

# RSC Chemical Biology

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: S. Pomplun, E. van der Nol, Z. Luo, Q. Q. Gao, N. A. Haupt and S. Boecker, *RSC Chem. Biol.*, 2025, DOI: 10.1039/D5CB00303B.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

## Integration of Palladium-Catalyzed C–N Coupling into Self-Encoded Libraries for Accelerated Hit Discovery

Authors: Edith van der Nol<sup>1,2+</sup>, Zhenshuo Luo<sup>1+</sup>, Qing Qing Gao<sup>1+</sup>, Nils Alexander Haupt<sup>3</sup>, Sebastian Böcker<sup>3</sup>, Sebastian Pomplun<sup>1,2\*</sup>

Affiliations:

<sup>1</sup>LACDR, Leiden University; Leiden, 2333 CC, The Netherlands

<sup>2</sup>Oncode Institute; Utrecht, 3521 AL, The Netherlands

<sup>3</sup>Chair for Bioinformatics, Institute for Computer Science, Friedrich Schiller University Jena, Jena, 07737, Germany

\*Corresponding authors. Email: [s.j.pomplun@lacdr.leidenuniv.nl](mailto:s.j.pomplun@lacdr.leidenuniv.nl)

+These authors contributed equally to this work

### Abstract:

Affinity screenings with encoded libraries are transformative tools for rapid hit discovery from vast compound collections. Yet the adaptation of established chemical reactions to DNA-encoded libraries (DELs) remains challenging due to DNA-compatibility constraints and mismatches between barcode and chemical structure in case of incomplete reactions or side product formation. Recently, we introduced self-encoded libraries (SELs) as a barcode-free alternative to DELs. The SEL platform offers unmatched flexibility in reaction conditions and decodes screening hits directly from their chemical structure, avoiding the problem of mismatched barcode-compound pairs. Here, we expand the SEL platform to Buchwald-Hartwig aminations, enabling the construction of new high diversity SELs. We performed a thorough reaction condition optimization and tested a scope of > 170 different building blocks. We adapted our automated MS/MS-based decoding methodology SIRIUS-COMET to the resulting scaffolds, enabling accurate compound decoding from complex mixtures. A 25,725-member library was synthesized and screened all at once against carbonic anhydrase IX (CAIX), resulting in robust enrichment of hits with specific building block patterns and yielding several nanomolar-affinity binders. This work showcases the seamless integration of palladium-catalyzed cross-couplings into SELs, expanding the chemical space of this technology and accelerating hit discovery with high synthetic versatility.



## Introduction

View Article Online  
DOI: 10.1039/D5CB00303B

The identification of bioactive molecules is a crucial step in drug discovery, relying heavily on access to large and diverse collections of small molecules. Traditionally, high-throughput screening (HTS) involves the parallel screening of individual compounds but can be resource-intensive. Encoded-library techniques have emerged as a powerful alternative technology, enabling the rapid and cost-effective screening of millions or even billions of compounds simultaneously.<sup>1-3</sup> Among these, DNA-encoded libraries (DELs) in particular has become a widespread technology in both industry and academia, leading to the discovery of numerous bioactive compounds across a broad range of targets.<sup>1,2,4-6</sup>

Obtaining high-quality combinatorial libraries requires robust, high-yielding chemical reactions to maintain synthetic fidelity. Incorporating a broad range of chemical transformations grants access to a large chemical space and, in turn, increases the probability of discovering novel bioactive molecules. DELs have inherent challenges due to the synthesis of small molecules on DNA strands and although the number of DEL-compatible reactions continues to grow, it remains challenging to find robust conditions for several chemical transformations.<sup>7-9</sup> Among the palladium-catalyzed cross-coupling reactions described for DELs, Suzuki–Miyaura, Sonogashira and Heck couplings have been successfully implemented under mild, DNA-compatible conditions.<sup>10-13</sup>

The Buchwald-Hartwig amination has become a cornerstone transformation in medicinal chemistry due to the prevalence of arylated amines in pharmaceuticals and natural products.<sup>10,14</sup> However, its on-DNA implementation is hindered by several DNA degradation conditions, such as high temperatures, palladium-catalyst interactions, and exposure to organic solvents.<sup>12,15-18</sup> Additionally, the presence of air or water can lead to catalyst inactivation, further complicating its use in DELs. The reported examples of Buchwald-Hartwig on DNA often suffer from low yields and limited substrate scope.<sup>15,19-23</sup> In principle, the presence of incomplete reaction products is not inherently problematic for affinity selections, as also shown by Cernak et al., who used nanoscale synthesized libraries in affinity ranking experiments.<sup>24</sup>

However, in encoded libraries, such incompleteness leads to a mismatch between the molecular barcode and the corresponding truncated or deleted compound, resulting in misleading outcomes during hit decoding. Although strategies to mitigate these issues are under development, they inevitably introduce additional complexity to library synthesis.<sup>25</sup>

Recently, we introduced self-encoded libraries (SELs) as a barcode-free alternative to DELs.<sup>26</sup> SELs are prepared by solid-phase synthesis and therefore offer exceptional flexibility in reaction conditions and catalyst compatibility with broad tolerance for organic reagents and elevated temperatures. In the SEL discovery workflow, combinatorial libraries are screened in affinity selections against a protein of interest. Enriched hit compounds are detected using tandem mass spectrometry (MS/MS) and decoded by an automated software (SIRIUS-COMET). SIRIUS-COMET enables the screening of very large libraries (up to 750,000 members) by accurately annotating compound structures from their fragmentation spectra, even in the presence of hundreds of isobaric compounds. Given the absence of DNA-compatibility constraints for library synthesis, the SEL platform enables the implementation of established medicinal chemistry transformations, including those challenging for DELs. Since the hit compounds are decoded directly from their own structure, there are no problems related to barcode-compound mismatches.



In this study, we implement the Buchwald-Hartwig reaction in our SEL platform. Building on previous reports on solid-phase based Buchwald-Hartwig aminations<sup>27–29</sup>, we performed a thorough reaction condition optimization and a broad substrate scope. We expanded the SIRIUS-COMET decoding workflow to account for the resulting scaffolds and validated the annotation fidelity on ~600 compounds. With the building blocks validated in the substrate scope, we built a 25,725 membered SEL and benchmarked the whole discovery workflow in affinity selections against carbonic anhydrase IX (CAIX). Robust enrichment of hits with experimentally validated nanomolar affinities were achieved, showcasing the compatibility of the SEL platform with this new library design using Buchwald-Hartwig aminations.

