Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry 11 (2003) 1607–1613

BIOORGANIC & MEDICINAL CHEMISTRY

Perspective

Solution-Phase Synthesis of Combinatorial Libraries Designed to Modulate Protein–Protein or Protein–DNA Interactions

Dale L. Boger*

Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Received 6 December 2002; accepted 6 December 2002

Abstract—A short personal perspective on the development of an approach to the solution-phase synthesis of combinatorial libraries for modulating cellular signaling by inhibiting, promoting, or mimicking protein–protein or protein–DNA interactions is provided.
© 2003 Elsevier Science Ltd. All rights reserved.

Over the course of the last several years, we have developed technology for the solution-phase preparation of combinatorial libraries with several key applications in mind. Because this work was necessarily published at different intervals across a range of chemical or biological journals, the strategic decisions we made in pursuing this work and the intimate relationship between the technology development and the intended applications is not likely to be clear to the casual reader. Consequently, I would like to take the opportunity of this Perspective Articles series to summarize our contributions to solution-phase library synthesis highlighting its development in the context of our intended applications for modulating cellular signaling through inhibiting, promoting, or mimicking protein–protein or protein–DNA interactions. Despite the prevailing bias that such targets might not prove viable for small molecule intervention, the screening of our libraries against each such target has provided the first small molecule modulators of the protein–protein or protein–DNA interaction and validated the targets for small molecule therapeutic intervention.

Solution-Phase Combinatorial Library Synthesis

Combinatorial chemistry has undergone rapid development and has provided a new paradigm for drug discovery. As a consequence of its extension from peptide and oligonucleotide synthesis, most approaches have relied on solid-phase synthesis techniques. A complement to adapting solution-phase chemistry to polymer-supported combinatorial synthesis is the development of protocols for solution-phase combinatorial synthesis. We introduced a simple protocol that permits the multistep synthesis of chemical libraries employing liquid–liquid and liquid–solid extractions to remove unreacted starting materials, reagents and reagent by-products, providing the purified product (>95% pure) irrespective of the reaction efficiency (Fig. 1).[1–34] It has been implemented in formats for the synthesis of individual compounds[1–4] (1000-member libraries), modest sized libraries composed of small mixtures (1000- to 10,000-member libraries, 10–50 compounds/mixture),[6] or combinatorially assembled to provide large libraries including positional scanning or deletion synthesis libraries[8,13,21,34] (25,000- to 1,000,000-member libraries, 100–28,000 compounds/mixture).[7,8] This allows the protocol to be adopted in a format compatible with any screening objective. Thus, its implementation is convenient for either lead discovery or lead optimization and produces the library members on a scale (5–150 mg) that allows their repeated use in screening without resynthesis. It is this latter feature along with its technically nondemanding implementation that we consider its greatest attributes. We presently have 40,000 compounds in the small (1000-member) and medium-sized (1000–10,000-member) libraries and efforts are ongoing to expand this to approximately 1,000,000 compounds. Such libraries have been prepared that interfere with (antagonists) or mimic (agonists) extracellular or intracellular protein–protein interactions,[35–37] inhibit intracellular enzymes, or modulate protein–DNA interactions. The chemistry is applicable to non-natural and natural product scaffolds, and cyclic (depsi)peptides and possesses a scope that exceeds what one might initially imagine based on its simplicity.

*Corresponding author. Tel.: +1-858-784-7522; fax: +1-858-784-7550; e-mail: boger@scripps.edu

0968-0896/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0968-0896(03)00031-2

Highlights of the methodology developed for the preparation of chemical libraries include:

- A general solution-phase technique for the high-throughput synthesis of libraries in a variety of formats (individual compounds, small mixtures, large mixtures, positional scanning libraries, deletion synthesis libraries) was introduced that is applicable to both lead identification and lead optimization.
- Liquid–liquid or liquid–solid (ion-exchange resin) extractions were developed.[1,2,11]
- Immobilized reagents for use in the solution-phase synthesis of libraries were introduced.[2,10–12]
- The first use of size exclusion chromatography (filtration) for the high-throughput purification and isolation of chemical libraries was disclosed.[11]
- The first application of an intermolecular olefin metathesis reaction in the synthesis of chemical libraries was disclosed.[5,7–9]
- The approach is applicable to the multistep synthesis of selected natural products and this was highlighted with a 2640-membered distamycin library,[20] 132-membered CC-1065 library,[24] 40-membered HUN-7293 library,[18] and a prototype triostin A synthesis.[22]
- Deletion synthesis deconvolution of mixture libraries was introduced and assessed.[8]
- The performance of positional scanning deconvolution was and continues to be compared against deletion synthesis deconvolution, small mixture screening, and individual compound screening.[8,13,21,34]
- Convergent versus divergent strategies of library synthesis were introduced and developed (only applicable to solution-phase, not solid-phase, synthesis).[7–12]
- Agonists from antagonists design concepts implemented.[32]
- A rapid, high-throughput, high-resolution screen for DNA binding affinity and selectivity applicable for use in assaying libraries was developed (FID assay).[20,23,25–27]
- Current archived chemical libraries contain $>40{,}000$ compounds that are available for continued screening in new assays.

