The European Lead Factory: An updated HTS compound library for innovative drug discovery

Herman van Vlijmen [a,*], Jean-Yves Ortholand [b], Volkhart M.-J. Li [c], Jon S.B. de Vlieger [d]

[a] Janssen Research and Development, Turnhoutseweg 30, 2340 Beerse, Belgium
[b] Edelris, 60 Avenue Rockefeller, 69008 Lyon, France
[c] Bayer AG, 42096 Wuppertal, Germany
[d] Lygature, Jaarbeursplein 6, 3521 AL Utrecht, the Netherlands

Through the European Lead Factory model, industry-standard high-throughput screening and hit validation are made available to academia, small and medium-sized enterprises, charity organizations, patient foundations, and participating pharmaceutical companies. The compound collection used for screening is built from a unique diversity of sources. It brings together compounds from companies with different therapeutic area heritages and completely new compounds from library synthesis. This generates structural diversity and combines molecules with complementary physicochemical properties. In 2019, the screening library was updated to enable another 5 years of running innovative drug discovery projects. Here, we investigate the physicochemical and diversity properties of the updated compound collection. We show that it is highly diverse, drug-like, and complementary to commercial screening libraries.

Keywords: Compound library; High throughput screening; Public–private partnership; Library synthesis

Introduction

The concept of large-scale sharing of screening compound collections was pioneered in the European Lead Factory consortium (ELF).[1,2] The ELF, a public–private partnership, started in 2013 and funded within the framework of the Innovative Medicines Initiative (IMI), created a unique core screening library of over 300 000 compounds originating from participating pharma companies through sharing proprietary compounds securely from their screening decks.[3] This collection was continuously expanded over the course of the ELF by crowdsourcing ideas for novel scaffolds from across the European synthetic chemistry community.[4,5] These ideas were assessed for novelty, drug likeness, and tractability in a peer-reviewed process. The best ideas were then selected for synthesis and elaboration by five synthetic chemistry small and medium-sized enterprises (SMEs) and added to the library. By the end of the initial 5 years of the ELF project, this resulted in a library of over 500 000 compounds stored at a central liquid store to facilitate logistics.

A competitive, confidential peer-reviewed application process open to academic groups and drug discovery biotechs enabled access to this screening library. The selected projects (i.e., protein targets) were taken forward for assay optimization, screening, and hit validation at centers in The Netherlands and UK. HTS triage was executed in close cooperation between ELF scientists and the project owner,[6,7] and validated 'hit lists' of molecules were returned to the owner together with a sample of each hit for further work.

The first 5 years of the ELF demonstrated the effectiveness of the model with multiple publications,[8,9] spin-out companies created in Sweden and the UK, agreements to develop molecules with third-party organizations, and further grant funding to develop molecules for neglected diseases.[10]

* *Corresponding author.* van Vlijmen, H. (hvvlijme@its.jnj.com)

1359-6446/
https://doi.org/10.1016/j.drudis.2021.04.019

Five factors were crucial to this success: (i) an effective shared state-of-the-art screening library; (ii) secure compound sharing protocols and IT systems gave confidence to pharma companies to share proprietary compounds; (iii) a bespoke high-throughput screening (HTS) triage application to maintain target confidentiality and enable effective triage within the confidentiality model; (iv) Intellectual Property Rights are retained by the organization proposing the project; and (v) no up-front costs were incurred by the project proposers, with only milestone payments should the project move towards commercialization

Following the successful experiment of compound sharing within the first phase of the ELF, in which over 150 targets were screened, the effort was continued with the IMI-2 ESCulab project. ESCulab started in 2019 and continues to operate under the ELF banner to provide screening and hit validation services while adhering to these five crucial factors to ensure continuity of a validated and robust model, with the addition of target agnostic and high-content phenotypic screening. Here, we investigate the properties of the newly available ESCulab Compound Collection analogously to the previous analysis by Besnard *et al.*[3] In particular, we compare the compounds contributed by the pharma companies with those synthesized over the course of the ELF and sets of approved drugs and commercial screening compounds.

Esculab compound collection analysis

The ESCulab compound collection comprises 13 subsets. Six were contributed by pharmaceutical companies previously in the ELF (Janssen Pharmaceutica, AstraZeneca PLC, Sanofi S.A., Bayer AG, Merck KGaA, and UCB S.A.), two subsets were contributed by companies who joined at the start of ESCulab (Grünenthal GmbH and Servier), and the remaining five by the synthetic chemistry SME members (Edelris s.a.s., Mercachem, Taros, Sygnature, and Syncom) of the ELF consortium, who synthesized the library proposals selected by the peer-review panel over the course of the ELF. The subset sizes range from 25 000 to 50 000 for the pharma companies and are around 40 000 for each of the synthetic chemistry SMEs. The properties of each subset are shown in this analysis and anonymized by giving each subset a reference number. The analysis includes a set of 9609 drugs (the drugstore collection from ChEMBL26)[11] and the Maybridge commercial HTS for Drug Discovery screening collection of ~55 000 compounds (ThermoFisher Scientific).

The subsets were first compared using key physicochemical properties [molecular weight (MW), atomic logP (AlogP), Fsp3, topological polar surface area (TPSA), numbers of hydrogen bond donors and acceptors, and rotatable bonds] and the quantitative estimate of drug likeness (QED) score,[12] calculated with Pipeline Pilot.[13] The first seven parameters are key calculable properties of a molecule. Fsp3 is the fraction of sp^3-hybridized carbon atoms and is often used as a proxy for the 'three-dimensionality' of a molecule. Three-dimensionality has been implicated as a predictor of clinical success.[14–16] TPSA is the sum of the areas of all polar fragments in a molecule. The QED score is a molecular desirability score and includes key physicochemical properties, such as MW and structural alerts, and sums their weighted contributions to generate a score between 0 and 1, with 1 representing an ideally drug-like molecule.

