



UNIVERSITY OF LEEDS

This is a repository copy of *Construction of a Shape-Diverse Fragment Set: Design, Synthesis and Screen against Aurora-A Kinase*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/144401/>

Version: Accepted Version

Article:

Zhang, R, McIntyre, PJ, Collins, PM et al. (6 more authors) (2019) Construction of a Shape-Diverse Fragment Set: Design, Synthesis and Screen against Aurora-A Kinase. *Chemistry - A European Journal*, 25 (27). pp. 6831-6839. ISSN 0947-6539

<https://doi.org/10.1002/chem.201900815>

© 2019 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. This is the peer reviewed version of the following article: Zhang, R, McIntyre, PJ, Collins, PM et al. (6 more authors) (2019) Construction of a Shape-Diverse Fragment Set: Design, Synthesis and Screen against Aurora-A Kinase. *Chemistry - A European Journal*, 25 (27). pp. 6831-6839. ISSN 0947-6539, which has been published in final form at <https://doi.org/10.1002/chem.201900815>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Construction of a Shape-Diverse Fragment Set: Design, Synthesis and Screen against Aurora-A Kinase

Rong Zhang,^[a,b] Patrick J. McIntyre,^[c] Patrick M. Collins,^[d] Daniel J. Foley,^[a,b] Christopher Arter,^[a,b] Frank von Delft,^[d,e,f] Richard Bayliss,^{[a,g],*} Stuart Warriner^{[a,b],*} and Adam Nelson^{[a,b],*}

Abstract: Historically, chemists have explored chemical space in a highly uneven and unsystematic manner. As an example, the shape diversity of existing fragment sets does not generally reflect that of all theoretically possible fragments. To assess experimentally the added value of increased three dimensionality, a shape-diverse fragment set was designed and collated. The set was assembled using both commercially-available fragments, and by harnessing unified synthetic approaches to sp³-rich molecular scaffolds. The resulting set of 80 fragments was highly three-dimensional, and its shape diversity was significantly enriched by twenty synthesised fragments. The fragment set was screened by high-throughput protein crystallography against Aurora-A kinase, revealing four hits that targeted the binding site of allosteric regulators. In the longer term, it is envisaged that the fragment set could be screened against a range of functionally-diverse proteins, allowing the added value of more shape-diverse screening collections to be more fully assessed.

Introduction

The highly uneven and unsystematic way in which chemists have explored chemical space historically^[1] stems, in large part, from the narrow reaction toolkit that dominates bioactive small molecule discovery.^[2] As a result, medicinal chemists, have increased their attention on flatter and less polar small molecules.^[3] Yet, lead molecules and clinical candidates with a lower fraction of sp³-hybridised carbons (Fsp³) are less likely to be translated successfully into marketed drugs.^[4]

Within the fragment-based discovery community, the optimal shape diversity of fragment sets has been intensely debated.^[5] Existing fragment sets are dominated by flatter (generally heteroaromatic) molecules,^[5a] whose shape diversity is not representative of the fragments that are theoretically possible.^[6]

This low shape diversity has spurred the assembly of sets of fragments with high fractions of sp³-hybridised carbons.^[5c-h] It has been argued that the increased complexity of more three-dimensional fragments will tend to lower hit rates because of the reduced likelihood^[7] of complementing protein binding sites.^[5a-b,8] On the other hand, such fragments may offer opportunities for elaboration along distinctive vectors.^[5b] In addition, the flat nature of many existing fragments has been suggested to be sub-optimal for some target classes including protein-protein interactions.^[8] There have, however, been few studies in which the effect of fragment three-dimensionality has been investigated experimentally. In one study, the hits from 30 fragment screens were analysed:^[9] with a disparate range of target proteins, it was found that flatter fragments enjoyed significantly – but not dramatically – higher hit rates.

In this paper, we describe the design and collation of a shape-diverse fragment set. To complement commercially-available fragments, we decided to harness four unified approaches that we have developed for the synthesis of diverse, novel and sp³-rich molecular scaffolds.^[10-13] As a preliminary assessment of this fragment set, we performed a fragment screen against Aurora-A kinase using high-throughput protein crystallography. In the longer term, we envisage that the fragment set could be screened against many different target proteins, allowing the added value of more shape-diverse screening collections to be more fully assessed.

Results and Discussion

Design of a set of eighty shape-diverse fragments

Initially, a set of eighty shape-diverse fragments was designed (Figure 1) based on both commercially-available fragments and compounds that would likely to be accessible using our lead-oriented^[14] synthetic approaches. In total, four synthetic approaches were selected in which pairs of building blocks may be converted into diverse sp³-rich scaffolds: (1) Ir-catalysed allylic amination and cyclisation;^[10] (2) cyclisation of α -allyl α -amino ester derivatives;^[11] (3) cyclic sulfamidate ring-opening and intramolecular Mitsunobu reaction;^[12] and (4) cyclic sulfamidate ring-opening, Au-catalysed reaction and Ugi reaction.^[13] A total of 63 tangible scaffolds was designed that had been, or were related to those that had been, previously prepared. Envisaged syntheses of some exemplar tangible scaffolds are summarised in Figure 2. The tangible scaffolds were virtually decorated once or twice using a set of small capping groups (Supporting Information). Following removal of compounds with undesired functionality (e.g. aldehydes or terminal alkenes), a virtual library of 66 814 compounds was obtained.

The virtual compounds and the ZINC database^[15] of commercially-available compounds were filtered for molecular size. The heavy atom filter was selected to maximise the

[a] Dr R. Zhang, Dr Daniel J. Foley, Christopher Arter, Prof. R. Bayliss, Dr S. Warriner and Prof. A. Nelson
Astbury Centre for Structural Molecular Biology,
University of Leeds, Leeds, LS2 9JT, UK
E-mail: a.s.nelson@leeds.ac.uk

[b] Dr R. Zhang, Dr S. Warriner and Prof. A. Nelson
School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK

[c] P. J. McIntyre, Department of Molecular and Cell Biology, Henry
Wellcome Building, University of Leicester, Leicester, LE1 9HN, UK

[d] Dr P. M. Collins and Prof. F. von Delft,
Diamond Light Source Ltd, Harwell Science and Innovation
Campus, Didcot, OX11 0DE, UK

[e] Prof. F. von Delft, Structural Genomics Consortium, Nuffield
Department of Medicine, University of Oxford, Oxford, OX3 7DQ,
UK

[f] Prof. F. von Delft, Department of Biochemistry, University of
Johannesburg, Auckland Park, 2006, South Africa

[g] Prof. R. Bayliss, School of Molecular and Cellular Biology, University
of Leeds, Leeds, LS2 9JT, UK

Supporting information for this article is given via a link at the end of the document. ((Please delete this text if not appropriate))

represented tangible scaffolds in the virtual fragment set. With a 18-22 heavy atom range, all but one of the tangible scaffolds was represented in the virtual fragment set. In contrast, significantly more tangible scaffolds were excluded with either a 16-20 (3 excluded scaffolds) or a 20-24 (7 excluded scaffolds) heavy atom range. At this stage, 100,000 of the filtered ZINC compounds were selected at random.

