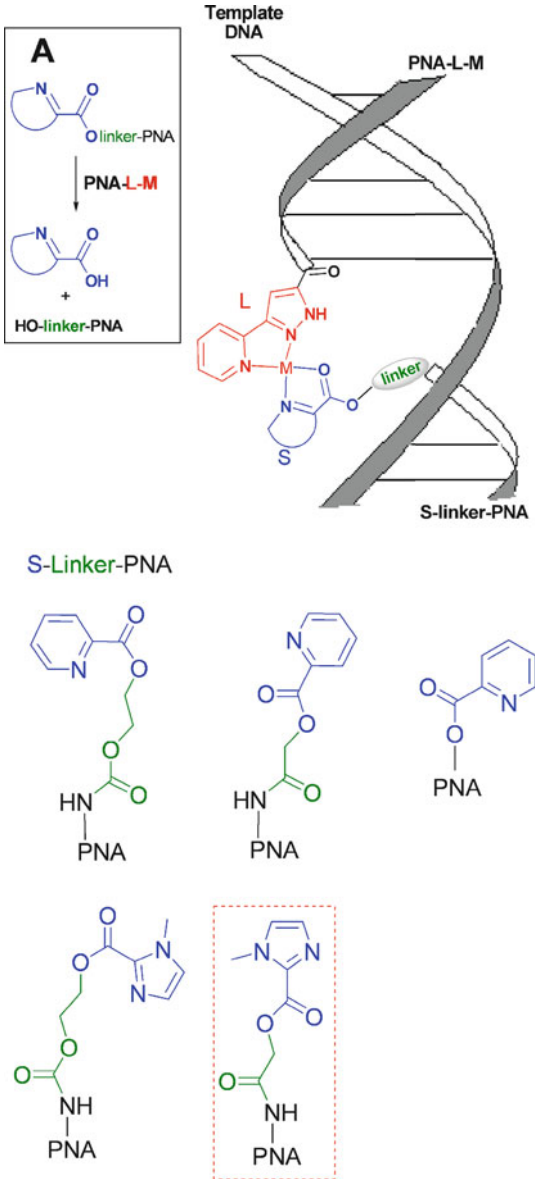


Figure 18 Structure of a complex S:M-L formed upon mixing S-linker-PNA, PNA-L-M, and a template DNA (M = Cu^{2+}). DNA-templated ester hydrolysis catalyzed by the metal complex is shown in insert A [37,38]. Reproduced from [38] with permission from the American Chemical Society; copyright 2005.

PNA-L-Cu^{2+} bind neighboring sites on a template nucleic acid, which brings the reacting groups (S and L-Cu^{2+}) in proximity to each other and accelerates the hydrolysis of the substrate ester (S). Structures of substrates (S-linker-PNA) prepared are shown in Figure 18. Among all studied S-linker-PNAs the best selectivity (V_0^{cat} / V_0^{uncat}) in the templated hydrolysis reaction was observed for the PNA, which contains the N-methyl-2-imidazolyl fragment (shown in the red box in Figure 18).

This reaction was applied for the amplified detection of single-stranded (Figure 18) and double-stranded nucleic acids by using MALDI-TOF mass spectrometry [38,39], which is a highly sensitive and mild method of analysis of short DNA and PNA probes. The time-of-flight (TOF) detector provides excellent resolution. Monocharged ions are usually detected, while peaks corresponding to doubly charged species as well as gas phase monocharged dimers have very low intensity.

The lowest DNA amount which could be detected by using the nucleic acid templated catalytic reaction was found to be 10 fmol.

6 Metal Ion-Triggered Cellular Uptake of Peptide Nucleic Acids

Fast progress in the application of PNA is still hampered by its poor bioavailability. Chemical modification of PNA may effectively enhance cellular uptake, but the main uptake route is often endosomal. Reagents thus have to escape the endosomal compartment, in order to enter the cellular compartments of action: the cytoplasm and/or the nucleus. Adding yet another level of complexity to the problem of cellular delivery is the fact that the real challenge in view of a pharmaceutical application is to deliver the compounds to cells in an organ.