Subsequently, the diversity of the library was investigated. We evaluated the similarity between the ESCulab collection and the commercial Maybridge library. Subsets from the contributing ESCulab partners were also compared with each other to establish their pairwise similarity and understand whether subsets are occupying similar or diverse chemical space. The diversity was assessed using a nearest neighbor approach described more fully below.

Lastly, we focused on the scaffolds of the library compounds. For every compound, the Bemis–Murcko scaffold was generated, in which all aliphatic sidechains are trimmed while preserving scaffold atom types and double bonds attached to rings.[17] Scaffold diversity was assessed using Shannon entropy and scaffold recovery curves.[18] Relative scaffold complexity in the different compound sets was evaluated using the metric proposed by Xu.[19]

Physicochemical properties

The distributions of most calculated parameters are shown below in box-and-whisker plots (Fig. 1). In addition, Table S1 in the supplemental information online lists the means and standard deviations of all properties. Subsets 100–107 represent the eight pharma subsets, 110–114 those from the five synthetic chemistry SMEs, and Drugs and MB the drugs and Maybridge collections, respectively. All the ESCulab subset distributions fall largely within the key lipophilicity and MWs described in the 'Lipinski Guidelines'.[20] It is apparent that the mean MWs of the SME subsets are higher than those of the pharma industry and Drugs/MB subsets, whereas the standard deviations are slightly lower. The synthetic strategies used to generate the SME subsets are likely to underlie this difference. Each library proposal included a synthetic scheme for a novel core scaffold having multiple points of diversity for decoration with diverse building blocks. This generated a library of a few hundred molecules. With two or three diverse substituents on a core scaffold, this immediately generates a significant MW. Conversely, a compound from a pharma screening deck could be a project compound from a library exploring small substitutions at a single point or a bespoke singleton and will not have this lower bound on MW. Considerable effort was put into the SME library design to restrain this effect and prevent excessive MW inflation, resulting in lower standard deviations.

From the AlogP predicted lipophilicity data, it is apparent that the efforts made by the ELF SMEs in controlling the upper limit of the lipophilicity were effective, with little difference between SME and pharma partner subsets apparent. Indeed, the subset with the lowest mean AlogP is an SME library. This is an important outcome given the higher MW of the SME molecules and the typically observed correlation between size and lipophilicity.

There is a slight trend of a higher TPSA of the SME compounds. This is likely to affect their ability to cross the blood–brain barrier, where a TPSA of less than ~90 $Å^2$ is associated with higher probability of crossing the barrier.[21] The attractiveness of the subset depends on whether a molecule targeting the central nervous system (CNS) or specifically not targeting the CNS is desired. When the ESCulab collection is viewed as a whole, molecules spanning a wide range of TPSA, MW, and AlogP are represented and the library is likely to generate hits for both CNS and peripheral targets.

