Chemistry & Biology
Article

Discovery of TNF Inhibitors from a DNA-Encoded Chemical Library based on Diels-Alder Cycloaddition

Fabian Buller,[1] Yixin Zhang,[1] Jörg Scheuermann,[1] Juliane Schäfer,[2,3] Peter Bühlmann,[2] and Dario Neri[1,*]
[1]Institute of Pharmaceutical Sciences, Department of Chemistry and Applied Biosciences, ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland
[2]Seminar for Statistics, Department of Mathematics, ETH Zurich, Rämistrasse 101, CH-8092 Zurich, Switzerland
[3]Basel Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, CH-4031 Basel, Switzerland
*Correspondence: neri@pharma.ethz.ch
DOI 10.1016/j.chembiol.2009.09.011

SUMMARY

DNA-encoded chemical libraries are promising tools for the discovery of ligands toward protein targets of pharmaceutical relevance. DNA-encoded small molecules can be enriched in affinity-based selections and their unique DNA "barcode" allows the amplification and identification by high-throughput sequencing. We describe selection experiments using a DNA-encoded 4000-compound library generated by Diels-Alder cycloadditions. High-throughput sequencing enabled the identification and relative quantification of library members before and after selection. Sequence enrichment profiles corresponding to the "bar-coded" library members were validated by affinity measurements of single compounds. We were able to affinity mature trypsin inhibitors and identify a series of albumin binders for the conjugation of pharmaceuticals. Furthermore, we discovered a ligand for the antiapoptotic Bcl-xL protein and a class of tumor necrosis factor (TNF) binders that completely inhibited TNF-mediated killing of L-M fibroblasts in vitro.

INTRODUCTION

The de novo identification of small organic molecules capable of specific binding to protein targets is a central problem in chemistry, biology, and medicine. Conventional ligand discovery methods in the pharmaceutical industry typically rely either on high-throughput robotic screening procedures (with thousands to millions of compounds being assayed separately), or on the chemical improvement of candidate molecules with certain biological activities (Hajduk and Greer, 2007; Pellecchia et al., 2008; Rees et al., 2004; Schreiber, 2000). Both approaches can be expensive in terms of library-associated costs, logistics, and availability of target protein in sufficient amounts. Furthermore, one typically needs to develop a suitable assay (e.g., enzymatic activity, ligand displacement) for screening purposes.

Affinity-based selection methods represent a novel avenue for the discovery of binding molecules from large libraries of chemical compounds (Makara and Athanasopoulos, 2005). Molecular entities capable of selective recognition of a target protein of interest are typically isolated from a mixture of chemical compounds by an affinity-capture procedure, followed by a physical separation of the bound and unbound fractions. For example, mixtures of compounds can be incubated with a target protein in solution and preferential binders can be coeluted with the protein in gel-filtration procedures and identified by liquid-chromatography tandem mass spectrometry (MS) (Annis et al., 2004; Muckenschnabel et al., 2004). This procedure is increasingly being used in drug discovery programs (Zehender and Mayr, 2007), but mixture size may be limited by MS detection and compound solubility (Annis et al., 2004; Muckenschnabel et al., 2004).

In principle, alternatives to MS identification of binding molecules can be considered. For example, tagging individual chemical compounds with unique DNA fragments serving as amplifiable identification "bar-codes" may allow the facile screening of very large libraries of small organic molecules (Brenner and Lerner, 1992; Clark et al., 2009; Gartner et al., 2004; Halpin and Harbury, 2004a; Mannocci et al., 2008; Melkko et al., 2007a, 2004; Scheuermann et al., 2006). The screening of DNA-encoded chemical libraries does not require the set-up of a protein-specific functional assay (because molecules are selected only on the basis of their ability to bind to an immobilized protein) and can be performed when the amount of target protein is limited (less than a milligram of protein per selection) (Melkko et al., 2006). The technology is compatible with the use of low concentrations of individual molecules, whose identity can be revealed using polymerase chain reaction (PCR) procedures.

DNA-encoded chemical libraries can be synthesized in a combinatorial fashion, because the DNA fragment itself can facilitate library construction via different strategies. In self-assembling chemical libraries (Melkko et al., 2004), each of two complementary DNA strands is conjugated to a small molecule, and the dual-molecule display allows the discovery of two ligands in close proximity, in analogy to fragment-based approaches (Blundell et al., 2002; Shuker et al., 1996). Other strategies for library design include DNA template-assisted chemistry, using assisting complementary oligonucleotides (Li and Liu, 2004) or DNA three-way junctions to direct two chemical reactants (Hansen et al., 2009). Alternatively, the solid-phase oligonucleotide-based separation of DNA conjugates and subsequent chemical

reactions (Halpin and Harbury, 2004a, 2004b; Halpin et al., 2004), or the alternated stepwise assembly of chemical moieties and corresponding DNA codes (Buller et al., 2008; Mannocci et al., 2008), can be employed for DNA-encoded chemical library synthesis.

DNA-encoded library construction relies on methodologies to synthesize DNA-conjugates at high yields and with good purities while preserving the integrity of the oligonucleotides. Until now many libraries are constructed using amide bond-forming reactions (Halpin and Harbury, 2004a; Mannocci et al., 2008; Melkko et al., 2004), yet other DNA-compatible reactions could be considered. We recently described the synthesis and characterization of a DNA-encoded chemical library, consisting of 4000 compounds generated by Diels-Alder cycloaddition reactions, each covalently attached to unique DNA fragments (Buller et al., 2008). In this article, we present the results of affinity-based selections using this 4000-member DNA-encoded chemical library against different protein targets. DNA fragments in the protein bound fraction were PCR amplified and the resulting amplicon mixtures were decoded with high-throughput sequencing technology (''454 technology'') (Margulies et al., 2005) to assess the relative abundance of library members before and after selection against a target protein of interest.

In this article, we describe the statistical evaluation of the sequencing data and subsequent hit validation by enzymatic assays or affinity measurements. Using high-throughput sequencing of the Diels-Alder library we identified ligands with nanomolar to micromolar affinities, toward streptavidin, human serum albumin, trypsin, Bcl-xL, and TNF, demonstrating the applicability of DNA-encoded chemical libraries for the de novo discovery of small molecules against a given protein target. The discovery of a novel class of TNF binders capable of complete inhibition of TNF-mediated cell killing in vitro suggests that it should be possible to drug this proinflammatory cytokine with small molecules. At the clinical level, TNF blockade has so far only been accomplished with therapeutic proteins (Lin et al., 2008), providing benefit to patients affected by chronic inflammatory conditions and generating yearly sales of more than 10 billion dollars. To our knowledge, although micromolar disruptors of TNF integrity have previously been reported (He et al., 2005), this is the first time that a TNF ligand was shown to completely inhibit TNF-mediated cell killing in vitro.

RESULTS

Library Synthesis, Affinity-Based Selections, and Decoding

Synthesis and characterization of the 4000-member DNA-encoded chemical library generated by Diels-Alder cycloaddition reactions has been described in our previous publication (Buller et al., 2008). In short, the synthesis was based on a two-step split-and-pool strategy, with an initial coupling of twenty 2,4-hexadiene derivatives of carboxylic acids via amide bond formation to 42-mer 5′ amino-modified oligonucleotides carrying distinctive six-base codes (Figure 1A). As depicted in Figure 1A, we performed cycloaddition reactions with a pool of 20 diene DNA-conjugates and 200 maleimides (generated from the corresponding amines). After hybridization of a 44-mer oligonucleotide fragment containing an eight-base code for the second building block and a subsequent fill-in reaction mediated by Klenow DNA polymerase, reactions were pooled to yield a 4000-member library (Figure 1A), which can be stably stored frozen at −20°C and used for selections.

For selections, the DNA-encoded chemical library was incubated with the target protein immobilized on sepharose resin, or with empty resin (tris[hydroxymethyl]aminomethane [Tris] modified sepharose). Streptavidin, human serum albumin, bovine trypsin, human Bcl-xL, and a murine TNF fusion protein (EDB-TNF) were used as target antigens. Selections were performed at a total DNA concentration of 10 nM, corresponding to 2.5 pM concentration of individual library members. After repetitive washing cycles binding library members were captured on the resin as shown in Figure 1B. After selections, DNA-conjugates retained on the resin were PCR amplified with the primers 5′-GCC TCC CTC GCG CCA TCA GGG AGC TTG TGA ATT CTG G-3′ and 5′-GCC TTG CCA GCC CGC TCA GGT AGT CGG ATC CGA CCA C-3′, consisting of a 18 bp complementary region to the library oligonucleotide tags and 19 bp overhang introducing the adaptor sites A and B, respectively, which are essential for emulsion PCR and high-throughput sequencing using the 454 technology. The general structure of oligonucleotides in the amplicon mixture is depicted in Figure 1C. High-throughput sequencing yielded up to 63,000 sequences for each selection experiment, which were subsequently analyzed by counting the frequency of each combination of the six-base code 1 and eight-base code 2 and normalization to 50,000 sequences (see Figure 1C, Experimental Procedures).