## Results

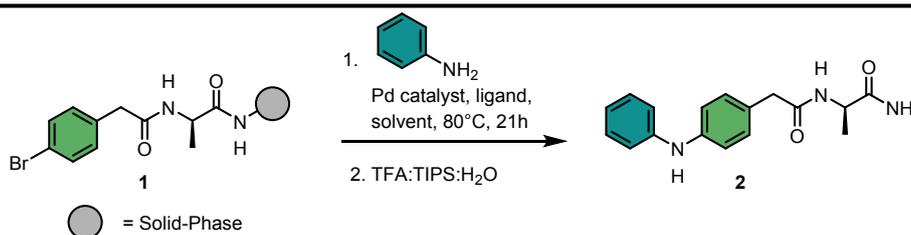
### Optimization of synthesis conditions for solid-phase Buchwald-Hartwig reactions

We started the condition optimization with the model reaction between solid-phase bound aryl bromide **1** and aniline (Table 1). For determining the conversion efficiency, we cleaved each compound from the solid support and analyzed the resulting mixture by LC-MS.<sup>25</sup> In a first iteration we kept the Pd source (PdCl<sub>2</sub>), base (K<sub>2</sub>CO<sub>3</sub>) and solvent composition (DMF:H<sub>2</sub>O, 9:1) constant and tested a set of different phosphine-based ligands. XPhos and tBuXPhos resulted in moderate conversions of 43% and 40%, respectively, to **2**, while the other ligands resulted in poor yields (Table 1, entries 1-6). Changing the base to Na<sub>2</sub>CO<sub>3</sub>, DBU or K<sub>2</sub>CO<sub>3</sub>, decreased the reaction conversion (Table 1, entries 7-9). Notably, by switching the palladium source to Pd(OAc)<sub>2</sub> a substantially better conversion was observed (67%), and could even be improved (80%), using higher catalyst loading (Table 1, entries 10-11). A small solvent screen showed that DMF, dioxane, toluene and anisole gave all comparably good conversions (Table 1, entries 12-14). Based on the conversions and green chemistry guidelines, we prioritized the most environmentally benign option, anisole.<sup>30</sup> Performing solid-phase reaction under fully inert conditions, especially when aiming at a combinatorial set up, is very challenging. As a feasible set up we performed the reactions in sealed Eppendorf tubes, purged with N<sub>2</sub>. Compared to performing the reactions under air, the reproducibility of the reaction outcomes improved substantially with this set up (data not shown).

View Article Online  
DOI: 10.1039/D5CB00303B

RSC Chemical Biology Accepted Manuscript



**Table 1. Optimization of the reaction between solid-phase coupled 2-(4-bromophenyl)acetic acid and aniline<sup>a</sup>**

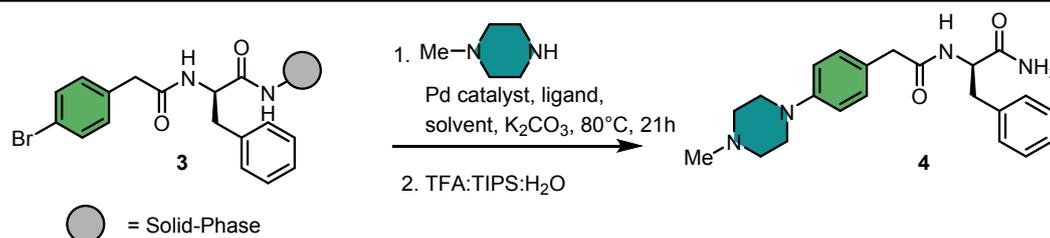
Entry	Pd catalyst (mol%)	Ligand (mol%)	Base (2 eq)	Solvent (9:1)	Conversion (%) <sup>b</sup>
1	<b>PdCl<sub>2</sub> (10)</b>	<b>Xphos (20)</b>	<b>K<sub>2</sub>CO<sub>3</sub></b>	<b>DMF:H<sub>2</sub>O</b>	<b>43</b>
2	PdCl <sub>2</sub> (10)	Johnphos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	11
3	PdCl <sub>2</sub> (10)	CyJohnphos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	2
4	PdCl <sub>2</sub> (10)	Davephos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	26
5	PdCl <sub>2</sub> (10)	tBuXphos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	40
6	PdCl <sub>2</sub> (10)	Sphos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	22
7	PdCl <sub>2</sub> (10)	Xphos (20)	Na <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	26
8	PdCl <sub>2</sub> (10)	Xphos (20)	DBU	DMF:H <sub>2</sub> O	0
9	PdCl <sub>2</sub> (10)	Xphos (20)	KOtBu	DMF:H <sub>2</sub> O	17
10	Pd(OAc) <sub>2</sub> (10)	Xphos (20)	K <sub>2</sub> CO <sub>3</sub>	DMF:H <sub>2</sub> O	67
<b>11</b>	<b>Pd(OAc)<sub>2</sub> (20)</b>	<b>Xphos (40)</b>	<b>K<sub>2</sub>CO<sub>3</sub></b>	<b>DMF:H<sub>2</sub>O</b>	<b>80</b>
12	Pd(OAc) <sub>2</sub> (20)	Xphos (40)	K <sub>2</sub> CO <sub>3</sub>	Toluene:H <sub>2</sub> O	78
13	Pd(OAc) <sub>2</sub> (20)	Xphos (40)	K <sub>2</sub> CO <sub>3</sub>	Dioxane:H <sub>2</sub> O	76
<b>14</b>	<b>Pd(OAc)<sub>2</sub> (20)</b>	<b>Xphos (40)</b>	<b>K<sub>2</sub>CO<sub>3</sub></b>	<b>Anisole:H<sub>2</sub>O</b>	<b>83</b>

<sup>a</sup> Reaction conditions: **1** (5 μmol, 1 eq.), aniline (10 μmol, 2 eq.), base (10 μmol, 2 eq.), solvent:H<sub>2</sub>O (9:1, 100 μL), 80°C, 21 hrs. <sup>b</sup> Conversion was determined by LC-MS analysis through peak integration of UV-absorption at 254 nm.

To evaluate the performance of the established conditions with aliphatic amines, we focused on the reaction with 1-methylpiperazine as a model substrate, given that piperidine and piperazine are among the most prevalent nitrogen heterocycles in FDA-approved drugs<sup>31</sup>. When testing the reaction between model substrate **3** with 1-methylpiperazine under the optimized conditions for aniline, we obtained a conversion to product of only 3% (Table 2, entry 1) which required further improvement of reaction conditions. We first screened 10 different phosphine ligands (Table 2, entries 1-10). Nine of the ten ligands resulted in conversions between 0 and 7% and only tBuXPhos afforded a moderate conversion of 30% (Table 2, entry 10). Increasing the catalyst loading further improved the conversion to 41% (Table 2, entry 11). Another minor improvement to 46% was obtained by using Pd<sub>2</sub>(dba)<sub>3</sub> (Table 2, entries 12-14). A subsequent solvent screening using different protic and aprotic solvents revealed toluene to be the optimal choice, leading to conversions of 61% (Table 2, entries 15-19). A similar performance was retained at slightly reduced catalyst loadings (Table 2, entry 20). While these results confirm the more challenging nature of Buchwald-Hartwig reactions for combinatorial library synthesis, the optimized conditions gave conversions over 50%, making it suitable for integration in our SEL platform.



