RESEARCH FRONT

CSIRO PUBLISHING

Review

www.publish.csiro.au/journals/ajc

Aust. J. Chem. **2011**, *64*, 671–680

Targeting Nucleic Acids using Dynamic Combinatorial Chemistry

Chandramathi R. Sherman Durai[A] *and Margaret M. Harding*[A,B]

[A]School of Chemistry, The University of New South Wales, NSW 2052, Australia.
[B]Corresponding author. Email: harding@unsw.edu.au

Dynamic combinatorial chemistry (DCC) is a powerful method for the identification of novel ligands for the molecular recognition of receptor molecules. The method relies on self-assembly processes to generate libraries of compounds under reversible conditions, allowing a receptor molecule to select the optimal binding ligand from the mixture. However, while DCC is now an established field of chemistry, there are limited examples of the application of DCC to nucleic acids. The requirement to conduct experiments under physiologically relevant conditions, and avoid reaction with, or denaturation of, the target nucleic acid secondary structure, limits the choice of the reversible chemistry, and presents restrictions on the building block design. This review will summarize recent examples of applications of DCC to the recognition of nucleic acids. Studies with duplex DNA, quadruplex DNA, and RNA have utilized mainly thiol disulfide libraries, although applications of imine libraries, in combination with metal coordination, have been reported. The use of thiol disulfide libraries produces lead compounds with limited biostability, and hence design of stable analogues or mimics is required for many applications.

Manuscript received: 14 January 2011.
Manuscript accepted: 2 March 2011.

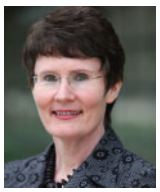
Introduction

The molecular recognition of DNA and RNA is fundamental to many biochemical processes related to transcription, regulation, and gene expression.[1–3] The flexibility of these biomolecules leads to a diversity of secondary structures, whose formation and stability is highly dependent on the base sequence and environment.[4,5] The study of small molecules that recognize and bind with high affinity and selectivity to duplex, triplex, and quadruplex DNA, as well as hairpins, bulges, and RNA loops has attracted significant interest because of the involvement of many of these structures in disease.[5–11]

Medicinal and synthetic chemists have incorporated the features present in natural products to design new ligands in order to elucidate the rules that govern nucleic acid recognition.[5,12,13] While significant progress has been made in the design of synthetic molecules that bind to the minor groove of DNA,[12] the design of sequence-specific compounds that bind to the major or minor grooves remains challenging. Similarly, while good progress has been made in the design of synthetic ligands that stabilize triplex and quadruplex DNA,[9,14–16] an improved understanding of how to target the unique molecular recognition features of different DNA secondary structures is still required. The recognition of RNA is much less advanced than DNA, with the unique folded structure requiring the ability to recognize structural parameters that arise from irregularities such as bulges, loops, and mismatches.[6]

Dynamic combinatorial chemistry (DCC) is a new approach to understanding molecular recognition. In DCC, building blocks that incorporate functional groups that are able to undergo reversible reactions, are equilibrated to generate a

Chandramathi R. Sherman Durai graduated with M.Sc. (2002) and M.Phil. (2005) degrees from the Manonmaniam Sundaranar University, Tirunelveli, India. Currently she is pursuing her Ph.D. under the supervision of Professor Margaret M. Harding at the University of New South Wales, Sydney, Australia. Chandramathi's research is focussed on the design and synthesis of DNA-binding compounds using dynamic combinatorial chemistry.



Margaret M. Harding holds B.Sc. (Honours) (1982), Ph.D. (1987), and D.Sc. (2002) degrees from the University of Sydney and is currently Pro Vice-Chancellor (Research) at the University of New South Wales. She held postdoctoral positions with Professor Jean-Marie Lehn at the Université Louis Pasteur, Strasbourg (1986–1988) and Professor Dudley Williams at the University of Cambridge (1988–1989) followed by an academic appointment at the University of Sydney (1990–2005). In 2005 she was appointed as Professor of Chemistry and the inaugural Dean of Graduate Research at the University of New South Wales. Current research interests are on antifreeze proteins and glycoproteins, DNA recognition, and new synthetic DNA-binding molecules.

10.1071/CH11023 0004-9425/11/060671

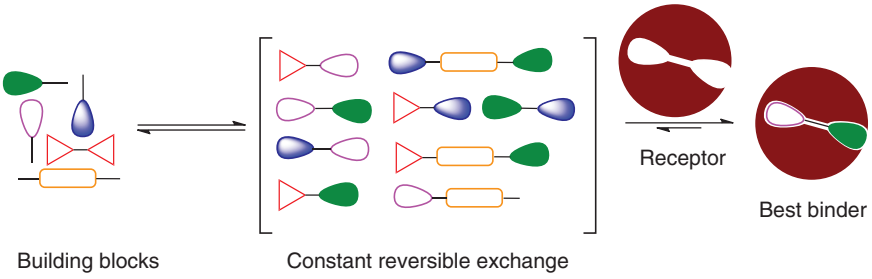


Fig. 1. Schematic of dynamic combinatorial chemistry showing generation of a dynamic combinatorial library and selection of best binder by target molecule.

dynamic combinatorial library (DCL) that comprises all possible combinations of the building blocks (Fig. 1).[17–19] As a result of this reversibility, all library members interconvert to give a distribution that is under thermodynamic control. Addition of a target molecule that selectively binds to one receptor in the library serves to increase the concentration of that receptor at the expense of others in the DCL, and is followed by isolation and identification of the amplified receptor. The scope and applications of DCC are the subject of several excellent reviews.[17,19–24]

This review will focus on the potential of DCC to identify new synthetic compounds for the recognition of nucleic acids. While an enormous number of studies have demonstrated the great potential of DCC to identify unexpected and novel receptor molecules, there have been limited studies with nucleic acids. The scope and limitations of DCC are discussed, as well as opportunities for identification of new lead compounds for improved understanding of the molecular recognition of nucleic acids.

DCC Reaction Conditions and Analysis

The choice of the functional groups that result in the reversible reactions required to generate a DCL and the need for aqueous solubility, present several challenges for the design of the building blocks for applications with nucleic acids. The functional group(s) need to undergo chemistry that is compatible with the experimental conditions required to maintain the secondary structure of the target nucleic acid structure, including the conditions required to initiate and quench the reversible reaction, and allow isolation of the selected product. The building blocks as well as the library members need to be soluble in aqueous solution in the approximate pH range 5–8, which can be difficult to achieve with aromatic chromophores that are typically present in intercalators and threaders. The reversible chemistry must also be compatible with the functional groups present in nucleic acids, notably the exocyclic amino groups, the phosphodiester backbone, and the C-4 and C-6 positions of the pyrimidines.[4]

Table 1 summarizes the reported applications of DCC with nucleic acid targets. Studies have been performed to identify small molecules with affinity for duplex DNA, quadruplex DNA, and RNA. In addition, several examples of DCC applied to identify modified oligonucleotides that are stabilized by appended groups have been reported. As shown in Table 1, almost all examples have utilized the reversible thiol-disulfide oxidation reaction, which can generally be conducted in aqueous media, in the physiological pH range, initiated and modulated by oxidized and reduced glutathione (GSSG/GSH). Under these conditions, there are no competitive reactions of the building blocks with the phosphate backbone or nucleophilic or electrophilic sites present in the nucleic bases. An interesting advantage of utilizing GSSG/GSH to facilitate the redox reaction is that the reagents also participate in the DCC reactions and thus add additional diversity to the constituents of the library.