To capture the shape diversity of the combined fragment sets, a set of reference shapes was generated using an established protocol.^[16] In brief, a representative conformer^[17] was initially generated for each compound using CORINA.^[18] Then, a randomly chosen fragment was selected to be the first reference shape, and its shape similarity to all other fragments assessed by maximising the overlapping volume of the overlaid molecules using ROCS.^[19] Shape-similar fragments (with shape Tanimoto > 0.8) were removed, and the most shape dissimilar fragment was selected as the second reference shape. The process was repeated until no fragments remained, resulting in a set of 2,477 reference shapes.

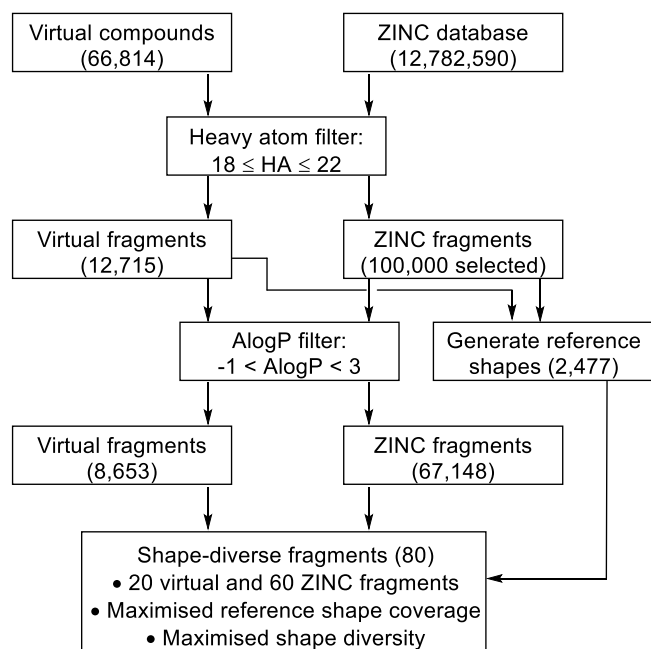


Figure 1. Overview of approach for the selection of eighty shape-diverse fragments. A virtual library of 66 814 compounds was enumerated by decoration of tangible scaffolds (see Figure 2). The resulting virtual compounds and the ZINC database of commercially-available compounds were filtered for molecular size, and a set of reference shapes was generated. After filtration for lipophilicity, a set of fragments was selected to maximise shape coverage and diversity and to include both virtual (20) and commercially-available (60) compounds.

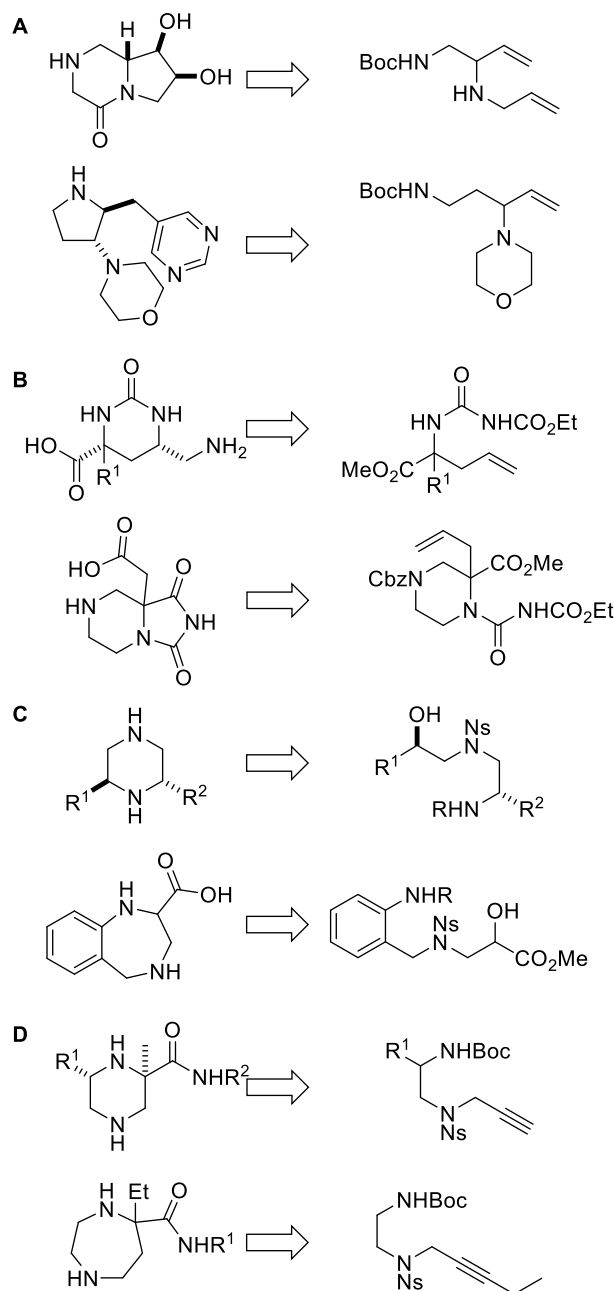


Figure 2. Exemplar envisaged syntheses of scaffolds from cyclisation precursors. Reactions between pairs of building blocks would yield cyclisation precursors which would be cyclised to yield molecular scaffolds. Underpinning synthetic approaches: Ir-catalysed allylic amination and cyclisation (Panel A); cyclisation of α -allyl α -amino ester derivatives (Panel B); cyclic sulfamidate ring-opening and intramolecular Mitsunobu reaction (Panel C); and cyclic sulfamidate ring-opening, Au-catalysed reaction and Ugi reaction (Panel D).

The shapes of the fragments were then described in terms of the set of reference shapes.^[16] After maximising the volume of overlap using ROCS, fragments were considered to be similar to each reference shape if the shape Tanimoto was greater than 0.7. The shape of each fragment could then be captured in a 2,477-bit binary fingerprint in which individual bits were set to 1 for the similar reference shapes.

