ChemPubSoc Europe

DOI: 10.1002/cmdc.201800193

CHEMMEDCHEM Communications

A DNA-Encoded Library of Chemical Compounds Based on Common Scaffolding Structures Reveals the Impact of Ligand Geometry on Protein Recognition

Nicholas Favalli,[a] Stefan Biendl,[a] Marco Hartmann,[a] Jacopo Piazzi,[b] Filippo Sladojevich,[c] Susanne Gräslund,[d, e] Peter J. Brown,[d] Katja Näreoja,[e] Herwig Schüler,[e] Jörg Scheuermann,[a] Raphael Franzini,*[a, f] and Dario Neri*[a]

A DNA-encoded chemical library (DECL) with 1.2 million compounds was synthesized by combinatorial reaction of seven central scaffolds with two sets of 343×492 building blocks. Library screening by affinity capture revealed that for some target proteins, the chemical nature of building blocks dominated the selection results, whereas for other proteins, the central scaffold also crucially contributed to ligand affinity. Molecules based on a 3,5-bis(aminomethyl)benzoic acid core structure were found to bind human serum albumin with a K_d value of 6 nM, while compounds with the same substituents on an equidistant but flexible L-lysine scaffold showed 140-fold lower affinity. A 18 nM tankyrase-1 binder featured L-lysine as linking moiety, while molecules based on D-Lysine or (2*S*,4*S*)-amino-L-proline showed no detectable binding to the target. This work suggests that central scaffolds which predispose the orientation of chemical building blocks toward the protein target may enhance the screening productivity of encoded libraries.

The encoding of compounds with DNA tags, serving as amplifiable identification barcodes, allows the facile construction and screening of large combinatorial chemical libraries.[1] The successful identification of binding molecules from DNA-encoded chemical libraries depends not only on the number and characteristics of the building blocks used for library construction, but also on library design.[1] We previously reported a DNA-encoded chemical library based on a conserved (*S*)-2,3-diaminopropanoic acid scaffold, which had been combinatorially reacted with two sets of carboxylic acids that yielded nanomolar binders against various target proteins, including serum albumins and tankyrase-1.[2] These results motivated us to investigate how the geometry, stereochemistry, and rigidity of a central scaffold influences the outcome of screening results and the binding affinity of selected compounds.[3] Herein we describe the synthesis and characterization of a DNA-encoded chemical library[4] obtained by the combinatorial modification of seven central scaffolds, each bearing both an amine and an azide moiety. The use of these orthogonal coupling sites for two diverse sets of building blocks (343 and 492 building blocks, respectively), eventually yielded a library with an overall size of 1.2 million compounds. Affinity measurements performed on selected hit compounds indicate that, for certain targets, the geometry and rigidity of the central scaffold can have a strong impact on the dissociation constants of the corresponding ligands.[5]

In this study, we explored whether subtle chemical variations of the central molecular "scaffold" impact the affinity and specificity of the discovered protein ligands. For this reason, we coupled seven trifunctional carboxylic acid derivatives to oligonucleotides, to enable the subsequent coupling and encoding of two sets of building blocks, resulting in a combinatorial library of $7\times343\times492=1181292$ compounds (Figure 1). The scaffolds included two stereo-defined protected derivatives of (*R*)-2-azido-3-aminopropionic acid (**1**) and (*S*)-2-azido-3-aminopropionic acid (**2**), the rigidified (2*S*,4*S*)-azido-L-proline (**5**) and (2*S*,4*R*)-azido-L-proline (**6**), ε-azido-D-lysine (**3**), ε-azido-L-lysine (**4**) and 3-(aminomethyl)-5-(azidomethyl)benzoic acid (**7**). The scaffolds each contained a protected primary amine group and an azide moiety. Using a split-and-pool protocol,[7] featuring the reaction of amines with carboxylic acids[8] or the copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC)[9] of azide derivatives with terminal alkynes, we constructed a library of structurally related compounds. Screening experiments were performed with biotinylated proteins immobilized on streptavidin-coated magnetic beads.[10] Details of library synthesis and of library encoding procedures can be found in Supporting Information Figure 1.1. Figure 2a shows the results of a library selection performed with biotinylated human serum albumin