Krämer and coworkers have observed that the biometal zinc(II) enhances cellular uptake of PNA when the latter is modified with the chelator 2,2':6',6''-terpyridine (tpy) [40]. Zinc(II) is a metal ion with highly tissue-specific biodistribution. Zinc-rich tissues include pancreas, testis, certain cancer types, and in particular the central nervous system [42].

Figure 19 compares the uptake by Hela cells of PNA of varying length, sequence and modification in the absence and presence of Zn^{2+} ions. Coordination of Zn^{2+} to the terpyridine moiety *in vitro* is confirmed by spectrophotometry. Cellular uptake was monitored by flow cytometry, using PNA samples labeled with a rhodamine (and other) fluorophores.

Obviously, attachment of terpyridine modification alone significantly enhances cellular uptake. Another up to 5,5-fold enhancement is found in the presence of Zn^{2+}. Confocal microscopy indicates – independently of the presence or absence of Zn^{2+}– strong enrichment of PNA in the nucleus (Figure 20) [40]. The lack of cellular import at 4°C may indicate an active transport process. The reason for enhanced uptake of the Zn-bound probe is not clear. Possibly, Zn coordination alters the conformation of the probe, making it a more potent trigger of receptor-mediated processes.

In a related series of experiments, terpyridine-PNA was further modified by dianionic (and potentially chelating) iminodiacetate (IDA) moiety. To reduce synthetic complexity in the preparation of the triply (tpy, IDA, fluorophore) functionalized molecule, an IDA-substituted coumarin fluorophore was introduced. IDA-PNA-tpy displayed very low cellular uptake, possibly due to increased hydrophilicity and charge repulsion with anionic cell surface components. IDA-PNA-tpy is a strong Zn^{2+} binder *in vitro* that potentially forms a tpy-IDA bis-chelate with the metal ion.