The approach avoids the disadvantages of solid-supported synthesis including its more restrictive scale, the required functionalized substrates and solid supports, compatible spacer linkers, and the requirements for orthogonal attachment/detachment chemistries. It does not require specialized protocols for monitoring each step of multistep syntheses, allows the purification of intermediates, and provides the final pure products directly for use in binding or functional assays. We described extensions of these studies for generating symmetrical or unsymmetrical chemical libraries suitable for probing receptor and protein homo- and heterodimerization events (Fig. 2). Thus, the preparation and dimerization linkage of iminodiacetic acid diamides can be conducted in a reaction sequence that requires only three steps. In addition to the multiplication of the diversity that arises through the combinatorial dimerization linkage of the iminodiacetic acid diamides, the solution-phase synthesis of the intermediates permits their direct linkage which would be precluded by solid-phase synthesis techniques. As such, the strategy is uniquely suited for taking advantage of such dimerization (convergent) strategies utilizing a limited number of synthetic steps. This modular approach to the generation of libraries is especially well-suited for the discovery of antagonists or agonists of receptor and protein homo- and heterodimerization. Simple binders (i.e., **3**) can serve as antagonists of ligand-induced receptor or protein dimerization. Covalently linked symmetrical dimers (i.e., **4**) can be used to promote receptor or protein homodimerization whereas unsymmetrical dimers can be utilized to promote receptor or protein heterodimerization. Thus, both antagonists and agonists may be developed depending on the therapeutic application.[41,45]

Target Protein–Protein Interactions Mediating Cellular Signaling

Cell growth, differentiation, migration, and apoptosis are regulated in part by growth factors or cytokines. These factors are unable to penetrate the cell membrane and exert their effects by binding to cell surface receptors. In many instances, such receptors are activated by ligand-induced dimerization or oligomerization.[38–44] In addition, several components of the intracellular signal transduction pathways are also regulated by dimerization. For instance, certain cytoplasmic signal transduction molecules dimerize after activation, and the active form of a transcription factor is often a dimer.[42–45] Thus, protein dimerization has emerged as a general mechanism for the initiation and downstream regulation of signal transduction. The targets we are addressing constitute prototypical examples of these events in signal transduction. Thus, the targets being pursued were chosen not only for their therapeutic importance, but also because each constitutes a distinct stage for modulating cellular signaling by controlling protein–protein interactions.

Effective inhibitors of angiogenesis (new blood vessel growth) and tumor growth have been discovered that act by disruption of the binding of matrix metalloproteinase 2 (MMP2) to the cell surface integrin $\alpha_v\beta_3$ validating the target for therapeutic intervention for the treatment of cancer and providing the first small molecule lead structures (Fig. 3).[30,31] The first small molecule inhibitors of Myc/Max dimerization have been discovered that act at the bHLHZip dimerization interface enlisting a novel FRET assay that were shown to disrupt the binding of this oncogenic transcription factor to DNA and to inhibit Myc induced conversion of normal to transformed fibroblasts validating an important small molecule target for the treatment of cancer (Fig. 4).[33] Promising inhibitors of LEF-1/β-catenin mediated gene transcription have been identified in a luciferase reporter assay (TOPFLASH).[29] These inhibitors and their mechanism of action (inhibition of LEF-1/β-catenin

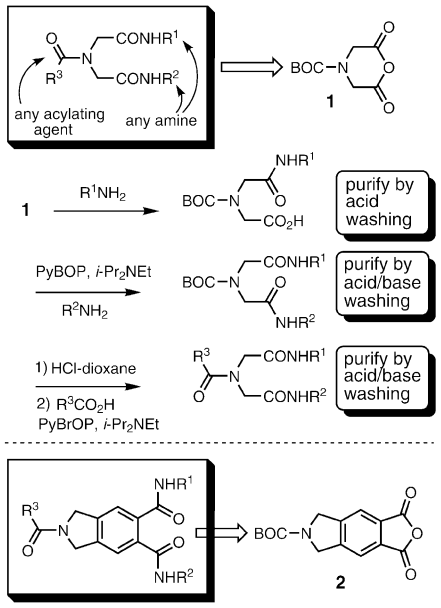

Figure 1. Solution-phase synthesis of libraries.

