Fragment-Based Hit Identification – Thinking in 3D

Kristian Birchall[1], Justin Bower[2], Paul Brennan[3], Nathan Brown[4], Tim Chapman[1], Ian Gilbert[5], Swen Hoelder[4], Allan Jordan[6], Steve V. Ley[7], Andy Merritt[1], David Miller[8], Andrew D. Morley*[5], Angelo Pugliese[2], Martin Swarbrick[9] and Paul G. Wyatt[5]

[1]Medical Research Council Technology, 1-3 Burtonhole Lane, Mill Hill, London, NW7 1AD, U.K.
[2]The Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, U.K.
[3]Structural Genomics Consortium, Old Road Campus Research Building, Roosevelt Drive, Oxford, OX3 7DQ, U.K.
[4]The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, U.K.
[5]University of Dundee, College of Life Sciences, 149 Nethergate, Dundee, DD1 4HN, U.K.
[6]The Paterson Institute for Cancer Research, The University of Manchester, Wilmslow Road, Manchester, M20 4BX, U.K.
[7]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K.
[8]University College London, Gower Street, London, WC1E 6BT, UK
[9]Cancer Research Technology Ltd, CRT Discovery Laboratories, Jonas Webb Building, Babraham Research Campus, Cambridge, CB22 3AT, U.K.

The identification of high quality hits in the early phases of drug discovery is essential if projects are going to have a realistic chance of progressing into clinical development and delivering marketed drugs. As the pharmaceutical industry goes through unprecedented change, there are increasing opportunities to collaborate via pre-competitive networks in order to marshal multi-functional resource and knowledge to drive impactful, innovative science. The 3D Fragment Consortium is developing fragment screening libraries with enhanced three dimensional characteristics and evaluating their effect on the quality of Fragment Based Hit Identification projects.

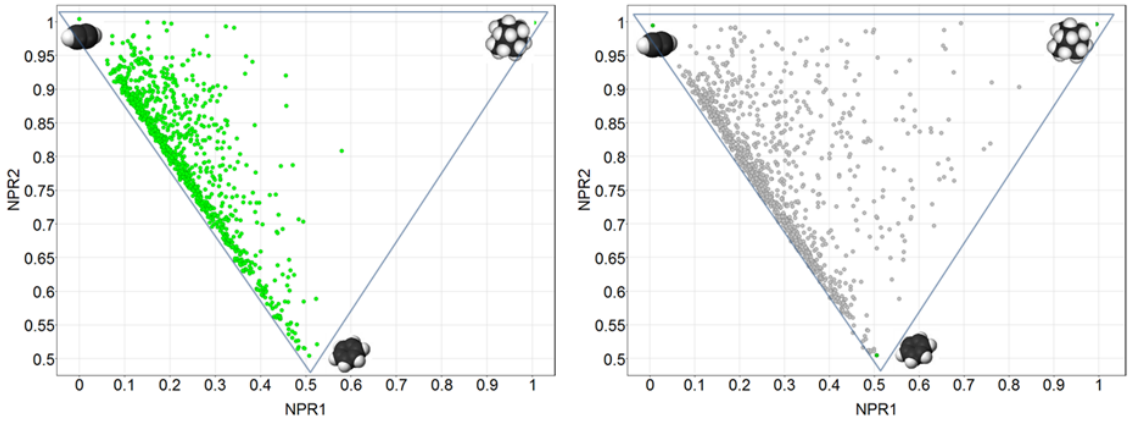
Introduction

Fragment screening is an established method to generate high quality hits.[1] The successful approval of Vemurafenib in 2011 for late stage melanoma, validates the approach to support the discovery of marketed drugs.[2] Small (typically 1000 member) libraries of lower molecular weight (<300 Daltons) allow for more efficient sampling of chemical space.[3] Screening is mostly undertaken using biophysical techniques as conventional biochemical techniques are often not sufficiently sensitive to identify the modest (though efficient) affinity of fragments for the target protein. Analysis suggests that projects deploying fragment screening generate smaller, less lipophilic hits and leads than those identified through the use of other screening techniques, such as high throughput screening.[4] Drastic recent changes across the drug discovery sector has seen a shift to smaller, leaner research organisations, making "traditional" high throughput screening prohibitive for these smaller enterprises due to high costs of generating suitably sized screening collections. Subsequently, Fragment Based Hit Identification (FBHI) has been implemented by several recently formed drug discovery groups, both in academia and in the biotech sector, focusing on the ability to rapidly progress modestly active hits in a proficient manner.