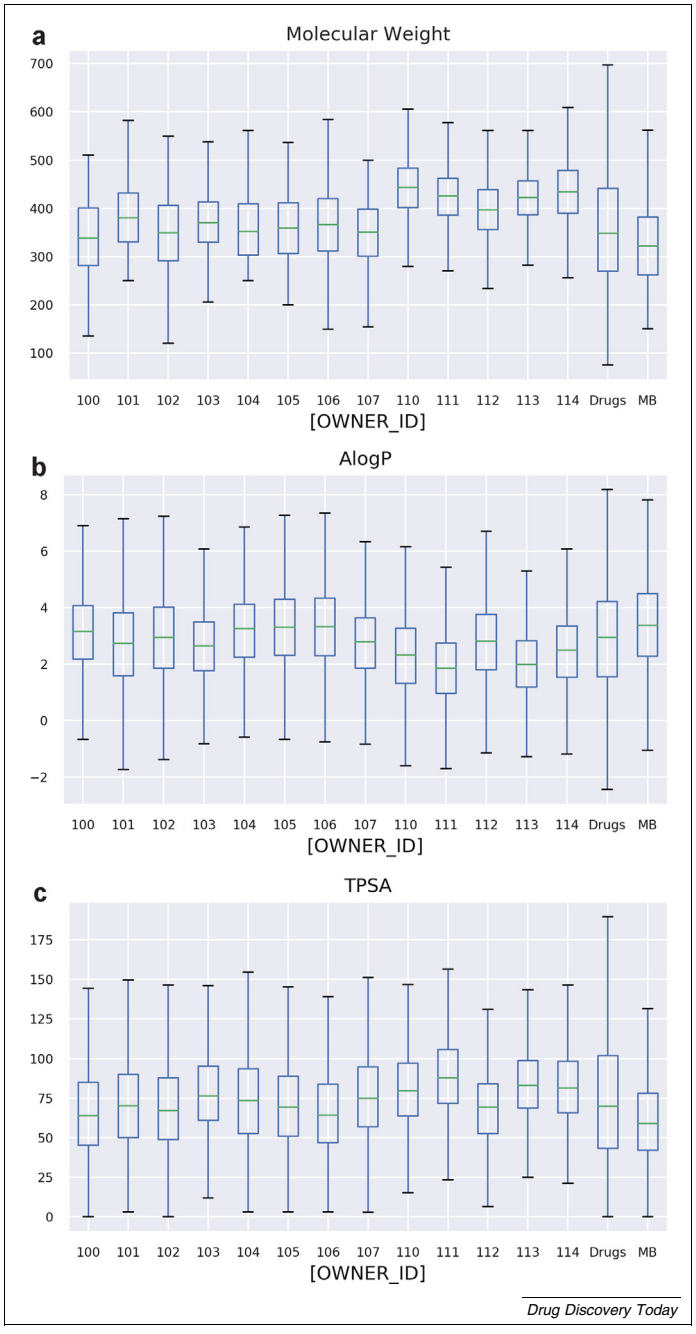

FIG. 1
Box-and-whisker plots showing key physicochemical properties of the ESCulab subsets and the Drugs and Maybridge (MB) sets. Subset owners are anonymized with pharma partners 100–107 and SME partners 110–114. The central horizontal line of each box shows the median value of the property distribution, and the top and bottom horizontal lines of each box indicate the first and third quartile values, respectively. Whiskers indicate the range of the data. (a) Molecular weight; (b) AlogP; (c) TPSA; (d) Fsp^3; and (e) quantitative estimate of drug likeness (QED).

The Fsp3 of the SME subsets is generally higher than that of the pharma subsets. Two library design objectives incorporated by the SMEs are likely to underlie this: (i) controlling lipophilicity; and (ii) generating '3D' molecules. The desirability of '3D' molecules was indicated in the 'Escaping Flatland' article,[15] although the importance of this has been questioned.[22] Sp2

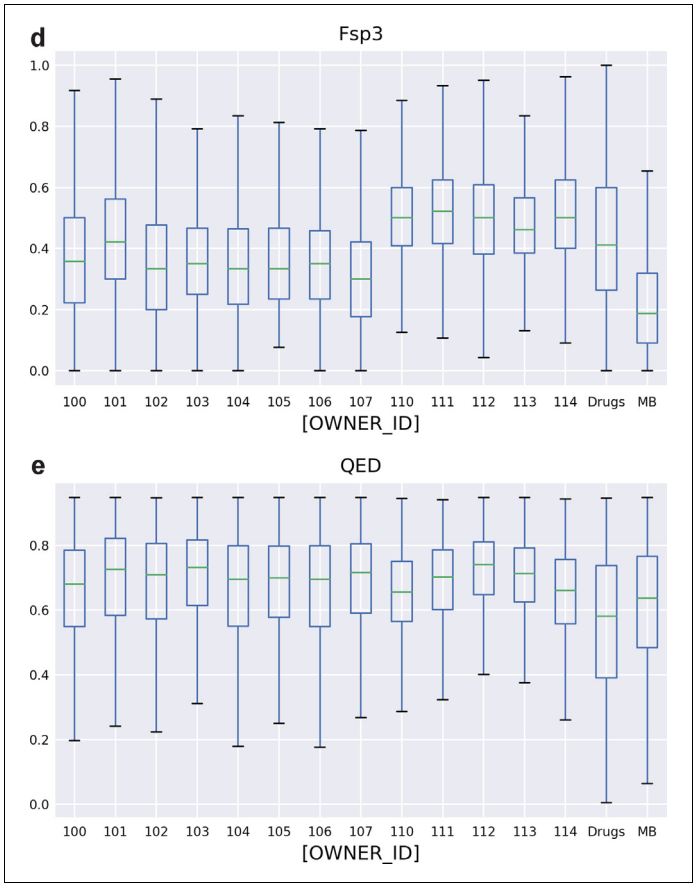

Fig. 1 (*continued*)

hybridization is typically associated with aromatic rings, which are flat and often lipophilic, so the interplay of all these library design aims will bias the SME compounds towards sp^3-rich structures. Regardless of the importance of sp^3 richness in developability of a molecule, including it as a library design feature will likely have the effect of increasing the diversity with respect to the pharmaceutical industry subsets.

Nine examples of scaffolds from the SME collection are shown in Fig. 2[23–31] and other examples can be found elsewhere in the literature.[4,5] These cores usually have at least two points of diversity and are elaborated with diverse substituents. High Fsp^3 is a feature shared by the scaffolds resulting in out-of-plane bond vectors emerging from the scaffolds. The scaffolds span a range of MWs and typically have at least one, usually defined, chiral center. Over 250 different libraries have been synthesized resulting in a diverse collection to complement the more typical HTS library compounds contributed by the pharma companies. These SME libraries include peptides, macrocycles, and natural product analogs.

The QED score provides a simple summary value of the attractiveness of a compound for a drug discovery program. The ESCulab compounds are generally high scoring, with most compounds having a QED score >0.55. The scores for some compounds are below this and inspection of some of these indicated that they were largely the high MW compounds, and this factor has a large influence on the QED score.