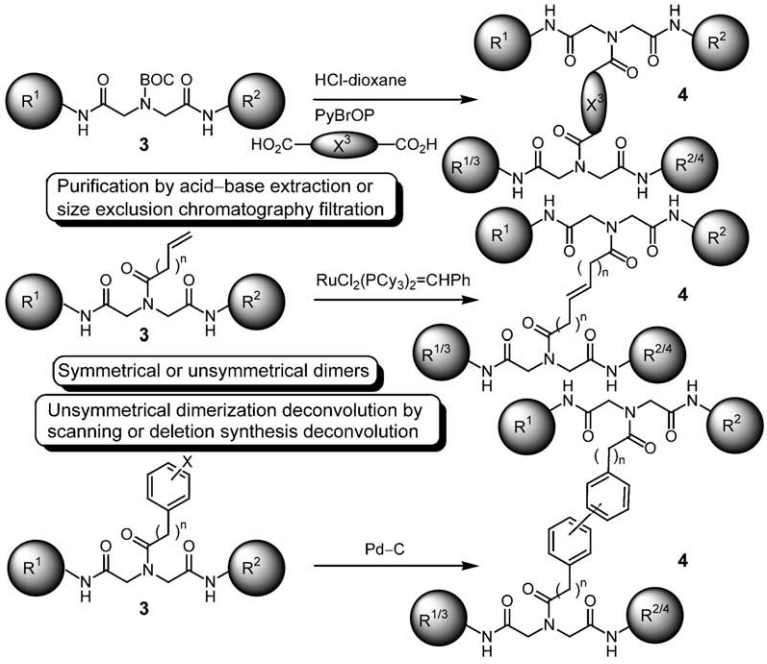

Figure 2. Homo- or heterodimerization of libraries for protein–protein dimerization.

binding) are being more fully explored with the development of a LEF-1/β-catenin FRET assay. The first inhibitors of the binding of the signaling adapter protein Paxillin to the short cytoplasmic tail of the integrin α_4 ($\alpha_4\beta_1$ = VLA4, Very Late Antigen 4) have been discovered and shown to inhibit cell migration validating a new target for the treatment of chronic inflammatory diseases, asthma, and multiple sclerosis (Fig. 5).[34] Finally, the remarkable discovery of relatively small erythropoietin (EPO) agonists that act by promoting homodimerization of the cell surface receptor (Fig. 6) has been made in an approach that could provide conventional therapeutic replacements for the recombinant human protein (7 billion/yr market) used for the treatment of anemias resulting from cancer, AIDS and other clinical disease states.[32]

The targets were carefully chosen not only for their therapeutic importance, but such that each represents a different prototypical extracellular or intracellular signaling event involving protein–protein or protein–DNA interactions:

