

The Nucleic Acid Ligand

A New Tool for Molecular Recognition

Selective binding through molecular recognition is the basis of many important techniques in chemical analysis and separations, including those that employ enzymes, antibodies, or selective chelators. Reagents such as cyclodextrins that have preferential interaction with one isomer or enantiomer over another, often combined with selectivity based on size or shape, have been used to accomplish chiral or isomeric separations. Affinity-based separations depend on specific molecular recognition between immobilized ligands and their receptors, which are the target analytes. Chemical sensors often use selective binders that are immobilized at the sensor surface to generate an analyte-dependent signal.

**Linda B. McGown and
Melissa J. Joseph**

Duke University

**J. Bruce Pitner, Glenn P.
Vonk, and C. Preston Linn**
Becton Dickinson Research Center

Oligonucleotide ligands provide specific and high-affinity binding with selected target molecules

A recent entry in the field of selective binding and molecular recognition is the nucleic acid ligand, an oligonucleotide that exhibits high-affinity specific binding with selected target molecules (excluding hybridization interactions such as double-strand formation through base pairing). These ligands have ranged from 8 to 120 nucleotides in length, corresponding to a molecular weight range of ~ 3000–40,000. The ligands are typically truncated to a “consensus” region, which is the

minimal sequence needed for binding to the target and is usually 15–50 bases in length.

Individual sequences that have high binding affinities for a target analyte are selected from oligonucleotide libraries of as many as 10^{15} random sequences by a selective, iterative enrichment process (1, 2). The selective binding affinity toward the target is thought to arise from specific interactions such as hydrogen bonding or association with the phosphate groups of the ligand, or “aptamer” (2). These interactions are facilitated by the sequence-specific, 3-D structure of the ligand, which provides a rigid scaffold for the arrangement of functionalities of the ligand. Examples of 3-D structures (Figure 1) include the stem-loop/bulge (3), the pseudoknot (4), the helix (not shown), the hairpin (1), and the G-quartet (5–7) structure, which is characteristic of a thrombin-binding ligand.

Nucleic acid ligand selection

Because the probability that a given sequence will form a stable, 3-D structure

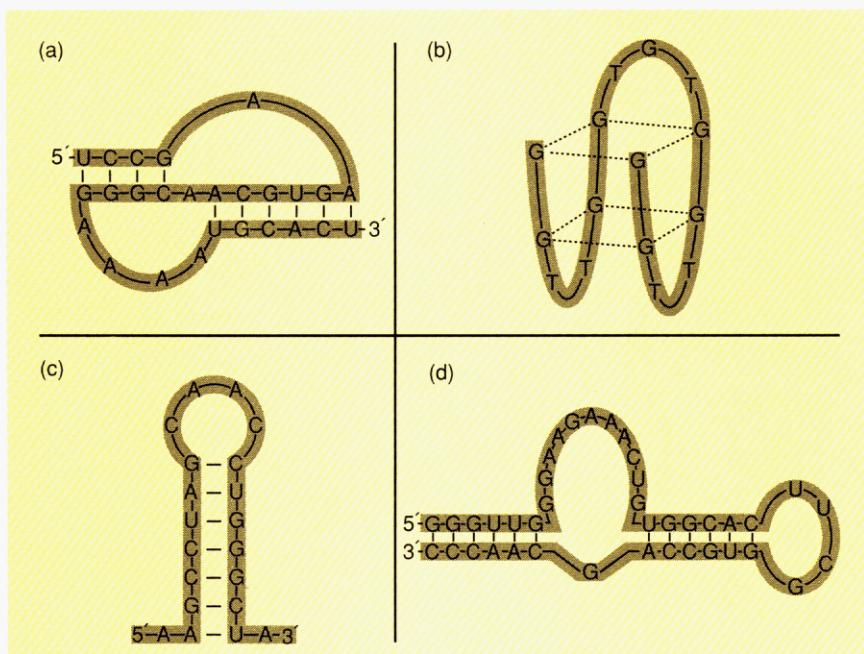


Figure 1. Structures of some nucleic acid ligands.