**Table 2. Optimization of the reaction between solid-phase coupled 2-(4-bromophenyl)acetic acid and 1-methylpiperazine<sup>a</sup>**



Entry	Pd catalyst (mol%)	Ligand (mol%)	Solvent (9:1)	Conversion (%)
1	Pd(OAc) <sub>2</sub> (20)	XPhos (40)	Anisole:H <sub>2</sub> O	3
2	Pd(OAc) <sub>2</sub> (20)	BINAP (20)	Anisole:H <sub>2</sub> O	0
3	Pd(OAc) <sub>2</sub> (20)	XantPhos (20)	Anisole:H <sub>2</sub> O	1
4	Pd(OAc) <sub>2</sub> (20)	SPhos (40)	Anisole:H <sub>2</sub> O	2
5	Pd(OAc) <sub>2</sub> (20)	RuPhos (40)	Anisole:H <sub>2</sub> O	4
6	Pd(OAc) <sub>2</sub> (20)	JohnPhos (40)	Anisole:H <sub>2</sub> O	7
7	Pd(OAc) <sub>2</sub> (20)	CyJohnPhos (40)	Anisole:H <sub>2</sub> O	1
8	Pd(OAc) <sub>2</sub> (20)	DavePhos (40)	Anisole:H <sub>2</sub> O	2
9	Pd(OAc) <sub>2</sub> (20)	BrettPhos (40)	Anisole:H <sub>2</sub> O	4
<b>10</b>	<b>Pd(OAc)<sub>2</sub> (20)</b>	<b>tBuXPhos (40)</b>	<b>Anisole:H<sub>2</sub>O</b>	<b>30</b>
<b>11</b>	<b>Pd(OAc)<sub>2</sub> (50)</b>	<b>tBuXPhos (100)</b>	<b>Anisole:H<sub>2</sub>O</b>	<b>41</b>
12	PdCl <sub>2</sub> (50)	tBuXPhos (100)	Anisole: H <sub>2</sub> O	18
13	APC (25)	tBuXPhos (100)	Anisole: H <sub>2</sub> O	29
<b>14</b>	<b>Pd<sub>2</sub>(dba)<sub>3</sub> (25)</b>	<b>tBuXPhos (100)</b>	<b>Anisole: H<sub>2</sub>O</b>	<b>46</b>
15	Pd <sub>2</sub> (dba) <sub>3</sub> (25)	tBuXPhos (100)	DMF: H <sub>2</sub> O	59
16	Pd <sub>2</sub> (dba) <sub>3</sub> (25)	tBuXPhos (100)	Dioxane: H <sub>2</sub> O	7
17	Pd <sub>2</sub> (dba) <sub>3</sub> (25)	tBuXPhos (100)	THF: H <sub>2</sub> O	14
18	Pd <sub>2</sub> (dba) <sub>3</sub> (25)	tBuXPhos (100)	Dibutylether: H <sub>2</sub> O	51
<b>19</b>	<b>Pd<sub>2</sub>(dba)<sub>3</sub> (25)</b>	<b>tBuXPhos (100)</b>	<b>Toluene: H<sub>2</sub>O</b>	<b>61</b>
<b>20</b>	<b>Pd<sub>2</sub>(dba)<sub>3</sub> (20)</b>	<b>tBuXPhos (80)</b>	<b>Toluene: H<sub>2</sub>O</b>	<b>60</b>

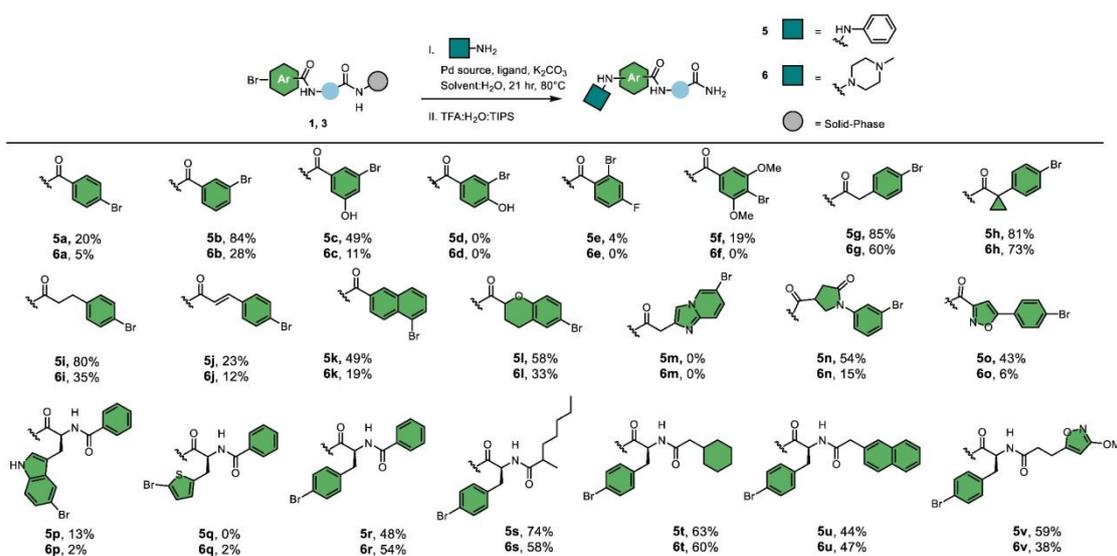
<sup>a</sup> Reaction conditions: **1** (10 μmol, 1 eq.), 1-methylpiperazine (20 μmol, 2 eq.), solvent:H<sub>2</sub>O (9:1, 300 μL), K<sub>2</sub>CO<sub>3</sub> (20 μmol, 2 eq.), 80 °C, 18 h. <sup>b</sup> Conversion was determined by LC-MS analysis through peak integration of UV-absorption at 254 nm.