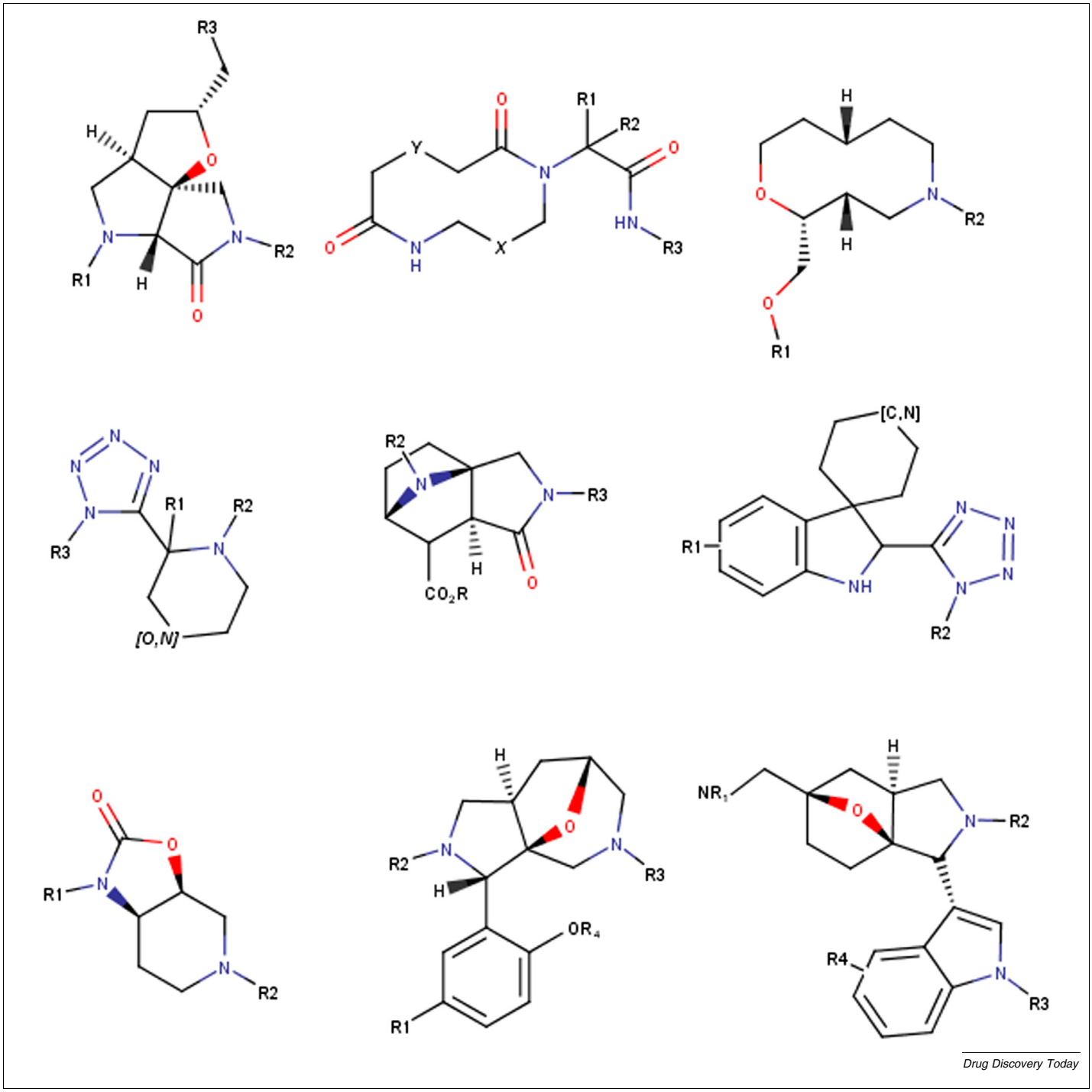

FIG. 2
Example scaffolds from the small and medium-sized enterprise (SME) library synthesis campaigns. The scaffolds are sp^3 rich, resulting in out-of-plane bond vectors from the scaffold and a highly '3-dimensional' structure.

MW and AlogP in screening collections are generally correlated, but it is of interest to understand whether there is any difference in behavior between the SME and pharma subsets, particularly given the design aims of the SMEs to stay within the typical drug-like properties while working with somewhat higher MW molecules. Fig. 3 shows joint density plots of AlogP versus MW for each compound owner and for the Drugs and MB sets. It is apparent that the SME subsets have a more symmetrical distribution and that correlation between the two parameters is low, whereas for example the correlation is more apparent in the MB collection and in pharma subsets 100, 102, and 106. This suggests that the multiobjective library design process adopted by the SMEs was effective in controlling lipophilicity.