(a) Pseudoknot (RNA ligand for HIV-1 reverse transcriptase [4]), (b) G-quartet (DNA ligand for thrombin [5]), (c) hairpin (RNA ligand for Bacteriophage T4 polymerase [1]), and (d) stem-loop bulge (RNA ligand for ATP [3]).

with a high binding affinity for a particular target molecule is very low, it is necessary to select ligands from a very large (10^{15} sequence) pool to maximize the chances of success. Consequently, the development of techniques to generate large, random DNA or RNA sequence libraries and to isolate molecules with specific binding affinities from these libraries is critical to the development of nucleic acid ligands as important binding reagents (1, 2).

For very short oligomers (15–25 nucleotides), all possible sequences may be included in the initial pool. However, such short strands may not fully represent the structure space needed to provide the desired binding properties. As the length of the oligonucleotide increases, the number of possible sequences increases exponentially by y^N , where y is the number of different oligonucleotides and N is the number of random positions.

With four different nucleotides (i.e., 4^N), it's not feasible to include all sequences with greater than ~ 25 random positions in the initial pool for a given selection process. Therefore, alternative strategies for generating the initial sequence pool are used, such as a "shotgun"

approach in which the entire range of possibilities is randomly sampled, or a more focused strategy in which the sequences are clustered about a particular sequence that has been identified as having the desirable binding properties (8, 9). Although the use of longer random oligonucleotides offers the advantages of easier generation of random sequence pools and more comprehensive spanning of the structural space, it also increases the likelihood of side reactions and of errors in the amplification process that may terminate strand replication.

Once the initial sequence pool has been generated, ligands with the desired binding characteristics are isolated by iterative in vitro processes that have variously been referred to as "systematic evolution of ligands by exponential enrichment" (SELEX) (1), in vitro selection (2), directed molecular evolution (10), or "evolution in a test tube" (11). Figure 2 summarizes the steps in these methods. The sequence pool is commonly passed over a support, such as an affinity column, to which the target molecule or macromolecule is attached. Numerous cycles of this partition procedure are repeated, each followed by polymerase chain reaction

amplification of the sequences that are highly retained on the support. If RNA is used, it is transcribed from the DNA template using a suitable promoter sequence and an RNA polymerase.

Table 1 (References 12–20) lists some of the ligands that have been isolated by ligand selection processes, with both RNA and DNA ligands represented. The targets include proteins and enzymes as well as a variety of small molecules. Some ligands exhibit stereoselectivity, such as the ligand that binds to agarose-bound D-tryptophan but not to L-tryptophan. Many of the ligands to macromolecules are notable for their ability to inhibit the action of their target macromolecule.

In contrast to methods such as SELEX, which screen entire libraries in parallel, alternative combinatorial methods successively fix positions in a biopolymer such as an oligonucleotide. Libraries are prepared in which the first position in the oligonucleotide chain is varied among the four possible bases. The rest of the positions in the polymer are allowed to vary randomly. The four different libraries, one for each base in the first position, are screened for binding activity. The identity of the first position is then fixed to that of the library exhibiting the tightest binding in the screening assays. The process is repeated at the second position, and so on progressively down the oligonucleotide chain. Generally, these methods screen a smaller number of molecules. However, the synthetic flexibility of these combinatorial methods permits a broader menu of structures uncompromised by the enzymatic requirements of permutational methods such as SELEX. Synthetic combinatorial methods have most recently yielded a DNA ligand that inhibits infection by HIV in vitro (21).

