ELSEVIER

Lead discovery by DNA-encoded chemical libraries

Samu Melkko[1,2], Christoph E. Dumelin[1], Jörg Scheuermann[1] and Dario Neri[1]

[1] Department of Chemistry and Applied Biosciences, ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland
[2] Philochem AG, c/o ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland

The isolation of specific binding molecules is a central problem in the drug discovery process and might be useful for the elucidation of the biological function of proteins identified in genome and proteome research. Libraries of organic molecules, conjugated covalently to DNA tags that serve as identification bar codes, have been proposed recently as a way to identify ligands of target proteins of choice efficiently. Here, we analyze the different strategies for constructing DNA-encoded chemical libraries, and the potential and challenges of this promising technology.

Introduction

Identification of molecules that bind specifically to target proteins of interest is a formidable challenge. Technologies that facilitate the isolation of binding molecules have profound implications for pharmaceutical research because most drug development programs rely on the ability to isolate small organic compounds that bind to a given protein. With an aging population and an increased understanding of the mechanisms of disease at a molecular level, biomedical scientists are facing the demand for more and better drugs. Additionally, elucidation of the biological function of proteins will, in many cases, require access to specific ligands (an approach that is often termed 'Chemical Genetics' [1]). Techniques for the general, fast, inexpensive isolation of small, organic, binding molecules are lacking at present.

Traditional high-throughput screening (HTS) procedures require the storage and handling of hundreds of thousands of chemical compounds and, typically, they rely on a biochemical assay [2]. Alternatives to HTS that have been introduced recently include dynamic combinatorial chemistry [3,4], small-molecule microarrays [5], fragment-based lead discovery [6,7] and DNA-encoded chemical libraries [8,9]. DNA-encoded chemical libraries are collections of small organic molecules that are conjugated covalently to DNA tags that serve as identification bar codes. In this review we discuss the architectures of the different kinds of DNA-encoded chemical libraries, the synthesis of such libraries, and the opportunities and the challenges of this technology.

Analogies to phage-display technology

Conceptually, DNA-encoded chemical libraries share features with biological and biochemical display technologies such as phage display [10], ribosome display [11] and CIS display [12], all of which include a physical link between a polypeptide and the gene that encodes it. In antibody phage-display technology, phage particles display antibody fragments as a fusion protein to the minor coat protein pIII on the tip of the phage and the phage particle contains a DNA plasmid carrying the gene for the displayed fusion protein (Figure 1a). Using genetic terminology, the antibody fragment corresponds to a 'phenotype' whereas the corresponding gene corresponds to a 'genotype'. Antibody phage libraries contain large collections of antibody fragments with different binding specificities (typically $>10^9$ different clones) [13], with variations introduced into their complementary determining regions [14]. The isolation of specific antibodies to a given target protein relies on affinity selection procedures. Phage particles that bind to the target antigen are captured on immobilized antigen, thus, rescuing the genetic information of the antibodies with the desired binding property. Eluted phage is used to infect bacteria, which can be used either to express and screen the selected antibody fragments or to produce a secondary phage library in which antibodies with the desired binding properties are enriched relative to the starting library. Typically, after 2–4 rounds of selection, binding specificities against a given target

Corresponding author. Melkko, S. (samu.melkko@philochem.ch)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2007.04.007

