



Cite this: *Green Chem.*, 2023, **25**, 1380

## 'Chemistry at the speed of sound': automated 1536-well nanoscale synthesis of 16 scaffolds in parallel†

Li Gao,<sup>‡a</sup> Shabnam Shaabani,<sup>‡a</sup> Atilio Reyes Romero,<sup>‡a</sup> Ruixue Xu,<sup>a</sup> Maryam Ahmadianmoghaddam<sup>a</sup> and Alexander Dömling<sup>a\*</sup>

Screening of large and diverse libraries is the 'bread and butter' in the first phase of the discovery of novel drugs. However, maintenance and periodic renewal of high-quality large compound collections pose considerable logistic, environmental and monetary problems. Here, we exercise an alternative, the 'on-the-fly' synthesis of large and diverse libraries on a nanoscale in a highly automated fashion. For the first time, we show the feasibility of the synthesis of a large library based on 16 different chemistries in parallel on several 384-well plates using the acoustic dispensing ejection (ADE) technology platform. In contrast to combinatorial chemistry, we produced 16 scaffolds at the same time and in a sparse matrix fashion, and each compound was produced by a random combination of diverse large building blocks. The synthesis, analytics, resynthesis of selected compounds, and chemoinformatic analysis of the library are described. The advantages of the herein described automated nanoscale synthesis approach include great library diversity, absence of library storage logistics, superior economics, speed of synthesis by automation, increased safety, and hence sustainable chemistry.

Received 15th November 2022,  
Accepted 13th January 2023

DOI: 10.1039/d2gc04312b

rsc.li/greenchem

## Introduction

Large and diverse compound libraries are used in high throughput screening to discover novel biologically active matter.<sup>1</sup> Often hits discovered from such compound libraries are the starting point of a medicinal chemistry optimization project which can lead to a drug for unmet medical needs.<sup>2</sup> Large pharmaceutical compound libraries are usually assembled in-house by contract research organization (CRO) synthesis and through purchase from vendors. However, the assembly by synthesis is lengthy and by purchase expensive and often biased to the limited chemical space of vendors. The quality of a screening library is determined from the chemical space covered, the physicochemical properties, and its size.<sup>1</sup> Strategies to maintain and enhance the overall quality of screening collections are highly sought after to stay competitive and work in proprietary protectable chemical space.<sup>3</sup> Contemporary approaches towards a large compound screening space include DNA-encoded chemical library (DEL),<sup>4</sup> frag-

ment-based drug discovery (FBDD),<sup>5</sup> *de novo* structure-based drug design (SBDD),<sup>6</sup> diversity-oriented synthesis (DOS),<sup>7</sup> and virtual screening (Fig. 1).<sup>8</sup> All the technologies have their advantages and disadvantages.<sup>2</sup> A popular way to assemble compound libraries uses combinatorial chemistry (CC).<sup>9</sup> In combinatorial chemistry, all possible combinations of available building blocks are combined around one chemical scaffold, and a dense matrix is formed. Such libraries often lack sufficient diversity and monotonous substituent placement is seen, when building around one or a few scaffolds. DEL technology makes perfect use of combinatorial chemistry by applying systematic variations of building blocks to create extensive libraries in a tiny volume. Combinatorial chemistry and DEL were successful at delivering starting points to develop drugs for clinical trials and the market.<sup>4</sup> DOS aims to produce libraries based on skeletal diversity, often leaning on natural product-derived motifs by using modular syntheses involving a few steps ('build-couple-pair') and typically from common stereochemically complex precursors. In contrast, fragment libraries have already repeatedly proven their value by delivering clinical drugs.<sup>5</sup> However, fragment library screening and hit optimization are demanding because fragments typically weakly bind to their target. Hence, significant effort has to go into stepwise building of more complex compounds to yield the required potency. Often fragment linking technologies are a not-straightforward undertaking and fail to exhibit

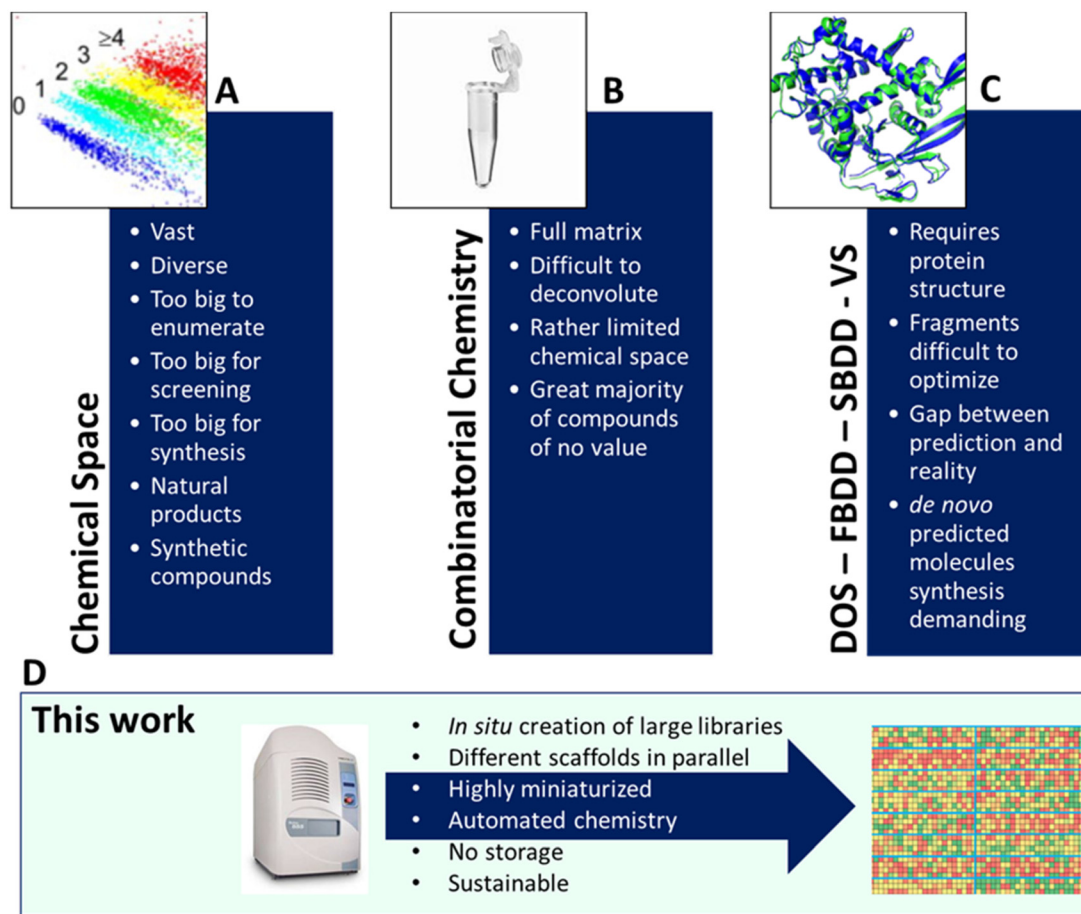
<sup>a</sup>Department of Drug Design, University of Groningen, Groningen, The Netherlands

<sup>b</sup>CATRIN, Department of Innovative Chemistry, Palacký University Olomouc, Olomouc, Czech Republic. E-mail: alexander.domling@upo.l.cz

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d2gc04312b>

‡ These authors contributed equally.





**Fig. 1** Chemical space and contemporary library synthesis strategies. (A) The vastness of chemical space makes systematic exploitation very demanding. (B) Combinatorial chemistry and DNA-encoded library (DEL). (C) Fragment-based drug design (FBDD), structure-based drug design (SBDD), diversity-oriented synthesis (DOS) and virtual screening (VS). (D) Herein pursued approach involving HT miniaturized multiple chemistries at once using rapid acoustic droplet ejection technology.

the expected gains of binding potency.<sup>10</sup> Virtual screening (VS) is making considerable progress and ultra-large libraries can be screened to obtain novel inhibitors.<sup>11</sup> However, physics-based *in silico* screening of billion-sized libraries in an exhaustive manner is cost and time prohibitive. While machine learning can be combined with VS to feasibly tackle even larger chemical space, in the end synthesis and screening of the multi-step hits of non-commercially accessible compounds comprise the true bottleneck.<sup>12</sup> Due to the major progress in protein crystallography, structure-activity-relationship (SAR) by crystallography in a high throughput (HT) mode is feasible at a shorter timeframe.<sup>13</sup> Primarily static fragment libraries have been used for enzyme active sites or even protein-protein interactions.<sup>14</sup> Routine times of around 2–3 weeks from soaking to multi cocrystal structure determinations are now possible. Recently, the integration of HT synthesis and HT protein crystallography has emerged as a promising novel approach to shorten the ubiquitous ‘design-make-test-analyze’ cycle (DMTA).<sup>15</sup>

While there is no ‘one-size-fits-all’ library method, we are reporting an alternative technology involving miniaturized and automated nanoscale syntheses applied to many different

chemistries, and multicomponent reactions (MCRs), in parallel. In classical approaches, one scaffold is elaborated at a time. Here we produce exemplary libraries based on 16 different scaffolds at the same time, thus yielding unprecedented scaffold and library diversity. Moreover, due to the miniaturization aspect, the approach reduces the ecological footprint of chemistry enormously while increasing the speed of high quality library production. The rationale for choosing MCR is that this reaction class is well known to provide excellent scaffold diversity, based on many building blocks which are commercially available and have tremendous functional group compatibility. Indeed, the scaffold diversity of MCR compares favorably with sequential syntheses in the pharmaceutical industry relying on only a few well-characterized reactions.<sup>16</sup>

## Results

Several well-defined, robust MCRs are selected for synthesis based on previous experience in our laboratory (Table 1).<sup>17</sup>