Chemical selectivity and stability

The binding of nucleic acid ligands to target molecules can be chemically selective as well as stereospecific. One selection experiment produced an RNA pool that bound to D-tryptophan rather than L-tryptophan, a molecule differing at only one stereocenter, by a greater than nine-fold preference (13). One individual clone from this pool had 670-fold greater affinity for D-tryptophan than for L-trypto-

phan. Even greater degrees of selectivity can be achieved by incorporating target discrimination explicitly in the selection strategy. For example, to encourage selectivity of one molecule over another, another partition step can be added to the selection process shown in Figure 2 in which the column is “washed” with the undesired molecule prior to elution with the target. Such a counter-selection process has yielded an RNA ligand that has > 10,000-fold selectivity for theophylline over caffeine, molecules that differ by only one methyl group (12).

Limitations on the stability of nucleic acid ligands are an important consideration in their use as analytical reagents. This is particularly important for RNA, which is readily degraded by ribonucleases in samples of biological origin. Stability of DNA ligands is more of a concern when they are used as therapeutics and extended *in vivo* stability is needed. In practice, DNA may be handled routinely in most laboratories without exceptional precautions. However, RNA should be handled with gloves to limit contamination with nucleases present on the skin, and “nuclease-free” reagents should be used.

To increase the stability of nucleic acid ligands, they can be constructed from

chemically modified derivatives with structures similar to DNA or RNA. Because most enzymatic degradation of RNA occurs through intramolecular participation of the 2' hydroxyl on the ribose sugar of pyrimidine nucleotides, substitution of this functionality with fluorine, amino, or alkoxy substituents greatly enhances the stability of these oligomers (22). Other modifications include 2'-O-methyl derivatives, carbocyclic ribose analogues, thiophosphates, and modifications of the pyrimidine or purine bases. In one recent study, a 2'-amino pyrimidine modification extended the half-life of RNA in both serum and urine from a few minutes to several hours (23). DNA may also be stabilized through chemical modification (24).

Nucleic acid ligand structures

The structure of nucleic acid ligands can be studied by a number of techniques, including X-ray crystallography, NMR, circular dichroism (CD), UV-vis and IR absorption spectroscopies, and fluorescence probe. Comparisons can be made among different sequences as well as between the ligands and longer, double-stranded DNA or RNA. For example, Figure 3 shows the CD spectra of three different DNAs: a

thrombin-binding DNA ligand that contains the G-quartet structure (5'-GGTTGGTGGTGG-3'), a DNA oligomer with the same base content as the ligand in a “scrambled” sequence that does not promote intramolecular G-quartet formation (5'-GGTGGTGGTGTGGT-3'), and duplex (double-stranded) DNA (dsDNA) (25). The CD spectrum of the thrombin-binding ligand is clearly distinct from the other two DNAs because of the unique intramolecular G-quartet. It is this G-quartet structure that underlies the unique high affinity of the thrombin-binding ligand for thrombin.

Interactions with nucleic acid indicator dyes

Indicator dyes that bind to double-stranded helical DNA or RNA may also bind to single-stranded ligands. For example, dyes such as oxazole yellow (YO), its homo dimer (YOYO) (26), and other related probes have been found to associate with the shorter, single-stranded ligands. These probes are essentially nonfluorescent in bulk aqueous solution, but they develop intense fluorescence upon association with double-stranded DNA or RNA. In the thrombin-binding ligand and the scrambled-sequence oligomer, YOYO and YO exhibit strong fluorescence, despite the absence of intercalation sites provided by dsDNA. Excitonic coupling leads to an induced CD spectrum of YOYO in the nucleic acids, which is further evidence of binding. At high dye loadings in dsDNA, intramolecular dimerization between surface-bound YO groups of a single, folded YOYO gives rise to a -/+ bisignate CD spectrum (27). In both the thrombin-binding ligand and the scrambled-sequence oligomer, a +/- bisignate CD spectrum is observed for YOYO at a dye loading below the threshold that was reported for excitonic coupling in the dsDNA (25). This indicates differences between the binding of YOYO in the single-stranded and double-stranded DNAs.

