DNA-targeted metallodrugs: an untapped source for artificial gene editing technology

Nicolò Zuin Fantoni,[a] Tom Brown[a] and Andrew Kellett*[b]

[a] Dr. N. Zuin Fantoni, Prof. Dr. Tom Brown
Department of Chemistry, Chemistry Research Laboratory,
University of Oxford,
12 Mansfield Road, Oxford OX1 3TA, UK.
[b] Prof. Dr. A. Kellett
School of Chemical Sciences and National Institute for Cellular Biotechnology and Nano Research Facility
Dublin City University
Glasnevin, Dublin 9, Ireland
E-mail: andrew.kellett@dcu.ie

Abstract: DNA binding metal complexes are synonymous with anticancer drug discovery. Given the array of structural and chemical reactivity properties available through careful design, metal complexes have been directed to bind nucleic acid structures through covalent or non-covalent binding modes. Several recognition modes—including crosslinking, intercalation, and oxidation—are central to the clinical success of broad-spectrum anticancer metallodrugs. However, recent progress in nucleic acid click chemistry coupled with advancement in our understanding of metal complex-nucleic acid interactions has opened up new avenues in genetic engineering and targeted therapies. Several of these applications are enabled by the hybridisation of oligonucleotide or polyamine probes to discrete metal complexes which facilitate sitespecificreactivity at the nucleic acid interface under the guidance ofthe probe. This review focuses on recent advancements in hybrid design and, by way of an introduction to this topic, we provide a detailed overview of nucleic acid structures and metal complexnucleic acid interactions. Our aim is to provide readers with an insight on the rational design of metal complexes with DNA recognition properties and an understanding of how the sequence-specific targeting of these interactions can be achieved for gene engineering applications.

1. Introduction

The discovery of DNA and its role as the biological carrier of genetic information has revolutionised several disciplines and attracted interest in the design of specific probes that target its structure and alter its function.[1–3] Nucleic acids are polymers composed of repeating nucleobases linked together by a phospho(deoxy)ribose backbone. In duplex DNA, two of these polynucleotide strands coil around each other in an antiparallel fashion exploiting the definite recognition between nucleobases and yielding the double helical structure. Rigidity of the double helix is provided by the inter-chain hydrogen bond formation and π-π stacking interactions between successive nucleobases. DNA bases bond non-covalently according to Chargaff's rules whereby purines and pyrimidines pair in a consistent 1:1 ratio (adenine:thymine and guanine:cytosine).[4]

In nature, DNA can adopt three biologically-active double-helical structures—right-handed A- and B-DNA along with left-handed Z-DNA (Figure 1)—with B-DNA being the most commonly found in physiological environments. In B-DNA the 2′-deoxyribose ring is locked in the C2′-endo configuration and the π-π stacking interactions between nucleobases orient them perpendicular to the axis of the double helix.[5] The glycosidic bond is generally locked in the *anti*-configuration and its asymmetrical edges cause unequal grooves in the DNA duplex called major and minor groove.[6,7] In the A-form, the double helix is more compacted along the helical axis (due to the C3′ endo conformation that reduces intra-nucleobase distance) and has a wider diameter. A-DNA is usually formed in anhydrous environments, in some protein-DNA complexes and for RNA-based double helices (due to the steric restrictions imposed by the 2′-hydroxyl group on the ribose).[8] Z-DNA is promoted *in vivo* by the torsional strain induced during DNA under-wounding by processive enzymes (*i.e.* polymerases and helicases)[9,10] or *in vitro* by B-DNA treatment with high salt concentrations or with *di*-nuclear complexes binding the guanine (N7′)-position.[11] The Z-DNA duplex is left-handed and characterised by a single narrow groove similar to the B-DNA minor groove. The alternating sugar pucker (C2′-endo for pyrimidines and C3′ endo for purines) and glycosidic bond (*anti* for pyrimidines and *syn* for purines) provide a zig-zag arrangement to the backbone, hence its name Z-DNA.

The distinctive chemical and structural features of nucleic acids expose several possible sites of interaction with small molecule probes and polymeric sequences.[12] Firstly, the charge on the phosphodiester backbone provides an ideal binding site for positively charged compounds such as metal complexes, polycationic molecular probes or amino acid chains.[13] In addition, the hydrophobic environment provided by the nucleobases allows binding of molecular probes by non-covalent intermolecular interactions such as π-π stacking interactions between these bases afforded by planar aromatic ligands.[14,15] Interactions with

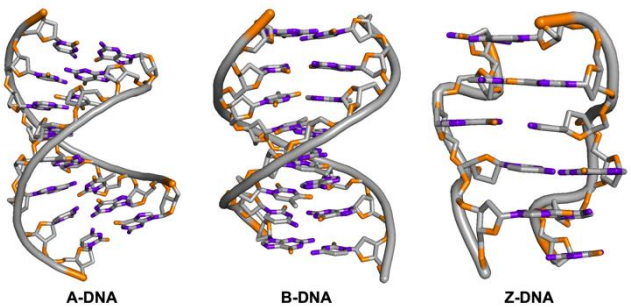

Figure. 1 Main three forms adopted by nucleic acids double helices: A-DNA (PDB: 1VJ4), B-DNA (PDB: 1BNA) and Z-DNA (PDB: 2DCG).

DNA bases can arise also through hydrogen bond with the nucleobase heteroatoms enabling carefully designed polymeric strands (*e.g*. triplex forming oligonucleotides or polyamide sequences) to recognise and bind specific genomic sequences.[16] DNA-binding by these probes often results in a modification of the DNA double helical structure whereby the degree is dependent on type and extent of the interaction.

In specific conditions various types of reactions (alkylation, metallation or oxidation) with nucleic acids can be promoted by small molecules resulting in a permanent alteration of the structure and function of the genome. The reactivity of these compounds with DNA has therapeutical relevance in that they can alter replication, protein expression, and cell metabolism. In this context metal complexes have attracted interest since their binding modes and reactivities can be tuned according to coordination number, shape and charge of the inorganic scaffold.[17] For example, cisplatin (*cis*-[$Pt(NH_3)_2Cl_2$]) and its derivatives react with the nitrogen (N7) on purine bases yielding 1,2-intrastrand crosslink. This chemical modification results into a kink of the DNA helix which ultimately leads to inhibition of DNA replication and transcription. The biological activity of cisplatin chemotypes quickly drew attention for their application as chemotherapeutic agents and nowadays cisplatin still represents the flagship in cancer treatment. Similarly, complexes with other metals have been analyzed for their reactivity with nucleic acids. Several ruthenium, iron and copper complexes have been reported to bind DNA and promote, under specific reacting conditions, the production of reactive oxygen species (ROS) capable of DNA oxidation and cleavage. DNA oxidation by ROS is a particularly attractive approach for permanent gene silencing as the oxidatively cleaved DNA strands are not easily repaired by ligase enzymes and damage often results in sequence deletion. Alternatively, other metals—such as cerium (Ce^{4+}) and zirconium (Zr^{4+})—can hydrolytically cleave oligonucleotide strands yielding intact 5′-phosphate and 3′-hydroxyl termini and offering an alternative to enzymatic nucleases for gene editing and manipulation.

Combining the chemical reactivity of metal complexes with probes that recognize and bind specific nucleic acid sites or sequences is an important research goal for developing precise diagnostic tools and precision gene-directed therapeutics. In this review we summarise and explore the applications of metallodrug-DNA interactions with particular focus on artificial gene editing technology. This review is will describe: (a) nucleic acid structure and function; (b) non-covalent and covalent DNA modifiers; (c) clinical metal complexes interacting with DNA; (d) metal complexes as artificial DNA scissors; (e) DNA footprinting; (f) protein engineering; and (g) gene editing technologies and the rise of artificial gene editing using metal complexes. The aim of this review is to provide readers with an insight on the rational design of metal complexes with nucleic acid recognition properties and to sequence-specifically target these interactions.

Nicolò Zuin Fantoni was born in Padova, Italy. He studied chemistry at the University of Padova receiving his MSc degree in 2014 for work in bioinorganic chemistry. In 2015 he joined the Kellett group in Dublin City University (DCU) as a Marie Curie PhD student within the ClickGene ITN project. Here, he worked on developing new copper polypyridyl complexes as stabilised artificial metallo nucleases (AMN) and linking them to triplex forming oligonucleotides by CuAAC and SPAAC reactions. In 2019 he joined the Brown group in the University of Oxford where he is working on the design of biocompatible triazole linkages.

Tom Brown is Professor of Nucleic Acid Chemistry in the Departments of Chemistry and Oncology at Oxford University. His research interests centre on oligonucleotide synthesis and applications of oligonucleotide chemistry in biology and medicine (diagnos-tics and therapeutics). He is co-inventor of several technologies for genetic analysis and co-founder of three Biotech companies including ATDBio (synthesis of modified oligonucleotides). He has published over 400 research papers and many patents. Awards including the Royal Society of Chemistry (RSC) Josef Loschmidt prize, the RSC award for Nucleic Acid Chemistry, the RSC prize for Interdisciplinary Re-search, Chemistry World entrepreneur of the year for 2014, and UK BBSRC research council Innovator of the Year for 2016. He was also presented with a life-time award for external engagement and promoting impact by Oxford University. Tom is a Fellow of the Royal Society of Edinburgh and a Fellow of the Royal Society of Chemistry. He is former President of the Chemistry Biology Interface Division of the RSC, Editor-in-Chief of the RSC Book series on Chemical Biology, and was co-Chair of the 2018 EuCheMS congress in Liverpool, UK. He is currently Associate Head of the Chemistry Department at Oxford University with responsibility for Research.

Andrew Kellett is Associate Professor of Inorganic and Medicinal Chemistry in the School of Chemical Sciences at Dublin City University (DCU). His group are based at the National Institute for Cellular Biotechnology (NICB) and Nano-Bioanalytical Research Facility and focus on the discovery of metallodrug-nucleic acid interactions for biochemical and therapeutic applications. He was coordinator of the Horizon 2020 Marie Skłodowska-Curie Innovative Training Network (ITN), ClickGene, and the newly funded NATURE-ETN Marie Skłodowska-Curie ITN. His lab is supported by Science Foundation Ireland, the Synthesis and Solid-State Pharmaceutical Centre (SSPC) , CÚRAM—an SFI Research Centre Designing The Next Generation of 'Smart' Medical Devices, and the Irish Research Council (IRC). He completed his B.Sc. in chemistry from Maynooth University and his PhD under the guidance of Prof. Michael Devereux and Dr. Maureen Walsh at the Technological University Dublin (TUD) and was then appointed as A.F. Graves Postdoctoral Fellow at the Focas Research Institute in 2008 and started his indpendent career at DCU in 2011.

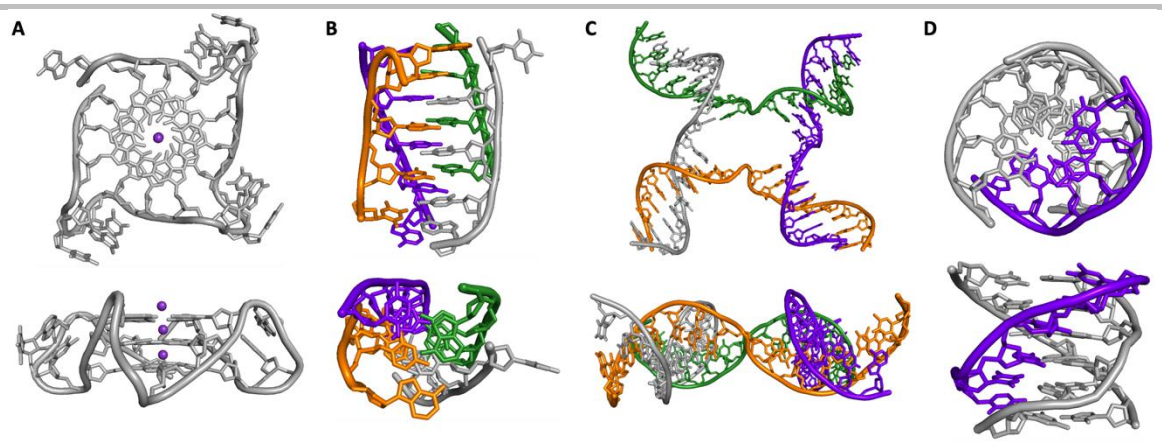

Figure. 2 Higher order DNA structures **A** G-Quadruplex (PDB: 1KF1); **B** I-motif (PDB: 1BQJ); **C** the Holliday junction (PDB: 467D); **D** triple helix (PDB: 1BWG).

2. Tertiary structures beyond the double helix

In addition to the three canonical DNA duplexes found in nature (A-, B- and Z-DNA), nucleic acids form a variety of tertiary structures with biological relevance such as G-quadruplexes, i-motifs, the Holliday junction, and triple helical structures (also known as H-DNA; Figure 2).[18] Small-molecule probes can designed to specifically target these tertiary structures or the structures themselves can be hybridised with reactive small molecules and used to target specific sites on the mammalian genome. In this section we will briefly describe these oligonucleotide tertiary structures that are found in cellular environments.

2.1. G-quadruplex DNA

G-quadruplexes are four-stranded structures formed by guanine-rich sequences where the Watson-Crick edge of each G-nucleobase is paired with the Hoogsteen edge of the adjacent guanine. These structures can exist in different topologies caused by the association of one to four G-rich strands stabilised by the presence of a central divalent cation (*e.g.* magnesium(II); Figure 2A).[19] G-quadruplexes are very common in telomeric DNA repeats suggesting a regulatory role in telomeric control of cell replication. The length of telomeres is usually shortened during cell replication until the Hayflick limit is reached and cell apoptosis is initiated.[20] Misregulation of this process, and of telomerase activity, can cause cell immortality and tumorigenesis.[21] Thus, developing compounds that specifically target these structures has particular relevance for understanding the involvement of G-quadruplexes in cellular pathways and for therapeutic purposes.

2.2. I-motif DNA

I-motifs are four stranded, intercalated structures formed by C-rich sequences extruded from G-rich regions. These motifs are formed by the stacked intercalation of hemiprotonated C^+•C base pairs and are usually less stable at physiological pH than their G-quadruplex counterparts due to requirement for slightly acidic environments (Figure 2B).[22] C-rich sequences are common within the human genome and i-motifs have been associated to telomeric, centromeric and promoter regions of proto-oncogenes suggesting possible regulatory functions in replication and transcription of DNA. Recently the presence of i-motifs and their regulatory roles were identified in the nuclei of human cells *in vivo*.[23]

2.3. The Holliday junction

Holliday junctions (HJ) are cross-shaped, branched nucleic acid structures characterized by four annealed dsDNA sequences (Figure 2C). This crossover arrangement forms during homologous recombination (HR) at sequences containing NCC trinucleotide cores (provided the N nucleobase is not thymine).[24] Two configurations are possible: the open-X structure which features four extended arms which are a result of the negative charge of the phosphate backbone on each of the nucleic acid strands repelling each other, and; the anti-parallel stacked-X conformation which is more compact and does not allow branch migration to occur. Holliday junctions are interesting as possible targets for combination anticancer therapies that selectively inhibit homologous recombination and increase PARP1 (poly(ADP-ribose) polymerase) sensitivity. In addition, Holliday junctions might have a role as alternative mechanisms for the extension of telomeres in that HR is involved in the process. The development of HJ-specific probes has therefore diagnostic potential in uncovering the role of Holliday junctions as an alternative mechanism for the lengthening of telomeres in immortalized cells lacking telomerases.[25][26]

2.4. DNA triplexes

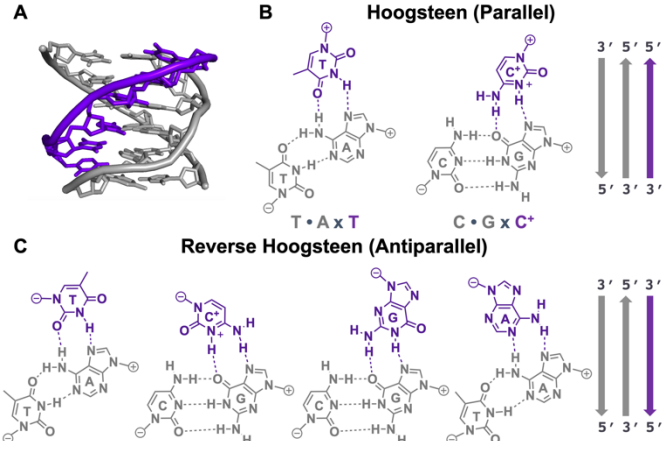

Fig. 3 A Triplex solution structure obtained by NMR studies (PDB: 1BWG); **B** Hoogsteen (parallel) base paired triplet schemes along with cartoon representation of the TFO strand orientation; **C** Reverse Hoogsteen (antiparallel) base triplet along with cartoon representation of the TFO strand orientation. (Hoogsteen and Watson-Crick bonding are represented by "x" and "•", respectively).

