Query-guided protein–protein interaction inhibitor discovery†

Sergio Celis, a d‡ Fruszina Hobor, a b c‡ Thomas James, a b c‡ Gail J. Bartlett, a d‡
Amaury A. Ibarra, e‡ Deborah K. Shoemark, e f Zsófia Hegedüs, a b c‡
Kristina Hetherington, a b Derek N. Woolfson, b d‡ Richard B. Sessions, b e f
Thomas A. Edwards, b a c‡ David M. Andrews, b g Adam Nelson b * a b
and Andrew J. Wilson b * a b

Protein–protein interactions (PPIs) are central to biological mechanisms, and can serve as compelling targets for drug discovery. Yet, the discovery of small molecule inhibitors of PPIs remains challenging given the large and typically shallow topography of the interacting protein surfaces. Here, we describe a general approach to the discovery of orthosteric PPI inhibitors that mimic specific secondary protein structures. Initially, hot residues at protein–protein interfaces are identified in silico or from experimental data, and incorporated into secondary structure-based queries. Virtual libraries of small molecules are then shape-matched against the queries, and promising ligands docked to target proteins. The approach is exemplified experimentally using two unrelated PPIs that are mediated by an α-helix (p53/hDM2) and a β-strand (GRAP/SHANK1-PDZ). In each case, selective PPI inhibitors are discovered with low μM activity as determined by a combination of fluorescence anisotropy and 1H–15N HSQC experiments. In addition, hit expansion yields a series of PPI inhibitors with defined structure–activity relationships. It is envisaged that the generality of the approach will enable discovery of inhibitors of a wide range of unrelated secondary structure-mediated PPIs.

Introduction

The discovery of small-molecule modulators of protein–protein interactions (PPIs) is a central challenge in both chemical biology and medicinal chemistry.1–4 For a number of PPI targets, potent PPI inhibitors and stabilizers have now been successfully discovered, for example by optimisation of high-throughput screening hits, fragment-based discovery approaches or virtual methods targeting p53/hDM2, 5–10 the BCL-2 family,11–15 and other interactions.16–21 However, the paucity of small-molecule PPI inhibitors that have progressed as clinical candidates,2 would be of significant importance for designing orthosteric peptidomimetic inhibitors.22–25 Although PPIs are known to involve shallow, relatively large interfaces of varied topography, the identification of hot-spots26–28 – i.e., amino-acid residues that contribute significantly to binding – can provide focus for ligand design efforts. Moreover, a significant proportion of the protein–protein interactome involves short linear peptide sequences that adopt secondary structures, providing promising templates for design of orthosteric peptidomimetic inhibitors.25,26

Design approaches can be categorized into four groups.2 Class A mimetics are peptides with a limited number of modified amino acids introduced to optimize properties (e.g. stapled peptides). Class B mimetics are peptidic in nature but with more dramatic changes to structure; they include foldamers and mimic secondary structure topology. Class C mimetics are topographical mimics and are more small-molecule like; a core scaffold projects groups to mimic the orientation and composition of hot-spot residues. Class D mimetics are functional mimics and do not necessarily have a connection to structure.25–29 Class C mimetics – also termed proteomimetics – have successfully been developed to inhibit a range of α-helix-mediated PPIs, in some cases with selectivity and cellular activity.25–29 However, the proteomimetic approach has not been...

© 2021 The Author(s). Published by the Royal Society of Chemistry

Chem. Sci.
results in the development of small-molecule proteomimetic inhibitors of PPIs which small-molecules can potentially mimic any secondary structure. The discovery workflow is implemented through several stages that combine established computational tools and experimental validation. Initially a query is built that incorporates the key secondary structural motif and hot residues from the PPI. A virtual library of small molecules is then shape-matched against the query, and promising compounds docked against the target protein. Candidate inhibitors are then subjected to experimental screening and characterisation, enabling selection of hits for: (i) clustering, near neighbour expansion and further screening by reiteration of step II and III of the workflow; or (ii) further development.