We have found that the fluorescent dyes Hoechst 33234 and 33258 also bind to the single-stranded DNA ligands. These two dyes are similar compounds and minor groove binders in dsDNA. Their association with the single-stranded oligomers is weaker than was observed for the intercalating YO and YOYO dyes,

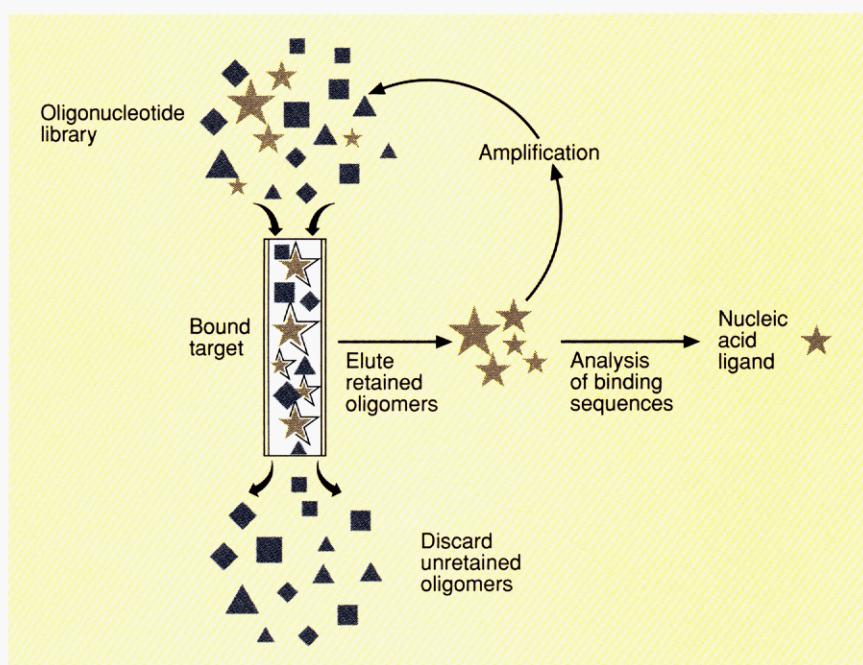


Figure 2. General summary of ligand selection processes.

Random sequence library undergoes partitioning for selection of binders for immobilized target; retained sequences are repetitively eluted and cycled through the selection and amplification processes to isolate families of nucleic acid ligands for the target.

Table 1. Some nucleic acid ligands and their target molecules

Target	Ligand type ^a	Ligand structure ^b	Reference
Small molecules			
Organic dyes	DNA		2
Theophylline	RNA	Hairpin with bulge	12
D-Tryptophan ^c	RNA		13
ATP	RNA	Stem-loop/bulge	3
L-Citralline/ L-Arginine ^d	RNA		14
Arginine	RNA		15
Cyanocobalamin (vitamin B-12)	RNA	Pseudoknot	16
Biological cofactors (FAD, FMN, NAD ⁺ , NMN ⁺)	RNA	Hairpin/bulge	17
Macromolecules			
Human thrombin	DNA	G-quartet	5
Bacteriophage T4 polymerase	RNA	Hairpin	1
Antipeptide antibody	RNA	Hairpin	18
Basic fibroblast growth factor	RNA		19
HIV-1 reverse transcriptase	RNA	Pseudoknot	4
<i>E. coli</i> RhD factor	RNA	Hairpin	20

^a All are single strands.^b If known.^c The ligand was selective for D-trp-agarose over L-trp-agarose.^d The L-arginine-binding ligand was evolved from the pool of L-citrullene-binding ligands and did not retain any affinity for L-citrullene.

which suggests the absence of an analogue to a minor groove-binding site or a suitable alternative in the single-stranded structures.

Further investigation of indicator dyes that bind to duplex nucleic acids will improve our understanding of the conformation and binding interactions of nucleic acid ligands and may lead to using these dyes as indicators of the ligands and their target analytes.