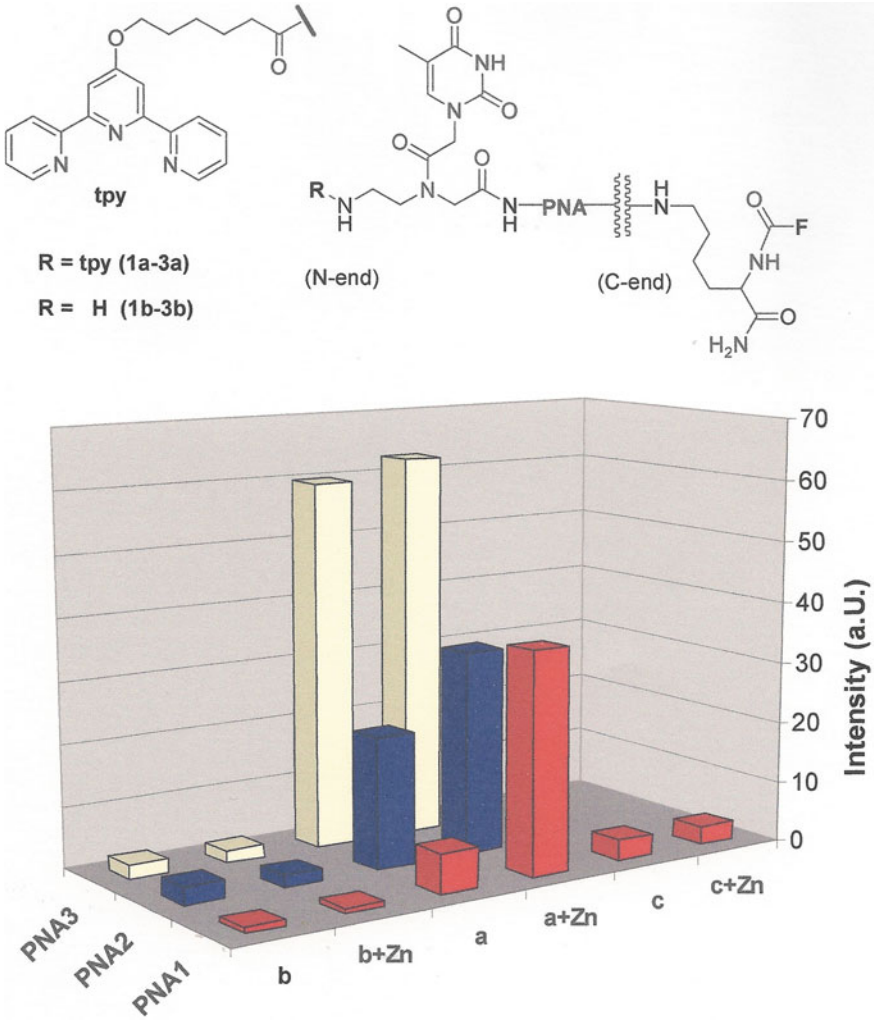


Figure 19 Effect of extracellular Zn^{2+} on the uptake of various PNAs by Hela cells, quantified by flow cytometry.
PNA 1a (N) tpy-TCACAACTAkkk-F_1(C);
PNA 1b (N) TCACAACTAkkk-F_1(C)
PNA 1c (N) F_1-TCACAACTAkkkk (C)
PNA 1d-e (N) tpy-TCACAACTAkkk-$F_{2\text{-}3}$(C);
PNA 2a (N) tpy-TACACAACTkkk-F_1(C);
PNA 2b (N) TACACAACTkkk-F_1(C)
PNA 3a (N) tpy-TCCTCGCCCTTGCTCACCATkkk-F_1(C)
PNA 3b (N) TCCTCGCCCTTGCTCACCATkkk-F_1(C)
F_1= tetratmethylrhodamine, F_2= Nile red, F_3= coumarin 343, k = lysine. Reproduced from [40] with permission from the American Chemical Society; copyright 2006.

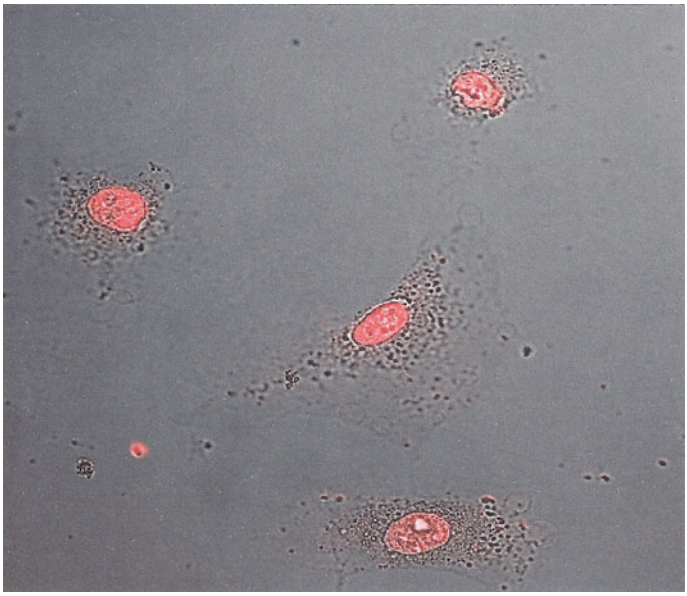


Figure 20 Confocal microscopy of living Hela cells after incubation with PNA 2a in the presence of Zn (overlay of DIC and rhodamine images). PNA 2a strongly accumulates in the nucleus. Reproduced from [40] with permission from the American Chemical Society; copyright 2006.

Remarkably, extracellular Zn^{2+} (2,5 μM) increases cell uptake of IDA-PNA-tpy up to 36-fold (Figure 21) [41]. Again, the probe enriches in the nucleus.

The effect is dependent on both IDA and terpyridine modification. The control substances IDA-PNA that lack the terpyridine moiety show very little cellular accumulation both in the absence and presence of Zn^{2+}. The antisense effect of internalized IDA-PNA-tpy was explored in a standard luciferase assay. Hela cells were transfected with recombinant plasmids carrying the luciferase gene, and intracellular PNA binding to a complementary mRNA sequence triggered a splicing correction that led to the biosynthesis of the enzyme luciferase. The latter was quantified by a chemoluminescence assay. Again, IDA-PNA-tpy probes were compared in the absence and presence of extracellular Zn^{2+}, and a 2,5-fold enhancement of luciferase expression by Zn^{2+} was found for the same PNA sequence that is internalized 36-times more effectively by Zn^{2+} [41].