Here, we exemplify a general approach for the discovery of small-molecule proteomimetic inhibitors of PPIs (Fig. 1) which we term: Query-Guided PPI Inhibitor Discovery. Hot-spot residues from a protein–protein interface are built into secondary structure-based queries. Assessment of the shape similarity of ligands to the query, in this case using FastROCS, enables small-molecule prioritisation. To provide additional confidence prior to experimental studies, further docking is performed. We demonstrate the power of the approach using two unrelated exemplar PPI targets: (i) p53/hDM2, an α-helix-mediated PPI that is a clinically relevant oncology target; and (ii) GKAP/SHANK1-PDZ, a β-strand-mediated PPI which plays a key role at the synaptic junction and is representative of PDZ-mediated interactions, which remain challenging for small-molecule inhibitor discovery. For both targets, we used this approach to virtually screen ~4 million compounds for which physical samples were available in AstraZeneca’s compound collection and then validated the approach experimentally using biophysical screens. Hit-expansion for the p53/hDM2 inhibitor identified a series of PPI inhibitors with defined structure–activity relationships demonstrating the approach identifies developable compounds.

Results and discussion

Discovery of small-molecule p53/hDM2 inhibitors

We used the p53/hDM2 interaction as an exemplar PPI to benchmark our discovery approach. The p53/hDM2 interaction involves thirteen residues of the disordered N-terminal transactivation domain of p53 (p5317–29 henceforth referred to as p53), which folds into a helix and docks into a hydrophobic cleft on hDM2 (Fig. 2a). Three key side chains – Phe19, Trp23 and Leu26 – located on one face of the helix (at the i, i + 4 and i + 7 positions), have been identified as making a dominant contribution to the binding free energy of association (Fig. 2a and b). These hot residues were used to construct queries for structure similarity searching. First, the p53 peptide was excised from a structure of p53/hDM2 (PDB: 1YCR), and a query was generated that contained the three hot-spot residues linked by a hydrocarbon backbone (Fig. 2c). A second query was generated in which a peptide backbone was retained (Fig. 2d).

A virtual library of 42 million conformers based on the AstraZeneca screening collection was shape-matched against both queries using FastROCS. For each query, the top 1000 hits were clustered, based on the sum of the scoring functions for 3D shape match and surface complementarity, were selected. These conformers were then docked rigidly to hDM2 (OEDocking; ©2019, OpenEye Scientific Software, Inc.), and scored with a hybrid function that captured both shape similarity to the bound p53 peptide and shape complementarity to the hDM2 binding site (Fig. 1, stage II, see Fig. S1† for more detailed workflow including clustering near neighbour expansion and further screening which can be reiterated as appropriate, and, ESI† for methodology). Next, the virtual hits were triaged ahead of experimental evaluation (Fig. 1, stage III). First, we removed fragment-like hits, a decision informed by observations made in the deconstruction of the known inhibitor RG7112: this retrospective analysis suggested a fragment-based discovery approach would only have been possible by starting with fragments that retained at least two hot-spot binding groups. Second, a significant number of compounds were removed because they had flipped during docking (i.e., the docked pose no longer adopted a p53 mimicking orientation), and therefore did not fit with the design hypothesis. Third, the remaining hits were clustered (Tanimoto similarity >0.7), and one or two representatives were retained from each cluster (see ESI†); clusters with singleton hits were only retained if there was at least one near-neighbour in the main AstraZeneca screening hits.
For each query, ~100 compounds were prioritised for experimental evaluation, including representatives of the top 30 clusters that were predicted to interact with at least two hot residue binding pockets. In addition, 100 randomly selected compounds were also selected for analyses (see ESI†).

The ability of compounds to displace a fluorescently labelled p53 peptide (p53_{15-31}Flu (L31C), Flu = fluorescein-5-maleimide) was determined using a competition fluorescence anisotropy assay (see ESI Fig. S2† for assay and Fig. S3–S6† for % inhibition at 10 and 100 μM). For compounds that displayed significant activity at 100 μM, structurally similar near-neighbours were selected from the main AstraZeneca screening collection for hit expansion. For the query with a hydrocarbon backbone, 21 hits from 15 clusters were