Nucleic acid ligands as analytical reagents

The application of nucleic acid ligands to chemical analysis is a new area of investigation with only a few specific examples to date, primarily in clinical diagnostics. Yet the possibility of generating stable structures with unique conformations offers enormous potential for chemical sensing and separations. This is analogous to the use of enzymes and antibodies in recent years but offers the advantages of smaller, less cumbersome molecules that, once identified, are simple to manufacture and manipulate. These are important factors, particularly for immobilization of these

reagents at surfaces of sensors or chromatographic supports because the smaller ligands will reduce steric hindrance and increase surface coverage and their conformational stability will help maintain their selectivity and activity upon attachment to a surface. Furthermore, using reversible attachment methods based on hybridization offers exciting possibilities for replacement or renewal of ligands at sensor or chromatographic surfaces.

A unique combination of stereoselectivity and chemical recognition is possible with nucleic acid ligands, which are in some respects similar to cyclodextrins but have greater structural variety and lack the size exclusion imposed by the rigid cyclodextrin cavity. The structural motifs that provide very specific sensing of designated target molecules may also show more general selectivity for a variety of unrelated molecules, which could be used to develop new methods for chemical and chiral separations. On the other hand, binding affinity for molecules unrelated to the target analyte may lead to unanticipated sources of interference that must be investigated.

Nucleic acid ligands and antibodies

A close analogy to the nucleic acid ligand is the antibody. Antibodies are proteins that develop molecular recognition by *in vivo* exposure of the unspecified immunoglobulin to the target (or target–carrier complex) through natural or artificially induced immunogenic response. In many ways, the evolution of molecular recognition in the immune response is analogous to the selection of nucleic acid ligands. In both cases, molecular recognition arises from the 3-D structure of the host (antibody or ligand) and its specific physicochemical interactions with the target analyte.

Antibodies, or antibody fragments, have binding constants on the order of 10^6 – 10^{12} . They are much larger than the nucleic acid ligands; molecular weights range from $\sim 160,000$ for the protein to 25,000 for SFv antibody fragments. They may be polyclonal, composed of a heterogeneous mixture of immunoglobulins with binding affinities for several determinant structures on the target molecule, or monoclonal, which is a homogeneous, pure species of immunoglobulin with selected specificity for a unique determinant on the target. Although homogeneity of monoclonal antibodies is advantageous for reproducibility and predictability, polyclonal antibodies are frequently more effective in immunoassays.

Nucleic acid ligands offer several potential advantages over traditional antibody-based reagents because they are not derived from living organisms and can be reproducibly and accurately synthesized in a short time by automated processes. Covalent attachment of dyes to nucleic acid ligands is relatively simple and may be done with high specificity at one or more locations on the ligand.

Other chemical modifications for stabilization, increased activity, or covalent attachment are also relatively simple. In traditional antibody production methods, the target molecule must be large enough to elicit an immune response (molecular weights of 1000 will provide marginal immunogenicity, and above 10,000, the response is usually strong); small molecular targets must be attached to a large carrier molecule, such as albumin, to generate antibodies to the target. Non- or

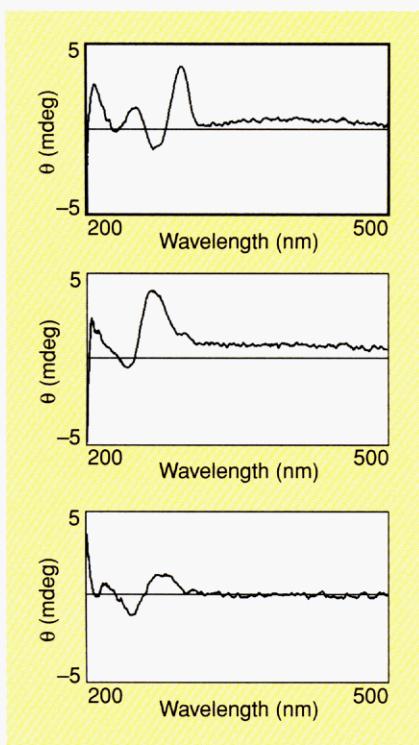


Figure 3. CD spectra of three different DNAs.