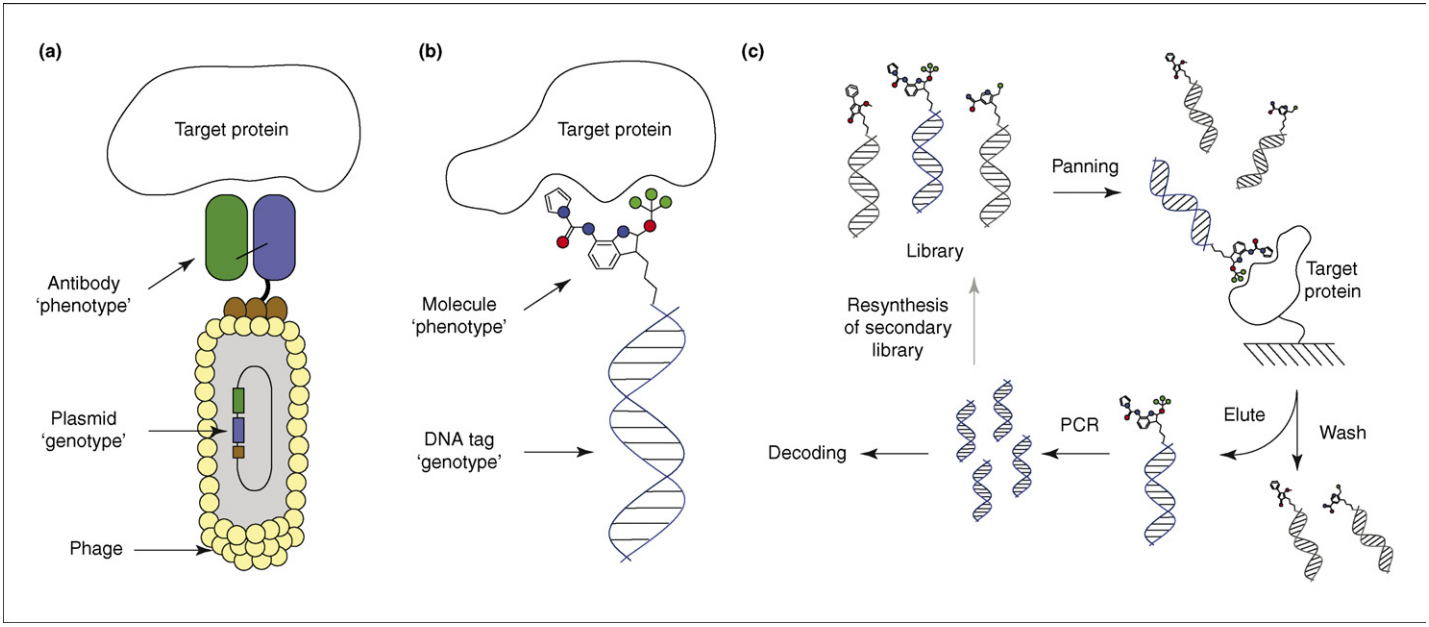

FIGURE 1

Common features of antibody phage display and DNA-encoded chemical libraries. (a) A phage particle displaying a scFv antibody fragment (potentially capable of binding a target protein) as a fusion protein to the minor coat protein pIII on the tip of the phage. The phage particle provides a physical link between the antibody fragment (the 'phenotype') and the gene encoding it ('genotype'). Green and blue colouring indicate heavy and light chain fragments of the scFv antibody and the corresponding genes, respectively. **(b)** A conjugate of an organic molecule (potentially capable of binding a target protein) and a DNA oligonucleotide. The DNA tag ('genotype') serves as an identification bar code, unambiguously encoding the molecule ('phenotype') to which it is attached. **(c)** Affinity selection ('panning') procedure for the isolation of binding compounds from a DNA-encoded chemical library. The library of encoded molecules is incubated with an immobilized target protein of choice. After affinity capture, non-binding library members are removed by washing procedures. After elution of the DNA-tagged binders, the enriched DNA moieties can be amplified by PCR, which enables the detection of very low amounts of template DNA. The subsequent decoding of the enriched DNA tags reveals the chemical identity of the molecules capable of binding to the target protein.

antigen are identified by screening individual antibody clones on a 96-well microtiter plate. Repeated rounds of selection and amplification result in an enrichment of antibodies with the desired binding properties that are present at low frequency in the starting library. Without repeated rounds of selection, identification of desired binders would require the screening a relatively large number of individual antibody clones (typically 10^3–10^5 clones).

DNA-encoded chemical library compounds share some of the characteristics of a phage particle that displays an antibody fragment (Figure 1b). The organic molecule, or putative binding moiety ('phenotype'), is linked physically to a DNA tag that carries the identity code ('genotype') of the molecule it is attached to. The selection procedure for DNA-encoded chemical libraries is also similar to a selection cycle with an antibody phage library (Figure 1c). The library of encoded molecules is incubated with the target protein of choice and, after affinity capture, non-binding library members are separated from binding library members by, for example, coupling the target protein to a solid support and removing the supernatant (that contains the non-binders). The stringency of selection is controlled using suitable washing conditions. After elution of the DNA-tagged binders, the enriched DNA moieties are amplified by polymerase chain reaction (PCR), which allows very low amounts of template DNA to be detected. Subsequent decoding of the enriched DNA uses either DNA sequencing or hybridization to oligonucleotide microarrays, depending on the architecture of the library and its size. In some libraries, the DNA tags not only serve as identification bar codes, but can also guide the synthesis of the displayed molecule. This procedure is reminiscent of the way a gene directs the synthesis of the corresponding polypeptide.

Alternate parallel synthesis of encoded libraries

It is probable that the concept for the construction of a DNA-encoded chemical library was proposed first by Sydney Brenner and Richard Lerner in 1992 [15]. The authors postulated the alternating stepwise synthesis of a polymer (e.g. a peptide) and an oligonucleotide sequence (serving as an identification bar code) on a common linker (e.g. a bead) in split and pool cycles (Figure 2). After affinity capture on a target protein, the population of DNA tags of the selected library members would be amplified by PCR and, in theory, utilized for enrichment of the bound molecules by serial hybridization steps to a subset of the library. In principle, the affinity-capture procedure could be repeated, possibly resulting in a further enrichment of the active library members. Finally, the structures of the chemical entities would be decoded by cloning and sequencing the PCR products. The feasibility of the orthogonal, solid-phase synthesis of peptides and oligonucleotides was demonstrated by attaching a test peptide (the pentapeptide leucine enkephalin) and an encoding DNA tag onto controlled-pore glass beads [16]. The peptide bound to a specific antibody and the corresponding DNA coding tag was amplified by PCR. To date, the only implementation of the technology at a library level has been published by scientists at Affymax, who have constructed a collection of $\sim 10^6$ heptapeptide sequences (based on seven amino acids) and their corresponding coding oligonucleotide tags on beads. The library was incubated with a fluorescently labelled anti-peptide