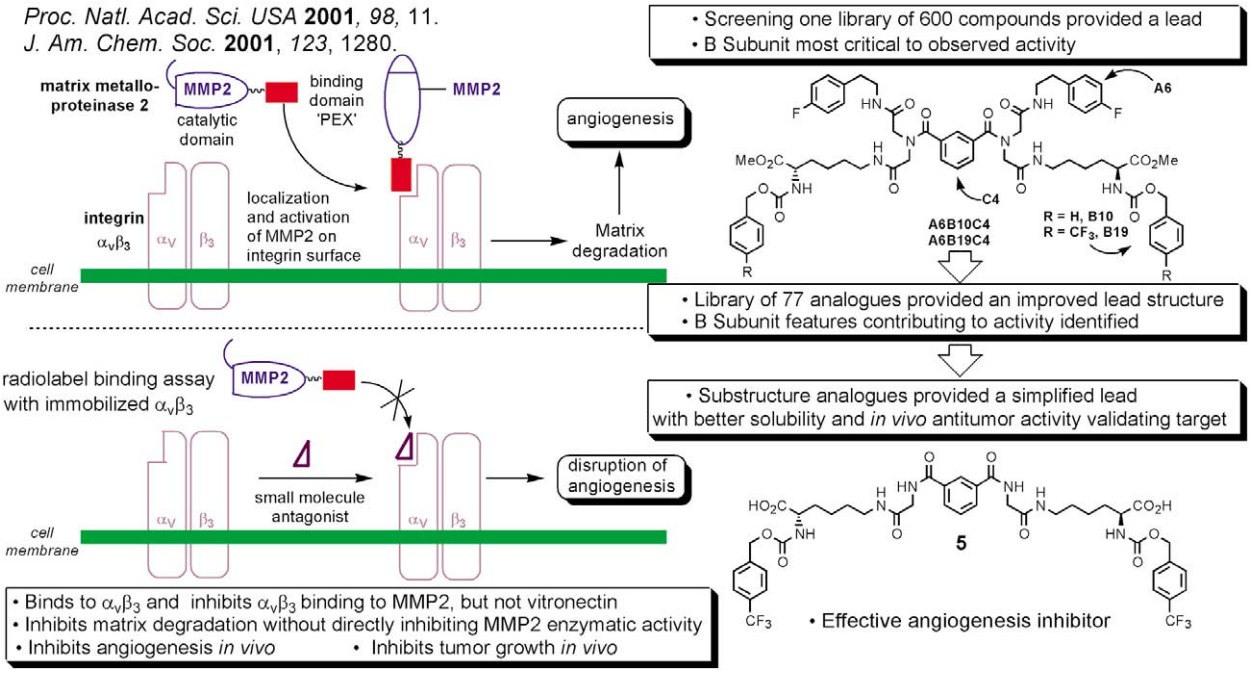

Figure 3. Disruption of angiogenesis by blocking MMP2 binding to integrin $\alpha_v\beta_3$.

- Promote a cell surface receptor homodimerization required for activation (EPO agonist)
- Inhibit an extracellular cell surface integrin–protein interaction (MMP2/$\alpha_v\beta_3$).
- Inhibit an intracellular cell surface integrin–protein interaction (Paxillin/$\alpha_4\beta_1$).
- Inhibit the intracellular heterodimerization of a transcription factor (Myc–Max, LEF-1/β-catenin).
- Inhibit a protein–DNA interaction of a transcription factor (LEF-1/β-catenin).

Protein–DNA Interactions: High-Throughput Synthesis and Screening Applied to the Discovery of Biologically Active DNA Binding Agents

We have also described the high-throughput synthesis and screening of DNA binding compounds[19–28] that are related to our interests in understanding and exploiting their properties. The approach integrates the solution-phase techniques for the synthesis of libraries with a technique we recently introduced for rapid high-throughput screening for DNA binding affinity or sequence selectivity (Fig. 7).[20,23] These techniques can be combined to rapidly explore and define the structural features responsible for the sequence selective DNA binding properties of known agents, to discover new paradigms for small molecule recognition of DNA (new bp codes), and to screen for compounds that selectively target consensus sequences of transcription factors for controlling aberrant gene transcription (i.e, LEF-1/β-catenin).

Highlights of these efforts to date include:

- Development of a high resolution, high throughput FID assay for assessing DNA binding affinity or selectivity.[20,23,25]

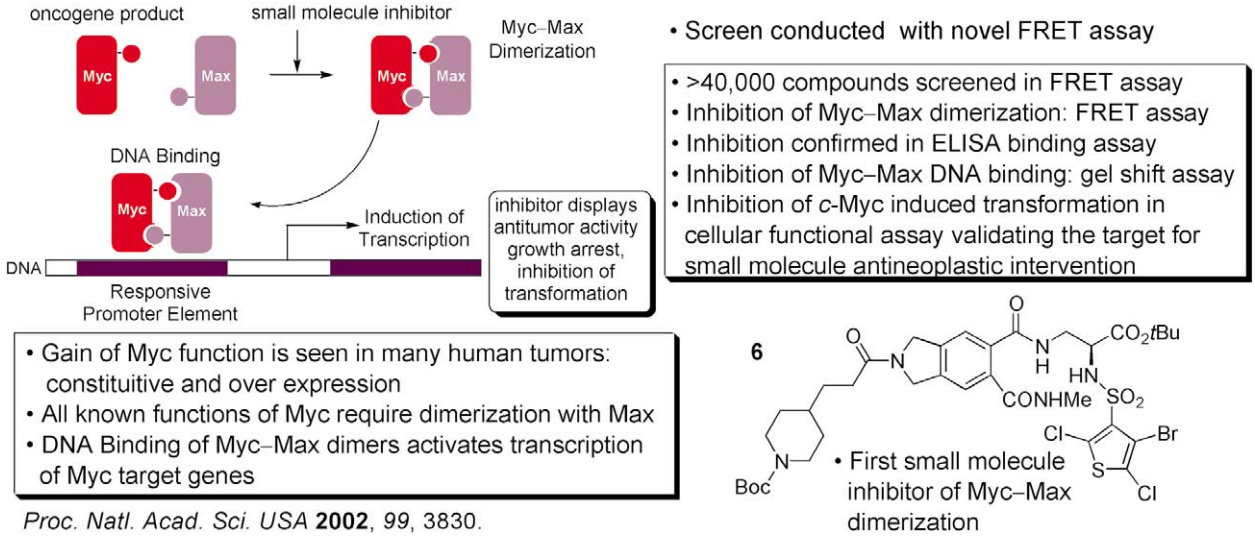

Figure 4. Inhibition of Myc–Max (transcription factor) heterodimerization and abberant gene transcription: antitumor target validation.