Thrombin-binding DNA ligand (top), a DNA oligomer of the same length and composition but different sequence (middle), and duplex DNA (bottom).

poorly immunogenic analytes that are problematic for antibody-based methods may be targeted by nucleic acid ligands, although the isolation of a highly selective ligand with a high binding affinity is by no means guaranteed for a given target analyte.

Like antibodies, nucleic acid ligands can be immobilized on electrodes or optical fibers for highly selective chemical sensing. Because the ligands are smaller and their 3-D structures are less complicated than those of antibodies, immobilization and subsequent binding interactions may encounter fewer steric hindrances and less degradation of binding activity. An important advantage of nucleic acid ligands as immobilized sensors is that they can easily be denatured to reverse binding and then regenerated simply by controlling buffer-ion concentrations. For example, the G-quartet structure can be controlled by altering the K^+ concentration. Antibody-based sensors generally require more drastic conditions, such as low pH, for regeneration, and they lose binding ability after repeated cycles.

Detection methods

The use of labeled nucleic acid ligands offers sensitive and simple methods for measuring binding to specific analytes. The techniques that have been developed for immunochemical analyses, including heterogeneous (separation based) and homogeneous (nonseparation) methods, are generally suitable for nucleic acid ligands as well. Detection can be accomplished with radiolabels in heterogeneous techniques, but using radioactive materials is generally discouraged because of their inherent danger and instability. Fluorescent labels are a less hazardous alternative that can be used in either heterogeneous or homogeneous analyses, through measurements of intensity, lifetime, anisotropy, or energy transfer. The use of fluorescent-labeled ligands allows direct signal generation without the need to separate bound from free labeled ligand.

One of the most successful homogeneous techniques is fluorescence polarization analysis. The change in signal upon binding of the analyte to the binding agent is related to the increase in size of the labeled moiety upon binding and results in a corresponding change in the effective rotational rate of the labeled moiety. Polarization analysis is an excellent example of a case in which nucleic acid ligands offer

important advantages over antibodies, as illustrated in Figure 4.

In immunoassay techniques, the change in size of the antibody upon binding of the analyte is often small because the antibody itself is large. Therefore, labeled analyte compound, rather than labeled antibody, is generally used. This necessitates a competitive determination scheme in which the unlabeled analyte competes with the labeled analyte reagent for antibody binding sites. In contrast, the relatively small nucleic acid ligands will experience a proportionally greater increase in effective size upon analyte binding, making it possible to label the ligands and measure the increase in their fluorescence polarization upon binding to the analyte in a direct, rather than a competitive, analysis (28).

The future

The exploration of nucleic acid ligands as reagents for chemical analysis will encompass a multitude of research objectives. New ligands for wide-ranging target analytes, both large and small, will continue to be identified through SELEX and other selection techniques. Innovations in the methodology for generating random libraries and identifying binding sequences will be pursued. Fundamental