Two methods have been commonly used to detect and analyze the amplified products in DCC studies with nucleic acids (Fig. 2). In the most common method (Fig. 2a), the biotinylated oligonucleotide target molecule is immobilized onto streptavidin-functionalized magnetic beads, by the strong biotin–streptavidin interaction, allowing ready separation of the DNA from the DCL members.[25,26] Upon denaturing the DNA, and removing the beads, the compounds bound to the DNA are identified. In the second method, termed resin-bound DCC (RB-DCC) (Fig. 2b), a library of resin-bound monomers are combined with an identical library of monomers in solution to generate a resin-bound DCL.[27] Addition of a fluorescently tagged target oligonucleotide, followed by analysis of the beads by fluorescence microscopy provides ready identification of the selected library members. An important feature of RB-DCC is that the solution phase and resin-bound components are in competition for target binding, which needs to be taken into account in the analysis of the beads, as diminished fluorescence may be a result of a strong solution binder being washed out of solution. Equilibrium dialysis has been used in DCC studies involving metal ions with both DNA and RNA.[28]

Studies with Nucleic Acids

Duplex DNA

Miller and colleagues were the first group to report the selection and amplification of novel DNA-binding compounds using DCC.[29,30] The DCL of 36 metal complexes was generated from a library of salicylaldimine ligands and zinc(II), and was performed in 10×10^{-3} M Tris·HCl, 100×10^{-3} M KCl, 1% dimethyl sulfoxide (DMSO) at pH 7.5. The DCL was screened against poly d(AT) immobilized on a cellulose resin. While this study was important in establishing the principle of identification of novel DNA-binding compounds, the formation of unstable imines in water, partial hydrolysis of the metal complexes, and participation of the buffer in the reactions complicated the analysis of the results.

Balasubramanian and colleagues designed several thiol-functionalized polyamides **1**, **2**, and **3** based on distamycin (Fig. 3) with the goal of identifying the optimum number of heteroaromatic units for binding to an AT-rich duplex DNA sequence identified in the promoter region of oncogene c-kit, a possible site for intervention of transcriptional regulation.[25] Distamycin-like polyamides, which contain *N*-methylpyrrole and *N*-methylimidazole functional groups, can bind to the minor groove of double-stranded DNA with an affinity similar to

Table 1. Summary of dynamic combinatorial chemistry (DCC) reactions and conditions used to study recognition of nucleic acid targets

Nucleic acid	Sequence		DCC reactions and conditions			Ref.
	Target	Control	Start. materials	Buffer	Initiation	
Duplex DNA	Poly d(AT)	Poly d(AT) in the absence of Zn^{2+}	Salicyl-aldimines Zn^{2+}	Tris·HCl, KCl, 1% DMSO (pH 7.5)	Incubate 25°C	[29,30]
	d(CTTTTATTTTG)·(GAAAATAAAAC)[A]	Random duplex; sequence not given	RSH	Tris·HCl, KCl (pH 7.4)	GSH GSSG	[25]
	d(CGCGAAATTTCGCG)·(GCGCTTTAAAGCGC) & d(CGTACGGCCGTACG)·(GCATGCCGGCATGC)	–	RSH R_1SSR_2	Tris·HCl, NaCl, $MgCl_2$ (pH 7.4)	GSH GSSG	[33]
	d(CCATGATATC)·(GGTACTATAG)[B]	d(TCTAGACGTC)·(AGATCTGCAG)	RS-resin	Phosphate (pH 7.4)	RSH	[27]
Quadruplex DNA	$d(GTTAGG)_5$ [A]	$d(AGTTAG)_5$ [A]	RSH	Tris·HCl, KCl (pH 7.4)	GSH GSSG	[26]
	$d(GTTAGG)_5$ [A]	Random duplex; sequence not given	RSH	Tris·HCl, KCl (pH 7.4)	GSH GSSG	[25]
	d(CGGGCGGGCGCGAGGGAGGGG) & d(TGAGGGTGGGTAGGGTGGGTAA)[A]	d(GGCATAGTGCGTGGGCGTTAGC)·(CCGTATCACGCACCCGCAATCG)[A]	RSH	Tris·HCl, KCl (pH 7.4)	GSH GSSG	[38]
	$d(GTTAGG)_5$ [A]	$d(AGTTAG)_5$ [A]	R_1SSR_2	Tris·HCl KCl (pH 7.4)	Cystamine hydrochloride	[40]
RNA	UAGUCUUUCGAGACUA	TAGTCTTTCGAGACTA	Salicyl-amides, Cu^{2+}	H15-Mg[C] (pH 7.5)	Equilibrium dialysis	[28]
	UUUUUUAGGGAAGAUCUGGCC–UUCCCACAAGGGAAGGCCAGGGAAU[D]	RNA in the presence of R-S-resin only	RS-resin (Cys S-^{t}Bu protected)	Phosphate (pH 7.4)	RSH	[35]
	$CCG(CUG)_{10}CGG$[D]	$GGG(CUG)_{109}GGG$[E]	RS-resin (Cys S-^{t}Bu protected)	Phosphate $MgCl_2$ (pH 7.2)	RSH	[42]
Oligo conjugates	(ACGCGU)·(ACGCGU)[F]	TTTCGU[F]	RCHO, oligo-NH_2	Phosphate (pH 8.0)	Equilibrium dialysis	[43]
	CACUGGGAUC	GAUCCCAGUG[G]	RCHO, oligo-NH_2	Phosphate, NaCl, KCl, $MgCl_2$ (pH 6.0)	Equilibrium dialysis	[44]
	CCAGAUUUGAGCCUGG–GAGCUCUCUGG	GGGAGGACGAAGCGGACGC–AGAAGACACGCCCGA	RCHO, RNH_2	Phosphate, NaCl, KCl, $MgCl_2$ (pH 6.0)	Incubate	[45]

[A] 5′-Biotinylated oligonucleotide.
[B] 5′-TAMRA (tetramethyl-6-carboxyrhodamine) labelled with a six-carbon spacer separating DNA and fluorophore.
[C] Na-HEPES (25×10^{-3} M), HEPES (25×10^{-3} M), KCl (135×10^{-3} M), and $MgCl_2$ (15×10^{-3} M) at pH 7.5.
[D] Labelled at 5′-end with Cy-3 water-soluble cyanine dye.
[E] Labelled with fluorescein amidite dye.
[F] Oligonucleotide strands contain 2′-deoxy-2′-aminouridines at 3′ ends.
[G] Oligonucleotide strands contain 2′-deoxy-2′-aminouridines at U3 and U9.