Clearly, extracellular Zn^{2+} triggers control of gene expression by modified PNA, but relative to Zn-enhanced uptake only with poor efficiency. *In vitro* experiments indicate slow release of IDA-tpy-PNA bound Zn even by strong chelators, combined with slow hybridization to complementary nucleic acids. This prototype of a zinc-responsive antisense reagent might be a too strong zinc binder so that antisense activity is not effectively restored, even in the zinc-sequestering intracellular medium.

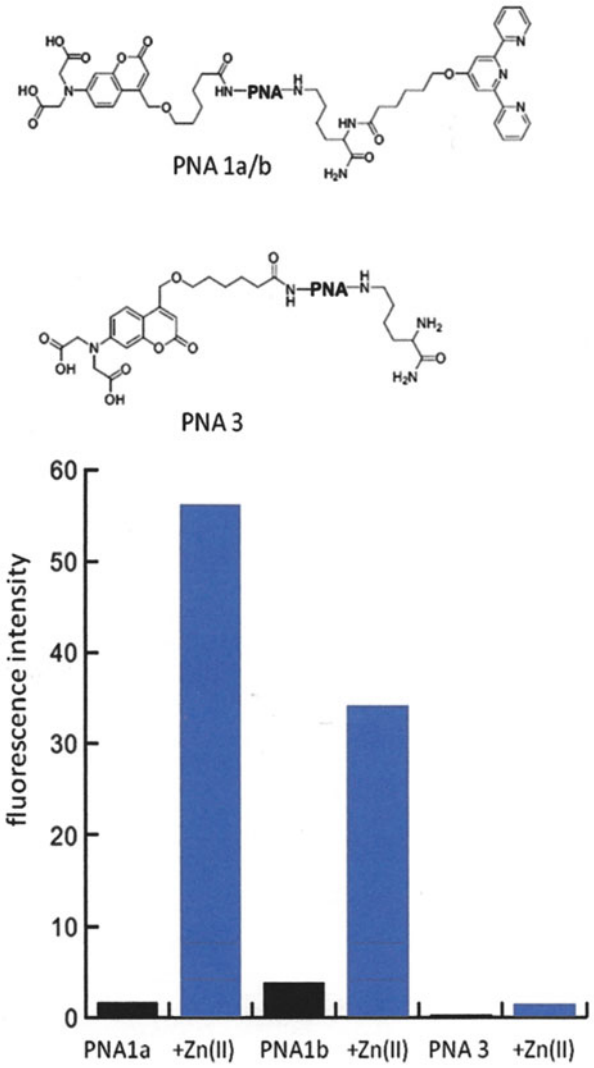


Figure 21 Effect of extracellular Zn^{2+} on the uptake of IDA-PNA-tpy (PNA 1a/b) and IDA-PNA (PNA 3) by Hela cells, quantified by flow cytometry. PNA 1a and 1b are both 18-mers but have different sequences. The sequence of PNA 3 is identical to that of PNA 1a.

Zinc-enhanced cellular delivery and gene expression is a first step toward a stimulus-controlled, tissue-selective uptake activity and of PNA reagents. While the level of free zinc in the extracellular space of tissues is generally very low, local concentrations in the micromolar range have been detected after stimulated efflux from specialized cells in, for example, brain and pancreas.