---

**Fig. 2** Discovery of small-molecule inhibitors of the p53/hDM2 interaction: (a) close-up of the p53/hDM2 interaction structure (PDB ID: 1YCR), p53 (cyan), with key side chains Phe19, Trp23 and Leu26 highlighted, docks into the hDM2 cleft with the hDM2 surface defined by the Phe19 (violet), Trp23 (magenta) and Leu 26 (green) pockets, into which each corresponding hydrophobic amino acid projects; (b) ribbon representation of the p53_{17-29} transactivation domain with key side chains Phe19, Trp23 and Leu 26 highlighted together with the primary sequence (below); (c) query that incorporates the hot residues and a "hydrocarbon" backbone; (d) query that incorporates the hot residues and a peptidic backbone. (e) Overlay of exemplar shape-matched hit compounds and queries; (f) structures of the exemplar hit compounds as representatives of the most populated clusters; (g) ligand efficiency plot for the 37 hits obtained from the computational workflow (LE = 1.4 x IC_{50}/nHA, nHA: number of non-hydrogen atoms, IC_{50} obtained by fluorescence anisotropy competition: 150 nM hDM2 and p53_{15-31}Flu, 40 mM phosphate, pH 7.4, 200 mM NaCl and 0.02 mg ml^{-1} BSA); (h) mapping of the chemical shift perturbations in hDM2 for hit compounds from cluster A, cluster B and cluster C. Colour variation is associated with a chemical shift perturbation that goes from 0 ppm (white) to 0.2 ppm (red).
Fig. 3 Development of a GKAP query for inhibitor screening against the GKAP/SHANK1-PDZ interaction; (a) close-up of the GKAP/SHANK1 PDZ interaction structure (PDB ID: 1Q3P),

gKAP (cyan), with key side chains Thr3 and Leu5 highlighted, interacts with SHANK1 PDZ domain mainly through polar contacts (orange area); however, hydrophobic effects play a significant role in binding (magenta area); (b) cartoon representation of the GKAP1–5 peptide with key side chains Thr3 and Leu5 highlighted together with the primary sequence (below); (c) X-ray crystal structure of Ac-Glu-Ala-Gln-Thr-Arg-Phe peptide (L6F) bound to SHANK1-PDZ (PDB ID: 7A00) illustrating good correspondence with the position of key recognition groups observed for the wild-type sequence (PDB ID: 1Q3P); (d) close up of the interactions between the C-terminus of the L6F GKAP peptide and SHANK1 PDZ (H-bonds orange, other contacts yellow dashed lines); (e) close up of the interactions between the C-terminus of the wt GKAP peptide and SHANK1 PDZ; (f) binding of wild type (left) and L6F (right) GKAP peptide to SHANK1, monitored by ITC [25 °C in 20 mM Tris, 150 mM NaCl, pH 7.5, heats of peptide dilution were subtracted from each measurement raw data] with data analysed using Microcal Origin 8 and fitted to a one-binding site model; (g) GKAP query whereby the Thr3 side chain is retained together with a Phe in lieu of a Leu side chain alongside key backbone donor (NH) and acceptor (CO) groups; (h) structure of the hit compounds Z-1 and Z-2 identified from single point screening work; (i) overlay of the compound Z-1 and the query (grey); (j) fluorescence anisotropy competition assay for compound Z-1 (FITC-GKAP 50 nM, SHANK1–PDZ 1 μM, pH 7.4, 20 mM Tris, 150 mM NaCl, 0.01% Triton-X-100 buffer); (k) expansion of the 1H–15N HSQC spectra of 15N labelled SHANK1 in the absence (red) and presence of compound Z-1 (compound : protein molar ratio 1:1 yellow, 2:1 green, 4:1 blue). Peaks indicated with asterisk undergo significant changes in intensity upon binding (SHANK1-PDZ 50 μM, pH 7.4, 5 mM Tris, 100 mM NaCl buffer).
identified, and were complemented with 100 near-neighbours. For the query with the peptide backbone, 3 hits from 3 clusters were identified, and were complemented with 11 near-neighbours. These 135 compounds were evaluated at 10 μM and 100 μM in the competition fluorescence anisotropy assay (not shown). Subsequently, based on visual inspection and team discussion of how well they matched the initial FastROCS shape-match and docking, alongside medicinal chemistry and synthetic assessment, 63 compounds from 9 clusters (e.g. Fig. 2e and f) were selected for IC_{50} determination (Fig. S6†). A number of these compounds had mediocre inhibitory potency that fell outside the assay window leaving 37 compounds for which IC_{50} values could be determined, allowing an assessment of their ligand efficiency (Fig. 2g), i.e. the binding energy per non-hydrogen atom of a ligand to its binding partner which is commonly used in drug discovery programmes to assist in identifying compounds with optimal combinations of properties.52-53