### Building block scope

View Article Online  
DOI: 10.1039/D5CB00303B

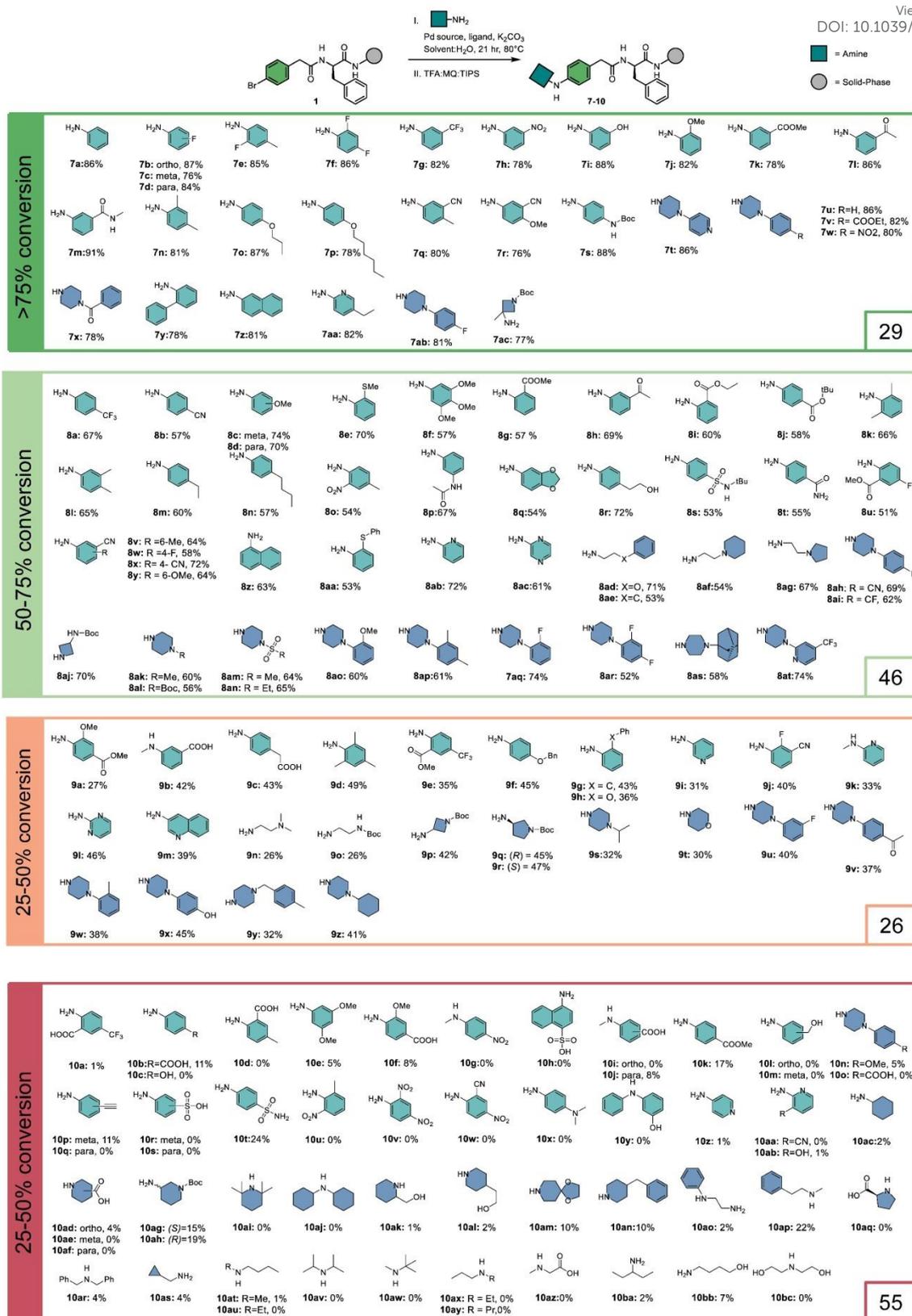
With the optimized conditions for aromatic and aliphatic amines, we evaluated the reaction scope on 22 different solid-phase bound aryl bromides, reacting each with aniline and 1-methylpiperazine. Looking ahead to the planned combinatorial library synthesis, we defined a threshold of 50% conversion to consider a building block suitable for library incorporation. From the 22 aryl bromides investigated (Figure 1), nine resulted in conversions > 50% when reacted with aniline and five when reacted with 1-methylpiperazine. The majority of substituted benzoic acids (**5a-f,6a-f**) gave unsatisfactory results. Bromophenylacetic acid variants (**5g, 6g**) and bromophenylpropanoic acids (**5i-j, 6i-j**) were well-tolerated. The bicyclic structures (**5k-n, 6k-n**) showed a general preference for the coupling with aromatic amines compared to the reaction with aliphatic amines. Furthermore, bromophenylalanine derivatives bearing different carboxylic acid decorators (**5r-v, 6r-v**) generally resulted in reasonable conversions.



**Figure 1. Scope of aryl bromides** <sup>a</sup>Compounds **5** were reacted with starting material **1** and compounds **6** with starting material **3**. <sup>b</sup>Conversion was determined by LC-MS analysis through peak integration of UV-absorption at 254 nm.

Next, we tested the reaction between substrate **1** and 156 aromatic and aliphatic amine building blocks (Figure 2). A broad range of aromatic amines bearing alkyl, alkoxy, fluoro, trifluoromethyl, nitrile, ester, and acetyl substituents were converted to the desired products with conversions > 50%. In contrast, amines containing alkyne, carboxylic acid, sulfonic acid, free sulfonamide, and nitro substituents mostly resulted in low conversions. From the heterocyclic amines, 2-aminopyridine and 2-aminopyrazine underwent efficient coupling, whereas the 3-amino- and 4-aminopyridine isomers or substituted derivatives gave poor conversion, likely due to their reduced nucleophilicity. Several piperazine analogs and a few other cyclic- and primary aliphatic amines met the conversion threshold. Among the variants that displayed lower conversions, steric hindrance seemed to be the most prominent common feature. Overall, 75 building blocks showed conversion over 50%, providing a set of diverse building blocks for library synthesis.





**Figure 2. Scope of 156 amines** <sup>a</sup>The investigated amines are categorized based on conversion. <sup>b</sup>The turquoise colored amine used the reaction conditions from Table 1, entry 14. The dark blue amine used the reaction conditions from Table 2, entry 20. <sup>c</sup>Conversion was determined by LC-MS analysis through peak integration of UV-absorption at 254 nm.



## Library decoding

View Article Online  
DOI: 10.1039/D5CB00303B

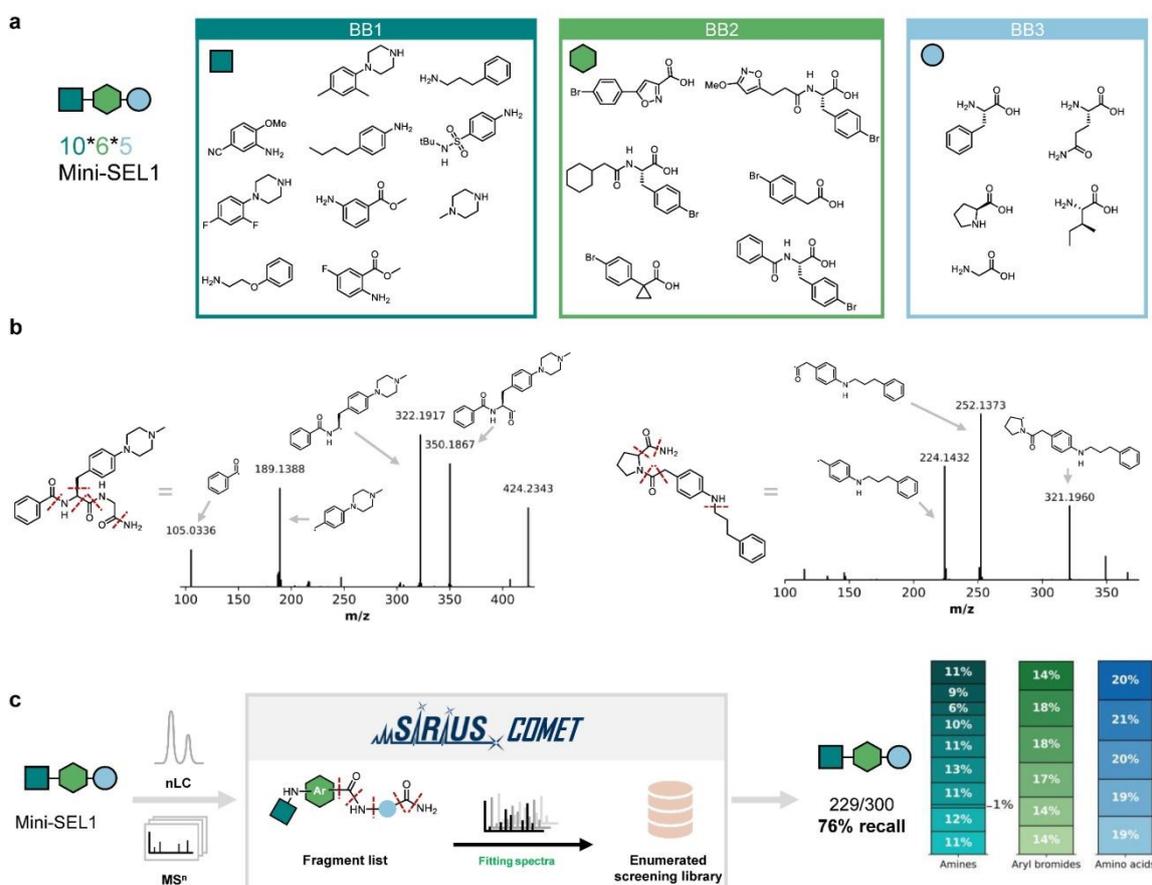
The SEL workflow entails screening a large library all at once against a target of interest via affinity selection. As opposed to classical AS-MS,<sup>32–35</sup> no particular attention to the presence of isobaric compounds in a library is needed. This is possible thanks to our previously developed computational tool, SIRIUS-COMET, capable of annotating compounds from MS/MS data and annotate their structures.<sup>26</sup>