Abbreviations

Bpy	2,2'-bipyridine
CD	circular dichroism
c-PNA	cyclic peptide nucleic acid
cyclen	1,4,7,10-tetraazacyclododecane
DIC	differential interference contrast
dT	thymidine
DTPA	diethylenetriamine pentaacetic acid
ESI	electrospray ionization
EXAFS	extended X-ray absorption fine structure
FRET	fluorescence resonance energy transfer
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography
HQ	8-hydroxyquinoline
ICP-MS	inductively coupled plasma mass spectrometry
IDA	iminodiacetate
L	ligand
MALDI-TOF	matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry
MRI	magnetic resonance imaging
NADI	naphthalenediimide
PNA	peptide nucleic acid
SEC	size exclusion chromatography
TAR	trans activation responsive element
Tm	melting point
TOF	time-of-flight
tpy	2,2':6',6''-terpyridine
U	uracil

References

1. P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1499.
2. (a) P. E. Nielsen, *Chem. Biodiv.* **2010**, *7*, 786–804; (b) P. E. Nielsen, *Acc. Chem. Res.***1999**, *32*, 624.
3. M. Egholm, O. Buchardt, L. Christensen, S. Behrens, M. C. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–567.
4. P. E. Nielsen, M. Egholm, O. Buchardt, *J. Mol. Recognit*. **1994**, *7*, 165.
5. D. Y. Cherny, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, R. H. Berg, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1667.
6. H. Yin, Q. Lu, M. Wood, *Mol. Ther*. **2008**, *16*, 38.
7. P. Sazani, F. Gemignani, S.-H. Kang, M. A. Maier, M. Manoharan, M. Persmark, D. Bortner, R. Kole, *Nat. Biotechnol*. **2002**, *20*, 1228–1231.
8. P. E. Nielsen, *Gene Therapy*, **2005**, *12*, 956–957.