[a] N. Favalli, S. Biendl, M. Hartmann, Dr. J. Scheuermann, Prof. Dr. R. Franzini, Prof. Dr. D. Neri
Institute of Pharmaceutical Sciences, ETH Zürich, Vladimir-Prelog-Weg 4, 8093 Zürich (Switzerland)
E-mail: neri@pharma.ethz.ch
raphael.franzini@utah.edu

[b] Dr. J. Piazzi
Philochem AG, Liebernstrasse 3, 8112 Otelfingen (Switzerland)

[c] Dr. F. Sladojevich
Roche Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La, Roche Ltd., Grenzacherstrasse 124, 4070 Basel (Switzerland)

[d] Dr. S. Gräslund, Dr. P. J. Brown
Structural Genomics Consortium (SGC), University of Toronto, Toronto, ON, M5G 1L7 (Canada)

[e] Dr. S. Gräslund, Dr. K. Näreoja, Prof. Dr. H. Schüler
Department of Structural Biology, Department of Medical Biochemistry and Biophysics (MBB), Karolinska Institutet, Scheeles väg 2, 17177 Stockholm (Sweden)

[f] Prof. Dr. R. Franzini
College of Pharmacy, University of Utah, 30 South 2000 East, Salt Lake City, UT 84112 (USA)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/cmdc.201800193.

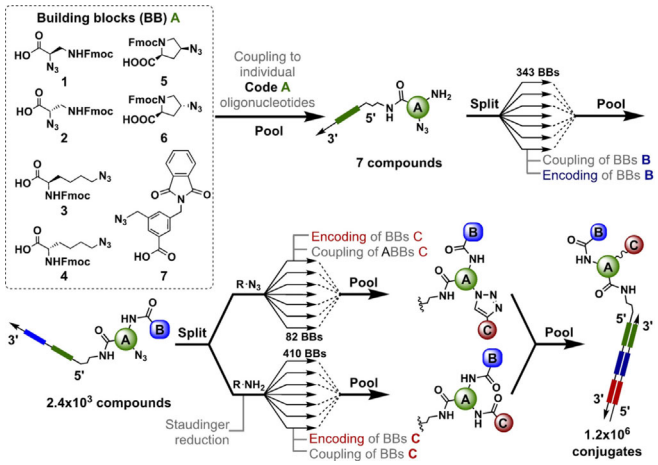


Figure 1. Representation of the library synthesis scheme. Seven building blocks (**1**–**7**), carrying protected amine functions and azides were coupled to amino-tagged oligonucleotides, each containing a central sequence ("code"), which unambiguously identifies the corresponding building block. After deprotection and using a split-and-pool procedure, carboxylic acids were coupled to the amines, and the corresponding 343 building blocks "B" encoded by a ligation procedure. In the final synthesis step the azide moiety of the central scaffolds was either reacted with terminal alkynes or converted into a primary amine, allowing the formation of an amide bond with carboxylic acids. The final encoding step for building blocks "C" was performed using partially complementary oligonucleotides and Klenow polymerization, as described.[6]

(HSA), followed by high-throughput sequencing for the identification and relative quantification of the DNA barcodes.[10] These selection results are displayed in pseudo-four-dimensional space, using three dimensions for definition of the identity of the three sets of building blocks (codes A, B and C), while spheres of different colors represent the fourth dimension (corresponding to the number of sequence counts for individual compounds) above a threshold of 3000 counts. The plot indicates that the most highly enriched library members all contained the 3,5-bis(aminomethyl)benzoic acid scaffold (code A = 7), whereas there was no such preference in the preselected library and selections with a set of control proteins used in our laboratory (Supporting Information Figures 6.2 and 6.3). Figure 2b presents a graphical display of the selection counts for the library members, arranged with seven code A structures on the *x*-axis and the 343 × 492 = 168 756 combinations of code B and code C on the *y*-axis. Compound **A7/B66/C292** is the most enriched compound (Figure 2). The compound consisted of the 3,4-di(aminomethyl)benzoic acid scaffold, with 3-(3,4,5-trimethoxyphenyl)propanoic acid (**B66**), and 2,2-difluoro-1,3-benzodioxole-5-carboxylic acid (**C292**). The same 3-(3,4,5-trimethoxyphenyl)propanoic acid found as building block **B66** in the second reaction cycle position of compound **A7/B66/C292**, was also found in another combination. In particular the 3-(3,4,5-trimethoxyphenyl)propanoic acid moiety was found as building block **C56** in the third reaction cycle position of a second highly enriched combination (**A7/B329/C56**). In a different representation of selection results, keeping code B constant (i.e., restricting the analysis to compounds based on the **B66** building block), the preferential en-

