ELSEVIER

monitor

Design and synthesis of a novel DNA-encoded chemical library

Medicinal chemists often search for small molecules that can bind to specific targets. A solution chosen when the target protein displays enzymatic activity, or the researchers have labeled ligands that can be used in displacement assays, is the use of large libraries of chemical compounds that can be screened individually in order to identify novel ligands. These high-throughput screening approaches can be expensive, both in terms of target protein requirements and costs in synthesizing the library, such as in chemical synthesis and the use of robotics. In this regard, the use of DNA fragments as amplifiable 'bar-codes' for the identification of chemical compounds in a library represents an attractive avenue for the synthesis and screening of large combinatorial libraries [1]. There are a number of different strategies that can be conceived for constructing DNA-encoded chemical libraries. One such approach is the use of self-assembling chemical libraries, where each of the two complementary DNA strands carries a different chemical moiety ('dual pharmacophore chemical libraries') [2]. A further example of this approach is the use of individual compounds covalently attached to unique DNA fragments. These 'single pharmacophore chemical libraries' can be generated by a number of synthetic approaches, for example, the oligonucleotide-assisted separation of DNA derivatives followed by chemical reactions [3]. The formation of DNA-encoded chemical libraries is facilitated by the ability to perform chemical reactions at high yields and purities, using synthetic methodologies that work well in water, thus preserving the integrity of DNA. Recent work [4] describes the construction of a DNA-encoded chemical library of 4000 compounds attached to individual double-stranded DNA fragments, carrying unique DNA sequences that allow identification of the displayed chemical moieties. The synthesis of the library was carried out using the sequential steps of reaction of dienes to distinct aminomodified synthetic oligonucleotides, pool and split, Diels–Alder reaction with maleimide derivatives, and encoding of the cycloaddition reaction by hybridization of partially complementary oligonucleotides, followed by Klenow-mediated DNA polymerization. This provided the final compounds in a double-stranded DNA format. In this way, a 4000-member library, based on a two-step split-and-pool strategy, was completed. More specifically, this library comprised twenty 2,4-hexadiene derivatives carrying a carboxylic acid moiety that was coupled to aminotagged oligonucleotides carrying distinctive six base codes. The 20 diene DNA-conjugates were then pooled, split into 200 aliquots and cycloaddition reactions were then performed with 200 maleimides. Finally, the second code was introduced to each DNA-encoded compound by hybridization with partially complementary oligonucleotides and subsequent Klenow polymerization. The 200 reactions were pooled in 20 vials and analyzed to confirm the purity and efficiency of the Klenow encoding procedure. Finally, the library was pooled in a single vial and aliquoted to a final concentration of 100 nM (total DNA content).

This methodology has the advantage that DNA encoding enables the use of individual library members at very low concentrations, analogous to peptide and antibody phage display libraries, in which rare binders can be efficiently recovered and identified [5]. The availability of high-throughput sequencing technologies, such as 454 sequencing [6], allows for profiling of the relative abundance of individual library members before and after a single panning step, thus providing direct information about the relative enrichment rate for all compounds in the library. In summary, this research has provided for the design, synthesis, and characterization of a DNA-encoded chemical library containing 4000 compounds individually coupled to DNA fragments, which serve as identification bar codes. The use of a Diels–Alder strategy with 2,4-hexadiene modified oligonucleotides and a variety of different maleimides gave cycloaddition products in good yields and purities. The library is compatible with decoding strategies based on ultra-high-throughput sequencing technologies (e.g., 454 sequencing, Roche), which enables the relative quantification of individual library members before and after selection against target proteins of interest. Further work in order to apply to other chemical systems, and hence provide small organic molecules that demonstrate protein specificity, is therefore warranted.