H-DNA is a triple helical arrangement where a third nucleic acid strand, called triplex forming oligonucleotide (TFO), recognises and binds a specific dsDNA sequence by hydrogen bond formation at the Hoogsteen face (major groove; Figure 3A). TFO binding is ruled by a specific code where presence of a polypurinic target represents an essential requirement to form concerted hydrogen-bonds among the three nucleobases.[27] Only purines can establish an extra pair of hydrogen bonds required for recognition by the third nucleobase on the TFO. The nucleotide content of the TFO defines orientation to the purinic target and pyrimidine-rich sequences prefer parallel (Hoogsteen; Figure 3B) binding, whereas polypurinic strands favour antiparallel base pairing (reverse Hoogsteen; Figure 3C). Parallel triplexes are more stable than antiparallel due to the isomorphism of T•AT and C^+•GC triplets, which resent of minor backbone tension compared to the reverse Hoogsteen base pairs. However, parallel TFOs require acidic environments in order to ensure formation of stable hydrogen bonds by complete cytosine protonation in the C^+•GC triplet.[28] In recent years, considerable indirect evidence of triplex formation and function *in vivo* was found by immunofluorescence using triplex-specific monoclonal antibodies. However, further studies are required to delineate if these structures arise *in vivo* from DNA triplexes or from RNA recognition of dsDNA sequences (*cf.* ref. 24).[29] The high abundancy of triplex target sites upstream of eukaryotic genes suggests a possible role of these structures in transcriptional regulation and led to the design of various TFOs with modified backbones and stabilizing moieties aimed to induce triplex-mediated gene-silencing.[30–32] The unique base-base recognition properties of TFOs have also been extensively exploited in combination strategies to develop small molecule-oligonucleotide hybrids for directed mutagenesis and cleavage.[33]

3. An introduction to non-covalent and covalent small molecule DNA interactions

The amphipatic nature and polymorphism of nucleic acids make DNA an exceptional substrate for the development of small molecules and biomaterials capable of interacting and reacting at specific genetic loci.[12] Chemotypes that interact with DNA usually fall within two classes: non-covalent and covalent (Figure 4). Non-covalent interactions are reversible and can be further divided into external electrostatic binding, groove binding and intercalation (including threading intercalation).[14,15]

The external electrostatic interaction is a non-covalent binding mode that occurs between small, positively charged molecules and the negative charge on the phosphate backbone of DNA. Cationic species such as metal ions, polyamines or charged opioid scaffolds diffuse along DNA interacting non-specifically with the phosphate groups on the minor groove.[13] The electrostatic interaction is followed by insignificant perturbation of the DNA structure and it usually represents a pre-association step in binding of metal complexes to nucleic acids.[34] Small molecules capable of electrostatic interaction have found applications as DNA condensation agents or stabilizer of tertiary nucleic acid structures such as triple helices.[35,36]

Groove binders interact either with the minor (e.g. netropsin, polyamides) or major (e.g. majority of DNA-binding proteins, methyl green) groove of DNA mainly through hydrogen bonding. This interaction together with other van der Waals and specific electrostatic contacts provide sequence selective binding and a series of small molecules such as DAPI, Hoechst 33258, netropsin and distamycin show sequence bias for the minor

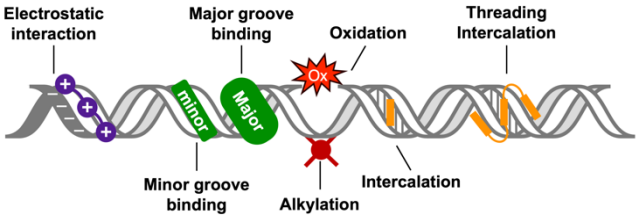

Figure. 4 Different modes of drug-DNA binding, from left to right: electrostatic interaction, minor and major groove binding, covalent modification (*i.e.* alkylation or oxidation), intercalation and threading intercalation.

groove of AT-rich regions of DNA.[14] Specificity arises from the narrow gap within the minor groove of AT tracts, which are quite different from to GC-rich sequences where the geometry is affected by bulky amino groups of guanine residues.[37,38] These differences have been exploited to target groove binders modified with specific gene editors to direct nucleobase-specific modification.[16,39]

Intercalation is a type of binding characteristic of planar aromatic molecules. These chemotypes slide between adjacent base pairs interacting with the nucleobases through hydrophobic and π-π stacking interactions.[34,40,41] As a consequence, intercalators are usually less sequence-selective than groove binders, but specific binding was recently shown for aromatic rings with particular modifications or multinuclear intercalative species.[15,42] Intercalation causes elongation of the DNA structure which is followed by an increase in the helix pitch and helical unwinding.[43] The affinity and perturbation properties of intercalators have been widely used to develop various stabilizing agents for DNA tertiary structure [44], probes[25,45,46], mutagens and developmental anticancer agents[15,47,48]. Threading intercalation is an alternative type of intercalation observed for flat aromatic systems flanked by bulky substituents. This type of binding has slower association and dissociation constants compared to canonical intercalators in that one of the bulky groups has to 'thread' through the stacked nucleobases in order for binding to occur.[49,50]

Covalent interactions are non-reversible and are often preceded by a non-covalent pre-association step. Here, binding involves a chemical reaction with DNA which generates a permanent modification on the genome altering its structure and function. Various covalent modifications can be introduced at the nucleic acid interface and they are usually divided in nucleotide alkylation, platination, 'metallation' and oxidation. Alkylating agents are among the oldest anticancer drugs and they include nitrogen and sulphur mustards, nitrosoureas, triazenes and ethyleneimines. These chemotypes react with the nucleophilic moieties of biological materials—for example the N7- or N1-atoms on guanine or adenine nucleobases, respectively—causing the formation of covalent adducts that inhibit DNA replication and RNA transcription.[51,52] Similarly, metal complexes can undergo ligand substitution and form non-reversible metal-nucleobase adducts. Among these, cisplatin (and derivatives) are best known for interacting with DNA in this manner, but complexes of other metal ions such as Ni(II) and Ru(II) can also bind directly to DNA.[53,54] In particular, some Pt(IV) and Ru(II) complexes were designed to undergo DNA metallation upon photoactivation of the complex thereby introducing an controlled cytotoxic mechanism that may overcome the general toxicity usually associated to DNA alkylating agents.[55–58]

In the case of DNA oxidation, a covalent modification is caused by interaction of reactive oxygen species (ROS) or metal-oxo species. ROS production is usually catalysed by a redox active metal that activates an oxygen molecule through reduction. This process generates reactive intermediates such as superoxide ($O_2^{\bullet-}$) and hydrogen peroxide (H_2O_2) which can further react producing hydroxyl radicals ($^{\bullet}OH$) and other metal-oxo species.[59] Reactivity of ROS (or metal-oxo species) at the nucleic acid interface usually generates a variety of mutagenic modifications dependent on the nature of the radical and of the nucleobase. Conversely, interaction with the sugar ring usually results in hydrogen abstraction and in a cascade of radical reactions that end in strand excision.[60]

Overall, a wide variety of metal complexes has been designed and tested for their binding and reactivity at the DNA interface. In several cases, the structural and DNA-recognition properties of organic ligands have augmented with the wide range of coordination environments, accessible redox potentials, and thermodynamic and kinetic properties of metal centres.[61–64] This synergy facilitates inorganic compounds to target nucleic acids for applications in the fields of anticancer therapy, DNA footprinting, gene editing and protein engineering.[63,65,66]

4. Clinical metal complexes that interact with DNA

4.1. Platinum derivatives

Although a wide range of metal complexes have been developed for anticancer purposes, cisplatin and its derivatives still represent class-leading chemotherapeutics. Cisplatin (*cis*-$[Pt(NH_3)_2Cl_2]$) was first synthesized in 1844 by Michele Peyrone and its *cis*-geometry was assigned by Alfred Werner in 1863 where it was distinguished from the *trans*-isomer.[67,68] However, it was only in 1965 when the interaction of this complex with DNA and its antiproliferative effects were serendipitously discovered by Barnett Rosenburg during electric field studies on *E. coli*.[3,69] The discovery of cell growth inhibition by cisplatin—produced at the platinum electrode—quickly led to its application in anticancer therapy. Cisplatin is commonly administered intravenously in saline buffer for the treatment of many forms of cancer including testicular, melanoma, bladder, non-small-cell-lung carcinomas, and in combination therapy, against ovarian cancer.[70,71] Once in the circulatory system, cisplatin is carried to cells where permeation into the cytosol occurs either through passive diffusion—due to its neutral charge—or by active transport ostensibly using copper transporter (hCTR) proteins.[72] The low chloride concentration within the cell (~3-20 mM) causes hydrolysis of the Cl^- substituents and activates the drug first to *cis*-$[Pt(NH_3)_2Cl(H_2O)]^+$ and then to the aqua complex *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$.[73] In this form, cisplatin is highly reactive towards DNA where it binds with the N7 group on purine bases yielding 1,2-intrastrand cross-links (47-50% 1,2-GG and 23-28 % 1,2-AG). Formation of the *cis*-Pt/DNA adducts causes a kink in the helical structure which inhibits polymerase binding, prevents DNA replication and transcription and ultimately leads to cell death.[74–76] Although the *trans* isomer, *trans*-$[Pt(NH_3)_2Cl_2]$, can coordinate to a single nucleobase, it lacks significant chemotherapeutic potential which emphasises the primary role of intrastrand cross-linking for antitumoral activity. Inspired by cisplatin, an extensive range of platinum derivatives have been prepared based on the *cis* chemotype but only carboplatin (1989) and oxaliplatin (2002) have received worldwide FDA approval for clinical use (Figure 5A) while nedaplatin, lobaplatin and heptaplatin (Figure 5B) are approved in Japan, China and South Korea, respectively.[64] Nonetheless, the therapeutic potency of cisplatin derivatives is paralleled with intrinsic or acquired resistance factors, bioaccumulation, and dose-limiting toxicity (i.e. nausea, ototoxicity, neurotoxicity, myelosuppression and nephrotoxicity).[77] With the aim of introducing molecular interactions not accessible to monomeric complexes, a series of polynuclear platinum complexes (PPCs) with structural similarities to the *trans*-Pt(II) chemotype were designed and

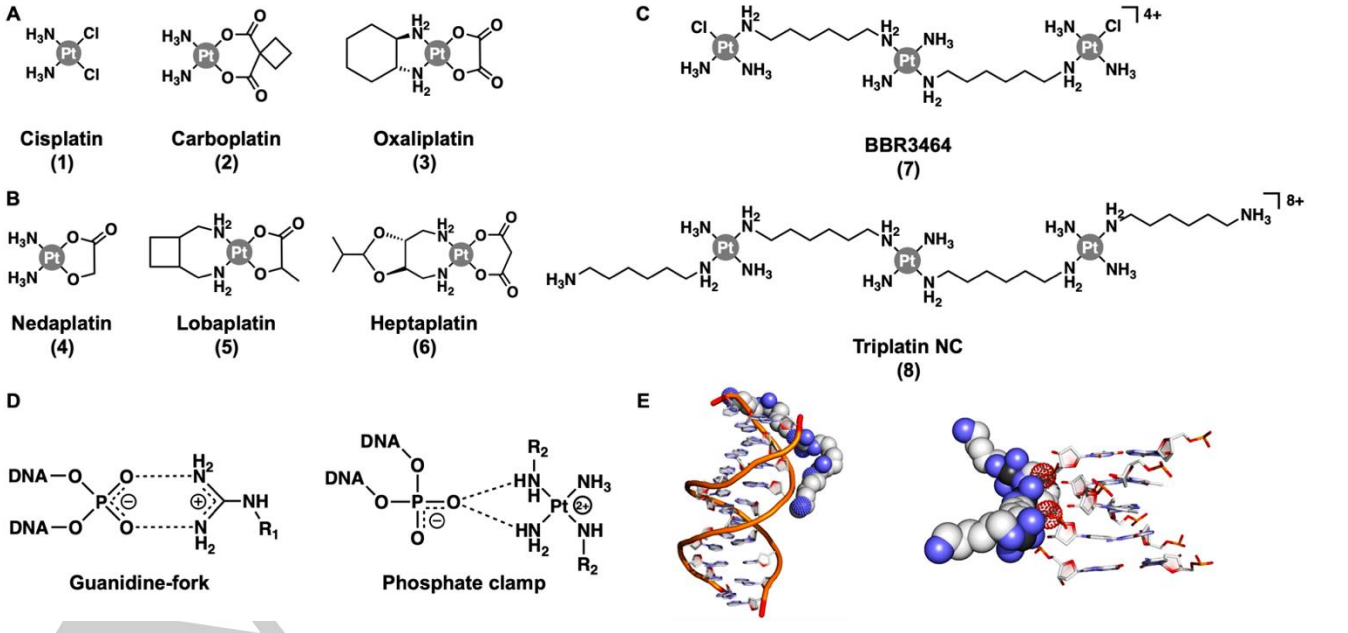

Figure. 5 Molecular structure of cisplatin derivatives **A** FDA approved (cisplatin, carboplatin, oxaliplatin) or **B** employed in single markets (nedaplatin, lobaplatin, heptaplatin); **C** Transplatin polynuclear derivatives with cytotoxic activity $[\{trans\text{-}PtCl(NH_3)_2\}_2\text{-}\mu\text{-}\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_2)_2\}]^{4+}$ (BBR3464), and $[\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_3)\}_2\text{-}\mu\text{-}\{trans\text{-}Pt(NH_3)_2(NH_2(CH_2)_6NH_2)_2\}]^{6+}$ (TriplatinNC); **D** Comparison in binding interaction of a guanidino-fork clamping by arginine residues (left) and the phosphate clamp of TriplatinNC chemotypes (right); E) Interaction modes of TriplatinNC with DNA: backbone tracking (left) and groove spanning (right).

synthesised in the Farrell group (Figure 5C).[78] These platinum complexes are characterised by *di*- or *tri*-nuclear transplatin cores connected through alkanediamine linkers of various length. In contrast to the 1,3-intrastrand and 1,2-intrastrand cross-links caused by transplatin and cisplatin on neighbouring nucleobases, these chemotypes promote 1,2-, 1,3-, 1,4- and 1,6- interstrand crosslinks (IXLs) where platination sites on opposite strands are separated by up to five base pairs.[79,80] The interstrand adducts induce very small conformational distortions in the DNA structure with 10° bending of the helix axis and 9° unwinding of the duplex for the 1,3- and 1,4-crosslinks.[81] For these reasons, the antitumoral activity by polynuclear platinum complexes might arise from the obstruction that interstrand adduct formation imposes on DNA and RNA polymerase machineries. This suggests a pharmacological mechanism that differs from cisplatin where the kink in the double helix promotes Pt-adduct recognition by HMG protein protecting the 'alkylated' site from excision repair mechanisms.[78] Similar to *cis*-Pt(II) complexes, DNA-crosslinking proceeds via hydrolysis of the chloride groups and is preceded by pre-association with the phosphate backbone. In the case of BBR3464—the only non-cisplatin chemotype to enter phase II clinical trials—the presence of a charged central moiety aids pre-association allowing the formation of directional isomers depending on the nature of the crosslink (1,2-IXL forms in the 3′-3′ direction; 1,6-IXL forms in the 5′-5′ direction while 1,4-IXL show no preferentiality).[82]

In order to explore the biological consequences and pre-association mechanism of PCCs further, a series of substitution inert non-covalent Pt(II) complexes were prepared by substituting chloride ligands with NH_3 or 'dangling am(m)ines'. The presence of these substitution-inert moieties introduces a distinctive type of 'non-covalent' DNA-binding compared to the covalent interactions by canonical *cis*- and *trans*-Pt(II) derivatives.[79] Here, the am(m)ine groups hydrogen bond to the oxygen atoms of the phosphate backbone (OP) establishing N-O-N motifs—called the phosphate clamp—which resembles, to a certain extent, the OP-clamping of guanidine groups on arginine residues (Figure 5D).[83,84] This type of binding induces either groove spanning or backbone tracking (Figure 5E) depending on the presence of poly[d(A-T)$_2$] or poly[d(G-C)$_2$] sequences, respectively. Both types of interaction are followed by a perturbation of the DNA structure in that: a greater axial bend and axial path length shortening ratio are induced; DNA is slightly bent in the major groove and; the minor groove width profile is modestly impacted.[85] Despite the non-covalent nature of this binding mode, the distortion induced in the double helix resembles the structural perturbation caused by covalent bifunctional 1,2-intrastrand adducts with cisplatin. In addition, TriplatinNC exhibits potent antitumoral activity toward various cell lines within the low micromolar range emphasising that covalent interactions are not a prerequisite for cytotoxic platinum(II) complexes.[83,85]

4.2. Clinically validated non-platinum chemotherapeutics

Systematic toxicity of cisplatin derivatives sparked new research to discover complexes with different pharmacokinetic and pharmacodynamic profiles. Among these, ruthenium complexes attracted considerable interest in that their therapeutic activity (including anticancer activity) was first recognised by Dwyer in 1950 but largely forgotten until the discovery of cisplatin.[86,87] In the last three decades however a wide range of ruthenium complexes have been synthesized and investigated as potential chemotherapeutics with some showing activity against cisplatin-resistant tumours and with three chemotypes entering clinical trials.[88–90] NAMI-A—a Ru(III) coordination compound of formula [ImH][*trans*-$RuCl_4$(DMSO-S)(Im)] (where Im = imidazole)—was the first ruthenium drug undergoing human testing. Phase I trials showed several side effects including—but not limited to—mild hematologic toxicity, quite disabling nausea, vomiting, and diarrhoea. Due to these results, new Phase I/II clinical trials were performed in combination therapy with gemcitabine, a nucleoside derivative with chemotherapeutic efficacy on non-small cell lung cancer. However, lack of efficacy results and the toxicity profile caused suspension of NAMI-A from clinical trials.[91] KP1019/KP1339—the $IndH^+$ and Na^+ salts, respectively, of an octahedral Ru(III) species with general formula [*trans*-$RuCl_4(Ind)_2]^-$ (where Ind = indazole)—was the second Ru(III) complex entering Phase I trials. This drug showed promising results including disease stabilization for 8-10 weeks on patients with solid tumours and only mild toxicities.[92] Both NAMI-A and KP1019 mechanism of action is supposed to start from activation by reduction to the Ru(II) derivative *in vivo*. While DNA has been proposed as one of the biological targets for KP1019, the antimetastatic activity of NAMI-A is however associated with its combined effects on angiogenesis control and antiinvasive properties towards tumour cells and blood vessels.[89,93]

Recently, the photochemical properties of some Ru(II) complexes—and also Pt(IV)[56,70], Rh(III)[94], Os(II)[95] and Ir(III)[96] complexes—has increased interest in their application in photodynamic therapy (PDT), photochemotherapy (PCT) and photoactivated chemotherapy (PACT).[54,90,97,98] In these cases, cytotoxicity is activated upon irradiation of the soft tissue with a penetrating red light. This pharmacodynamic profile has particular relevance because it allows restriction of the therapeutic effect to the diseased tissue preventing onset of side toxicities. One such example is the racemic mixture *rac*-[Ru(dmb)$_2$(IP-3T)]Cl_2 (TLD1433, where IP-3T = imidazo[4,5-f][1,10]-phenanthroline (IP) ligand appended to an α-terthienyl (3T) as the organic chromophore and dmb = 4,4′-dimethyl-2,2′-bipyridine), which was the first Ru(II)-based photosensitizer to enter human clinical trials (ClinicalTrials.gov, identifier NCT03053635). This complex has a long-lived triplet ^{3}IL and ^{3}ILCT states which allow $^1O_2$ sensitization and electron-transfer, respectively, producing extremely potent photocytotoxic effects. In addition, the bright luminescence given by the ^{3}MLCT states provides this photosensitizer with theragnostic capacity.[90] DNA damage studies by this complex highlighted that upon irradiation single strand cleavage is triggered at 500 nM loading with onset of double strand cleavage at higher concentrations.[99] Interestingly, damage was enhanced in the presence of glutathione (GSH), a reductant known to play an essential role in platinum chemotherapeutic detoxification.[100] Promoted double strand cleavage by TLD1433 in the presence of GSH and other important antioxidants gives prominence to a unique cytotoxic mechanism which can overcome the limitations of platinum therapeutics.