**Table 1** Structures of 16 MCR scaffolds for miniaturized high throughput chemistry. The reaction schemes, components, and names are given. The building block classes R's are shown in different colors and numbers to indicate their repeated use in different MCRs. Stereoscopic views of each scaffold obtained with Pymol are shown with methyl substituents. Blue = nitrogen, red = oxygen, white = hydrogen, and other colors = carbons. Representative references of these 16 MCR scaffolds are cited as ref. 8–23 in the ESI.†

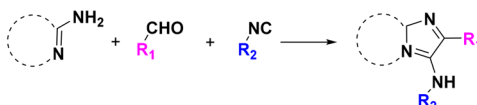
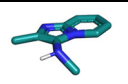
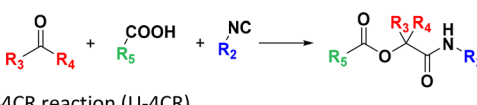
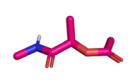
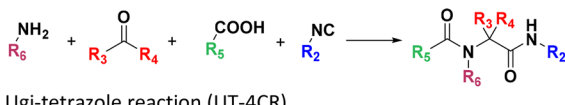
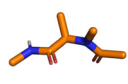
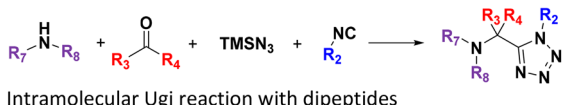
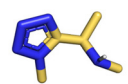
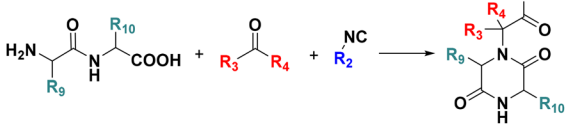
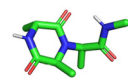
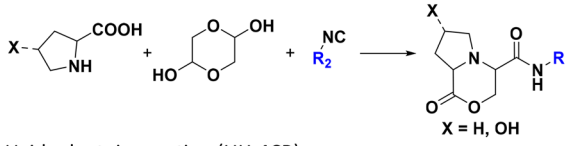
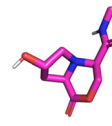
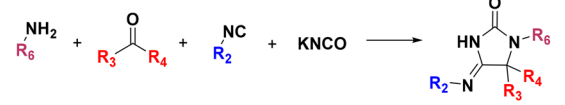
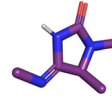
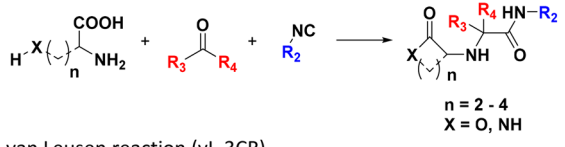
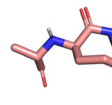
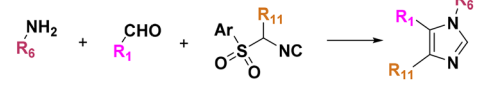
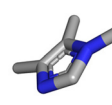

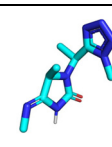
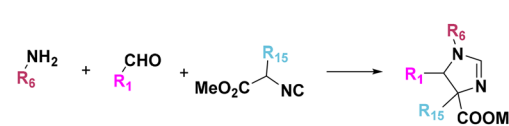
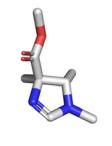
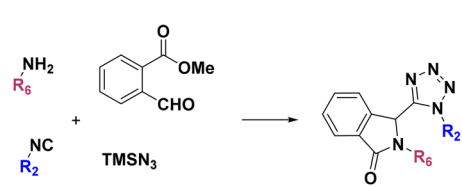
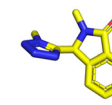
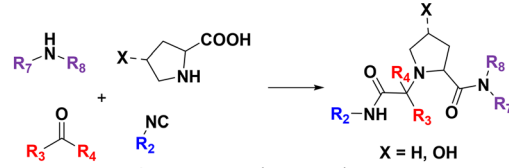
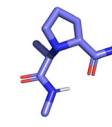
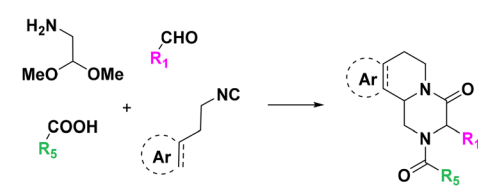
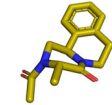
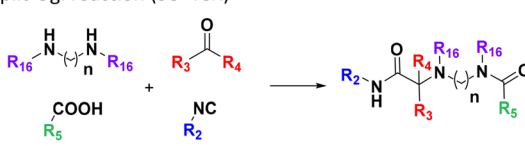
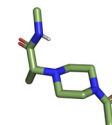
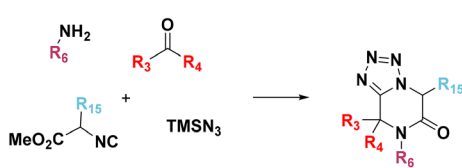
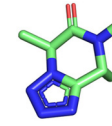
No.	Reaction scheme	3D view
1	Groebcke-Blackburn-Bienaymé reaction (GBB-3CR) 	
2	Passerini reaction (P-3CR) 	
3	Ugi-4CR reaction (U-4CR) 	
4	Ugi-tetrazole reaction (UT-4CR) 	
5	Intramolecular Ugi reaction with dipeptides 	
6	Intramolecular Ugi-type reaction of proline and hydroxy acetaldehyde (U-5C-3CR) 	
7	Ugi-hydantoin reaction (UH-4CR) 	
8	Intramolecular Ugi reaction of alpha-amino acids with nucleophilic side chains (U-5C-3CR) 	
9	van Leusen reaction (vL-3CR) 	



Table 1 (Contd.)

No.	Reaction scheme	3D view
10	Ugi-hydantoin reaction involving tetrazole amines (UH-4CR) 	
11	Orru reaction (OR-3CR) 	
12	Intramolecular Ugi-tetrazole reaction using methyl 2-formylbenzoate (iUT-4CR) 	
13	U-5C-4CR reaction with additional amine 	
14	Ugi-Pictet-Spengler sequence (UPS-4CR) 	
15	Split-Ugi reaction (SU-4CR) 	
16	Intramolecular Ugi-tetrazole reaction using methyl isocynoacetates (iUT-4CR) 	

All reactions have in common the use of isocyanides (Isocyanide-based Multicomponent Reactions, IMCRs).<sup>18</sup> The reactions were chosen to reflect a balanced 3D-scaffold and -substituent space ranging from flat heteroaromatic (Table 1:

reactions (1), (4), (7) and (9)), non-flat mono heterocyclic (Table 1: reactions (5), (8), (11) and (13)), bicyclic not annulated (Table 1: reactions (10) and (12)), bicyclic annulated (Table 1: reactions (6) and (16)), and polycyclic (Table 1: reac-



tion (14)) to non-cyclic scaffolds (Table 1: reactions (2), (3) and (15)).

The synthesis was performed on an Echo 555 acoustic dispensing platform. In acoustic droplet ejection (ADE), precise nanodroplets are formed by applying acoustic waves to the stock solutions of the building blocks in the source plate and transported above to the inverted destination plate, where the reaction is taking place.<sup>19</sup> A total of 336 building blocks and reagents were used here, including cyclic 5- and 6-membered aromatic amidines (36), aldehydes (52) and ketones (15), isocyanides (64), carboxylic acids (71), 78 amines (46 primary and 32 secondary amines), all fitting into one 384-well source plate (Fig. 2, ESI, section 2.3†). Due to their rapid evaporation, volatile solvents such as methanol, THF or DCM are not suitable for ADE technology. Hence, we chose ethylene glycol in most cases as a ‘transporter solvent’ for stock solution preparation, or 2-methoxyethanol or dimethoxyethane when the building block was not soluble in ethylene glycol. All soluble building blocks were prepared as 0.5 M solutions or otherwise diluted to 0.25 M or 0.16 M solutions. The stock solutions were kept in 384-well source plates sealed with tape and kept at  $-20\text{ }^{\circ}\text{C}$  before combining them in a corresponding 384-well destination plate using an Echo 555 liquid handler. The scale of each reaction per well was 300, 375, or 500 nanomoles directed by the number of components/reagents transferred. Each of the

16 MCR scaffolds was designed for 96 different products to investigate the substrate scope and limitations by choosing the corresponding building blocks. Therefore, four destination 384-well plates were filled by the Echo 555 liquid handler to generate 1536 reactions in total. In a sparse matrix approach, 1536 reactions were performed, however, reflecting only a small fraction of  $\sim 0.1$  per mille of the possible combinations based on the 16 different chemistries (Table 1). Indeed, the 332 building blocks can be theoretically combined to give 16 446 250 products. Once the transfer of the starting material to the destination plates was complete ( $\sim 150$  min for each 384-well plate), 10  $\mu\text{L}$  of the appropriate solvent for each reaction was added using a multichannel pipette (ESI, section 2.5†). For example, MeOH was added to GBB-3CR, Ugi-4CR, and UT-4CR (Table 1: reactions (1), (3) and (4)) and  $\text{CHCl}_3$  was added to Passerini-3CR (Table 1: reaction (2)). The destination plates were sealed and placed for 12 h at  $21\text{ }^{\circ}\text{C}$  on an orbital shaker. After desealing, the solvent (MeOH and  $\text{CHCl}_3$ ) was evaporated for further analytics. To the intramolecular Ugi reaction with dipeptides, trifluoroethanol (TFE) was added (Table 1: (5)). One scaffold (Table 1: reaction (14)) is in fact a two-step reaction sequence consisting of an initial Ugi-4CR, followed by an acid-catalyzed Pictet-Spengler reaction.<sup>20</sup> In this case, 10  $\mu\text{L}$  of MeOH was added to the initial Ugi reaction. The destination plate was sealed and placed for 12 h at  $21\text{ }^{\circ}\text{C}$