While fragment libraries are believed to efficiently sample chemical space, it has been suggested that an imbalance exists regarding success rates across all target classes, although to our knowledge no sufficiently detailed analyses have been undertaken to corroborate this assumption. Evaluation of a number of fragment libraries shows that they are predominantly populated with (hetero)aromatic derived chemotypes, which may predispose their success for certain targets. Both GSK[5] and Pfizer[6] have published data showing the improvements in profile and project progression by, for example, increasing the proportion of sp^3 centres contained in molecules or reducing LogP, although some of the statistics in these reviews has recently been questioned.[7] Whilst most fragment libraries are diverse having been selected to contain a good balance of properties,[8] they all tend to have limited shape diversity and this "flatness" could explain why they are less successful in identifying hits for certain targets which may require alternative representations, substitution vectors and key binding functionalities to interact with these proteins. We have compared typical fragment

libraries and the fragments embedded in compounds which have been evaluated in humans.[9] The results of this analysis are depicted in Figure 1. This analysis uses principal moments of inertia (PMI)[10] as a simple way to calculate and evaluate the three dimensional diversity. Molecules have their lowest energy three dimensional conformer generated and this is used to calculate the proportion of "rod-like", "disc-like" and "sphere-like" characteristics. These values are sorted by ascending magnitude, then normalised, dividing the two lowest values by the highest to generate normalised PMI ratios (NPR1/2) that can be plotted in two dimensions to provide the triangular output shown in Figure 1. Fragments contained within clinically evaluated compounds appear to have greater three dimensional conformations than those in the typical fragment library. This may provide a reason for the limited tractability for certain target classes for the latter.

Figure 1 PMI of a typical fragment library (green) and ZINC InMan subset fragmented by RECAP (Grey)



Building libraries around fragments with greater three dimensionality is an area of intense current debate. Molecules of this type are expected to contain different substituent vectors which should generate alternative pharmacophoric relationships to those of flatter, sp^2 rich, molecules. Chemists intuitively reason that compounds with greater three dimensional characteristics will be more complex as a result of higher numbers of sp^3 centres, stereochemical relationships *etc.* and will possess lower hit rates in fragment screens (the latter point based on the complexity theory proposed by Hann and co-workers[3]). However, three dimensional conformers can be generated by molecules comprised almost entirely of sp^2 centred atoms and conversely flat conformers can be derived from fragments with a large number of chiral centres. A number of approaches have been published that allow molecular complexity to be calculated. Using methods that are independent of atom, bond types and connectivity provide a means to assess overall characteristics free from chemical bias.[11] A key goal is to try and broaden the diversity of a library without significantly increasing the complexity of the selected fragments. Subsequently, complexity calculations can be helpful in guiding library design. Recent disclosures from Evotec suggest that although flatter fragments tend to have higher hit rates, the difference to more shaped fragments is not significant.[12] A recent analysis by Jorgensen and co-workers showed that addition of a methyl group can produce significant potency (and ligand efficiency) increases, primarily through conformational changes, generating more shaped analogues in the specific highlighted cases.[13] In addition, work by AstraZeneca and Astex on BACE showed the significant improvement in potency and ligand efficiency by dearomatisation and introduction of a methyl substituent in order to preferentially adopt a lower energy conformer with the correct vector to access the S1 and S3 pockets.[14] The principal of screening flexible ligands at high concentration in combination with X-ray crystallography as a start point to improving ligand effciency through constraning the linker has shown to be feasible.[1] However, these types of transformations are not always successful and having a percentage of conformationally constrained fragments in your screening library could provide excellent start points to allow rapid progression in certain instances. Achieving a suitable balance of diverse properties and characteristics is an important aspect to try and address as part of building a library such as we are describing and delivering outputs with increased overall quality. It is unlikely that a constrained core scaffold will be maintained throughout the entirity of the project, without some attempt to modify it being made, either to improve compound quality or solve issues as they are identified as projects