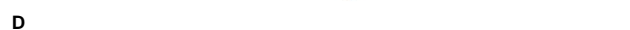
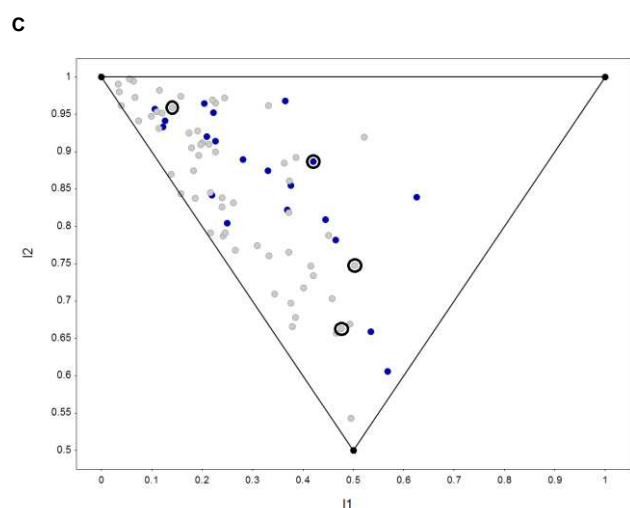
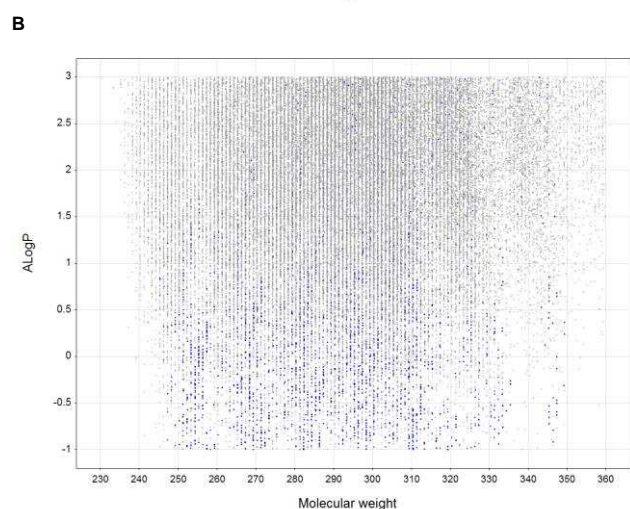
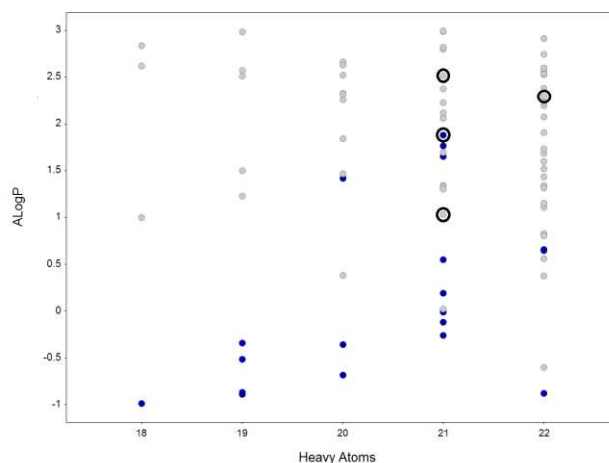
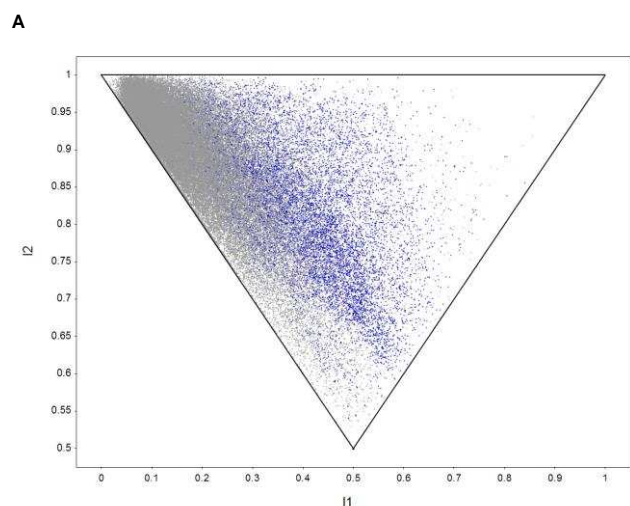


Figure 3. Shape diversity and molecular properties of fragment sets. Panels A and B: PMI plot (Panel A) and molecular properties (Panel B) of the set of 8,653 virtual fragments (blue) and 67,148 ZINC fragments (grey) that had been filtered for molecular size ($18 \leq \text{HA} \leq 22$) and lipophilicity ($-1 < \text{AlogP} < 3$). Panels C and D: PMI plot (Panel C) and molecular properties (Panel D) of the designed fragment set of 20 synthesised (blue) and the 60 purchased (grey) fragments. Fragment hits that were identified by high-throughput protein crystallography are highlighted (black circle).

A set of eighty shape-diverse fragments was then selected from the combined fragment set (Figure 3, Panels A and B) that had additionally been filtered for lipophilicity ($-1 < \text{AlogP} < 3$). The objectives were: (1) to select fragments from both sets (with, ideally, 20 fragments to be synthesised, and 60 to be purchased); (2) to maximise the number of reference shapes covered (i.e. that were similar to at least one selected fragment); and (3) to maximise the shape diversity of the selected fragments (by pairwise comparison of their shape fingerprints^[16]). A simulated annealing protocol^[20] was used to select the fragment set. In brief, eighty fragments were initially chosen at random, and the fit to the three objectives was captured in terms of a desirability score. Some of the fragments within the chosen set were then replaced at random – initially up to 30% of the fragments, but fewer fragments at later stages – and a new desirability score calculated. If a better desirability score was obtained, then the revised set became the new chosen set of fragments; otherwise, to enable local minima to be escaped, the probability of the new set being chosen was determined by a switch probability. The process was completed until a fragment set with an optimised desirability score was obtained.

Acquisition of a shape-diverse fragment set

During the course of our investigation, it became apparent that a few of the selected fragments would be difficult to obtain. First, we encountered some synthetic problems that were difficult to solve (see below); and second, some of the ZINC compounds were difficult to obtain in practice from commercial suppliers. At three points, therefore, the designed fragment set was revised: fragments that had already been obtained were retained, but the other fragments were revised to optimise the desirability score of

the new set. This pragmatic approach enabled us to secure a shape-diverse fragment set that met our three objectives.