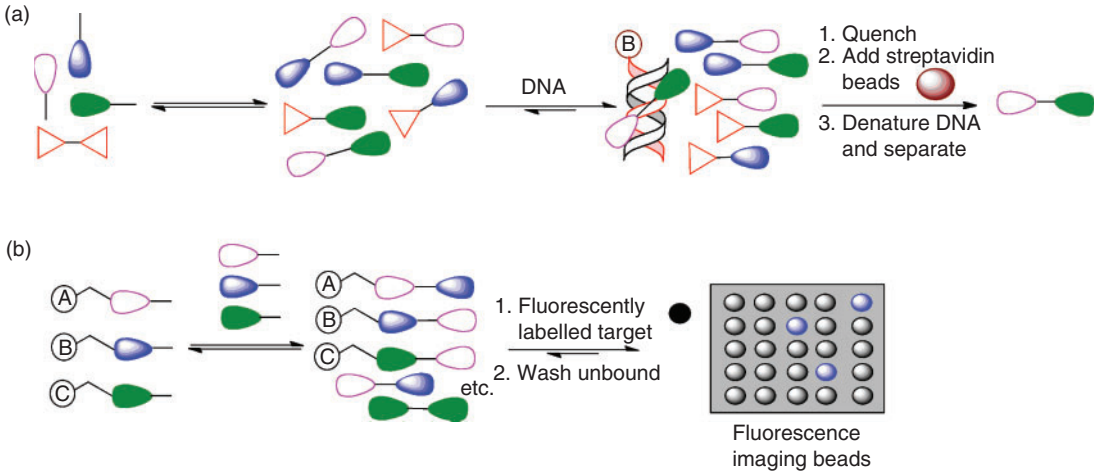


Fig. 2. Schematic of formation of a dynamic combinatorial library formed from thiol and disulfide building blocks and detection of amplified products in the presence of a nucleic acid target using (a) biotinylated probes and streptavidin beads;[25,26] and (b) resin-bound dynamic combinatorial chemistry.[27]

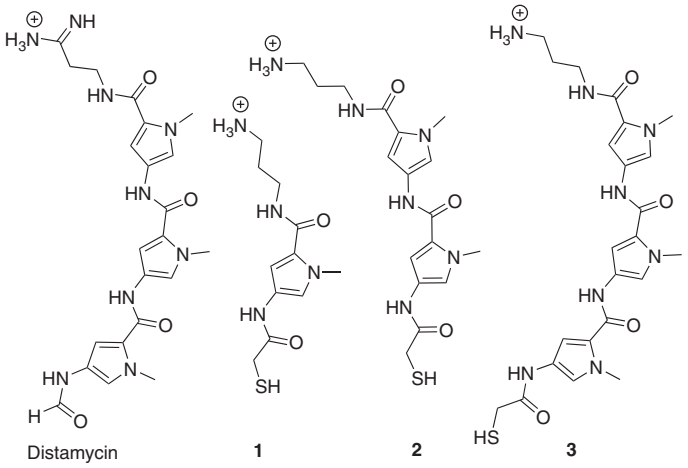


Fig. 3. Structures of thiol-functionalized polyamides designed to mimic distamycin.[25]

naturally occurring DNA-binding proteins in a sequence-specific manner.[12,31,32] In the presence of an 11-mer duplex, the heterodisulfide formed from oxidation of **2** and **3**, and the homodisulfide of **3** were amplified.

Thiol derivatives of Hoechst33258, a known DNA minor groove binder with high AT selectivity were studied using an equilibrium shift assay.[33] Thiol, carboxylated, and alkylated derivatives of Hoechst33258 as well as the glutathione derivatives were studied and the DCL mixtures were analyzed in the presence of DNA sequences containing an A_3T_3 binding motif. The ligands connecting Bis-Hoechst33258 units showed selective binding towards the DNA sequence with two A_3T_3 binding sites as well as three-way junction DNA.

McNaughton and Miller illustrated the concept of RB-DCC by application to the detection of DNA-bisintercalators based on the naturally occurring quinoxaline depsipeptides echinomycin and triostin A, which is stabilized by an intramolecular disulfide bond (Fig. 4a).[27] Several studies of synthetic variants have shown that the composition of the depsipeptide linker has a profound effect on the sequence selectivity of the bisintercalators.[34] A family of nine cysteine (Cys)-containing quinoline tripeptides (Fig. 4b) was oxidized to form a library of 45 unique disulfide dimers. The library was screened for binding against a DNA sequence reported to be preferably bound by triostin A (5′-TCTAGACGTC-3′) and a sequence reported to be preferentially bound by a synthetic analogue (5′-CCATGATATC-3′). Both oligonucleotide sequences were labelled at the 5′-end with the fluorophore tetramethyl-6-carboxyrhodamin (TAMRA). The serine bisquinoline disulfide **4** (Fig. 4c) was identified as a lead compound, and further independent DNA-binding experiments confirmed the high binding affinity of **4** for DNA. While this report demonstrated the concept of RB-DCC with a relatively small library of components, the authors highlighted the potential of RB-DCC for significantly larger libraries by the use of microarrays and beads as mixtures using an encoding system.[27] This approach has been used to generate a library of around 11000 members for detection of RNA-selective small molecules[35] (see later section on RNA).

Quadruplex DNA

The identification of small molecules that bind selectively and with strong affinity to quadruplex DNA has attracted significant interest in recent years.[8,9,36,37] G-quadruplex formation has been linked to telomer formation and has implications in cancer biology and ageing, and the stabilization of quadruplexes has the potential to control gene expression.[8,9,11]

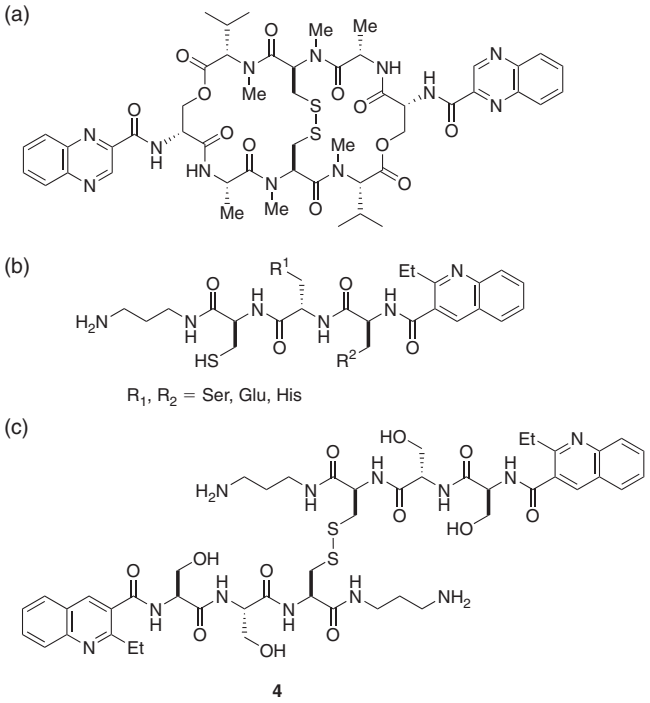


Fig. 4. (a) Structure of the natural bisintercalator triostin A. (b) Peptide mimics of triostin A used for generation of a dynamic combinatorial library of solution and resin-bound disulfides. (c) Highest affinity ligand **4** for $d(CT_4AT_4G)\cdot(GA_4TA_4C)$ identified using resin-bound dynamic combinatorial chemistry.[27]