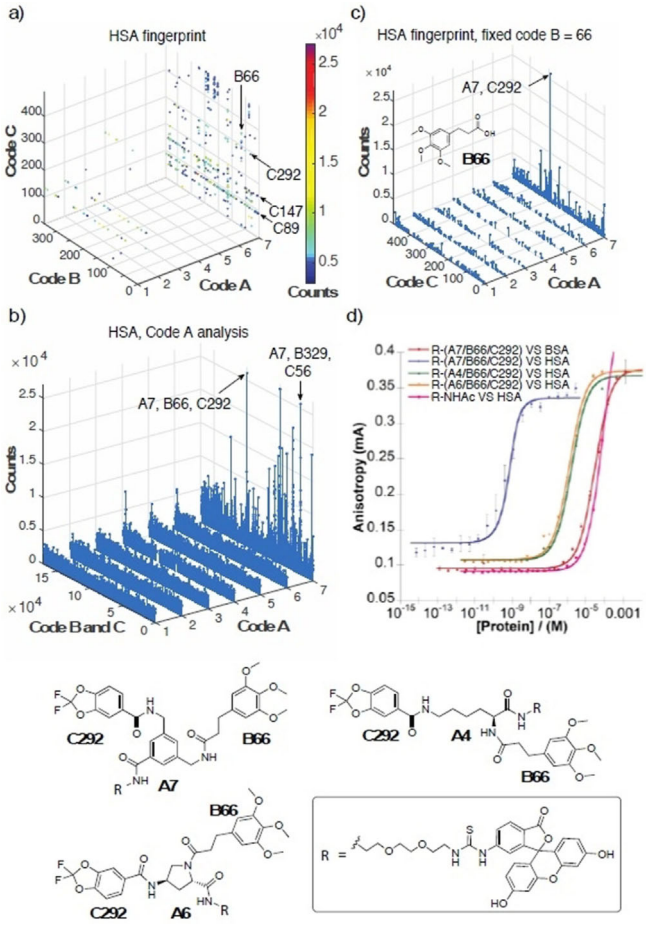


Figure 2. Results of library selections against human serum albumin (HSA). a) Selection fingerprint. The individual library members are unambiguously identified by their codes A (ranging between 1 and 7), B (ranging between 1 and 343) and C (ranging between 1 and 492). The number of sequence counts for each compound is displayed as spheres of a different color, with a cutoff threshold set at 3000 counts. b) The code A analysis shows how the different scaffolds (code A, ranging between 1 and 7) display different enrichment factors for each building block A and B combination (Code B × Code C, ranging between 1 and 1.7×10^5). c) Selection fingerprint obtained fixing the code **B** = 66 (corresponding to 3-(3,4,5-trimethoxyphenyl)propanoic acid) with variable codes A (ranging between 1 and 7) and C (ranging between 1 and 492). This analysis shows **A7/B66/C292** as the most enriched combination of building blocks against HSA. d) Fluorescence polarization (FP) measurement of FITC conjugates of the most enriched combination of building blocks (**B66/C292**), featuring the preferred scaffold **A7** (red and blue curves) or scaffolds **A4** (green) and **A6** (orange). A FITC conjugate of **A7/B66/C292** was also tested against bovine serum albumin (BSA) and revealed a double-digit micromolar dissociation constant (K_d) against that protein (red curve). The acetylated PEG-fluorescein derivative showed no detectable binding to HSA (pink curve).

richment of library members based on the **A7** scaffold (and of the **A7/B66/C292** in particular) is visible (Figure 2c). The dissociation constant (K_d) of a fluorescein derivative of **A7/B66/C292** determined by concentration-dependent fluorescence polarization experiments was found to be 7 nM. In contrast, exchanging the **A7** scaffold by two alternative diamines drastically decreased the affinity for HSA, in agreement with the selection results (K_d(A4/B66/C292) = 1.6 μM, K_d(A6/B66/C292) = 1.2 μM). The fluorophore-labeled and acetylated PEG2-diamino