High-Throughput Sequencing Enrichment Profiles

High-throughput sequencing results of the 4000 compound library can be visualized in three-dimensional matrices, with 20 × 200 code combinations in the xy plane and sequence counts for individual library codes on the z axis (Figure 2). As expected, high-throughput sequencing of the naive library PCR product before selection showed no significant enrichment of specific codes. All sequence frequencies were lower than 42 counts, indicating a comparable abundance of the corresponding oligonucleotide-compound conjugates (Figure 2A; for a statistical analysis of the results, see below). To verify library quality, a test selection against streptavidin was carried out. The sequence evaluation of the selection results indicated a strong enrichment of compounds with code 2 = **200** (Figure 2B), which corresponds to a D-desthiobiotin moiety as second building block. This chemical structure is known to bind to streptavidin in the nanomolar range (Table 1) (Dumelin et al., 2006; Hirsch et al., 2002; Mannocci et al., 2008). The selection against immobilized trypsin showed a striking enrichment of many compounds containing code 2 = **6**, which share the benzamidine moiety (Figure 2C, Table 1). The panning experiment of the library against the Tris-quenched resin (see above) led to the enrichment of conjugates with code 2 of **120**, **123**, **136**, **175** (see Figure S1 available online). These compounds are structurally related and were omitted for ligand identification.

Panning of the library against human serum albumin resulted in several hits as indicated by the enrichment profile depicted in Figure 2D. Several enriched compounds were chosen for resynthesis and affinity determination, including structures with code 2 = **199**, encoding a 4-(p-iodophenyl)butyl moiety, previously discovered by DNA-encoded chemical library

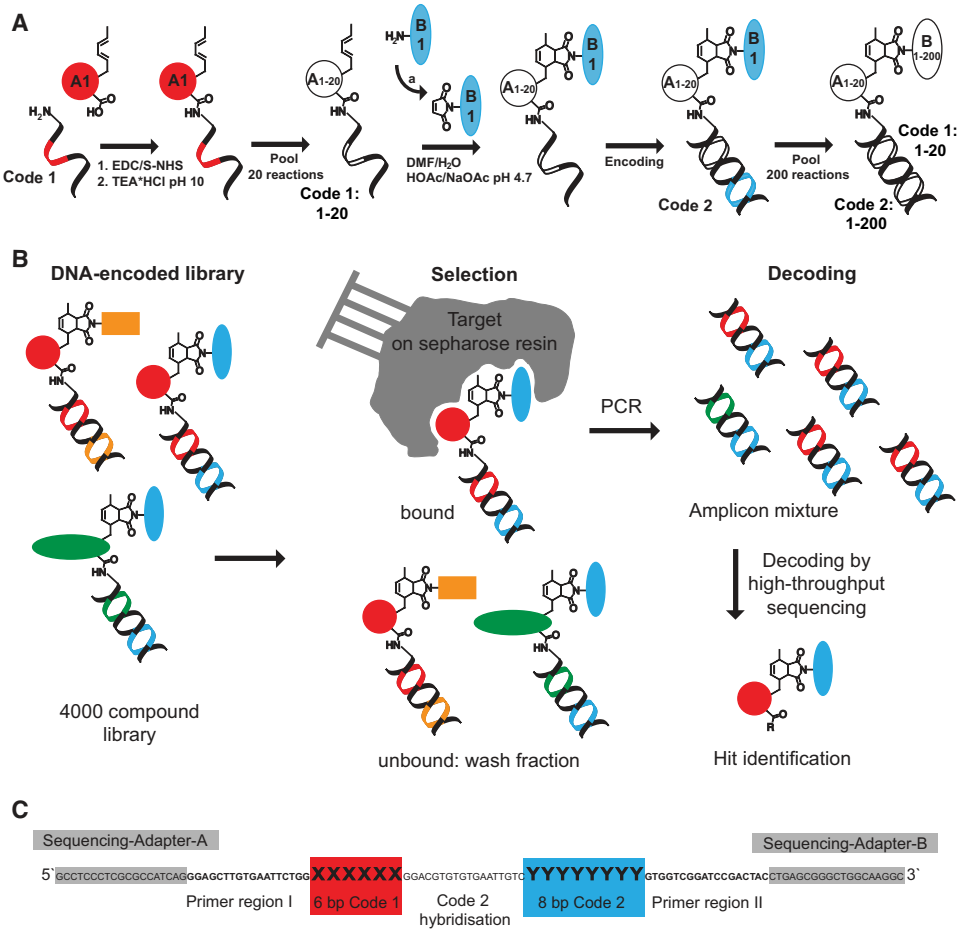


Figure 1. Ligand Discovery by DNA-Encoded Chemical Libraries

(A) Synthesis scheme of a 4000-member DNA-encoded chemical library using Diels-Alder cycloadditions. After coupling of twenty 2,4-hexadiene derivatives (A 1-20) via amide bond formation to 42-mer amino-tagged oligonucleotides carrying code 1, the oligonucleotide conjugates were pooled. Cycloaddition reactions were performed with 200 maleimides (B 1-200), which were derived from a condensation reaction of the corresponding amines and maleic anhydride (a). The second building block B was encoded through hybridization and fill-in reaction mediated by Klenow DNA polymerase with an oligonucleotide fragment carrying code 2. Finally, reactions were pooled yielding a 4000-member library.

(B) Selection for binders using a DNA-encoded chemical library. The 4000-compound library was panned against a target protein immobilized on sepharose resin. Nonbinding conjugates were washed away before PCR amplification of the retained DNA conjugates. The resulting amplicon mixture was decoded using high-throughput sequencing, allowing the relative quantification of all library members.

(C) General oligonucleotide sequence used for high-throughput sequencing.

technology and displaying submicromolar affinity toward human serum albumin (Figure 2D, Table 1) (Dumelin et al., 2008). The selection experiment against Bcl-xL revealed a few codes with frequencies of more than 40 counts, two of which were chosen for hit validation (Figure 2E, Table 1). A selection against human matrix metalloproteinase 3 however showed no significant enrichment of any sequence (Figure S1), indicating that no binding structures were present in the DNA-encoded 4000-compound library. In the selection against EDB-TNF, we chose the two most enriched DNA conjugates (**10-171**, 45 counts and **5-69**, 41 counts) for hit validation (Figure 2F).

Statistical Analysis of High-Throughput Sequencing Data

To our knowledge, a statistical analysis of the library decoding by high-throughput sequencing has so far not been described in the literature. In order to identify significantly enriched ligands after selection, we compared the distribution of sequence counts for members of a DNA-encoded chemical library in the absence of selection pressure (PCR amplification of the library before selection) with the sequence profiles in presence of selection pressure (e.g., after capture on an immobilized target protein).

High-throughput sequencing is based on random sampling of beads resulting from emulsion PCR. Therefore, we consider the negative binomial (NB) distribution as a good model for the analysis of the sequence counts, because this distribution describes an experiment consisting of a series of independent events. Specifically, sequence counts Y_i ($i = 1,\ldots,4000$) for the individual compounds in the library, obtained prior to selection experiments, are assumed to be independent and identically distributed (iid) samples from the NB distribution with mean and