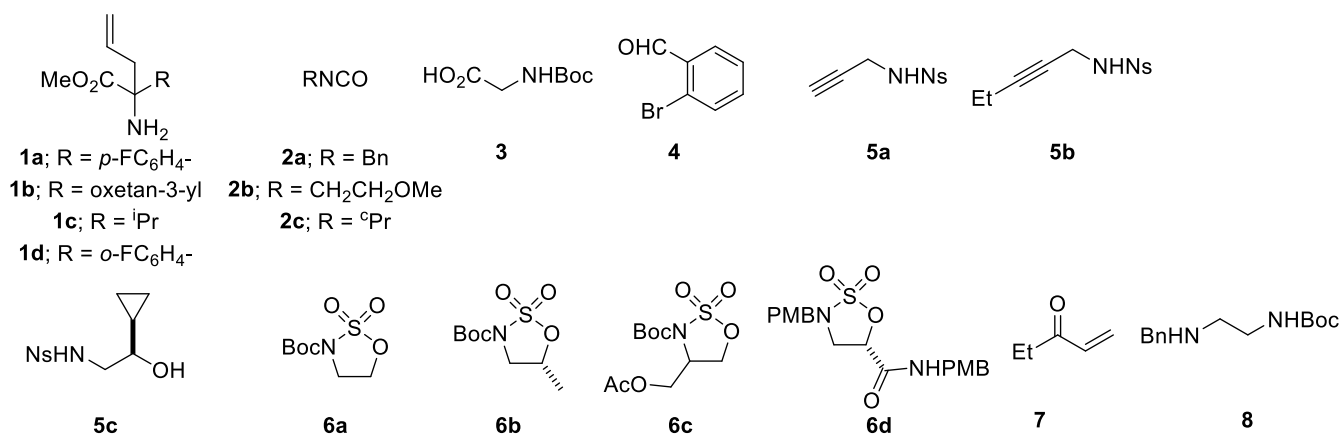
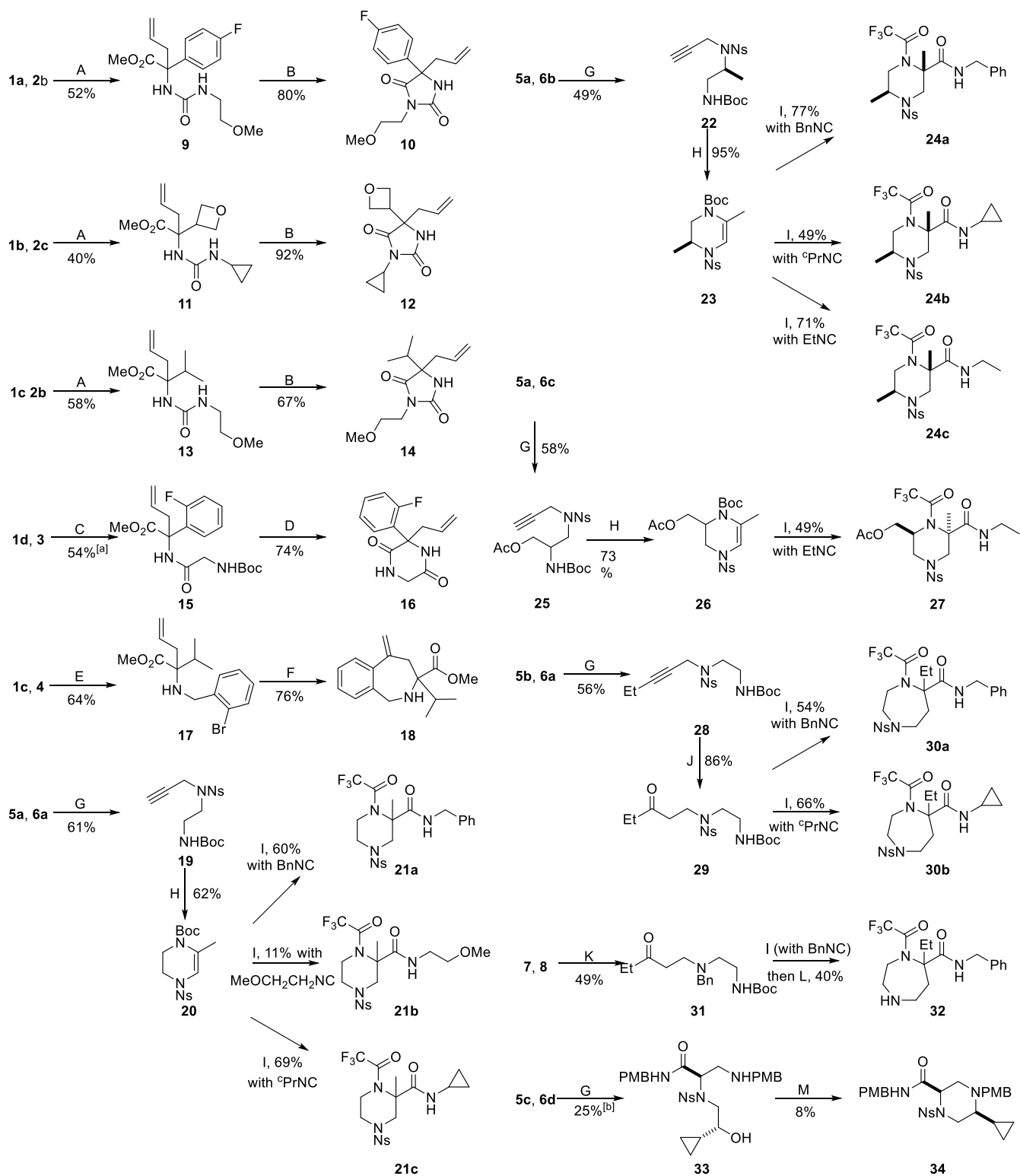


Figure 4. Structures of the building blocks used.



Scheme 1. Synthesis of scaffolds from pairs of building blocks. Methods: A: Et₃N, CH₂Cl₂; B: NaOtBu, toluene, 100 °C; C: TBTU, Et₃N, CH₂Cl₂; D: (a) TFA, CH₂Cl₂; (b) CsCO₃, DMF, 120-160 °C; E: NaBH(OAc)₃, THF, 45 °C; F: 10 mol% Pd(PPh₃)₄, Et₃N, MeCN, μw, 125 °C; G: NaH, DMF; H: 5 mol% Au(PPh₃)Cl, 5 mol% AgSbF₆, dioxane, 100 °C; I: (a) TFA, CH₂Cl₂; (b) RNC, EtOH; J: 1 mol% Au(IPr)Cl, 1 mol% AgSbF₆, dioxane-H₂O, Δ; K: EtOH; L: NH₄HCO₂, Pd(OH)₂/C, EtOH, 70 °C; M: PPh₃, DEAD, THF. [a] Yield over 2 steps from the corresponding α-amino α-aryl acid. [b] The diastereomeric product was also isolated in 19% yield. IPr, 1,3-bis(2,6-diisopropylphenyl)imidazol-2-ylidene.