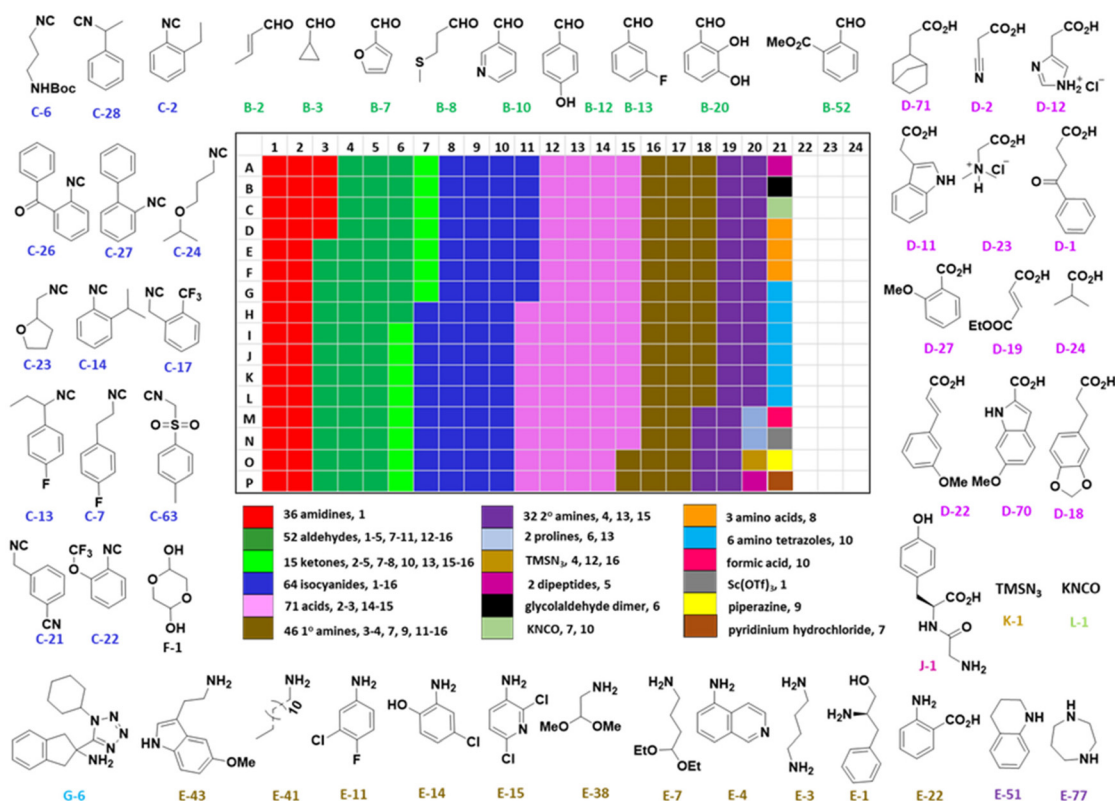


Fig. 2 Selected representatives of the 336 building blocks, and reagents distributed on a 384-well source plate. The colors indicate the different building block classes. The total number of each building block and the reaction number where each building block is used are mentioned (e.g. below the heat map “36 amidines, 1” means 36 amidines are used in reaction (1), Table 1).



on an orbital shaker. After desealing to let the MeOH evaporate (~3 h evaporation time), 10  $\mu$ L of formic acid was added as a catalyst for the Pictet–Spengler reaction. The reaction plate was sealed again and kept at 60 °C for 2 h. The plate was kept at –20 °C for further analytics, after desealing and formic acid evaporation (~3 h), (Table 1: reaction (14)). The experimental details of each scaffold of the 16 MCR reactions on the plates are described in the ESI (section 2.5).†

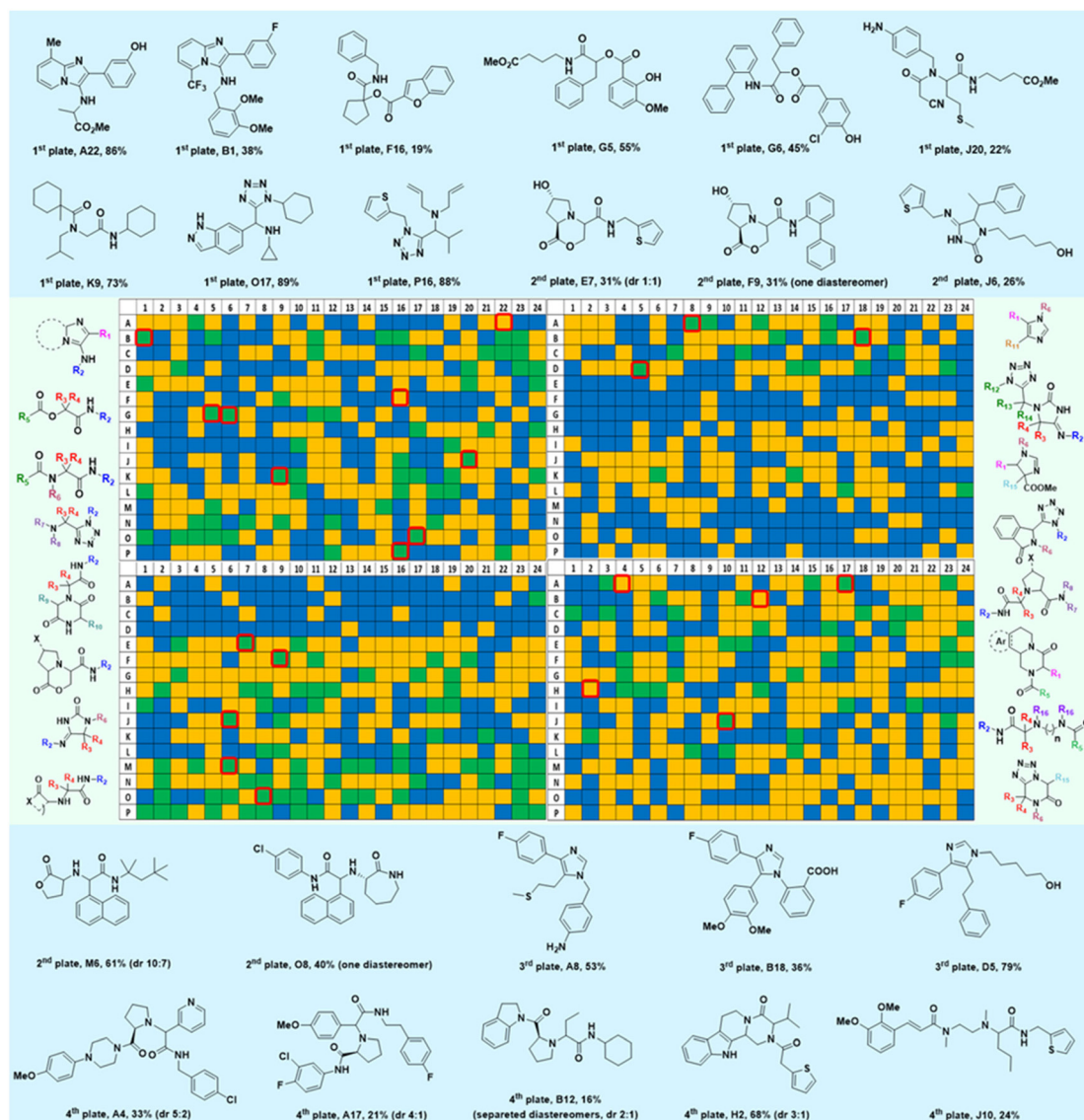
The building blocks were chosen to optimize diversity in terms of electronic structures (electron donating–withdrawing), and steric (small, large, cyclic, acyclic) and orthogonal (halogen, hydroxyl, amine, carboxylic acid, heterocycle) functional groups (Fig. 2).

For analytics of the 1536 reactions, 100  $\mu$ L of ethylene glycol was added to each well using a multichannel pipette.

The reaction wells were automatically analyzed by SFC-MS overnight using our previously described python script (Fig. 3).<sup>21</sup>

Mass spectrometry analysis of the reactions gives information on the performance of the different chemistries and the reactivity of the different building blocks and combinations under the performed reaction conditions (Fig. 4). It was found that amongst the 16 performed chemistries, 13 performed well, producing the products in major or medium quantities. Exemplary reactions are discussed in the following.

The GBB-3CR (Table 1: reaction (1)) is a popular reaction in medicinal chemistry which was used in the discovery of tool compounds for the first bromodomain of human BRD4<sup>22</sup> and was instrumental in the development of the clinical autotaxin inhibitor Ziritaxesta® that is used for the treatment of idio-



**Fig. 3** 1536 reactions based on 16 different chemistries for the parallel assembly of a highly diverse library of small molecules. The generic scaffolds are shown in the green boxes left and right of the 1536 well plate. The MS-analysis is plotted on the 1536-well plate based on a three-color code (blue = no product formation, yellow = medium product formation, green = major product formation). The structures of the randomly resynthesized compounds (red boxes) from each scaffold on a mmol scale are shown in blue boxes together with their isolated yields.