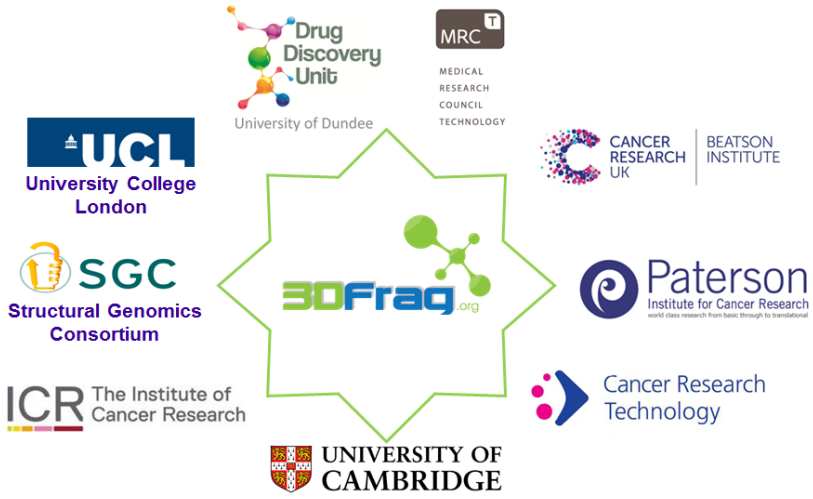
progress through the discovery phases. Robust synthetic routes to attractive scaffolds need to be available and developed to allow timely evolution.

Establishing the 3D Fragment Consortium

To address the limited shape diversity available within typical current fragment libraries along with a desire to increase project success rates and overall compound profile quality, a number of UK not-for-profit drug discovery groups, spanning a range of therapeutic *foci*, have come together to form the 3D Fragment Consortium.[15] The goal is to build a shared library of between 500 and 3000 fragments that includes components found in many existing libraries whilst adding a complementary set of compounds possessing greater three dimensionality. By focusing on the highlighted strategy, the project aims to evaluate whether these latter compounds possess; lower overall hit rates; improved hit rates for target classes with low tractability; superior ligand efficiencies, compared to current libraries. The consortium's diverse capabilities allow it to prosecute targets deploying a variety of biophysical and other screening methods as well as evaluating the library across a diverse set of biological targets. The small size and typically weak potency of fragment hits should limit any intellectual property concerns arising from using a shared library and allow a broad data set to be compiled and analysed. A key principle is to share aspects of the screening output from this joint library to assess if the 3D fragment approach will be beneficial, as highlighted above and to explore whether areas of chemical space are associated with particular target classes

The core consortium membership is shown in Figure 2. An additional aim of this project is to develop networks involving scientists outside the consortium, such as pharma, biotech, commercial and academic groups. This is believed to be an excellent opportunity to bring together a large number of UK drug discovery groups and help improve impact and success rate for the sector through appropriate knowledge and capability sharing.