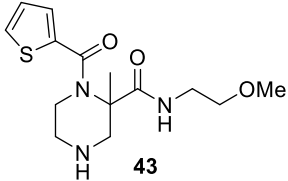
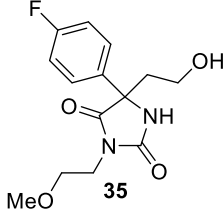
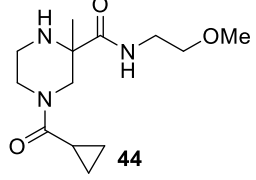
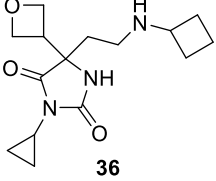
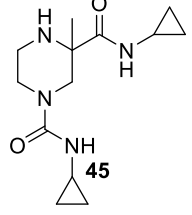
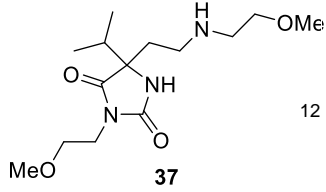
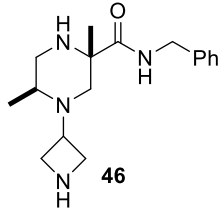
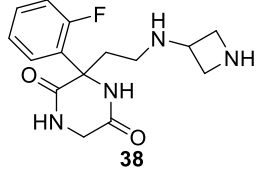
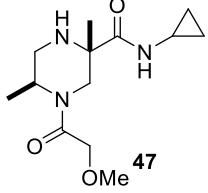
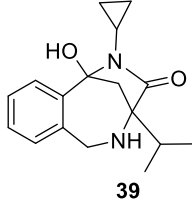
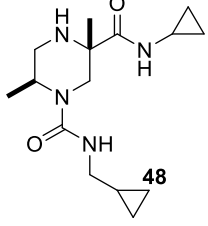
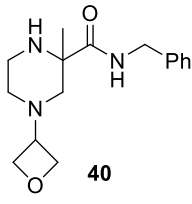
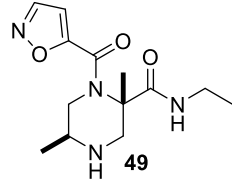
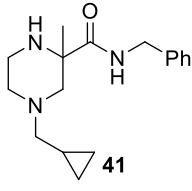
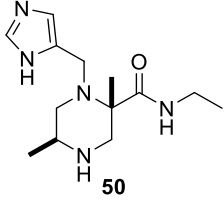
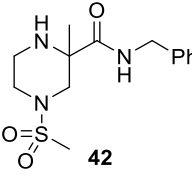
The required scaffolds were prepared from pairs of building blocks (Figure 4 and Scheme 1) in two stages: the pairs of building blocks were initially combined to yield a cyclisation precursor which was then cyclised (often via an intermediate) to yield a scaffold.

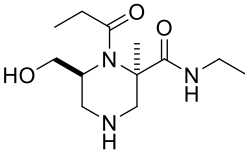
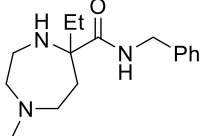
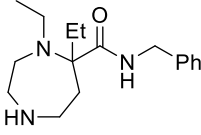
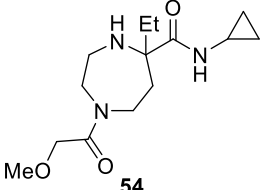
Five scaffolds were prepared by cyclisation of α -allyl α -amino ester precursors (Scheme 1).^[11] The ureas **9**, **11** and **13** were prepared by reaction between an α -allyl α -amino ester **1** and an isocyanate **2**; base-mediated cyclisation then gave the hydantoins **10**, **12** and **14**. Alternatively, acylation of the α -allyl α -amino ester **1d** gave the amide **15** which was Boc-protected and cyclised to give the diketopiperazine **16**. Finally, intramolecular Heck reaction of the *o*-bromobenzyl-substituted α -amino ester **17**, prepared by reductive amination, gave the benzo-fused azepane **18**.

Ten further scaffolds were prepared using Ugi reactions (Scheme 1).^[13] The required cyclisation precursors (**19**, **22**, **25** and **28**) were generally prepared by ring-opening of a cyclic sulfamidate **6** with a propargylic sulfonamide **5**. The outcome of the reaction of these cyclisation precursors with 5 mol% Au(PPh₃)Cl and 5 mol% AgSbF₆ in dioxane at 100 °C depended on substitution.^[13a] Thus, with the terminal alkynes, **19**, **22** and **25**, the Boc-protected amine cyclised onto proximal alkyne carbon to yield the corresponding tetrahydropyrazines **20**, **23** and **26**. In contrast, the 1,2-disubstituted alkyne **28** underwent hydration at the remote alkyne carbon to yield the open-chain β -keto amine derivative **29**. The related β -keto amine derivative **31** was also prepared by conjugate addition of the amine **8** to the enone **7**.^[13b] After Boc protection, the tetrahydropyrazines

20, **23** and **26** and the β -keto amine derivatives **29** and **31** were all competent substrates in Ugi reactions: subsequent reaction with an isocyanide in ethanol to yielded either substituted 2-piperazinyl carboxamides (**21a-c**, **24a-c** and **27**) or 2-(1,4-diazepanyl) carboxamides (**30a-b** or **32**). Notably, with the chiral tetrahydropyrazines **23** and **26**, the corresponding Ugi reaction products were obtained, like related examples,^[13b] with high diastereoselectivity.

The synthesis of the piperazine **34** (Scheme 1) emphasised some of the shortcomings of a lead-oriented synthesis of diverse saturated nitrogen heterocycles.^[12] First, treatment of the sulfonamide **5c** with sodium hydride, and reaction with the cyclic sulfamidate **6d**, was poorly diastereoselective, presumably stemming from epimerisation by enolisation: the amino alcohol **33** was nonetheless obtained in 25% yield. However, the intramolecular Mitsunobu reaction of **33** was extremely poor, and the piperazine **34** was obtained in just 8% yield. Related syntheses of similarly substituted piperazines have also been disappointing.^[12] This lead-oriented approach was not, therefore, ultimately exploited in fragment syntheses. Finally, Ir-catalysed amination^[10] was also not ultimately exploited. The required Ir-catalysed substitutions of allylic carbonates with amine building blocks were either low yielding or were poorly regioselective. For example, with azetidine as nucleophile, the unwanted terminal regioisomer was obtained predominantly. In these cases, to ensure that a shape-diverse fragment set was assembled, the targeted fragments were revised (see above).