4.3. Bleomycin

Another clinically established metallodrug is metallobleomycin—a family of anticancer antibiotics produced by *Streptomyces Verticillus*—where cytotoxicity is induced through a redox active metal capable of catalysing oxidative DNA damage.[101] These molecules are characterized by a common glycopeptide core, but they differ in the pendant sugar and positively charged tail modifications (Figure 6). Bleomycins bind DNA sequence-specifically via their N3- and N4- pyrimidine interactions with reciprocal guanine nucleobases.[102,103] DNA binding is aided by cooperative interaction of the bithiazole tail through nucleobase intercalation or minor groove recognition. Finally, the pendant positively charged tail interact electrostatically with DNA enhancing the overall binding affinity.[104,105,102] Despite the sugar modification not contributing in DNA binding and recognition, it was proposed to affect solubility and cell permeability.[106,107] DNA cleavage by bleomycin is promoted by the metal-binding domain which comprises a β-aminoalanine-pyrimidine-β-hydroxyhistidine moiety. Here, a redox active metal ion (primarily Fe^{2+}, but also Cu^{2+}, Mn^{2+} or Co^{3+}) is coordinated by five nitrogen ions leaving the sixth coordination site open for interaction with an oxygen molecule.[108] In the presence of reductant, the metal bound domain promotes production of reactive oxygen species (ROS) and abstraction of the C4′ hydrogen from the deoxyribose ring of DNA. Depending on O_2 availability the 4′-radical intermediate further reacts with H_2O or O_2 leading to a cascade of reactions that result either in formation of 4′-oxidized abasic site or strand

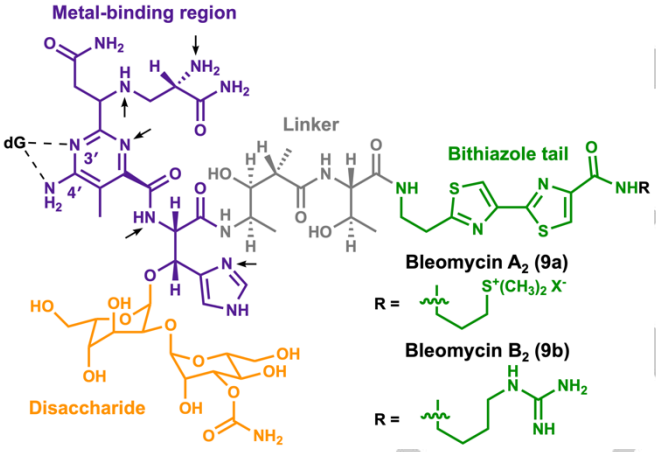

Figure. 6 Molecular structure of bleomycin. The metal binding region (purple) has five nitrogen atoms (indicated by arrows) available to coordinate a metal ion. The linker region (grey) and bithiazole tail (green) aids DNA recognition and binding while the disaccharide moiety (orange) is supposed to enhance solubility and cellular uptake.

excision, respectively.[109–111] Nowadays, bleomycins are used clinically as a mixture of different congeners (60% bleomycin A_2 and 30% bleomycin B_2, denoted as Blenoxane, Figure 6) or in combination with other therapeutics for the treatment of several types of tumors, notably squamous cell carcinomas and malignant lymphomas. Their therapeutic efficacy is associated with low myelosuppression and immunosuppression levels but is limited by dose-dependent pneumonitis.[101]

5. Artificial metallo-nucleases

5.1. Cleavage mechanism by [Cu(1,10-phenantroline)$_2$]$^{2+}$

In parallel to natural compounds such as bleomycins, transition metal complexes have been developed and studied for their DNA

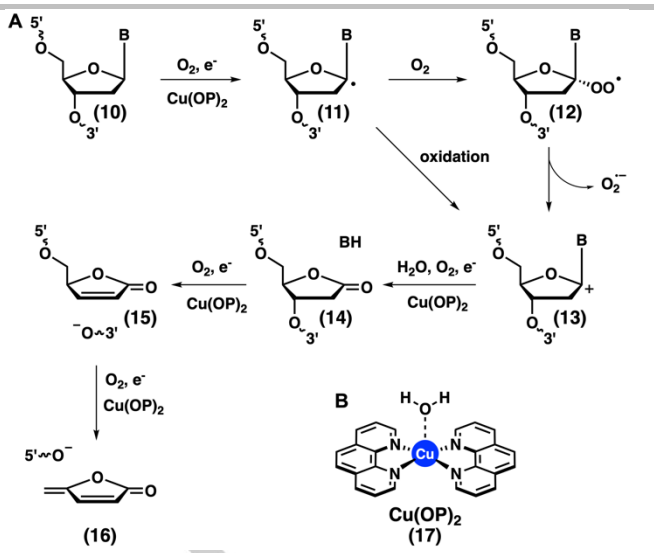

Figure. 7 A Radical oxidation mechanism by CuPhen resulting in direct strand cleavage; **B** Molecular structure of the [Cu(1,10-phenantroline)$_2$(H_2O)]$^{2+}$.

binding properties and artificial metallo nuclease (AMNs) activity. Complexes of various redox active metals such as Cu, Fe, Mn, Ni, Co, Ru and Rh can promote oxidative DNA damage at nucleobase and/or deoxyribose moieties. Among these, Cu-AMNs have been extensively investigated due to their accessible redox states, variety of coordination geometries and endogenous nature of the metal.[63,112,113] DNA binding and oxidation by AMN has therefore been exploited for the development of potential antibacterial, antifungal and anticancer drugs, for DNA footprinting, gene silencing, and protein engineering. [Cu(1,10-phenantroline)$_2$]$^{2+}$ (CuPhen, Figure 7B) is the most widely studied AMN for both its cleavage mechanism and biological potential.[114–117] In the presence of exogenous reductant and oxidant, CuPhen induces direct single or double strand cleavage—arising from proximate single strand breaks—at the minor groove of poly[d(A-T)$_2$] sequences.[118] In the reduced form, CuPhen semi-intercalates DNA promoting ROS which causes abstraction primarily of the C1′ hydrogen (C5′ and C4′ at a lower extent) from the deoxyribose ring.[119,120] Seminal work by Greenberg *et al*. on the oxidative mechanism by CuPhen identified that after H-atom abstraction, direct strand breakage proceeds through formation of a peroxyl radical intermediate.[115,121] Here, expulsion of superoxide generates an oxidised C1′ cation intermediate that in presence of other two CuPhen undergoes β-elimination and conversion from 3′-furanone to the free 5-methylene furanone (Figure 7A). The interaction of CuPhen with deoxyguanosine was further studied by the Kellett group and shown to generate 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxo-dG) lesions within superhelical plasmid DNA using an enzyme-linked immunosorbent assay (ELISA).[122,123]

5.2. CuPhen-modified chemotypes

Inspired by CuPhen, various structural modifications were introduced in the complex scaffold aimed at improving its stability, DNA recognition and reactivity (Figure 8).[123,124] Carboxylate ligands were introduced to afford a series of mononuclear and bridged dinuclear complexes with higher anticancer activities (low μM range) than cisplatin against various cell lines (MCF-7, DU145, SK-OV-3, HT29). Interestingly copper(II) bis-phen complexes with pendant phthalate showed nuclease activity in absence of exogenous oxidant or reductant and were the first ‘self-activating’

complexes of their kind reported in the literature (Figure 8A).[59,125] A dinuclear complex of this series—Cu-Oda, [{Cu(Phen)$_2$}$_2$(μ-Oda)]$^{2+}$—showed discrimination of AT/AT from TA/TA steps of dodecamer palindromic sequences through induction of B- to Z-like DNA change or by intercalative binding.[42] In order to introduce stronger DNA interactions and sequence recognition, one phenanthroline ligand was substituted with phenazine intercalators of extended aromaticity to yield a class of heteroleptic complexes with general formula Cu-Phen-*N,N′* (where *N,N′* = dipyrido[3,2-*f*:2',3'-*h*]quinoxaline, DPQ; dipyrido[3,2-*a*:2',3'-*c*]phenazine, DPPZ and benzo[*i*]dipyrido[3,2-*a*:2',3'-*c*]phenazine, DPPN; Figure 8B). These complexes showed binding constants for calf thymus DNA (ctDNA) within ~10^7 M (bp)$^{-1}$ and are currently the best binders among the CuPhen derivatives reported in the literature.[123,124]

5.3. Stabilised artificial metallo-nucleases

Despite their prominent binding profile, Cu(II) bis-phenanthrene systems tend to form new species in solution and the second coordinated ligand dissociates yielding the inactive monocoordinated complex.[126,127] In order to tackle this stability issue various strategies have been employed and covalent bridging linkers were used to tether two phenanthrene ligands together or polypyridyl scaffolds were coordinated to form 'metal-caged' complexes.

In the first strategy serinol bridges were introduced between either the C2 or C3 (2-Clip-Phen and 3-Clip-Phen, respectively) carbon of Phen in order to favor the 2:1 Phen:Cu ratio (Figure 8C).[128] Here, position and nature of the linker significantly influenced the activity of the compounds and 3-Clip-Phen was 60-fold more efficient as AMN than CuPhen whereas 2-Clip-Phen was only twice more effective in DNA cleavage. The high activity of 3-Clip-Phen was attributed by DFT calculations to low structural rearrangement energy barriers between Cu(II)/Cu(I) oxidation states which allows the complex to retain a certain planarity ideal for intercalation and minor groove binding.[129]

insight in the mechanism of oxygen binding and activation of several Cu-enzymes such as hemocyanin, tyrosinase and multicopper oxidase.[130–132] More recently however these chemotypes were studied for their capability to promote oxidation of bio-organic substrates and for their application as artificial metallo-nucleases.[133] In an effort to combine the stabilising properties of polypyridyl scaffolds with the high recognition and binding provided by phenanthrene derivatives, Cu-TPMA-*N,N′*[134] and Cu-DPA-*N,N′*[135] (where TPMA = *tris*-(2-pyridylmethyl)amine); DPA = *di*-(2-pycolylamine); and *N,N'* = phenanthrene ligands) complexes were recently synthesised and studied for their DNA binding and cleavage properties. The flexible tetracoordination of TPMA allows facile rearrangement of the ligand donors around the copper center during the Cu(II)/Cu(I) catalytical cycles and leaves two additional coordination sites on the copper center for binding by the phenanthrene DNA intercalator. Binding assays conducted on Cu-TPMA-*N,N'* showed that extending the aromaticity on the phenanthrene ligands enhances recognition by the compounds against dsDNA. Topoisomerase I-mediated relaxation assays identified binding by Cu-TPMA-*N,N'* to plasmid DNA occurs through intercalation with DPQ and DPPZ derivatives exhibiting the greatest unwinding activity within the series. However, binding by the TPMA complexes (K_{app} ~10^5 M) is significantly lower compared to previously reported Cu-phenazine derivatives (K_{app} ~10^7 M) likely due to the steric hindrance imposed by a *tris*-polypyridyl ligand scaffold. In this context, the less hindered *bis*-polypyridyl ligand DPA provided a greater balance between complex stabilization and high-affinity DNA binding with K_{app} values rising to 10^6 and 10^7 M bp^{-1} for DPQ and DPPZ complexes, respectively, within the Cu-DPA-*N,N'* series. Both Cu-TPMA-*N,N'* and Cu-DPA-*N,N'* compounds are efficient AMNs with an overall trend in nuclease activity dependent on the phenanthrene intercalator and following the general trend: Phen ≃ DPQ >> DPPZ. Cleavage by these compounds occurs within the minor groove of DNA by a superoxide-mediated mechanism that departs from classical

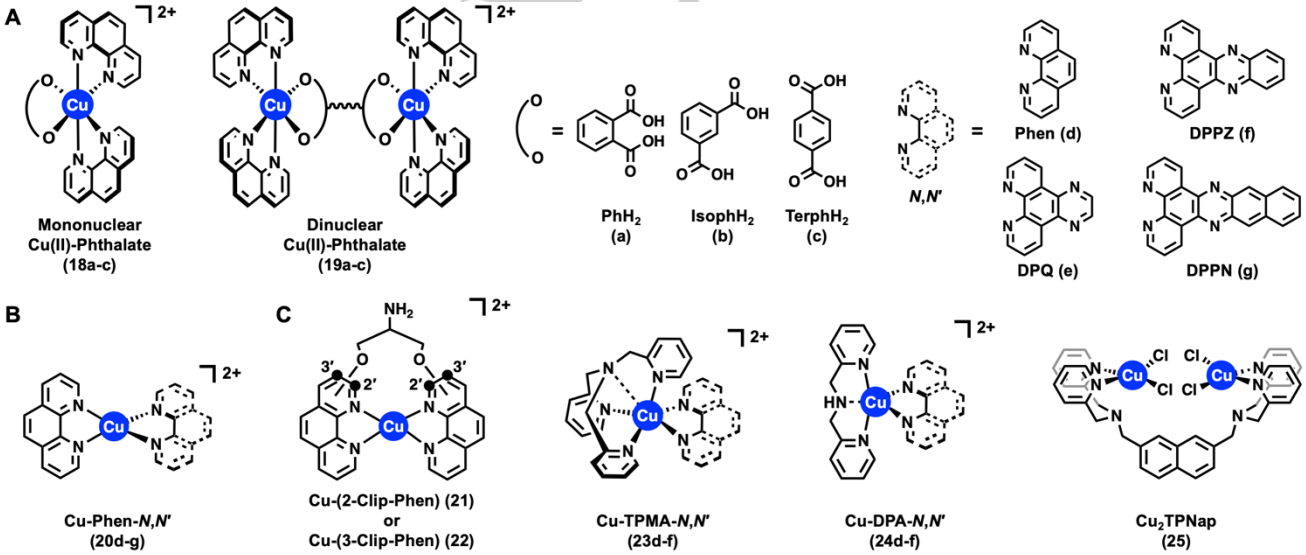

Figure. 8 Molecular structures of **A** mononuclear and dinuclear 'self-activating' Cu(II) phthalate complexes; **B** mononuclear heteroleptic Cu-Phen-Phenazine complexes; **C** 'stabilized' Cu-Clip-Phen and Cu-polypyridyl (Cu-TPMA-*N,N'*, Cu-DPA-*N,N'* and Cu$_2$TPNap) derivatives.

In the second approach, polypyridyl ligands have been used to afford stable complexes where the metal is 'caged' within an environment that mimics the histidine-rich active site of copper proteins. Hence, these complexes have been used to gain an

Fenton-type or Haber-Weiss processes. The potential chemotherapeutic application of Cu-TPMA-Phen was analysed against the human-neuroblastoma derived cell line NB100.[Ref needed here Nicoló] After 24 h of continuous exposure, the

calculated EC_{50} of the complex was 4.2 μM and two thirds of the population underwent apoptosis while ~5 % was in necrosis. Apoptosis represents a therapeutic advantage over necrosis in that the latter causes inflammation leading to unwanted toxicity toward surrounding healthy tissues. In addition, NB100 treatment with Cu-TPMA-Phen resulted in significant alteration of the fatty acid membrane composition with an increase of saturated fatty acids (SFA) and a decrease of the monounsaturated counterparts (MUFA) promoted by complex-mediated lipid peroxidation. Promising cytotoxic activity was also observed by Cu-DPA-*N*,*N*' complexes against four different human pancreatic cell lines (MiaPaCa-2, Panc-1, HPAC, and a patient-derived model PIN-127). The cytotoxicity of these compounds was compared to oxaliplatin—one of the main chemotherapeutics used to treat pancreatic ductal adenocarcinoma. Overall, Cu-DPA-Phen and Cu-DPA-DPQ showed similar *in vitro* cytotoxicity to oxaliplatin while Cu-DPA-DPPZ outperformed oxaliplatin activity with IC_{50} values falling in the nanomolar range. The enhanced solution stability, unique DNA oxidation properties, and promising cytotoxic profile of copper polypyridyl complexes appended with phenanthrene intercalators therefore opens new prospects for the further development and application of these compounds as new chemotherapeutics.