Figure 2 3D Fragment Consortium Members

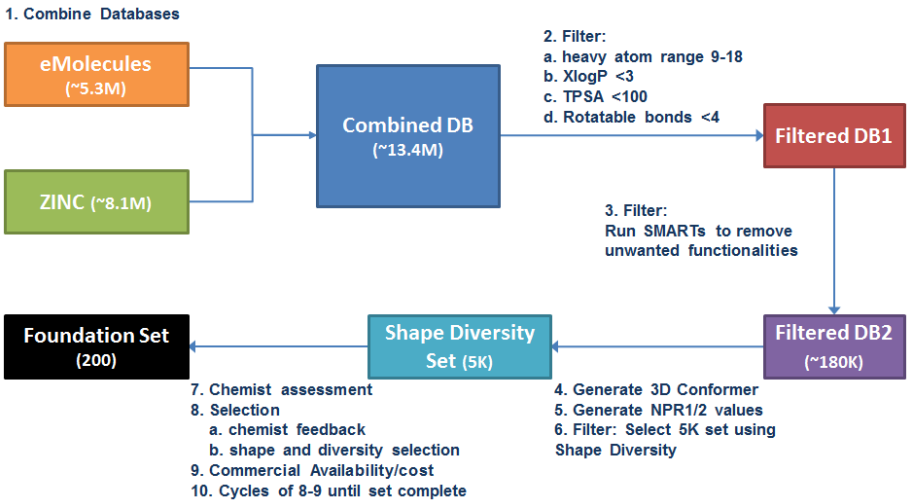


The 3D Fragment Consortium Foundation Set

There are a variety of definitions of what constitutes a fragment; to help with the selection of compounds, the consortium set their own guideline criteria (Figure 3 and supplementary information) based around other published ideas and work.[8, 16] A number of potential sources of fragments are available for consideration: commercial compounds; synthetic molecules, from various literature sources; and *de novo* designed fragments. To kick start the project, the consortium selected and purchased an initial foundation library of ~200 commercial fragments. The aim of generating this common fragment set with diverse shape profile was to generate a uniform data set to evaluate across different targets and classes to provide a starting point to

allow prioritisation and selection of new compounds for procurement/synthesis. The selection process for the library is shown in Figure 3.

Figure 3 Selection process to identify the consortium's foundation library

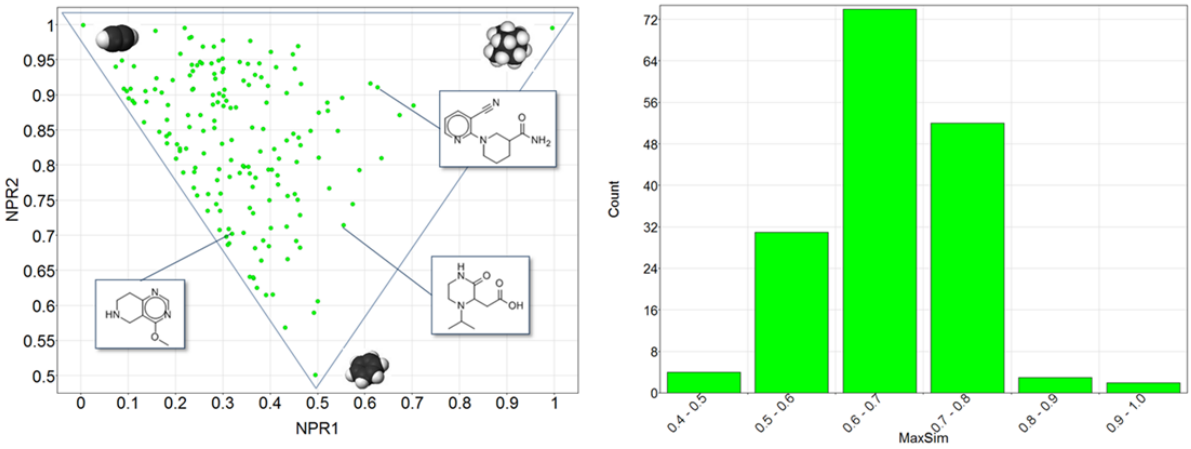


The eMolecules[17] and ZINC[18] databases were combined to give 13.4 million compounds. Filtering on heavy atom count (non hydrogen atoms) and removing molecules with unwanted functional groups left 180 thousand compounds. Three-dimensional conformations were generated using Accelrys's Conformer and the shape descriptors NPR1 and NPR2 were calculated.[10] A subset of five thousand compounds was selected on the basis of shape diversity. A medicinal chemist from each consortium institute visualised each compound in this set and defined whether it was an attractive fragment for screening purposes in their view. The selection of compounds prefered by a majority of chemists was further refined by diversity selection and commercial availability to deliver a final collection of 200 compounds. This workflow suffered from expected attrition at various stages. These included removal of >90% of commercial samples in early evaluation processes, as they did not meet criteria agreed by consortium members. This generated a similar conclusion to those recently reported by chemists at GSK.[19] Although commercial samples meeting these proposed profile guidelines cover the majority of the PMI triangle, >70% have an NPR1 + NPR2 value <1.1, characteristic of flat molecules. Fragments with increased three dimensionality occur at a much lower frequency. A number of interesting commercial scaffolds with useful shape characterisitcs were identified as part of this process, but feedback across the consortium highlighted derivitisation would be necessary for them to be suitable for screening purposes (e.g. poly ionic compounds). Out of the selected foundation set, approximately 10% of the library could not be delivered by the suppliers. Another 5% of those delivered to the consortium failed QC when evaluated (poor solubility or decomposition by NMR assessment).