Mapping of chemical shift perturbations by 1H-15N HSQC for representatives from each promising cluster confirmed a bona fide interaction with hDM2 (Fig. 2h and S7–S9†). We used ligand efficiency scores and chemical accessibility to eventually select compound A1 (Fig. 2) as a hit upon which to carry out structure activity relationship (SAR) analysis (see later and Fig. 4). At IC_{50} = 7.9 (±0.9) μM, hit compound A1 with MW = 444 is only one order of magnitude lower in inhibitory potency than Nutlin-3 (IC_{50} = 0.6 ± 0.04 μM; MW = 581, see ESI Fig. S8†). However, it has a comparable ligand efficiency of 0.24 (versus 0.22 for Nutlin-3) and significant potential for further derivatization and optimization. Compound A1 was also shown to induce shifts in the 1H-15N HSQC spectrum and when mapped onto the residue assignments, the affected resonances corresponded well to those affected upon binding of Nutlin-3, lending support to the design hypothesis. Compound A1 was also tested in a FA competition assay for BID/MCL-1 (Fig. S10†), a further α-helix mediated PPI: no inhibition was observed, indicating the compound A1 to be selective and confirming it as a good choice for further development.

Application of the discovery workflow to GKAP/SHANK1-PDZ

Subsequently, we applied the same workflow to the GKAP/SHANK1-PDZ interaction (Fig. 3). The GKAP PDZ-binding motif (Ac-Glu-Ala-Gln-Thr-Arg-Leu-CO2H), henceforth referred to as GKAP contains the known Type-I PDZ recognition motif:54 Thr-Xxx-Leu, with the C-terminal carboxylate making three essential hydrogen bonds with backbone amide hydrogen atoms in a loop on SHANK1 (Fig. 3a, orange area, PDB ID: 1Q3P).55 In our prior studies, computational modelling using BudeAlaScan55,56 and experimental analyses confirmed the Thr and Leu amino acids together with the terminal carboxylic acid as crucial for binding (Fig. 3b).55,56 Although SHANK1 has a defined hydrophobic cavity in which the Leu side chain of GKAP is accommodated (Fig. 3a, magenta area), it represents a significantly more challenging target for small-molecule mimicry than p53/hDM2. As a β-strand mediated PPI, the GKAP/SHANK1 interface relies on both side-chain contacts and an extensive hydrogen bond network that plays an active role in molecular recognition (Fig. 3a, d and e, H-bonds are shown as orange dashed lines); these features are not readily reproduced by typical (often hydrophobic) small molecule scaffolds. Thus, for query generation, the key hot residues of GKAP together with backbone heteroatoms were used. However, when our workflow was implemented with a query bearing these features, we identified only flexible peptide-based inhibitor candidates (not shown).

We reasoned that the C-terminal leucine, with more side-chain degrees of freedom, might have contributed to the moderate results in our in silico structure similarity analyses. The plasticity of PDZ domains allows the accommodation of various hydrophobic side chains at the C-terminus of the peptide; for SHANK1, Leu dominates for C-terminal carboxylates, however Phe has been observed to dominate for non-C-terminal sequences.57,58 Therefore, we considered Phe, with its more rigid hydrophobic amino acid sidechain, as a suitable surrogate in this position of the query. Gratifyingly, preparation of the Ac-Glu-Ala-Gln-Thr-Arg-Phe-CO2H peptide and assessment by isothermal titration calorimetry confirmed the Phe variant peptide to be tolerated (K_{d} = 5 (±0.5) μM compared with K_{d} = 2 (±0.3) μM for the GKAP sequence) and as an effective inhibitor (Fig. 3f and S11† for fluorescence competition anisotropy assay data using FAM-Ahx-Glu-Ala-Gln-Thr-Arg-Leu-CO2H as tracer). We obtained a peptide/protein co-crystal structure confirming the Phe to be a viable substitution for the C-terminal Leu (PDB ID: 7A00). The structure (Table S1† for Data collection and refinement statistics) reveals that the L6F variant peptide binds in the PDZ-binding site and reproduces many of the key recognition features and similar orientation of side chains, albeit with subtle differences (Fig. 3c–e).