DNA recognition and oxygen activation by AMNs are strongly dependent on both the nuclearity and the coordination environment surrounding copper.[136] While some mononuclear Cu-AMNs are considered promiscuous in their biological activities, *di*- and *tri*-nuclear complexes were found to recognize distinct structural features of nucleic acids. Two Cu-polypyridyl compounds, $[Cu_2(D^1)(H_2O)_2]^{4+}$ and $[Cu_3(TP1)(H_2O)_3(NO_3)_2]^{4+}$ (where D^1 = dinucleating ligand with two tris-(2-pyridylmethyl)-amine units covalently linked in their 5-pyridyl positions by a -CH_2CH_2- bridge and TP1 = 2,2′,2″-*tris*(*di*-picolylamino)triethylamine) bind and cleave junctions between single and double stranded DNA in a sequence specific fashion. Here, site-specific cleavage is afforded by the coordination of at least one copper to G-rich strands.[137,138] Indeed, guanine is known to have the highest coordination affinity to transition metals and is the most easily oxidised nucleobase followed by adenine, thymine, and cytosine.[138,139] The same group also showed modification of the linker spacer between copper cores can switch oxidation selectivity from the usually targeted deoxyribose ring to guanine nucleobases of frayed DNA sequences.[140] Similarly, [Cu_2(*tetra*-(2-pyridyl)-NMe-naphthalene)Cl_4] (Cu_2TPNap) was shown to target the major groove of dsDNA and induces single strand breaks in the absence of added reductant.[141] The coordination environment and nuclearity are known to control the type of Cu_nO_2 coordination geometry. Therefore, the oxygen activation mechanism and the type of radical produced ultimately results in a wide variety of DNA oxidation products.[142] In this context, Cu-AMNs provide not only an alternative pharmacological mechanism to state-of-the-art chemotherapeutics, but are also invaluable tools to probe and modify the underlying genetic structure useful for a range of biochemical functions.

6. DNA footprinting

Developed in 1978 by Galas and Schmitz, DNA footprinting is a technique used for studying sequence-selectivity of DNA binding compounds as it exploits the protection from cleavage provided by a ligand bound to a specific site. The method is based on a combination of Maxam-Gilbert DNA sequencing with DNAse-protected fragment isolation.[143] Footprinting templates are radiolabelled strands of between 50-200 base pairs that are cleaved both in the presence and absence of a binding agent. The region bound to the ligand is protected from cleavage and creates a gap or “footprint” in the ladder of digested products when these are resolved through polyacrylamide gel electrophoresis (PAGE). The nuclease agent should therefore cleave DNA in an aspecific fashion in order to provide an even distribution of cleaved fragments. DNase I is the most common nuclease used in footprinting experiments but it generates an erratic ladder of cleavage products, as its efficiency is affected by global and local DNA structure.[144,145] In this context mononuclear chemical nucleases can provide an exceptional alternative tool to enzymatic nucleases due to their aspecific cleavage. Various complexes such as methidiumpropyl-EDTA Fe(II) (MPE), uranyl photocleavers, and Cu-Phen have been employed in DNA footprinting.[146] Using these tools, footprinting enables the elucidation of sequence-specific binding by various small molecules, proteins, and triplex-forming oligonucleotides.[Refs?] Nicolo, should we mention something here about the advantages of Cu-Phen (resolution)? Any other advantages would be good since the review is focused on AMNs.

7. Protein Engineering

In parallel to anticancer and footprinting applications, AMNs were recently employed in the field of protein engineering to generate libraries of highly diverse mutant proteins. For this purpose, the cleavage action of a nuclease (*i.e.* DNase I) is usually used to generate various 5′-phosphorylated and 3′-hydroxylated fragments that can be further manipulated and shuffled into new gene variants.[147] This process is known as DNA shuffling and was first introduced by Stemmer in 1994 (Figure 9A).[148] This technique has major commercial importance for the production of high performance proteins in areas such as affinity antibody development and it involves digestion of a parent gene by DNase I followed by reannealing of the generated DNA fragments in the

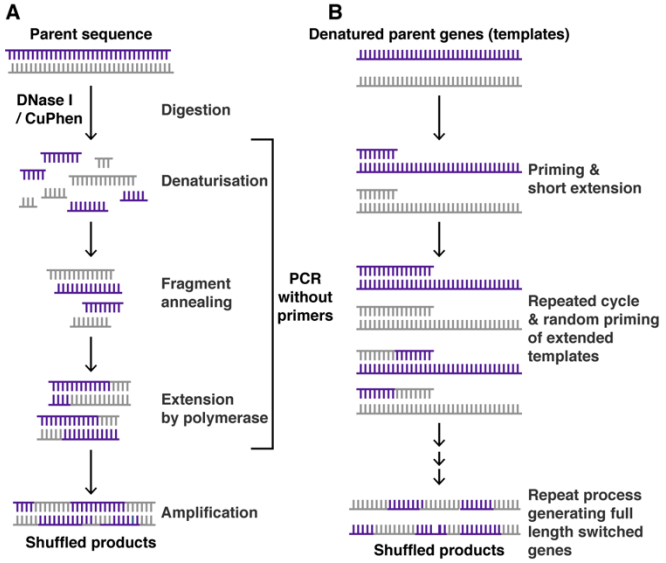

Figure. 9 A DNA shuffling through primerless PCR; **B** staggered extension process (StEP).

presence of a primerless DNA polymerase. The shuffled genes are then integrated within a larger plasmid and transformed into *E. coli* bacteria for translation in the corresponding mutant antibody proteins. Recombinant DNA libraries are then sequenced, and antigen-binding affinity of antibody variants accurately identified using technology such as surface plasmon resonance. Alternatively, engineered protein libraries can be obtained through the staggered extension process (StEP). This method—first developed in 1998 by Arnold who subsequently shared the Nobel prize in 2018 with Smith and Winter—prepares full-length recombinant genes using primers and repeated cycles of denaturation, fast annealing, and polymerase catalysed extension (Figure 9B).[149] This experimental design allows the growing fragments to anneal to a different template within each cycle and to acquire sequence information of different parental genes.

Recently, we exploited Cu-Phen mediated cleavage and DNA shuffling to yield a recombinant library of antibodies for prostate-specific antigen (PSA).[150] Despite the oxidative cleavage effects by CuPhen, a fraction of fragments suitable for PCR was generated through direct strand breakage where promoted β-elimination yields intact 5′-adducts through C1′ abstraction. Additionally, the Taq DNA polymerase I used in the experiment lacked a functional 3′ → 5′ exonuclease, but its functioning N-terminal 5′ → 3′ exonuclease suggested possible amplification through the nick translation process. DNA shuffling of the fragments restricted using CuPhen and DNase I produced 60 clones with affinity to PSA and five of the top ten best-performing clones were derived from the CuPhen digested library. In addition, the top-performing antibodies showed high levels of modification suggesting CuPhen allows accessibility to a more diverse gene mutation library than DNase I. Results highlighted in this work showed the potential that Cu-AMNs can have within the field of protein engineering and their mutagenesis can be directed by conjugating AMNs to selective organic-based DNA binding agents.[66]

8. Gene editing

Since double strand breaks of genomic DNA are toxic to cells, biological systems possess various mechanisms to repair lesions, namely: non-homologous end joining (NHEJ) and homologous directed repair (HDR). In NHEJ, two cleaved ends of the damaged duplex are resealed together by a ligase enzyme without the requirement for templates. NHEJ can often result in the introduction of insertion or deletions that alters gene function. Conversely, HDR is less error-prone since the repair is based on the annealing of the damaged sequence to a matching template.[151] In HDR the 3′-end of one strand recognizes and invades a specific dsDNA sequence forming a Holliday junction structure. Here, the invading strand operates as a primer initiating DNA synthesis and damage repair. Therefore, both processes have interesting applications in genome engineering in that single base modifications, but also exogenous DNA sequences, can be introduced in the genetic code resulting into gene disruption, mutation or in the creation of specific genome-edits.[152] Hence, research has focused on developing gene-editing tools where cleavage by promiscuous nuclease machinery is directed by sequence-specific DNA binding domains (Figure 10).

8.1. Enzymatic nucleases

8.1.1. ZFNs

Zinc finger nucleases (ZFNs) are the oldest gene-editing technology and consist of the cleavage domain of *Fok*I—a Type IIS restriction endonuclease—and a polypeptide chain that recognises specific DNA sequences (Figure 10A). The polypeptide is an ensemble of zinc finger domains, also known as Cys_2His_2 due to the coordination environment of zinc which aids structure folding and stabilisation. Each finger consists of ~30 amino acids folded into a α-helix and a β-sheet. Sequence specific binding to DNA is provided by the interaction between aminoacidic side chains on the α-helix domain (mainly -1, 3 and 6 from the start of the α-helix) and the nucleobases in the major groove.[153] Each zinc finger can bind 3 bp with recognition of at

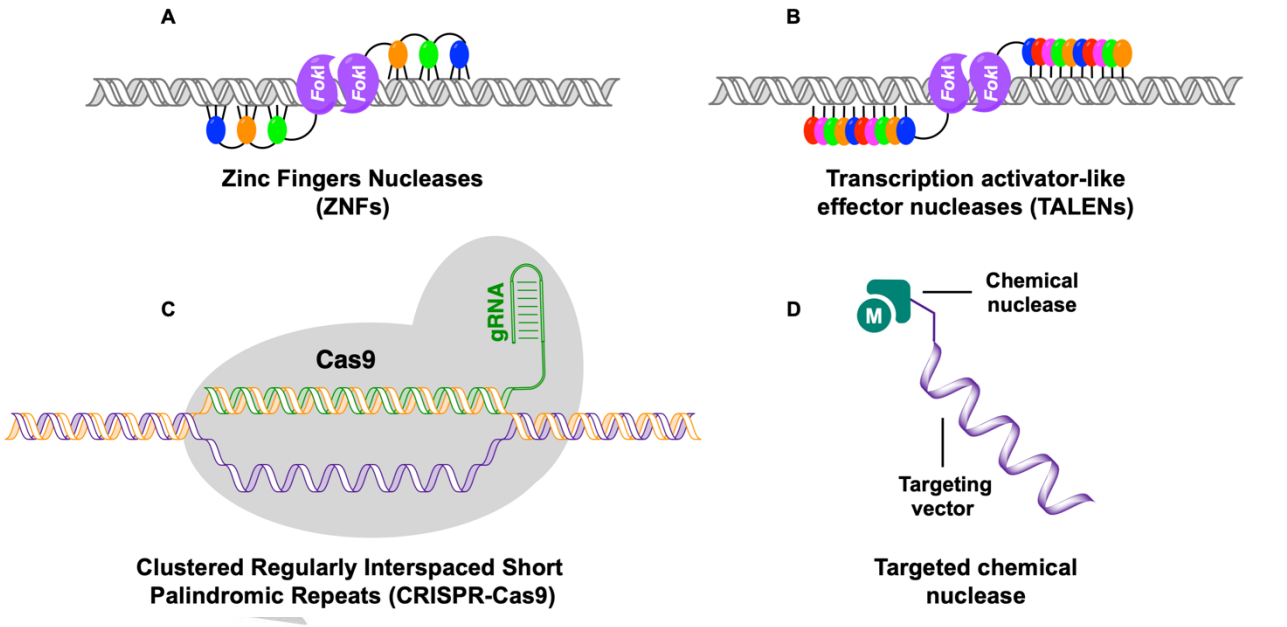

Figure. 10 Cartoon representation of sequence-targeted enzymatic (Zinc finger nuclease, ZNF (**A**); Transcription activator-like effector nuclease, TALEN (**B**); Clustered regularly interspaced short palindromic repeat, CRISPR-Cas9 (**C**)) and chemical nucleases (**D**).

least 9 bp is required for stable hybridisation to DNA. ZF domains have been developed with specificity to nearly all the 64 possible nucleotide triplets and linked together to assemble combinations capable of targeting a broad range of DNA sequences.[154] Cleavage by ZFN is promoted by the non-specific nuclease domain of *Fok*I after dimerization of two ZFN hybrids on opposite strands. The cleavage moiety is linked to the C terminus of the peptide and both linker length and distance between the two cleavage agents are important in the effectiveness of the gene-editor. In cells it was found that the best performing spacer length was 6 bp for ZFNs with short linkers (n = 4).[155,156] However, the complex design of the zinc finger has somewhat limited the application of ZFNs in genome engineering and opened the field to simpler systems such as TALENs and CRISPRs.

8.1.2. TALENs

Transcription activator-like effector nucleases (TALENs) are a class of gene-editing tools introduced as an advancement to zinc finger nucleases. Despite both editors sharing the same cleavage domain (*Fok*I), in TALENs specific recognition is provided by a DNA binding protein: the transcription activator-like effector (TALE; Fig. 10B). TALEs are a class of proteins secreted by plant-pathogenic bacteria (*Xanthomonas*) that bind and activate genes that aid infection.[157] These proteins are constituted by repeats of 33-35 amino acids mostly conserved among the modules apart from the 12th and 13th residues—known as repeat variable *di*-residues (RVD)—which control base pair recognition. In the RVD, amino acid 13 binds to the nucleobase while residue 12 side chain folds back to promote stabilizing interactions with the other residues in the module.[158,159] TALE single base interaction provides a more flexible design than triplet recognition by ZF domains and TALE repeat combinations have been reported to virtually recognize any DNA sequence. Canonical TALENs share the same *Fok*I cleavage domain common in ZFNs (hence requiring dimerization) but several variants with nucleases, transcriptional activators and recombinases were developed in the literature.[160–162] In contrast to ZFNs, TALENs require some additional protein on both ends of the DNA binding domain. Two modules with similar structure but different sequences to the TALE repeats are located upstream and are responsible for the contacts with a T nucleobase (the preferential base at the start of the binding site) while additional residues separate the C termini of TALE from the FokI domain, probably to aid folding and stability.[160,163]

8.1.3. CRISPR-Cas

CRISPR-Cas systems are recognised as the state-of-art in gene engineering due to their simple design, generation and modification compared to ZFNs and TALENs. As such, the 2020 Nobel prize in chemistry was recently awarded to Emmanuelle Charpentier and Jennifer Doudna for developing this genome editing method. CRISPR-Cas originates from an adaptive immune system common in prokaryotes such as bacteria and archaea.[164] These organisms integrate deleterious exogenous DNA (*e.g.* from hostile viruses and plasmids) within their genome in between copies of repeated sequences called clustered regularly interspaced short palindromic repeats (CRISPR). The process generates a cellular memory of the past invader and if the same exogenous DNA penetrates the cell, the prokaryote will recognize it and degrade it.[165] In detail, the cellular immune response starts with transcription of the CRISPR site to form pre-crRNA and with generation of a trans-activating CRISPR-RNA (tracrRNA) from a genomic locus upstream the CRISPR gene. Part of pre-crRNA and its complementary tracrRNA site anneal to form a dsRNA which elicit cleavage by RNAse III. This process yields the crRNA:tracrRNA complex which is then bound by a Cas enzyme to generate the active gene-editor. tracrRNA aids hybridisation of the RNA complex to the nuclease machinery while crRNA acts as a homing vector directing cleavage of the protein to the target sequence (Figure 10C).[166] In addition to crRNA, the Cas enzyme has intrinsic sequence selectivity provided by a DNA-binding domain that recognizes a specific nucleobase combination called proto-spacer adjacent motif (PAM). Various Cas nucleases with different PAM bias were found in prokaryotes and the most commonly used *Streptococcus pyogenes* SpCas9 binds 5′-NGG-3′ sequences through specific interaction between two arginine residues and the guanine nucleobases on the noncomplementary strand. When the CRISPR-Cas9 complex is formed the enzyme skims the genome searching for its PAM site. Here, arginine-guanine interaction unwinds DNA and aids identification of complementarity between the flanking DNA and sgRNA.[167] At the targeted loci, the cooperative effect of HNH and RuvC-like domains of Cas9 promotes blunt, hydrolytic cleavage of the complementary and non-complementary strands, respectively.[168]

Uncovering this adaptive immune mechanism quickly led to applications of CRISPR-Cas in gene-editing and ever since these systems have been engineered in order to improve their design, increase their binding specificity or widen their applications.[152,169] Jinek *et al*. joined the tracrRNA and crRNA to form a single guide RNA (sgRNA) that efficiently directed Cas9 cleavage.[164] Similarly both Smith and co-workers and Brown and co-workers used azide-alkyne chemical ligation to join a fixed tracrRNA strand to the variable crRNA sequence as an efficient tool for the high-throughput generation of wide sgRNA libraries.[170] In both cases, the triazole modification was placed at the loop structure of the upper stem and here the click linker length was proved to be non-essential for CRISPR-Cas9 activity. To further reduce the synthetic burden in the production of multiple variable crRNAs, Brown and co-workers placed the triazole modification also close to the protospacer adjacent motif (PAM) therefore dividing the design of the sgRNA in a fixed 79 mer tracrRNA and a variable 20 mer crRNA. In this case the triazole spacer length is however essential for efficient gene editing. Importantly, the long linker used significantly impaired Cas9-mediated cleavage whereas the biocompatible triazole—formed by cycloaddition between 5'-azide and a 3'-propargyl group—showed *in vitro* gene editing efficiency comparable to *in vitro* transcribed (IVT) sgRNA.

In parallel to the homing agent, the enzyme was modified and its activity optimized for different applications. Cas enzymes from other prokaryotes (*i.e.* Cas9 variants, Cpf1, Cas13) have been employed as editors due to their different sizes, PAM requirements and substrate preferences. The smaller sizes of NmCas9 (*Neisseria meningitides*), SaCas9 (*Staphylococcus aureus*), CjCas9 (*Campylobacter jejuni*) allow the therapeutic delivery challenge posed by packaging the large SpCas9 variant within Adeno associated viruses (AAV) to be overcome. As a trade-off to their reduced size, these nucleases however require longer and complex PAM sequences limiting the flexibility of their application.[171] Another class of Cas enzymes—the Cpf1 variants—cleave DNA in a different fashion to Cas9 at the 3′ site of 5′-TTN-3′ PAM yielding staggered rather than blunt ends.[172] The generated 5′-overhangs are of particular interest because

they can be exploited for introducing exogenous DNA fragments through complementation and ligation of the "sticky" ends.