The PMI and maximum similarity[20] profiles for the resultant library of ~170 are shown in Figure 4, the latter used as a straightforward way to visualise and assess library diversity. Nearest neighbour fingerprint-based similarities were calculated and the frequency (count) plotted. A library with a high degree of internal similarity will have a higher scoring (right shift) histogram, which could be useful for screening targets where key binding interactions are already identified and mimicry is the goal. Conversely, a lower scoring (left shift) histogram, could be more useful for general screening across wide ranging targets/classes. If the majority of compounds within a library have a nearest neighbour similarity <0.8, then it can be considered to be reasonably diverse[21] and this is supported by eye inspection of the foundation set by consortium chemists. The complexity score for the library was also calculated using the method of Nilakatan.[11] The mean complexity score for the foundation library is 35.6 compared to the library shown in left hand plot of Figure 1, which has a mean complexity score of 38. This demonstrates that libraries possessing greater three dimensional charcteristics can be generated without resorting to using only complex scaffolds. The mean heavy atom count for both

libraries is ~13, so the size of molecules isn't influencing the complexity score. The availability of this foundation library helps provide a benchmark against which to measure new compounds and ideas, prior to their selection and inclusion and the data also helps ensure the correct balance of properties can be maintained as the library grows.

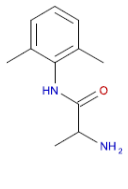
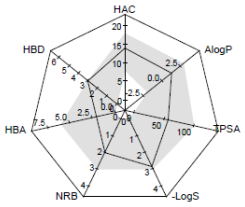
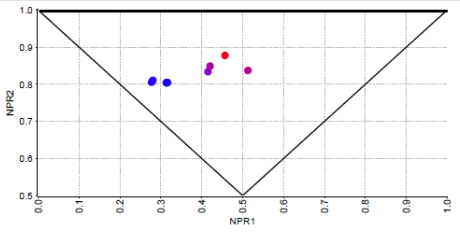
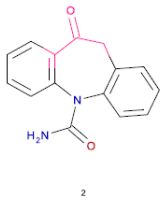
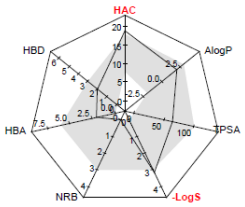
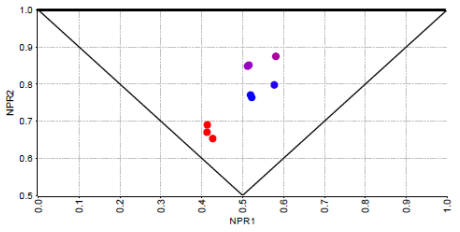
Figure 4 PMI and MaxSim distribution for the foundation library of ~170 compounds



3DFIT: an online tool to evaluate fragments

In order to enable collaboration both within and beyond the consortium, a major effort was undertaken to develop a web based tool, named 3DFIT (3D Fragment Idea Tool). This allows users to assess the shape and predicted physiochemical properties of molecules, using a readily accessible Pipeline Pilot interface.[22] Instant feedback is provided (Figure 5) showing where compounds (commercial or synthetic ideas) align within three dimensional space along with providing other key data to facilitate decision making. For each compound up to nine low energy conformers are generated and plotted on the PMI triangle. These are colour coded (blue for lowest energy to red for highest calculated energy conformer) to allow easy comparison. As can be seen compounds can generate diverse shapes within the energy range and molecules with strong three dimensional conformations can be generated from compounds consisting almost entirely of sp^2 centred atoms. The consortium has collated a set of ~500 SMARTs to flag structural motifs that it considers could be detrimental (reactive, safety related issues etc.). Key physicochemical properties are calculated and depicted as a radar plot. The consortium guidelines for acceptable parameters are highlighted within the shaded region and those parameters outside these boundaries are flagged for consideration, as can be seen for the second example in Figure 5. 3DFIT allows rapid feedback to chemists to assess their compound ideas for potential inclusion into the 3D fragment library and highlights aspects of the molecules where improvements to profiles could be considered.