The potential of combining several distinct recognition units to recognize a human telomeric quadruplex formed from $d(GTTACG)_5$ was investigated using the thiols shown in Fig. 5.[26] The hydrophobic acridone unit **5** was designed to intercalate in the terminal G tetrad of the quadruplex, and the thiol derivative of the tetrapeptide FRHR **6** has been reported to have quadruplex recognition properties. In the presence of GSH, a library of nine components was generated, including adducts of **5** and **6** with GSH. The disulfide formed from the oxidation of **5** with **6**, as well as the hexapeptide dimer of **6**, showed greater binding affinity with G-quadruplex DNA than other species present in the library. The amplification of the hexapeptide is a particularly interesting result, as there are no other reports of short peptides that bind quadruplex DNA with high affinity.

The DNA-binding properties of the DCL generated from thiol functionalized polyamides based on distamycin (see Fig. 3) was also contrasted with the quadruplex DNA.[25] These experiments showed that the homodimers formed from **2** and **3** have a much greater affinity for duplex DNA than for quadruplex DNA.

DCC has provided important insight into the requirements for specificity in G-quadruplex recognition by oxazole-peptide macrocycles.[38] Based on previous results that indicated that both the number and length of simple alkylamine side chains appended to the macrocycle are strong determinants of quadruplex affinity,[39] the library building blocks included derivatives of *p*-benzylic thiols and neutral carbohydrate derivatives with different potential for H-bonding and electrostatic interactions (Fig. 6). The DCL was screened against two intramolecular quadruplex forming sequences (c-Kit21, c-Myc22) and a 22-mer duplex DNA for comparison. In the presence of c-Kit21 and c-Myc22, and the charged building blocks (Fig. 6b), the two

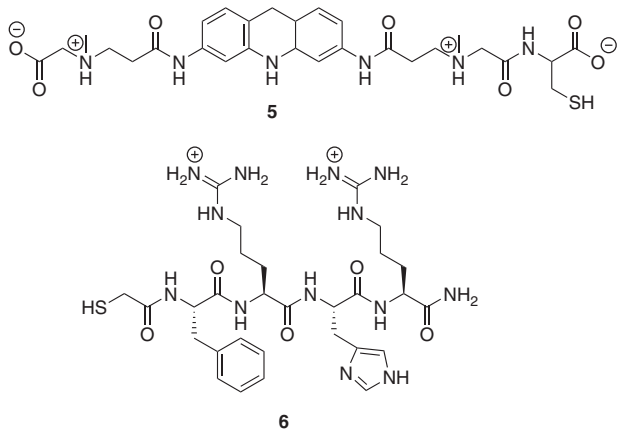


Fig. 5. Acridone **5** and peptide thiol building block **6** used to generate a dynamic combinatorial chemistry in the presence of GSH and quadruplex DNA.[26]

guanidinium disufides formed from **7** and **8**, and **7** and **9**, were amplified, demonstrating that the geometry and/or the hydrogen bonding potential of the side chains play a role, and it is not just overall charge that determines affinity. Of interest was the fact that under the same conditions in the presence of DNA there was no evidence for any binding. In the case of the carbohydrate building blocks (Fig. 6c), there was also no perturbation of the DCL composition in the presence of DNA. However, the disulfide formed from **7** and **10** was most strongly amplified in the presence of c-Kit21. This disulfide as well as the disulfide formed from **7** and **11** bound to c-Myc22 with similar K_d values.

In contrast to the examples above, Nielsen and Ulven employed disulfide scrambling to generate a DCL by the use of a central scaffold designed to carry and equilibrate with several side chains (Fig. 7).[40] Three aromatic scaffolds, each carrying positively charged side-chains, were designed based on the structures of the majority of G-quadruplex ligands, which generally contain a central aromatic scaffold with one or several positively charged side chains. Cystamine hydrochloride was used to initiate the reaction, and the DCL was generated from the scaffolds and a large excess of the positively charged side chains in order to minimize oligomerization of the scaffolds. The reaction was conveniently monitored by the UV of the chromophores, which was outside the absorbance range of the side chains. From the equilibrium distribution of 18 species, the acridine **13** was extracted with high selectivity with a lower amount of the acridine **12** also extracted.

RNA

The rules that govern the design and synthesis of small molecules that bind with high sequence specificity and affinity to RNA are much less advanced than for DNA. The majority of RNA-binding molecules are natural products based on derivatives of amino glycosides,[7] but some synthetic peptide threaders have been successfully designed and synthesized to interact with RNA by a threading mechanism.[41]

The first example of DCC applied to RNA involved the use of metal ion coordination as the reversible chemistry to generate the DCL.[28] Metal ion coordination with ligands provides significant diversity in a DCL as a range of different stoichiometries and geometries of the metal complexes can result depending on the coordination preference of the metal. However, the lability of metal–ligand coordination required to

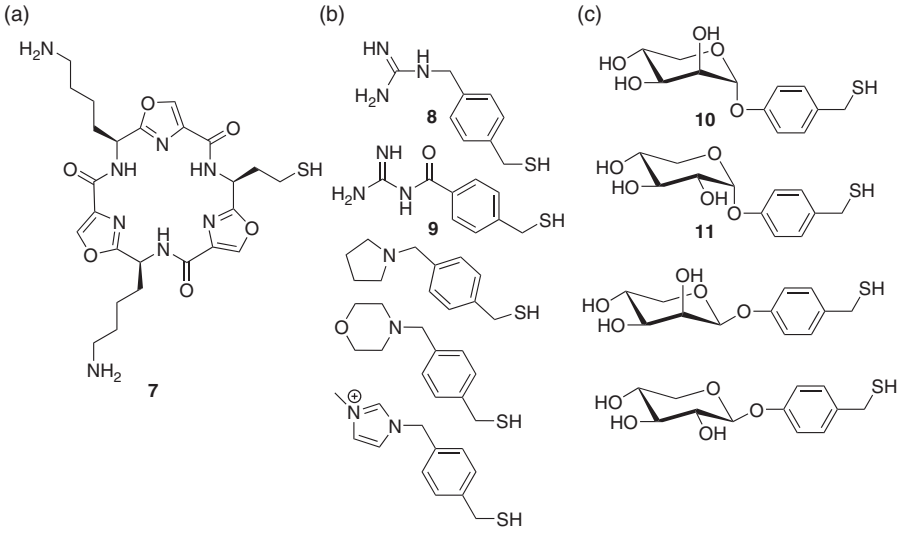


Fig. 6. Structures of (a) oxazole-based peptide macrocycle **7**, (b) cationic benzylic thiols, and (c) neutral carboydrate benzylic thiols used to generate a dynamic combinatorial library for recognition of quadruplex DNA targets.[38]