Application of the in silico screening workflow to a modified query with the Phe substitution (Fig. 3g) identified a number of small-molecule candidate inhibitors (e.g. Fig. 3h, i and S12†). These candidates were screened in a fluorescence anisotropy competition assay at 30 and 300 μM following a similar approach to that pursued for the p53/hDM2 interaction, although a smaller number of less potent hits were obtained. Compound Z-1 had an IC_{50} ~ 300 μM in a fluorescence anisotropy competition assay (Fig. 3j). The moderate potency is deceiving; the low molecular weight (294 Da) and promising ligand efficiency (LE: 0.24) render it a promising starting point for subsequent optimization. 1H-15N HSQC perturbation shifts indicated interaction of the ligand with the SHANK1-PDZ domain (Fig. 3k and S13†), however in these analyses the absence of assignments prevented mapping of a putative binding site. Compound Z-1 has significant similarity to a recently-reported small-molecule ligand for SHANK3: compound Z-2.59 We synthesized and tested compound Z-2 in the fluorescence anisotropy competition and 1H-15N HSQC assays (see ESI†); these experiments confirmed that compound Z-2 also acts as...
Fig. 4 Development of p53/hDM2 inhibitors, (a) synthetic route to hit A1; (b) selected chemical structures and associated IC₅₀ values (fluorescence anisotropy competition: 150 nM hDM2 and p53₁₅₋₃₅Flu, 40 mM phosphate, pH 7.4, 200 mM NaCl and 0.02 mg ml⁻¹ BSA) of developed inhibitors (increasing colour intensity of the wedges denotes increased potency, *reduced anisotropy in the assay attributed to solubility †non-specific interactions observed); (c) ligand efficiency plot for selected inhibitors shows exploration of the chemical space starting from selected hit A1 (green sphere). Coloured shapes are assigned to describe inhibitors from the first (triangle), second (square, diamond and inverse triangle) and third series (star). (d) Mapping of chemical shift perturbations in hDM2 for selected inhibitors (750 MHz, 100 mM phosphate, pH 7.4, 2.5% glycerol, 1 mM DTT, increasing concentration of compound was titrated into 50 μM hDM2 and CSP was determined at 1:2 protein:compound ratio). Colour variation is associated with a chemical shift perturbation from 0 ppm (white) to 0.2 ppm (red).
a GKAP/SHANK1-PDZ inhibitor (IC₅₀ ~ 110 μM in competition assay, see ESI Fig. S14†).

Query-guided hit identification provides useful starting points for PPI inhibitor elaboration

To illustrate the application of our approach to the identification of useful starting ligands for further optimization, we carried out structure–activity relationship analysis on one of the p53/hDM2 inhibitor series. After a synthetic route was established (Fig. 4a and Schemes S1–S3†), the original hit A1 was resynthesized together with a series of derivatives (Fig. 4b for selected compounds). For the first series, full competition anisotropy experiments and NMR titrations (Fig. S15 and Table S2† for competition data and Fig. S16† for representative HSQC spectra) were performed for these (racemic) compounds, revealing compounds A1, 8a and 8b to be the most potent in the series (Fig. 4b). The loss of activity upon exchange of the 5-chloro-2-thioiphenyl ring for a benzene ring (compounds 9a–9d) and the hydrophobic amide linker length (compound 7c) point to a key role of these groups as determinants of inhibition. Indole ring modifications and pyridine-to-benzene substitution were tolerated; however, an aromatic group in this position was necessary (analogs 7b and 8b versus analogs 7d and 8c). Although we screened multiple crystallisation conditions, crystals of the A1/hDM2 complex were not obtained. Therefore, to rationalise the initial SAR and further investigate the binding mode of A1, we carried out molecular dynamics (MD) simulations (Fig. 5). The root mean square deviations (RMSD) of the atomic positions of the ligand over a 10 nanosecond simulation of the R-A1/hDM2 complex revealed important features. The indole and thiophene rings of R-A1 formed a fixed core occupying the Trp23 and Phe19 pockets, respectively. In contrast, the (pyridin-2-yl)ethyl moiety appeared to flip between two different positions, as a consequence of rotation about the CO-Cα carbon bond and the inherent flexibility of the ethyl linker (Fig. 5c and d). Interestingly, only one of those positions occupied the shallower Leu26 pocket, which corresponds to an extended conformation of the inhibitor. This behaviour was replicated in a longer 200 ns MD simulation (Fig. S17a and Movie R-A1†). For the S-enantiomer the initial docked poses fit less well to the design hypothesis; for one pose the thiophene was placed in the W23 site whilst the other pose placed the indole in the W23 site, however, in both simulations the remaining arms were more dynamic indicating a poor correlation with the design hypothesis (Fig. S17b, c, Movies S-A1-I and II†). Given the extended conformation of R-A1 was shown to best overlay with the three p53 hot-spot residues, we assumed it to be the most active conformation. Therefore, we developed a second series of inhibitors focusing on stabilisation of the extended conformation.