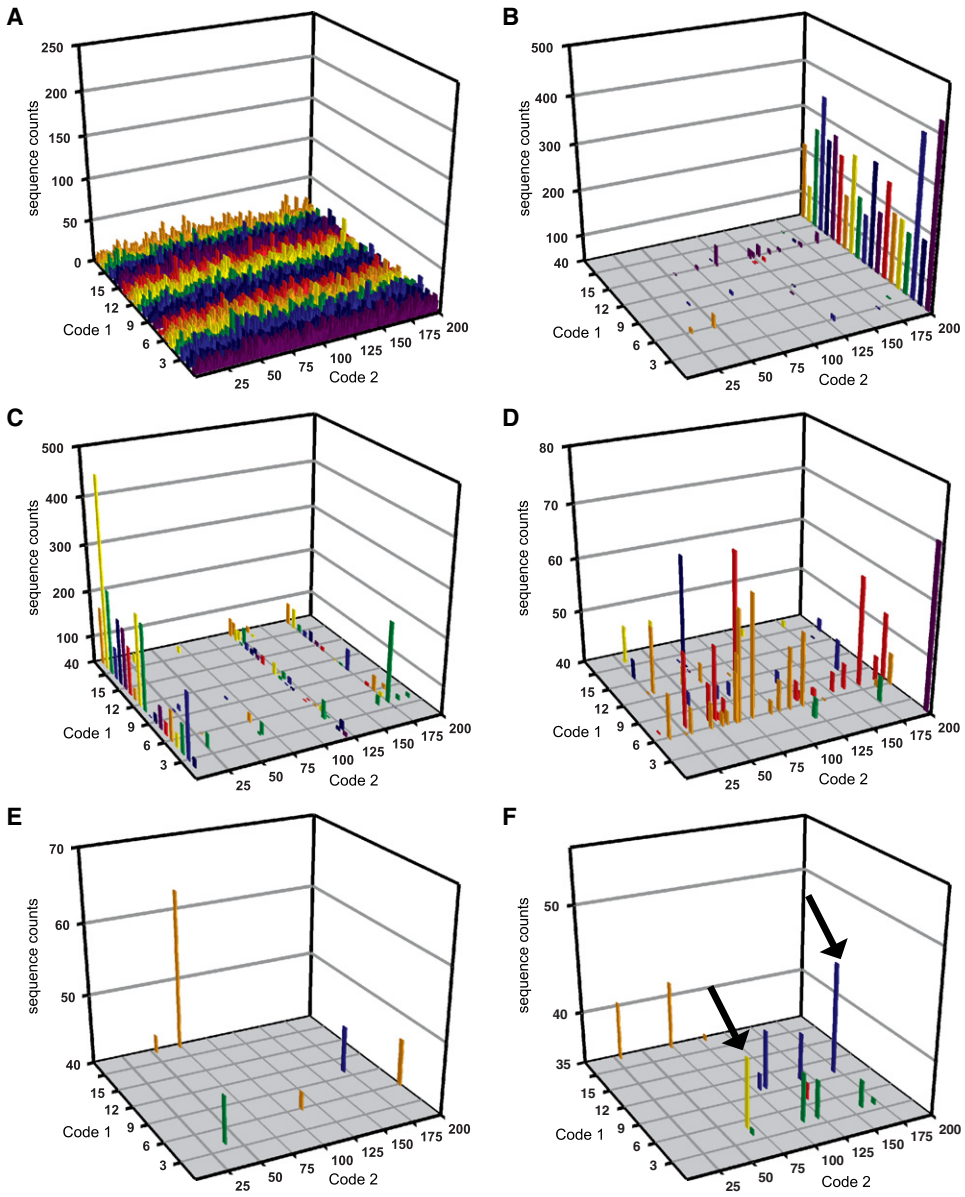


Figure 2. Sequence Enrichment Profiles after High-Throughput Sequencing of the DNA-Encoded Chemical Library

All 20 × 200 code combinations are plotted in the xy plane while the sequence counts for individual library codes are displayed on the z axis.

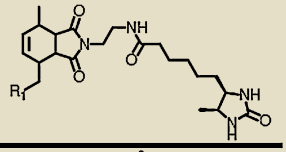
(A) Sequencing of the naive library (before selection).

(B–F) Selection experiments with the 4000-compound library against streptavidin (B), trypsin (C), human serum albumin (D), Bcl-xL (E), and EDB-TNF (F). Selections against Bcl-xL and EDB-TNF showed an enrichment of the background binders with code 2 equal to 120, 123, 136, 175, which are omitted from the sequence enrichment profile for clarity (see Figure S1). The two arrows indicate TNF binders that were selected for further chemical optimization procedures.

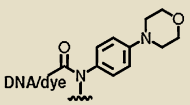
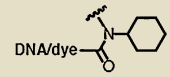
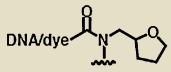
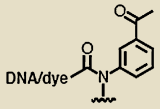
overdispersion parameters $\theta = (m,k)$. Note that the parameter k represents the amount of overdispersion in Y relative to the popular Poisson distribution, and when $k \rightarrow 0$ the distribution of Y becomes Poisson. We employed the method of maximum likelihood in order to fit an NB model to the observed sequence counts. Figure 3A displays the high-throughput sequencing of the DNA-encoded chemical library before selection. All 4000 data points are shown by the histogram, whereas the solid curve corresponds to the fitted NB probability density function with mean 12.445 and overdispersion parameter 5.628. Similarly, Figure 3B shows a quantile-quantile plot of the sample quantiles versus the theoretical quantiles, indicating that the NB distribution fits the data very well.

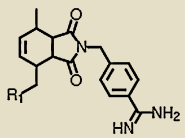
Enrichment of DNA codes in the affinity-based selections was addressed by calculating a one-sided p value for each of the 4000 library codes, using as a null hypothesis the NB probability density function based on the library PCR before selection (i.e., the appropriate null hypothesis if sequence counts corresponding to DNA-conjugates come from the equimolar mixture). A false discovery rate multiplicity correction, level $q = 0.1$, was applied

Table 1. Summary of Selection Results and Ligand Binding

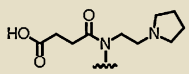
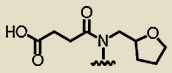
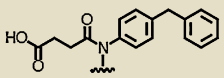
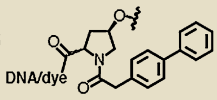
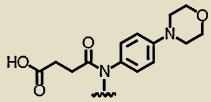


streptavidin

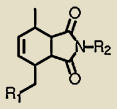
Compound	R_1	Kd[a] [nM]	counts[b]	p[c]
1-200		185	426	<0.001
5-200		278	191	<0.001
3-200		284	388	<0.001
9-200		422	274	<0.001



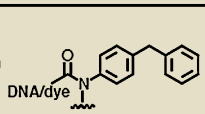
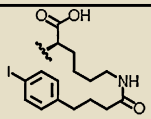
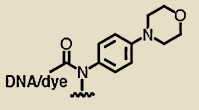
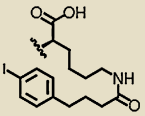
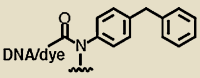
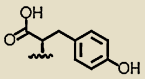
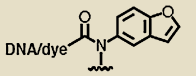
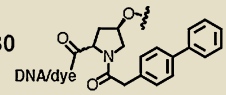
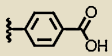
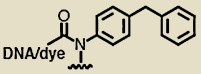
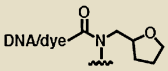
trypsin

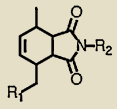
Compound	R_1	Ki[d] [μM]	counts[b]	p[c]
12-6		11	235	<0.001
3-6		20	181	<0.001
6-6		30	105	<0.001
19-6		49	425	<0.001
1-6		220	34	n.s.

[a] Measured by fluorescence polarization.
[b] Sequence counts of the DNA-conjugate after selection.
[c] adjusted p-value for the Benjamini & Hochberg (1995) FDR controlling procedure
[d] Measured by trypsin-inhibition of the carboxylic acid.

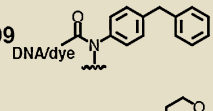
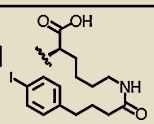
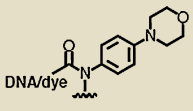
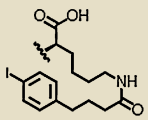
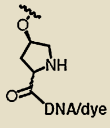
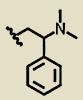


human serum albumin

Compound	R_1	R_2	Kd[a] [μM]	counts[b]	p[c]
6-199			0.44	46	0.039
1-199			0.79	70	<0.001
6-79			12	63	0.002
7-70			17	69	<0.001
19-30			19	47	0.034
6-86			155	29	n.s.
3-74			n.a.	12	n.s.



Bcl-xL

Compound	R_1	R_2	Kd[a] [μM]	counts[b]	p[c]
6-199			10	47	0.025
1-199			78	10	n.s.
20-74			n.a.	63	0.001

to the 4000 resulting p values in order to control the expected proportion of erroneously rejected null hypotheses (Benjamini and Hochberg, 1995). This yielded for each selection a list of significant candidate ligands, from which compounds were chosen for further hit validation. As an example, Figures 3C and 3D show the analysis of the high-throughput sequencing of the trypsin selection. The decoding of the mixture exhibits considerable deviations from the null distribution with notionally equal abundance of all 4000 library members, indicating that enriched DNA conjugates are present after selection.