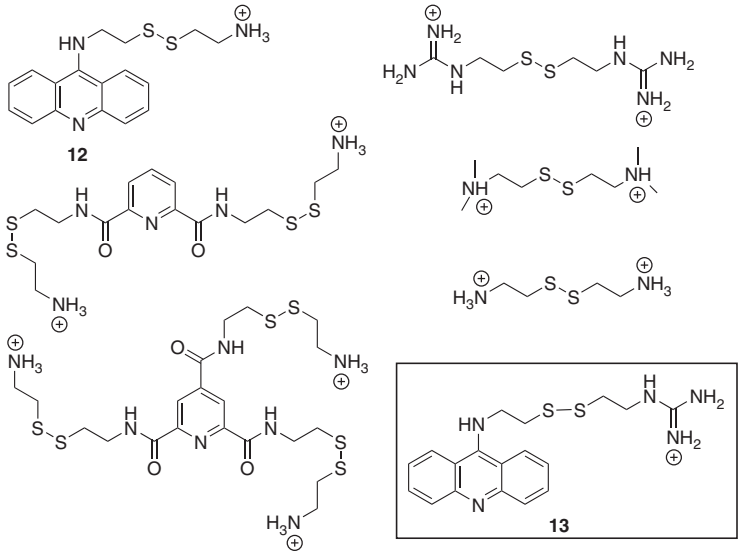


Fig. 7. Structures of aromatic scaffolds, and side chains used to generate a dynamic combinatorial library in the presence of the scrambling initiator, cystamine hydrochloride, for selective extraction of G-quadruplex DNA; acridine **13** identified as strongest binder to $d(AGT_2AG)_5$.[40]

generate the DCL presents challenges for isolation of the stable, selected metal complex. Karan and Miller used derivatives of salicylamide ligands for the construction of a RNA binding library in the presence of metal ions. The ligands included a variable position for incorporation of potential RNA-binding moieties.[28] It was assumed that square-planar mono and bis (salicylamide) complexes would form in the presence of Cu^{2+}, with other coordination modes possible, giving at equilibrium, a minimum of 27 metal complexes. The library of metal complexes was equilibrated in the presence of an RNA hairpin derived from the GTP-binding P7 helix from the *Pneumocystis carinii* Group I intron, as well as the homologous DNA sequence for comparison. Using equilibrium dialysis and a 3500 molecular weight cut-off membrane, in the absence of Cu^{2+} none of the ligands bound to either RNA or DNA. However, in the presence of Cu^{2+}, rather surprisingly, a neutral peptide ligand was selected in preference to the positively charged peptide ligands. UV titrations confirmed the strong affinity of the selected metal complex for the RNA hairpin, which exhibited greater than 300-fold selectivity over the DNA sequence. It should be noted however, that while binding to DNA was detected in the DCC experiment, there was no evidence for binding by UV, possibly as a result of non-specific binding or because the binding mode of the metal complex with DNA could not be detected by UV. While this communication demonstrated the potential of DCC to identify RNA binding molecules, the use of paramagnetic metal complexes, and the inability to fully characterize the overall shape and stoichiometry of the bound metal complex, highlights some of the difficulties in using metal ion coordination in DCC with nucleic acids.

In a series of papers, Miller and colleagues have demonstrated the very high potential of RB-DCC to identify novel RNA binding molecules (Fig. 8).[10,35,42] The 22 nucleotide hairpin sequence in the HIV-1 frameshift stem-loop RNA was targeted

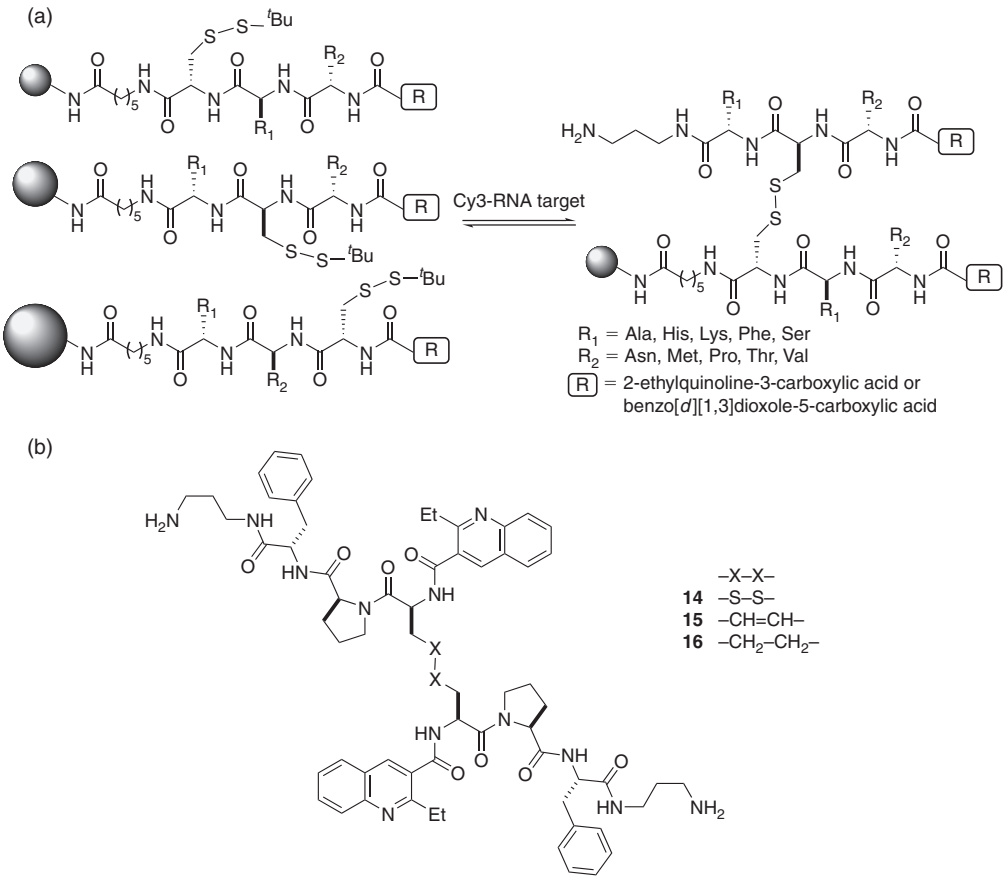


Fig. 8. (a) Design of the 11325-member resin-bound dynamic combinatorial chemistry using three different bead sizes to encode the Cys position for studies with RNA,[42] and (b) the lead compound **14** detected using this assay for HIV-1 frameshift stem loop RNA,[35] and hydrocarbon isosteres **15** and **16**.[10]