In order to evaluate the decoding software and optimize parameters for this new Buchwald-Hartwig based library, we first synthesized two defined test libraries of 300 compounds each (miniSEL1 and miniSEL2, Figure 3a and Figure S1). Each library consists of a scaffold with three variable positions: an amino acid building blocks (BB3) coupled to an aryl bromides (BB2) and subsequently reacted with amines (BB1) using the established conditions. Each library was prepared via split-and-pool synthesis, bulk purified and then measured as a mixture by nanoLC-MS/MS. First, we analyzed a few selected spectra manually to observe overall trends in fragmentation of the library compounds. The most prominent and consistent fragmentation observed across scaffolds is the amide bond between the amino acid building block and the aryl bromide (see Figure 3b for selected examples). Also, the terminal amide resulting from solid-phase synthesis was fragmented in all selected compounds. The C-N bond formed by the Buchwald-Hartwig coupling fragmented only sporadically with a low signal intensity. Increasing the fragmentation energy improved fragment intensity, but not enough to utilize it as a consistent fragment for library decoding (Figure S2).

For the full decoding of our libraries, we used SIRIUS-COMET<sup>26</sup>, a software platform that enables accurate annotation and decoding of combinatorial libraries and at the same time effectively reducing the number of false positive annotations. A fragment list based on the selected most prominent scaffold fragmentations is used to filter MS/MS scans based on a minimum number of matching peaks. The remaining spectra were then annotated using EPIMETHEUS, the ranking and substructure annotation module within the SIRIUS software, using the enumerated library database (represented as SMILES). Analysis of the MS/MS spectra of miniSEL1 and 2 with SIRIUS-COMET resulted in the correct annotation of 229/300 library members for miniSEL1 (76%), and 218/300 (73%) for miniSEL2. An even distribution among the building blocks was observed, except for 4-butylaniline, indicating that combinations of the optimized building blocks were generally well tolerated (Figure 3c and Figure S3). These results confirm the successful synthesis under combinatorial settings of most library members, forming a robust basis for the preparation of larger libraries.

As SELs are decoded by their chemical structure, we investigated whether common side products could be detected and characterized. In this way, side products could even become part of the enumerated library and contribute to greater diversity. While analyzing the building block scope, we often observed conversions to the following recurring side products; unreacted starting material **1** and its hydrolyzed and proto-dehalogenated derivative. While we had previously confirmed the correct synthesis of the majority of all library compounds for miniSEL1 and 2, we wanted to assess whether we could also detect and annotate these combinatorial side products. For miniSEL1 74 out of 90 possible side products were identified in the MS/MS spectra (Figure S5) and 16/45 for miniSEL2 (Figure S6). These results highlight that SELs can, in principle, identify any structure formed during synthesis, provided that the underlying reaction pathways are known. Whether these side products will actually be productive and detectable in real screening campaigns will depend on each individual quantity, but, notably, many of these structures would be incorrectly assigned in conventional DEL screenings and would require extensive off-DNA validation to determine the correct hit structure.





**Figure 3. Library decoding using SIRIUS-COMET** a) MiniSEL1 was synthesized via split-and-pool synthesis on solid support using selected amino acid, aryl bromide and amine building blocks. b) Examples of LC-MS/MS traces. The highest intensity fragments were used as input in SIRIUS-COMET. A generalized overview of the fragments is depicted in Figure S7. c) The LC-MS/MS data was imported into SIRIUS-COMET. MS/MS spectra were filtered based on the minimum number of matching peaks from the fragment list. Compounds are then annotated based on a database containing all library members. The workflow results in good recall rates considering when considering the top ranked annotated structures by EPIMETHEUS and building blocks are evenly distributed among the annotated structures.

### Library synthesis and selection against CAIX

With an established method for combinatorial synthesis and decoding of Buchwald-Hartwig based libraries, the preparation of a large SEL library was initiated. Using 49 amino acid building blocks, 7 aryl bromides and 75 amines, we assembled a SEL with 25,725 members (Figure 4a). Depending on the type of amine, we used the respective optimized conditions for the Buchwald-Hartwig amination step. For specific exact masses the library contains up to 28 isobaric compounds but each of those features a distinct combination of building blocks, likely enabling structure annotation through MS/MS fragmentation (Figure 4b). We computationally enumerated the entire library and visualized compliance with the Lipinski parameters: molecular weight (MW < 500 Da), XlogP (> 5), hydrogen bond donors



(HBD < 5), hydrogen bond acceptors (HBA < 10), and topological polar surface area (TPSA < 140) (Figure 4c).