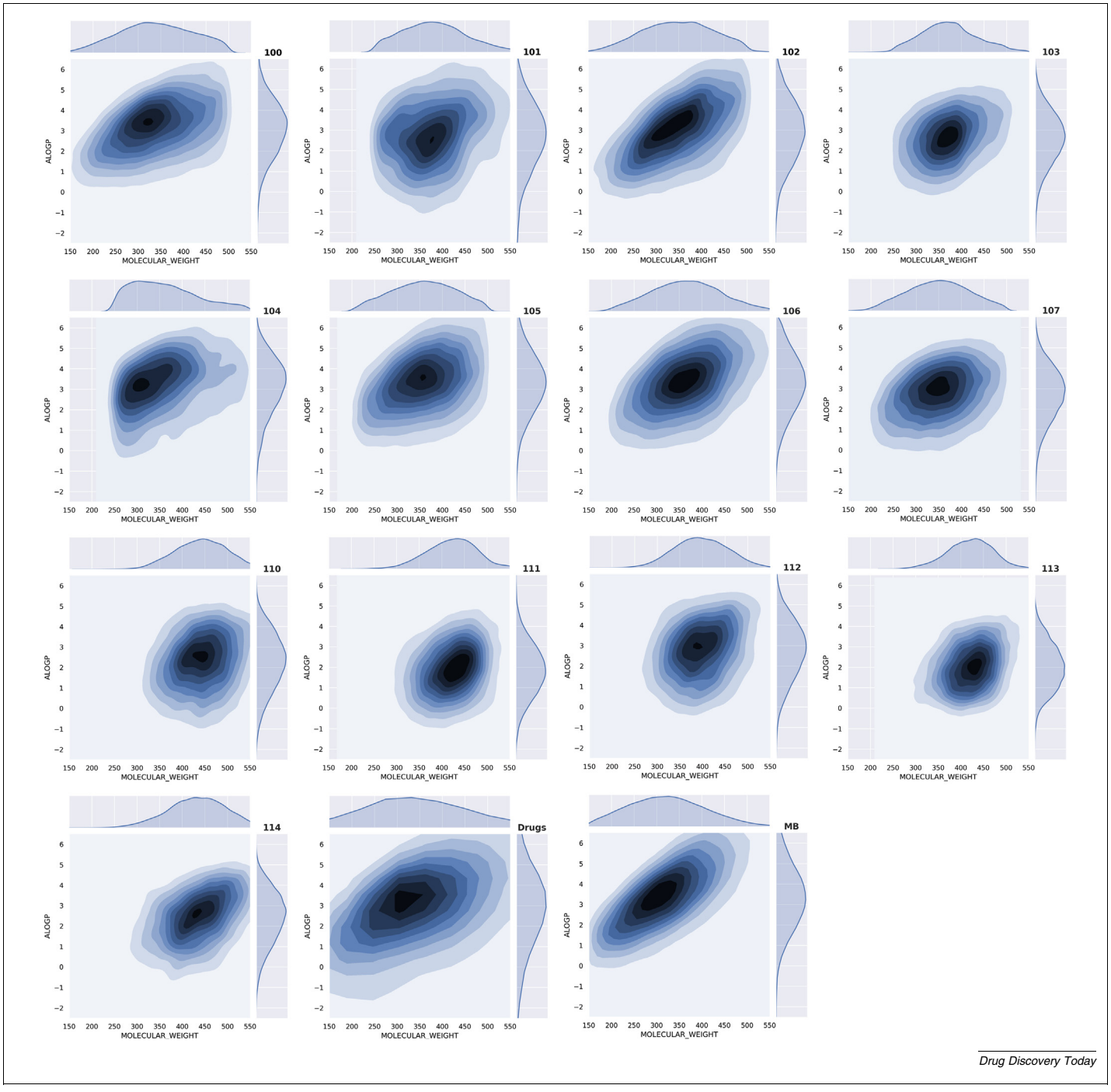

FIG. 3
Joint distributions of molecular weight (MW; *x* axis) and AlogP (*y* axis). The subset identifiers are shown at the top right of every plot.

Diversity and scaffold analysis

The diversity of the 13 subsets in ESCulab was assessed with a nearest neighbor approach. The Tanimoto distance from each molecule in subset A was compared with all the molecules in subset B using ECFP6 fingerprints, and the highest similarity for each subset A molecule was recorded. If a molecule in subset A had a similarity of at least 0.6 to a compound in subset B, the number of neighbors for subset A was incremented by one, and this was repeated for all subset A molecules. This calculation was then repeated for all pairwise combinations of subsets, including the 'self similarity' of molecules within subsets. This algorithm establishes whether there are near neighbors within and between subsets and whether the subsets explore diverse chemical space. Near neighbors are useful in interpreting structure–activity relationships in the screen and building confidence that a hit is genuine. Exploration of diverse chemical space is useful to maximize the likelihood of finding novel hits or compounds active towards novel targets.

The subset similarities are shown in the heatmap in Fig. 4. Each subset Owner ID is shown on the *x* and *y* axes and the num-