Point mutations in Cas9 enzymes can be used to deactivate either one or both of the cleaving domains resulting in nickase or catalytically inactive variants.[173] Cas systems that can induce only single strand breaks (SSB) are particularly interesting as a solution to off-target effects often linked to gene editing with CRISPR-Cas technologies. In this case dimerization of two hybrids with sequence specificity on opposite strands is required for cleavage to occur.[174] On the other hand, catalytically inactive CRISPR-Cas can be employed to recognize, bind and deactivate specific genes of interest. Alternatively, the carboxy or amino terminus of 'dead' Cas (dCas) can be fused to fluorescent probes, transcription activators or epigenetic modifiers.[166] The latter case has particular therapeutic relevance because most of known human pathogenic mutations are represented by point modification—also known as single-nucleotide polymorphism (SNPs). Base editors crossbred to CRISPR-Cas systems can induce pinpointed epigenetic conversions of CG bp to AT bp (and *vice versa*) without requirement for deleterious strand breaks, hence reversing SNPs and restoring gene function.[173,175–177]

8.2. Synthetic chemical nucleases

For many years scientists have tried to develop chimeric systems that act as biomimetics of enzymatic nucleases exploiting the cooperation of a sequence-specific vector with a small molecule that promotes reactivity and cleavage at the DNA interface (Figure 10D). Strand scission is promoted either through hydrolysis of the phosphate backbone or by oxidative damage of the deoxyribose moiety by ROS production.[112,178] The type of DNA cleavage mechanism adopted by metal complexes is highly dependent on the nature of the metal centre and on its coordination environment. Hydrolytic strand scissions are particularly relevant for the introduction of genomic modifications as they yield 3′-OH and 5′-PO_4 products that are biocompatible for enzymatic manipulation. Conversely, DNA oxidation can provide a wide range of deoxyribose and nucleobase modifications—dependent on the reaction site and nature of the radical—which can result either in epigenetic modifications, DNA damage or gene knockdown. This type of cleavage has been extensively employed to prepare DNA structural probes, footprinting agents as well as gene-silencers.[178,179]

To achieve sequence specific cleavage the chemical nuclease can be either covalently bound to the homing agent or recognize specific secondary structures induced by hybridisation of the vector to DNA. In the covalent approach, the gene-editor is directly bound to the homing agent and sequence specific cleavage is caused by enhanced local concentration of the chemical nuclease at the scission site. For this purpose, a wide range of alkylating agents and complexes (*i.e.* Ce^{4+}, Fe^{2+}, Zr^{4+}, Rh^{2+} and Cu^{2+}) have been conjugated to various sequence-directing groups such as proteins, polyamides and polynucleotides (Figure 11).[180] This review will focus mainly on Fe(II)-EDTA and Cu-Phen hybrid systems.

8.2.1. AMN conjugation to sequence-specific binders

Fe(II)-EDTA was the first complex modified for targeted DNA oxidation by conjugation with the intercalator methydium-propylamine. Despite this derivative showing enhanced cleavage activity in the presence of dithiothreitol and efficiency comparable to bleomycin, it lacked any sequence specificity.[181] To restrict AMN activity to defined sequences, Dervan *et al*. linked the antibiotic distamycin (DE)—an oligopeptide containing three *N*-methylpyrrole carboxamides—to the complex by peptide chemistry.[182] Distamycin binds the minor groove of poly[d(A-T)$_2$] tracts through a combination of electrostatic interactions, hydrogen bonds and van der Waals contacts between the polyamide *N*-methylpyrrole groups and DNA nucleobases.[16,39,183] The Fe(II)-EDTA-DE derivative had higher specificity than bleomycin, cleaving fewer sites on restriction fragments from pBR322 and predominantly at 5′-ATTT-3′.[182] Extension of the targeting moiety to five *N*-methylpyrrole carboxamides enhanced DNA recognition and cleavage efficiency at 5′-TTTTTA-3′ sequences.[184] In parallel to the iron-EDTA conjugates, Cu-Phen and derivatives were conjugated to distamycin to direct nuclease action by the complex. Similar to free Fe(II)-EDTA, Cu-Phen complexes induce promiscuous DNA cleavage which, once tethered to distamycin, is restricted to nucleobases close to poly[d(A-T)$_2$] tracts. Among the Cu-Phen hybrids, 3-Clip-Phen had the highest cleavage efficiency suggesting relevance of the serinol bridge in oxygen activation. In this case DNA damage was slower but more efficient than the cleavage induced by the untethered AMN. In addition, while free $[Cu(3\text{-Clip-Phen})]^{2+}$ induce radical damage at C1′, C4′ and C5′ positions of the deoxyribose ring, in the case of the hybrid C1′ oxidation is reduced probably due to steric hindrance that promotes interaction at the edge of the minor groove.[39,185] Inspired by DNA recognition of distamycin and to expand specificity of polyamides to other sequences, Dervan *et al.* synthesized a series of hairpin pyrrole-imidazole polymers capable of recognizing defined Watson-Crick

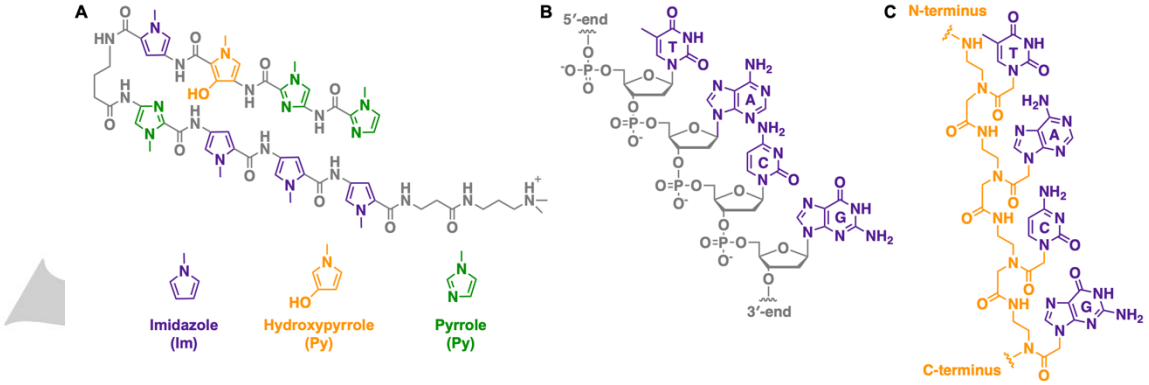

Figure. 11 Molecular structure of **A** polyamide, **B** DNA and **C** PNA based homing agents. Binding specificity by polyamides is provided by pairs of noncanonical amino acids capable of recognizing individual base pair combinations. In the case of DNA and PNA sequence recognition is given by base pairing of the nucleobases (purple).

Table 1 Examples of hybrid chemical editors where reactivity is targeted by DNA-specific binders, oligonucleotides or structural recognition (Ox = oxidative; Mut = mutation; Hyd = hydrolitic; Cov = covalent; Non-Cov = non-covalent; ss = single strand; ds = double strand).

Group	Hybrid	Cleavage	Approach	Substrate	Reference
AMN-DNA binder hybrids					
Dervan	Fe(II)-EDTA-EtBr	Ox	Cov	dsDNA	*J. Am. Chem. Soc.* 1982, *104*, 313.[178]
Dervan	Fe(II)-EDTA-Distamycin	Ox	Cov	dsDNA	*J. Am. Chem. Soc.* 1982, *104*, 6861.[179]
Meunier	3-Clip-Phen-Distamycin	Ox	Cov	dsDNA	*Nucleic Acid Res.* 2000, *28*, 4856.[39]
AMN-oligonucleotide (ON) hybrids					
Dervan	Fe(II)-EDTA-ON	Ox	Cov	ssDNA	*Proc. Natl. Acad. Sci. U.S.A.* 1985, *82*, 968.[180]
Orgel	Fe(II)-EDTA-ON	Ox	Cov	ssDNA	*Proc. Natl. Acad. Sci. U.S.A.* 1985, *82*, 963.[181]
Sigman	Cu(II)-Phen-ON	Ox	Cov	ssDNA	*Proc. Natl. Acad. Sci. U.S.A.* 1986, *83*, 7147.[182]
Dervan	Fe(II)-EDTA-ON	Ox	Cov	dsDNA	*Science* 1987, *283*, 645.[183]
Hélène	Cu(II)-Phen-ON	Ox	Cov	dsDNA	*Proc. Natl. Acad. Sci. U.S.A.* 1989, *86*, 9702.[184]
Glazer	Psoralen-ON	Mut	Cov	dsDNA	*Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 7879.[185]
Kodama	Ce(IV)-iminodiacetate-ON	Hyd	Cov	ssDNA	*Supramol. Chem.* 1994, *4*, 31.[186]
Krämer	Zr(IV)-tris(hydroxymethyl)-aminomethane-ON	Hyd	Cov	ssDNA	*Inorg. Chem.* 2003, *42*, 8618.[187]
Ganesh	Cu(II)/Co(III)- metallodesferal-ON	Ox	Cov	ss/dsDNA	*Biochim. Biophys. Acta* 1994, *1201*, 454.[188]
Meunier	Mn(III)-porphyrin-ON	Ox	Cov	dsDNA	*Nucleic Acids Res.* 1995, 23, 3894.[189]
Zarytova	Bleomycin-ON	Ox	Cov	ss/dsDNA	*Russ. Chem. Rev.* 1996, *65*, 355.[190]
Structure-specific AMNs					
Komiyama	Ce(IV)-EDTA	Hyd	Non-Cov	dsDNA	Chem. Soc. Rev 2011, 40, 5657.[191]
Hélène	Fe(II)-EDTA-BQQ	Ox	Non-Cov	H-DNA	Chem. Biol. 1999, 6, 771.[192]

base pairs through the minor groove of DNA (Figure 11A). In this strategy unsymmetrical ring pairs of three non-canonical amino acids (imidazole, pyrrole and hydroxypyrrole) are combined to distinguish the four base pair combinations (AT, TA, GC, CG) through the specific stereochemistry of their hydrogen bonds, differences in steric bulk and electronic potential surfaces. Later, the high DNA binding specificity of these homing agents were exploited to deliver various chemical gene silencers and site-selective alkylation and cleavage of DNA was achieved for a class of polyamides with 1-(chloromethyl)-5-hydroxy-1,2-dihydro-3H-benz[e]indole (seco-CBI) moiety.[16,186–188] The overall neutral backbone of polyamides is attractive for cell permeability and delivery but their difficult design and short sequence recognition restrict their targeting when compared to other vectors.

8.2.2. AMN-oligonucleotide hybrids

Oligonucleotides can be programmed to specifically bind long DNA regions and their recognition through canonical base pairing provides an easier design of the homing agent compared to polyamides. In this case, hydrogen bonds established between nucleobases of the vector—either a peptide- (PNA) or deoxyribonucleic acid strand (Figure 10B and C)—target direct reactivity of the chemical nuclease to the major groove of dsDNA.[189] Here, PNAs can induce various secondary structures dependent on vector length and sequence composition while deoxyribonucleic strands forms triple helical arrangements.[28,190]

Much elegant work has been dedicated to designing chimeric agents with oligonucleotides (ON) and sequence-specific reactivity was reported for conjugates carrying photo-inducible crosslinkers, alkylating groups and cleaving agents.[191–195] Fe(II)-EDTA was the first complex to be linked to a 19-nucleotide sequence targeting pBR322 by integration of a thymidine-EDTA-triethylester through phosphoramidite solid phase synthesis (Figure 12A).[196] Chu and Orgel also conjugated the ferrous-EDTA complex to an oligonucleotide but they used a different synthetic approach and directly attached it to the 5′-phosphate termini through a ethylenediamine linker (Figure 12B).[197] In the case of Dervan's hybrid, cleavage of a 167 bp amplicon occurred site-specifically over a range of 16 nucleotides from the site of hybridization suggesting that damage is produced by diffusible radicals, namely $^{\bullet}OH$.[196] The first example of $[Cu(Phen)]^{2+}$-ON hybrid was reported by Sigman *et al*. who conjugated glycylamido-1,10-phenanthroline to oligonucleotides through their phosphate-imidazolide intermediates. In this case targeted cleavage was promoting through hydroxyl radical production by $[Cu(Phen)]^{2+}$ in the presence of H_2O_2 but damage was restricted to fewer sites compared to iron-agents indicating that the ROS produced is not diffusible (Figure 12C).[198] Despite the above systems promoting targeted DNA cleavage, an essential requirement in the experimental design relied on disruption of the pre-existing duplex structure to allow hybridisation of the hybrid.

In 1987 Dervan and Moser were the first to exploit triplex

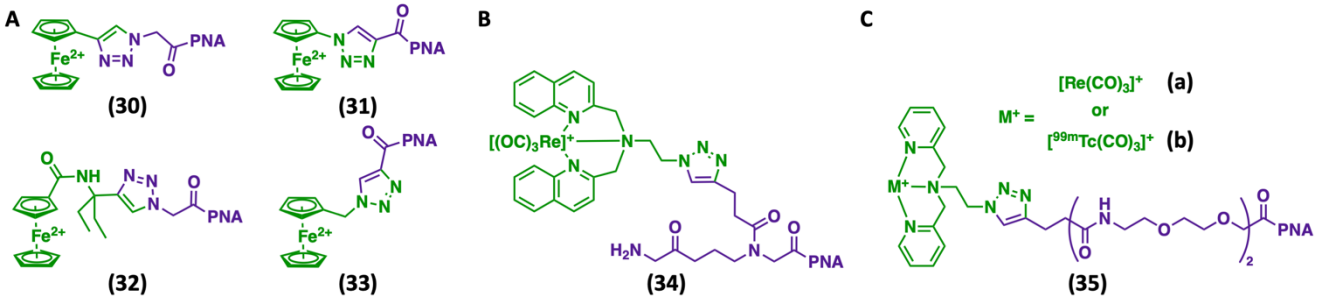

Figure. 13 PNA-hybrids developed by Metzler-Nolte and *co*-workers using click-chemistry as targeted electrochemical (A), fluorescent or radioactive (B and C) sensors.

formation and Hoogsteen recognition of polypurine sequences by polypyrimidine agents to develop nine Fe(II)-EDTA-homopyrimidine probes capable of causing localized double strand breaks in plasmid DNA.[199–201] In this work, sequence specificity by the targeted chemical nuclease relied significantly on: *i*) buffer composition in that presence of organic solvents induce transition from B- to A-DNA facilitating triplex structure formation; *ii*) acidic pH sensitivity promoting protonation of cytosine bases and triplex stabilization but causing parallel $(EDTA)^{4-}$ quenching and AMN deactivation; *iii*) probe length, temperature and sequence similarities; and *iv*) presence of added cations (i.e. $[Co(NH_3)_6]^{3+}$ or spermine) to reduce phosphate backbone repulsion and stabilize the H-DNA structure.[199] When positioned at the 5′-end of the TFO, the AMN moiety was exposed to the major groove and promoted oxidation of the polypurinic strand toward the 5′-side, indicating parallel directionality with respect to the target strand. Alternatively, internal modifications induced damage of the polypyrimidine sequence for the target purinic sequence is protected by the sugar-phosphate backbone of the TFO. In both cases, damage was distributed over a range of up to 16 bp due to the diffusible hydroxyl radical.[199] Later work by Hélène reported sequence specific photocrosslinking by a duplex-targeted *p*-azidophenacyl-octathymidylate derivative. Here, conversion of crosslinks into chain breaks under alkaline treatment allowed to identify a parallel orientation of the octathymidylate to the target strand.[202] Similarly, a phenanthroline-oligonucleotide conjugate reported by the same group induced site-specific cleavage of Simian virus 40 (SV40) DNA in the presence of copper ions and reductant (Figure 12D). The efficiency of double strand cleavage (70%) was higher than the one reported for Fe(II)-EDTA hybrid (25%).[200,203] In contrast the hybrid reported by Dervan, oxidative damage by the Cu(Phen)-ON was performed at the minor groove of DNA. This behaviour suggested that intercalation of phenanthroline occurs in the major groove while coordination of copper through the minor groove locks the complex close to its cleavage site. In addition, cleavage efficacy was enhanced when a second free Phen intercalator was added to form the active $[Cu(Phen)_2]^{2+}$ complex whereas damage was completely inhibited in the absence of spermine even at high salt concentrations (1 M NaCl).[192,203] With the aim of improving targeted damage, the Cu(Phen)-ON was modified on the opposite end (5'- or 3'- end) with an acridine derivative capable of dsDNA intercalation and triple helix stabilization.[192]

Inspired by these systems, reactivity by other chemical agents has been targeted to DNA by oligonucleotide vectors. Psoralen-TFO hybrids were reported to generate triplex mediated adducts upon irradiation both *in vitro* and *in vivo* and these systems have been employed as probes for chromatin structure, inhibitors of transcription and site-specific gene editors.[204–206] Chimeric conjugates of Zr^{4+} and Ce^{4+} with peptide nucleic acids or oligonucleotides, respectively, induced site-selective hydrolysis of single-stranded DNA.[207–209] In both cases hydrolytic cleavage was mediated by the strong Lewis acids coordinating first to a phosphate group and then triggering scission through water activation. More recent AMN-TFOs include hybrids with bleomycin A_5, Mn(III)-porphyrin and Cu(II)/Co(III)-metallodispheral complexes. All of these chimeric agents showed efficient cleavages of short dsDNA sequences.[210–212] Despite various strategies being employed to induce specific cleavage, conjugation of AMNs to oligonucleotide vectors usually requires complex procedures and usage of linkers easily cleavable in cellular conditions. In addition, most of the aforementioned work

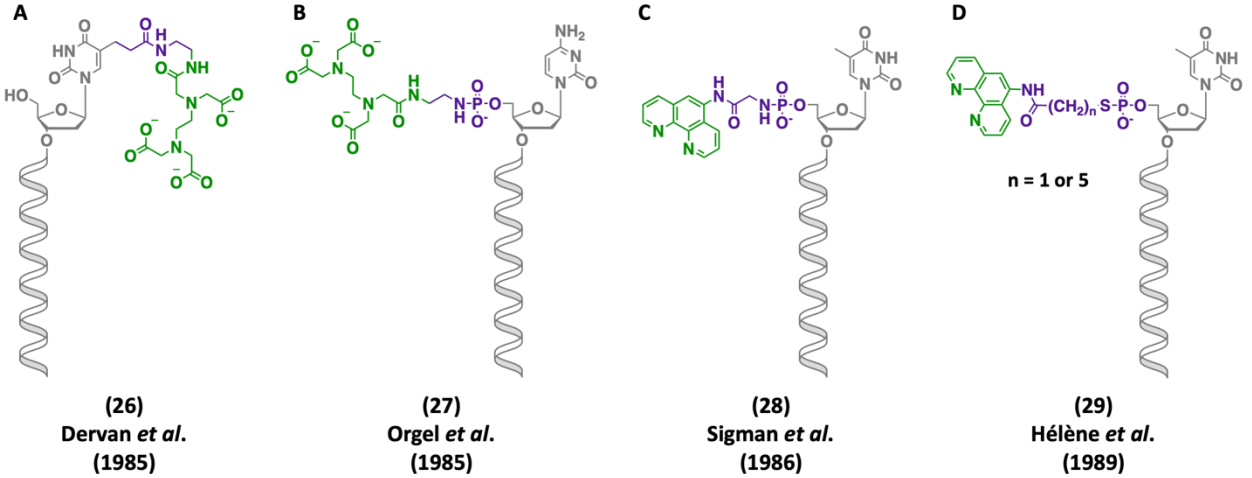

Figure. 12 Molecular structure of the state-of-art targeted-AMNs where an oligonucleotide vector has been coupled to either EDTA (**A** and **B**) or Phen (**C** and **D**) through peptide linkers (highlighted in purple).