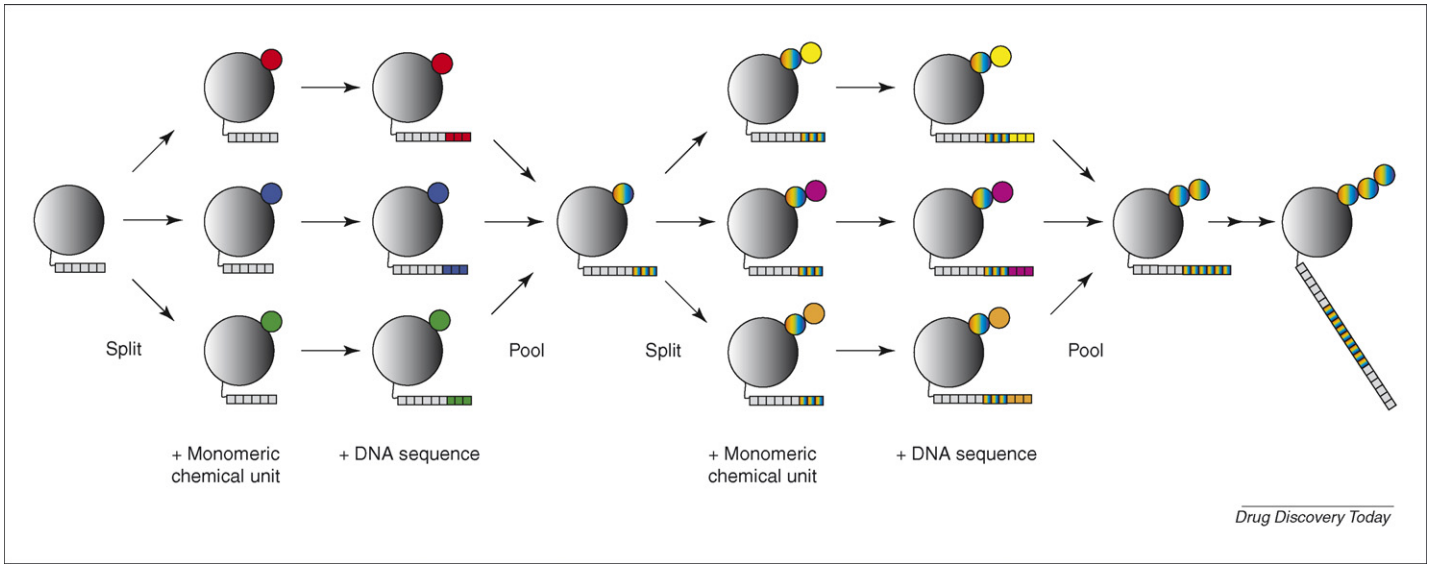

FIGURE 2

Construction of an encoded combinatorial chemical library by alternating parallel combinatorial synthesis as proposed by Brenner and Lerner [15]. The addition of monomeric chemical units (coloured spheres) to a growing polymer chain, alternated with the stepwise orthogonal synthesis of a DNA 'bar-code' (blocks in colour matching that of the sphere) leads to the synthesis of DNA-encoded beads displaying polymers, which can be probed for binding to suitable target proteins of interest. The coding DNA sequence is flanked by constant DNA stretches, enabling the PCR amplification of the DNA tag.

antibody, and binders were selected successfully by fluorescent-assisted cell sorting [17]. In the original paper, Brenner and Lerner suggested that the alternate synthesis of chemical compounds and oligonucleotides could also be implemented in the absence of beads. This approach has been followed by the companies NuEvolution [18,19] and Praecis [20] by using enzyme-catalyzed ligation of coding DNA blocks rather than chemical synthesis of oligonucleotides. These companies have filed patent applications for the technology but, to our knowledge, scientific articles have not been published on this topic.

Library synthesis assisted by sequence-encoded routing

Pehr Harbury and co-workers have developed a variation of split-and-pool library synthesis in which the DNA moiety does not serve merely as a coding tag, but also directs the step-wise library synthesis of polymers [21] (Figure 3a). For library construction, a collection of 340-mer single-stranded oligonucleotides with six, variable, 20-base sequences (that serve as codes) are loaded on a series of ten anticodon columns. After code-specific separation of the oligonucleotide library by hybridization, the anticodon columns are disconnected, the library oligonucleotides transferred to diethylaminoethyl (DEAE) Sepharose columns and the first chemical monomers conjugated in parallel to the library oligonucleotides. After elution of the conjugates from the columns, the eluates are pooled and separated on another series of further anticodon columns, allowing the code-specific addition of the second chemical monomer. The advantage of such a DNA-directed split-and-pool synthesis is the possibility to resynthesize a secondary library after PCR amplification of the DNA tag-enriched library compounds, so allowing several rounds of selection. The DNA-routing strategy has been put into practice by constructing an *N*-acetylated pentapeptide library with a complexity of 10^6 [22]. Acylated leucine enkephalin was included in the library as a positive control, and affinity capture selections were performed against an anti leucine-enkephalin antibody (as in [16]). Two rounds of selection and library resynthesis were performed. Sequencing of ~70 clones from the input and from each round of selection showed a clear enrichment of amino acid sequences that correspond to leucine enkephalin after each selection cycle. Although a polypeptide library has been used in this example, the method is amenable to other chemical monomers provided that suitable coupling chemistries are available for efficient, solid-phase synthesis on DNA [23].