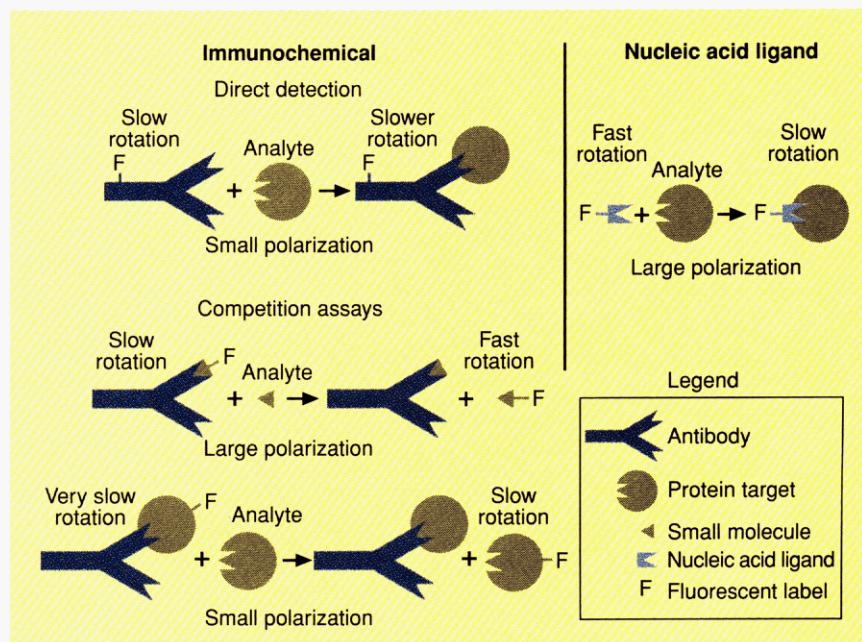
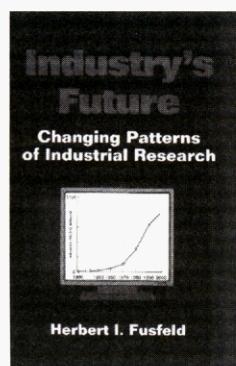


Figure 4. Comparison of polarization detection in immunochemical antibody-based methods and nucleic acid ligand-based methods.

For clarity, only one binding site interaction is shown for a given antibody.

Just Published!

Industry's Future: Changing Patterns of Industrial Research



This fascinating volume provides the readers with an understanding of the dynamic processes that make industrial research a principal driving force

in creating technical change, producing economic growth, and strengthening the technical institutions in society.

It describes and analyzes the major factors that shape the conduct and organization of industrial research, including the internationalization of R&D, restructuring of industry, declining defense expenditures and the pressure to develop access to sources of technical change outside the corporation.

The volume presents approaches for improved industry relations with government and academia, discusses complex conditions for conducting industrial research effectively and analyzes potential new conditions that can shape future industrial research. The experiences of specific corporations with modern management of technology are also described.

Herbert I. Fusfeld

384 pages (1994)
Clothbound: ISBN 0-8412-2983-X
\$39.95
Paperbound: ISBN 0-8412-2984-8
\$24.95

Order from

American Chemical Society
Distribution Office Dept. 74
1155 Sixteenth Street, NW
Washington, DC 20036

Or Call TOLL FREE 1-800-227-5558
(in Washington, D.C. 872-4363)
and use your credit card!
FAX: 202-872-6067.

ACS Publications Catalog now available on
internet: gopher.acsinfo.acs.org

ACS PUBLICATIONS
Essential Resources for the Chemical Sciences

Report

studies of the nature of molecular recognition by the ligands, and the dependence of binding strength and selectivity on the sequence, structure, and conformation of the ligands, will be investigated, as will the effects of experimental conditions and chemical modifications. Further identification and characterization of structural motifs and physicochemical interactions may lead to more rational and efficient approaches to the design of new ligands.

Applications of nucleic acid ligands will expand beyond clinical diagnostics and therapeutic monitoring to a broader arena of analytical chemistry. Immobilization chemistry, including reversible attachment and denaturation schemes, will facilitate the use of ligands as reagents at sensor or chromatographic support surfaces. Explorations of chiral recognition may lead to applications in the separation of enantiomers. Development of novel detection strategies will play a key role in the utilization of nucleic acid ligands to maximize the effectiveness of these reagents. Because methods for isolation, characterization, and modification of nucleic acid ligands are still at the embryonic stage, new properties and applications of these uniquely versatile reagents will unfold with further study.

The authors are grateful to Bob Hanson of Becton Dickinson for designing the figure graphics and to Larry Gold and Barry Polisky of NeXstar Pharmaceuticals for their helpful comments.