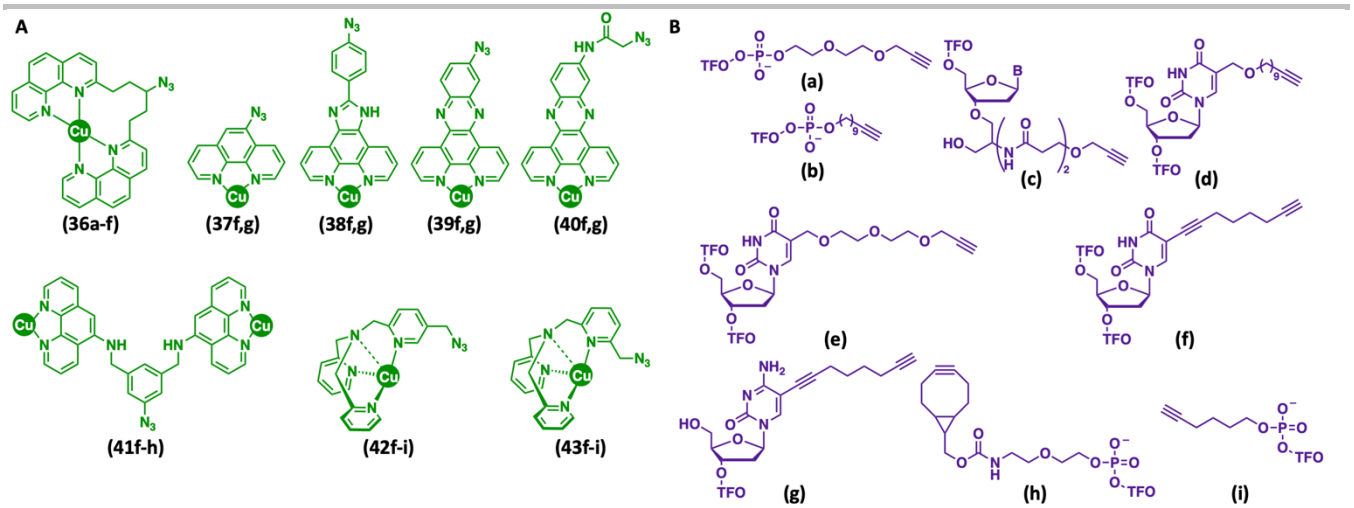

Figure. 14 A Clicked AMN-TFO hybrids developed by Hocek and co-workers (**36**) and Kellett and co-workers (**37**-**43**) as sequence targeted gene-knockout agents. **B** Alkyne spacers used to conjugate the AMN to the oligonucleotide strand.

on TFO-hybrids relied on the presence of polycationic species for triplex formation.

8.2.3 A Click-Chemistry approach for the generation of targeted AMN-TFO hybrids

Nicolo – there seems to be a problem with references in this section. They are missing.

In the context of sequence targeted metallo-nucleases, the capability of efficiently linking metal complexes with different reactivity to various sequences recognizing multiple genomic sites is of paramount importance for the quick and efficient generation and screening of wide libraries of AMN-TFO hybrids. For this reason, azide-alkyne cycloadditions provide an attractive alternative to the methods previously reported as *i.*) most molecules can be easily modified with azide or alkyne groups; *ii.*) organic azides and alkynes have usually high chemical stability; and *iii.*) azide and alkyne functionalities can only react with each other and are chemically orthogonal towards most functional moieties of biological substrates.

Wide libraries of small-molecule oligonucleotide hybrids have been generated using click conjugations either to provide nucleic acids with defined physico-chemical properties or to exploit their nucleobase recognition for directing different types of chemical reactivity to specific sequences. To increase cellular uptake RNA was modified by click chemistry with different azide modified receptor-binding molecules including 3-cholesterol, 3'-folate, and 3'-anandamide.[Ref] The anandamide-RNA hybrids had high cellular uptake and the modification permitted efficient siRNA delivery into difficult-to-transfect neuronal model RBL-2H3 cells. Application of this system to human immune cells (BJAB) showed similar silencing effects as those achieved using cationic transfection reagents but representing a nontoxic alternative to the latter class. Similarly, metal complexes have been conjugated to nucleic acid derivatives by click chemistry enabling the efficient generation of hybrid libraries with diverse functionalities dependent on the metal centre and the shape of the coordinating scaffold. Hüsken *et al*. generated a small library of ferrocenyl bioconjugates with peptide nucleic acids (PNA) where different electro-chemical potentials—provided by four ferrocene chemotypes—could be encoded into PNA oligomer sequences using the same CuAAC synthetic conditions.[Ref] The wide potential difference (> 60 mV) exhibited by the four clicked ferrocenyl-PNA hybrids is distinguishable by electrochemical methods (*i.e*. differential pulse voltammetry) and provides a "four-potential" labelling with electrochemical probes that resembles the "four-colour" detection with different fluorophores in classical DNA analysis. This strategy opens potential applications for PNA in the field of electrochemical nucleic acid biosensors. Click chemistry was exploited by the same group to conjugate a $[Re(CO)_3(L–N_3)]^+$ (where L-N_3 = (2-azido-N,N-bis((quinolin-2-yl)methyl)ethanamine) and a $[^{99m}Tc(CO)_3(DPA\text{-}N_3)]$+ (where DPA-$N_3$ = di-(2-pycolyl)amine) complex to PNA as a fluorescent or radioactive probe, respectively.[Ref] The rhenium hybrid was detectable in living cells providing effective bioimaging at a concentration of 10 μM by fluorescent microscopy. In addition, a Re-PNA conjugate targeting the eGFP in genetically modified HeLa cells sequence decreased eGFP expression by 18% after incubation, suggesting the possible application of the hybrid as an antisense agent. When cells were treated with a conjugate containing a mismatched PNA sequence, no change in protein expression was observed further supporting the antisense effect hypotesis. Biodistribution analysis of the Technetium-PNA hybrid in Wistar rat or mouse models (NMRI nu/nu) identified fast blood clearance and low accumulation in the kidneys.

Recently, the ClickGene consortium[Ref – website] applied click chemistry to conjugate various copper-based AMNs to triplex forming oligonucleotides. Using this strategy wide libraries of TFO-targeted AMNs were efficiently generated and evaluated for their copper-complex structure-activity relationship and sequence-directed oxidative damage. The artificial nuclease Clip-Phen machinery was coupled by Hocek *et al*. to different positions (i.e. 5′ -end, **36a** and **36b**; 3′ -end, **36c**; or internal, **36d-f**) of TFO probes using various linkers differing in length, flexibility and polarity (Figure 14). The AMN-TFOs damage was assessed using a target and an off-target—non-complementary by Hoogsteen base pairing—dsDNA sequences. Under optimized conditions, the hybrids with Clip-Phen linked internally or at the 5'-end of TFOs by a flexible linker induced significant cleavage of the target duplex (up to 34%) with no off-target effects. The absence of clear foot-printing cleavage indicated that diffusible radicals are produced in close proximity to the target DNA by the copper source and DNA-oxidation occurs sequence-specifically rather than with single-nucleotide precision.

In parallel, two studies carried out by Kellett *et al.* focused on conjugating either phenanthrene (**37**-**41**) or TPMA (**42**-**43**) derivatives (Figure 14) to TFOs of various length by click chemistry. The extended aromaticity of phenanthrene dsDNA intercalators covalently attached to the TFO significantly stabilized the triplex structure upon hybridisation and increased the its melting temperature with ΔT_M up to 22 °C. When a fluorophore-labelled dsDNA target was treated with a Cu-AMN-TFO (where **39** is the AMN scaffold) a small extent sequence-specific oxidative damage was identified. However, high-resolution cleavage at the single nucleotide level was not, at this point, achieved. To enhance cleavage selectivity a bis-Phen di-copper compound (**41**) was linked to three of the most promising TFO sequences. In this case, cleavage was assessed in the presence and absence of ancillary intercalators—such as Phen, DPQ and DPPZ—coordinated to both copper centers. Although limited cleavage discrimination between an *on*- and *off*-target sequence was achieved upon treatment with $Cu(NO_3)_2$, selective damage markedly increased upon complexation of **41** to Cu-Phen or Cu-DPQ. In these conditions, the target duplex was completely ablated without any apparent damage of the off-target sequence. This result further supports previous findings where an additional ancillary intercalator enhanced cleavage efficacy by a $[Cu(Phen)_2]^{2+}$-ON type hybrid.

Promising sequence selectivity was also achieved for Cu-TPMA-TFOs hybrids (**42**-**43**). Similar to the Clip-Phen-TFO derivatives, the best sequence discrimination was achieved when the metallo-nuclease was covalently linked at an internal position of the TFO sequence (**42f**). In this case, target duplex depletion was higher than off-target damage (analysed by gel densitometry 72% versus 24% knockdown at 100 eq. Na-L-asc, respectively). Interestingly, when the oxidative damage of the hybrids was analysed by real time PCR, DNA cleavage by the clicked AMN was identified also in the absence of exogenous reductant. On the contrary, no effect on duplex integrity occurred when the target DNA was treated with free Cu-TPMA in the absence of reductant. In the presence of reductant, cleavage by the hybrids where the nuclease was covalently attached to the probe strand was considerably higher than in the case of the free AMN. All together, these results prove the role of the TFO strand in directing the oxidative action of the AMN copper complex. To understand the oxidative damage mechanism by the hybrids, a larger DNA construct—the closed circular pCSanDI-HYG plasmid—was treated with an internal AMN-modified hybrid recognizing a specific sequence of its genome. Cleavage experiments conducted in the absence or presence of ROS scavengers identified conversion of supercoiled (SC) plasmid DNA to the nicked open circular (OC) and cleaved linear (L) form through an oxidative mechanism caused (primarily) by production of superoxide type ($O_2^{\bullet-}$) radicals. In contrast to the phenanthrene derivatives, conjugation of the polypicolyl ligands to TFOs caused a decrease of triplex stabilization probably caused by the bulk of the ligand. Nonetheless, increase in triplex stability was observed upon coordination of the organic scaffold to copper (ΔTM up to ~6 °C) and augmented when the AMN-TFO was modified with intercalating thiazole orange (TO) groups (ΔTM up to 18.4 °C).

Although AMN-TFOs do not yet show the same single-nucleotide precision of cleavage found in enzymatic nucleases, the extent and oxidative nature of their damage render these hybrids highly attractive sequence-selective knockdown agents. Nonetheless, to effectively apply this technology for *in vitro* editing, the TFO probe stability to endo or exo-nucleases should be increased by inclusion of DNA backbone or ribose modifications such as phosphorothioate backbones, 2'-*O*-methyl, or inverted dT functionalities.

8.2.4 Targeting AMNs by non-covalent structural recognition

Targeted cleavage can also be achieved using a non-covalent approach. Here, a biphasic mechanism is involved: *i.*) in the first event, the vector binds and forms a specific type of secondary or tertiary structures; and ii.) in the second event a conformation specific molecule binds the target site and promotes reactivity. Recently, the non-covalent approach has been reported for a class of compounds (ARCUT) that showed higher hydrolysis rates for single-stranded DNA (ssDNA) than for dsDNA. In this case two oligonucleotides were used for dsDNA invasion or to form a gap structure on ssDNA where hydrolysis by a Ce(IV)/EDTA complex was directed toward the ssDNA/dsDNA junctions. Cleavage bias by Ce(IV)/EDTA is due to its different binding constant for ssDNA which is 100 fold higher than for dsDNA.[213,214] The stronger binding originates from the flexible structure of ssDNA that allows a conformational change where three or more phosphodiester groups simultaneously coordinate to the Ce(IV) centre. Complexation to EDTA was essential for discrimination as the 'naked' metal was found to hydrolyse both ssDNA and dsDNA with similar efficiencies. These systems have been successfully employed to introduce genome edits both *in vitro* and *in vivo*.[215] Similarly, formation of triple helical structure has been exploited to direct cleavage by triplex-specific agents. Pentacyclic benzoquinoquinoxaline derivatives (BQQ) bind weakly to double-helical DNA but were shown to strongly stabilize triple helical structures. Therefore, the specific recognition and strong structural discrimination of these intercalators has been exploited to efficiently direct the nuclease action of Fe(II)-EDTA and structure-specific cleavage was achieved both on short linear and plasmid DNA sequences.[216,217]

9. Summary and Outlook

Nicolo – some references are missing here too.

The structural features of nucleic acids offer immense flexibility for the design of small molecules capable of site-specific recognition. Binding by these compounds often results in an alteration of the DNA structure where the extent and reversibility is dependent on the nature of the binding interaction. Due to recent advancements in preparative nucleic acid chemistry, our knowledge in the design of compounds capable of binding or modifying specific genomic sites is now enabling the design of highly selective diagnostic probes and personalised therapies.

In this context, metal complexes are of particular interest due to the variety of binding modes and reactivities they offer. These features are dependent on the coordination number, shape and charge of the inorganic scaffold. Starting with classical binders, cisplatin and derivatives were shown to covalently bind the N7-nitrogen atom of purine bases yielding 1,2-intrastrand crosslinks thereby kinking the duplex helical axis. More recently, research has shifted to probing non-covalent DNA binding modes directed by polynuclear platinum complexes (PPCs). Unlike cisplatin, these complexes are substitution inert and contain a *trans*-Pt(II) geometry where NH_3 or 'dangling am(m)ines' replace base-reactive chlorides groups. Here, binding occurs via hydrogen bond interactions between the am(m)ine groups and the oxygens of the phosphate backbone to produce a N-O-N

clamp. Both the mononuclear *cis*-Pt(II) and PPCs show high antitumoral activity with cisplatin and several other mononuclear derivatives widely applied anticancer agents. In parallel, complexes with different metals have been designed and studied with the aim of developing anticancer drugs with lower systematic toxicity and different pharmaco-kinetic and -dynamic profiles than cisplatin. One such agent is a photoactive ruthenium(II) compound called TLD1433 which is capable of $^1O_2$ sensitization and electron-transfer upon light-irradiation and was shown to induce single strand DNA cleavage at nanomolar concentrations. TLD1433 has entered human clinical trials due to its extremely potent cytotoxicity.[Nicolo – please see if there is an update on this trial]

Copper phenanthrene complexes such as $[Cu(1,10\text{-phenantroline})_2]^{2+}$ can cause oxidative DNA strand cleavage by ROS production upon DNA intercalation and reduction of the copper center with exogenous reductants (*i.e*. Na-*L*-ascorbate). However, the low stability of copper-phenanthrene compounds in solution—where the second coordinated ligand dissociates yielding the inactive monocoordinated complex—limits their therapeutical application. To aid complex stabilization, intensive design efforts have been carried out on this ligand scaffold. One approach involved the introduction of a serinol bridge to covalently link two 1,10-pheanthroline ligands to favor the 2:1 Phen:Cu ratio required for efficient DNA cleavage activity. More recently, polypyridyl ligands were used in order to 'cage' the copper metal center while also providing space for a coordinated ancillary *N*,*N*-phenanthrene intercalator.[ref, ref] Within this class, three Cu-DPA-Phenanthrene complexes recently isolated showed outstanding cytotoxicity against various pancreatic cancer cell lines with activities, in some cases, surpassing the reference drug oxaliplatin.[ref] Sequence selective interactions by copper complexes have also been demonstrated in recent years. By increasing the nuclearity of polypyridyl copper complexes, enhanced oxygen activation and cleavage at ssDNA/dsDNA junctions could be afforded. Another example was recently identified using dinuclear complex—Cu-Oda, $[\{Cu(Phen)_2\}_2(\mu\text{-Oda})]^{2+}$—which showed discrimination of AT/AT from TA/TA steps of dodecamer palindromic sequences through the induction of a B- to Z-like DNA modification.[42]

The application of metal complexes capable of interaction with nucleic acids also expands beyond therapeutical purposes. Oxidative-mediated DNA excision induced by several copper complexes provides an exceptional alternative to enzymatic nucleases in DNA footprinting assays where aspecific cleavage is needed to generate even distribution of high-resolution cleaved fragments. Recently, Cu-Phen has also been employed in protein engineering to generate recombinant libraries of prostate specific antigen (PSA) antibodies by DNA shuffling. Here, Cu-Phen afforded accessibility to more gene mutations than DNAse I thereby demonstrating a possible application for directed evolution by DNA shuffling.