DNA-templated organic synthesis

David Liu and co-workers have developed a strategy for the DNA-templated synthesis of libraries of macrocycles in solution in a single reaction tube [24]. Template oligonucleotides (48-mers) with three variable regions (10-, 11- and 12-bases long) were hybridized with four reactive chemical groups displayed onto oligonucleotides complementary to the first variable region of the template oligonucleotide. After sequence-specific hybridization, the chemical moiety (an amino acid) was transferred to a reactive group on the extremity of the template oligonucleotide. Following cleavage of the amino acid from the parental oligonucleotide (which was removed subsequently), four new reactive oligonucleotide derivates were hybridized to the second variable region of the template oligonucleotide. The corresponding chemical moieties were then transferred to the chemical moieties that had been added in the previous synthesis step. After a third round of sequence-specific addition of four further chemical moieties, macrocyclization was induced (Figure 3b). Reaction yields of 5% are reported after the three-step synthesis. The specificity of the chemical coupling relies on the high local concentration of the reactants that is caused by site-specific hybridization of the reagent oligonucleotides with the template oligonucleotide [25–27]. The resulting DNA-encoded, 64-member macrocycle library was spiked with an oligonucleotide conjugated to phenyl sulphonamide, a nanomolar binder to carbonic anhydrase. After affinity capture to

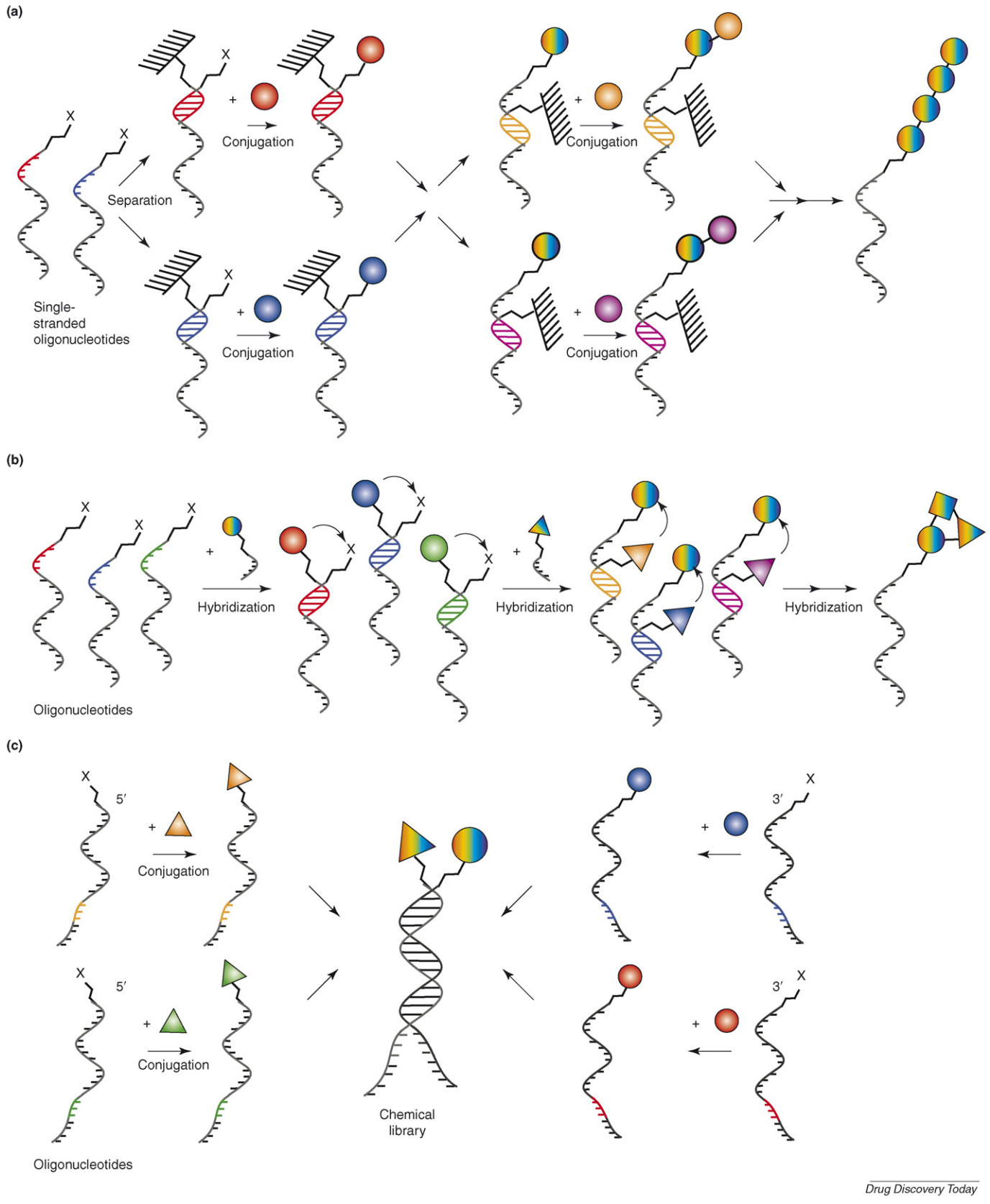

FIGURE 3

Different avenues for the synthesis of DNA-encoded chemical libraries. (a) Sequence-encoded routing enables the solid-phase synthesis of polymers in DNA-directed split and pool cycles. Single-stranded oligonucleotides, which contain variable coding sequences (different coding sequences are indicated by