Table 1. Synthesis of fragments.			9	21b	J (41%), H (91%)		
Entry	Scaffold	Method ^[a] (Yield) ^[l]	Fragment				
1	10	A (49%), B (68%)		10	21b	H,J,B (43%)	
2	12	A (7%), C (87%)		11	21c	H,J (16%), B (61%)	
3	14	A (71%), C (>98%)		12	24a	H,C (62%), B (>98%), E (84%)	
4	16	D,C (23%), E (60%)		13	24b	H,I,B (19%)	
5	18	F (46%), D (49%) ^[b]		14	24b	H,J,B (26%)	
6	21a	H,C (98%), B (78%)		15	24c	B (52%), J (54%), H (18%)	
7	21a	H,C,B (24%)		16	24c	B (52%), C (14%), H (59%)	
8	21a	H,I,B (18%)					

17	27	B (53%), J (51%), H (89%)	
18	32	K (36%), B (82%)	
19	30a	B (34%), C (64%), H (68%)	
20	30b	H (98%), J (78%), B (27%)	

[a] Typical methods (see Supporting Information): A: O₃, CH₂Cl₂, -40 or -78 °C; B: NaBH₄, MeOH; C: amine, aldehyde or ketone, NaBH(OAc)₃; D: (a) NMO, 1 mol% K₂OsO₄•H₂O; (b) NaIO₄, MeOH-H₂O; E: TFA, CH₂Cl₂; F: (a) NaOH, DMF, 120 °C; (b) RNH₂, TBTU, Et₃N, DMF; G: NaBH(OAc)₃; H: PhSH, K₂CO₃, MeCN; I: RSO₂Cl, CH₂Cl₂; J: acid chloride or isocyanate, Et₃N, DMAP, CH₂Cl₂; K: formaldehyde, Zn, AcOH, dioxane-H₂O. [b] The product existed predominantly (>98%) in the ring-closed form.

The synthesis of the 20 fragments that contributed to the final shape-diverse fragment set is summarised in Table 1. In most cases, it was initially necessary to unmask functionality in the starting scaffold, for example by deprotection or by oxidative alkene cleavage. Four different reactions were exploited in decoration steps: reductive amination, sulfonylation, acylation, and urea formation. The synthesis of 13 of the fragments was influenced by the propensity of trifluoroacetylated Ugi products

to undergo base-catalysed hydantoin formation:^[13b] to minimise this undesired reaction pathway, trifluoroacetylated α -amino esters were deprotected by treatment with sodium borohydride in methanol.

The shapes and molecular properties of the compounds in the final fragment set are presented in Figure 3 (Panels C and D). The fragment set is significantly more three-dimensional than typical commercial fragment sets,^[5a] with only two of the eighty fragments lying close to the rod-disk axis of the PMI plot (Panel C). Fragment sets with distinctive three dimensionality have previously also already been designed through judicious selection of commercially-available compounds.^[5]

The shape diversity of our fragment set is significantly enriched by the 20 synthesised fragments which are more three-dimensional, and are thus complementary to, the 60 purchased fragments (Figure 3, Panel C). We note that, whilst the 20 synthesised and 60 purchased fragments have a similar mean number of rings ($\mu = 2.4$ and 2.2 respectively), the synthesised fragments generally have fewer aromatic rings ($\mu = 0.6$ compared with $\mu = 1.5$ for the purchased compounds). The synthesised fragments are generally based on saturated nitrogen heterocycles that bear at least two carbon-based substituents. We note that all of the synthesised fragments have at least one stereogenic centre ($\mu = 1.3$ compared with $\mu = 0.6$ for the purchased fragments), and many bear small (3- or 4-membered) cyclic substituents ($\mu = 0.8$ compared with $\mu = 0.1$ for the purchased compounds). These features were presumably important to complement the shape diversity of commercially available fragments.

Screen of the designed fragment set against Aurora-A kinase

The shape-diverse fragment set was screened against Aurora-A kinase by high-throughput protein X-ray crystallography. Aurora-A kinase regulates mitotic spindle assembly by reversible phosphorylation of microtubule-associated proteins, and is a promising target for anticancer therapy.^[21] Protein crystals were soaked with the 80 individual fragments,^[22] picked and then subjected to automated X-ray diffraction. Fragment hits were identified through detection of additional electron density,^[23] and inspection of the polar interactions with the protein.