Finally, by directing the cleavage activity of metal complexes using (covalently or non-covalently bound) homing agents, a potential alternative to the state-of-art enzymatic nucleases were developed. In contrast to the enzymatic nucleases, targeted-AMNs cleave DNA using an oxidative mechanism and therefore offer the potential for permanent gene knockout. The small size of the metal complex might also aid cellular transfection compared to bulky enzymatic nucleases. In addition, the careful design of the sequence-specific probe enables guidance of a metal complex to either the minor or major groove of dsDNA. As such, covalently linked polyamide probes have been used to direct the chemical reactivity of various small-molecule compounds to the minor groove of dsDNA. Alternatively, PNA and DNA have been used to generate antisense and antigene hybrids where sequence recognition arises from Watson-Crick or Hoogsteen base pairing with nucleobases in the major groove. In parallel, click chemistry has enabled the generation of chimeric AMN-TFOs with desirable targeted effects with limited *off*-target cleavage effects. Three recent examples of copper AMN-TFOs have been developed and these include: clip-Phen modified TFOs[ref]; polypyridyl-modified TFO [ref]; and intercalating phenanthrene-modified TFOs [ref]. All three systems have distinct advantages in either maintaining a desired 2:1 phenthroline:copper ratio needed for DNA cleavage (clip-Phen TFOs), for caging the copper centre and facilitating efficient ROS-mediated strand scission (polypyridyl-modified TFO), and in providing dual targeted cleavage with enhanced of triplex stability where the repulsion energy exerted by introducing a third strand in the minor groove is minimised (phenanthrene-TFOs). Another important class of targeted chemical nucleases are Ce(IV)/EDTA complexes that were recently shown to bind and hydrolytically cleave ssDNA/dsDNA junctions and have been used in combination with PNA to successfully introduce genome edits both *in vitro* and *in vivo*.[ref]

In summary, the chemical reactivity and wide flexibility of metal complex design together with their coupling to sequence specific probes for directed applications clearly demonstrates a range of untapped possibilities of these compounds in diverse biological fields including chemotherapy, protein engineering, DNA footprinting and gene editing.

Acknowledgements

AK and TB acknowledge funding from the Marie Skłodowska-Curie Innovative Training Network (ITN) ClickGene (H2020-MSCA-ITN-2014-642023). TB acknowledges funding from the UK BBSRC (grants BB/J001694/2 and BB/R008655/1). [AK to add SFI]

Keywords: metal complexes • nucleic acids • artificial metallo nucleases • DNA targeting • inorganic chemotherapeutics

[1] L. H. Hurley, *Nat. Rev. Cancer* **2002**, *2*, 188–200.
[2] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.
[3] B. Rosenberg, L. V. Camp, T. Krigas, *Nature* **1965**, *205*, 698–699.
[4] E. Chargaff, *Science* **1971**, *172*, 637–642.
[5] S. Harteis, S. Schneider, *Int. J. Mol. Sci.* **2014**, *15*, 12335–12363.
[6] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature* **1980**, *287*, 755–758.
[7] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179–2183.
[8] P. S. Ho, M. Carter, in *DNA Replication - Current Advances* (Ed.: H. Seligmann), IntechOpen, London, United Kingdom, **2011**, pp. 4–28.
[9] A. Rich, S. Zhang, *Nat. Rev. Genet.* **2003**, *4*, 566–572.
[10] A. Herbert, *Commun. Biol.* **2019**, *2*, 1–10.
[11] A. Medina-Molner, B. Spingler, *Chem. Commun.* **2012**, *48*, 1961–1963.
[12] C. J. Cardin, in *Structure and Bonding*, Springer, Berlin, Heidelberg, **2019**, pp. 1–35.

[13] V. A. Bloomfield, D. M. Crothers, I. Tinoco, *Nucleic Acids: Structures, Properties, and Function.*, University Science Books, Sausalito, **2000**.
[14] P. Lincoln, L. M. Wilhelmsson, B. Nordén, in *DNA-Targeting Molecules as Therapeutic Agents* (Ed.: M.J. Waring), Cambridge, United Kingdom, **2018**, pp. 45–73.
[15] H.-K. Liu, P. J. Sadler, *Acc. Chem. Res.* **2011**, *44*, 349–59.
[16] P. B. Dervan, *Isr. J. Chem.* **2019**, *59*, 71–83.
[17] A. Kellett, Z. Molphy, C. Slator, V. McKee, N. P. Farrell, *Chem. Soc. Rev.* **2019**, *48*, 971–988.
[18] L. Stefan, B. Bertrand, P. Richard, P. L. Gendre, F. Denat, M. Picquet, D. Monchaud, *ChemBioChem* **2012**, *13*, 1905–1912.
[19] S. Neidle, S. Balasubramanian, Eds. , *Quadruplex Nucleic Acids*, Royal Society Of Chemistry, Cambridge, United Kingdom, **2007**.
[20] D. R. Boer, A. Canals, M. Coll, *Dalton Trans.* **2008**, 399–414.
[21] H. J. Lipps, D. Rhodes, *Trends Cell Biol.* **2009**, *19*, 414–422.
[22] M. Guéron, J.-L. Leroy, *Curr. Opin. Struct. Biol.* **2000**, *10*, 326–331.
[23] M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger, D. Christ, *Nat. Chem.* **2018**, *10*, 631–637.
[24] P. A. Khuu, A. R. Voth, F. A. Hays, P. S. Ho, *J. Mol. Recognit.* **2006**, *19*, 234–242.
[25] A. L. Brogden, N. H. Hopcroft, M. Searcey, C. J. Cardin, *Angew. Chem. Int. Ed.* **2007**, *46*, 3850–3854.
[26] L. A. Howell, Z. A. E. Waller, R. Bowater, M. O'Connell, M. Searcey, *Chem. Commun.* **2011**, *47*, 8262–8264.
[27] K. R. Fox, T. Brown, *Biochem. Soc. Trans.* **2011**, *39*, 629–634.
[28] K. R. Fox, T. Brown, D. A. Rusling, in *DNA-Targeting Molecules as Therapeutic Agents* (Ed.: M.J. Waring), Royal Society Of Chemistry, Cambridge, **2018**, pp. 1–32.
[29] F. A. Buske, J. S. Mattick, T. L. Bailey, *RNA Biol.* **2011**, *8*, 427–439.
[30] Y. Taniguchi, S. Sasaki, *Org. Biomol. Chem.* **2012**, *10*, 8336–8341.
[31] A. R. Morgan, R. D. Wells, *J. Mol. Biol.* **1968**, *37*, 63–80.
[32] J. C. Hanvey, M. Shimizu, R. D. Wells, *Nucleic Acids Res.* **1990**, *18*, 157–161.
[33] M. Faria, C. Giovannangeli, *J. Gene Med.* **2001**, *3*, 299–310.
[34] A. Kellett, Z. Molphy, C. Slator, V. McKee, N. P. Farrell, *Chem. Soc. Rev.* **2019**, *48*, 971–988.
[35] N. McStay, A. M. Reilly, N. Gathergood, A. Kellett, *ChemPlusChem* **2019**, *84*, 38–42.
[36] N. McStay, Z. Molphy, A. Coughlan, A. Cafolla, V. McKee, N. Gathergood, A. Kellett, *Nucleic Acids Res.* **2017**, *45*, 527–540.
[37] W. D. Wilson, F. A. Tanious, H. J. Barton, R. L. Jones, K. Fox, R. L. Wydra, L. Strekowski, *Biochemistry* **1990**, *29*, 8452–8461.
[38] U. Sehlstedt, S. K. Kim, B. Norden, *J. Am. Chem. Soc.* **1993**, *115*, 12258–12263.
[39] M. Pitié, C. J. Burrows, B. Meunier, *Nucleic Acids Res.* **2000**, *28*, 4856–4864.
[40] M. J. Waring, *J. Mol. Biol.* **1965**, *13*, 269–282.
[41] C. G. Reinhardt, T. R. Krugh, *Biochemistry* **1978**, *17*, 4845–4854.
[42] C. Slator, Z. Molphy, V. McKee, C. Long, T. Brown, A. Kellett, *Nucleic Acids Res.* **2018**, *46*, 2733–2750.
[43] H.-K. Liu, P. J. Sadler, *Acc. Chem. Res.* **2011**, *44*, 349–359.
[44] S. Walsh, A. H. El-Sagheer, T. Brown, *Chem. Sci.* **2018**, *9*, 7681–7687.
[45] A. R. Morgan, J. S. Lee, D. E. Pulleyblank, N. L. Murray, D. H. Evans, *Nucleic Acids Res.* **1979**, *7*, 547–565.
[46] J. P. Hall, P. M. Keane, H. Beer, K. Buchner, G. Winter, T. L. Sorensen, D. J. Cardin, J. A. Brazier, C. J. Cardin, *Nucleic Acids Res.* **2016**, *44*, 9472–9482.
[47] C. Slator, Z. Molphy, V. McKee, A. Kellett, *Redox Biol.* **2017**, *12*, 150–161.
[48] G. Toniolo, M. Louka, G. Menounou, N. Z. Fantoni, G. Mitrikas, E. K. Efthimiadou, A. Masi, M. Bortolotti, L. Polito, A. Bolognesi, A. Kellett, C. Ferreri, C. Chatgilialoglu, *ACS Omega* **2018**, *3*, 15952–15965.
[49] J. Andersson, P. Lincoln, *J. Phys. Chem. B* **2011**, *115*, 14768–14775.
[50] L. Wu, A. Reymer, C. Persson, K. Kazimierczuk, T. Brown, P. Lincoln, B. Nordén, M. Billeter, *Chem. Eur. J.* **2013**, *19*, 5401–5410.
[51] Institute of Medicine (US) Committee on the Survey of the Health Effects of Mustard Gas and Lewisite, in *Veterans at Risk The Health Effects of Mustard Gas and Lewisite* (Eds.: C.M. Pechura, D.P. Rall), National Academies Press (US), Washington D.C., **1993**, pp. 71–80.
[52] M. J. Waring, *Annu. Rev. Biochem.* **1981**, *50*, 159–192.
[53] J. G. Muller, S. J. Paikoff, S. E. Rokita, C. J. Burrows, *J. Inorg. Biochem.* **1994**, *54*, 199–206.
[54] C. Mari, V. Pierroz, S. Ferrari, G. Gasser, *Chem. Sci.* **2015**, *6*, 2660–2686.
[55] E. Wachter, A. Zamora, D. K. Heidary, J. Ruiz, E. C. Glazer, *Chem. Commun.* **2016**, *52*, 10121–10124.
[56] A. F. Westendorf, J. A. Woods, K. Korpis, N. J. Farrer, L. Salassa, K. Robinson, V. Appleyard, K. Murray, R. Grünert, A. M. Thompson, P. J. Sadler, P. J. Bednarski, *Mol. Cancer Ther.* **2012**, *11*, 1894–1904.
[57] E. Wachter, D. K. Heidary, B. S. Howerton, S. Parkin, E. C. Glazer, *Chem. Commun.* **2012**, *48*, 9649–9651.
[58] F. S. Mackay, J. A. Woods, H. Moseley, J. Ferguson, A. Dawson, S. Parsons, P. J. Sadler, *Chem. Eur. J.* **2006**, *12*, 3155–3161.
[59] A. Kellett, O. Howe, M. O'Connor, M. McCann, B. S. Creaven, S. McClean, A. Foltyn-Arfa Kia, A. Casey, M. Devereux, *Free Radic. Biol. Med.* **2012**, *53*, 564–576.
[60] J. Cadet, J. R. Wagner, *Cold Spring Harb. Perspect. Biol.* **2013**, *5:a012559*, 1–18.
[61] P. C. Bruijnincx, P. J. Sadler, *Curr. Opin. Chem. Biol.* **2008**, *12*, 197–206.
[62] C. Marzano, M. Pellei, F. Tisato, C. Santini, *Anticancer Agents Med. Chem.* **2009**, *9*, 185–211.
[63] C. Santini, M. Pellei, V. Gandin, M. Porchia, F. Tisato, C. Marzano, *Chem. Rev.* **2014**, *114*, 815–862.
[64] N. P. E. Barry, P. J. Sadler, *Chem. Commun.* **2013**, *49*, 5106–31.
[65] A. Spassky, D. S. Sigman, *Biochemistry* **1985**, *24*, 8050–8056.
[66] N. Fantoni, T. Lauria, A. Kellett, *Synlett* **2015**, *26*, 2623–2626.
[67] G. B. Kauffman, R. Pentimalli, S. Doldi, M. D. Hall, *Platin. Met. Rev.* **2010**, *54*, 250–256.
[68] G. H. W. Milburn, M. R. Truter, *J. Chem. Soc. A* **1966**, 1609–1616.
[69] F. Muggia, R. Leone, A. Bonetti, *Anticancer Res.* **2014**, *34*, 417–417.
[70] H. Shi, C. Imberti, P. J. Sadler, *Inorg. Chem. Front.* **2019**, *6*, 1623–1638.
[71] L. Kelland, *Nat. Rev. Cancer* **2007**, *7*, 573–584.
[72] X. Du, X. Wang, H. Li, H. Sun, *Metallomics* **2012**, *4*, 679–685.
[73] T. C. Johnstone, K. Suntharalingam, S. J. Lippard, *Chem. Rev.* **2016**, *116*, 3436–3486.
[74] M. J. Hannon, *Chem. Soc. Rev.* **2007**, *36*, 280–295.
[75] A. Ratanaphan, S. Wasiksiri, B. Canyuk, P. Prasertsan, *Cancer Biol. Ther.* **2009**, *8*, 890–898.
[76] B. Lippert, *Coord. Chem. Rev.* **1999**, *182*, 263–295.
[77] R. Oun, Y. E. Moussa, N. J. Wheate, *Dalton Trans.* **2018**, *47*, 6645–6653.
[78] V. Brabec, J. Kasparkova, V. Menon, N. P. Farrell, in *Metallo-Drugs: Development and Action of Anticancer Agents* (Eds.: A. Sigel, H. Sigel, E. Freisinger, R.K.O. Sigel), De Gruyter, Berlin, Boston, **2018**, pp. 43–68.
[79] J. B. Mangrum, N. P. Farrell, *Chem. Commun.* **2010**, *46*, 6640–6650.
[80] N. Farrell, S. Spinelli, in *Uses of Inorganic Chemistry in Medicine* (Ed.: N.P. Farrell), Royal Society Of Chemistry, Cambridge, **1999**, pp. 124–134.
[81] J. Kasparkova, N. Farrell, V. Brabec, *J. Biol. Chem.* **2000**, *275*, 15789–15798.
[82] M. S. Davies, D. S. Thomas, A. Hegmans, S. J. Berners-Price, N. Farrell, *Inorg. Chem.* **2002**, *41*, 1101–1109.
[83] N. M. P. Rosa, F. H. do C. Ferreira, N. P. Farrell, L. A. S. Costa, *Front. Chem.* **2019**, *7:307*, 1–12.
[84] S. Komeda, T. Moulaei, K. K. Woods, M. Chikuma, N. P. Farrell, L. D. Williams, *J. Am. Chem. Soc.* **2006**, *128*, 16092–16103.
[85] a. Prisecaru, Z. Molphy, R. G. Kipping, E. J. Peterson, Y. Qu, a. Kellett, N. P. Farrell, *Nucleic Acids Res.* **2014**, *42*, 13474–13487.
[86] A. Levina, A. Mitra, P. A. Lay, *Metallomics* **2009**, *1*, 458–470.
[87] E. S. Antonarakis, A. Emadi, *Cancer Chemother. Pharmacol.* **2010**, *66*, 1–9.
[88] C. X. Zhang, S. J. Lippard, *Curr. Opin. Chem. Biol.* **2003**, *7*, 481–489.
[89] N. P. E. Barry, P. J. Sadler, *Chem. Commun.* **2013**, *49*, 5106–5131.