Translation of DNA into a 13 000 small-molecule macrocycle library suitable for *in vitro* selection

Molecules with biological activity emerge in living systems through cycles of translation, selection, and amplification with mutation. Scientists have adopted features of biological evolution to create DNA, RNA, and protein molecules with tailor-made binding or catalytic properties. The benefits to the discovery of functional molecules by applying translation and selection-based methods has provided the impetus to develop new approaches to addressing translating DNA sequences into structures not necessarily compatible with polymerase enzymes or ribosomal machinery. For example, a 'DNA display' method in which resin-bound DNA hybridization directs split-and-pool combinatorial synthesis, as well as the use of DNA display to generate libraries of linear

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2010.06.013

peptides and peptoids has been developed [7]. Macrocycles are of interest for the development of biologically active small molecules, partly because of the increased rigidity that can decrease the entropic cost of their binding to biological targets, which results in higher potential binding affinities and specificities than those of corresponding linear compounds, as well as the potential for higher bioavailability, membrane permeability, and resistance to *in vivo* degradation than their linear counterparts [8]. Recent work [9] discloses the DNA-templated synthesis of a large library of synthetic macrocycles suitable for *in vitro* selection. Through these endeavors, a more robust and efficient library synthesis route to these compounds was developed. The synthesis of the DNA-templated synthesis comprising 13 824 macrocycles was completed. The synthetic material thus produced was sufficient for hundreds of *in vitro* selections against biological targets of interest. The authors [9] synthesized the complete library of 13 824 DNA-templated macrocycles using all eight validated scaffolds and 36 amino acid building blocks through three rounds of peptide coupling reactions and split–pool DNA synthesis, followed by cyclization. Since the structural elucidation of any library member surviving selection requires the identification of the starting scaffold as well as the three building blocks, it was necessary to encode the identity of the scaffold in each template sequence. As the scaffold-encoding sequence does not direct any chemical reactions through DNA hybridization, it can be of minimal length, and be placed anywhere within the template between the two PCR primer binding sites. In this work, the scaffold was encoded. This DNA-templated macrocycle library provided a starting point for the discovery of functional synthetic macrocycles through *in vitro* selections against biological targets. Biology-inspired approaches to the discovery of functional synthetic molecules bring with it the efficiency of *in vitro* selection, the sensitivity of DNA amplification, and the ease of DNA sequence analysis with structures that can only be accessed through synthetic organic chemistry. Through this work, experimental validation of the ability of a new codon set to mediate DNA-templated reactions efficiently and in a sequence-specific manner was demonstrated. This library may represent one of the largest collections of medium-sized synthetic macrocycles reported to date. Sufficient material of the completed macrocycle library was generated for many *in vitro* selections against biological targets of interest. Further work is warranted to facilitate future DNA-templated synthesis efforts.

1 Brenner, S. and Lerner, R.A. (1992) Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5381–5383

2 Melkko, S. *et al.* (2004) Encoded self-assembling chemical libraries. *Nat. Biotechnol.* 22, 568–574

3 Halpin, D.R. and Harbury, P.B. (2004) DNA display. II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol.* 2, 1022–1030

4 Buller, F. *et al.* (2008) Design and synthesis of a novel DNA-encoded chemical library using Diels–Alder cycloadditions. *Bioorg. Med. Chem. Lett.* 18, 5926–5931

5 Silacci, M. *et al.* (2005) Design, construction, and characterization of a large synthetic human antibody phage display library. *Proteomics* 5, 2340–2350

6 Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380

7 Wrenn, S.J. *et al.* (2007) Synthetic ligands discovered by in vitro selection. *J. Am. Chem. Soc.* 129, 13137–13143

8 Driggers, E.M. *et al.* (2008) The exploration of macrocycles for drug discovery—an underexploited structural class. *Nat. Rev. Drug Discov.* 7, 608–624

9 Tse, B.N. *et al.* (2008) Translation of DNA into a library of 13 000 synthetic small-molecule macrocycles suitable for in vitro selection. *J. Am. Chem. Soc.* 130, 15611–15626

Paul Edwards
mepauledwards@fsmail.net