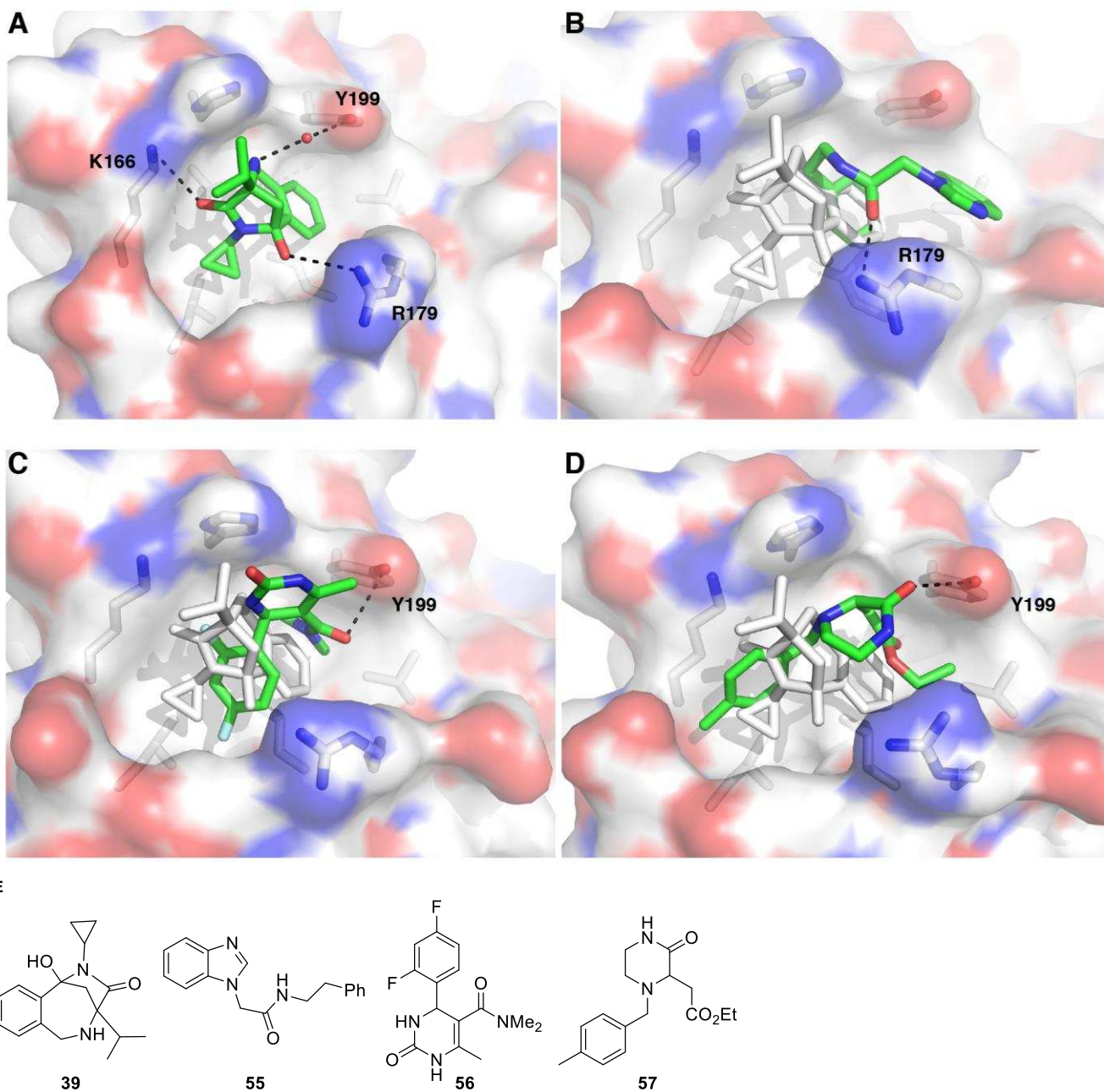


Figure 5. Hits from fragment screen against Aurora-A kinase. Panels A: Direct and through-water polar interactions (black) between protein residues and the fragment hit **39**. Panels B-D: Polar interactions (black) between protein residues and fragment hits (**55**, Panel B; **56**, Panel C; **57**, Panel D) in the context of the fragment hit **39** (grey). Panel E: Structures of the fragment hits.

The fragment screen revealed four hits: the synthesised fragment **39**, and the purchased fragments **55-57**. The representative conformers of these four hits are very shape-diverse (Figure 3, Panel C). The fragments all bound in a pocket that is critical for the interaction with TPX2, a microtubule-associated protein that activates Aurora-A (Figure 5).^[24] This pocket, which is equivalent to the PIF pocket of PDK1, has been targeted by allosteric inhibitors of Aurora-A: the small molecule AurkinA, and a single domain antibody VNAR-D03.^[25] The pocket also dominated the hits obtained from a crystallography-based screen of a commercial fragment library.^[26] The synthesised fragment hit **39** made three polar interactions: direct interactions with K166 and R179 and a unique water-mediated interaction with Y199 (Panel A). In contrast, the three purchased fragments each made one direct interaction with the protein: with R179 (hit **55**) or Y199 (hits **56** and **57**). Notably, the phenyl ring of **55** and the benzo-fused ring of **39** had a common binding mode that suggests how the fragments might subsequently be merged. We determined the dose-dependent inhibition of Aurora-A kinase by the fragment **39** (IC₅₀: ~8 mM) (Supporting Information).

To investigate **39** as a starting point for ligand discovery, we prepared and evaluated by protein crystallography some analogues in which the benzo-fused ring, the cyclopropyl substituent and/or the iso-propyl substituent had been modified (Supporting Information). The fragment hit **39** was also identified by throughput X-ray crystallography (see above for details of the screen of the original 80 shape-diverse fragments). Gratifyingly, the binding mode of the exemplar analogue **58** was extremely similar to that of the original fragment hit **39** (Figure 6).

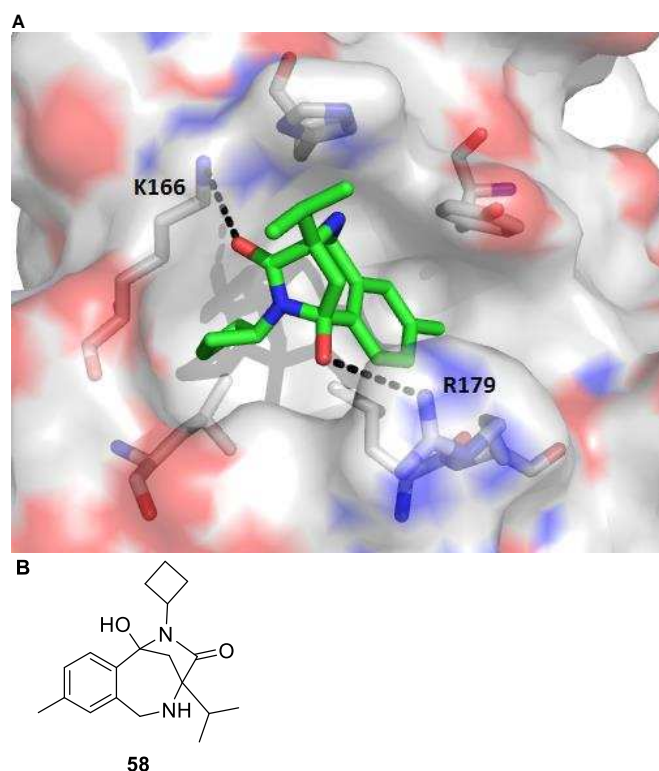


Figure 6. Follow-up to the fragment hit **39**. Panel A: Interaction between the analogue **58** and Aurora-A kinase. Panel B: Structure of the exemplar analogue **58** that also interacted with Aurora-A kinase.

Conclusions

We have designed and assembled a set of 80 shape-diverse fragments. The fragment set was highly distinctive from most existing sets: in particular, it was highly three-dimensional, with just two fragments close to the rod-disk axis of a PMI plot. The shape diversity of the set was enriched by the 20 synthesised fragments that were, on average, significantly more three-dimensional than the purchased fragments.

Our fragment set will be a useful resource for assessing experimentally the added value of high shape diversity. Indeed, we specifically chose to assemble a fragment set (rather than, for example, a set of shape-diverse lead-like compounds) to maximise the efficiency of coverage of the possible shape space. We have already performed a fragment screen against Aurora-A kinase by high-throughput protein crystallography: the screen revealed four hits that targeted the binding site of allosteric regulators of the enzyme. The shape-diverse fragments are available to the scientific community for screening at XChem, Diamond's high-throughput crystallography facility. The added value of the fragments will only be captured following the completion of fragment screens against a broad range of functionally-diverse proteins.

Acknowledgements

We thank EPSRC (EP/N025652/1, DTP studentship and Doctoral Prize Fellowship), MRC CASE industrial studentship (MR/K016903) and LifeArc (formerly MRC Technology) for funding. We thank OpenEye for an academic software license, and Diamond Light Source for access to beamline i04-1 (proposal Ib14331) that contributed to the results presented here.