carbonic anhydrase, the enrichment of the sulphonamide conjugate was demonstrated by performing PCR on the eluate and input, and subjecting the products to restriction cutting and gel analysis. Only the oligonucleotides that are conjugated to sulphonamide contain the recognition site of the restriction enzyme used. As for the DNA-routing approach developed by the Harbury group, DNA-templated organic synthesis allows the possibility of resynthesizing a secondary library after a round of affinity capture. Furthermore, the induction of a high, local concentration of chemical reactants through the hybridization of conjugated oligonucleotides has allowed the discovery of new chemical reactions [28]. The list of reactions that are compatible with DNA-templated synthesis is long and includes S_N2 reactions, reductive aminations, amino acylations, Wittig olefinations and cycloadditions [27]. The technology is being commercialized by the company Ensemble Discovery.

Encoded self-assembling chemical libraries

The constitution of large collections of DNA-encoded molecules through the self-assembly of new molecular entities has been proposed independently by our group [29] and the group of Andrew Hamilton [30]. For example, the self-assembly (heterodimerization) of two libraries containing 1000 members will yield 10^6 combinations (i.e. 10^6 chemical entities). For library construction, we use 48-mer oligonucleotides that contain a variable, six-nucleotide coding region that is flanked by constant DNA sequences, and a chemical modification on an extremity (e.g. a primary amino group) that is suitable for conjugation to organic molecules. The constant DNA sequence adjacent to the site of attachment of the organic molecule lends itself to the hybridization with a second, complementary sublibrary, and both constant sequences are required for PCR amplification of selected codes. The synthesis of library compounds requires a single chemical reaction. Because the sublibraries typically contain a few hundred compounds, the conjugates can be purified individually by HPLC, if required. The two sublibraries can be mixed by hybridization so that every member of one sublibrary forms a stable heterodimer with those of the second sublibrary (Figure 3c). After affinity capture on a target protein of choice, the codes of the enriched compounds are amplified by PCR and either subcloned and sequenced or decoded on microarrays. Successful selections provide direct information about binding fragments that contribute synergistically to the binding of a target protein [31]. As with fragment-based lead discovery, this information must be translated to the design of a ligand molecule by coupling the identified fragments with a suitable chemical linker. The characteristics of the linker (e.g. length, flexibility, geometry, chemical nature and solubility) influence the binding affinity and the chemical properties of the resulting binder (Figure 4a).

Different formats of encoded self-assembling chemical (ESAC) libraries have been used and proposed for the construction of very large libraries (Figure 4b). Our group has isolated several ligands of target proteins such as streptavidin (K_D 2 nM) [32], human serum albumin (K_D 3 μM) [33] and calmodulin (K_D 6 μM) (S. Melkko, C.E. Dumelin, J. Scheuermann and D. Neri, unpublished data) with small, single-pharmacophore libraries (620 library members) (Figure 4bi). Moreover, the affinity maturation of known lead compounds with ESAC libraries (Figure 4bii) has led to the improvement of the binding affinity of ligands to carbonic anhydrase (IC_{50} improved from 1 μM to 25 nM) [29], human serum albumin (K_D improved from 146 μM to 4 μM) [29] and trypsin (IC_{50} improved from 90 μM to 98 nM) [34]. We have constructed a duplex library, as depicted in Figure 4biii, with a complexity of 70000 that is suitable for the *de novo* isolation of binders. Obviously, triplex (Figure 4biv) and quadruplex libraries (Figure 4bv) of oligonucleotide conjugates allow the construction of even larger libraries. Triplexes [35] and quadruplexes of oligonucleotides [36] can be stable because of base stacking and Hoogstein hydrogen-bonding. Notably, Andrew Hamilton's group has generated quadruplex assemblies of oligonucleotide conjugates that orient four organic fragments for the recognition of cytochrome c [37]. However, there are challenges associated with working with triplex and quadruplex ESAC libraries, notably efficient decoding procedures and the site-directed synthesis of multivalent molecules with suitable organic scaffolds.