Figure 5 3DFIT Output for non-foundation library compounds (tocainide and oxacarbazepine) showing compound structure, predicted compound properties (plotted in relationship to consortium guidelines) plus visualisation of the lowest energy conformers using principal moments of inertia scaling

Molecule	Radar plot	Properties table	PMI plot
		HAC 14 ALogP 1.45 TPSA 55.1 -LogS 2.86 NRB 2 HBA 2 HBD 3	
		HAC 19 ALogP 2.13 TPSA 63.4 -LogS 3.12 NRB 0 HBA 2 HBD 2	

The ideas are captured to a linked database and more detailed off line analysis can be undertaken, refining the original idea, if necessary, prior to finalising selection of compounds for purchase/synthesis. This also ensures that the compounds selected allow generation of a diverse library. 3DFIT has just completed its initial beta testing and several additional concepts and suggestions are being considered to increase the value of the tool further. It also acts as an invaluable communication device across the consortium and the wider synthetic community and is also used as a learning capability to help those less familiar with medicinal chemistry concepts to appreciate the impact of different functionalities and substitution patterns on compound profiles.

Fragmentation of Bioactive Chemical Space

In parallel with the initiatives described above, the consortium has undertaken thorough analyses of known biologically relevant chemical space by evaluating the ChEMBL[23] database and the Zinc "InMan"[9] subset plus a set of patent compounds kindly provided through a collaboration with SureChem.[24] Molecules were fragmented using the RECAP algorithmn,[25] based on predefined cleavage rules, converted to chemically sensible structures, filtered against consortium agreed parameters and evaluated by a number of chemists to determine which of the resultant molecules were broadly attractive for further consideration. This process generated approximately 7000 "shaped" scaffolds. Whilst the majority are not suitable for fragment screening themselves, due to the nature of the embedded functional groups generated through the RECAP process or small size, they do provide a powerful set of biologically relevant templates. These were derivitised, through enumeration processes, generating a virtual set of analogues, which was combined with a preliminary output from 3DFIT. The total set was compared against the foundation set to select compounds for synthesis prioritisation. Expanding the foundation library using this strategy, ensures good diversity can be maintained as the library grows and also allows compounds awaiting synthesis to be replaced should superior fragments subsequently be identified.

Synthesis of new 3D Fragments

In parallel with the initiatives described above, the consortium has undertaken analyses of published biologically relevant chemical space, which will be published shortly. The consortium believe that these data sets will be invaluable, having identified a number of interesting non-commercial scaffolds to produce highly attractive fragment to supplement the library. We've concluded that bespoke synthesis, rather than expansion through acquisition of currently available commercial fragment sized compounds is the most appropriate way to develop the library to attain the desired profile. The project has used various methods and resource to rapidly deliver >60 additional compounds into the library. These include novel fragments being synthesised internally by consortium members. The first wave of these are currently being included to support future screening. It is intended that this is the start of a process that will grow and develop over time. The consortium has also successfully secured: post doctoral funding; PhD student support with reknowned academics to make 3D Fragments; short-term secondees from other pharma companies to assist project delivery; in house chemistry co-operation from the member institutes and students working in discovery labs as part of their training programmes as well as summer studentships. In addition, consortium members are assessing in-house

inventories for building blocks and scaffolds to share in a pre-competitive manner to further enrich the joint library. The need to synthesise novel molecules that expand biologically relevant chemical space demonstrates a significant role that academic synthetic chemistry can play in facilitating target evaluation and generating the most appropriate start points for drug discovery programmes. A number of groups are devising new and innovative methodologies (i.e. CH activation, cascade reactions, enzymatic functionalization) and techniques (e.g. flow and photochemistry) which can be harnessed to facilitate expansion of drug discovery relevant chemical space. Many interesting scaffolds are being generated in academic labs, the majority of which are more three dimensional than many commercial fragments. Many will be of high value to drug discovers and the use of 3DFIT acts as an interface to enhance the quality chemistry or derivatives that would be considered attractive. The approach synergises with the substantial investment in world-class synthetic chemistry methodology development and initiatives (i.e. Dial-a-Molecule) by the EPSRC and others.