Keywords: molecular shape • fragments • diversity-oriented synthesis

- [1] A. H. Lipkus, Q. Yuan, K. A. Lucas, S. A. Funk, W. F. Bartelt, R. J. Schenck, A. J. Trippe, *J. Org. Chem.* **2008**, 73, 4443-4451.
- [2] (a) D. G. Brown, J. Boström, *J. Med. Chem.* **2016**, 59, 4443-4458; (b) T. W. J. Cooper, I. B. Campbell, S. J. F. Macdonald, *Angew. Chem. Int. Ed.* **2010**, 49, 8082-8091.
- [3] W. P. Walters, J. Green, J. R. Weiss, M. A. Murcko, *J. Med. Chem.* **2011**, 54, 6405-6416.
- [4] F. Lovering, J. Bikker, C. Humblet, *J. Med. Chem.* **2009**, 52, 6752-6756.
- [5] (a) A. D. Morley, A. Pugliese, K. Birchall, J. Bower, P. Brennan, N. Brown, T. Chapman, M. Drysdale, I. H. Gilbert, S. Hoelder, A. Jordon, S. V. Ley, A. Merritt, D. Miller, M.E. Swarbrick, P. G. Wyatt, *Drug Discov. Today* **2013**, 18, 1221-1227; (b) C. W. Murray, D. C. Rees, *Angew. Chem. Int. Ed.* **2016**, 55, 488-492; (c) B. Over, S. Wetzels, C. Grütter, Y. Nakai, S. Renner, D. Rauh, H. Waldmann, *Nat. Chem.* **2013**, 5, 21-28; (d) H. Prescher, G. Koch, T. Schuhmann, P. Ertl, A. Bussenault, M. Glick, I. Dix, F. Petersen, D. E. Lizos, *Bioorg. Med. Chem.* **2017**, 25, 921-925; (e) D. G. Twigg, N. Kondo, S. L. Mitchell, W. R. J. D. Galloway, H. F. Sore, A. Madin, D. R. Spring, *Angew. Chem. Int. Ed.* **2016**, 55, 12479-12483; (f) A. W. Hung, A. Ramek, Y. Wang, T. Kaya, J. A. Wilson, P. A. Clemons, D. W. Young, *Proc. Natl. Acad. Sci. USA* **2011**, 108, 6799-6804; (g) O. A. Davis, R. A. Croft, *J. A. Bull. Chem. Commun.* **2015**, 51, 15446-15449; (h) H. Hassan, S. P. Marsden, A. Nelson, *Bioorg. Med. Chem.* **2018**, 26, 3030-3033.
- [6] (a) L. Ruddigkeit, R. van Deursen, L. C. Blum, J. L. Reymond, *J. Chem. Inf. Model.* **2012**, 52, 2864-2875; R. Visini, M. Awale, J.-L. Reymond, *J. Chem. Inf. Model.* **2017**, 57, 700-709.
- [7] A. R. Leach, M. M. Hann, *Curr. Opin. Chem. Biol.* **2011**, 15, 489-496.

-
- [8] D. C. Fry, C. Wartchow, B. Graves, C. Janson, C. Lukacs, U. Kammlott, C. Belunis, S. Palme, C. Klein, B. Vu, *ACS Med. Chem. Lett.* **2013**, *4*, 660-665.
- [9] R. J. Hall, P. N. Mortenson, C. W. Murray, *Prog. Biophys. Mol. Biol.* **2014**, *116*, 82-91.
- [10] R. G. Doveston, P. Tosatti, M. Dow, D. J. Foley, H. Y. Li, A. J. Campbell, D. House, I. Churcher, S. P. Marsden, A. Nelson, *Org. Biomol. Chem.* **2015**, *13*, 859-865.
- [11] D. J. Foley, R. G. Doveston, I. Churcher, A. Nelson, S. P. Marsden, *Chem. Commun.* **2015**, *51*, 11174-11177.
- [12] T. James, P. MacLellan, G. M. Burslem, I. Simpson, J. A. Grant, S. Warriner, V. Sridharan, A. Nelson, *Org. Biomol. Chem.* **2014**, *12*, 2584-2591.
- [13] (a) T. James, I. Simpson, J. A. Grant, V. Sridharan, A. Nelson, *Org. Lett.* **2013**, *15*, 6094-6097; J. D. Firth, R. Zhang, R. Morgentin, R. Guilleux, T. Kalliokoski, S. Warriner, R. Foster, S. P. Marsden, A. Nelson, *Synthesis* **2015**, *47*, 2391-2406.
- [14] A. Nadin, C. Hattotuagama, I. Churcher, *Angew. Chem. Int. Ed.* **2012**, *51*, 1114-1122.
- [15] J. J. Irwin, T. Stirling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, *J. Chem. Inf. Model.* **2012**, *52*, 1757-1768.
- [16] J. A. Haigh, B. T. Pickup, J. A. Grant, A. Nicholls, *J. Chem. Inf. Model.* **2005**, *45*, 673-684.
- [17] It has previously been shown that the number of reference shapes is almost independent of whether single or multiple conformers are used.
- [18] J. Sadowski, J. Gasteiger, G. Klebe, *J. Chem. Inf. Model.* **1994**, *34*, 1000-1008.
- [19] (a) J. A. Grant, M. A. Gallardo, B. Pickup, *J. Comp. Chem.* **1996**, *17*, 1653; (b) ROCS version 3.2.1. OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.
- [20] S. Kirkpatrick, C. D. Gelatt Jr, M. P. Vecchi, *Science* **1983**, *220*, 671-680.
- [21] H. T. Ma, R. Y. Poon, *Biochem. J.* **2011**, *435*, 17-31.
- [22] P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M. Pearce, F. von Delft, *Acta Cryst.* **2017**, *D73*, 246-255.
- [23] (a) N. M. Pearce, A. R. Bradley, T. Krojer, B. D. Marsden, C. M. Deane, F. von Delft, *Structural Dynamics* **2017**, *4*, 032104; (b) N. M. Pearce et al, *Nat. Commun.* **2017**, *8*, 15123.
- [24] (a) R. Bayliss, T. Sardon, I. Vernos, E. Conti, *Mol. Cell.* **2003**, *12*, 851-862; (b) R. Bayliss, S. G. Burgess, P. J. McIntyre, *FEBS J.* **2017**, doi: 10.1111/febs.14069.
- [25] (a) M. Janeček, M. Rossmann, P. Sharma, A. Emery, D. J. Huggins, S. R. Stockwell, J. E. Stokes, Y. S. Tan, E. G. Almeida, B. Hardwick, A. J. Narvaez, M. Hyvönen, D. R. Spring, G. J. McKenzie, A. R. Venkitaraman, *Sci. Rep.* **2016**, *6*, 28528; (b) S. G. Burgess, A. Oleksy, T. Cavazza, M. W. Richards, I. Vernos, D. Matthews, R. Bayliss, *Open Biol.* **2016**, *6*, 160089; (c) J. O. Schulze, G. Saladino, K. Busschots, S. Neimanis, M. N. Lisa, P. M. Alzari, F. L. Gervasio, R. M. Biondi, *Cell Chem. Biol.* **2016**, *23*, 1193-1205.
- [26] P. J. McIntyre, P. M. Collins, L. Vrzal, K. Birchall, L. H. Arnold, C. Mpamhanga, P. J. Coombs, S. G. Burgess, M. W. Richards, A. Winter, V. Veverka, F. von Delft, A. Merritt and R. Bayliss, *ACS Chem. Biol.* **2017**, *12*, 2906-2914.
-