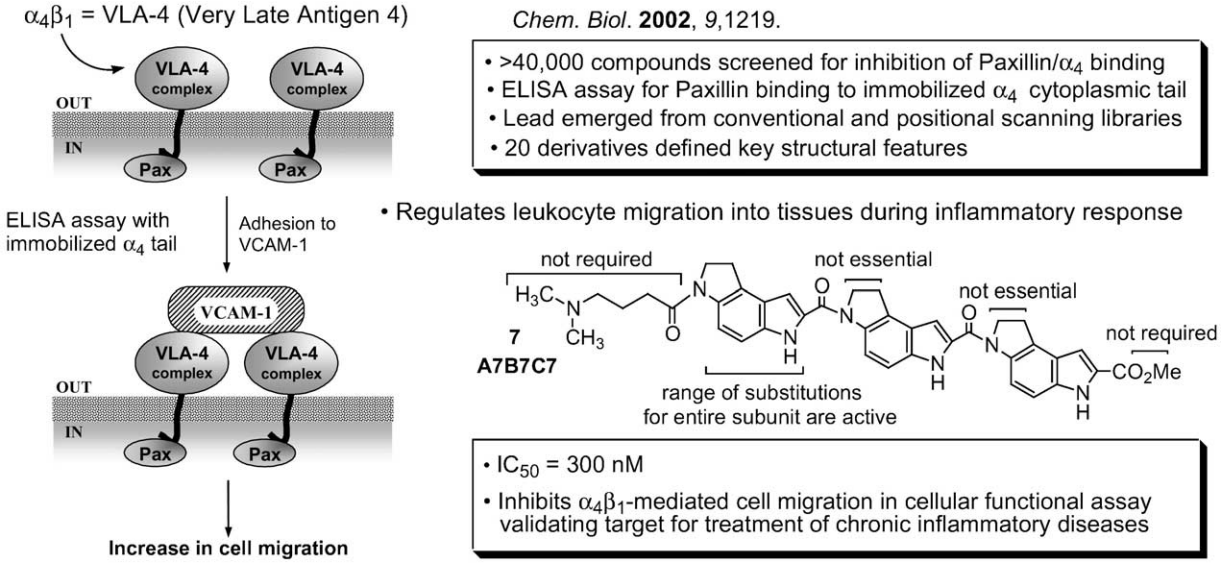

Figure 5. Inhibition of cell migration by blocking intracellular Paxillin/α_4 binding.

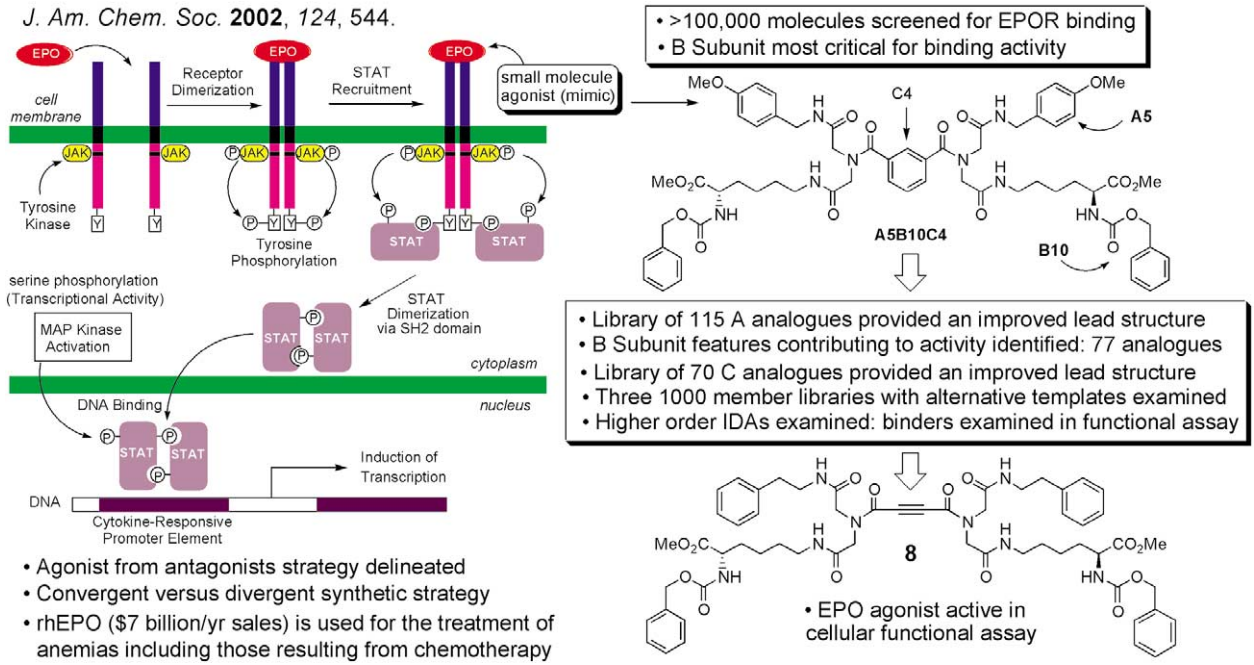

Figure 6. EPO agonists (mimics) that function by promoting EPOR dimerization.