[90] S. Monro, K. L. Colón, H. Yin, J. Roque, P. Konda, S. Gujar, R. P. Thummel, L. Lilge, C. G. Cameron, S. A. McFarland, *Chem. Rev.* **2019**, *119*, 797–828.
[91] S. Leijen, S. A. Burgers, P. Baas, D. Pluim, M. Tibben, E. van Werkhoven, E. Alessio, G. Sava, J. H. Beijnen, J. H. M. Schellens, *Invest. New Drugs* **2015**, *33*, 201–214.
[92] C. G. Hartinger, S. Zorbas-Seifried, M. A. Jakupec, B. Kynast, H. Zorbas, B. K. Keppler, *J. Inorg. Biochem.* **2006**, *100*, 891–904.
[93] E. Alessio, L. Messori, *Molecules* **2019**, *24*, 1995.
[94] S. Bonnet, *Comment Inorg. Chem.* **2015**, *35*, 179–213.
[95] S. Lazic, P. Kaspler, G. Shi, S. Monro, T. Sainuddin, S. Forward, K. Kasimova, R. Hennigar, A. Mandel, S. McFarland, L. Lilge, *Photochem. Photobiol.* **2017**, *93*, 1248–1258.
[96] H. Huang, S. Banerjee, K. Qiu, P. Zhang, O. Blacque, T. Malcomson, M. J. Paterson, G. J. Clarkson, M. Staniforth, V. G. Stavros, G. Gasser, H. Chao, P. J. Sadler, *Nat. Chem.* **2019**, 1–8.
[97] F. Harris, L. Pierpoint, *Med. Res. Rev.* **2012**, *28*, 1292–1327.
[98] J. Liu, C. Zhang, T. W. Rees, L. Ke, L. Ji, H. Chao, *Coord. Chem. Rev.* **2018**, *363*, 17–28.
[99] G. Shi, S. Monro, R. Hennigar, J. Colpitts, J. Fong, K. Kasimova, H. Yin, R. DeCoste, C. Spencer, L. Chamberlain, A. Mandel, L. Lilge, S. A. McFarland, *Coord. Chem. Rev.* **2015**, *282–283*, 127–138.
[100] B. S. Howerton, D. K. Heidary, E. C. Glazer, *J. Am. Chem. Soc.* **2012**, *134*, 8324–8327.
[101] S. M. Hecht, *J. Nat. Prod.* **2000**, *63*, 158–168.
[102] W. Wu, D. E. Vanderwall, J. Stubbe, J. W. Kozarich, C. J. Turner, *J. Am. Chem. Soc.* **1994**, *116*, 10843–10844.
[103] A. D. D'Andrea, W. A. Haseltine, *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 3608–3612.
[104] A. T. Abraham, X. Zhou, S. M. Hecht, *J. Am. Chem. Soc.* **2001**, *123*, 5167–5175.
[105] L. F. Povirk, M. Hogan, N. Dattagupta, *Biochemistry* **1979**, *18*, 96–101.
[106] J. Chen, J. Stubbe, *Nat. Rev. Cancer* **2005**, *5*, 102–112.
[107] S. A. Kane, S. M. Hecht, *Prog. Nucleic Acid Res. Mol. Biol.* **1994**, *49*, 313–352.
[108] C. Avendaño, J. C. Menéndez, in *Medicinal Chemistry of Anticancer Drugs (Second Edition)* (Eds.: C. Avendaño, J.C. Menéndez), Elsevier, Boston, **2015**, pp. 133–195.
[109] J. C. Wu, J. W. Kozarich, J. Stubbe, *Biochemistry* **1985**, *24*, 7562–7568.
[110] L. E. Rabow, J. Stubbe, J. W. Kozarich, *J. Am. Chem. Soc.* **1990**, *112*, 3196–3203.
[111] M. Pitié, G. Pratviel, *Chem. Rev.* **2010**, *110*, 1018–1059.
[112] F. Mancin, P. Scrimin, P. Tecilla, U. Tonellato, *Chem. Commun.* **2005**, *0*, 2540–2548.
[113] D. S. Sigman, *Biochemistry* **1990**, *29*, 9097–9105.
[114] T. Hirohama, Y. Kuranuki, E. Ebina, T. Sugizaki, H. Arii, M. Chikira, P. Tamil Selvi, M. Palaniandavar, *J. Inorg. Biochem.* **2005**, *99*, 1205–1219.
[115] B. C. Bales, M. Pitié, B. Meunier, M. M. Greenberg, *J. Am. Chem. Soc.* **2002**, *124*, 9062–9063.
[116] T. Oyoshi, H. Sugiyama, *J. Am. Chem. Soc.* **2000**, *122*, 6313–6314.
[117] A. Zaid, J.-S. Sun, C.-H. Nguyen, E. Bisagni, T. Garestier, D. S. Grierson, R. Zain, *ChemBioChem* **2004**, *5*, 1550–1557.
[118] D. S. Sigman, D. R. Graham, V. D'Aurora, A. M. Stern, *J. Biol. Chem.* **1979**, *254*, 12269–12272.
[119] M. M. Meijler, O. Zelenko, D. S. Sigman, *J. Am. Chem. Soc.* **1997**, *119*, 1135–1136.
[120] D. S. Sigman, T. W. Bruice, A. Mazumder, C. L. Sutton, *Acc. Chem. Res.* **1993**, *26*, 98–104.
[121] T. Chen, M. M. Greenberg, *J. Am. Chem. Soc.* **1998**, *120*, 3815–3816.
[122] T. Gimisis, C. Chatgilialoglu, in *Encyclopedia of Radicals in Chemistry, Biology and Materials*, John Wiley & Sons, Inc., **2012**, pp. 1–26.
[123] Z. Molphy, C. Slator, C. Chatgilialoglu, A. Kellett, *Front. Chem.* **2015**, *3*, 1–9.
[124] Z. Molphy, A. Prisecaru, C. Slator, N. Barron, M. McCann, J. Colleran, D. Chandran, N. Gathergood, A. Kellett, *Inorg. Chem.* **2014**, *53*, 5392–5404.
[125] A. Kellett, M. O'Connor, M. McCann, M. McNamara, P. Lynch, G. Rosair, V. McKee, B. Creaven, M. Walsh, S. McClean, A. Foltyn, D. O'Shea, O. Howe, M. Devereux, *Dalton Trans.* **2011**, *40*, 1024–1027.
[126] B. C. Bales, T. Kodama, Y. N. Weledji, M. Pitié, B. Meunier, M. M. Greenberg, *Nucleic Acids Res.* **2005**, *33*, 5371–5379.
[127] M. Pitié, B. Meunier, *Bioconjug. Chem.* **1998**, *9*, 604–611.
[128] A. Robertazzi, A. V. Vargiu, A. Magistrato, P. Ruggerone, P. Carloni, P. De Hoog, J. Reedijk, *J. Phys. Chem. B* **2009**, *113*, 10881–10890.
[129] P. de Hoog, M. J. Louwerse, P. Gamez, M. Pitié, E. J. Baerends, B. Meunier, J. Reedijk, *Eur. J. Inorg. Chem.* **2008**, *2008*, 612–619.
[130] A. W. Schaefer, M. T. Kieber-Emmons, S. M. Adam, K. D. Karlin, E. I. Solomon, *J. Am. Chem. Soc.* **2017**, *139*, 7958–7973.
[131] G. B. Wijeratne, S. Hematian, M. A. Siegler, K. D. Karlin, *J. Am. Chem. Soc.* **2017**, *139*, 13276–13279.
[132] K. D. Karlin, J. C. Hayes, Y. Gultneh, R. W. Cruse, J. W. McKown, J. P. Hutchinson, J. Zubieta, *J. Am. Chem. Soc.* **1984**, *106*, 2121–2128.
[133] S. T. Frey, H. H. J. Sun, N. N. Murthy, K. D. Karlin, *Inorganica Chim. Acta* **1996**, *242*, 329–338.
[134] N. Zuin Fantoni, Z. Molphy, C. Slator, G. Menounou, G. Toniolo, G. Mitrikas, V. McKee, C. Chatgilialoglu, A. Kellett, *Chem. Eur. J.* **2019**, *25*, 221–237.
[135] N. Z. Fantoni, Z. Molphy, S. O'Carroll, G. Menounou, G. Mitrikas, M. Krokidis, C. Chatgilialoglu, J. Colleran, A. Banasiak, M. Clynes, S. Roche, S. Kelly, V. McKee, A. Kellett, *Chemistry – A European Journal* **n.d.**, *n/a*, DOI 10.1002/chem.202001996.
[136] T. Ito, S. Thyagarajan, K. D. Karlin, S. E. Rokita, *Chem. Commun.* **2005**, 4812–4814.
[137] K. J. Humphreys, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.* **2002**, *124*, 6009–6019.
[138] K. J. Humphreys, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.* **2002**, *124*, 8055–8066.
[139] L. Li, N. N. Murthy, J. Telser, Lev. N. Zakharov, G. P. A. Yap, A. L. Rheingold, K. D. Karlin, S. E. Rokita, *Inorg. Chem.* **2006**, *45*, 7144–7159.
[140] L. Li, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.* **2005**, *127*, 520–521.
[141] Z. Molphy, D. Montagner, S. S. Bhat, C. Slator, C. Long, A. Erxleben, A. Kellett, *Nucleic Acids Res.* **2018**, *46*, 9918–9931.
[142] L. Q. Hatcher, K. D. Karlin, *J. Biol. Inorg. Chem.* **2004**, *9*, 669–683.
[143] D. J. Galas, a Schmitz, *Nucleic Acids Res.* **1978**, *5*, 3157–3170.
[144] H. R. Drew, *J. Mol. Biol.* **1984**, *176*, 535–557.
[145] H. R. Drew, A. A. Travers, *Cell* **1984**, *37*, 491–502.
[146] K. R. Fox, Ed. , *Drug-DNA Interaction Protocols*, Humana Press, Totowa, New Jersey, **2010**.
[147] D. Suck, C. Oefner, *Nature* **1986**, *321*, 620–625.
[148] W. P. Stemmer, *Rapid Evolution of a Protein in Vitro by DNA Shuffling.*, **1994**.
[149] H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat Biotechnol* **1998**, *16*, 258–261.
[150] R. Larragy, J. Fitzgerald, A. Prisecaru, V. McKee, P. Leonard, A. Kellett, *Chem. Commun.* **2015**, *51*, 12908–12911.
[151] P. Mali, L. Cheng, *Stem Cells* **2012**, *30*, 75–81.
[152] D. M. Thurtle-Schmidt, T.-W. Lo, *Biochem. Mol. Biol. Educ.* **2018**, *46*, 195–205.
[153] C. O. Pabo, E. Peisach, R. A. Grant, *Annu. Rev. Biochem.* **2001**, *70*, 313–340.
[154] T. Gaj, C. A. Gersbach, C. F. Barbas, *Trends Biotechnol.* **2013**, *31*, 397–405.
[155] J. Smith, M. Bibikova, F. G. Whitby, A. R. Reddy, S. Chandrasegaran, D. Carroll, *Nucleic Acids Res.* **2000**, *28*, 3361–3369.
[156] M. Bibikova, D. Carroll, D. J. Segal, J. K. Trautman, J. Smith, Y.-G. Kim, S. Chandrasegaran, *Mol. Cell Biol.* **2001**, *21*, 289–297.
[157] A. J. Bogdanove, D. F. Voytas, *Science* **2011**, *333*, 1843–1846.
[158] J. Boch, H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt, U. Bonas, *Science* **2009**, *326*, 1509–1512.
[159] M. J. Moscou, A. J. Bogdanove, *Science* **2009**, *326*, 1501–1501.
[160] J. C. Miller, S. Tan, G. Qiao, K. A. Barlow, J. Wang, D. F. Xia, X. Meng, D. E. Paschon, E. Leung, S. J. Hinkley, G. P. Dulay, K. L. Hua, I. Ankoudinova, G. J. Cost, F. D. Urnov, H. S. Zhang, M. C. Holmes, L. Zhang, P. D. Gregory, E. J. Rebar, *Nat. Biotechnol.* **2011**, *29*, 143–148.
[161] J. F. Lutz, Z. Zarafshani, *Adv. Drug Deliv. Rev.* **2008**, *60*, 958–970.

[162] A. C. Mercer, T. Gaj, R. P. Fuller, C. F. Barbas, *Nucleic Acids Res.* **2012**, *40*, 11163–11172.
[163] M. Christian, T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang, A. Hummel, A. J. Bogdanove, D. F. Voytas, *Genetics* **2010**, *186*, 757–761.
[164] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816–821.
[165] R. Barrangou, P. Horvath, *Nat. Microbiol.* **2017**, *2*, 1–9.
[166] H. Wang, M. La Russa, L. S. Qi, *Annu. Rev. Biochem.* **2016**, *85*, 227–264.
[167] C. Anders, O. Niewoehner, A. Duerst, M. Jinek, *Nature* **2014**, *513*, 569–573.
[168] T. Yamano, H. Nishimasu, B. Zetsche, H. Hirano, I. M. Slaymaker, Y. Li, I. Fedorova, T. Nakane, K. S. Makarova, E. V. Koonin, R. Ishitani, F. Zhang, O. Nureki, *Cell* **2016**, *165*, 949–962.
[169] C. Fellmann, B. G. Gowen, P.-C. Lin, J. A. Doudna, J. E. Corn, *Nat. Rev. Drug. Discov.* **2017**, *16*, 89–100.
[170] L. Taemaitree, A. Shivalingam, A. H. El-Sagheer, T. Brown, *Nat. Commun.* **2019**, *10:1610*, 1–8.
[171] M. Adli, *Nat. Commun.* **2018**, *9:1911*, 1–13.
[172] B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, F. Zhang, *Cell* **2015**, *163*, 759–771.
[173] H. A. Rees, W.-H. Yeh, D. R. Liu, *Nat. Commun.* **2019**, *10:2212*, 1–12.
[174] F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, *Nat. Protoc.* **2013**, *8*, 2281–2308.
[175] H. A. Rees, D. R. Liu, *Nat. Rev. Gen.* **2018**, *19*, 770–788.
[176] W.-H. Yeh, H. Chiang, H. A. Rees, A. S. B. Edge, D. R. Liu, *Nat. Commun.* **2018**, *9:2184*, 1–10.
[177] A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, *Nature* **2016**, *533*, 420–424.
[178] A. Erxleben, *Coord. Chem. Rev.* **2018**, *360*, 92–121.
[179] C. Liu, M. Wang, T. Zhang, H. Sun, *Coord. Chem. Rev.* **2004**, *248*, 147–168.
[180] Z. Yu, J. Cowan, *Curr. Opin. Chem. Biol.* **2018**, *43*, 37–42.
[181] R. P. Hertzberg, P. B. Dervan, *J. Am. Chem. Soc.* **1982**, *104*, 313–315.
[182] P. G. Shultz, J. S. Taylor, P. B. Dervan, *J. Am. Chem. Soc.* **1982**, *104*, 6861–6863.
[183] P. B. Dervan, *Science* **1986**, *232*, 464–471.
[184] P. G. Schultz, P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6834–6837.
[185] M. Pitié, J. D. Van Horn, D. Brion, C. J. Burrows, B. Meunier, *Bioconjug. Chem.* **2000**, *11*, 892–900.
[186] A. Y. Chang, P. B. Dervan, *J. Am. Chem. Soc.* **2000**, *122*, 4856–4864.
[187] N. R. Wurtz, P. B. Dervan, *Chem. Biol.* **2000**, *7*, 153–161.
[188] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, *391*, 468–471.
[189] C. Giovannangeli, C. Hélène, *Nat. Biotechnol.* **2000**, *18*, 1245–1246.
[190] P. E. Nielsen, M. Egholm, *Curr. Issues Molec. Biol.* **1999**, *1*, 89–104.
[191] M. Boidot-Forget, M. Chassignol, M. Takasugi, N. T. Thuong, C. Hélène, *Gene* **1988**, *72*, 361–371.
[192] J. C. François, T. Saison-Behmoaras, M. Chassignol, N. T. Thuong, C. Helene, *J. Biol. Chem.* **1989**, *264*, 5891–5898.
[193] A. Majumdar, A. Khorlin, N. Dyatkina, F.-L. M. Lin, J. Powell, J. Liu, Z. Fei, Y. Khripine, K. A. Watanabe, J. George, P. M. Glazer, M. M. Seidman, *Nat. Genet.* **1998**, *20*, 212–214.
[194] I. V. Panyutin, A. N. Luu, I. G. Panyutin, R. D. Neumann, *Radiat. Res.* **2001**, *156*, 158–166.
[195] M. P. Knauert, P. M. Glazer, *Hum. Mol. Genet.* **2001**, *10*, 2243–2251.
[196] G. B. Dreyer, P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968–972.
[197] B. C. Chu, L. E. Orgel, *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 963–967.
[198] C. H. Chen, D. S. Sigman, *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 7147–7151.
[199] H. Moser, P. Dervan, *Science* **1987**, *238*, 645–650.
[200] S. A. Strobel, H. E. Moser, P. B. Dervan, *J. Am. Chem. Soc.* **1988**, *110*, 7927–7929.
[201] S. A. Strobel, P. B. Dervan, *Science* **1990**, *249*, 73–75.
[202] D. Praseuth, L. Perrouault, T. L. Doan, M. Chassignol, N. Thuong, C. Hélène, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349–1353.
[203] J. C. François, T. Saison-Behmoaras, C. Barbier, M. Chassignol, N. T. Thuong, C. Hélène, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9702–9706.
[204] P. A. Havre, E. J. Gunther, F. P. Gasparro, P. M. Glazer, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7879–7883.
[205] G. Wang, D. D. Levy, M. M. Seidman, P. M. Glazer, *Mol. Cell Biol.* **1995**, *15*, 1759–1768.
[206] N. Puri, A. Majumdar, B. Cuenoud, F. Natt, P. Martin, A. Boyd, P. S. Miller, M. M. Seidman, *J. Biol. Chem.* **2001**, *276*, 28991–28998.
[207] M. Komiyama, T. Shiiba, Y. Takahashi, N. Takeda, K. Matsumura, T. Kodama, *Supramol. Chem.* **1994**, *4*, 31–34.
[208] M. Komiyama, N. Takeda, Y. Takahashi, H. Uchida, T. Shiiba, T. Kodama, M. Yashiro, *J. Chem. Soc. Perkin Trans.* **1995**, *2*, 269–274.
[209] F. H. Zelder, A. A. Mokhir, R. Krämer, *Inorg. Chem.* **2003**, *42*, 8618–8620.
[210] R. R. Joshi, K. N. Ganesh, *Biochim. Biophys. Acta* **1994**, *1201*, 454–460.
[211] P. Bigey, G. Pratviel, B. Meunier, *Nucleic Acids Res.* **1995**, *23*, 3894–3900.
[212] D. S. Sergeyev, V. F. Zarytova, *Russ. Chem. Rev.* **1996**, *65*, 355–378.
[213] M. Komiyama, S. Kina, K. Matsumura, J. Sumaoka, S. Tobey, V. M. Lynch, E. Anslyn, *J. Am. Chem. Soc.* **2002**, *124*, 13731–13736.
[214] W. Chen, T. Igawa, J. Sumaoka, M. Komiyama, *Chem. Lett.* **2004**, *33*, 300–301.
[215] Y. Aiba, J. Sumaoka, M. Komiyama, *Chem. Soc. Rev.* **2011**, *40*, 5657–5668.
[216] R. Zain, C. Marchand, J. Sun, C. H. Nguyen, E. Bisagni, T. Garestier, C. Hélène, *Chem. Biol.* **1999**, *6*, 771–777.
[217] C. Marchand, C. H. Nguyen, B. Ward, J.-S. Sun, E. Bisagni, T. Garestier, C. Hélène, *Chem. Eur. J.* **2000**, *6*, 1559–1563.