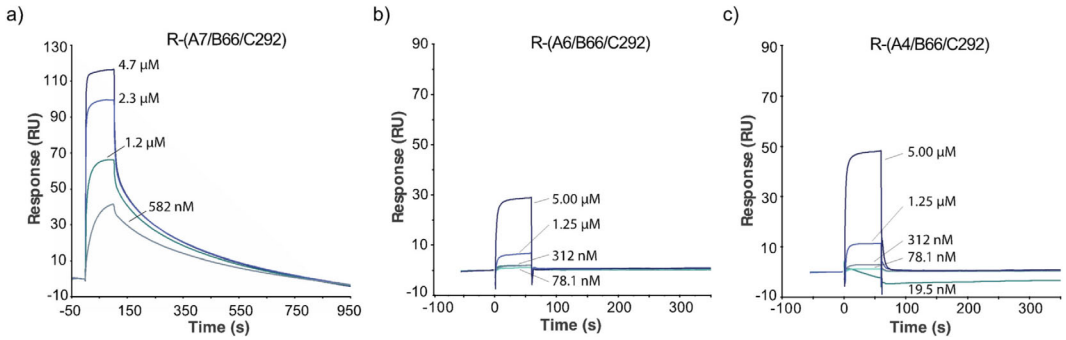


Figure 3. Binding properties of compound **A7/B66/C292** analyzed by surface plasmon resonance (SPR) against HSA immobilized on a BIAcore chip (CM5, 5770 RU). a) SPR profile of compound **R-(A7/B66/C292)** at various concentrations (5 μM, 1.25 μM, 312 nM, 78.1 nM, 19.5 nM). An overall fitting of the resulting curves (4.7 μM, 2.3 μM, 1.2 μM, and 582 nM) revealed a $k_{off} = 7.4 \times 10^{-3}\ s^{-1}$ and a $k_{on} = 2.0 \times 10^{5}\ s^{-1}\ M^{-1}$, resulting in a $K_d = 36.6$ nM b) SPR profile of compound **R-(A6/B66/C292)** at different concentrations (5 μM, 1.25 μM, 312 nM, 78.1 nM, 19.5 nM). c) SPR profile of compound **R-(A4/B66/C292)** at different concentrations (5 μM, 1.25 μM, 312 nM, 78.1 nM, 19.5 nM).

linker exhibited a residual dissociation constant higher than 100 μM (Figure 2d).

We confirmed the fluorescence polarization findings by studying the interaction of compounds **A7/B66/C292**, **A6/B66/C292** and **A4/B66/C292** with HSA using surface plasmon resonance on a BIAcore instrument. Figure 3 shows that compound **A7/B66/C292** bound to its cognate target with a kinetic dissociation constant (k_{off}) of $4.5 \times 10^{-3}\ s^{-1}$. In contrast, the structural analogues featuring L-lysine and (2*S*,4*R*)-amino-L-proline as a central scaffold did not bind to HSA under the same experimental conditions (Figure 3), confirming the role played by the **A7** 3,5-bis(aminomethyl)benzoic acid core structure.

Library selections performed for human tankyrase-1 (TNKS1) revealed a distinctive pattern of enriched compounds, featuring a preference for linker **A4** (L-lysine) and building block **B101** (thymine-1-acetic acid) (Figure 4a). The observed fingerprint was observed when selections were performed at various concentrations of Tween 20, providing confidence about the reproducibility of the screening procedure[1h] (Supporting Information Figure 6.1). A plot of selection results, emphasizing sequence counts for various library members featuring the **B101** building block, highlighted the preferential enrichment of compounds with linker **A4** and with certain conserved structural features of **C** building blocks (Figure 4b,c). Synthesis of amide derivatives the most enriched compound **A4/B101/C491** revealed a high-affinity binding to the cognate TNKS1 protein immobilized on a BIAcore chip ($K_d = 15 \pm 8$ nM), while the use of D-lysine (**A3/B101/C491**) or of (2*S*,4*S*)-amino-L-proline (**A5/B101/C491**) as linker did not result in any detectable binding by SPR analysis (Figure 4d–f).

We also explored the impact on affinity constants to TNKS1 for chemical modifications at the site, originally occupied by the linkage to DNA (Table 1 and Supporting Information Figure 5.3). While a simple amide derivative exhibited the best dissociation constant ($K_d = 15$ nM), the corresponding carboxylic acid or amides derived from 3-aminopropan-1-ol showed a substantial decrease in binding affinity.