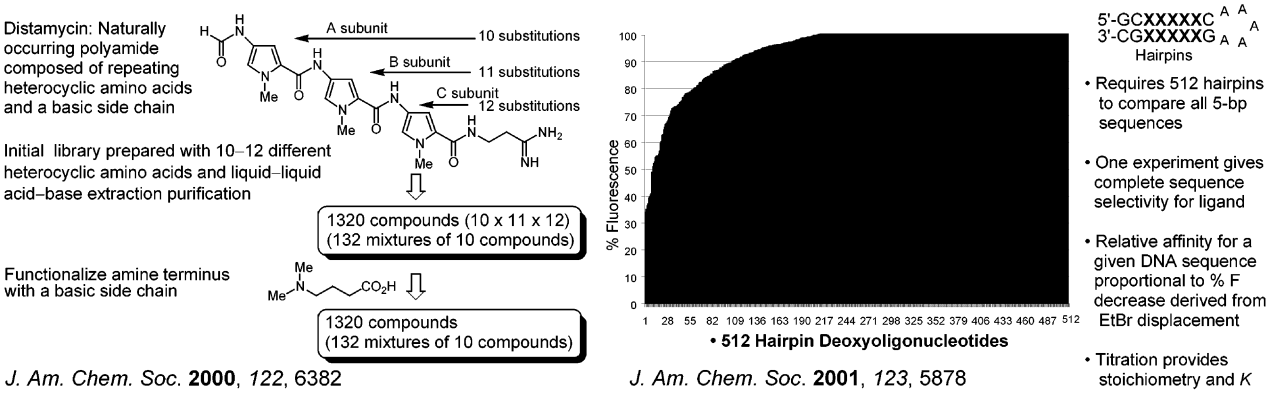

Figure 7. Solution-phase synthesis of DNA minor groove binding ligands and FID assay for DNA binding affinity and selectivity.

- Disclosure of some of the first libraries of DNA binding compounds (distamycin, CC-1065).[20–22,24]
- Discovery of a general hairpin versus extended DNA binding of a substituted β-alanine linked polyamide.[26]
- The first characterization of cooperative extended 2:1 side-by-side parallel (vs antiparallel) DNA binding with a novel class of iminodiacetic acid (IDA) linked polyamides.[27]

Acknowledgements

We gratefully acknowledge the financial support of the National Institutes of Health which has supported various aspects of our work (protein–protein interactions, CA78045; DNA binding compounds, CA41986; cyclic (depsi)peptides, CA41101; angiogenesis inhibitors, EY14174), the Department of Army (DAMD17-02-1-0135), and The Skaggs Institute for Chemical Biology.

References and Notes

1. Boger, D. L.; Tarby, C. M.; Caporale, L. H. Generalized dipeptidomimetic template: solution-phase parallel synthesis of combinatorial libraries. *J. Am. Chem. Soc.* **1996**, *118*, 2109.
2. Cheng, S.; Comer, D. D.; Williams, J. P.; Boger, D. L. Novel solution-phase strategy for the synthesis of chemical libraries containing small organic molecules. *J. Am. Chem. Soc.* **1996**, *118*, 2567.
3. Cheng, S.; Tarby, C. M.; Comer, D. D.; Williams, J. P.; Caporale, L. H.; Boger, D. L. A solution-phase strategy for the synthesis of chemical libraries containing small organic molecules: a universal and dipeptide mimetic template. *Bioorg. Med. Chem.* **1996**, *4*, 727.
4. Review: Tarby, C. M.; Cheng, S.; Boger, D. L. Solution phase strategy for the synthesis of chemical libraries containing small organic molecules: a general depeptide mimetic and a flexible universal template, In *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*; Chaiken, I. M., Janda, K. D., Eds.; ACS: Washington, 1996, pp 81–98.
5. Boger, D. L.; Chai, W.; Ozer, R. S.; Andersson, C.-M. Solution-phase combinatorial synthesis via the olefin metathesis reaction. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 463.