For the second series, we introduced modifications at three different positions of the hit structure A1: methyl substituents on the α carbon of the CO-Cα bond, cyclic constraints at the ethyl linker moiety and halogen, and, carboxylic acid or amide substituents at the C2-indole position (Fig. 4b, S15, S16 and Table S2†). These modifications were anticipated to increase steric congestion adjacent to the CO–Cα bond, restricting the accessible conformations in favour of the extended conformation with concomitant improvement in inhibitory potency (hypotheses supported by MD simulations Fig. S18–S21† see below). Compounds anti-10, 11a and 11b were identified as the most potent inhibitors from this second series (see Fig. S22 and S23† for syn/anti assignment). Although there is still a two- to five-fold difference in potency in comparison to Nutlin-3, the improved ligand efficiency (0.28 11a and 0.26 for anti-10 and 11b versus 0.22 for Nutlin-3) corroborates their potential as hits for development of drug-like p53/hDM2 inhibitors. The potential for further development of this series is underscored by the fact that it is possible to increase potency without marked increase in heavy atom count. Restriction of the CO–Cα bond with a gem-dimethyl group in 12 had no effect upon inhibition,
(IC50 ~ 8 μM), whereas single methyl analogues resulted in a significant variation in inhibitory activity. Diastereomer anti-10 inhibits the p53/hDM2 interaction with an IC50 five times smaller than syn-10. Resolution of the anti-10 enantiomers by chiral HPLC (>99.0% ee, see Fig S24 and 25†) allowed us to test them separately, with both showing comparable potency (3.1 μM for anti-10, 4.6 μM for (-)-anti-10 and 2.9 μM for (+)-anti-10). MD simulations were performed to explore behaviour of the 4 stereoisomers of 10 to determine which best fits the p53 binding site of hDM2. Although more complex behaviour was observed, the simulations validate the hypothesis that restriction of the conformational mobility improves inhibitor potency (Fig. S18 and S19, Movies RR-10, SS-10, RS-10 and SR10†). Compound 7a with a constrained ethyl linker appeared to have good inhibitory potency, however, although HSQC analyses were consistent with effective binding, a lower anisotropy was observed in the assay which we attribute to poor solubility. Nonetheless, one of the 8 stereoisomers was subjected to MD simulation; this was generated by adding the cyclopropyl ring to R-A1 in the only position compatible with the bound conformation and the constraints of the binding site. Bound 7a shows particularly low mobility (Fig. S20 and Movie 7a†) consistent with the design hypothesis.

Modifications at the C2-indole position also induced a notable increase in potency; derivatisation with a bromine in 11a and a carboxylic acid in 11b. For derivatives 11a and 11d, the significant difference in activity (1 μM versus 13 μM) was attributed to a regioisomer effect arising from the presence of both the bromine and chlorine substituents. In the absence of the bromine, the regioisomers A1 and 8a (6 μM versus 6 μM) elicited similar potencies, pointing to a potential synergy between the orientation of the indole in the Trp23 pocket and the orientation of the other two hot-spot mimicking groups. Inhibitor 11b was also observed to have increased potency; MD simulations place the 5-chloroindole ring deep in the Trp23 pocket, projecting the C2-indole carboxylate toward the edge of the α2 helix of hDM2 (see Fig. S21 Movies R11b and S11b†). This would permit additional hydrogen bonds, although a water mediated interaction has been observed for C2-strand-mediated interactions with the dynamic hDM2 surface when compared to p53) are in the AstraZeneca compound collection, but were not scored in the top 1000 hits identified for experimental screening – this represents an excellent hit rate. Whilst proteomimetics have been used to inhibit a range of intracellular α-helix-mediated PPIs, there remains a need to broaden the methodology to other targets and use the conceptual framework to identify small molecules that can be developed further.63-68 This study also emphasizes that identification of non-α-helix mediated PPI inhibitors e.g. β-strand-mediated PPIs is more difficult, although the identification of inhibitors for a strand-mediated interaction with a PDZ domain, albeit with a significantly lower hit-rate represents the first steps towards such a goal. Similarly, pharmacophore-based virtual screening approaches whereby key hot-spot residues have been used as an anchor69 have also recently been introduced, but to date brought to bear only on targets with a dominant hot residue. Taken together, our new approach harnesses the best of both worlds to identify and prioritise small-molecules that mimic diverse secondary structures and inhibit PPIs broadening the proteomimetic concept beyond α-helix mediated PPIs and extending it to genuine small-molecule ligands.