When selections were performed with targets for which a specific building block dominates the ligand enrichment procedure, different fingerprint patterns were observed. Figure 5

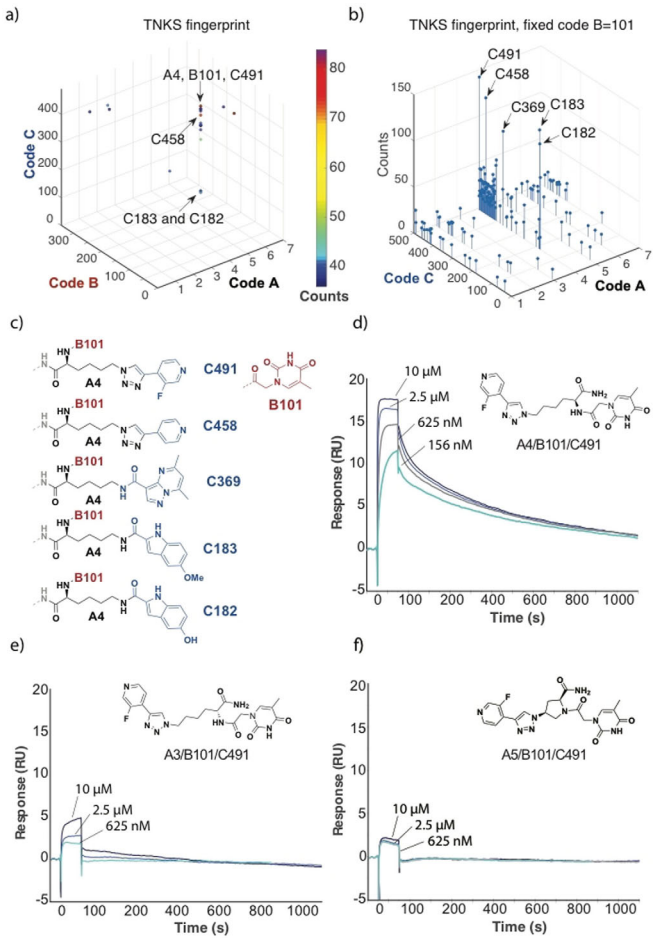
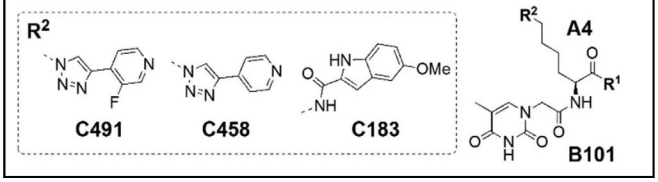


Figure 4. Results of library selections against human tankyrase-1 (TNKS1). a) Selection fingerprint, revealing a preferential enrichment of compounds with **A4** and **B101**. b) Plot of sequence counts for library members, featuring **B101** as preferred building block. The plot reveals a preferential enrichment of library members with linker **A4** and building blocks **C491**, **C453**, **C369**, **C183** and **C182**. The structures of the five most enriched compounds is shown in panel (c). The BIAcore profiles at various concentrations of ligand for the amide derivatives of **A4/B101/C491**, **A3/B101/C491**, and **A5/B101/C491** are shown in panels (d), (e), and (f). Fitting of the sensorgrams for **A4/B101/C491** in panel (d) yielded a $k_{off} = (2.1 \pm 0.8) \times 10^{-3}\ s^{-1}$ and a $k_{on} = (1.7 \pm 0.2) \times 10^{5}\ s^{-1}\ M^{-1}$, corresponding to a $K_d = 15 \pm 8$ nM.

Table 1. Impact of chemical modification at the site of DNA coupling on dissociation constants.

Compound ID[a]	R^1	R^2	K_d[b]
45 (A4/B101/C491)	OH	C491	259 nM
44 (A4/B101/C458)	OH	C458	307 nM
54 (A4/B101/183)	OH	C183	258 nM
46 (A4/B101/C491)	NH_2	C491	15 nM
47 (A4/B101/C491)	$NH(CH_2)_3OH$	C491	6.0 μM
55 (A4/B101/183)	$NH(CH_2)_3OH$	C183	2.4 μM

[a] The complete structures are provided in the Supporting Information. [b] The SPR profiles are shown in Supporting Information Figure 5.3.



shows the results of selections performed against carbonic anhydrase IX, a tumor-associated cell-surface marker.[11] This enzyme can by efficiently targeted by ligands containing aromatic or heteroaromatic sulfonamides.[12] Indeed, an efficient enrichment of building blocks **B128**, **B340** and **C410** was visible in the selection fingerprints. In this case, the role of the central scaffold was less important, as sulfonamide derivatives for all seven building blocks "A" could be efficiently enriched.