by screening against a DCL with a theoretical size of 11325 members.[35] The library was constructed from 150-resin attached Cys-containing building blocks and an identical set of solution-phase building blocks (Fig. 8a). Dimers formed from three of the building blocks were identified as possible RNA-binding compounds, and following synthesis of the pure dimers and measurement of affinity constants, the sequence-selective, high affinity dimer **14** formed from a single Cys-building block was identified (Fig. 8b). In an independent study, the same library was screened for target compounds able to inhibit muscular dystrophy type 1 (MBNL1) binding to (CUG) repeat RNA.[42] Four lead compounds were identified and inhibited the interaction of $GGG(CUG)_{109}GGG$ RNA with MBNL1 in vitro with K_i values in the low micromolar range.

In a very recent study, the olefin and hydrocarbon isosteres of **14**, compounds **15** and **16** respectively (Fig. 8b), were synthesized and tested for activity, as a first step to improving the biostability of the disulfide **14**.[10] The olefin analogue **15** had similar activity to the parent disulfide, and thus provides the basis for production of compounds that may be suitable for cellular assays of frameshifting.

Stabilized Oligonucleotide Conjugates

Rayner and colleagues have applied imine libraries, formed from an oligonucleotide functionalized with an amino group with a set of aldehydes, to the identification of covalently appended residues that stabilize oligonucleotides (Fig. 9).[43] The library design involved incorporation of the reactive amine group as the 2′-amino-2′-deoxynucleotide **17**. The higher nucleophilicity of the amino group at this position (pK_a 6.2) ensured selective reaction at this position only and no reaction at the amino groups in the nucleic bases. The dynamic mixture of conjugated duplexes was generated in aqueous solution, the reaction was quenched with $NaBH_3CN$ to give the stable amines, and the products were analyzed by HPLC. A significant enrichment of the nalidixic conjugated product **18** was obtained compared with the other possible products. These initial results with the DNA model system were applied to a tertiary structured RNA complex formed by a loop–loop interaction between a RNA hairpin aptamer and its target the transactivation-responsive (TAR) RNA hairpin element of HIV-1.[43] A 14-nucleotide version of the aptamer and a 27-nucleotide form of TAR (miniTAR) were used for the study. The nalidixic derivative **18** was amplified by 20% in the presence of miniTAR.

A later study extended this result to the incorporation of the appended stabilized ligand within the oligonucleotide ligand,[44] and allowed entry into DCC combined with systematic evolution of ligands by exponential enrichment (SELEX) for the in vitro selection of modified aptamers; i.e., structured RNA or RNA oligonucleotides that display specific target-binding or catalytic properties.[44,45] Using this approach, conjugated RNA aptamers that bind tightly to the TAR element of HIV-1 were identified.

Very recently, DCC has been applied to the selection of triplex-forming oligonucleotides (TFO).[46] TFO designed to bind to a DNA target and stabilize triple-helix formation were selected from a DCL of amines and polyamines.

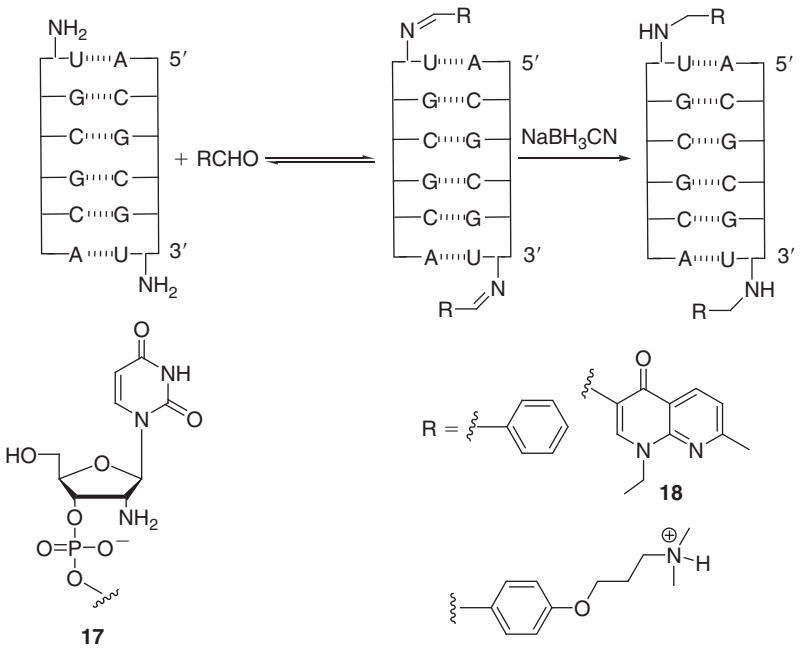


Fig. 9. Selection of 3′-appended residues that stabilize a DNA duplex using dynamic combinatorial chemistry.[43]

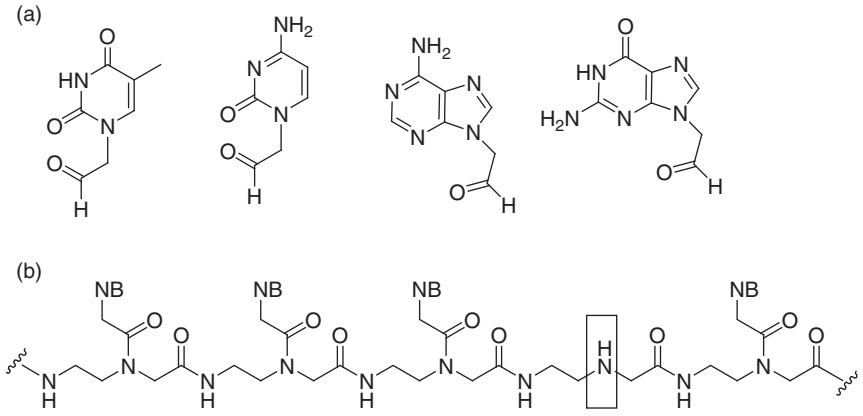


Fig. 10. (a) Aldehyde modified nucleobases and (b) the general structure of a modified blank peptide nucleic acid (PNA) strand (vacant site with free amine indicated in box) used to generate a DNA/iminium PNA dynamic combinatorial library, which is reduced to a DNA/PNA duplex.[47]

DNA Analysis

Very recently, DCC has been applied to single-nucleotide polymorphism (SNP) in DNA analysis.[47] While there are several methods available for SNP analysis, DCC was an attractive method for non-enzymatic genotyping of genomic DNA. The imine chemistry and approach is similar to that used for the identification of stabilized oligonucleotide conjugates shown in Fig. 9, but relied on the novel use of the amine group in a peptide nucleic acid (PNA) backbone (Fig. 10b). PNA is a peptide mimic of DNA in which the sugar-phosphate backbone of DNA is replaced by repeating *N*-(2-aminoethyl)glycine units linked by peptide bonds. The amino groups in PNA were treated with the aldehyde functionalized nucleic bases (Fig. 10a), followed by reduction and analysis, for SNP analysis.