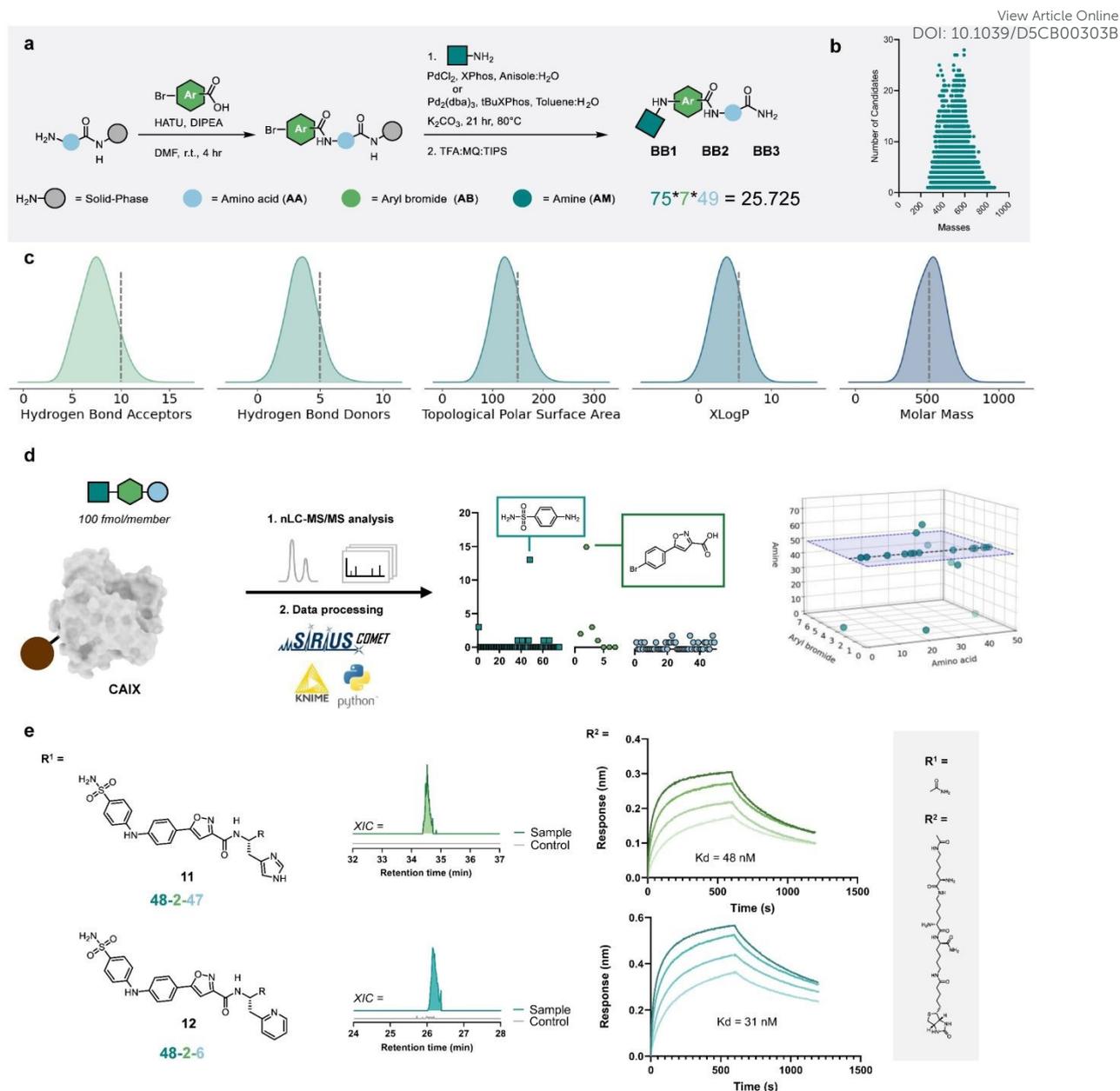
Finally, we wanted to validate the entire discovery workflow including affinity selection, tandem mass spectrometry and hit decoding. To this end, we performed an affinity selection experiment against CAIX. CAIX is a therapeutically relevant target for cancer treatment, particularly in hypoxic tumors, due to its role in tumor cell survival and pH regulation in the tumor microenvironment. High affinity CAIX ligands are enabling tools for radionuclide cancer therapy.<sup>36</sup> Moreover, CAIX is a preferred scaffold for benchmarking library selections due to its predictable binding profile to aromatic sulfonamides<sup>25,36</sup>. In addition, other members of the carbonic anhydrase family have frequently served as benchmark targets for the validation of ligand discovery strategies such as *in situ* click chemistry<sup>37</sup> and direct to biology screening<sup>38,39</sup>.

The affinity selection was performed on biotin-CAIX, immobilized on streptavidin-coated magnetic beads and incubated with the library (using 100 fmol per member). After washing away non-binders, protein-bound hits were eluted under denaturing conditions and analyzed with our nanoLC-MS/MS SIRIUS-COMET workflow. The selection in parallel was performed on unfunctionalized magnetic streptavidin beads. Non-specific binders, appearing in both selections, were computationally removed from our analysis by performing a background subtraction. Only compounds displaying at least three-fold higher intensity in the CAIX sample were considered as hits.

After background subtraction and SIRIUS-COMET structure annotation, we obtained 21 candidate structures. When analyzing building block frequencies, we found a significant enrichment of 4-aminobenzenesulfonamide, appearing in 13 annotated structures as first ranked (Figure 4d). Given the known affinity of 4-aminobenzenesulfonamide for CAIX, these 13 hits can be considered as real binders with high confidence. For four of the remaining eight candidates, the sulfonamide building block was still present, but proposed with a lower rank. This outcome indicates that optimizing the ranking algorithm could improve hit detection in the future. The remaining four candidates in the manual analysis showed unclear enrichment and/or overall poor structure annotation and were not considered as hits. Interestingly, the sulfonamide always occurred in combination with 5-(4-bromophenyl)isoxazole-3-carboxylic acid, indicating that this combination is important for binding (Figure 4d). The building block enrichment (proportion of sulfonamide in hits vs expected random distribution) has a strong statistical significance with a p value of  $p < 0.0001$  (Figure S5).

From the hits we selected two compounds for experimental binding validation (Figure 4e). We prepared variants of both compounds (**11** and **12**) with a biotin tag and tested their binding via biolayer interferometry (BLI), which demonstrated low nanomolar binding affinities with  $K_D$  values of 48 nM and 31 nM, respectively. We resynthesized the untagged version of compound **12** to further confirm and validate the results of our affinity selection. While the nanoLC retention time and MS/MS spectra matched the original hit perfectly (Figure S18-19), NMR characterization was hindered by a proto-dehalogenated side product formed during synthesis. Despite multiple attempts, this byproduct could not be separated from the target compound (Figure S20-21). This difficulty illustrates a common bottleneck in combinatorial chemistry, where the transition from a mixture-based "hit" to a purified, isolated compound is not always straightforward. Altogether, the CAIX selection demonstrates that Buchwald-Hartwig aminations can be incorporated in SELs and enable discovery campaigns from large libraries.





**Figure 4 Library synthesis and the identification of ligands against CAIX** **a)** The SEL design consists of an amino acid and the optimized Buchwald-Hartwig cross coupling between an aryl bromide and amine. Building blocks with crude purity over 50% were selected. **b)** The scatter plot shows the mass degeneracy at 5 ppm mass accuracy. **c)** The distribution plot of druglike properties show that the average distribution abides Lipinski's rule of 5. **d)** AS-MS results plotted based on the building block frequencies from the best ranked annotations plotted as 2D and 3D scatter plots. The x, y and z axis represent the three building blocks. Showing that the combination of 5-(4-Bromophenyl) isoxazole-3-carboxylic acid and 4-aminobenzenesulfonamide is significantly enriched ( $p < 0.0001$ , see Figure S4), indicated by the black dotted line. **e)** Structure of two selected hits. The extracted ion chromatogram using 5 ppm accuracy shows selective enrichment for CAIX compared to the streptavidin magnetic beads control. Binding affinity was determined by BLI association/dissociation curves of the biotinylated version of hit compounds.



## Discussion and Conclusion

View Article Online  
DOI: 10.1039/D5CB00303B

Here, we developed procedures for solid-phase Buchwald-Hartwig aminations, enabling the construction of large combinatorial libraries. We optimized reaction conditions for aromatic and aliphatic amines and tested a building block scope of 22 aryl bromides and 156 amines. Building blocks with incorporations > 50% were included in a final SEL of 25,725 members. We adapted our previously developed SIRIUS-COMET software to enable the automated structure annotation of MS/MS spectra of complex samples of compounds resulting from the combinatorial Buchwald-Hartwig amination. Finally, we demonstrated the application in AS-MS/MS against CAIX, resulting in robust hit enrichment and two validated binders with nanomolar affinities (48 nM and 31 nM), highlighting the potential of SELs for discovering novel hits.

Compared to Suzuki cross couplings that we previously developed<sup>26</sup>, establishing the Buchwald-Hartwig amination for combinatorial synthesis was more challenging, exhibited a narrower substrate scope and lower yields. The Buchwald-Hartwig amination reported here required tailored reaction conditions based on the type of amine, resulting in distinct reaction protocols for primary aromatic and cyclic aliphatic amines. Primary aromatic amines were generally well tolerated, while aliphatic amines showed poor conversion and probably due to their susceptibility to  $\beta$ -hydride elimination. Cyclic amines outperformed their acyclic counterparts. Both findings are consistent with previous reaction performed in DEL optimization.<sup>15,40</sup>

A remaining challenge for DELs is the identification of true hits after the selection arising from incorrectly encoded library members, due to truncations or side product formation. In contrast, SELs enable direct decoding of compounds from their own structure, thereby eliminating issues related to barcode-compound mismatches. Moreover, SELs also offer the potential to detect and characterize side products that may arise during the synthesis.