By analysing current biologically relevant chemical space and selecting diverse sets of compounds, regions of the library that are not currently represented can be identified by the consortium to determine if these areas require populating. Computational assessment and evaluation of the Generated DataBases (GDB)[26] of Jean Louis Reymond plus structural databases such as PDB[27]/CREDO[28] also potentially offer highly valuable information to support the latter stages of the library build when specific regions of the library will require expanding. Once target compounds have been identified, the consortium envisage engagement with academic groups to define routes and proposals to synthesise these interesting fragments. The consortium is keen to also work with computational chemistry and chemoinformaticians to build on the analyses and selection processes it has undertaken to increase the quality of the library over time. Developing the interface between multiple chemistry functions (e.g. synthetic, medicinal and computational chemistry) will ensure the highest priority compounds are delivered in the most efficient manner.

Conclusion

The team believe that the library they are building will be a highly valuable asset that could extensively impact medicinal chemistry and drug discovery. The perception that screening fragments with a greater degree of three dimensionality will significantly reduce hit rates is a valid concern. However, we believe the strategy will build a library with broader coverage of biologically relevant chemical space compared to current fragment collections and the hits that it will identify will be more tractable and offer higher quality start points to progress projects. The aim of the project is to test and hopefully validate this hypothesis. Provision of suitable libraries and screening facilities to cost effectively support the new cutting edge biology emanating from small academic groups is still missing and subsequently this is limiting opportunities to identify and progress innovative early phase drug discovery projects. Recent evaluation of fragment screening outputs against downstream success rates means the library could be efficiently used as a tractability assessment for new targets towards small molecule modulation.[29] In addition, screening the library against novel targets would allow suggestions for library supplementation to be fed back to develop chemical space as these biologies come on line.

The project, however, is a high risk concept that is currently the focus for much debate across the sector. Owing to the considerable cost and high attrition rate in drug discovery, the need to collaborate, share knowledge, resource, equipment and assets becomes ever more apparent, if the sector is going to remain competitive and improve its impact and success rates. The building of a pre-competitive forum is seen by many as a pivotal arena to facilitate these goals as the industry becomes more diluted. However, for this to successfully evolve and generate adequate momentum requires all the key parties to actively support the concept and its delivery. In order for this environment to develop, interested parties need to embrace the model of willingly providing information and support, in order to receive significantly more in return. Industry wide focus to deliver internal milestone driven project activities with limited resource, on schedule, results in inadequate support being available for innovative “blue sky” work such as the undertaking described here. Getting groups to agree an acceptable balance of participation and provision is subsequently challenging, which has meant progress has not been as rapid as originally hoped. Identifying suitable mechanisms to support key multi-disciplinary projects with the potential to significantly impact this pivotal industry is essential in helping it to stabilise, allow medium to long term development and delivery. The learning through the generation and diverse testing of a library of this type would provide be a significant resource for the pharmaceutical industry, particularly if it can be evaluated alongside other related activities.

This project is still in its infancy, having made progress in the planning stage, providing a strong foundation on which to build and it is now keen to move to a more sustained delivery phase. Although the project team are

enthusiastic to continue and attempt to answer a key question being raised across the sector, it has become clear that to complete the library build will require broader internal and external support. The team are looking to achieve this through core initiatives alongside identification of interested parties to share and support the workload and help the project advance. It would be interested in hearing views from across the industry regarding how some of these challenges could be tackled.