Collectively, the results of this study suggest that the chemical nature of central scaffolds, defining the orientation and flexibility of substituents blocks pointing toward the protein target of interest, may represent an important determinant of binding affinity. The examples of HSA and TNKS1 binders revealed that subtle differences in the chemical nature of central scaffolds can lead to substantial variations (i.e., >100-fold differences) in binding affinity for the cognate target protein of interest. The nanomolar binders to HSA and to TNKS1 described herein may be useful for serum half-life prolongation purposes,[13] and the TNKS1 binders as chemical probes for the biological characterization of TNKS1 function,[14] although the selectivity to other poly(ADP-ribose) polymerases remains to be established.[15] We anticipate that new DNA-encoded chemical libraries will be designed in the future, exploiting novel designs for the spatial arrangement of building blocks.

Acknowledgements

Financial support from the ETH Zürich, the Swiss National Science Foundation (Grants 310030B_163479), Sinergia (CRS112_160699), and from the ERC Advanced Grant "ZAUBERKUGEL" (Grant agreement 670603) is gratefully acknowledged. The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) (ULTRA-DD grant no. 115766), Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and the Wellcome Trust.

Conflict of interest

D.N. is a cofounder and shareholder of Philochem AG, and J.S. is a board member of Philochem AG.

Keywords: bifunctional ligands · DNA-encoded chemical libraries · human serum albumin · tankyrase-1

[1] a) S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381–5383; b) C. Arico-Muendel, Z. Zhu, H. Dickson, D. Parks, J. Keicher, J. Deng, L. Aquilani, F. Coppo, T. Graybill, K. Lind, A. Peat, M. Thomson, *Antimicrob. Agents Chemother.* **2015**, *59*, 3450–3459; c) N. Winssinger, *Chimia* **2013**, *67*, 340–348; d) P. Blakskjaer, T. Heitner, N. J. Hansen, *Curr. Opin. Chem. Biol.* **2015**, *26*, 62–71; e) J. W. Cuozzo, P. A. Centrella, D. Gikunju, S. Habeshian, C. D. Hupp, A. D. Keefe, E. A. Sigel, H. H. Soutter, H. A. Thomson, Y. Zhang, M. A. Clark, *ChemBioChem* **2017**, *18*, 864–871; f) M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri, J. Scheuermann, *Nat. Chem.* **2015**, *7*, 241–249; g) F. Samain, T.

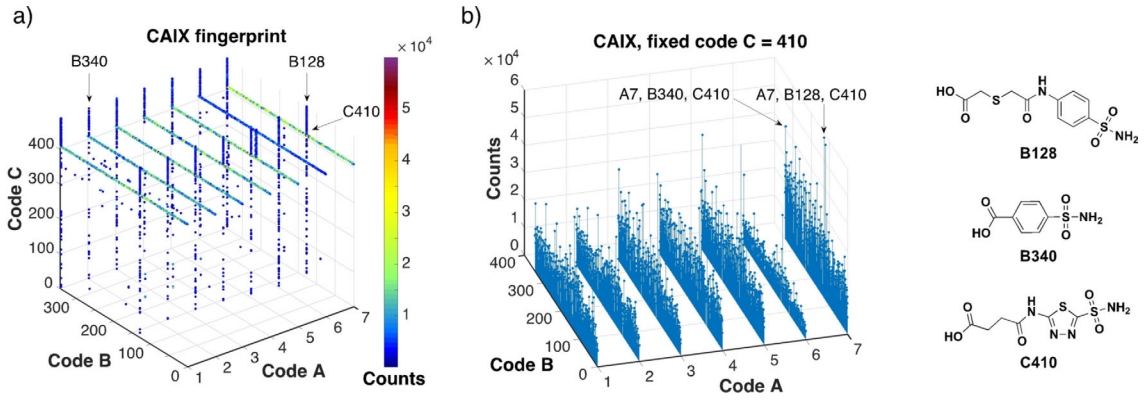


Figure 5. Results of library selections against carbonic anhydrase IX (CAIX). a) Selection fingerprint. The individual library members are unambiguously identified by their codes A (ranging between 1 and 7), B (ranging between 1 and 343) and C (ranging between 1 and 492). The number of sequence counts for each compound is displayed as spheres of different colors, with a cutoff threshold set at 3000 counts. b) Selection fingerprint obtained fixing the code **C** = 410 (corresponding to acetazolamide) with variable codes A (ranging between 1 and 7) and B (ranging between 1 and 343). This picture shows how the scaffold does not affect the affinity for CAIX.