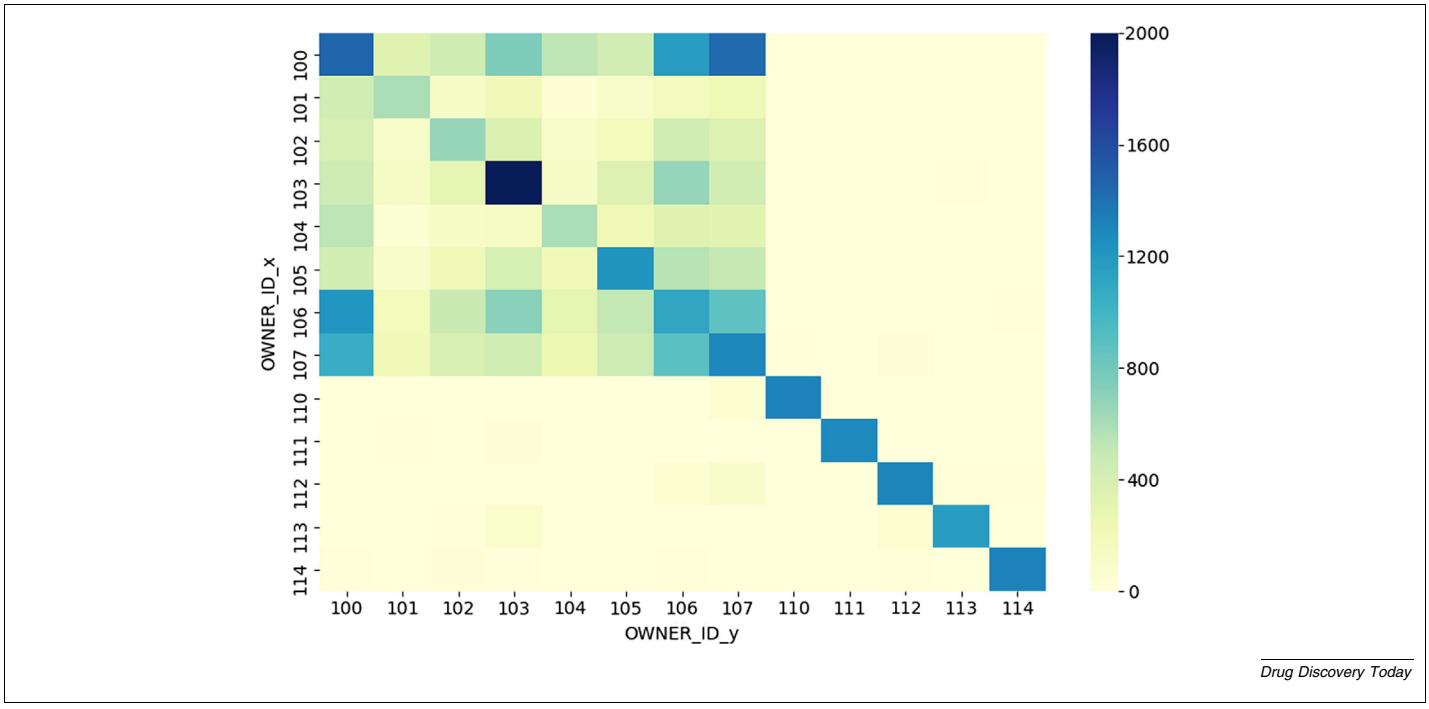

FIG. 4

Heatmap showing a similarity analysis of the 13 ESCulab subsets. The color of each cell represents the number of compounds sharing a Tanimoto similarity of ≥0.6 between the two subsets defined on the *x* and *y* axes. The off-diagonal cells represent the number of similar compounds between different subsets (i.e., the intersubset similarity). The leading diagonal shows the intrasubset similarity (i.e., the number of compounds within each subset sharing a Tanimoto similarity of ≥0.6). Intrasubset similarities on the leading diagonal are scaled by a factor of 1/15.

ber of compounds sharing a Tanimoto similarity of at least 0.6 between each subset is shown according to the color gradient. To facilitate interpretation of the heatmap, the values on the leading diagonal (i.e., number of similar compounds within a subset) have been multiplied by a scale factor of 1/15. This maps them into a similar range as the off-diagonal cells and allows interpretation of the color gradient.

It is immediately apparent that the subsets have more near neighbors within the subsets than between the subsets, with around half the molecules within each subset having at least one near neighbor with a similarity of 0.6 or more. This is not surprising because many of the pharma company partners selected compounds specifically to enable SAR analysis in the screen. The SME subsets comprise combinatorial libraries comprising a few hundred compounds. The common scaffold will immediately introduce some structural similarity and one or two common building blocks in the enumeration will likely achieve the 0.6 threshold. This can be useful in an HTS triage and in early hit expansion because the near neighbors will give a large amount of 'off-the-shelf' SAR straight out of the screen.

The subsets showing the highest intersubset similarity had around 1500 compounds in common (subsets 100 and 106, and subsets 100 and 107). This represents a rather small proportion of the overall library showing near neighbor similarity. It is notable that the SME subsets show little or no similarity to any other subset. This emphasizes the novelty of these molecules with respect to subsets drawn from eight pharma companies and that they explore novel chemical space not generally explored in standard HTS subsets, while still having attractive drug-like properties. This combination of near neighbors within subsets and high diversity between subsets ensures exploration of diverse chemical space and interpretable SAR in the screening output.

Analogously, the similarity between the ESCulab collection and the commercial MB library was calculated by counting the number of compounds in MB with a Tanimoto similarity higher than 0.6 using the same ECFP6 fingerprints. Out of the ~500 000 compounds, there were 1611 with more than two neighbors in MB, 1545 with two neighbors, and 5369 with one neighbor. This means that <2% of the ESCulab library had any similar compound in the MB collection.

We further analyzed the library by focusing on the scaffolds of the compounds. Two aspects of the Bemis–Murcko scaffolds in the collection were evaluated: their diversity and complexity. On average, there were approximately 2.5 compounds per scaffold, which indicates a high level of scaffold diversity in the collection. The distribution of compounds over scaffolds is summarized in scaffold recovery curves, in which the cumulative percentage of compounds versus scaffolds is shown. Curves that are closer to the diagonal have more scaffold diversity. Fig. S1a in the supplemental information online shows these curves for all subsets in ESCulab and the MB collection. The F_{50} metric is defined as the percentage of scaffolds at which 50% of the compounds is covered. There are significant differences between the subsets and, interestingly, the three subsets with the lowest F_{50} are from a pharma company (100), an SME set (112), and the commercial MB collection. Fig. S1b in the supplemental information online focuses on the first 1% of the scaffolds, and it is

apparent that all collections have a small percentage of scaffolds that contain a relatively large number of compounds. For instance, the first 0.2% of highest populated scaffolds contain between 6 and 20% of the compounds.

Another metric for scaffold diversity is the scaled Shannon entropy SSE, which ranges from 0 (all compounds have the same scaffold) to 1.0 (all compounds have a different scaffold). SSE values for the full compound subsets are relatively similar, ranging from 0.85 to 0.95 (Fig. S2a in the supplemental information online). When only considering the 20 and 100 most populated scaffolds, we see some clear differences: the MB collection has a few scaffolds with very many compounds, and the collections from the SMEs (110–114) tend to have their compounds more evenly distributed over scaffolds than the pharma subsets (Fig. S2b, c in the supplemental information online).

Another property of the scaffolds that can be quantified is their complexity. One metric of scaffold complexity was defined by Xu[19] and includes for each scaffold the smallest set of smallest rings (sssr), the number of heavy atoms, the number of bonds, and the sum of the heavy atomic numbers. These four parameters are compared to the largest value of each in the full compound data set, resulting in a value between 0 and 1, with higher values for higher complexity. The distribution of scaffold complexities of the subsets is shown in Fig. S3 in the supplemental information online. Two features stand out: the MB collection generally has lower scaffold complexity and the SME subsets consistently have higher scaffold complexity, confirming the design criteria for these newly synthesized libraries.