Future research will be needed to further expand the chemical space by incorporating a more diverse set of aryl bromides and amines. A significant variability in ligand performance in our optimizations would suggest that further optimization into newly developed ligands for the incorporation of primary amines<sup>41</sup>, secondary amines<sup>42,43</sup> and heterocycles<sup>44-46</sup> would increase our substrate scope. Since we have no DNA constraints, we can likely directly apply novel reported reaction conditions to solid-support, offering a distinct advantage over DEL-based reaction optimization.

To decode isobaric compounds in larger libraries (beyond 100,000 members), additional MS/MS fragment information on the introduced C-N is required. Under the current conditions, this bond fragments only sporadically, limiting confident structural assignment for larger libraries, as information about the amine building block is missing. By further optimizing MS/MS-based fragmentation in terms of collision energy and structure annotations, we could potentially enhance the decoding accuracy.

The modest library size shown in this study still exceeds the size of pools (typically 500 – 1000 members) that can be screened by classical affinity selection mass spectrometry screens, relying on identification by MS1. DELs usually achieve substantially larger library sizes. However, we are confident that our SEL platform continues expanding and multiple libraries with different scaffolds will overall contribute to a broad chemical diversity.

The selection against CAIX yielded 13 hits, where the enrichment was dependent on the combination of 4-aminobenzenesulfonamide and 5-(4-bromophenyl)isoxazole-3-carboxylic acid. CAIX is known to bind to aromatic sulfonamides, but adjacent BBs can play an important role<sup>36</sup>. The dependence on these two building blocks showcase the importance of developing diverse chemistry of different library scaffolds.



From the 343 potential sulfonamide binders, only a fraction was identified. This outcome reflects the constraints arising from building-block orientation in combination with the applied washing protocol. While our washing steps efficiently remove non-binders also ligands with fast dissociation rates might be lost. Consequently, at 100 fmol per library member, the assay is biased toward the identification of low-nanomolar binders. Fine-tuning the library concentration and washing parameters could allow systematic modulation of this affinity threshold, an aspect to be explored in subsequent studies.

Metal residues in the library might pose a challenge in specific screening campaigns. The solid support based synthesis enables extensive washing after the cross-coupling step, likely removing the catalyst from the system. After acidic cleavage from the resin the library is also bulk purified on a C18 column, leading to further removal of undesired components. In this study we did not take additional steps for quantitative metal removal. We acknowledge that in specific cases the presence of palladium might represent a limitation and we suggest in these cases to perform additional resin washes using potent metal chelators.

Overall, this work expands the chemical space accessible to self-encoded libraries and illustrates how established reactions can be integrated into the SEL platform for early drug discovery workflows.

### Author contributions

Conceptualization: EvdN, ZL, QQG, SP

Synthesis: EvdN, ZL, QQG

Methodology: EvdN, NAH, SB, SP

Assays: EvdN

Software: NAH, SB

Visualization: EvdN, SP

Supervision: EvdN, SP

Funding acquisition: SP, SB

Writing – original draft: EvdN, SP

Writing – review & editing: EvdN, NAH, SB, SP

**Conflicts of interest:** SP, SB, EvdN and NAH have filed a patent application for the methodology described here. The remaining authors declare no competing interests.

**Data availability:** The raw LC-MS/MS files and the files needed to run SIRIUS:COMET have been deposited in the Zenodo database (10.5281/zenodo.17632652). Source data are provided with this paper.

All other data is available in the main text, the supplementary materials and from the corresponding author(s) upon request.

**Acknowledgements:** SJP acknowledges funding via the ERC Starting Grant (101039354). The Pomplun Lab gratefully acknowledges financial support from Mr. H. J. M. Roels through a donation to the Oncode Institute and KWF's financial support of the Oncode Institute.



## References

View Article Online  
DOI: 10.1039/D5CB00303B

1. Gironda-Martínez, A., Donckele, E. J., Samain, F. & Neri, D. DNA-Encoded Chemical Libraries: A Comprehensive Review with Successful Stories and Future Challenges. *ACS Pharmacol. Transl. Sci.* **4**, 1265–1279 (2021).
2. Peterson, A. A. & Liu, D. R. Small-molecule discovery through DNA-encoded libraries. *Nat. Rev. Drug Discov.* **22**, 699–722 (2023).
3. Favalli, N., Bassi, G., Scheuermann, J. & Neri, D. DNA-encoded chemical libraries – achievements and remaining challenges. *FEBS Lett.* **592**, 2168–2180 (2018).
4. Yang, X., Childs-Disney, J. L., Paegel, M. & Disney, M. D. DNA-Encoded Libraries and Their Application to RNA. *Isr. J. Chem.* **63**, e202300073 (2023).
5. Chen, Q. & Israel, D. I. Affinity-based DEL Selections with Different Target Types: Overview and Achievements. in *DNA-encoded Library Technology for Drug Discovery* 1–16 (Royal Society of Chemistry, 2025). doi:10.1039/9781788016032-00001.
6. Sibai, R. E., Farahat, Z. E. M., Qasem, H. H. & Hassan, H. The power of DNA-encoded chemical libraries in the battle against drug-resistant bacteria. *RSC Adv.* **15**, 14001–14029 (2025).
7. Ma, P. *et al.* Evolution of chemistry and selection technology for DNA-encoded library. *Acta Pharm. Sin. B* **14**, 492–516 (2024).
8. Gao, Y. *et al.* Recent advances in DNA-encoded libraries. *Chemical Communications* **61**, 10952–10968 (2025).
9. Götte, K., Chines, S. & Brunschweiler, A. Reaction development for DNA-encoded library technology: From evolution to revolution? *Tetrahedron Lett.* **61**, 151889 (2020).
10. Ruiz-Castillo, P. & Buchwald, S. L. Applications of Palladium-Catalyzed C-N Cross-Coupling Reactions. *Chem. Rev.* **116**, 12564–12649 (2016).
11. Shi, Y., Wu, Y. R., Yu, J. Q., Zhang, W. N. & Zhuang, C. L. DNA-encoded libraries (DELs): a review of on-DNA chemistries and their output. *RSC Adv.* **11**, 2359–2376 (2021).
12. Fitzgerald, P. R. & Paegel, B. M. DNA-Encoded Chemistry: Drug Discovery from a Few Good Reactions. *Chem. Rev.* **121**, 7155–7177 (2021).
13. Fair, R. J., Walsh, R. T. & Hupp, C. D. The expanding reaction toolkit for DNA-encoded libraries. *Bioorg. Med. Chem. Lett.* **51**, 128339 (2021).
14. Dorel, R., Grugel, C. P. & Haydl, A. M. The Buchwald–Hartwig Amination After 25 Years. *Angewandte Chemie International Edition* **58**, 17118–17129 (2019).