Ekblad, G. Mikutis, N. Zhong, M. Zimmermann, A. Nauer, D. Bajic, W. Decurtins, J. Scheuermann, P. J. Brown, J. Hall, S. Graslund, H. Schuler, D. Neri, R. M. Franzini, *J. Med. Chem.* **2015**, *58*, 5143 – 5149; h) M. Leimbacher, Y. Zhang, L. Mannocci, M. Stravs, T. Geppert, J. Scheuermann, G. Schneider, D. Neri, *Chem. Eur. J.* **2012**, *18*, 7729 – 7737; i) R. M. Franzini, D. Neri, J. Scheuermann, *Acc. Chem. Res.* **2014**, *47*, 1247 – 1255; j) H. Salamon, M. Klika Skopic, K. Jung, O. Bugain, A. Brunschweiger, *ACS Chem. Biol.* **2016**, *11*, 296 – 307; k) D. Neri, R. Lerner, *Annu. Rev. Biochem.* **2018**, *87*, 24.

[2] R. M. Franzini, T. Ekblad, N. Zhong, M. Wichert, W. Decurtins, A. Nauer, M. Zimmermann, F. Samain, J. Scheuermann, P. J. Brown, J. Hall, S. Graslund, H. Schuler, D. Neri, *Angew. Chem. Int. Ed.* **2015**, *54*, 3927 – 3931; *Angew. Chem.* **2015**, *127*, 3999 – 4003.

[3] A. L. Satz, *ACS Comb. Sci.* **2016**, *18*, 415 – 424.

[4] a) A. L. Satz, J. Cai, Y. Chen, R. Goodnow, F. Gruber, A. Kowalczyk, A. Petersen, G. Naderi-Oboodi, L. Orzechowski, Q. Strebel, *Bioconjugate Chem.* **2015**, *26*, 1623 – 1632; b) M. A. Clark, R. A. Acharya, C. C. Arico-Muendel, S. L. Belyanskaya, D. R. Benjamin, N. R. Carlson, P. A. Centrella, C. H. Chiu, S. P. Creaser, J. W. Cuozzo, C. P. Davie, Y. Ding, G. J. Franklin, K. D. Franzen, M. L. Gefter, S. P. Hale, N. J. Hansen, D. I. Israel, J. Jiang, M. J. Kavarana, M. S. Kelley, C. S. Kollmann, F. Li, K. Lind, S. Mataruse, P. F. Medeiros, J. A. Messer, P. Myers, H. O'Keefe, M. C. Oliff, C. E. Rise, A. L. Satz, S. R. Skinner, J. L. Svendsen, L. Tang, K. van Vloten, R. W. Wagner, G. Yao, B. Zhao, B. A. Morgan, *Nat. Chem. Biol.* **2009**, *5*, 647 – 654.

[5] a) J. P. Daguer, C. Zambaldo, M. Ciobanu, P. Morieux, S. Barluenga, N. Winssinger, *Chem. Sci.* **2015**, *6*, 739 – 744; b) H. Deng, J. Zhou, F. S. Sundersingh, J. Summerfield, D. Somers, J. A. Messer, A. L. Satz, N. Ancellin, C. C. Arico-Muendel, K. L. Sargent Bedard, A. Beljean, S. L. Belyanskaya, R. Bingham, S. E. Smith, E. Boursier, P. Carter, P. A. Centrella, M. A. Clark, C. W. Chung, C. P. Davie, J. L. Delorey, Y. Ding, G. J. Franklin, L. C. Grady, K. Herry, C. Hobbs, C. S. Kollmann, B. A. Morgan, L. J. Pothier Kaushansky, Q. Zhou, *ACS Med. Chem. Lett.* **2015**, *6*, 919 – 924; c) L. Encinas, H. O'Keefe, M. Neu, M. J. Remuinan, A. M. Patel, A. Guardia, C. P. Davie, N. Perez-Macias, H. Yang, M. A. Convery, J. A. Messer, E. Perez-Herran, P. A. Centrella, D. Alvarez-Gomez, M. A. Clark, S. Huss, G. K. O'Donovan, F. Ortega-Muro, W. McDowell, P. Castaneda, C. C. Arico-Muendel, S. Pajk, J. Rullas, I. Angulo-Barturen, E. Alvarez-Ruiz, A. Mendoza-Losana, L. Ballell Pages, J. Castro-Pichel, G. Evindar, *J. Med. Chem.* **2014**, *57*, 1276 – 1288.