9. A. Mokhir, R. Stiebing, R. Krämer, *Bioorg. Med. Chem. Lett*. **2003**, *13*, 1399–1401.
10. I. Boll, L. Kovbasyuk, R. Krämer, T. Oeser, A. Mokhir, *Bioorg. Med. Chem. Lett*. **2006**, *16*, 2781–2785.
11. C. Li, X. Li, X. Liu, H.-B. Kraatz, *Anal. Chem*. **2010**, *82*, 1166–1169.
12. L. Gao, C. Li, X. Li, H.-B. Kraatz, *Chem. Commun*. **2010**, *46*, 6344–6346.
13. M. J. Belousoff, G. Gasser, B. Graham, Y. Tor, L. Spiccia, *J. Biol. Inorg. Chem*. **2009**, *14*, 287–300.
14. M. Shionoya, E. Kimura, M. Shiro, *J. Am. Chem. Soc*. **1993**, *115*, 6730–6737.
15. S. Aoki, E. Kimura, *Chem. Rev*. **2004**, *104*, 769–787.
16. A. Mokhir, R. Krämer, *Bioconj. Chem*. **2003**, *14*, 877–883
17. A. Mokhir, R. Krämer, H. Wolf, *J. Am. Chem. Soc*. **2004**, *126*, 6208–6209.
18. N. Graf, M. Göritz, R. Krämer, *Angew. Chem*. **2006**, *45*, 4013–4015.
19. Z. Wang, J. H. Lee, Y. Lu, *Chem. Comm.* **2008**, 6005–6007.
20. D.-L. Popescu, T. J. Parolin, C. Achim, *J. Am. Chem. Soc*. **2003**, *125*, 6345–6355.
21. R. M. Watson, Y. A. Skorik, G. K. Patra, C. Achim, *J. Am. Chem. Soc*. **2005**, *127*, 14628–14639.
22. A. Küsel, J. Zhang, M. A. Gil, A. C. Stückl, W. Meyer-Klaucke, F. Meyer, U. Diederichsen, *Eur. J. Inorg. Chem.* **2005**, 4317–4324.
23. J. Kovács, T. Rödler, A. Mokhir, *Angew. Chem., Int. Ed*. **2006**, *45*, 7815–7817.
24. J. Kovács, E. Jentzsch, A. Mokhir, *Inorg. Chem*. **2008**, *47*, 11965–11971.
25. J. Kovács, A. Mokhir, *Bioorg. Med. Chem. Lett*. **2008**, *18*, 5722–5724.
26. B. Datta, C. Schmitt, B. A. Armitage, *J. Am. Chem. Soc*. **2003**, *123*, 4111–4118.
27. J. C. Verheijen, G. A. van der Marel, J. H. van Boom, N. Metzler-Nolte, *Bioconj. Chem*. **2000**, *11*, 741–743.
28. A. Maurer, H.-B. Kraatz, N. Metzler-Nolte, *Eur. J. Inorg. Chem.* **2005**, 3207–3210.
29. G. Gasser, N. Hüsken, S. D. Köster, N. Metzler-Nolte, *Chem. Comm.* **2008**, 3675–3677.
30. A. M. Sosniak, G. Gasser, N. Metzler-Nolte, *Org. Biomol. Chem*. **2009**, *7*, 4992–5000.
31. N. Hüsken, G. Gasser, S. D. Köster, N. Metzler-Nolte, *Bioconj. Chem*. **2009**, *20*, 1578–1586.
32. N. Hüsken, M. Gebala, W. Schuhmann, N. Metzler-Nolte, *ChemBioChem*, **2010**, *11*, 1754.
33. E. Ferri, D. Donghi, M. Panigati, G. Prencipe, L. D'Alfonso, I. Zanoni, C. Baldoli, S. Maiorana, G. D'Alfonso, E. Licandro, *Chem. Comm*. **2010**, *46*, 6255–6257.
34. G. Gasser, K. Jäger, M. Zenker, R. Bergmann, J. Steinbach, H. Stephan, N. Metzler-Nolte, *J. Inorg. Biochem*. **2010**, *104*, 1133–1140.
35. C. Xavier, C. Giannini, L. Gano, S. Maiorana, R. Alberto, I. Santos, *J. Biol. Inorg. Chem*. **2008**, *13*, 1335–1344.
36. R. Krüger, K. Braun, R. Pipkorn, W. D. Lehmann, *J. Anal. Atom. Spectr*. **2004**, *19*, 852–857.
37. J. Brunner, A. Mokhir, R. Krämer, *J. Am. Chem. Soc*. **2003**, *125*, 12410–12411.
38. I. Boll, R. Krämer, J. Brunner, A. Mokhir, *J. Am. Chem. Soc*. **2005**, *127*, 7849–7856.
39. I. Boll, E. Jentzsch, R. Krämer, A. Mokhir, *Chem. Comm.* **2006**, 3447–3449.
40. A. Fuessl, A. Schleifenbaum, M. Goeritz, A. Riddell, C. Schultz, R. Krämer, *J. Am. Chem. Soc*. **2006**, *128*, 5986–5987.
41. A. Fuessl, Ph.D. Dissertation, Universität Heidelberg, **2007**
42. (a) C. J. Frederickson, *Int. Rev. Neurobiol.* **1989**, 31, 145–238; (b) M. P. Cuajungco, G. J. Lees, *Neurobiol. Dis*. **1997**, 4, 137–169. (c) C. J. Frederickson, S. W. Suh, D. Silva, C. J. Frederickson, R. B. Thompson, *J. Nutr.* **2000**, *130*, 1471S–1483S.
43. P. D. Zalewski, S. H. Millard, I. J. Forbes, O. Kapaniris, A. Slavotinek, W. H. Betts, A. D. Ward, S. F. Lincoln, I. J. Mahadevan, *Histochem. Cytochem*. **1994**, *42*, 877–884.
44. (a) P. D. Zalewski, X. Jian, L. L. Soon, W. G. Breed, R. F. Seamark, S. F. Lincoln, A. D. Ward, F. Z. Sun, *Reprod. Fertil. Dev*. **1996**, *8*, 1097–1105.
45. E. F. Margalioth, J. G. Schenker, M. Chevion, *Cancer* **1983**, *52*, 868–872.
46. J. Wolford, M. Kidd, J. E. Penner-Hahn, T. V. O'Halloran, ICBIC-12, 2005, http://www.u-mich.edu/~icbic/Program/All_Presentations.htm

Index

A. Sigel, H. Sigel, and R.K.O. Sigel, *Interplay between Metal Ions and Nucleic Acids*, Metal Ions in Life Sciences 10, DOI 10.1007/978-94-007-2172-2,