Discussion

The ESCulab compound collection draws compounds from eight pharma companies with diverse therapeutic area heritage and combines these with up to 200 000 entirely novel compounds, synthesized over the course of the ELF project and specifically designed to explore novel, drug-like chemical space distinct from the pharma subsets. The construction of the library highlights some key considerations when bringing multiple subsets into a single screening collection. The pharma subsets were selected by each company internally; thus, care had to be taken to remove accidental duplication in the combined library. Circular fingerprints for the compounds were generated within the firewalls of the companies using a common Pipeline Pilot protocol distributed to the companies and sent to an independent 'honest broker' group. This enabled deduplication (at fingerprint level) of compounds from different companies by the independent group without the compound structures being made public and this facilitated the willingness to share information. The pharma companies were able to design their subsets to include at least moderately near neighbors to build confidence in hits identified in the screen and generate SAR for machine-learning algorithms. This was necessary because, in general, random overlap between different company collections appears to be low.[32]

Expansion of the set with the additional 200 000 compounds also required careful planning. Library proposals were compared for 2D chemical similarity with commercially available compounds, with the core pharma set, and with libraries previously made by the consortium to ensure novelty. This was facilitated by bespoke web services to the ELF compound database. This was supplemented by comparison with public databases and commercially available molecules to eliminate the possibility of wasteful duplication of synthetic effort. Expert manual review to ensure synthetic tractability was also important. Another key step was a multiobjective library design to restrain physicochemical properties within acceptable ranges for orally bioavailable compounds. This was particularly important given the number of substitution points on the scaffolds. The effectiveness of this step is clear from the boxplots shown in Fig. 1 and the MW/AlogP joint plots in Fig. 3. As a result of these measures, compounds from all the subsets fall largely within guidelines for drug likeness but show distinct variations in calculated properties.

A strength of the compound-sharing model is the diversity of sources from which the library is generated. It brings together compounds from companies with very different therapeutic area heritages while combining 'traditional' pharma molecules with those generated by library synthesis. Not only does this generate structural diversity, but also combines molecules with complementary physicochemical properties. Consequently, the potential for finding highly novel, sp^3-rich molecules in the SME subset is complemented by pharmaceutical subsets that explore lower MWs and will be enriched in CNS-penetrant compounds. The ESCulab library is now available for SMEs and academic groups in Europe to be screened in target-focused, target-agnostic, and high-content imaging assays via the European Lead Factory website (www.europeanleadfactory.eu).

Concluding remarks

The ESCulab compound library of the ELF is an attractive HTS screening resource, containing highly diverse and drug-like compound sets from the screening collections of participating pharma companies and from crowdsourced newly designed libraries synthesized by SMEs.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The ELF has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement N° 115489, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. This project has also received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement N° 806948: 'ESCulab: European Screening Centre; Unique Library for Attractive Biology'. The JU receives support from the European Union's Horizon 2020 research and innovation program and EFPIA and Medicines for Malaria Venture. We thank Andrew D. Pannifer for help in the analysis of the results and suggestions for the manuscript, and Guillaume Paillard and Philip Cochrane for help in extracting the data from the Honest Data Broker system. The contributions from many scientists across several organizations throughout Europe, working

together within the European Lead Factory Consortium, have been instrumental to the success of the endeavor and are greatly appreciated (Universities of Leeds, Leiden, Nottingham, Groningen, Nijmegen, Amsterdam, Duisburg, Copenhagen and the Max Planck Institute and Lead Discovery Center, both in Dortmund).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.drudis.2021.04.019.