Discussion and future perspectives

In recent years the field of DNA-encoded chemical libraries has been revitalized by diverse, novel approaches to library construction. However, theoretical considerations apply to all of these libraries. Similar to phage-display libraries, affinity selection with DNA-encoded chemical libraries can be performed in one reaction tube with standard laboratory equipment. The complexity of such libraries can be high. For example, assuming that 10000 copies of any library member are present in a affinity-capture experiment, 1 ml of library compounds (200 nM total concentration) corresponds to a library size of $>10^{10}$, which is orders of magnitude larger than libraries that are accessible with traditional HTS. It is

different colours), are separated on anticodon columns by hybridization. Chemical monomers (coloured spheres) are then conjugated in parallel to the library oligonucleotides [e.g. by 9-fluorenylmethylmethoxycarbonyl (Fmoc)-based peptide synthesis, which is compatible with the presence of unprotected DNA [23]). The identity of a monomer at a given position is determined by the code-specific separation (sequence-encoded routing) of oligonucleotide populations (thus, chemical monomers and corresponding codes are in the same colour). In this particular representation, the end-product is a library of 4-mers. **(b)** Library construction via DNA-templated organic synthesis enables the synthesis of repertoires of macrocycles in single reaction tubes. Template oligonucleotides contain variable coding sequences that can hybridize to reagent oligonucleotides, carrying reactive chemical groups (coloured spheres, triangles and diamonds). The transfer of these reactive chemical groups relies on the spatial proximity mediated by the sequence-specific hybridization of reagent and template oligonucleotides. Chemical monomers and corresponding codes are in the same colour. **(c)** The construction of encoded self-assembling chemical libraries features the parallel conjugation of chemical compounds (coloured spheres and triangles) to oligonucleotides that contain a variable sequence (encoding the organic molecule that the oligonucleotide is attached to) and a constant domain (enabling heterodimerization with complementary oligonucleotide conjugates). Typical conjugation reactions include the coupling of amino-modified oligonucleotides (with a primary amino group either on the 5′ or the 3′ end of the oligonucleotide; typical synthesis scale = 5 nmol) to organic molecules with either a carboxylic acid, an *N*-hydroxysuccinimide ester, or an isothiocyano group (resulting in stable amide bonds or thiourea bonds, respectively). The resulting library can be large, as it originates from the combinatorial self-assembly of smaller sublibraries. In (a–c), rainbow colouring of chemical moieties means that this particular entity is present as a mixture of compounds. The X at the extremity of oligonucleotides represents a reactive chemical group, which might be, for example, a primary amino group.

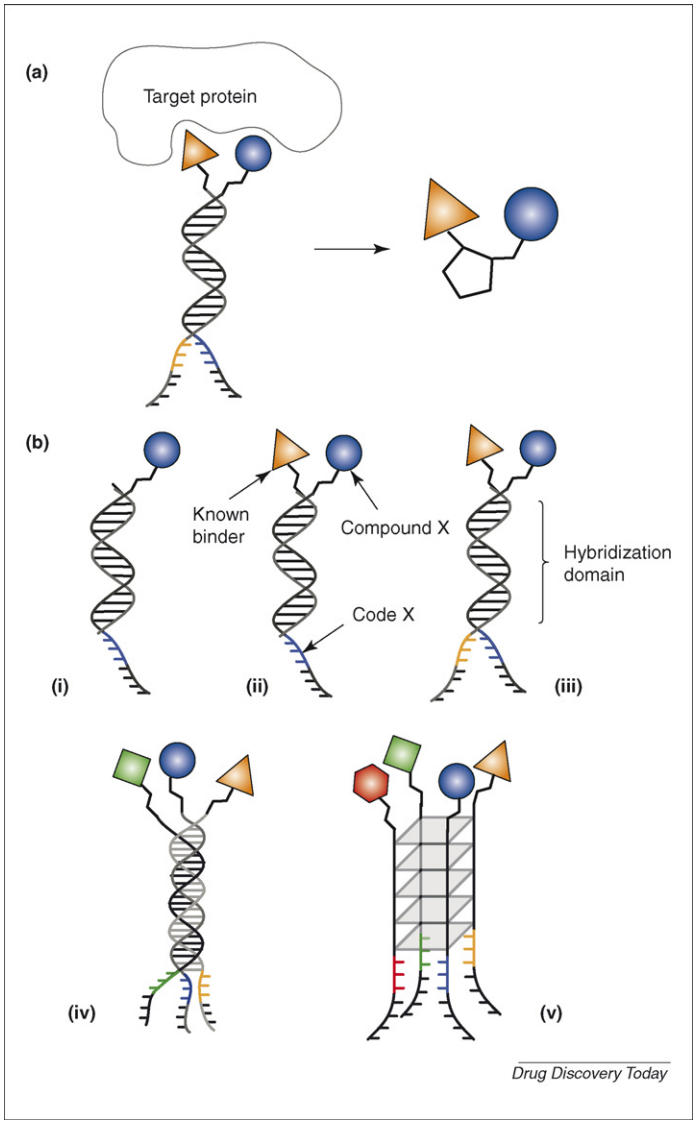

FIGURE 4

Different formats of encoded self-assembling chemical libraries. (a) The aim of affinity selections with ESAC libraries is the identification of molecules that can bind to adjacent epitopes on a given target protein. For most applications, the molecules have to be cross-linked using a suitable chemical linker. **(b)** The different formats of ESAC libraries: (i) Oligonucleotide-compound conjugate libraries can be used in single-pharmacophore selections using relatively small libraries (typically <1000 compounds). (ii) A known binder to a target protein (a lead compound) can be improved (i.e. affinity matured) by conjugating the known binder to a shorter oligonucleotide containing only a hybridization domain, pairing it with a small ESAC library and selecting pharmacophores that might interact with adjacent binding sites on the target protein. The *de novo* isolation of ligands to target proteins of choice can be facilitated by using libraries of large size, which are formed by the self-assembly of oligonucleotide-compound conjugates in either (iii) duplex, (iv) triplex, (v) or quadruplex format. Spheres triangles, diamonds and hexagons represent chemical compounds (typically library compounds) conjugated to DNA oligonucleotides. The colours used for the geometric objects (representing the chemical compounds) are also used to highlight DNA codes that serve as identifiers for the chemical compounds to which the oligonucleotides are attached.