**Conclusions**

In summary, we have used query-guided inhibitor discovery to identify small-molecule inhibitors of the p53/hDM2 and GKA1/SHANK1 interactions. For each interaction, a computational workflow that involved FastROCS matching to secondary structure queries identified candidate inhibitors. Experimental screening of subsets of these candidates identified multiple genuine inhibitors demonstrating the approach to be valid for improving hit rate. Characterization of the most-promising compounds using two orthogonal assays identified inhibitors for both targets with promising ligand efficiency, which bound at the anticipated binding site and could be further developed. For instance starting with 200 compounds for experimental screening against p53/hDM2 we obtained IC50 values for 37 compounds. This demonstrates the strong performance of the computational workflow considering that high-throughput or fragment screening typically require >1 million or 10^3 compounds respectively to be screened to obtain similar numbers of hits to take forward. Although the approach is less likely to identify inhibitors that rely on induced conformational changes – indeed Nutlin-like compounds (which are known to have subtly different interactions with the dynamic hDM2 surface when compared to p53) are in the AstraZeneca compound collection, but were not scored in the top 1000 hits identified for experimental screening – this represents an excellent hit rate. Whilst proteomimetics have been used to inhibit a range of intracellular α-helix-mediated PPIs, there remains a need to broaden the methodology to other targets and use the conceptual framework to identify small molecules that can be developed further.63-68 This study also emphasizes that identification of non-α-helix mediated PPI inhibitors e.g. β-strand-mediated PPIs is more difficult, although the identification of inhibitors for a strand-mediated interaction with a PDZ domain, albeit with a significantly lower hit-rate represents the first steps towards such a goal. Similarly, pharmacophore-based virtual screening approaches whereby key hot-spot residues have been used as an anchor69 have also recently been introduced, but to date brought to bear only on targets with a dominant hot residue. Taken together, our new approach harnesses the best of both worlds to identify and prioritise small-molecules that mimic diverse secondary structures and inhibit PPIs broadening the proteomimetic concept beyond α-helix mediated PPIs and extending it to genuine small-molecule ligands.

**Author contributions**

S. C., F. H. and T. J. contributed equally to this work. D. N. W., R. B. S., T. A. E., D. M. A., A. N. and A. J. W., conceived and designed the research program, S. C., F. H., T. J., A. A. I., G. J. B., D. K. S., Z. H. and K. H. designed studies and performed research. The
manuscript was written by S. C., A. N and A. J. W. with contributions from all authors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by EPSRC (EP/N013573/1 and EP/KO39292/1) and the BBSRC/EPSRC-funded Synthetic Biology Research Centre, BrisSynBio (BB/L01386X/1). AJW holds a Royal Society Leverhulme Trust Senior Fellowship (SRF/R1/191087), AN holds an EPSRC Established Career Fellowship (EP/S015860/1). The authors wish to thank Kirsten Spence and Pallavi Ramsahye for protein production. We acknowledge Diamond Light Source for time on Beamline I03, I04 and I24 under Proposal mx19248 and the authors would like to thank the staff of these beamlines for assistance with data collection.