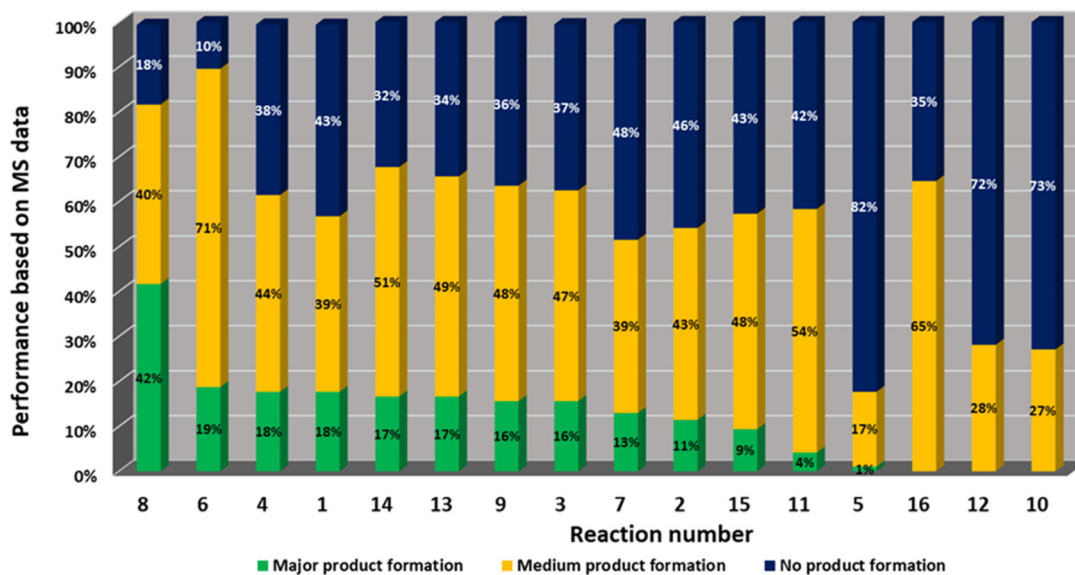


Fig. 4 Statistical reaction analysis: performance of 16 MCR reactions based on MS analysis in descending order.

pathic pulmonary fibrosis.<sup>23</sup> A recent comprehensive review reveals that GBB-3CR is a versatile MCR variation of the Ugi reaction, which can be performed under a broad range of conditions. However, there is no 'one-size-fits-all' procedure, but the exact reaction conditions (catalyst, temperature, solvent) have to be adjusted to the building blocks used. For example, more than 20 different Lewis and Bronsted acid catalysts have been reported to be of use in the GBB-3CR.<sup>24</sup> In particular, the nature of the five or six membered aromatic amidine plays a crucial role and the yields can vary largely depending on the reaction conditions. Nonetheless, we choose the following generalized conditions for the nanoscale syntheses during non-optimal reaction conditions for all building block combinations: 0.166 M of all starting materials in ethylene glycol, 10 mol% Sc(OTf)<sub>3</sub>, room temperature, 12 h.

In the 96 GBB-3CRs, product formation was indicated by MS analysis in 57%, and in 18% as the major product. Analysis of the amidine structures reveals that 2-aminopyridine with electron-withdrawing groups (NO<sub>2</sub>, CF<sub>3</sub>, CN, I) reacted well (Fig. 5). Aminotriazoles (A-21, A-26) on the other hand reacted not at all or showed medium reaction success. This can be explained by the general electron-deficient nature of the triazole heterocycles and the mechanism of the GBB-3CR which requires the condensation of the exocyclic triazole nitrogen with the aldehyde component. Not surprisingly, triazole-based GBB-3CR reactions, previously, were performed under elevated temperatures or microwave assistance.<sup>25</sup>

The high success rate (major and medium product formation 57%) was rather surprising, in view of the strong dependency of the GBB-3CR performance on the nature of the Bronsted or Lewis acid catalysts, the solvent and the temperature, besides the electronic and steric influences of the nature of the building blocks.<sup>24</sup>

The intramolecular Ugi reaction of  $\alpha$ -amino acids (U-5C-3CR, Table 1, reaction (8)) with suitable nucleophilic

side chains is an interesting Ugi variation for several reasons.<sup>26</sup> First, it utilizes chiral pool  $\alpha$ -amino acids which direct the newly formed stereocenter at a high diastereomeric ratio. Second, the nucleophilic  $\alpha$ -amino acid side chain is involved in the formation of 5- to 7-membered rings. Third, some members of this family, are known to exhibit potent biological activity, *e.g.*, GABA mimics with potent *in vivo* anticonvulsant activity.<sup>27</sup>

For this Ugi MCR variation lysine, penta homo-serine, ornithine, and homo-cysteine have been described as multi-functional substrates in different publications, but no comparative study was performed ever.<sup>28–32</sup> In our HT synthesis we used available lysine, penta homo-serine, and ornithine, yielding the corresponding 7- and 6-membered lactones and 5-membered  $\delta$ -lactam, respectively. Interestingly, the 7-membered lactones were formed best, followed by the 5-membered  $\delta$ -lactams and the 6-membered lactones (Fig. 6).

Amongst the 16 performed chemistries, the diketopiperazine (DKP) formation using dipeptides in the Ugi reaction, the Orru 3-component reaction, the methyl 2-formylbenzoate UT-4CR and the tetrazolo ketopiperazine resulted in mostly medium or no product formation (Fig. 4). Potential reasons for their poor performance are discussed here. The intramolecular Ugi reaction of dipeptides leading to DKP formation (Table 1, reaction (5)) was described only one time in the literature.<sup>33</sup> A closer inspection of the work reveals that very specific conditions were applied, a mixture of [Bmim]PF<sub>6</sub> and TFE (CF<sub>3</sub>CH<sub>2</sub>OH) was used as a solvent system and the reaction was performed under microwave conditions. Ethylene glycol at room temperature were the conditions used here and are far from the optimized conditions yielding the DKP. The Orru 3-component reaction (Table 1, reaction (11)) of  $\alpha$ -acidic isocyanides, primary amines and aldehydes has been often described but is rather solvent specific and mostly involved the



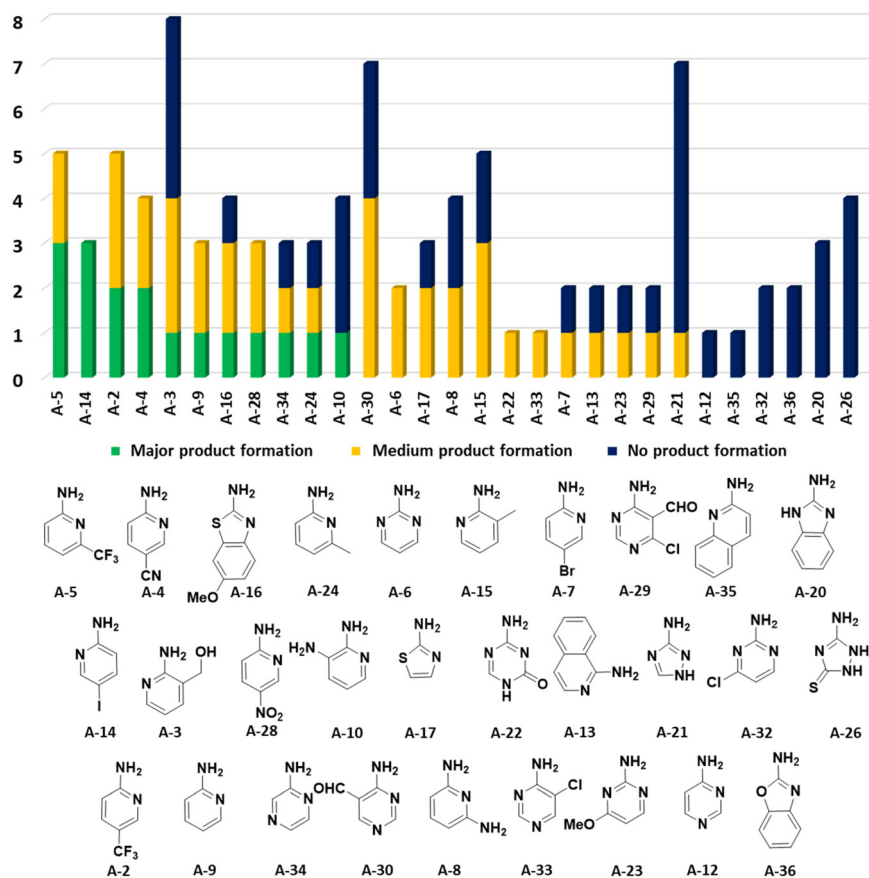


Fig. 5 Performance of the cyclic amidine building blocks in the GBB-3CR. The number of reactions is indicated on the Y-axis and the nature of the amidine on the X-axis.