References

- (1) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505.
- (2) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818.
- (3) Sasanfar, M.; Szostak, J. W. *Science* **1993**, *364*, 550.
- (4) Tuerk, C.; MacDougal, S.; Gold, L. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6988.
- (5) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564.
- (6) Wang, K. Y.; McCurdy, S.; Shea, R. G.; Swaminathan, S.; Bolton, P. H. *Biochemistry* **1993**, *32*, 1899.
- (7) Macaya, R. F.; Schultz, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745.
- (8) Wilson, C.; Szostak, J. W. *Nature* **1995**, *374*, 777.
- (9) Nieuwlandt, D.; Wecker, M.; Gold, L. *Biochemistry* **1995**, *34*, 5651.
- (10) Joyce, G. F. *Sci. Am.* **1992**, *269*, 90.
- (11) Schmidt, K. F. *Sci. News* **1993**, *144*, 90.
- (12) Jenison, R. D. et al. *Science* **1994**, *263*, 1425.
- (13) Famulok, M.; Szostak, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 3990.
- (14) Famulok, M. *J. Am. Chem. Soc.* **1994**, *116*, 1698.
- (15) Connell, G. J.; Illangesekare, M.; Yarus, M. *Biochemistry* **1993**, *32*, 5497.
- (16) Lorsch, J. R.; Szostak, J. W. *Biochemistry* **1994**, *33*, 973.
- (17) Burgstaller, P.; Famulok, M. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1084.
- (18) Tsai, D. E.; Kenan, D. J.; Keene, J. D. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8864.
- (19) Jelinek, D.; Lynott, C. K.; Rifkin, D. B.; Janic, N. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11227.
- (20) Schneider, D.; Gold, L.; Platt, T. *FASEB J.* **1993**, *7*, 201.
- (21) Wyatt, J. R. et al. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1356.
- (22) Piecken, W. A.; Olsen, D. B.; Benseler, F.; Aurup, H.; Eckstein, F. *Science* **1991**, *253*, 314.
- (23) Lin, Y.; Qiu, Q.; Gill, S. C.; Jayasena, S. *Nucleic Acids Res.* **1994**, *22*, 5229.
- (24) Shaw, J. P.; Kent, K.; Fishbach, J.; Froehler, B. *Nucleic Acids Res.* **1991**, *19*, 747.
- (25) Joseph, M. J.; Taylor, J. C.; McGown, L. B.; Pitner, J. B.; Linn, C. P., submitted for publication in *J. Phys. Chem.*
- (26) Rye, H. S. et al. *Nucleic Acids Res.* **1992**, *20*, 2803.
- (27) Larsson, A.; Carlsson, C.; Jonsson, M.; Albinsson, B. *J. Am. Chem. Soc.* **1994**, *116*, 8459.
- (28) Hsieh, H. V. *Biophys. J.* **1995**, *68*, A298.

Suggested reading

- Edgington, S. M. *Bio/Technology* **1993**, *11*, 285.
The RNA World; Gesteland, R. F.; Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1993; Chapters 19 and 20.
Burke, J. M.; Berzal-Herranz, A. *FASEB J.* **1993**, *7*, 106-12.
Ellington, A. D. *Curr. Biol.* **1994**, *4*, 427.
Klug, S. J.; Famulok, M. *Mol. Biol. Rep.* **1994**, *20*, 97.

Linda B. McGown, professor of chemistry, performs research in fluorescence lifetime techniques, DNA detection and sequencing, molecular probe techniques, and organized media. Melissa J. Joseph recently completed her Ph.D. in chemistry at Duke University and is now a research chemist at Lorillard Tobacco Company in Greensboro, NC. Glenn P. Vonk uses organic chemistry to develop novel systems for biochemical analysis; J. Bruce Pitner performs research in bioconjugation, molecular recognition, and combinatorial chemistry; and C. Preston Linn develops novel chemical detection methods in the Molecular Biology Department at Becton Dickinson Research Center. Address comments to McGown at Department of Chemistry, Box 90346, Duke University, Durham, NC 27708-0346.