reasonable to assume that, as with phage-display technology, larger library sizes will yield higher-affinity ligands [38] provided that suitable decoding methods are available. Although the concentration of individual library members can be low, PCR amplification allows recovery and identification of the corresponding DNA codes. The concentration of antigen needed for efficient capture must be greater than the desired dissociation constant but (as with phage display) this corresponds to a few micrograms of protein in most cases. Because binders to given target proteins are isolated with affinity capture without an activity assay, it should be possible to work with a wider range of target proteins for which screening assays are not available.

There are numerous challenges regarding the applicability of the different kinds of DNA-encoded chemical libraries. These include the reaction efficiencies for conjugating chemical moieties to DNA oligonucleotides and the resulting yields of library compounds that display the desired chemical entity. A large fraction of unconjugated and/or truncated library members might complicate selection and decoding procedures. The choice of selection conditions is important to discriminate non-binders from binders, and to discriminate ligands with different binding affinities. By analogy to phage-display protocols, we have found that stringency of washing, the addition of competitor molecules and variations in the concentration of target antigen [31] have strong influences on the outcome of selection experiments. For example, in our hands, the ability to discriminate between ligands with different binding affinities (dissociation constants ranging from mM to nM) depends on the coating densities of the target protein on sepharose beads [32]. In addition to library synthesis, decoding the selected library might be the most significant challenge when working with DNA-encoded libraries. The number of codes to be sequenced depends on: (i) the complexity of the library; (ii) the enrichment factor of selected compounds over non-binders in a round of affinity capture; and (iii) whether it is sufficient to observe enrichment of the most favoured codes (with the aim of combining the most favoured moieties) or whether the most favourable combination of codes for the best binder should be detected. In the latter case, one might make the following 'Gedankenexperiment': if one assumes a library complexity of 10^6 and an enrichment factor of 100 of good binders versus non-binders in a round of selection then, statistically, $\sim 10^5$ sequences are required to identify preferential binding compounds with suitable confidence. The number of sequences to be read rises with increasing library size and with lower enrichment factors. The recent development of high-throughput sequencing technologies with the possibility of acquiring up to hundreds of thousands of sequences in a single run [39] might be an attractive methodology for decoding DNA-encoded libraries. At present, however, both 'low-throughput' sequencing and microarray technology allow detection of the enrichment codes that are enriched preferentially in the affinity capture procedures, but not which combination is the most favourable. In practice, the binding properties of the selected moieties must be tested individually after decoding. Another possibility is to perform a step-wise screening approach, for example by identifying a starting binder by varying only one module (e.g. with an approach depicted in Figure 4bi) and by later varying another module while keeping the first moiety isolated constant (e.g. using the approach depicted in Figure 4bii). Furthermore, the

resynthesis of a secondary library after one round of affinity selection might make it possible to limit the number of sequences that must be read to identify enriched binders. However, it is probably misleading to state that multiple rounds of affinity selection and enrichment of active compounds correspond to 'directed evolution' because a necessary feature of Darwinian evolution (i.e. the formation of new species by either mutation or recombination based on the parental population) is missing. By contrast, biochemical display technologies [40] might incorporate features of variation through either inaccuracies in PCR [41] or molecular breeding [42]. The development of truly evolutionary schemes for the directed evolution of organic molecules is a challenge for the future.

Conflict of interest statement

The authors declare competing financial interests. ESAC technology is covered by a patent application, which was licensed from the ETHZ to Philogen, under a share of revenues agreement. D.N. owns shares of Philogen and consults for this company. S.M. has been employed by Philochem, a daughter company of Philogen, since November 2006.

Acknowledgements

We acknowledge financial support from the ETH Zurich, the Bundesamt für Bildung und Wissenschaft for the EU Project 'STROMA', the Swiss National Science Foundation and the EU FP6 'Immuno PDT'.