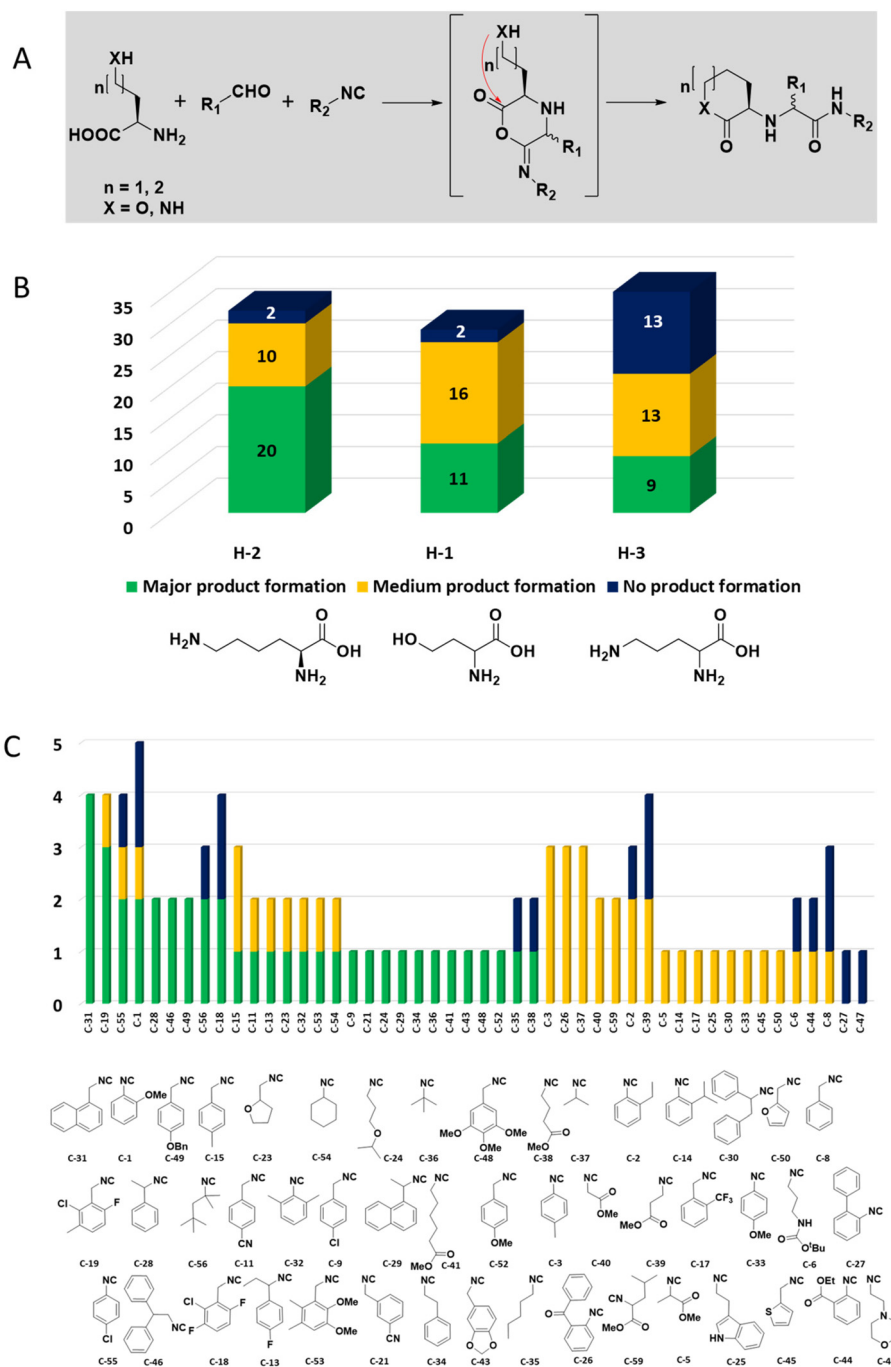
use of chlorinated solvents and often  $\text{Ag}^{\text{I}}$  catalysis.<sup>34</sup> The intramolecular Ugi tetrazole reaction involving *o*-formylbenzoic acid methylester (Table 1, reaction (12)), described twice, also employs drastic reaction conditions (refluxing solvent, or NaOEt) to accomplish lactam ring formation.<sup>35</sup> Under the herein applied conditions no product formation was observed. Finally, the successful transformation into fused tetrazolo ketopiperazine (Table 1, reaction (16)) required prolonged refluxing conditions in methanol.<sup>36</sup>

### Scalability of the 16 reactions

Given the highly miniaturized nature of ADE, reproducibility on a larger scale needs to be addressed, which is done here through the scale-up of a selection of wells to validate translation into more traditional conditions. To show the scalability of the nanoscale reactions, we randomly resynthesized 22 compounds with featured structures chosen from the 16 scaffolds in the 1536-plate. All products were resynthesized on a 1 mmol scale and the yields were determined after chromatographic purification by weighing (Fig. 3). All the resynthesized compounds are unprecedented and based on known chemistries. J20 (1<sup>st</sup> plate) incorporates the cyanoacetamide moiety, which is a popular motif in covalent inhibitors, for example in the Janus kinase inhibitor Tofacitinib.<sup>37</sup>

G5 and G6 (1<sup>st</sup> plate) feature free phenolic hydroxyl groups which are perfectly compatible functional groups under the mild reaction conditions of the P-3CR. Likewise, esters (F16, G5, G6, J20), aniline- $\text{NH}_2$  (J20), alkenes (P16), primary aliphatic hydroxyl (J6), secondary hydroxyl (E7, F9), secondary amine (O8, O17), tertiary amine (A4, A17, B12, J10, O17, F16, E7, F9), thioether (J20, A8), and free carboxylic acid (B18) can be incorporated problem-free into the products not requiring tedious protecting group chemistries. Particularly noteworthy is the unprecedented smooth incorporation of the free carboxylic acid (B18) and free aniline (J20) without any protecting group manipulations, resulting from the anthranilic acid E-22 and 4-aminobenzylamine building block E-26 in the vL-3CR (Table 1, reaction (9)). Several heterocycles are featured in the resynthesized compounds, including thiophene (P16), indole (H2), tetrahydroindole (B12), piperazine (A4, H2), pyrrolidine (A4, A17, B12, E7, F9), imidazole (A8, B18, D5), and  $\gamma$ -lactones (M6, O8),  $\delta$ -lactam (O8), tetrazole (P16, O17), imidazopyridine (A22, B1), benzofurane (F16), morpholine (E7, F9), azaindole (O17), and pyridine (A4), underscoring the broad usefulness of IMCR in producing diverse libraries. In summary all 22 nanoscale reactions selected for resynthesis could be synthesized and characterized on a mmol scale as well.





**Fig. 6** Intramolecular U-5C-3CR performance of three different  $\alpha$ -amino acids with nucleophilic side chains. (A) Synthesis scheme with the key  $\alpha$ -adduct intermediate. (B) Statistics of the performance by the  $\alpha$ -amino acid building block as analyzed by the peak height of direct MS injection. (C) Reactivity/performance of the isocyanides in descending order.

### Chemoinformatic analysis

The calculation and analysis of physicochemical descriptors is an important step in library design and can help to avoid compounds with poor properties, leading to poor solubility, absorption, issues with screening assays, or causing pan assay interference (PAIN).<sup>3</sup> To investigate the physicochemical and biopharmaceutical diversity of the 1536 compounds based on

16 scaffolds, we conducted a comparative chemoinformatic analysis with Food and Drug Administration (FDA) approved drugs (ESI, section 5.1†).<sup>38,39</sup> The calculated and analyzed descriptors include molecular weight (MW), partition coefficient (cLogP), number of hydrogen bond acceptors/donors (NHA/NHD), total polar surface area (TPSA), and number of rotatable bonds (NumRotBonds) and partition of a chemical compound between the lipid and aqueous phases (cLogD).



Initially we investigated the dependency between  $c\text{Log } P$  and MW which are the first indicators of solubility and membrane permeability during hit optimization. Using our random-compound-generator algorithm, all scaffolds produce more lipophilic compounds as the MW increases (Fig. S7†), except for some compounds clustered at the top of the chart as the polarity increases. Noteworthy, 33% reside in a preferred area of MW < 500 Dalton and  $c\text{Log } P$  2–4. Interestingly, certain scaffolds are narrowly distributed (6, 14), whereas others show a broad distribution (3, 4, 16). Next, we conducted a more in-depth diversity study by assessing the distribution of the other five properties (Fig. S8†) and quantifying the statistical significance between them (Fig. S9, ESI, sections 5.2 and 5.8†). This resulted in a total of 952 possible pair-wise comparisons for each property between the 1536 compounds based on 16 scaffolds and the FDA compounds. More than half of these comparisons showed highly significant statistical differences except for  $c\text{Log } D$ . This implies a broad spectrum of molecular diversity in terms of MW, membrane permeability, flexibility, polarity, NHA and NHD which are crucial parameters for binding ligand/receptor interaction and aqueous solubility. Furthermore, we studied the drug likeliness profile of each molecule with Lipinski, Ghose<sup>40</sup> and Muegge scores<sup>41</sup> (ESI, section 5.3†). Importantly, our results confirm that the 11 scaffolds produce more than 50% of compounds that pass all three biopharmaceutical filters (Fig. S10 and S11†), and hence ‘druglike’ molecules.

We also assessed the Central Nervous System Multi-Parameter Optimization (CNS MPO) desirability score as illustrated in Fig. S12.†<sup>42</sup> We differentiated the compounds between high CNS MPO (>5) and low CNS MPO ( $\leq 2$ ) scores (Fig. S11A†).<sup>43</sup> Thus, compounds with CNS MPO > 5 could cross the blood–brain barrier (BBB) and exhibit pharmacological activity in the central nervous system (CNS). Most scaffolds display low CNS MPO scores in at least 75% of the compounds, because they were not designed specifically for CNS penetration (Fig. S13A†). Interestingly, scaffold 16 displayed a higher CNS MPO score (31%) than the FDA drugs (26%), while scaffold 10 returned the lowest score (0%). These data indicate that reaction (16) could produce molecules highly penetrable through the BBB.

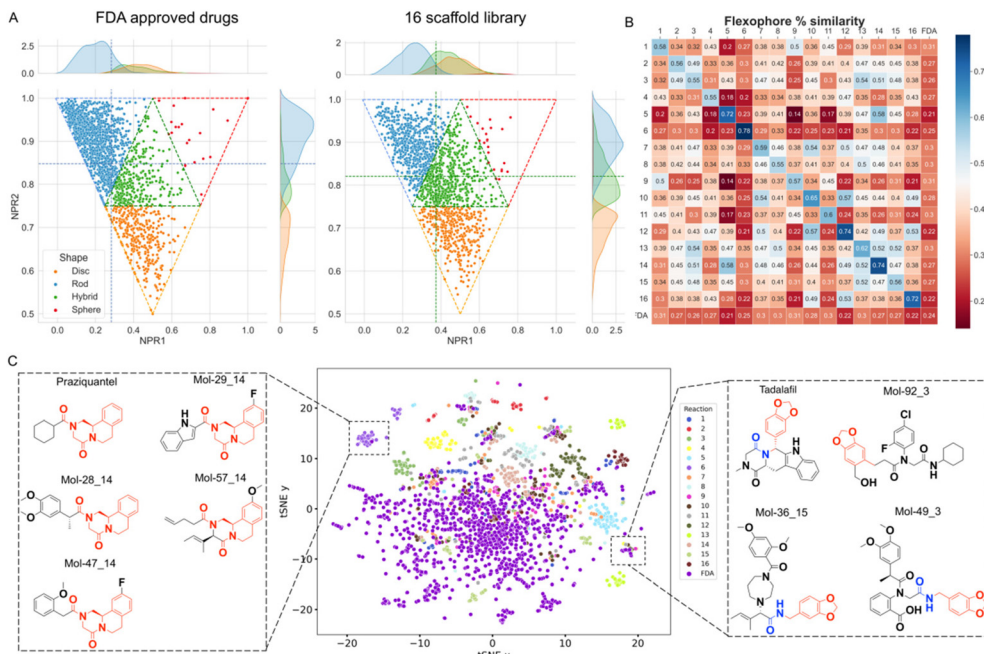
As shown by PCA analysis in Fig. S13B,† both structural flexibility (NumRotBond) and basicity (basic  $pK_a$ ) were found to be the most representative variables in influencing BBB penetration of scaffold 16 over scaffold 10. To investigate even further which properties positively affect scaffold 16 CNS MPO score, we conducted a principal component analysis (PCA) (ESI, section 5.4†) employing TPSA, MW, NumRotBonds,  $c\text{Log } D$ ,  $c\text{Log } P$  and Strongest basic  $pK_a$  as descriptors. The first three principal components described 93.1% of the variance within the data (Table S1†) with the first component carrying a high load from MW (0.963) followed by  $\text{Log } P$  (0.820), NumRotbonds (0.791),  $\text{Log } D$  (0.767) and the strongest basic  $pK_a$  (0.570) (Table S2†). The data points are centered around the origin with no dispersion or observable outliers. Interestingly, a clear separation of the molecules of reaction