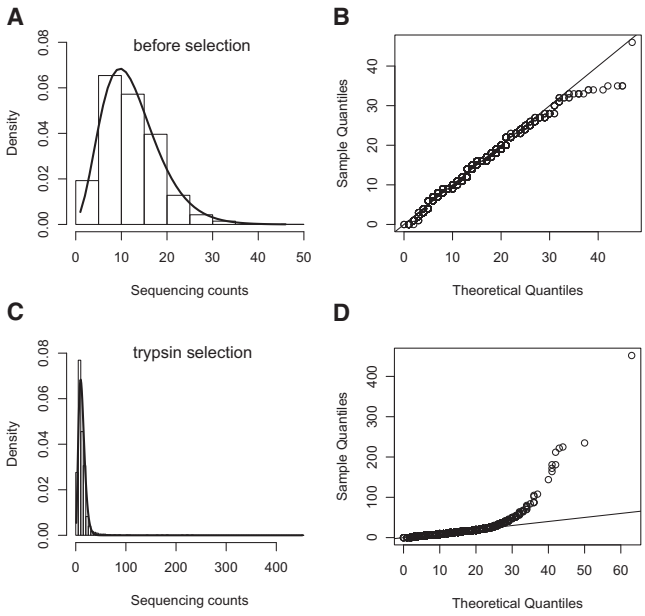


Figure 3. Statistical Analysis of the High-Throughput Sequencing Decoding

(A) Histogram of the sequence counts obtained from high-throughput sequencing of the DNA-encoded library before selection. The solid curve is the fitted NB probability density function.

(B) Associated quantile-quantile plot of sample versus theoretical NB quantiles, indicating good model fit.

(C) Histogram of the sequence counts after affinity-based selection against trypsin. The solid curve is the same fitted NB probability density function as in (A), i.e., prior to imposing selection pressure.

(D) Quantile-quantile plot of sample versus theoretical NB quantiles, where deviations from the diagonal reveal interesting candidate compounds.

Hit Validation by Affinity and Inhibition Measurements

To evaluate the affinity of preferentially enriched compound-DNA conjugates, we resynthesized individual compounds as triethylene glycol-fluorescein-conjugates and measured the dissociation constants by fluorescence polarization at 25°C.

Streptavidin ligands showed dissociation constants between 185 nM and 422 nM, whereas the corresponding DNA conjugates were enriched with up to 426 sequence counts, resulting in adjusted p values lower than 0.001 for all 20 conjugates carrying a d-desthiobiotin moiety (Figure 4A). Structures, dissociation constants and sequence counts are given in Table 1.

Compounds selected against human serum albumin showed dissociation constants as low as 440 nM (compound **6-199**) and the selection process also allowed to identify a series of novel binders in the 10 to 20 μM range, such as compound **6-79** with a K_d of 12 μM (Figure 4B). For comparison, the nonenriched compound **6-86** and **3-74** showed no specific binding to human serum albumin, reflecting the difference between binders and nonbinders in the selection and decoding process (Table 1).

Affinity-based selections against Bcl-xL lead to the enrichment of several DNA conjugates (Figure 2E). Polarization measurements with the fluorescein conjugate of compound **6-199** revealed a K_d of 10 μM, whereas the nonsignificant analog **1-199** was interacting with the protein in the range of 100 μM and compound **20-74** was nonbinding (Figure 4C, Table 1). The affinity cut-off in this selection procedure is therefore likely to be in the lower micromolar range. For comparison, the **BH3I-1** compound described previously binds to Bcl-xL with a K_d of 4.4 μM (Figure S3) (Degterev et al., 2001).

The trypsin selection revealed a pronounced enrichment of compounds containing the benzamidine moiety that were resynthesized and measured in an inhibition assay using Z-GGR-AMC as substrate (Figure 4D, Supplemental Data). This led to the identification of compound **12-6** with a K_i of 11 μM and 235 sequence counts, whereas compound **1-6** with 34 sequence counts showed a K_i of 220 μM (Table 1).

Two significantly enriched DNA conjugates in the TNF selection were synthesized as fluorescein conjugates revealing a K_d of 20 μM for compound **10-171** and a K_d of more than 50 μM for compound **5-69**, as measured by fluorescence polarization against the trimeric EDB-TNF protein (Figure 4E). The fusion protein EDB-TNF was chosen because of its good expression yields in HEK293 cells, high stability, ease of purification, and biological activity. For comparison, a series of nonenriched library analogs was synthesized, showing weaker or no binding in this assay (Table 2, Figure 4E, Supplemental Data). The binders were TNF specific, because no binding of the ligands toward recombinant EDB was observed (Figure S2). We prepared analogs **1a-b** and **2a-c**, based on the dichloro-benzophenone moiety, in order to increase affinity, reduce stereochemical complexity and introduce more hydrophilic residues, leading to compounds **1b** (K_d = 10 μM) and **2c** (K_d = 15 μM; Table 2, Figure 4E and Figure S2). As depicted in Figure 4F, the free carboxylic acid analog of this compound showed a complete inhibition of recombinant murine TNF induced apoptosis in murine L-M fibroblasts cells (Halin et al., 2003) at concentrations higher than 300 μM, whereas nonrelated compounds (N1-N5) showed no inhibition of TNF-mediated killing. Importantly, the compound did not exhibit any toxicity toward L-M fibroblasts up to 500 μM concentrations (Figure 4F).

DISCUSSION

We have synthesized and characterized a high-quality DNA-encoded chemical library containing 4000 members generated via Diels-Alder cycloaddition (Buller et al., 2008). High-yield reactions were a key factor in the construction of the library, underlined by the analytical profiles previously reported by the group on the purity of library members (Buller et al., 2008) and by the statistical analysis of sequence counts to display a homogenous distribution of different library members. Affinity-based selections using the library revealed that DNA-encoded compounds specific to certain protein targets, with binding affinities lower than 20 μM, could be enriched using affinity capture methodologies. High-sequencing counts (above 40 counts) led to the direct identification of ligands, with a very low false-positive rate (only compound **20-74** was later identified to be a false-positive in the Bcl-xL selection). Even though factors of solid-phase mediated selections, e.g., the influence of matrix material and protein density, might impede an exact correlation between sequence counts and affinity, the technology allows a clear differentiation between nonbinders and ligands. Thus, we could discover novel ligands toward human-serum albumin, trypsin, Bcl-xL, and TNF, which were then characterized by fluorescence polarization

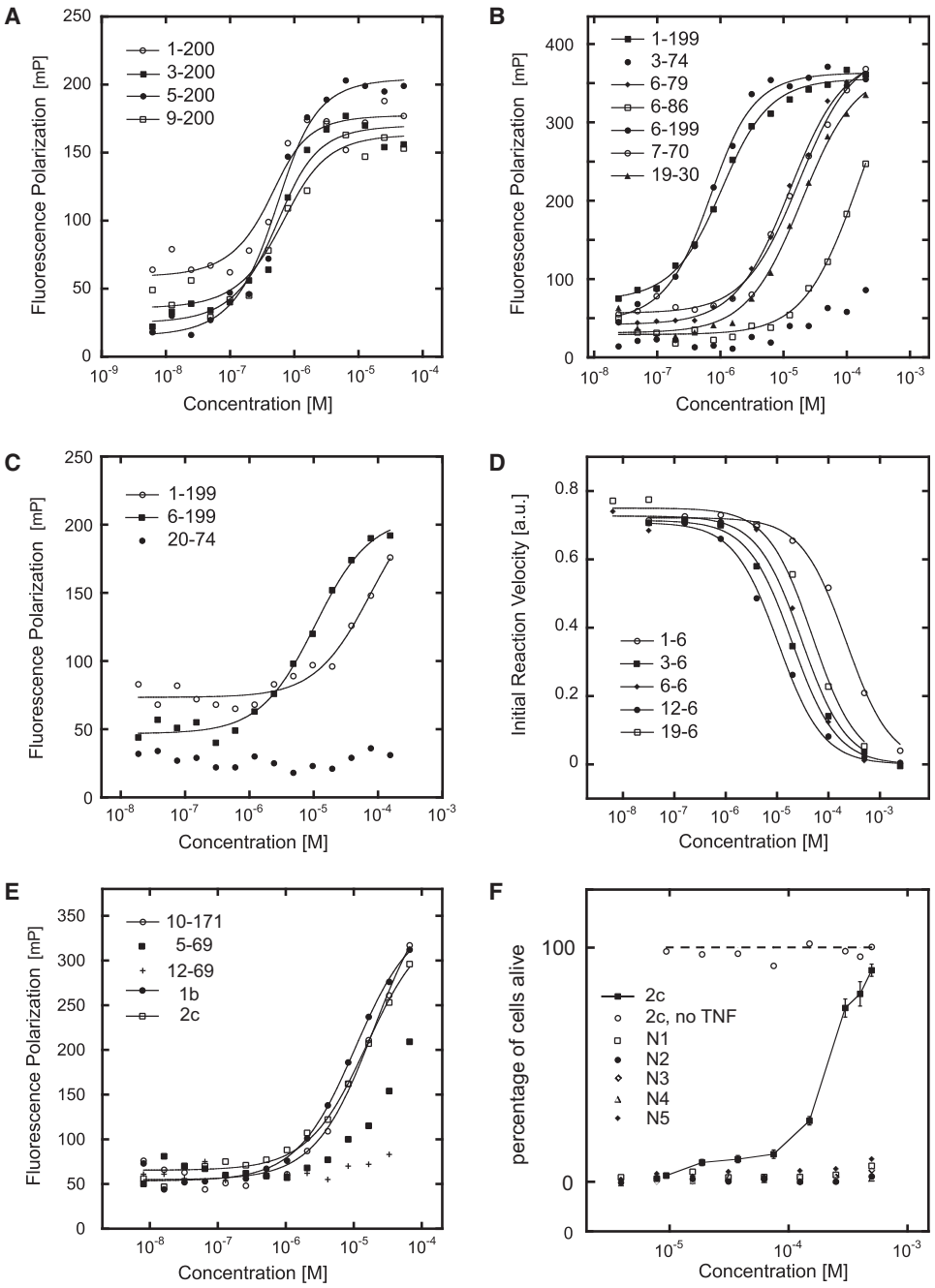
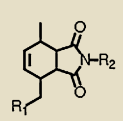