References

1 Strausberg, R.L. and Schreiber, S.L. (2003) From knowing to controlling: a path from genomics to drugs using small molecule probes. *Science* 300, 294–295
2 Keseru, G.M. and Makara, G.M. (2006) Hit discovery and hit-to-lead approaches. *Drug Discov. Today* 11, 741–748
3 Lehn, J.M. and Eliseev, A.V. (2001) Dynamic combinatorial chemistry. *Science* 291, 2331–2332
4 Otto, S. *et al.* (2002) Dynamic combinatorial chemistry. *Drug Discov. Today* 7, 117–125
5 Uttamchandani, M. *et al.* (2005) Small molecule microarrays: recent advances and applications. *Curr. Opin. Chem. Biol.* 9, 4–13
6 Carr, R.A. *et al.* (2005) Fragment-based lead discovery: leads by design. *Drug Discov. Today* 10, 987–992
7 Erlanson, D.A. (2006) Fragment-based lead discovery: a chemical update. *Curr. Opin. Biotechnol.* 17, 643–652
8 Dumelin, C.E. *et al.* (2006) DNA-encoded chemical libraries. *QSAR Comb. Sci.* 25, 1081–1087
9 Scheuermann, J. *et al.* (2006) DNA-encoded chemical libraries. *J. Biotechnol.* 126, 568–581
10 Winter, G. *et al.* (1994) Making antibodies by phage display technology. *Annu. Rev. Immunol.* 12, 433–455
11 Schaffitzel, C. *et al.* (1999) Ribosome display: an *in vitro* method for selection and evolution of antibodies from libraries. *J. Immunol. Methods* 231, 119–135
12 Odegrip, R. *et al.* (2004) CIS display: *In vitro* selection of peptides from libraries of protein-DNA complexes. *Proc. Natl. Acad. Sci. U. S. A.* 101, 2806–2810
13 Hoogenboom, H.R. (2005) Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.* 23, 1105–1116
14 Silacci, M. *et al.* (2005) Design, construction, and characterization of a large synthetic human antibody phage display library. *Proteomics* 5, 2340–2350
15 Brenner, S. and Lerner, R.A. (1992) Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5381–5383
16 Nielsen, J. *et al.* (1993) Synthetic methods for the implementation of encoded combinatorial chemistry. *J. Am. Chem. Soc.* 115, 9812–9813
17 Needels, M.C. *et al.* (1993) Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc. Natl. Acad. Sci. U. S. A.* 90, 10700–10704
18 Franch, T. *et al.* (2006) Ligational encoding of small molecules. US2006246450
19 Pedersen, H. *et al.* (2002) Templated molecules and methods for using such molecules. WO02103008
20 Morgan, B. *et al.* (2005) Methods for synthesis of encoded libraries. WO2005058479
21 Halpin, D.R. and Harbury, P.B. (2004) DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol.* 2, E173
22 Halpin, D.R. and Harbury, P.B. (2004) DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol.* 2, E174
23 Halpin, D.R. *et al.* (2004) DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biol.* 2, E175
24 Gartner, Z.J. *et al.* (2004) DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 305, 1601–1605
25 Gartner, Z.J. *et al.* (2002) Multistep small-molecule synthesis programmed by DNA templates. *J. Am. Chem. Soc.* 124, 10304–10306
26 Sakurai, K. *et al.* (2005) DNA-templated functional group transformations enable sequence-programmed synthesis using small-molecule reagents. *J. Am. Chem. Soc.* 127, 1660–1661
27 Li, X. and Liu, D.R. (2004) DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew. Chem. Int. Ed. Engl.* 43, 4848–4870
28 Kanan, M.W. *et al.* (2004) Reaction discovery enabled by DNA-templated synthesis and *in vitro* selection. *Nature* 431, 545–549
29 Melkko, S. *et al.* (2004) Encoded self-assembling chemical libraries. *Nat. Biotechnol.* 22, 568–574
30 Sprinz, K.I. *et al.* (2005) Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments. *Bioorg. Med. Chem. Lett.* 15, 3908–3911
31 Melkko, S. *et al.* (2006) On the magnitude of the chelate effect for the recognition of proteins by pharmacophores scaffolded by self-assembling oligonucleotides. *Chem. Biol.* 13, 225–231
32 Dumelin, C.E. *et al.* (2006) Selection of streptavidin binders from a DNA-encoded chemical library. *Bioconjug. Chem.* 17, 366–370
33 Dumelin, C.E. and Neri, D. (2006) Binding molecules and uses thereof. UK-0621973,7
34 Melkko, S. *et al.* Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. *Angew. Chem. Int. Ed. Engl.* (in press)
35 Aich, P. *et al.* (1998) Thermodynamic and kinetic studies of the formation of triple helices between purine-rich deoxyribo-oligonucleotides and the promoter region of the human c-src proto-oncogene. *Nucleic Acids Res.* 26, 4173–4177
36 Burge, S. *et al.* (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* 34, 5402–5415
37 Tagore, D.M. *et al.* (2007) Protein recognition and denaturation by self-assembling fragments on a DNA quadruplex scaffold. *Angew. Chem. Int. Ed. Engl.* 46, 223–225
38 Griffiths, A.D. *et al.* (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* 13, 3245–3260
39 Margulies, M. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380
40 Amstutz, P. *et al.* (2001) *In vitro* display technologies: novel developments and applications. *Curr. Opin. Biotechnol.* 12, 400–405
41 Zaccolo, M. and Gherardi, E. (1999) The effect of high-frequency random mutagenesis on *in vitro* protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* 285, 775–783
42 Minshull, J. and Stemmer, W.P. (1999) Protein evolution by molecular breeding. *Curr. Opin. Chem. Biol.* 3, 284–290