6. Boger, D. L.; Ozer, R. S.; Andersson, C.-M. Generation of symmetrical compound libraries by solution-phase combinatorial chemistry. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1903.
7. Boger, D. L.; Chai, W. Solution-phase combinatorial synthesis: convergent multiplication of diversity via the olefin metathesis reaction. *Tetrahedron* **1998**, *54*, 3955.
8. Boger, D. L.; Chai, W.; Jin, Q. Multistep convergent solution-phase combinatorial synthesis and deletion synthesis deconvolution. *J. Am. Chem. Soc.* **1998**, *120*, 7220.
9. Boger, D. L.; Ducray, P.; Chai, W.; Jiang, W.; Goldberg, J. Higher order iminodiacetic acid libraries for probing protein–protein interactions. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2339.
10. Boger, D. L.; Goldberg, J.; Jiang, W.; Chai, W.; Ducray, P.; Lee, J. K.; Ozer, R. S.; Andersson, C.-M. Higher order iminodiacetic acid libraries for probing protein–protein interactions. *Bioorg. Med. Chem.* **1998**, *6*, 1347.
11. Boger, D. L.; Goldberg, J.; Andersson, C.-M. Solution phase combinatorial synthesis of biaryl libraries employing heterogeneous conditions for catalysis, isolation and size exclusion chromatography for purification. *J. Org. Chem.* **1999**, *64*, 2422.
12. Boger, D. L.; Jiang, W.; Goldberg, J. Convergent solution-phase synthesis of combinatorial libraries through rigid biaryl or diacetylene couplings. *J. Org. Chem.* **1999**, *64*, 7094.
13. Boger, D. L.; Lee, J. K.; Jin, Q. Two comparisons of the performance of positional scanning and deletion synthesis for the identification of active constituents in mixture combinatorial libraries. *J. Org. Chem.* **2000**, *65*, 1467.
14. Boger, D. L., Solution-Phase Synthesis of Combinatorial Libraries: To Bead or Not to Bead; In *Perspective Articles on the Utility and Application of Solution-Phase Combinatorial Chemistry*; Baldino, C. M., Eds.; *Comb. Chem.* **2000**, *2*, 89.
15. Review: Boger, D. L.; Goldberg, J. Multip-Step Solution-Phase Combinatorial Chemistry; In *Combinatorial Chemistry: A Practical Approach*; Fenniri, H. Ed.; Oxford University Press: Oxford, 2000; pp 303–326.
16. Boger, D. L.; Chen, Y.; Foster, C. A. Synthesis and evaluation of aza HUN-7293. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1741.
17. Review: Boger, D. L.; Goldberg, J. Cytokine receptor dimerization and activation: prospects for small molecule agonists. *Bioorg. Med. Chem.* **2001**, *9*, 557.
18. Chen, Y.; Bilban, M.; Foster, C. A.; Boger, D. L. Solution-phase parallel synthesis of a pharmacophore library of HUN-7293 analogues: a general chemical mutagenesis approach to defining structure–function properties of cyclic (depsi)peptides. *J. Am. Chem. Soc.* **2002**, *124*, 5431.
19. Boger, D. L.; Fink, B. E.; Hedrick, M. P. A new class of highly cytotoxic diketopiperazines. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1019.
20. Boger, D. L.; Fink, B. E.; Hedrick, M. P. Total synthesis of distamycin A and 2640 analogs: a solution-phase combinatorial apporach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 6382.
21. Boger, D. L.; Dechantsreiter, M. A.; Ishii, T.; Fink, B. E.; Hedrick, M. P. Assessment of solution phase positional scanning libraries based on distamycin A for the discovery of new DNA binding agents. *Bioorg. Med. Chem.* **2000**, *8*, 2049.
22. Boger, D. L.; Lee, J. K. Development of a solution-phase synthesis of minor groove binding bis-intercalators based on triostin A suitable for combinatorial synthesis. *J. Org. Chem.* **2000**, *65*, 5996.
23. Boger, D. L.; Fink, B. E.; Brunette, S. R.; Tse, W. C.; Hedrick, M. P. A simple, high resolution method for establishing DNA binding affinity and sequenxe selectivity. *J. Am. Chem. Soc.* **2001**, *123*, 5878.
24. Boger, D. L.; Schmitt, H.; Fink, B. E.; Hedrick, M. P. Parallel synthesis and evaluation of 132 (+)-1,2,3,9a-tetrahydrocyclopropa[*c*]benz[*e*]indol-4-one (CBI) analogues of CC-1065 and the duocarmycins defining the contribution of the DNA biding domain. *J. Org. Chem.* **2001**, *66*, 6654.
25. Boger, D. L.; Tse, W. C. Thiazole orange as the fluorescent intercalator in a high resolution FID assay for determining DNA binding affinity and sequence selectivity of small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2511.
26. Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L. Hairpin versus extended DNA binding of a substituted β-alanine linked polyamide. *J. Am. Chem. Soc.* **2002**, *124*, 2148.
27. Woods, C. R.; Ishii, T.; Boger, D. L. Synthesis and DNA binding properties of iminodiactic acid (IDA) linked polyamides: characterization of cooperative extended 2:1 side-by-side parallel binding. *J. Am. Chem. Soc.* **2002**, *124*, 10676.
28. Woods, C. R.; Faucher, N.; Eschgfaller, B.; Bair, K. W.; Boger, D. L. Synthesis and DNA binding properties of saturated distamycin analogues. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647.
29. Boger, D. L.; Goldberg, J.; Satoh, S.; Cohen, S. B.; Vogt, P. K. Non amide-based combinatorial libraries derived from N-BOC iminodiacetic acid: solution-phase synthesis of a 150-membered piperazinone library with activity against LEF-1/β-catenin mediated transcription. *Helv. Chim. Acta* **2000**, *83*, 1825.
30. Silletti, S.; Kessler, T.; Goldberg, J.; Boger, D. L.; Cheresh, D. A. Selective disruption of MMP2 binding to integrin $\alpha_v\beta_3$ by a novel organic molecule inhibitors angiogenesis and tumor growth in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 119.
31. Boger, D. L.; Goldberg, J.; Silletti, S.; Kessler, T.; Cheresh, D. A. Identification of a novel class of small-molecule anti-angiogenic agents through screening of combinatorial libraries which function by inhibiting the binding and localization of proteinase MMP2 to integrin $\alpha_v\beta_3$. *J. Am. Chem. Soc.* **2001**, *123*, 1280.
32. Goldberg, J.; Jin, Q.; Ambroise, Y.; Satoh, S.; Desharnais, J.; Capps, K.; Boger, D. L. Erythropoietin mimetics derived from solution phase combinatorial libraries. *J. Am. Chem. Soc.* **2002**, *124*, 544.
33. Berg, T.; Cohen, S. B.; Desharnais, J.; Sonderegger, C.; Maslyar, D. J.; Goldberg, J.; Boger, D. L.; Vogt, P. K. Small molecule antagonists of Myc/Max dimerization inhibit Myc-indiced transformation of chicken embryo fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3830.
34. Ambroise, Y.; Yuspan, B.; Ginsberg, M. H.; Boger, D. L. Inhibitors of cell migration that inhibit intracellular paxillin/α_4 binding: a well documented use of positional scanning libraries. *Chem. Biol.* **2002**, *9*, 1219.
35. Toogood, P. L. Inhibition of protein–protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, *45*, 1543.
36. Cochran, A. G. Antagonists of protein–protein interactions. *Chem. Biol.* **2000**, *7*, R85.
37. Zutshi, R.; Brickner, M.; Chmielewski, J. Inhibition of the assembly of protein–protein interfaces. *Curr. Opin. Chem. Biol.* **1998**, *2*, 62.
38. Heldin, C.-H. Dimerization of cell surface receptors in signal transduction. *Cell* **1995**, *80*, 213.
39. Lemmon, M. A.; Schlessinger, J. Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biol. Sci.* **1994**, *19*, 459.
40. Stahl, N.; Yancopoulos, G. D. The alphas, betas, and kinases of cytokine receptor complexes. *Cell* **1993**, *74*, 587.
41. Austin, D. J.; Crabtree, G. R.; Schreiber, S. L. Proximity versus allostery: the role of regulated protein dimerization in biology. *Chem. Biol.* **1994**, *1*, 131. Belshaw, P. J.; Ho, S. N.;

Crabtree, G. R.; Schreiber, S. L. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4604. For a recent disclosure, see: Koide, K.; Finkelstein, J. M.; Ball, Z.; Verdine, G. L. A synthetic library of cell permeable molecules. *J. Am. Chem. Soc.* **2001**, *123*, 398, and references cited therein.

42. Ihle, J. N.; Witthuhn, B. A.; Quelle, F. W.; Yamamoto, K.; Thierfelder, W. E.; Kreider, B.; Silvennoinen, O. Signaling by the cytokine receptor superfamily: JAKs and STATs. *TIBS* **1994**, *19*, 222.

43. Lamb, P.; Seidel, H. M.; Stein, R. B.; Rosen, J. The role of JAKs and STATs in transcriptional regulation by cytokines. *Annu. Rep. Med. Chem.* **1996**, *31*, 269.

44. Brivanlou, A. H.; Darnell, J. E., Jr. Signal transduction and the control of gene expression. *Science* **2002**, *295*, 813.

45. Seed, B. Making agonists of antagonists. *Chem. Biol.* **1994**, *1*, 125.