(16) from reaction (10) is visible. Surprisingly, the highest contribution to the first (27%) and second largest components (33%) is from the MW and TPSA (Table S3†). Taking it together, we hypothesized that scaffold 16 is advantaged by less conformational freedom since tetrazole is *ortho*-fused with the piperazinone ring. We speculate that the general structural rigidity contributes importantly to the BBB. For example, the scaffold of 16 molecules consists of two *ortho*-fused rings, a tetrazole, and a piperazinone. As a result, they are subject to less freedom in conformational space rather than those of scaffold 10. Additionally, scaffold 16 contains fewer exit points than scaffold 10 (4 vs. 6) where polar and bulky functional groups can be installed. This allows scaffold 16 molecules to more easily permeate the dense phospholipid layer of the BBB.

Next, we performed an exhaustive shape analysis. The complementarity of the molecule and receptor shape is an important prerequisite for efficient binding. Normalized ratios of the three main moments of inertia (NPR) address the diversity of receptor shapes of a general screening library and were calculated (ESI, section 5.5†). On the basis of the abundance and arrangement of rings, alkyl chains or internal cyclization a scaffold library can be intuitively classified into three categories, namely spheroidal, discoidal, and flat.<sup>44</sup> As shown in Fig. 7A, the molecules adopt specific molecular shapes. The molecule count presented in Table S6† shows that FDA approved drugs primarily tend to assume a rod-like shape (65%), followed by discoidal (19%), hybrid (16%) and spherical (1%) shapes. On the other hand, our library based on 16 scaffolds assume a rod (41%), hybrid (34%), disc (24%) and sphere (2%) shape with more compounds in the last three spatial regions. Overall, the barycenter, calculated from the averages of the kernel densities for both moments of inertia, confirms that the FDA approved drugs tend to be more rod-like whilst the 16 scaffold library populate a region of hybrid forms. To investigate the different trends, we conducted a comparative non-parametric Kolmogorov–Smirnov test to elucidate whether the cumulative distribution function (CDF) values for each shape belong to the same distribution (ESI, section 5.8†). The differences among the rod, sphere and disc shapes are highly statistically relevant ( $p$ -values < 0.001, Table S7†), indicating that our library can produce indeed higher molecular shape diversity than the FDA approved drugs, with more spherical and hybrid shapes and less rod-like or discoidal shapes.

Understanding the pharmacophore distribution of a chemical space is an additional way to assess the shape diversity of libraries. For this, we investigated the pharmacophoric diversity of our scaffolds using Flexophore descriptors (Fig. 7B, ESI, section 5.6†).<sup>45</sup> The highest values of structural similarity are distributed along the diagonal of the matrix, except for the FDA group. This is because each of the 16 scaffolds has been designed to share one common scaffold, while the majority of FDA approved drugs are not built on common scaffolds. Noteworthy, 51% of the cross-comparisons between different scaffolds in the array show pharmacophoric similarity below the average calculated over the entire matrix (0.36) (Table S4†).





**Fig. 7** Chemoinformatic analysis and chemical diversity of our 16 scaffold library produced by ADE technology. (A) Averaged normalized ratios of principal moments of inertia (NPR) of FDA approved drugs and our library. The kernel density distribution of the two principal moments is displayed. The average shape between all points is identified by the intersection of the density values calculated on both axes (dashed lines). Molecules occupying the "rod", "sphere", "hybrid", and "disc" triangles are colored as blue, red, green, and orange dots. (B) Pharmacophoric similarity matrix indicated on a divergent scale from blue (high similarity) to brick red (low similarity). (C) *t*-Distributed stochastic neighbor embedding (*t*SNE) chart showing the diversity of our library versus the FDA approved drugs. The two boxes contain clusters of similar molecules to the FDA approved drugs Praziquantel and Tadalafil. Examples of structures corresponding to clusters densely populated by the reference group (purple) and the 16 scaffold library are framed in the scatter plot.

These data indicate a wide pharmacophoric diversity and a lower propensity to yield similar molecules with different reactions. This implies that our library based on the selection of 16 scaffolds leads indeed to highly dissimilar molecules.

Yet another way to assess the structural similarity between the 16 scaffolds and FDA drugs is by *t*-distributed stochastic neighbor embedding (*t*SNE). To find any similar substructures between our library and the FDA approved drugs, we have employed SkeletonSpheres descriptors (ESI, section 5.7<sup>†</sup>).<sup>46</sup> The data points corresponding to the 16 scaffolds are distributed more widely than the reference FDA group (Fig. 7C) indicating that our synthetic approach covers a chemical space not yet explored by current drugs. On the other hand, it is possible to discover close analogues of FDA approved drugs in our 16-scaffold library. An example is given by Mol-8, Mol-13, Mol-39 and Mol-91 from scaffold 14 sharing the same substructure with the anti-schistosomiasis drug Praziquantel, previously synthesized by Cao *et al.* employing the same reaction.<sup>20</sup>

Analysis of the neighbor tree map (Fig. S14<sup>†</sup>) highlights another potential therapeutical application based on the pharmacophoric similarity of our nano-scale library to FDA approved drugs. For example, Mol-87 from the Ugi-hydantoin reaction shares 0.92 Flexophore similarity with Doxapram, an analeptic drug used to stimulate respiration in alveolar hypoventilation. *Vice versa*, the closest FDA drugs to Mol-18, Mol-31

and Mol-76 obtained from the Orru reaction find approved applications as muscle relaxants to treat injuries (Chlorphenesin, 0.91), antihistamines for the treatment of the relief of nasal congestion (*e.g.*, Antazoline, 0.91), analgesic or anti-inflammatory (Mefenamic acid 0.90, Fenoprofen 0.91) and calcium channel blocker for containment of ischemic damage (Nimodipine 0.93). Interestingly, the van Leusen scaffold (Table 1: 9) retained the largest number of compounds like FDA drugs. For example, the pharmacophoric motifs that Mol-36 shares with its hits are a benzene ring and/or a basic center, embedding pi-stacking and ionic features, respectively. Several of these hits are prescribed as cardiostimulants like Mephentermine (0.99) or heartbeat correctors such as Mexiletine (0.94), anorectics for the treatment of obesity such as Phentermine (0.99) and Phendimetrazine (0.98), psychostimulants for the treatment of attention deficit hyperactivity disorder (*e.g.*, Amphetamine, Methamphetamine, Dextroamphetamine, all with 0.99 Flexophore similarity), and antidepressants (*e.g.*, Phenelzine, 0.93). Other isolated applications include anti-poisoning by organophosphate Pralidoxime (0.99), and treatment of arterial hypertension (Bethanidine 0.95) and appetite suppression by Phenmetrazine (0.97, now withdrawn for its abuse). Finally, Mol-2 is sharing 0.97 and 0.92 of Flexophore similarity to Methoxamine and Tramadol, for the treatment of hypotension and moderate to severe pain, respectively.



## Discussion

The physical presentation and exploration of chemical space is usually a sequential and slow process. We have previously introduced ADE in the organic chemistry of small molecules with the advantages of its extraordinary degree of automation and scale to perform many reactions at the nanoscale. We used ADE in focused projects to automatically scout and elaborate the scope and limitations of new compounds including isoquinolines<sup>47</sup> or quinazolines.<sup>48</sup> Moreover, we used ADE for the selective synthesis of libraries of compound classes, such as boronic acids,<sup>49</sup> indoles<sup>50</sup> or covalent inhibitors at the nanoscale.<sup>51</sup> We also described the *in situ* synthesis and screening of nanoscale GBB-scaffold based libraries for the discovery of potent protein–protein interaction antagonists in a ‘direct-to-biology’ approach.<sup>52</sup> On an unprecedented breath, here we describe for the first time the automated acoustic dispending-enabled nanoscale synthesis of multiple different scaffolds in parallel. In contrast to classical library synthesis approaches, not one scaffold is synthesized at one time, but 16 different scaffolds based on 16 different chemistries, considerably accelerating synthetic chemistry and widening the chemical space. The acceleration and miniaturization of synthetic chemistry are becoming significant parameters to discover novel drugs and materials in an age of big data and artificial intelligence driven discovery and manufacturing in a timely and economical fashion.<sup>53,54</sup> The analysis of the 1536 reactions provides insight into the scope and limitations of compatible building blocks and subsequently can be instrumental for reaction optimization.<sup>55</sup> ‘Chemistry at the speed of sound’ can be used for the generation of diverse compound libraries for the identification of biologically active compounds or material properties. Indeed, our chemoinformatic analysis of the 16 scaffolds supports the high diversity and drug-likeness of our *in situ* synthesized chemical space. While we employed here just 16 different chemistries, it is conceivable that the scaffold diversity can be further increased (*e.g.*  $48 \times 32 = 1536$ ) to produce even more diverse libraries ‘on the fly’ in a high density format. Library synthesis as shown here is practical, as for each scaffold type several derivatives are presented. In the case of the identified screening hits, neighbors from the same scaffold class allow a first rapid SAR. Our ADE approach also allows for a fast evaluation of the scope and limitations of multiple reactions in parallel. Under the reaction conditions used here, 12 out of 16 scaffolds worked well in more than 50% of the cases, whereas 4 scaffolds resulted in less than 30% successful transformations. Detailed analysis of the generated ‘big data’ can help to understand the compatibility of electronic and steric features in the different building blocks. High throughput analysis of more than 1000 reactions enables high throughput experimentation to optimize the reaction conditions for library synthesis or to deselect building blocks which repeatedly do not result in successful product formation. Another important advantage of this library synthesis approach is its sustainability. One of the 12 principles of green chemistry is ‘to better prevent waste than to treat or clean up