References

1 K. Kingwell, European Lead Factory hits its stride, Nat Rev Drug Discov 15 (2016) 221–222.
2 H. Laverty, K.M. Orrling, F. Giordanetto, M. Poinot, E. Ottow, T.W. Rijnders, et al., The European lead factory - an experiment in collaborative drug discovery, J Med Dev Sci 1 (2016) 20.
3 J. Besnard, P.S. Jones, A.L. Hopkins, A.D. Pannifer, The Joint European Compound Library: boosting precompetitive research, Drug Discov Today 20 (2015) 181–186.
4 A. Karawajczyk, F. Giordanetto, J. Benningshof, D. Hamza, T. Kalliokoski, K. Pouwer, et al., Expansion of chemical space for collaborative lead generation and drug discovery: the European Lead Factory Perspective, Drug Discov Today 20 (2015) 1310–1316.
5 R. Morgentin, M. Dow, A. Aimon, G. Karageorgis, T. Kalliokoski, D. Roche, et al., Translation of innovative chemistry into screening libraries: an exemplar partnership from the European Lead Factory, Drug Discov Today 23 (2018) 1578–1583.
6 P. Jones, S. McElroy, A. Morrison, A. Pannifer, The importance of triaging in determining the quality of output from high-throughput screening, Future Med Chem 7 (2015) 1847–1852.
7 G. Paillard, P. Cochrane, P.S. Jones, W.P. van Hoorn, A. Caracoti, H. van Vlijmen, et al., The ELF Honest Data Broker: informatics enabling public–private collaboration in a precompetitive arena, Drug Discov Today 21 (2016) 97–102.
8 E.D. Mock, M. Mustafa, O. Gunduz-Cinar, R. Cinar, G.N. Petrie, V. Kantae, et al., Discovery of a NAPE-PLD inhibitor that modulates emotional behavior in mice, Nat Chem Biol 16 (2020) 667–675.
9 E.M. Beck, E. Parnell, A. Cowley, A. Porter, J. Gillespie, J. Robinson, et al., Identification of a novel class of benzofuran oxoacetic acid-derived ligands that selectively activate cellular EPAC1, Cells 8 (2019) 1425.
10 European Lead Factory: Results. www.europeanleadfactory.eu/results. Published June 28, 2018 [Accessed April 2, 2021].
11 A. Gaulton, A. Hersey, M. Nowotka, A.P. Bento, J. Chambers, D. Mendez, et al., The ChEMBL database in 2017, Nucleic Acids Res 45 (2017) D945–D954.
12 G.R. Bickerton, G.V. Paolini, J. Besnard, S. Muresan, A.L. Hopkins, Quantifying the chemical beauty of drugs, Nat Chem 4 (2012) 90–98.
13 BIOVIA, Dassault Systèmes, BIOVIA Pipeline Pilot, Release 2020, San Diego: BIOVIA
14 D.C. Kombo, K. Tallapragada, R. Jain, J. Chewning, A.A. Mazurov, J.D. Speake, et al., 3D molecular descriptors important for clinical success, J Chem Inf Model 53 (2013) 327–342.
15 F. Lovering, J. Bikker, C. Humblet, Escape from flatland: Increasing saturation as an approach to improving clinical success, J Med Chem 52 (2009) 6752–6756.
16 W. Wei, S. Cherukupalli, L. Jing, X. Liu, P. Zhan, Fsp3: a new parameter for drug-likeness, Drug Discov Today 25 (2020) 1839–1845.
17 G.W. Bemis, M.A. Murcko, The properties of known drugs. 1. Molecular frameworks, J Med Chem 39 (1996) 2887–2893.
18 J.L. Medina-Franco, K. Martínez-Mayorga, A. Bender, T. Scior, Scaffold diversity analysis of compound data sets using an entropy-based measure, QSAR Comb Sci 28 (2009) 1551–1560.
19 J. Xu, A new approach to finding natural chemical structure classes, J Med Chem 45 (2002) 5311–5320.
20 C.A. Lipinski, Drug-like properties and the causes of poor solubility and poor permeability, J Pharmacol Toxicol Methods 44 (2000) 235–249.
21 H. van de Waterbeemd, G. Camenisch, G. Folkers, J.R. Chretien, O.A. Raevsky, Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors, J Drug Target 6 (1998) 151–165.
22 P.W. Kenny, C.A. Montanari, Inflation of correlation in the pursuit of drug-likeness, J Comput Aided Mol Des 27 (2013) 1–13.
23 T. Flagstad, C.M.G. Azevedo, G. Min, A. Willaume, R. Morgentin, T.E. Nielsen, M. H. Clausen, Petasis/Diels-Alder/Cyclization cascade reactions for the generation of scaffolds with multiple stereogenic centers and orthogonal handles for library production, Eur J Org Chem 36 (2018) 5023–5029.
24 E.M.M. Abdelraheem, R. Madhavachary, A. Rossetti, K. Kurpiewska, J. Kalinowska-Tłuścik, S. Shaabani, et al., Ugi multicomponent reaction based synthesis of medium-sized rings, Org Lett 19 (2017) 6176–6179.
25 P. Patil, R. Madhavachary, K. Kurpiewska, J. Kalinowska-Tłuścik, A. Dömling, De novo assembly of highly substituted morpholines and piperazines, Org Lett 19 (2017) 642–645.
26 A.T. Murray, E. Packard, A. Nortcliffe, W. Lewis, D. Hamza, G. Jones, et al., Synthesis of epibatidine analogues by pyrrole Diels-Alder reactions: rapid access to azabicyclo[2.2.1]heptane and 3,8-diazabicyclo[3.2.1]octane scaffolds for library synthesis, Eur J Org Chem 1 (2017) 138–148.
27 V. Estévez, L. Kloeters, N. Kwietniewska, E. Vicente-García, E. Ruijter, R. Orru, Ugi-type reactions of spirocyclic indolenines as a platform for compound library generation, Synlett 28 (2016) 376–380.
28 S.J. Cully, T.E. Storr, M.J. Rawling, I.R. Abeysena, D. Hamza, G. Jones, et al., Expedient synthesis of an atypical oxazolidinone compound library, Bioorg Med Chem 24 (2016) 5249–5257.
29 T. Flagstad, G. Min, K. Bonnet, R. Morgentin, D. Roche, M.H. Clausen, et al., Synthesis of sp^3-rich scaffolds for molecular libraries through complexity-generating cascade reactions, Org Biomol Chem 14 (2016) 4943–4946.
30 P. Wu, M.Å. Petersen, R. Petersen, T. Flagstad, R. Guilleux, M. Ohsten, et al., Tandem Mannich/Diels–Alder reactions for the synthesis of indole compound libraries, RSC Adv 6 (2016) 46654–46657.
31 B.D. Narhe, A.C. Breman, J. Padwal, D.A.L. Vandenput, J.M. Scheidt, J.C.J. Benningshof, et al., Piperidine and octahydropyrano[3,4-c] pyridine scaffolds for drug-like molecular libraries of the European Lead Factory, Bioorg Med Chem 25 (2017) 5160–5170.
32 T. Kogej, N. Blomberg, P.J. Greasley, S. Mundt, M.J. Vainio, J. Schamberger, et al., Big pharma screening collections: more of the same or unique libraries? The AstraZeneca–Bayer Pharma AG case, Drug Discov Today 18 (2013) 1014–1024.