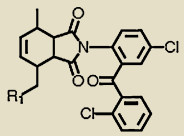
Figure 4. Binding Analysis by Fluorescence Polarization and Inhibition Measurements

Dissociation constants of selected compounds were determined by fluorescence polarization (A–C). Individual compounds identified in the high-throughput sequencing decoding were synthesized as fluorescein conjugates and incubated with different concentrations of the target proteins streptavidin (A), human serum albumin (B), and Bcl-xL (C). Binding molecules with a K_d as low as 184 nM to streptavidin (A), 440 nM to human serum albumin (B), and 10 μM to Bcl-xL (C) were identified. Trypsin inhibition (D) was measured by incubating 2 nM bovine pancreatic trypsin with varying concentrations of the identified ligands without oligonucleotide tag. Inhibitory constants of enriched compounds ranged from 11 μM (compound **12-6**) to 49 μM (compound **19-6**), whereas the nonenriched analog **1-6** exhibited a K_i of only 220 μM. TNF ligands showed a K_d as low as 10 μM, measured by fluorescence polarization against the trimeric EDB-TNF protein (E). Biological activity of **2c** (in the free carboxylic acid form, i.e., without fluorescein tag) was confirmed in a TNF cytotoxicity assay on murine L-M fibroblasts (F), providing complete inhibition of TNF-mediated cell killing (standard error of the mean is shown). The same agent did not show any toxicity to the cells when used in the absence of TNF (empty circles). By contrast, nonrelated compounds used as negative controls did not exhibit any inhibition of TNF-mediated killing (nonrelated compounds: naphthylacetic acid (N1), 2-(3-bromophenyl)acetic acid (N2), 3-(4-Methylpiperazin-1-yl)propan-1-amine (N3), 3-(phenylsulfonyl)propanoic acid (N4), 4-Morpholinoaniline (N5); dashed line: DMSO control with no TNF present). Structures and affinity constants can be found in Tables 1 and 2.

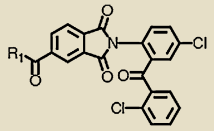
Table 2. TNF Selection Results and Ligand Binding



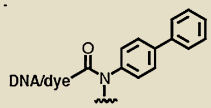
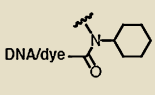
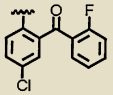
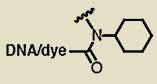
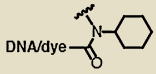
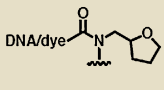
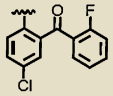
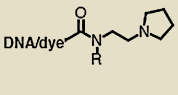
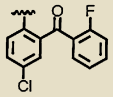
library compounds

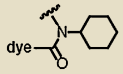
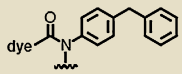
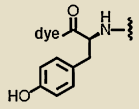
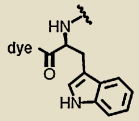


(1a-b)



(2a-c)

DNA-Code	R_1	R_2	Kd[a] [μM]	counts[b]	p[c]
10-171			20	45	0.047
5-69			>50	41	0.065
5-171			>70	15	n.s.
5-110			n.a.	21	n.s.
3-69			n.a.	14	n.s.
12-69			n.a.	16	n.s.

Compound	R_1	Kd[a] [μM]
1a		50
1b		10
2a		>50
2b		>70
2c		15

[a] Measured by fluorescence polarization.
[b] Sequence counts of the DNA-conjugate after selection.
[c] adjusted *p*-value for the Benjamini & Hochberg (1995) FDR controlling procedure

assays and, if appropriate, chemically optimized for inhibition assays.

In this article, we have shown that a NB distribution could be used to adequately describe sequence count profiles of the DNA-encoded chemical library before selection. By contrast, sequence counts after selection could be matched against the null hypothesis of no enrichment, thus offering a statistical basis for the identification of potential protein-specific ligands. As sequencing power of novel high-throughput methodologies increases (Pettersson et al., 2009), it might become possible to oversample larger libraries before and after selection, thus allowing similar statistical analyses of decoding results.

In contrast to phage display and ribosome display selection methodologies, in which multiple rounds of panning and amplification are needed for the identification of preferential binding molecules (Silacci et al., 2005; Winter et al., 1994; Zahnd et al., 2007), sequence-based enrichment profiles of DNA-encoded chemical libraries allow the identification of preferential binders without the need to reamplify/resynthesize the DNA-encoded chemical library.

Because the Diels-Alder DNA-encoded chemical library consists of two sets of independent building blocks, it could be expected that both chemical moieties in a given compound positively or negatively contribute to the overall binding performance. This is clearly exemplified in the analysis of the series of trypsin binders, which contain a benzamidine moiety. A series of simple benzamidine derivatives was previously shown to display K_i values to trypsin ranging between 11 and 220 μM (for example, a K_i = 90 μM was reported for 4-(aminomethyl)benzamidine) (Melkko et al., 2007b). We observed K_i values ranging between 11 and 49 μM (for preferentially enriched compounds; Figure 4D), in contrast to a K_i = 220 μM value for compound **1-6**, which was not preferentially enriched in the selection. We have also previously shown that micromolar benzamidine-based trypsin binders could be subjected to affinity maturation procedures using DNA encoded chemical libraries (Melkko et al., 2007b). In addition to the de novo identification of molecules endowed with biological activity, DNA-encoded chemistry might represent an efficient avenue for the isolation of building blocks that can be used in fragment-based drug discovery programs (Hajduk and Greer, 2007; Pellecchia et al., 2008; Rees et al., 2004).

Although for trypsin and streptavidin the best binders contained moieties, previously described by our group and others, that were known to contribute to specific protein recognition, novel molecular entities were identified against three proteins

of pharmaceutical interest: human serum albumin, Bcl-xL, and TNF. Albumin represents the most abundant protein in human plasma (45 mg ml^{-1}) and features a long circulatory half-life thanks to its size above the renal filtration threshold and its unique ability to interact with the neonatal FcRn receptor (Andersen et al., 2006). Albumin binding pharmaceuticals (Dennis et al., 2002; Nguyen et al., 2006; Syed et al., 1997) make it possible to adjust the exposure of the body to adequate concentrations of the therapeutic agents by increasing their serum half-life. Indeed, certain albumin binders, such as derivatives of 4-iodophenylbutyric acid, could lead to the improvement of pharmacokinetic profiles of therapeutic agents and of contrast agents for imaging applications (Dumelin et al., 2008). Certain molecules identified from the selection with the 4000-membered library, such as compound **6-79** (Kd 12 μM) or **7-70** (Kd 17 μM) (Figure 4B and Table 1), corresponded to novel chemical structures, which have so far not been associated with albumin binding.

Bcl-xL is a promising target in cancer therapy, because it is known to signal for cell survival upon heterodimerization with other proteins of the Bcl-2 family (Diaz et al., 1997; Marzo and Naval, 2008; Vander Heiden and Thompson, 1999). In principle, small-molecule inhibitors can disrupt Bcl-xL interactions with its binding partners, thus altering the proapoptotic balance within the cell. Recently, novel small-molecule ligands capable of inhibiting Bcl-xL protein-protein interactions in the nanomolar to micromolar range have been described (Azmi and Mohammad, 2009) and several compounds are currently in clinical development: For example, (-)-Gossypol, a polyphenolic aldehyde (molecular weight [MW] = 519 g/mol, K_i = 480 nM) (Wang et al., 2006) has been shown to be a potent inhibitor of cell growth and clinical trials are ongoing (Kitada et al., 2003; Ko et al., 2007; Van Poznak et al., 2001). The Bcl-xL inhibitor ABT-263 (MW = 975 g/mol, Ki = 0.5 nM) (Tse et al., 2008), which has been shown to have potent cytotoxicity in vitro, is currently tested in phase I/II clinical trials (Lock et al., 2008; Azmi and Mohammad, 2009). Our discovery of compound **6-199** as a Bcl-xL binder (Kd = 10 μM) might allow future affinity improvement, either by a medicinal chemical approach or by the construction of DNA-encoded affinity maturation libraries that contain this chemical moiety. Compound **6-199** was found to display an affinity to Bcl-xL that is comparable to the BH3I-1 ligand recently described in the literature (K_d = 4.4 μM; Figure S3) (Degterev et al., 2001). The BH3I-1 ligand was identified in a screening of 16,320 individual compounds, by competition measurements of a fluorescently-labeled peptide derived from the Bcl-xL interaction partner Bak. Future studies will elucidate whether **6-199** is capable of inhibiting protein-protein interactions with other Bcl-2 family members and can function in vivo.

The proinflammatory cytokine TNF represents a target of great pharmaceutical interest, because it is crucial for the establishment and maintenance of inflammation in multiple autoimmune diseases. TNF blockade has proven to be an efficient treatment of a number of human diseases such as rheumatoid arthritis, inflammatory bowel disease, psoriatic arthritis, psoriasis, and others (Lin et al., 2008). Three TNF antagonists are broadly used in the clinical practice: the monoclonal antibodies adalimumab and infliximab and a soluble receptor etanercept, but the current injectable protein therapies have also associated risks and limitations (Wong et al., 2008). The clinical use of an orally available drug, which interferes with TNF biology, is so far limited to TNF synthesis inhibitors or intracellular pathway inhibitors that antagonize nuclear factor κB (Palladino et al., 2003) The direct disruption of protein-protein interactions such as the TNF interaction with its receptors remains a difficult challenge in drug discovery (Wells and McClendon, 2007). He at al. were the first to discover a small molecule that acts directly on TNF, inhibiting its receptor interaction in the micromolar range by destabilization of the trimeric complex (He et al., 2005). Our DNA-encoded library selection against TNF led to the identification of novel lead structures binding to TNF in the low micromolar range, which proved to have a direct effect on cell survival in a TNF cytotoxicity assay, therefore presenting promising structures for future medicinal chemistry optimization.

Our identification of specific ligands from a 4000-compound DNA-encoded chemical library using high-throughput sequencing and the statistic analysis are encouraging for the construction of future libraries of more than 10^6 compounds: orthogonal synthetic procedures are necessary to react the first building block A with two chemical moieties B and C. Upon introduction of a second functional group besides the diene-moiety on the first building block A (e.g., free amino groups are compatible with the Diels-Alder reaction performed on oligonucleotides [data not shown]), the displayed chemical moieties A can be further diversified using two split and pool synthesis rounds. In this case, two codes for the last two reaction steps can be introduced in a sequential manner (Clark et al., 2009; Mannocci et al., 2008). Recent advances in high-throughput DNA sequencing indicate that it should be possible to sequence more than 1 million sequence tags per sequencing run (Pettersson et al., 2009). Thus, selections and decoding of DNA-encoded libraries containing millions of chemical compounds can be performed, and facilitate the identification of binding molecules for pharmaceutical applications.

SIGNIFICANCE

DNA-encoded chemical libraries represent a new tool for the identification of small organic molecules capable of specific binding to proteins of biomedical relevance, thus contributing to the pharmaceutical lead discovery process. In this article, we demonstrate that a DNA-encoded chemical library, constructed via Diels-Alder cycloaddition reactions, could be used in affinity-capture procedures for the identification of binders to four proteins chosen as targets for our selections. Decoding library composition by high-throughput sequencing enabled the relative quantification of library members before and after affinity capture and allowed the discovery of novel specific ligands. Statistical analysis of sequence data provided an objective avenue for the identification of candidate molecules that are significantly enriched after affinity-based selections using high-throughput DNA sequencing procedures. This work provides the basis for the construction and decoding of DNA-encoded chemical libraries of larger size (e.g., > 10^6 members), generated by split and pool procedures with three sets of building blocks, offering a fast and cheap alternative to conventional high-throughput small-molecule screening.

EXPERIMENTAL PROCEDURES

General Methods

All chemicals were purchased from Sigma-Aldrich-Fluka (Buchs, Switzerland), ABCR (Karlsruhe, Germany), VWR (Leicestershire, England), or Bachem (Bubendorf, Switzerland). Enzymes were purchased from New England Biolabs (Ipswich, MA) and high-pressure liquid chromatography (HPLC) grade lyophilized DNA oligonucleotides were purchased from IBA GmbH (Göttingen, Germany) or Sigma-Aldrich (Buchs, Switzerland). Streptavidin-sepharose slurry and CNBr-modified sepharose resin were purchased from GE Healthcare. Human MMP-3 catalytic domain (Scheuermann et al., 2008), Bcl-xL (Manion et al., 2004), and the ED-B domain of fibronectin (Carnemolla et al., 1996) were expressed as described previously. Recombinant mouse TNF was purchased from eBioscience (San Diego, CA). The EDB domain of fibronectin (Zardi et al., 1987) was fused to mouse TNF with a $(SSSSG)_3$ linker to improve solubility and facilitate purification. The resulting gene was cloned into the mammalian cell expression vector pcDNA3.1 (Invitrogen, Basel, Switzerland). The protein was expressed in HEK293 cells and purified via anion-exchange chromatography. Biological activity of the TNF constructs was confirmed as described previously (Halin et al., 2003). SpinX columns were purchased from Corning Costar Incorporated (Acton, MA). HPLC purifications were performed on a Waters 2795 Alliance system with a Synergi Polar-RP_{18} column (4 μm, 10 × 150 mm, Phenomenex, Torrance CA) using a linear gradient from 0% to 100% MeCN with 0.1% TFA. Electrospray ionization MS was performed on a Waters Quattro Micro spectrometer (Waters, Milford, MA). Statistical computation was carried out using R version 2.8.1 and the multtest package, available from the Comprehensive R Archive Network (http://cran.r-project.org/) under the terms of the GNU General Public License.

Preparation of Target Protein Resin

CNBr-modified sepharose (100 mg, GE Healthcare) was thoroughly washed with cold 1 mM HCl and incubated overnight at 4°C either with 100 mg/ml human serum albumin (in 100 mM $NaHCO_3$, 0.5 M NaCl [pH 9]), 1 mg/ml trypsin (in 100 mM $NaHCO_3$, 0.5 M NaCl [pH 8.3]), 1.6 mg/ml MMP-3 (in 100 mM $NaHCO_3$, 0.5 M NaCl [pH 8.3]), 1.4 mg/ml Bcl-xL (in phosphate-buffered saline [PBS] [pH 7.4]), 2.3 mg/ml EDB-TNF (in PBS [pH 7.4]), or with 1 M Tris/HCl (pH 8.9) (Tris-quenched resin). MMP-3 resin was prepared and used as described previously (Scheuermann et al., 2008). After coupling, resins were washed repeatedly with alternating buffers (100 mM HOAc/NaOAc, 0.5M NaCl [pH 4.7] and 0.1 M Tris/HCl, 0.5 M NaCl [pH 8]) using SpinX columns. Before selection, resins were preincubated with PBS and 0.2 mg/ml herring sperm DNA.

Selection for Binding

The DNA-encoded 4000-compound library (Buller et al., 2008) was diluted in PBS to a final concentration of 20 nM and 50 μl was added to 50 μl sepharose resin slurry, carrying the immobilized target protein, or to Tris-quenched sepharose resin. After incubation for 1 hr at room temperature, the mixture was transferred to a SpinX column, the supernatant was removed, and the resin was washed five times with 400 μl PBS containing 1 mM $MgCl_2$. Subsequently the resin was resuspended in 100 μl H_2O.

Decoding by High-Throughput Sequencing

After selections, the codes of the oligonucleotide-compound conjugates were amplified by PCR (50 μl reaction, 25 cycles of 1 min at 98°C, 1 min at 50°C, 30 s at 72°C, with Phusion High-Fidelity DNA Polymerase, New England Biolabs) using 5 μl resuspended resin as a template. For an analysis of the library before selection, the library was diluted to 1 nM and 5 μl was used as a template for PCR amplification. The PCR primers P1_HTS (*GCC TCC CTC GCG CCA TCA* GGG AGC TTG TGA ATT CTG G) and P2_HTS (*GCC TTG CCA GCC CGC TCA* GGT AGT CGG ATC CGA CCA C) contain at the 5′-extremity a 19-base domain (italicized) required for high-throughput sequencing with the 454 genome sequencer system. The PCR products were purified by agarose gel electrophoresis and subsequent ion-exchange cartridge purification and ethanol precipitation of the DNA. High-throughput sequencing was performed on a 454 Life Sciences-Roche GS 20 sequencer platform (sequencing service by Roche USA). Analysis of the sequences was performed by an in-house program written in C++. The frequency of each code was assigned to each individual compound-conjugate, after normalization to 50,000 sequences per experiment.

Affinity Measurements

Fluorescein conjugates (500 nM, Syntheses, see Supplemental Data) were incubated with increasing amounts of protein in PBS containing 1.7% dimethyl sulfoxide (DMSO), for 1 hr at 25°C. Fluorescence polarization was determined with a TECAN Polarion instrument. The change of polarization signal was plotted against the corresponding protein concentrations, and the K_D value was obtained by fitting with equation $y = a + b \times 0.5 \times (((x + const + c) - sqrt((x + const + c)^2 - 4 \times const \times x)))$ (const = 500 nM, total ligand concentration).

Trypsin Inhibition Assay

Bovine pancreatic trypsin (2 nM) in PBS (20mM NaH_2PO_4, 30 mM Na_2HPO_4, 100 mM NaCl [pH 7.4], containing 5% DMSO) was incubated with varying concentrations of inhibitors in 96-well microtiter plates (Nunc, Roskilde, Denmark) for 30 min at room temperature (Syntheses, see Supplemental Data). The reaction was started by the addition of the fluorogenic substrate Z-GGR-AMC (Bachem, Bubendorf, Switzerland) dissolved in PBS to an end concentration of 0.1 mM in a total volume of 100 μl. The change of fluorescence signal (ex: 383 nm; 455 nm; cutoff filter: 420 nm) was recorded over 10 min using a SpectraMax microplate reader (Molecular Devices). The rate of fluorescence signal increase over time (reaction velocity v) was plotted against the corresponding inhibitor concentrations, and the K_i value for the inhibitor was obtained by fitting with equation $v = v_{max}/(1+[I_0]/K_i)$ ($[I_o]$: inhibitor concentration).

TNF Inhibition Assay

TNF bioactivity was determined by performing a cytotoxicity assay on murine L-M fibroblasts (Corti et al., 1994; Halin et al., 2003). L-M fibroblasts were grown in DMEM containing 10% FCS and antibiotic-antimycotic solution (Invitrogen) and seeded in 96-well plates at 30,000 cells/well in 100 μl medium. Recombinant mouse TNF (final concentration of 3 pM trimer) was preincubated for 30 min at room temperature with varying concentrations of inhibitors in media containing actinomycin-D (final concentration of 2 μg/ml) and DMSO (2% of the final concentration) (Syntheses see Supplemental Data). After addition of the TNF compound mixture (50 μl), cells were incubated for 24 hr at 37°C, 5% CO_2. Next, a solution of 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (30 μL, CellTiter 96® AQueous Assay, Promega, Madison, WI) was added and plates were incubated for an additional 2 hr at 37°C, 5% CO_2. Plates were then read at 490nm using a VersaMax microplate reader (Molecular Devices).

SUPPLEMENTAL DATA

Supplemental Data include three figures and Supplemental Experimental Procedures and can be found with the article online at http://www.cell.com/chemistry-biology/supplemental/S1074-5521(09)00321-4.

ACKNOWLEDGMENTS

We thank Ilona Molnar and Francesca Bacchion for help in synthesis and protein purifications, and Eveline Trachsel for EDB-TNF cloning. Financial support from the ETH Zurich, the Swiss National Science Foundation, the Scholarship Fund of the Swiss Chemical Society (SSCI), Philochem AG, KTI (Grant Nr. 8868.1 PFDS-LS), the Gebert-Rüf Foundation, the European Union (ADAMANT Project), and DSM Nutritional Products Ltd. (Basel, Switzerland) is gratefully acknowledged. The authors are grateful to the Functional Genomics Center Zurich for help with the high-throughput sequencing technology, and to Luca Mannocci (Philochem AG) for reviewing the manuscript. D.N. is a cofounder and shareholder of Philochem AG, a company that has certain rights to the technology described in this article.

Received: April 7, 2009
Revised: September 15, 2009
Accepted: September 17, 2009
Published: October 30, 2009

REFERENCES

Andersen, J.T., Dee Qian, J., and Sandlie, I. (2006). The conserved histidine 166 residue of the human neonatal Fc receptor heavy chain is critical for the pH-dependent binding to albumin. Eur. J. Immunol. *36*, 3044–3051.

Annis, D.A., Nazef, N., Chuang, C.C., Scott, M.P., and Nash, H.M. (2004). A general technique to rank protein-ligand binding affinities and determine allosteric versus direct binding site competition in compound mixtures. J. Am. Chem. Soc. *126*, 15495–15503.

Azmi, A.S., and Mohammad, R.M. (2009). Non-peptidic small molecule inhibitors against Bcl-2 for cancer therapy. J. Cell. Physiol. *218*, 13–21.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B *57*, 289–300.

Blundell, T.L., Jhoti, H., and Abell, C. (2002). High-throughput crystallography for lead discovery in drug design. Nat. Rev. Drug Discov. *1*, 45–54.

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

Buller, F., Mannocci, L., Zhang, Y., Dumelin, C.E., Scheuermann, J., and Neri, D. (2008). Design and synthesis of a novel DNA-encoded chemical library using Diels-Alder cycloadditions. Bioorg. Med. Chem. Lett. *18*, 5926–5931.

Carnemolla, B., Neri, D., Castellani, P., Leprini, A., Neri, G., Pini, A., Winter, G., and Zardi, L. (1996). Phage antibodies with pan-species recognition of the oncofoetal angiogenesis marker fibronectin ED-B domain. Int. J. Cancer *68*, 397–405.

Clark, M.A., Acharya, R.A., Arico-Muendel, C.C., Belyanskaya, S.L., Benjamin, D.R., Carlson, N.R., Centrella, P.A., Chiu, C.H., Creaser, S.P., Cuozzo, J.W., et al. (2009). Design, synthesis and selection of DNA-encoded small-molecule libraries. Nat. Chem. Biol. *5*, 647–654.

Corti, A., Poiesi, C., Merli, S., and Cassani, G. (1994). Tumor necrosis factor (TNF) alpha quantification by ELISA and bioassay: effects of TNF alpha-soluble TNF receptor (p55) complex dissociation during assay incubations. J. Immunol. Methods *177*, 191–198.

Degterev, A., Lugovskoy, A., Cardone, M., Mulley, B., Wagner, G., Mitchison, T., and Yuan, J. (2001). Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-xL. Nat. Cell Biol. *3*, 173–182.

Dennis, M.S., Zhang, M., Meng, Y.G., Kadkhodayan, M., Kirchhofer, D., Combs, D., and Damico, L.A. (2002). Albumin binding as a general strategy for improving the pharmacokinetics of proteins. J. Biol. Chem. *277*, 35035–35043.

Diaz, J.L., Oltersdorf, T., Horne, W., McConnell, M., Wilson, G., Weeks, S., Garcia, T., and Fritz, L.C. (1997). A common binding site mediates heterodimerization and homodimerization of Bcl-2 family members. J. Biol. Chem. *272*, 11350–11355.

Dumelin, C.E., Scheuermann, J., Melkko, S., and Neri, D. (2006). Selection of streptavidin binders from a DNA-encoded chemical library. Bioconjug. Chem. *17*, 366–370.

Dumelin, C.E., Trussel, S., Buller, F., Trachsel, E., Bootz, F., Zhang, Y., Mannocci, L., Beck, S.C., Drumea-Mirancea, M., Seeliger, M.W., et al. (2008). A portable albumin binder from a DNA-encoded chemical library. Angew. Chem. Int. Ed. Engl. *47*, 3196–3201.

Gartner, Z.J., Tse, B.N., Grubina, R., Doyon, J.B., Snyder, T.M., and Liu, D.R. (2004). DNA-templated organic synthesis and selection of a library of macrocycles. Science *305*, 1601–1605.

Hajduk, P.J., and Greer, J. (2007). A decade of fragment-based drug design: strategic advances and lessons learned. Nat. Rev. Drug Discov. *6*, 211–219.

Halin, C., Gafner, V., Villani, M.E., Borsi, L., Berndt, A., Kosmehl, H., Zardi, L., and Neri, D. (2003). Synergistic therapeutic effects of a tumor targeting antibody fragment, fused to interleukin 12 and to tumor necrosis factor alpha. Cancer Res. *63*, 3202–3210.

Halpin, D.R., and Harbury, P.B. (2004a). DNA display I. Sequence-encoded routing of DNA populations. PLoS Biol. *2*, E173.

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

Halpin, D.R., Lee, J.A., Wrenn, S.J., and Harbury, P.B. (2004). DNA display III. Solid-phase organic synthesis on unprotected DNA. PLoS Biol. *2*, E175.

Hansen, M.H., Blakskjaer, P., Petersen, L.K., Hansen, T.H., Hojfeldt, J.W., Gothelf, K.V., and Hansen, N.J. (2009). A yoctoliter-scale DNA reactor for small-molecule evolution. J. Am. Chem. Soc. *131*, 1322–1327.

He, M.M., Smith, A.S., Oslob, J.D., Flanagan, W.M., Braisted, A.C., Whitty, A., Cancilla, M.T., Wang, J., Lugovskoy, A.A., Yoburn, J.C., et al. (2005). Small-molecule inhibition of TNF-alpha. Science *310*, 1022–1025.

Hirsch, J.D., Eslamizar, L., Filanoski, B.J., Malekzadeh, N., Haugland, R.P., and Beechem, J.M. (2002). Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. Anal. Biochem. *308*, 343–357.

Kitada, S., Leone, M., Sareth, S., Zhai, D., Reed, J.C., and Pellecchia, M. (2003). Discovery, characterization, and structure-activity relationships studies of proapoptotic polyphenols targeting B-cell lymphocyte/leukemia-2 proteins. J. Med. Chem. *46*, 4259–4264.

Ko, C.H., Shen, S.C., Yang, L.Y., Lin, C.W., and Chen, Y.C. (2007). Gossypol reduction of tumor growth through ROS-dependent mitochondria pathway in human colorectal carcinoma cells. Int. J. Cancer *121*, 1670–1679.

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.

Lin, J., Ziring, D., Desai, S., Kim, S., Wong, M., Korin, Y., Braun, J., Reed, E., Gjertson, D., and Singh, R.R. (2008). TNFalpha blockade in human diseases: an overview of efficacy and safety. Clin. Immunol. *126*, 13–30.

Lock, R., Carol, H., Houghton, P.J., Morton, C.L., Kolb, E.A., Gorlick, R., Reynolds, C.P., Maris, J.M., Keir, S.T., Wu, J., et al. (2008). Initial testing (stage 1) of the BH3 mimetic ABT-263 by the pediatric preclinical testing program. Pediatr. Blood Cancer *50*, 1181–1189.

Makara, G.M., and Athanasopoulos, J. (2005). Improving success rates for lead generation using affinity binding technologies. Curr. Opin. Biotechnol. *16*, 666–673.

Manion, M.K., O'Neill, J.W., Giedt, C.D., Kim, K.M., Zhang, K.Y., and Hockenbery, D.M. (2004). Bcl-XL mutations suppress cellular sensitivity to antimycin A. J. Biol. Chem. *279*, 2159–2165.

Mannocci, L., Zhang, Y., Scheuermann, J., Leimbacher, M., De Bellis, G., Rizzi, E., Dumelin, C., Melkko, S., and Neri, D. (2008). High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. Proc. Natl. Acad. Sci. USA *105*, 17670–17675.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376–380.

Marzo, I., and Naval, J. (2008). Bcl-2 family members as molecular targets in cancer therapy. Biochem. Pharmacol. *76*, 939–946.

Melkko, S., Scheuermann, J., Dumelin, C.E., and Neri, D. (2004). Encoded self-assembling chemical libraries. Nat. Biotechnol. *22*, 568–574.

Melkko, S., Dumelin, C.E., Scheuermann, J., and Neri, D. (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.

Melkko, S., Dumelin, C.E., Scheuermann, J., and Neri, D. (2007a). Lead discovery by DNA-encoded chemical libraries. Drug Discov. Today *12*, 465–471.

Melkko, S., Zhang, Y., Dumelin, C.E., Scheuermann, J., and Neri, D. (2007b). Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. Angew. Chem. Int. Ed. Engl. *46*, 4671–4674.

Muckenschnabel, I., Falchetto, R., Mayr, L.M., and Filipuzzi, I. (2004). SpeedScreen: label-free liquid chromatography-mass spectrometry-based high-throughput screening for the discovery of orphan protein ligands. Anal. Biochem. *324*, 241–249.

Nguyen, A., Reyes, A.E., 2nd, Zhang, M., McDonald, P., Wong, W.L., Damico, L.A., and Dennis, M.S. (2006). The pharmacokinetics of an albumin-binding Fab (AB.Fab) can be modulated as a function of affinity for albumin. Protein Eng. Des. Sel. *19*, 291–297.

Palladino, M.A., Bahjat, F.R., Theodorakis, E.A., and Moldawer, L.L. (2003). Anti-TNF-alpha therapies: the next generation. Nat. Rev. Drug Discov. *2*, 736–746.

Pellecchia, M., Bertini, I., Cowburn, D., Dalvit, C., Giralt, E., Jahnke, W., James, T.L., Homans, S.W., Kessler, H., Luchinat, C., et al. (2008). Perspectives on NMR in drug discovery: a technique comes of age. Nat. Rev. Drug Discov. *7*, 738–745.

Pettersson, E., Lundeberg, J., and Ahmadian, A. (2009). Generations of sequencing technologies. Genomics *93*, 105–111.

Rees, D.C., Congreve, M., Murray, C.W., and Carr, R. (2004). Fragment-based lead discovery. Nat. Rev. Drug Discov. *3*, 660–672.

Scheuermann, J., Dumelin, C.E., Melkko, S., and Neri, D. (2006). DNA-encoded chemical libraries. J. Biotechnol. *126*, 568–581.

Scheuermann, J., Dumelin, C.E., Melkko, S., Zhang, Y., Mannocci, L., Jaggi, M., Sobek, J., and Neri, D. (2008). DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors. Bioconjug. Chem. *19*, 778–785.

Schreiber, S.L. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. Science *287*, 1964–1969.

Shuker, S.B., Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. Science *274*, 1531–1534.

Silacci, M., Brack, S., Schirru, G., Marlind, J., Ettorre, A., Merlo, A., Viti, F., and Neri, D. (2005). Design, construction, and characterization of a large synthetic human antibody phage display library. Proteomics *5*, 2340–2350.

Syed, S., Schuyler, P.D., Kulczycky, M., and Sheffield, W.P. (1997). Potent antithrombin activity and delayed clearance from the circulation characterize recombinant hirudin genetically fused to albumin. Blood *89*, 3243–3252.

Tse, C., Shoemaker, A.R., Adickes, J., Anderson, M.G., Chen, J., Jin, S., Johnson, E.F., Marsh, K.C., Mitten, M.J., Nimmer, P., et al. (2008). ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. Cancer Res. *68*, 3421–3428.

Van Poznak, C., Seidman, A.D., Reidenberg, M.M., Moasser, M.M., Sklarin, N., Van Zee, K., Borgen, P., Gollub, M., Bacotti, D., Yao, T.J., et al. (2001). Oral gossypol in the treatment of patients with refractory metastatic breast cancer: a phase I/II clinical trial. Breast Cancer Res. Treat. *66*, 239–248.

Vander Heiden, M.G., and Thompson, C.B. (1999). Bcl-2 proteins: regulators of apoptosis or of mitochondrial homeostasis? Nat. Cell Biol. *1*, E209–E216.

Wang, G., Nikolovska-Coleska, Z., Yang, C.Y., Wang, R., Tang, G., Guo, J., Shangary, S., Qiu, S., Gao, W., Yang, D., et al. (2006). Structure-based design of potent small-molecule inhibitors of anti-apoptotic Bcl-2 proteins. J. Med. Chem. *49*, 6139–6142.

Wells, J.A., and McClendon, C.L. (2007). Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature *450*, 1001–1009.

Winter, G., Griffiths, A.D., Hawkins, R.E., and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. Annu. Rev. Immunol. *12*, 433–455.

Wong, M., Ziring, D., Korin, Y., Desai, S., Kim, S., Lin, J., Gjertson, D., Braun, J., Reed, E., and Singh, R.R. (2008). TNFalpha blockade in human diseases: mechanisms and future directions. Clin. Immunol. *126*, 121–136.

Zahnd, C., Amstutz, P., and Pluckthun, A. (2007). Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. Nat. Methods *4*, 269–279.

Zardi, L., Carnemolla, B., Siri, A., Petersen, T.E., Paolella, G., Sebastio, G., and Baralle, F.E. (1987). Transformed human cells produce a new fibronectin isoform by preferential alternative splicing of a previously unobserved exon. EMBO J. *6*, 2337–2342.

Zehender, H., and Mayr, L.M. (2007). Application of mass spectrometry technologies for the discovery of low-molecular weight modulators of enzymes and protein-protein interactions. Curr. Opin. Chem. Biol. *11*, 511–517.