waste after it is created’. Clearly, during the process of property optimization, unavoidably, hundreds or even thousands of derivatives have to be synthesized and tested. On a classical mmol scale, thousands of compounds are generated in the discovery phase, and the synthesis of 1536 compounds would use hundreds of kg of highly valuable building blocks, solvents and purification materials as well as consumables to produce 20 mg of each compound. Amazingly, our campaign of synthesizing 1536 novel compounds at the nanoscale used a total of ~7 mg of building blocks and ~20 mg of solvent. The high prices of building blocks are a true economic burden. For example, a 96-well plate of pre-weighed building blocks on a 0.2 mmol scale easily costs ~15 000 USD. However, high building block diversity is a prerequisite for the construction of diverse general screening libraries. Moreover, due to the fully automated synthesis set-up, the contact time of the bench chemist with the toxic solvent and chemical fumes is reduced to a minimum. ‘Chemistry at the speed of sound’ can therefore considerably help to reduce the environmental footprint of organic synthesis for drug discovery. Importantly, automated and fast microscale purification in support of high-throughput medicinal chemistry was recently described to avoid potential false positives through the screening of unpurified compounds.<sup>56</sup> Given the speed of ADE and the structural scaffold diversity of IMCR it is conceivable that many more chemistries can be performed in parallel. However, the true advantage of the herein described HT miniaturized high diversity chemistry will become apparent by the integration with HT screening and machine learning guidance to optimize compound properties to reduce the time and effort and increase the throughput and sustainability of the ubiquitous ‘design-make-test-analyze’ cycle.<sup>57</sup>

## Conflicts of interest

The authors declare no competing financial interest.

## Acknowledgements

Hylke Middel is acknowledged for his help with the automated MS analysis software. We thank Pravin Patil for the helpful discussions. Martin Stahorsky is acknowledged for the preparation of stock solutions. LG and RX were supported by the Chinese Scholarship Council Grant. SS was supported by a postdoctoral research fellowship from Kankerbestrijding (KWF Grant Agreement 10504). This project has received funding (to AD) from the NIH Grant (2R01GM097082-05), the European Lead Factory (IMI) under Grant Agreement 115489, the Qatar National Research Foundation (NPRP6-065-3-012), and the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie (ITN “Accelerated Early stage drug dIScovery”, Grant Agreement 675555; Cofunds ALERT (665250) and PROMINENT (754425)).



## References

- S. Dandapani, G. Rosse, N. Southall, J. M. Salvino and C. J. Thomas, Selecting, Acquiring, and Using Small Molecule Libraries for High-Throughput Screening, *Curr. Protoc. Chem. Biol.*, 2012, **4**, 177–191.
- M. Abou-Gharbia and W. E. Childers, Discovery of Innovative Therapeutics: Today's Realities and Tomorrow's Vision. 2. Pharma's Challenges and Their Commitment to Innovation, *J. Med. Chem.*, 2014, **57**, 5525–5553.
- M. Follmann, H. Briem, A. Steinmeyer, A. Hillisch, M. H. Schmitt, H. Haning and H. Meier, An approach towards enhancement of a screening library: The Next Generation Library Initiative (NGLI) at Bayer—against all odds?, *Drug Discovery Today*, 2019, **24**, 668–672.
- S. L. Belyanskaya, Y. Ding, J. F. Callahan, A. L. Lazaar and D. I. Israel, Discovering Drugs with DNA-Encoded Library Technology: From Concept to Clinic with an Inhibitor of Soluble Epoxide Hydrolase, *ChemBioChem*, 2017, **18**, 837–842.
- J. Osborne, S. Panova, M. Rapti, T. Urushima and H. Jhoti, Fragments: where are we now?, *Biochem. Soc. Trans.*, 2020, **48**, 271–280.
- A. C. Anderson, The process of structure-based drug design, *Chem. Biol.*, 2003, **10**, 787–797.
- S. L. Schreiber, Target-oriented and diversity-oriented organic synthesis in drug discovery, *Science*, 2000, **287**, 1964–1969.
- B. K. Shoichet, Virtual screening of chemical libraries, *Nature*, 2004, **432**, 862–865.
- R. Liu, X. Li and K. S. Lam, Combinatorial chemistry in drug discovery, *Curr. Opin. Chem. Biol.*, 2017, **38**, 117–126.
- H. S. Yu, K. Modugula, O. Ichihara, K. Kramschuster, S. Keng, R. Abel and L. Wang, General Theory of Fragment Linking in Molecular Design: Why Fragment Linking Rarely Succeeds and How to Improve Outcomes, *J. Chem. Theory Comput.*, 2021, **17**, 450–462.
- J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Alga, K. Tolmacheva, A. A. Tolmachev, B. K. Shoichet, B. L. Roth and J. J. Irwin, Ultra-large library docking for discovering new chemotypes, *Nature*, 2019, **566**, 224–229.
- Y. Ying, Y. Kun, P. R. Matthew, L. Karl, A. Robert, S. Brian and J. Steven, Efficient Exploration of Chemical Space with Docking and Deep-Learning, *Chem. Theory Comput.*, 2021, **17**, 7106–7119.
- N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane and F. von Delft, A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density, *Nat. Commun.*, 2017, **8**, 15123.
- C. Nichols, J. Ng, A. Keshu, G. Kelly, M. R. Conte, M. S. Marber, F. Fraternali and G. F. De Nicola, Mining the PDB for Tractable Cases Where X-ray Crystallography Combined with Fragment Screens Can Be Used to Systematically Design Protein–Protein Inhibitors: Two Test Cases Illustrated by IL1 $\beta$ -IL1R and p38 $\alpha$ -TAB1 Complexes, *J. Med. Chem.*, 2020, **63**, 7559–7568.
- F. Sutanto, S. Shaabani, R. Oerlemans, D. Eris, P. Patil, M. Hadian, M. Wang, M. E. Sharpe, M. R. Groves and A. Dömling, Combining High-Throughput Synthesis and High-Throughput Protein Crystallography for Accelerated Hit Identification, *Angew. Chem., Int. Ed.*, 2021, 18231–18239.
- C. G. Neochoritis, T. Zhao and A. Dömling, Tetrazoles via Multicomponent Reactions, *Chem. Rev.*, 2019, **119**, 1970–2042.
- A. Dömling, W. Wang and K. Wang, Chemistry and Biology Of Multicomponent Reactions, *Chem. Rev.*, 2012, **112**, 3083–3135.
- A. Dömling, Recent Developments in Isocyanide Based Multicomponent Reactions in Applied Chemistry, *Chem. Rev.*, 2006, **106**, 17–89.
- B. Hadimioglu, R. Stearns and R. Ellson, Moving Liquids with Sound: The Physics of Acoustic Droplet Ejection for Robust Laboratory Automation in Life Sciences, *J. Lab. Autom.*, 2016, **21**, 4–18.
- H. Cao, H. Liu and A. Dömling, Efficient multicomponent reaction synthesis of the schistosomiasis drug praziquantel, *Chem. – Eur. J.*, 2010, **16**, 12296–12298.
- A. Osipyan, S. Shaabani, R. Warmerdam, S. V. Shishkina, H. Boltz and A. Dömling, Automated, Accelerated Nanoscale Synthesis of Iminopyrrolidines, *Angew. Chem., Int. Ed.*, 2020, **59**, 12423–12427.
- M. R. McKeown, D. L. Shaw, H. Fu, S. Liu, X. Xu, J. J. Marineau, Y. Huang, X. Zhang, D. L. Buckley, A. Kadam, Z. Zhang, S. C. Blacklow, J. Qi, W. Zhang and J. E. Bradner, Biased multicomponent reactions to develop novel bromodomain inhibitors, *J. Med. Chem.*, 2014, **57**, 9019–9027.
- N. Desroy, C. Housseman, X. Bock, A. Joncour, N. Bienvenu, L. Chereil, V. Labeguere, E. Rondet, C. Peixoto, J. M. Grassot, O. Picolet, D. Annot, N. Triballeau, A. Monjardet, E. Wakselman, V. Roncoroni, S. Le Tallec, R. Blanque, C. Cottreaux, N. Vandervoort, T. Christophe, P. Mollat, M. Lamers, M. Auberval, B. Hrvacic, J. Ralic, L. Oste, E. van der Aar, R. Brys and B. Heckmann, Discovery of 2-[[2-Ethyl-6-[4-[2-(3-hydroxyazetididin-1-yl)-2-oxoethyl]piperazin-1-yl]-8-methylimidazo[1,2-a]pyridin-3-yl]methylamino]-4-(4-fluorophenyl)thiazole-5-carbonitrile (GLPG1690), a First-in-Class Autotaxin Inhibitor Undergoing Clinical Evaluation for the Treatment of Idiopathic Pulmonary Fibrosis, *J. Med. Chem.*, 2017, **60**, 3580–3590.
- A. Boltjes and A. Dömling, The Groebke-Blackburn-Bienaymé Reaction, *Eur. J. Org. Chem.*, 2019, 7007–7049.
- M. Aouali, D. Mhalla, F. Allouche, L. El Kaim, S. Tounsi, M. Trigui and F. Chabchoub, Synthesis, antimicrobial and antioxidant activities of imidazotriazoles and new multicomponent reaction toward 5-amino-1-phenyl[1,2,4]triazole derivatives, *Med. Chem. Res.*, 2015, **24**, 2732–2741.