15. Lu, X., Roberts, S. E., Franklin, G. J. & Davie, C. P. On-DNA Pd and Cu promoted C–N cross-coupling reactions. *Medchemcomm* **8**, 1614–1617 (2017). View Article Online  
DOI: 10.1039/D5CB00303B
16. Malone, M. L. & Paegel, B. M. What is a ‘DNA-Compatible’ Reaction? *ACS Comb. Sci* **18**, 52 (2016).
17. Liu, T. Z., Lin, T. F., Chiu, D. T. Y., Tsai, K. J. & Stern, A. Palladium or Platinum Exacerbates Hydroxyl Radical Mediated DNA Damage. *Free Radic. Biol. Med.* **23**, 155–161 (1997).
18. Rössler, S. L., Grob, N. M., Buchwald, S. L. & Pentelute, B. L. Abiotic peptides as carriers of information for the encoding of small-molecule library synthesis. *Science* (1979). **379**, 939–945 (2023).
19. Xu, H. *et al.* DNA-Encoded Libraries: Aryl Fluorosulfonates as Versatile Electrophiles Enabling Facile On-DNA Suzuki, Sonogashira, and Buchwald Reactions. *Advanced Science* **6**, 1901551 (2019).
20. Chen, Y. C. *et al.* C-N Coupling of DNA-Conjugated (Hetero)aryl Bromides and Chlorides for DNA-Encoded Chemical Library Synthesis. *Bioconjug. Chem.* **31**, 770–780 (2020).
21. De Pedro Beato, E. *et al.* Mild and Efficient Palladium-Mediated C-N Cross-Coupling Reaction between DNA-Conjugated Aryl Bromides and Aromatic Amines. *ACS Comb. Sci.* **21**, 69–74 (2019).
22. Chheda, P. R., Simmons, N., Schuman, D. P. & Shi, Z. Palladium-Mediated C–N Coupling of DNA-Conjugated (Hetero)aryl Halides with Aliphatic and (Hetero)aromatic Amines. *Org. Lett.* **24**, 3401–3406 (2022).
23. Graham, J. S., Hunter, J. H. & Waring, M. J. Micellar Buchwald-Hartwig Coupling of Aryl and Heteroarylamines for the Synthesis of DNA-Encoded Libraries. *Journal of Organic Chemistry* **86**, 17257–17264 (2021).
24. Gesmundo, N. J. *et al.* Nanoscale synthesis and affinity ranking. *Nature* **557**, 228–232 (2018).
25. Keller, M. *et al.* Highly pure DNA-encoded chemical libraries by dual-linker solid-phase synthesis. *Science* **384**, 1259–1265 (2024).
26. van der Nol, E. *et al.* Barcode-free hit discovery from massive libraries enabled by automated small molecule structure annotation. *Nature Communications* **2025 16:1 16**, 1–12 (2025).
27. Willoughby, C. A. & Chapman, K. T. Solid phase synthesis of aryl amines. *Tetrahedron Lett.* **37**, 7181–7184 (1996).
28. Weigand, K. & Pelka, S. Palladium-Catalyzed Animation of Aryl Halides on Solid Support. *Org. Lett.* **4**, 4689–4692 (2002).



29. Ward, Y. D. & Farina, V. Solid phase synthesis of aryl amines via palladium catalyzed amination of resin-bound aromatic bromides. *Tetrahedron Lett.* **37**, 6993–6996 (1996). View Article Online  
DOI: 10.1039/D5CB00303B
30. Prat, D. *et al.* CHEM21 selection guide of classical- and less classical-solvents. *Green Chemistry* **18**, 288–296 (2015).
31. Vitaku, M. E., Smith, D. T. & Njardarson, J. T. Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem* **57**, 10257–10274 (2014).
32. Muchiri, R. N. & van Breemen, R. B. Drug discovery from natural products using affinity selection-mass spectrometry. *Drug Discov. Today Technol.* **40**, 59–63 (2021).
33. Rizvi, N. F. *et al.* Discovery of Selective RNA-Binding Small Molecules by Affinity-Selection Mass Spectrometry. *ACS Chem. Biol.* **13**, 820–831 (2018).
34. Prudent, R., Annis, D. A., Dandliker, P. J., Ortholand, J. Y. & Roche, D. Exploring new targets and chemical space with affinity selection-mass spectrometry. *Nat. Rev. Chem.* **5**, 62–71 (2021).
35. Annis, D. A., Nickbarg, E., Yang, X., Ziebell, M. R. & Whitehurst, C. E. Affinity selection-mass spectrometry screening techniques for small molecule drug discovery. *Curr. Opin. Chem. Biol.* **11**, 518–526 (2007).
36. Oehler, S. *et al.* A DNA-encoded chemical library based on chiral 4-amino-proline enables stereospecific isozyme-selective protein recognition. *Nat. Chem.* **15**, 1431–1443 (2023).
37. Kugler, M. *et al.* Identification of specific carbonic anhydrase inhibitors via in situ click chemistry, phage-display and synthetic peptide libraries: comparison of the methods and structural study. *RSC Med. Chem.* **14**, 144–153 (2023).
38. Thomas, R. P. *et al.* A direct-to-biology high-throughput chemistry approach to reactive fragment screening. *Chem. Sci.* **12**, 12098–12106 (2021).
39. Thomas, R. P. *et al.* Reactive fragments targeting carboxylate residues employing direct to biology, high-throughput chemistry. *RSC Med. Chem.* **14**, 671–679 (2023).
40. Li, R., Medchemcomm, I., Ruff, Y. & Berst, F. Efficient copper-catalyzed amination of DNA-conjugated aryl iodides under mild aqueous conditions. *Medchemcomm* **9**, 1188–1193 (2018).
41. Zhang, L. *et al.* The terphenyl phosphine TRuPhos: a novel and efficient ligand for the palladium-catalyzed arylation of hindered primary alkyl amines with mechanism clarification. *J. Catal.* **451**, 116391 (2025).



42. Feng, K. *et al.* Development of a Deactivation-Resistant Dialkylbiarylphosphine Ligand for Pd-Catalyzed Arylation of Secondary Amines. *J. Am. Chem. Soc.* **146**, 26609–26615 (2024). View Article Online  
DOI: 10.1039/D5CB00303B
43. Park, N. H., Vinogradova, E. V., Surry, D. S. & Buchwald, S. L. Design of New Ligands for the Palladium-Catalyzed Arylation of  $\alpha$ -Branched Secondary Amines. *Angewandte Chemie International Edition* **54**, 8259–8262 (2015).
44. Su, M., Hoshiya, N. & Buchwald, S. L. Palladium-Catalyzed Amination of Unprotected Five-Membered Heterocyclic Bromides. *Org. Lett.* **16**, 832–835 (2014).
45. Charles, M. D., Schultz, P. & Buchwald, S. L. Efficient Pd-Catalyzed Amination of Heteroaryl Halides. *Org. Lett.* **7**, 3965–3968 (2005).
46. Zhou, F., Zhang, L., Hu, W., Yuan, B. & Shi, J. cheng. A general catalyst for Buchwald-Hartwig amination to prepare secondary five-membered heteroaryl amines with breaking the base barrier. *J. Catal.* **422**, 36–42 (2023).



**Data availability:** The raw LC-MS/MS files and the files needed to run SIRIUS:COMET have been deposited in the Zenodo database (10.5281/zenodo.17632652). Source data are provided with this paper.

All other data is available in the main text, the supplementary materials and from the corresponding author(s) upon request.