The synthesis of components for the generation of constitutional dynamic analogues of nucleic acids also deserves mention.[48] A range of bisfunctionalized nucleotide analogues containing hetero- and homotopic monomer units were synthesized for the generation of both main chain and side chain reversible polymeric mimics of nucleic acids.

Summary and Outlook

While DCC is now an established field of chemistry, the application to nucleic acids presents unique challenges. The design of building blocks with structural and functional features that may lead to the recognition of nucleic acid secondary structures requires chemistry that is compatible with biologically relevant conditions and methods for the isolation, detection, and characterization of the amplified molecules and binding mode(s) with the nucleic acid target molecule.

The highly sequence dependent conformational flexibility of nucleic acids, long-range effects, the involvement of bound water molecules, and the interplay between hydrophobicity and electrostatic potentials along the helix, are all important in molecular recognition. One of the major difficulties synthetic and medicinal chemists have faced in targeting nucleic acids

through traditional structure-based design is the conformationally flexible target of DNA or RNA. It is still not possible to accurately predict the shape and properties of a given DNA or RNA sequence, and in addition, significant conformational changes can often accompany binding. In this environment, DCC has the potential to offer enormous insights compared with other approaches, as all possible conformations of the nucleic target molecule can be accessed in the DCL and screened for the presence of binding ligands.

While a large number of reversible reactions are suitable for DCC,[17,19,20] in many cases these reactions cannot be performed under aqueous conditions, or involve reaction conditions and reagents that react with oligonucleotides, or denature nucleic acids. As shown in Table 1, almost all examples to date have utilized reversible thiol disulfide reactions for applications with nucleic acids. However, the major limitations of these studies are the instability of the selected products in biological media that includes significant concentrations of the biological redox system GSH/GSSG as well as other thiols.[49,50] Thus, the application of these compounds in medicine and biology requires the development of stable analogues of these lead compounds. The very recent report of stable olefin and hydrocarbon isosteres of the lead disulfide **14**, that showed comparable binding affinities and biological activities, validates this approach.[10] It should also be noted that the solution DCC studies with nucleic acids have used small numbers of building blocks. The size of the library has been shown to be important in detecting binders and larger DCL have been shown to be likely to produce better binders.[51]

Recent studies have identified new classes of non-classical heterocyclic DNA groove-binders that do not fit the established paradigm for minor groove binding, suggesting that new design features and modes of binding to DNA are yet to be discovered.[52–54] Similarly, DCC has significant potential to improve our understanding of factors that dictate groove location and specificity in linker design in DNA-bisintercalators and threaders.[55–57] The development of new reversible reactions that can be initiated and quenched under conditions compatible with nucleic acids, would expand the opportunities for nucleic acid recognition using DCC. Our recent studies have focussed on reversible aqueous metathesis reactions for potential applications in DCC as a step towards this goal.[58]

Acknowledgements

Financial support from the Australian Research Council Discovery Grant Programme (DP0984695) and the University of New South Wales is gratefully acknowledged.

References

[1] P. L. Privalov, A. I. Dragan, C. Crane-Robinson, K. J. Breslauer, D. P. Remeta, C. A. S. A. Minetti, *J. Mol. Biol.* **2007**, *365*, 1. doi:10.1016/J.JMB.2006.09.059

[2] D. Jantz, B. T. Amann, G. J. Gatto, Jr, J. M. Berg, *Chem. Rev.* **2004**, *104*, 789. doi:10.1021/CR020603O

[3] M. Tateno, K. Yamasaki, N. Amano, J. Kakinuma, H. Koike, M. D. Allen, M. Suzuki, *Biopolymers* **1997**, *44*, 335. doi:10.1002/(SICI)1097-0282(1997)44:4<335::AID-BIP3>3.0.CO;2-R

[4] G. M. Blackburn, M. J. Gait, *Nucleic Acids in Chemistry and Biology* **2006** (Royal Society of Chemistry: Cambridge).

[5] G. Song, J. Ren, *Chem. Commun.* **2010**, *46*, 7283. doi:10.1039/C0CC01312A

[6] S. Fulle, H. Gohlke, *J. Mol. Recognit.* **2010**, *23*, 220.

[7] X.-S. Ye, L.-H. Zhang, *Curr. Med. Chem.* **2002**, *9*, 929. doi:10.2174/0929867024606740

[8] Y. L. Jiang, Z. P. Liu, *Mini Rev. Med. Chem.* **2010**, *10*, 726. doi:10.2174/138955710791572433

[9] M. Franceschin, *Eur. J. Org. Chem.* **2009**, 2225. doi:10.1002/EJOC.200801196

[10] P. B. Palde, L. O. Ofori, P. C. Gareiss, J. Lerea, B. L. Miller, *J. Med. Chem.* **2010**, *53*, 6018. doi:10.1021/JM100231T

[11] Y. Wu, R. M. Brosh, Jr, *FEBS J.* **2010**, *277*, 3470. doi:10.1111/J.1742-4658.2010.07760.X

[12] R. M. Doss, M. M. Marques, S. Foister, D. M. Chenoweth, P. B. Dervan, *J. Am. Chem. Soc.* **2006**, *128*, 9074. doi:10.1021/JA0621795

[13] E. J. Fechter, B. Olenyuk, P. B. Dervan, *Angew. Chem. Int. Ed.* **2004**, *43*, 3591. doi:10.1002/ANIE.200454231

[14] I. Haq, J. Ladbury, *J. Mol. Recognit.* **2000**, *13*, 188. doi:10.1002/1099-1352(200007/08)13:4<188::AID-JMR503>3.0.CO;2-1

[15] K. R. Fox, T. Brown, *Q. Rev. Biophys.* **2005**, *38*, 311. doi:10.1017/S0033583506004197

[16] T. A. Winters, E. Pastwa, K. Datta, R. D. Neumann, *Curr. Genomics* **2003**, *4*, 275. doi:10.2174/1389202033490420

[17] J. M. Lehn, *Chemistry* **1999**, *5*, 2455. doi:10.1002/(SICI)1521-3765(19990903)5:9<2455::AID-CHEM2455>3.0.CO;2-H

[18] A. V. Eliseev, J. M. Lehn, *Comb. Chem. Biol.* **1999**, *243*, 159.