References

1. (a) M. Baker, Nature Reviews Drug Discovery, 2013, 12, 5-7 (b) R. E. Hubbard, J. B. Murray, Methods Enzymol. 2011, 493:509-531 (c) M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, J. Med. Chem. 2008, 51:3661-3680 (d) D. A. Erlanson, R. S. McDowell, T. O'Brien, J. Med. Chem., 2004, 47, 3463-82
2. C. W. Murray, M. L. Verdonk, D. C. Rees, Trends in Pharmacological Sciences; 2012, 1–9
3. M. M. Hann, A. R. Leach, G. Harper; J. Chem. Inf. Comput. Sci. **2001,** 41, 856-864
4. P. D. Leeson, S. A. St-Gallay; Nature Reviews Drug Discovery 2011, 10, 749-765
5. T. J. Ritchie, S. J. F. MacDonald; Drug Discovery Today 2009, 14, 1011–1020
6. F. Lovering, J. Bikker, C. Humblet; J. Med. Chem. 2009, 52, 6752–6756
7. Peter W. Kenny, Carlos A. Montanari, J. Comput. Aided Mol. Des. 2013, 27, 1-13
8. (a) M. Congreve, R. Carr, C. Murray, H. Jhoti; Drug Discovery Today, 2003, 8, Pages 876–877 (b) N. Blomberg, D. A. Cosgrove, P. W. Kenny, K. Kolmodin K, J. Comput. Aided Mol. Des. 2009, 23, 513-525
9. http://zinc.docking.org/subsets/zim
10. W. H. B. Sauer, M. K. Schwarz; J. Chem. Inf. Comput. Sci. 2003, 43, 987-1003
11. Nilakantan, R., Nunn, D.S., Greenblatt, L., Walker, G., Haraki, K., and Mobilio, D., A family of ring system-based structural fragments for use in structure-activity studies: database mining and recursive partitioning. J Chem Inf Mod *46*, 1069-1077.
12. D. Ullman, O. Barker, Fragments 2013, 4th RSC-BMCS meeting 3rd-5th March 2013
13. C. S. Leung , S. S. F. Leung , J. Tirado-Rives, W. L. Jorgensen, *J. Med. Chem.* 2012, 55, 4489–4500
14. P. D. Edwards, J. S. Albert, M. Sylvester, D. Aharony, D. Andisik, O. Callaghan, J. B. Campbell, R. A. Carr, G. Chessari, M. Congreve, M. Frederickson, R. H. A. Folmer, S. Geschwindner, G. Koether, K. Kolmodin, J. Krumrine, R. C. Mauger, C. W. Murray, L-L. Olsson, S. Patel, N. Spear, G. Tian, *J. Med. Chem.* **2007,** *50,* 5912-5925
15. http://www.3DFrag.org
16. (a) R. Brenk, A. Schipani, D. James, A. Krasowski, I. H. Gilbert, J. Frearson, P. G. Wyatt, ChemMedChem, 2008, 3, 435-444 (b) W. F. Lau, J. M. Withka, D. Hepworth, T. V. Magee, Y. J. Du, G. A. Bakken, M. D. Miller, Z. S. Hendsch, V. Thanabal, S. A. Kolodziej, L. Xing, Q. Hu, L. S. Narasimhan, R. Love, M. E. Charlton, S. Hughes, W. P. van Hoorn, J. E. Mills, J. Comput. Aided Mol. Des. 2011, 25, 621-636
17. http://www.emolecules.com/
18. http://zinc.docking.org/subsets/all-now
19. A. Nadin, C. Hattotuwagama, I. Churcher, Angew. Chem. Int. Ed.; 2012, 51, 1114 – 1122
20. Canvas, version 1.5, Schrodinger, LLC, New York, NY, 2012
21. (a) Matter, H., J. Med. Chem, 1997, **40**, 1219-1229 (b)Taylor, R., J. Chem. Inf. Comput. Sci., 1995, **35** ,59-67
22. http://www.accelrys.com/products/pipeline-pilot/
23. https://www.ebi.ac.uk/chembl/
24. https://www.surechem.com/
25. X. Q. Lewell , D. B. Judd , S. P. Watson , M. M. Hann, J. Chem. Inf. Comput. Sci. 1998, 38, 511–522
26. (a) L. Ruddigkeit, R. van Deursen, L. C. Blum and J.-L. Reymond; *J. Chem. Inf. Model.,* **2012**; *52*, 2864-2875. (b) L. Ruddigkeit, R. van Deursen, L. C. Blum and J.-L. Reymond; *J. Chem. Inf. Model.;* **2012**, *52*, 2864-2875.
27. http://www.pdb.org
28. http://www-cryst.bioc.cam.ac.uk/databases

29. F. N. B. Edfeldt, R. H. A. Folmer, A. L. Breeze Drug Discovery Today, 2011 ;16, 284-7