- 26 K. Khoury, M. K. Sinha, T. Nagashima, E. Herdtweck and A. J. A. C. Dömling, Efficient assembly of iminodicarboxamides by a “truly” four-component reaction, *Angew. Chem.*, 2012, **124**, 10426–10429.
- 27 B. Malawska, K. Kulig, J. Gajda, D. Szczeblewski, A. Musiał, K. Wieckowski, D. Maciag and J. P. Stables, Design, synthesis and pharmacological evaluation of alpha-substituted N-benzylamides of gamma-hydroxybutyric acid with potential GABA-ergic activity. Part 6. Search for new anticonvulsant compounds, *Acta Pol. Pharm.*, 2007, **64**, 127–137.
- 28 B. Beck, S. Srivastava, K. Khoury, E. Herdtweck and A. Dömling, One-pot multicomponent synthesis of two novel thiolactone scaffolds, *Mol. Divers.*, 2010, **14**, 479–491.
- 29 Y. B. Kim, S. J. Park, G. C. Geum, M. S. Jang, S. B. Kang, D. H. Lee and Y. S. Kim, An Efficient Synthesis of  $\alpha$ -Amino- $\delta$ -valerolactones by the Ugi Five-Center Three-Component Reaction, *Bull. Korean Chem. Soc.*, 2002, **23**, 1277–1320.
- 30 A. Demharter, W. Hörl, E. Herdtweck and I. Ugi, Synthesis of Chiral 1,1'-Iminodicarboxylic Acid Derivatives from  $\alpha$ -Amino Acids, Aldehydes, Isocyanides, and Alcohols by the Diastereoselective Five-Center-Four-Component Reaction, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 173–175.
- 31 I. Ugi, A. Demharter, W. Hörl and T. Schmid, Ugi reactions with trifunctional  $\alpha$ -amino acids, aldehydes, isocyanides and alcohols, *Tetrahedron*, 1996, **52**, 11657–11664.
- 32 B. Beck, S. Srivastava and A. Dömling, New end-on thiolactone scaffold by an isocyanide-based multicomponent reaction, *Heterocycles*, 2008, **73**, 177.
- 33 S. Cho, G. Keum, S. B. Kang, S. Y. Han and Y. Kim, An efficient synthesis of 2,5-diketopiperazine derivatives by the Ugi four-center three-component reaction, *Mol. Diversity*, 2003, **6**, 283–286.
- 34 R. S. Bon, B. van Vliet, N. E. Sprenkels, R. F. Schmitz, F. J. de Kanter, C. V. Stevens, M. Swart, F. M. Bickelhaupt, M. B. Groen and R. V. Orru, Multicomponent synthesis of 2-imidazolines, *Mol. Divers.*, 2005, **70**, 3542–3553.
- 35 S. Gunawan and C. Hulme, Bifunctional building blocks in the Ugi-azide condensation reaction: a general strategy toward exploration of new molecular diversity, *Org. Biomol. Chem.*, 2013, **11**, 6036–6046.
- 36 T. Nixey, M. Kelly and C. Hulme, The one-pot solution phase preparation of fused tetrazole-ketopiperazines, *Tetrahedron Lett.*, 2000, **41**, 8729–8733.
- 37 M. L. Vazquez, N. Kaila, J. W. Strohbach, J. D. Trzupke, M. F. Brown, M. E. Flanagan, M. J. Mitton-Fry, T. A. Johnson, R. E. TenBrink, E. P. Arnold, A. Basak, S. E. Heasley, S. Kwon, J. Langille, M. D. Parikh, S. H. Griffin, J. M. Casavant, B. A. Duclos, A. E. Fenwick, T. M. Harris, S. Han, N. Caspers, M. E. Dowty, X. Yang, M. E. Banker, M. Hegen, P. T. Symanowicz, L. Li, L. Wang, T. H. Lin, J. Jussif, J. D. Clark, J.-B. Telliez, R. P. Robinson and R. Unwalla, Identification of N-{cis-3-[Methyl(7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino]cyclobutyl}propane-1-sulfonamide (PF-04965842): A Selective JAK1 Clinical Candidate for the Treatment of Autoimmune Diseases, *J. Med. Chem.*, 2018, **61**, 1130–1152.
- 38 T. Sander, J. Freyss, M. von Korff and C. Rufener, DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis, *J. Chem. Inf. Model.*, 2015, **55**, 460–473.
- 39 D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang and J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, *Nucleic Acids Res.*, 2006, **34**, 668–672.
- 40 A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases, *J. Comb. Chem.*, 1999, **1**, 55–68.
- 41 I. Muegge, S. L. Heald and D. Brittelli, Simple Selection Criteria for Drug-like Chemical Matter, *J. Med. Chem.*, 2001, **44**, 1841–1846.
- 42 T. T. Wager, X. Hou, P. R. Verhoest and A. Villalobos, Central Nervous System Multiparameter Optimization Desirability: Application in Drug Discovery, *ACS Chem. Neurosci.*, 2016, **7**, 767–775.
- 43 Z. Rankovic, CNS Drug Design: Balancing Physicochemical Properties for Optimal Brain Exposure, *J. Med. Chem.*, 2015, **58**, 2584–2608.
- 44 W. H. B. Sauer and M. K. Schwarz, Molecular Shape Diversity of Combinatorial Libraries: A Prerequisite for Broad Bioactivity, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987–1003.
- 45 D. Schaller, D. Šribar, T. Noonan, L. Deng, T. N. Nguyen, S. Pach, D. Machalz, M. Bermudez and G. Wolber, Next generation 3D pharmacophore modeling, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2020, **10**, e1468.
- 46 C. Boss, J. Hazemann, T. Kimmerlin, M. von Korff, U. Lüthi, O. Peter, T. Sander and R. Siegrist, The screening compound collection: a key asset for drug discovery, *Chimia*, 2017, **71**, 667–677.
- 47 Y. Wang, S. Shaabani, M. Ahmadianmoghaddam, L. Gao, R. Xu, K. Kurpiewska, J. Kalinowska-Tluscik, J. Olechno, R. Ellson and M. Kossenjans, Acoustic droplet ejection enabled automated reaction scouting, *ACS Cent. Sci.*, 2019, **5**, 451–457.
- 48 M. Hadian, S. Shaabani, P. Patil, S. V. Shishkina, H. Böltz and A. J. G. C. Dömling, Sustainability by design: automated nanoscale 2,3,4-trisubstituted quinazoline diversity, *Green Chem.*, 2020, **22**, 2459–2467.
- 49 C. G. Neochoritis, S. Shaabani, M. Ahmadianmoghaddam, T. Zarganes-Tzitzikas, L. Gao, M. Novotná, T. Mitříková, A. R. Romero, M. I. Irianti and R. Xu, Rapid approach to complex boronic acids, *Sci. Adv.*, 2019, **5**, eaaw4607.
- 50 S. Shaabani, R. Xu, M. Ahmadianmoghaddam, L. Gao, M. Stahorsky, J. Olechno, R. Ellson, M. Kossenjans, V. Helan and A. J. G. C. Dömling, Automated and accelerated synthesis of indole derivatives on a nano-scale, *Green Chem.*, 2019, **21**, 225–232.
- 51 F. Sutanto, S. Shaabani, C. G. Neochoritis, T. Zarganes-Tzitzikas, P. Patil, E. Ghonchepour and A. Dömling,



- Multicomponent reaction-derived covalent inhibitor space, *Sci. Adv.*, 2021, 7, eabd9307.
- 52 K. Gao, S. Shaabani, R. Xu, T. Zarganes-Tzitzikas, L. Gao, M. Ahmadianmoghadam, M. R. Groves and A. Dömling, Nanoscale, automated, high throughput synthesis and screening for the accelerated discovery of protein modifiers, *RSC Med. Chem.*, 2021, 12, 809–818.
- 53 S. Lin, S. Dikler, W. D. Blincoe, R. D. Ferguson, R. P. Sheridan, Z. Peng, D. V. Conway, K. Zawatzky, H. Wang and T. J. S. Cernak, Mapping the dark space of chemical reactions with extended nanomole synthesis and MALDI-TOF MS, *Science*, 2018, 361, eaar6236.
- 54 N. J. Gesmundo, B. Sauvagnat, P. J. Curran, M. P. Richards, C. L. Andrews, P. J. Dandliker and T. J. N. Cernak, Nanoscale synthesis and affinity ranking, *Nature*, 2018, 557, 228–232.
- 55 M. Shevlin, Practical high-throughput experimentation for chemists, *ACS Med. Chem. Lett.*, 2017, 8, 601–607.
- 56 C. Barhate, A. F. Donnell, M. Davies, L. Li, Y. Zhang, F. Yang, R. Black, G. Zipp, Y. Zhang and C. Cavallaro, Microscale purification in support of high-throughput medicinal chemistry, *Chem. Commun.*, 2021, 57, 11037–11040.
- 57 A. T. Plowright, C. Johnstone, J. Kihlberg, J. Pettersson, G. Robb and R. A. Thompson, Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle, *Drug Discovery Today*, 2012, 17, 56–62.