Notes and references

1 S. Surade and T. L. Blundell, Chem. Biol., 2012, 19, 42–50.
2 V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, Nat. Chem., 2013, 5, 161–173.
3 M. R. Arkin, Y. Tang and J. A. Wells, Chem. Biol., 2014, 21, 1102–1114.
4 L.-G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld and C. Ottmann, Chem. Rev., 2014, 114, 4695–4748.
5 M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, Angew. Chem., Int. Ed., 2015, 54, 8896–8927.
6 D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, Nat. Rev. Drug Discovery, 2016, 15, 533–550.
7 L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotoulis and E. A. Liu, Science, 2004, 303, 844–848.
8 D. Sun, Z. Li, Y. Rew, M. Gribble, M. D. Barterberger, H. P. Beck, J. Canon, A. Chen, X. Chen, D. Chow, J. Deignan, J. Duquette, J. Eksterowicz, B. Fisher, B. M. Fox, J. Fu, A. Z. Gonzalez, F. Gonzalez-Lopez De Turiso, J. B. Houze, X. Huang, M. Jiang, L. Jin, F. Kayser, J. Liu, M.-C. Lo, A. M. Long, B. Lucas, L. R. McGee, J. McIntosh, J. Mihalic, J. D. Oliner, T. Osgood, M. L. Peterson, P. Roveto, A. Y. Saiki, P. Shaffer, M. Toteva, Y. Wang, Y. C. Wang, S. Wortman, P. Yakowee, X. Yan, Q. Ye, D. Yu, M. Yu, X. Zhao, J. Zhou, J. Zhu, S. H. Olson and J. C. Medina, J. Med. Chem., 2014, 57, 1454–1472.
9 P. Holzer, K. Masuya, P. Furet, J. Kallen, T. Valat-Stachyra, S. Ferretti, J. Berghausen, M. Bouisset-Leonard, N. Buschmann, C. Pissot-Soldermann, C. Rynn, S. Ruetz, S. Stutz, P. Chène, S. Jeay and F. Gessier, J. Med. Chem., 2015, 58, 6348–6358.
10 Y. Zhao, A. Aguilar, D. Bernard and S. Wang, J. Med. Chem., 2015, 58, 1038–1052.
11 T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, R. C. Armstrong, D. J. Augeri, B. A. Belli, M. Bruncko, T. L. Deckwerth, J. Dinges, P. J. Hajduk, M. K. Joseph, S. Kitada, S. J. Korsmeyer, A. R. Kunzer, A. Letai, C. Li, M. J. Mitten, D. G. Nettlesheim, S.-C. Ng, P. M. Nimmer, J. M. O’Connor, A. Oleksijew, A. M. Petros, J. C. Reed, W. Shen, S. K. Tahir, C. B. Thompson, K. J. Tomasselli, B. Wang, M. D. Wendt, H. Zhang, S. W. Fesik and S. H. Rosenberg, Nature, 2005, 435, 677–681.
12 A. Kotschy, Z. Szlavik, J. Murray, J. Davidson, A. L. Maragno, G. Le Toumelin-Braizat, M. Chanrion, G. L. Kelly, J.-N. Gong, D. M. Moujalled, A. Bruno, M. Csekei, A. Pazzal, Z. B. Szabo, S. Sipos, G. Radics, A. Prosznyak, B. Balint, L. Ondi, G. Blasko, A. Robertson, A. Surgenor, P. Dokurno, I. Chen, N. Matassova, J. Smith, C. Pedder, C. Graham, A. Studeny, G. Lysiak-Auvity, A.-M. Girard, F. Gravé, D. Segal, C. D. Riffkin, G. Pomilio, L. C. A. Galbraith, B. J. Aubrey, M. S. Brennan, M. J. Herold, C. Chang, G. Guascon, N. Cauquil, F. Melchiore, N. Guigal-Stephan, B. Lockhart, F. Colland, J. A. Hickman, A. W. Roberts, D. C. S. Huang, A. H. Wei, A. Strasser, G. Lessene and O. Geneste, Nature, 2016, 538, 477–482.
13 A. W. Roberts, M. S. Davids, J. M. Pagel, B. S. Kah, S. D. Puvvada, J. F. Gerecitano, T. J. Kipps, M. A. Anderson, J. R. Brown, L. Gressick, S. Wong, M. Dunbar, M. Zhu, M. B. Desai, E. Cerri, S. Heitner Enschede, R. A. Humerickhouse, W. G. Wierda and J. F. Seymour, N. Engl. J. Med., 2016, 374, 311–322.
14 A. E. Tron, M. A. Belmonte, A. Adam, B. M. Aquila, L. H. Boise, E. Chiarpin, J. Cidado, K. J. Embrey, E. Gangl, F. D. Gibbons, G. P. Gregory, D. Hargreaves, A. J. Hendrickx, J. W. Johannes, R. W. Johnstone, S. L. Kazmirski, J. G. Kettle, M. L. Lamb, S. M. Matulis, A. K. Nooka, M. J. Packer, B. Peng, P. B. Rawlins, D. W. Robbins, A. G. Schuller, N. Su, W. Yang, Q. Ye, X. Zheng, J. P. Secrist, E. A. Clark, M. Wilson, S. E. Pawell and A. W. Hird, Nat. Commun., 2018, 9, 5341.
15 A. Ashkenazi, W. J. Fairbrother, J. D. Leveryson and A. J. Souers, Nat. Rev. Drug Discovery, 2017, 16, 273–284.
16 J. Grembecka, S. He, A. