ACS chemical biology

ARTICLES

pubs.acs.org/acschemicalbiology

Selection of Carbonic Anhydrase IX Inhibitors from One Million DNA-Encoded Compounds

Fabian Buller,[†] Martina Steiner,[†] Katharina Frey,[†] Dennis Mircsof,[†] Jörg Scheuermann,[†] Markus Kalisch,[‡] Peter Bühlmann,[‡] Claudiu T. Supuran,[§] and Dario Neri[†,*]

[†]Institute of Pharmaceutical Sciences, Department of Chemistry and Applied Biosciences, ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland

[‡]Seminar for Statistics, ETH Zürich, Rämistrasse 101, CH-8092 Zurich, Switzerland

[§]Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Rm. 188, Via della Lastruccia 3, I-50019 Sesto Fiorentino, Firenze, Italy

Supporting Information

ABSTRACT: DNA-encoded chemical libraries, i.e., collections of compounds individually coupled to distinctive DNA fragments serving as amplifiable identification barcodes, represent a new tool for the *de novo* discovery of small molecule ligands to target proteins of pharmaceutical interest. Here, we describe the design and synthesis of a novel DNA-encoded chemical library containing one million small molecules. The library was synthesized by combinatorial assembly of three sets of chemical building blocks using Diels−Alder cycloadditions and by the stepwise build-up of the DNA barcodes. Model selections were performed to test library performance and to develop a statistical method for the analysis of high-throughput sequencing data. A library selection against carbonic anhydrase IX revealed a new class of submicromolar bis(sulfonamide) inhibitors. One of these inhibitors was synthesized in the absence of the DNA-tag and showed accumulation in hypoxic tumor tissue sections *in vitro* and tumor targeting *in vivo*.

DNA-encoded chemical libraries are increasingly used in Academia and in the Pharmaceutical Industry for the discovery of small molecule ligands against proteins of pharmaceutical interest.[1−3] In close analogy to display technologies of polypeptides (*e.g.*, Antibody Phage Display Technology)[4,5] the covalent linkage of small molecules to unique DNA-tags enables the affinity-based isolation of small molecule ligands against a target protein of interest from large libraries.[1−3] In most selection procedures, a DNA-encoded chemical library is incubated with a target protein immobilized on a solid support, and repetitive washing of the support allows the recovery of library members that bind the target protein (Figure 1). Ligands enriched in the selection procedure are identified by PCR amplification of the corresponding DNA-tag and by the subsequent sequencing of up to millions of PCR amplicons either by Roche's 454 or by the Illumina technology.[6,7] The relative abundance of library codes found in the sequencing experiment correlates to those small molecules that are retained on the target protein in the selection experiment.

In contrast to conventional screening technologies of small molecule libraries (*e.g.*, by the robotic screening of single compounds in a functional assay or fragment-based screening of small molecule libraries),[8,9] DNA-encoded chemical libraries offer an alternative strategy for ligand discovery: (i) affinity-based selection does not require assay development and can in principle be performed on any isolated protein; (ii) miniaturization of selection conditions (μL volume) enables the parallel selection of up to millions of small molecules with minimal amount of target protein (μg quantities of protein) and (iii) once libraries have been synthesized selections can be performed at low costs (i.e., cost of DNA sequencing reagents).[7,10−21]

At present, a number of different DNA-encoded chemical library formats are explored in terms of library size, synthesis strategy and number of chemical moieties displayed on the DNA-strand.[1] Here, we focus on DNA-encoded chemical libraries in which only one chemical moiety is displayed on the DNA strand (single pharmacophore library), but the simultaneous display of pairs of chemical moieties at the 5′ and 3′ end of complementary DNA strands has also been described (encoded self-assembling chemical libraries).[13,16,18] Single pharmacophore DNA-encoded chemical libraries can be synthesized by the combinatorial assembly of chemical building blocks using split-and-pool synthesis[6,19,21,22] or by the use of complementary oligonucleotide derivatives which facilitate chemical reactions (DNA-templated synthesis).[14,15,17,23]

Received: October 29, 2010
Accepted: December 27, 2010
Published: December 27, 2010

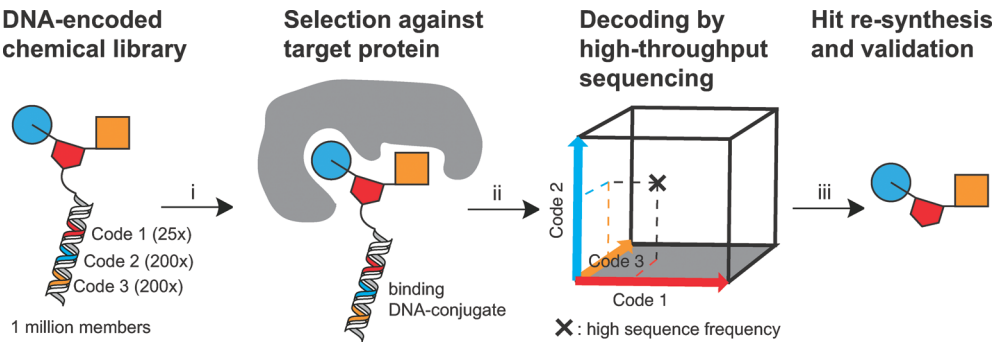


Figure 1. Selection of binders using a DNA-encoded chemical library. (i) The one million compound library consisting of three chemical building blocks is panned on a target protein immobilized on a solid support. Nonbinding conjugates are washed away before (ii) PCR amplification and Illumina high-throughput sequencing of the retained DNA-conjugates. (iii) High code frequencies observed in the decoding experiment correspond to hit-structures that are resynthesized for further characterization.

The split-and-pool synthesis of DNA-encoded chemical libraries typically relies on high-yield and DNA-compatible reactions for the coupling of organic molecules to oligonucleotide derivatives. Single reactions can be carried out in small scale (down to nanomole quantities of DNA-conjugates) with large excess of reagents, as the DNA-conjugates can be purified by precipitation or immobilization on positively charged resin.[6,22,24] Libraries of up to thousands of small molecules consisting of two sets of chemical building blocks have been described using amide bond formation or Diels–Alder cycloaddition reactions.[6,12,22] Additionally, a library constructed by DNA-templated synthesis with thousands of macrocyclic structures [15] and a library synthesized through the hybridization of DNA hairpin loops yielding linear peptidic structures have also been reported.[17]

DNA-encoded chemical libraries have been used in selection experiments against kinases,[14,21] proteases [13,19] and other targets of pharmaceutical interest (*e.g.*, BcL-xL,[12] TNF,[10] HSA [25]) yielding ligands with dissociation constants ranging from low μM to single-digit nM values. DNA-encoded chemical libraries have also been used to improve known lead structures. The incorporation of a lead structure during library synthesis, with the aim to optimize a starting compound with suboptimal affinity to the target, has been experimentally demonstrated with the successful isolation of highly potent and selective serine protease inhibitors.[13,19]

The recent availability of sequencing technologies, which permit the simultaneous analysis of millions of DNA sequences, allows the decoding of large DNA-encoded chemical libraries.[7,21] This development allowed increasing DNA-encoded chemical library diversity by the introduction of additional sets of building blocks within a split-and-pool synthesis. Barry Morgan and colleagues have recently described two libraries based on cyanuric chloride substitutions with three and four sets of chemical building blocks yielding hundreds of millions of encoded small molecules.[21]

In this article, we describe the synthesis of a new DNA-encoded chemical library with one million compounds based on Diels–Alder cycloadditions and amine modifications. A novel library synthesis strategy allowed the assembly of three sets of chemical building blocks and three DNA-coding regions. We performed selections against the model target streptavidin to validate library performance and derive statistical tools for high-throughput sequencing data analysis. Furthermore, we selected from one million DNA-encoded compounds potent inhibitors of carbonic anhydrase IX (CA IX), an enzyme upregulated in cancer.[26] In particular, a bis(sulfonamide) CA IX inhibitor allowed the detection of the cognate target protein at hypoxic sites on tumor tissue sections, with a performance which rivaled the one of monoclonal antibodies in similar immunochemical procedures. Importantly, the bis(sulfonamide) CA IX binding compound exhibited the ability to selectively localize to neoplastic lesions in mice carrying xenografted LS174T and SK-RC-52 tumors.

RESULTS AND DISCUSSION

Design and Synthesis of a DNA-Encoded Chemical Library. DNA encoding facilitates the construction and screening of small molecule libraries of unprecedented size. A split-and-pool synthesis approach with alternating steps for the stepwise assembly of chemical moieties and for the synthesis of the corresponding DNA barcodes was chosen for library construction (Figure 2). We used amide-bond formation and Diels–Alder cycloadditions as key synthetic steps, since these reactions had previously been successfully used by our group for the construction of smaller DNA-encoded chemical libraries and for the isolation of binders to proteins of pharmaceutical interest.[6,22] The one-million compound library was constructed by the following sequential chemical steps: (i) coupling of 25 scaffold molecules with both amino- and (2*E*,4*E*)-hexadiene-functionalities to amino-modified oligonucleotides carrying code 1. The resulting 25 oligonucleotide conjugates were then pooled and split into 200 reaction vessels; (ii) in each vessel, the free amino-group was modified with a reactive building block and the new functionality was encoded by hybridization of a partially double-stranded DNA-fragment followed by Klenow-fragment mediated polymerization (iii).[6] These 200 encoded reaction products were then pooled and split to 200 new reaction vessels; (iv) in each vessel, a Diels–Alder reaction between the (2*E*,4*E*)-hexadiene group and maleimides as dienophiles was carried out before the final encoding step, achieved by ligation to a partially double-stranded DNA fragment carrying code 3. The 200 reactions were then pooled to yield a library with one million DNA encoded members (v) (Figure 2).

In detail, the 25 initial DNA derivatives were synthesized by the coupling of a (2*E*,4*E*)-hexadiene-modified hydroxyproline derivative and Fmoc-protected amino acids to individual amino-tagged oligonucleotides (Figure 2 and Supporting Figure 1), which carried a 9 bp "code" region and two hybridization domains for subsequent encoding and PCR amplification steps.

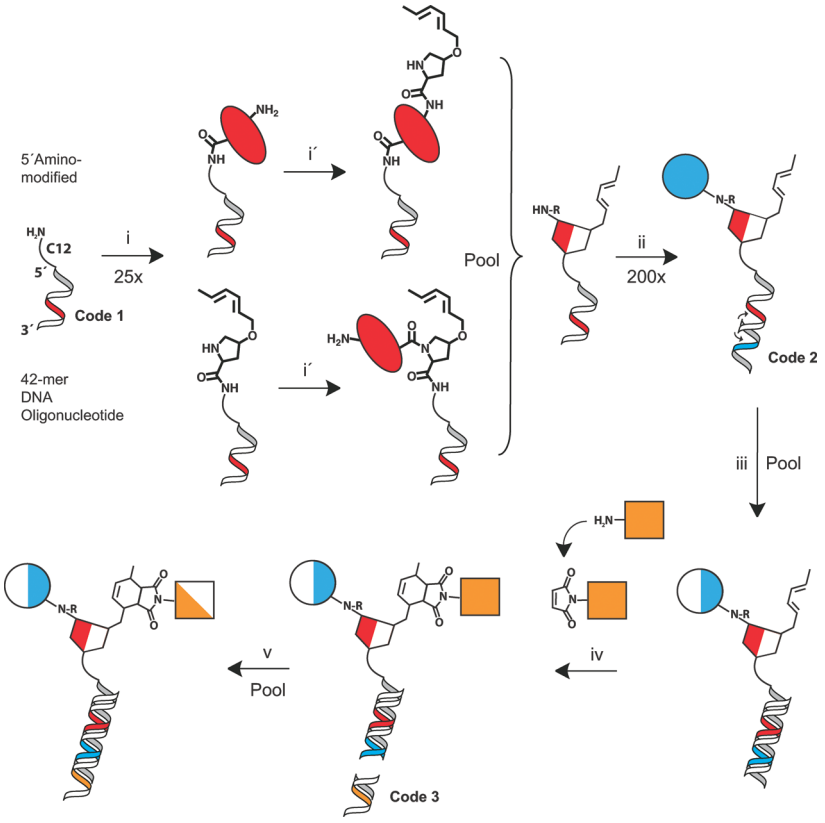


Figure 2. Scheme of a three-step split-and-pool synthesis of a DNA-encoded chemical library. Twenty-five scaffold molecules with an amino- and a (2*E*,4*E*)-hexadiene functionality are conjugated to amino-modified oligonucleotides in a two step process using a set of Fmoc-protected amino acid building blocks (red, i or i′) and a diene-hydroxyproline derivative (i or i′). (ii) The primary or secondary amino group is modified with reactive building blocks (blue, *e.g.*, activated carboxylic acids, sulfonyl chlorides), and (iii) the second reaction is encoded by elongation of the DNA-tag. (iv) Maleimides are prepared (*e.g.*, from the corresponding amines, orange) and used in 200 Diels−Alder cycloaddition reactions. The reaction is encoded by ligation of a short DNA fragment carrying code 3. (v) The 200 encoded reactions are then pooled to yield a DNA-encoded chemical library containing one million members.

After deprotection, the amino group was modified with sets of carbonyl chlorides, sulfonyl chlorides, isothiocyanates, carboxylic acid anhydrides and carboxylic acids (Supporting Figure 3). Finally, high-yield Diels−Alder cycloadditions with reactive maleimido derivatives were performed (Supporting Figure 4).[22] The performance of the individual reaction steps could not be analytically monitored in the split-and-pool library synthesis, but was assessed in model reactions and was confirmed by the satisfactory recovery of "positive control" chemical moieties (*e.g.*, desthiobiotin) incorporated during library synthesis and detected after library selection on the cognate binding proteins (*e.g.*, streptavidin selection, see below).

The Diels−Alder reaction was encoded by ligation of 5′phosphorylated, partially double-stranded oligonucleotides carrying code 3 (8 bases), to *Bam*HI restricted DNA-conjugates. Compared to the ligation of a double-stranded DNA coding fragment, the new encoding procedure allowed to use only one oligonucleotide per code which was annealed to a shorter constant complementary oligonucleotide (Figure 2), thus reducing costs.

The extension of split-and-pool synthesis from two to three building blocks features a dramatic increase in library size. Library members are slightly larger than what is required by the "rule-of-five" [27,28] (average molecular weight of 805 g/mol, ClogP <5 for most library members) (Supporting Figure 5), which is not a major limitation when targeting extracellular proteins.[29]

Affinity-Based Selections. In full analogy to antibody phage library selections,[30] we panned the one million compound library against proteins immobilized on sepharose resin. As negative control, sepharose resin without immobilized target protein ("empty resin") was used. After repeated washing cycles, the DNA-code of the encoded compounds retained on the resin was PCR amplified and subjected to high-throughput sequencing. A comparison of sequence counts of individual library members before and after selection allowed the identification of preferential binding molecules.

Selections were performed at a concentration of 2 nM (total library concentration, each library member has a theoretical concentration of 2 fM) in a volume of 100 μL (containing $\sim 1.2 \times 10^5$ molecules of each library member). To prevent unspecific binding of library members (*e.g.*, through their DNA-tag), we added herring sperm DNA to the selection mixture. After selection, DNA-tags were PCR amplified and DNA flanking sites necessary for Illumina high-throughput sequencing were introduced.[7]

Selection Decoding. To assess library performance, we sequenced the library before selection, after selection against streptavidin resin and after selection against empty resin. A total of 0.64 million compounds were detected from 1.8 million DNA-tags sequenced from the unselected library (from a random sampling, one would expect 0.84 million compounds) (Figure 3a). The absolute sequence count distribution (background distribution) (Figure 4a) can be described by a negative binomial distribution, in analogy to the analysis of the sequence data obtained with a 4'000 member library and 454 sequencing.[10] We estimated the

dx.doi.org/10.1021/cb1003477 |*ACS Chem. Biol.* 2011, 6, 336–344

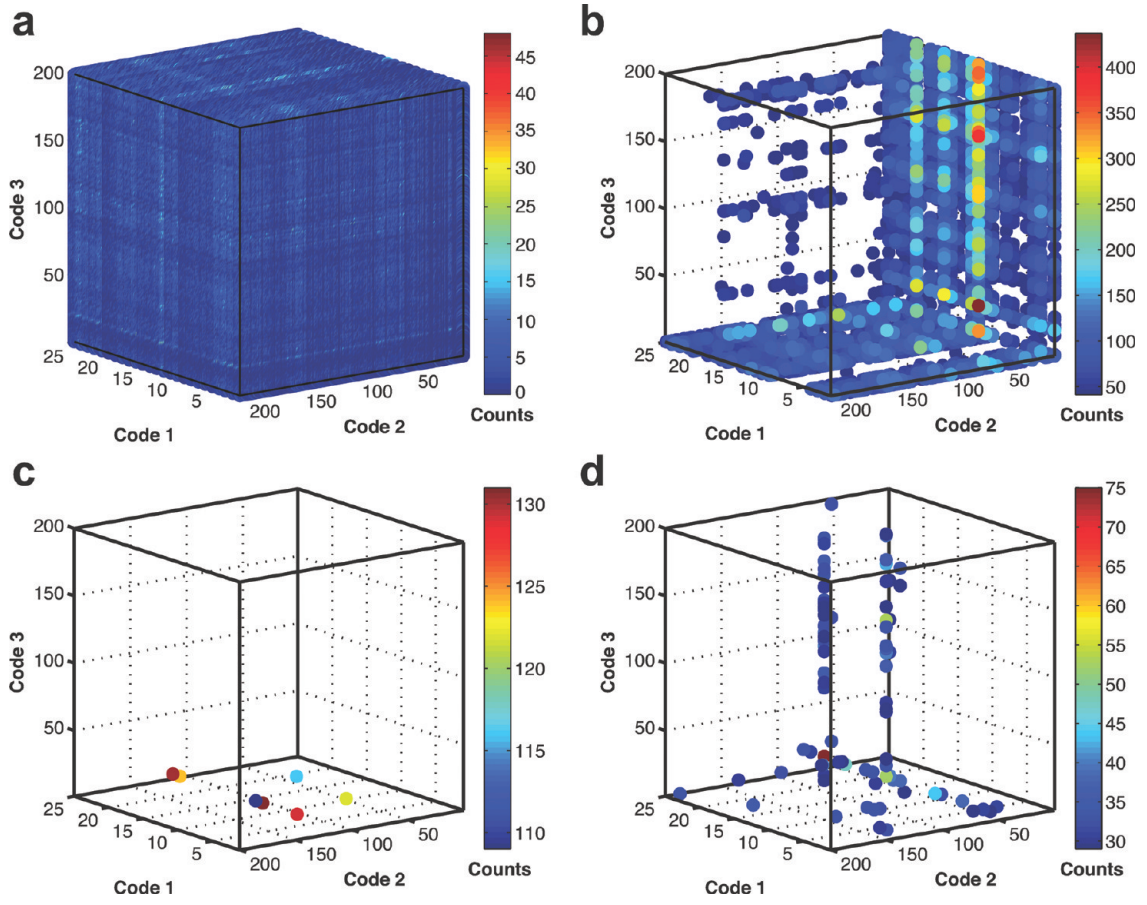


Figure 3. Decoding of the one million compound library before and after selection against target proteins using Illumina high-throughput sequencing. Sequence counts of library codes are plotted *vs* the code combination in pseudo-four-dimensional plots. The codes are represented on the axes, and the sequence counts are depicted by the color of the dots. (a) The unselected library before selection was analyzed with 1.8 million sequence tags. All DNA-conjugates identified are depicted indicating a flat background distribution of sequencing counts. (b) The sequence profile of DNA-conjugates enriched (*p*-value <0.05 for control of FWER) after a selection against streptavidin (1.4 million sequences analyzed). Two compound classes with code 2 = 1 or code 3 = 1 (plane [x-1-z] and [x-y-1]) were selected. Both compound classes include D-desthiobiotin structures, which are known to bind to streptavidin. (c) For comparison, the sequence count profile of a selection against sepharose resin with no protein immobilized shows that only a few compounds with Code 3 = 20 are retained on the solid support (4.0 million sequences analyzed, displayed are DNA-conjugates with *p*-value <0.05 for control of FWER). (d) The sequence profile after selection against CA IX (0.76 million sequences evaluated) reveals the preferential enrichment of compounds with code 2 = 83 and/or code 3 = 6 (displayed are DNA-conjugates with *p*-value <0.05 for control of FWER, selected structures are shown in Figure 6).

background distribution for each sequencing experiment by fitting a negative binomial distribution in a robust way.[31] Outliers in the selection experiments were detected based on deviations from the background distribution. Multiple testing was performed using the Bonferroni-Holm correction to control the Family-Wise-Error-Rate (FWER) with the level 0.05, yielding *p*-values for the library members enriched after selection.[32] As expected, before selection most library members were found to be evenly distributed and only four library members displayed more than 37 counts (i.e., *p*-values <0.05) (Figure 3a, 4a). For visualization, the sequence counts were plotted against the code-combination in pseudofour dimensional plots. The axes represent code 1, 2, and 3 while the sequence counts are depicted by the color of the dots (Figure 3a).

The sequence count profile after selection against streptavidin (Figure 3b) was dramatically different compared to the one of the unselected library (Figure 3a). The profile indicated the enrichment of two planes, with code-x-code-y-code-z being either x-1-z or x-y-1, representing two compound classes (Figure 3b). The compound classes contained at least one D-desthiobiotin moiety (code 2 = 1 or code 3 = 1), which is known to bind to streptavidin in the nanomolar range.[33] The significant enrichment of the two compound classes (*p*-value <0.05, i.e., counts above 40, average count of a library member was 1.4) (Figure 4b) served as a quality control for library synthesis (reaction step 2 and 3), selection conditions and for Illumina-based decoding. The sequence count profile for a library selection on empty resin, which is depicted in Figure 3c, was similar to the profile of the unselected library, except for a slight enrichment of DNA-conjugates with code 3 = 20 (i.e., 7 hits with *p*-values <0.05, which means more than 108 counts, average count of a DNA-conjugate was 4.0) (Figure 4c). This observation confirmed that false-positives (i.e., molecules interacting with the solid support) were not frequently found in the experimental conditions used for selections and can be easily identified by performing control selections.

Next, we performed an affinity-based selection against carbonic anhydrase IX (CA IX). The sequence count profile of the CA IX selection showed a fingerprint of enriched DNA-conjugates (Figure 3d) (i.e., 81 hits with *p*-values <0.05, which means more than 28 counts, average count of a library was 0.76). In detail, 73 out of 81 hit structures, identified in the CA IX selection

dx.doi.org/10.1021/cb1003477 |*ACS Chem. Biol.* 2011, 6, 336–344

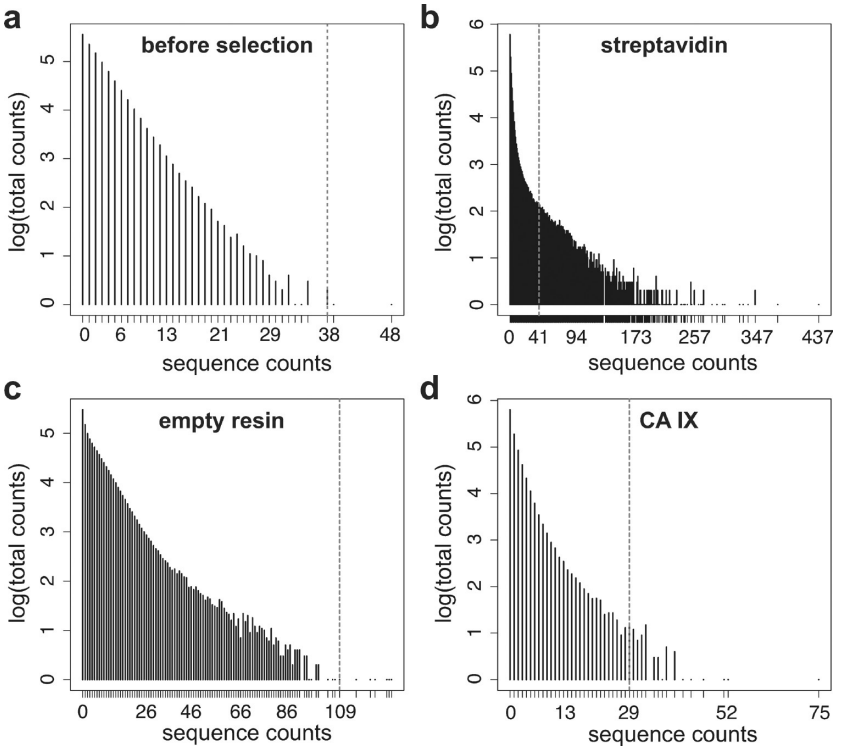
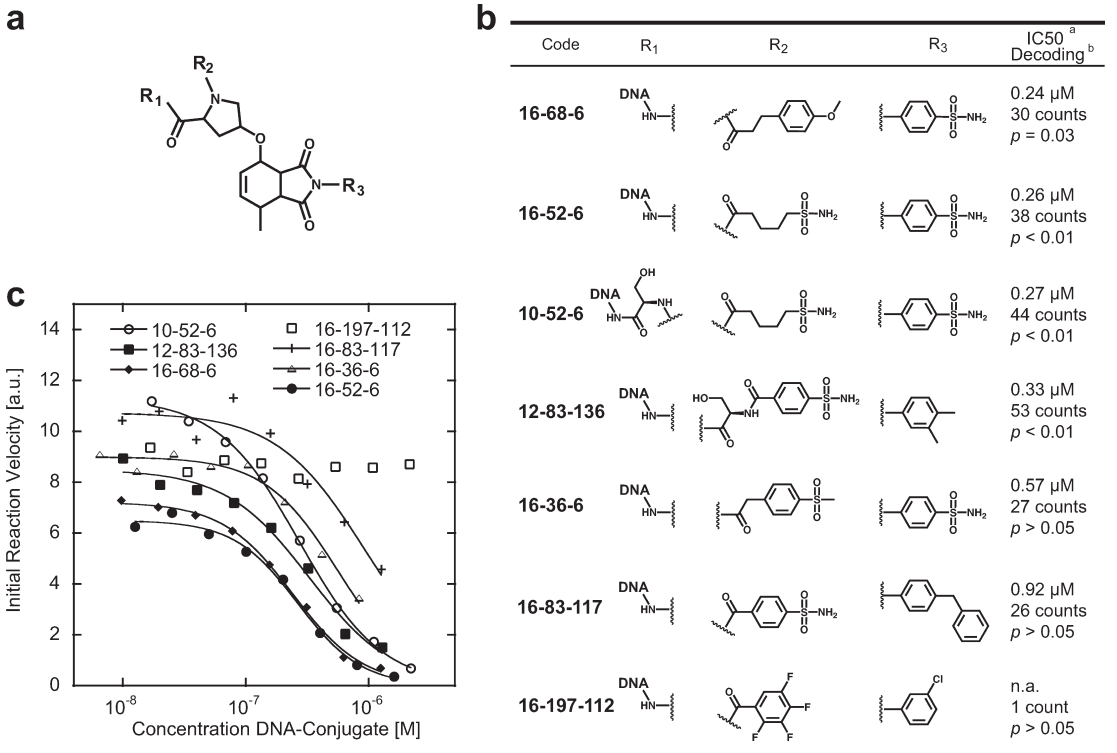


Figure 4. Statistical analysis of the high-throughput sequencing decoding. Logged histogram (log10) of the absolute sequence counts. Vertical lines (dashed) show the cutoff for significant values according to *p*-values with FWER smaller than 0.05 based on robust parameter estimation for a negative binomial density distribution and correction for multiple testing. (a) Amplification of the library before selection; (b) selection against streptavidin; (c) selection against empty resin; d: selection against carbonic anhydrase IX.



Code	R_1	R_2	R_3	IC50 [a] Decoding [b]
16-68-6	DNA			0.24 μM 30 counts $p = 0.03$
16-52-6	DNA			0.26 μM 38 counts $p < 0.01$
10-52-6	DNA			0.27 μM 44 counts $p < 0.01$
12-83-136	DNA			0.33 μM 53 counts $p < 0.01$
16-36-6	DNA			0.57 μM 27 counts $p > 0.05$
16-83-117	DNA			0.92 μM 26 counts $p > 0.05$
16-197-112	DNA			n.a. 1 count $p > 0.05$

Figure 5. DNA-based hit-validation of the CA IX selection. (a) General core structure; (b) table of structures with inhibition data [a] and selection results [b]; (c) inhibition of CA IX with DNA-conjugates was measured in an esterase assay using *p*-nitrophenyl acetate as substrate.

(Figure 3d), shared the common structural feature of a sulfonamide group with code 2 = 52, code 2 = 83 or code 3 = 6 (8 hits shared building blocks also found in the selection against empty resin and were not investigated further). Notably, six of the hit

dx.doi.org/10.1021/cb1003477 |*ACS Chem. Biol.* 2011, 6, 336–344

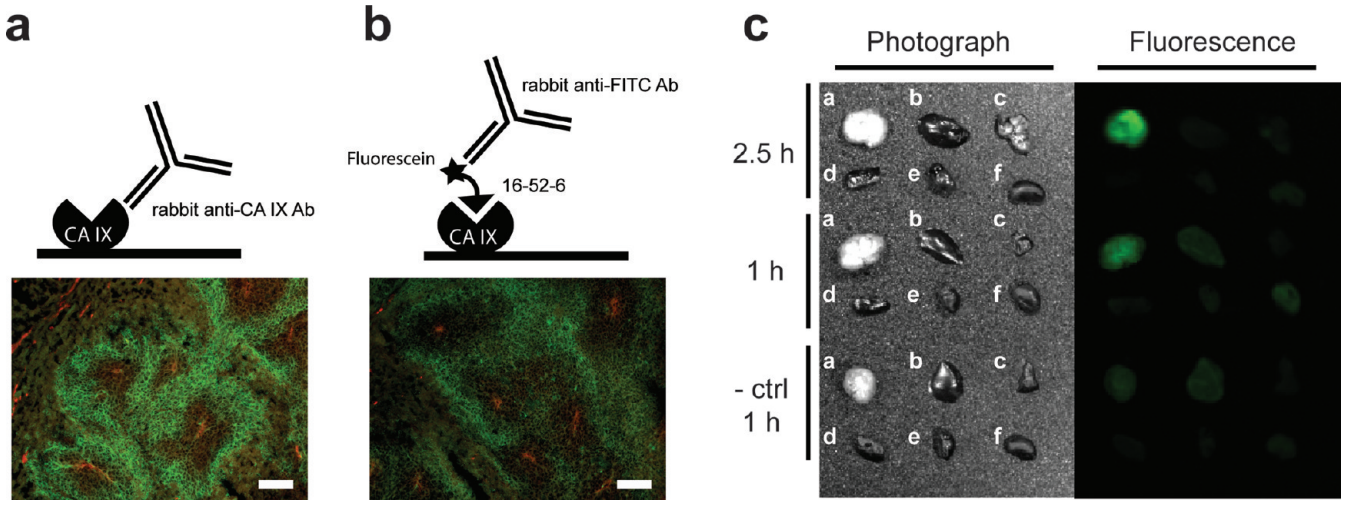


Figure 6. (a, b) Immunofluorescence analysis performed on human colorectal adenocarcinoma LS174T xenografted tumor tissue sections. Red staining represents endothelial cells (anti-CD31 staining), whereas green staining represents expression of CA IX detected by (a) polyclonal anti-CA IX antibody or (b) a complex of inhibitor fluorescein-16-52-6 and anti-fluorescein polyclonal IgG. Scale bars, 100 μm. (c) Fluorescence imaging of isolated organs performed 1 or 2.5 h after iv injection of 68 nmol of fluorescein-16-52-6 or 68 nmol of a fluorescein derivative without CA IX ligand (negative control) in nude mice bearing LS174T tumors. Fluorescence intensity was measured at λ_{ex} = 465 nm, λ_{em} = 520 nm with an exposure time of 1 s using an IVIS Spectrum (Xenogen, Caliper Life Sciences) instrument and showed *in vivo* tumor accumulation of the compound fluorescein-16-52-6: a) tumor, b) liver, c) lung, d) spleen, e) heart, f) kidney.

sequences encoded for bis(sulfonamide) structures with one aliphatic sulfonamide moiety (code 2=52) and one phenylsulfonamide moiety (code 3 = 6). Figure 5 shows structures of DNA-conjugates selected for further hit-validation.

Characterization of CA IX Inhibitors. Hit structures from library selections are typically resynthesized as individual compounds devoid of DNA and used in biochemical assays.[10,12,21] Considering that the affinity-based screening with libraries of three or more chemical building blocks may lead to hundreds of enriched compounds, larger efforts are required for conventional resynthesis of hit-compounds. Here, we synthesized compounds enriched in the CA IX selection (Figure 5a and b) as molecules tagged with a short DNA fragment. The DNA-based synthesis facilitated purification steps (HPLC and ethanol precipitation), quantification by absorbance measurements and analytical characterization by ESI-MS. However, we could not analyze the stereochemistry of the Diels−Alder products when coupled to DNA. Investigations of similar Diels−Alder reactions in the absence of the DNA-tag showed that *endo* diastereomers may be preferentially formed.[34] Figure 5c depicts carbonic anhydrase inhibition data of selected DNA-conjugates based on an esterase activity assay with p-nitrophenyl acetate as the substrate.[35,36] The compounds enriched in the selection experiments showed inhibition of CA IX with an IC50 in the range of 0.24 μM − 0.33 μM ($p < 0.05$), while a nonenriched library member (16−197−112, selection count =1) showed no detectable inhibition (IC50 > 10 μM) (Figure 5c). Selected compounds with *p*-values <0.05 tended to have smaller inhibition constants compared to less enriched compounds (*e.g.*, compound 16−36−6 and 16−83−117). However, a 1:1 correlation of sequence counts and inhibition constants was not expected because of (i) statistical reasons, (ii) variations of synthesis yields for different members present in the library and (iii) selection conditions used (*e.g.*, higher selection pressure might lead to a better discrimination of binders in micromolar to nanomolar range). These findings were in good agreement with observations previously made with a number of smaller DNA-encoded chemical libraries.[6,10,12,19]

One of the CA IX inhibitors (16-52-6; see also Supporting Figure 7) was selected for further hit-validation. As several bis(sulfonamide) compounds were highly enriched in the selection experiment, we wanted to explore whether this novel compound class could exhibit beneficial properties for CA IX tumor targeting. Immunofluorescence histology analysis is a technique commonly used in the antibody field to verify specific antigen binding. Figure 6a shows immunofluorescent staining of colorectal LS174T adenocarcinoma tumor sections with a commercially available antibody specific to carbonic anhydrase IX. CA IX overexpression is linked to cells under hypoxic conditions at a distance of 100 to 200 μm from the nearest oxygen-supplying blood vessel and commonly observed in solid tumors.[37−39] A distinct staining pattern (shown in green, detected with a fluorescently labeled secondary antibody) was observed at a distance of ∼100 μm from the tumor blood vessels (shown in red, detected with a primary antibody against CD31). We used the small molecule ligand 16-52-6, specific to CA IX, as a fluorescein conjugate, in order to study its binding properties on the tumor section, upon addition of an anti-fluorescein secondary antibody for fluorescence detection (shown in green). The fluorescein moiety was introduced via a triethylene glycol linker instead of the DNA-tag, and the conjugate showed submicromolar inhibition of CA IX (IC50 = 0.65 μM; K_d = 0.60 μM; Supporting Figure 8). In close analogy to the use of an antibody specific to CA IX as primary detection reagent (Figure 6a), the fluorescein-16-52-6 conjugate clearly delineated hypoxic regions around tumor blood vessels (Figure 6b), confirming the ability of the CA IX ligand to selectively recognize the cognate antigen in tumor tissue.

Based on the encouraging staining results obtained with fluorescein-16-52-6 on tumor sections, we studied the *in vivo* tumor targeting performance of this compound in nude mice bearing subcutaneous LS174T tumors (tumor volume 500 mm^3). Fluorescein-16-52-6 (68 nmol) or a fluorescein-derivative without CA IX ligand as negative control (68 nmol) were injected intravenously into mice, which were then sacrificed after

dx.doi.org/10.1021/cb1003477 |*ACS Chem. Biol.* 2011, 6, 336–344

1 or 2.5 h. Individual organs and tumors were analyzed by fluorescence measurements (Figure 6c). A strong tumor accumulation of fluorescein-16-52-6 was observed 1 and 2.5 h after injection, while the negative-control fluorescein conjugate exhibited no preferential accumulation in the neoplastic site (Figure 6c). Similar results were obtained *in vivo* using mice bearing the renal cell carcinoma xenograft SK-RC-52 (Supporting Figure 9).

Carbonic anhydrase IX has recently emerged as a potential target for cancer therapy.[26] RNAi experiments indicated that the simultaneous blockade of CA IX and CA XII leads to tumor growth retardation, as a result of impaired acidification of the tumor environment.[39] While classical inhibitors of carbonic anhydrases such as acetazolamide are clinically used as antiglaucoma drugs, the development of selective carbonic anhydrase inhibitors to the membrane bound isozymes CA IX and XII has been a matter of intense medicinal chemical research.[37] Recently, anti-cancer strategies including the development of bioreductive prodrugs,[40] membrane-impermeable glycoconjugates [41] or positively charged CA inhibitors,[42] and the covalent linkage of acetazolamide to an albumin-binding moiety [36] have been proposed to direct binding of the inhibitor to membrane bound carbonic anhydrase IX and XII. First *in vivo* data monitoring hypoxia response with fluorescently labeled sulfonamides in HT-29 colorectal tumors have been reported,[43] and growth retardation in an SK-RC-52 tumor model with an albumin-binding acetazolamide derivative has been shown.[36] At a clinical level, CA IX targeting is currently being investigated with WX-G250, an IgG1 chimeric monoclonal antibody specific to CA IX, for the treatment of renal cell carcinoma.[44,45]

The selection of bis(sulfonamide) compounds indicates the simultaneous engagement of both sulfonamide moieties in the binding event. A small molecule inhibitor with two sulfonamide moieties has been previously studied with a different carbonic anhydrase (CA II), and crystallographic evidence revealed that one sulfonamide engaged zinc in the active site, while a second sulfonamide moiety interacted with residues at the entrance to the active site cleft.[46] The recently solved crystal structure of CA IX reveals the homodimeric nature of this protein,[47] which is consitent with size-exclusion chromatographic data.[48] The two active sites in the CA IX homodimer are located at a distance of approximately 30 Å, which makes bridging of the two active sites by a bis(sulfonamide) inhibitor virtually impossible. However, polar sites for a productive interaction with the second sulfonamide moiety can be found in close proximity to the entrance of the active site.[47]

The compound isolated from the DNA-encoded chemical library showed specific tumor accumulation in two different tumor models. Tumor pretargeting studies with a bispecific monoclonal antibody with high-affinity for CA IX and a radiolabeled peptide showed excellent tumor:organ ratios over a range of time points after injection.[49] Additional *in vivo* targeting experiments are required to investigate whether sulfonamide-based small molecule inhibitors can also be used to deliver toxic agents specifically to the tumor site.

In summary, we have described the design, synthesis and characterization of a DNA-encoded chemical library containing one million compounds coupled to DNA fragments, which serve as amplifiable identification bar codes. High-throughput sequencing of the library before and after selection showed the feasibility of selecting binding molecules out of a large pool of encoded compounds. The library was used to select ligands to carbonic anhydrase IX, and bis(sulfonamide) binders could be isolated, which exhibited binding constants in the submicromolar range and selective CA IX targeting in hypoxic tumors *in vivo*. Taken together, our large DNA-encoded chemical library promises its general applicability for the isolation of ligands against pharmaceutically relevant target proteins.

METHODS

DNA-Encoded Chemical Library Synthesis. Detailed synthesis procedures for the construction of the one million member DNA-encoded compound library can be found in the Supporting Information. In short, the library was constructed by the assembly of three chemical moieties in a split-and-pool synthesis: (i) 25 DNA-conjugated scaffold molecules containing an amino group and a (2*E*,4*E*)-hexadiene moiety, (ii) a set of carboxylic acids, acyl chlorides, carboxylic acid anhydrides, and isothiocyanates that were reacted with the amine function, and (iii) a set of maleimides and maleimides derived from their corresponding amines that were reacted in a Diels−Alder cycloaddition with the (2*E*,4*E*)-hexadiene moiety. The parallel extension of the DNA-tag after each chemical reaction enabled the unique encoding of each chemical structure present in the library.

Tumor Mouse Models. The human colorectal adenocarcinoma cell line LS174T (CL-188, ATCC) was maintained in DMEM, supplemented with 10% fetal bovine serum (FBS) and antibiotic-antimycotic solution at 37 °C in an atmosphere of 5% CO_2. The human renal cell carinoma (RCC) cell line SK-RC-52[50] was maintained in RPMI medium, supplemented as described above. Tumor cells (1×10^7) were injected subcutaneously into the left flank of 8−10 week old female Balb/c nu/nu mice (Charles River). Xenografted tumors were allowed to grow for 20 days (LS174T) or 17 days (SK-RC-52) to a size of typically 500 mg (LS174T) or 400 mg (SK-RC-52) for imaging applications.

Immunofluorescence on Frozen Tissue Sections. LS174T human colorectal adenocarcinoma xenografted tumor tissue was prepared as described previously.[48] Tissue sections (10 μm) were first fixed for 8 min in ice-cold acetone, rehydrated with PBS, and saturated with PBS containing 20% FCS. Sections were double stained for CA IX and CD31. Sections were incubated for 1 h with a complex of the fluorescein-labeled small molecule ligand 16-52-6 (fluorescein-16-52-6, 25 nM; for structure and synthesis see Supporting Information) and rabbit polyclonal anti-FITC-HRP IgG (10 μg/mL, no. 4510-7864, AbD Serotec). Primary rabbit polyclonal anti-CA IX antiserum (sc-25599; Santa Cruz Biotechnolgy) was used for staining a positive control section. CA IX staining was detected using goat anti-rabbit Alexa Fluor 594 (Invitrogen). Primary rat anti-mouse CD31 antibody (BD Biosciences) was detected with donkey anti-rat Alexa Fluor 488 (Invitrogen). Commercial binding reagents were diluted according to the manufacturer's recommendation. Rinsing with PBS was performed between all incubation steps. Finally, slides were mounted with fluorescence mounting medium (DAKO) and analyzed with a Zeiss Axioskop 2 mot fluorescence microscope (Carl Zeiss AG). Images were captured with an AxioCam MRC using AxioVision 4.7 image analysis software (Carl Zeiss AG).

Fluorescence Imaging. LS174T or SK-RC-52 tumor-bearing mice were injected iv into the tail vein with 68 nmol (80 μg) of fluorescein-16-52-6 (Supporting Information) at 1 h (LS174T and SK-RC-52) and 2.5 h (LS174T) prior to sacrifice. For comparison, LS174T or SK-RC-52 tumor-bearing mice were injected iv into the tail vein with 68 nmol (39 μg) amino-tri(ethylene glycol)-fluorescein (Supporting Information) at 1 h prior to sacrifice. CA IX targeting was analyzed with IVIS Spectrum (Xenogen, Caliper Life Sciences) either by imaging the whole animal or isolated organs with λ_{ex} = 465 nm, λ_{em} = 520 nm, exposure time = 1 s, f/stop = 1 and medium binning. For image analysis, Living Image software (Caliper Life Sciences) was used. Experiments were performed under a project license granted by the Veterinäramt des Kantons Zürich, Switzerland (169/2008).

■ ASSOCIATED CONTENT

Supporting Information. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*Tel: (+41)-44-6337401. E-mail: neri@pharma.ethz.ch.

■ ACKNOWLEDGMENT

We thank K. Zuberbühler for help with protein purification, Dr. K. Schwager for help with animal experiments and Dr. V. Mumprecht for help with fluorescence imaging experiments. 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) and the Gebert Rüf Foundation is gratefully acknowledged. The authors are grateful to the Department of Biosystems Science and Engineering at ETH Zurich for help with high-throughput sequencing technology, and to Dr. L. Mannocci (Philochem AG) for reviewing the manuscript.

■ REFERENCES

(1) Buller, F., Mannocci, L., Scheuermann, J., and Neri, D. (2010) Drug discovery with DNA-encoded chemical libraries. *Bioconjugate Chem. 21*, 1571–1580.

(2) Clark, M. A. (2010) Selecting chemicals: the emerging utility of DNA-encoded libraries. *Curr. Opin. Chem. Biol. 14*, 396–403.

(3) Scheuermann, J., and Neri, D. (2010) DNA-encoded chemical libraries: a tool for drug discovery and for chemical biology. *ChemBioChem 11*, 931–937.

(4) McCafferty, J., Griffiths, A. D., Winter, G., and Chiswell, D. J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 348*, 552–554.

(5) 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.

(6) 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. U.S.A. 105*, 17670–17675.

(7) Buller, F., Steiner, M., Scheuermann, J., Mannocci, L., Nissen, I., Kohler, M., Beisel, C., and Neri, D. (2010) High-throughput sequencing for the identification of binding molecules from DNA-encoded chemical libraries. *Bioorg. Med. Chem. Lett. 20*, 4188–4192.

(8) Mayr, L., and Bojanic, D. (2009) Novel trends in high-throughput screening. *Curr. Opin. Pharmacol. 9*, 580–588.

(9) Pellecchia, M., Bertini, I., Cowburn, D., Dalvit, C., Giralt, E., Jahnke, W., James, T. L., Homans, S. W., Kessler, H., Luchinat, C., Meyer, B., Oschkinat, H., Peng, J., Schwalbe, H., and Siegal, G. (2008) Perspectives on NMR in drug discovery: a technique comes of age. *Nat. Rev. Drug Discovery 7*, 738–745.

(10) Buller, F., Zhang, Y., Scheuermann, J., Schäfer, J., Bühlmann, P., and Neri, D. (2009) Discovery of TNF inhibitors from a DNA-encoded chemical library based on diels-alder cycloaddition. *Chem. Biol. 16*, 1075–1086.

(11) Wrenn, S. J., Weisinger, R. M., Halpin, D. R., and Harbury, P. B. (2007) Synthetic ligands discovered by in vitro selection. *J. Am. Chem. Soc. 129*, 13137–13143.

(12) Melkko, S., Mannocci, L., Dumelin, C. E., Villa, A., Sommavilla, R., Zhang, Y., Grutter, M. G., Keller, N., Jermutus, L., Jackson, R. H., Scheuermann, J., and Neri, D. (2010) Isolation of a small-molecule inhibitor of the antiapoptotic protein Bcl-xL from a DNA-encoded chemical library. *ChemMedChem 5*, 584–590.

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

(14) Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K., and Liu, D. R. (2010) In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J. Am. Chem. Soc. 132*, 11779–11791.

(15) Tse, B. N., Snyder, T. M., Shen, Y., and Liu, D. R. (2008) Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J. Am. Chem. Soc. 130*, 15611–15626.

(16) Scheuermann, J., Dumelin, C., 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. *Bioconjugate Chem. 19*, 778–785.

(17) 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.

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

(19) Mannocci, L., Melkko, S., Buller, F., Molnar, I., Bianke, J. P., Dumelin, C. E., Scheuermann, J., and Neri, D. (2010) Isolation of potent and specific trypsin inhibitors from a DNA-encoded chemical library. *Bioconjugate Chem. 21*, 1836–1841.

(20) Dumelin, C. E., Trussel, S., Buller, F., Trachsel, E., Bootz, F., Zhang, Y., Mannocci, L., Beck, S. C., Drumea-Mirancea, M., Seeliger, M. W., Baltes, C., Muggler, T., Kranz, F., Rudin, M., Melkko, S., Scheuermann, J., and Neri, D. (2008) A portable albumin binder from a DNA-encoded chemical library. *Angew. Chem., Int. Ed. Engl. 47*, 3196–3201.

(21) 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., Davie, C. P., Ding, Y., Franklin, G. J., Franzen, K. D., Gefter, M. L., Hale, S. P., Hansen, N. J., Israel, D. I., Jiang, J., Kavarana, M. J., Kelley, M. S., Kollmann, C. S., Li, F., Lind, K., Mataruse, S., Medeiros, P. F., Messer, J. A., Myers, P., O'Keefe, H., Oliff, M. C., Rise, C. E., Satz, A. L., Skinner, S. R., Svendsen, J. L., Tang, L., van Vloten, K., Wagner, R. W., Yao, G., Zhao, B., and Morgan, B. A. (2009) Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat. Chem. Biol. 5*, 647–654.

(22) 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.

(23) 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.

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

(25) Dumelin, C., Trüssel, S., Buller, F., Trachsel, E., Bootz, F., Zhang, Y., Mannocci, L., Beck, S., Drumea-Mirancea, M., Seeliger, M., Baltes, C., Müggler, T., Kranz, F., Rudin, M., Melkko, S., Scheuermann, J., and Neri, D. (2008) A portable albumin binder from a DNA-encoded chemical library. *Angew. Chem., Int. Ed. Engl. 47*, 3196–3201.

(26) De Simone, G., and Supuran, C. T. (2010) Carbonic anhydrase IX: Biochemical and crystallographic characterization of a novel anti-tumor target. *Biochim. Biophys. Acta 1804*, 404–409.

(27) Lipinski, C. A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods 44*, 235–249.

(28) Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev. 46*, 3–26.

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

(30) Hoogenboom, H. R. (2005) Selecting and screening recombinant antibody libraries. *Nat. Biotechnol. 23*, 1105–1116.

(31) Marazzi, A., and Yohai, V. (2010) Optimal robust estimates using the Hellinger distance. *Adv. Data Anal. Classification 4*, 169–179.

(32) Holm, S. (1979) A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Statist. 6*, 65–70.

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

(34) de Araujo, A. D., Palomo, J. M., Cramer, J., Seitz, O., Alexandrov, K., and Waldmann, H. (2006) Diels-Alder ligation of peptides and proteins. *Chem.—Eur. J. 12*, 6095–6109.

(35) Pocker, Y., and Stone, J. T. (1967) The catalytic versatility of erythrocyte carbonic anhydrase. 3. Kinetic studies of the enzyme-catalyzed hydrolysis of p-nitrophenyl acetate. *Biochemistry (Moscow) 6*, 668–678.

(36) Ahlskog, J. K., Dumelin, C. E., Trussel, S., Marlind, J., and Neri, D. (2009) In vivo targeting of tumor-associated carbonic anhydrases using acetazolamide derivatives. *Bioorg. Med. Chem. Lett. 19*, 4851–4856.

(37) Supuran, C. T. (2008) Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nat. Rev. Drug Discovery 7*, 168–181.

(38) Brahimi-Horn, M. C., Chiche, J., and Pouyssegur, J. (2007) Hypoxia and cancer. *J. Mol. Med. 85*, 1301–1307.

(39) Chiche, J., Ilc, K., Laferriere, J., Trottier, E., Dayan, F., Mazure, N. M., Brahimi-Horn, M. C., and Pouyssegur, J. (2009) Hypoxia-inducible carbonic anhydrase IX and XII promote tumor cell growth by counteracting acidosis through the regulation of the intracellular pH. *Cancer Res. 69*, 358–368.

(40) De Simone, G., Vitale, R. M., Di Fiore, A., Pedone, C., Scozzafava, A., Montero, J. L., Winum, J. Y., and Supuran, C. T. (2006) Carbonic anhydrase inhibitors: hypoxia-activatable sulfonamides incorporating disulfide bonds that target the tumor-associated isoform IX. *J. Med. Chem. 49*, 5544–5551.

(41) Winum, J. Y., Poulsen, S. A., and Supuran, C. T. (2009) Therapeutic applications of glycosidic carbonic anhydrase inhibitors. *Med. Res. Rev. 29*, 419–435.

(42) Pastorekova, S., Casini, A., Scozzafava, A., Vullo, D., Pastorek, J., and Supuran, C. T. (2004) Carbonic anhydrase inhibitors: the first selective, membrane-impermeant inhibitors targeting the tumor-associated isozyme IX. *Bioorg. Med. Chem. Lett. 14*, 869–873.

(43) Dubois, L., Lieuwes, N. G., Maresca, A., Thiry, A., Supuran, C. T., Scozzafava, A., Wouters, B. G., and Lambin, P. (2009) Imaging of CA IX with fluorescent labelled sulfonamides distinguishes hypoxic and (re)-oxygenated cells in a xenograft tumour model. *Radiother. Oncol. 92*, 423–428.

(44) van Dijk, J., Zegveld, S. T., Fleuren, G. J., and Warnaar, S. O. (1991) Localization of monoclonal antibody G250 and bispecific monoclonal antibody CD3/G250 in human renal-cell carcinoma xenografts: relative effects of size and affinity. *Int. J. Cancer 48*, 738–743.

(45) Bleumer, I., Knuth, A., Oosterwijk, E., Hofmann, R., Varga, Z., Lamers, C., Kruit, W., Melchior, S., Mala, C., Ullrich, S., De Mulder, P., Mulders, P. F., and Beck, J. (2004) A phase II trial of chimeric monoclonal antibody G250 for advanced renal cell carcinoma patients. *Br. J. Cancer 90*, 985–990.

(46) Casini, A., Abbate, F., Scozzafava, A., and Supuran, C. T. (2003) Carbonic anhydrase inhibitors: X-ray crystallographic structure of the adduct of human isozyme II with a bis-sulfonamide-two heads are better than one? *Bioorg. Med. Chem. Lett. 13*, 2759–2763.

(47) Alterio, V., Hilvo, M., Di Fiore, A., Supuran, C. T., Pan, P., Parkkila, S., Scaloni, A., Pastorek, J., Pastorekova, S., Pedone, C., Scozzafava, A., Monti, S. M., and De Simone, G. (2009) Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proc. Natl. Acad. Sci. U.S.A. 106*, 16233–16238.

(48) Ahlskog, J. K., Schliemann, C., Marlind, J., Qureshi, U., Ammar, A., Pedley, R. B., and Neri, D. (2009) Human monoclonal antibodies targeting carbonic anhydrase IX for the molecular imaging of hypoxic regions in solid tumours. *Br. J. Cancer 101*, 645–657.

(49) van Schaijk, F. G., Oosterwijk, E., Molkenboer-Kuenen, J. D., Soede, A. C., McBride, B. J., Goldenberg, D. M., Oyen, W. J., Corstens, F. H., and Boerman, O. C. (2005) Pretargeting with bispecific anti-renal cell carcinoma x anti-DTPA(In) antibody in 3 RCC models. *J. Nucl. Med. 46*, 495–501.

(50) Ebert, T., Bander, N. H., Finstad, C. L., Ramsawak, R. D., and Old, L. J. (1990) Establishment and characterization of human renal cancer and normal kidney cell lines. *Cancer Res. 50*, 5531–5536.