[6] L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, D. Neri, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17670 – 17675.

[7] R. A. Goodnow, Jr., C. E. Dumelin, A. D. Keefe, *Nat. Rev. Drug Discovery* **2017**, *16*, 131 – 147.

[8] Y. Li, E. Gabriele, F. Samain, N. Favalli, F. Sladojevich, J. Scheuermann, D. Neri, *ACS Comb. Sci.* **2016**, *18*, 438 – 443.

[9] M. M. Haque, X. Peng, *Sci. China Chem.* **2014**, *57*, 215 – 231.

[10] W. Decurtins, M. Wichert, R. M. Franzini, F. Buller, M. A. Stravs, Y. Zhang, D. Neri, J. Scheuermann, *Nat. Protoc.* **2016**, *11*, 764 – 780.

[11] a) S. Muselaers, P. Mulders, E. Oosterwijk, W. Oyen, O. Boerman, *Immunotherapy* **2013**, *5*, 489 – 495; b) C. T. Supuran, *Nat. Rev. Drug Discovery* **2008**, *7*, 168 – 181.

[12] V. Alterio, D. Esposito, S. M. Monti, C. T. Supuran, G. De Simone, *J. Enzyme Inhib. Med. Chem.* **2018**, *33*, 151 – 157.

[13] C. E. Dumelin, S. Trussel, F. Buller, E. Trachsel, F. Bootz, Y. Zhang, L. Mannocci, S. C. Beck, M. Drumea-Mirancea, M. W. Seeliger, C. Baltes, T. Muggler, F. Kranz, M. Rudin, S. Melkko, J. Scheuermann, D. Neri, *Angew. Chem. Int. Ed.* **2008**, *47*, 3196 – 3201; *Angew. Chem.* **2008**, *120*, 3240 – 3245.

[14] a) X. Li, H. Han, M. T. Zhou, B. Yang, A. P. Ta, N. Li, J. Chen, W. Wang, *Cell Rep.* **2017**, *20*, 737 – 749; b) X. Wu, F. Luo, J. Li, X. Zhong, K. Liu, *Int. J. Oncol.* **2016**, *48*, 1333 – 1340.

[15] E. Wahlberg, T. Karlberg, E. Kouznetsova, N. Markova, A. Macchiarulo, A. G. Thorsell, E. Pol, A. Frostell, T. Ekblad, D. Oncu, B. Kull, G. M. Robertson, R. Pellicciari, H. Schuler, J. Weigelt, *Nat. Biotechnol.* **2012**, *30*, 283 – 288.

Manuscript received: March 22, 2018

Accepted manuscript online: ■■ ▮▮, 0000

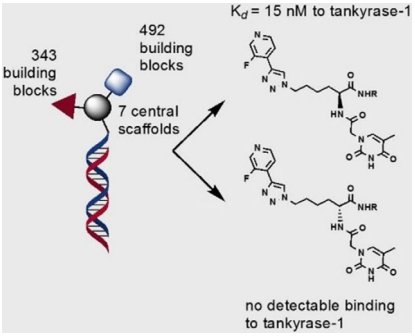
Version of record online: ■■ ▮▮, 0000

COMMUNICATIONS

N. Favalli, S. Biendl, M. Hartmann, J. Piazzi, F. Sladojevich, S. Gräslund, P. J. Brown, K. Näreoja, H. Schüler, J. Scheuermann, R. Franzini, D. Neri**

■■ – ■■

A DNA-Encoded Library of Chemical Compounds Based on Common Scaffolding Structures Reveals the Impact of Ligand Geometry on Protein Recognition



Advanced search function: A 1.2-million-compound DNA-encoded library was built on seven central scaffolds with two sets of 343×492 building blocks. Library screening by affinity capture revealed that for some target proteins, the chemical nature of the building blocks dominated the selection results, whereas for other proteins the central scaffold also crucially contributed to ligand affinity. This suggests that predisposing the orientation of chemical building blocks toward the protein target may enhance the screening productivity of encoded libraries.