[19] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, S. Otto, *Chem. Rev.* **2006**, *106*, 3652. doi:10.1021/CR020452P

[20] O. Ramström, J.-M. Lehn, *Nat. Rev. Drug Discov.* **2002**, *1*, 26. doi:10.1038/NRD704

[21] S. Otto, K. Severin, *Top. Curr. Chem.* **2007**, *277*, 267. doi:10.1007/128_2007_116

[22] S. Ladame, *Org. Biomol. Chem.* **2008**, *6*, 219. doi:10.1039/B714599C

[23] J. N. H. Reek, S. Otto (Eds), *Dynamic Combinatorial Chemistry: In Drug Discovery, Bioorganic Chemistry, and Materials Science* **2010** (Wiley-VCH: Weinheim).

[24] A. R. Hunt, S. Otto, *Chem. Commun.* **2011**, *47*, 847. doi:10.1039/C0CC03759A

[25] S. Ladame, A. M. Whitney, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2005**, *44*, 5736. doi:10.1002/ANIE.200501450

[26] A. M. Whitney, S. Ladame, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2004**, *43*, 1143. doi:10.1002/ANIE.200353069

[27] B. R. McNaughton, B. L. Miller, *Org. Lett.* **2006**, *8*, 1803. doi:10.1021/OL060330+

[28] C. Karan, B. L. Miller, *J. Am. Chem. Soc.* **2001**, *123*, 7455. doi:10.1021/JA010325V

[29] B. Klekota, M. H. Hammond, B. L. Miller, *Tetrahedron Lett.* **1997**, *38*, 8639. doi:10.1016/S0040-4039(97)10374-4

[30] B. Klekota, B. L. Miller, *Tetrahedron* **1999**, *55*, 11687. doi:10.1016/S0040-4020(99)00687-0

[31] J. Portugal, M. J. Waring, *Eur. J. Biochem.* **1987**, *167*, 281. doi:10.1111/J.1432-1033.1987.TB13334.X

[32] J. G. Pelton, D. E. Wemmer, *Biochemistry* **1988**, *27*, 8088. doi:10.1021/BI00421A018

[33] S. Tsujita, M. Tanada, T. Kataoka, S. Sasaki, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 68. doi:10.1016/J.BMCL.2006.09.089

[34] S. Dawson, J. P. Malkinson, D. Paumier, M. Searcey, *Nat. Prod. Rep.* **2007**, *24*, 109. doi:10.1039/B516347C

[35] B. R. McNaughton, P. C. Gareiss, B. L. Miller, *J. Am. Chem. Soc.* **2007**, *129*, 11306. doi:10.1021/JA072114H

[36] S. C. M. Teixeira, J. H. Thorpe, A. K. Todd, H. R. Powell, A. Adams, L. P. G. Wakelin, W. A. Denny, C. J. Cardin, *J. Mol. Biol.* **2002**, *323*, 167. doi:10.1016/S0022-2836(02)00923-3

[37] C. Rajput, R. Rutkaite, L. Swanson, I. Haq, J. A. Thomas, *Chemistry* **2006**, *12*, 4611. doi:10.1002/CHEM.200501349

[38] A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2008**, *47*, 2677. doi:10.1002/ANIE.200705589

[39] K. Jantos, R. Rodriguez, S. Ladame, S. Shirude Pravin, S. Balasubramanian, *J. Am. Chem. Soc.* **2006**, *128*, 13662. doi:10.1021/JA064713E

[40] M. C. Nielsen, T. Ulven, *Chemistry* **2008**, *14*, 9487. doi:10.1002/CHEM.200801109
[41] B. Gooch, P. A. Beal, *J. Am. Chem. Soc.* **2004**, *126*, 10603. doi:10.1021/JA047818V
[42] P. C. Gareiss, K. Sobczak, B. R. McNaughton, P. B. Palde, C. A. Thornton, B. L. Miller, *J. Am. Chem. Soc.* **2008**, *130*, 16254. doi:10.1021/JA804398Y
[43] A. Bugaut, J.-J. Toulime, B. Rayner, *Angew. Chem. Int. Ed.* **2004**, *43*, 3144. doi:10.1002/ANIE.200454041
[44] A. Bugaut, K. Bathany, J.-M. Schmitter, B. Rayner, *Tetrahedron Lett.* **2005**, *46*, 687. doi:10.1016/J.TETLET.2004.11.110
[45] A. Bugaut, J.-J. Toulmé, B. Rayner, *Org. Biomol. Chem.* **2006**, *4*, 4082. doi:10.1039/B610890C
[46] L. Azéma, K. Bathany, B. Rayner, *ChemBioChem* **2010**, *11*, 2513. doi:10.1002/CBIC.201000538
[47] F. R. Bowler, J. J. Diaz-Mochon, M. D. Swift, M. Bradley, *Angew. Chem. Int. Ed.* **2010**, 1809. doi:10.1002/ANIE.200905699
[48] D. T. Hickman, N. Sreenivasachary, J.-M. Lehn, *Helv. Chim. Acta* **2008**, *91*, 1. doi:10.1002/HLCA.200890022
[49] A. Krezel, W. Bal, *Acta Biochim. Pol.* **1999**, *46*, 567.
[50] C. K. Sen, *Curr. Top. Cell. Regul.* **2001**, *36*, 1. doi:10.1016/S0070-2137(01)80001-7
[51] R. F. Ludlow, S. Otto, *J. Am. Chem. Soc.* **2010**, *132*, 5984. doi:10.1021/JA1013689
[52] C. C. O'Hare, P. Uthe, H. Mackay, K. Blackmon, J. Jones, T. Brown, B. Nguyen, W. D. Wilson, M. Lee, J. A. Hartley, *Biochemistry* **2007**, *46*, 11661. doi:10.1021/BI701053A
[53] M. Munde, M. Lee, S. Neidle, R. Arafa, D. W. Boykin, Y. Liu, C. Bailly, W. D. Wilson, *J. Am. Chem. Soc.* **2007**, *129*, 5688. doi:10.1021/JA069003N
[54] M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, Y. Liu, L. Hu, M.-H. David-Cordonnier, A. Lansiaux, C. Bailly, D. W. Boykin, W. D. Wilson, *J. Am. Chem. Soc.* **2007**, *129*, 13732. doi:10.1021/JA074560A
[55] J. Lee, V. Guelev, S. Sorey, D. W. Hoffman, B. L. Iverson, *J. Am. Chem. Soc.* **2004**, *126*, 14036. doi:10.1021/JA046335O
[56] Y. Chu, S. Sorey, D. W. Hoffman, B. L. Iverson, *J. Am. Chem. Soc.* **2007**, *129*, 1304. doi:10.1021/JA066480X
[57] J. Portugal, D. J. Cashman, J. O. Trent, N. Ferrer-Miralles, T. Przewloka, I. Fokt, W. Priebe, J. B. Chaires, *J. Med. Chem.* **2005**, *48*, 8209. doi:10.1021/JM050902G
[58] L. Hunter, G. R. Condie, M. M. Harding, *Tetrahedron Lett.* **2010**, *51*, 5064. doi:10.1016/J.TETLET.2010.07.105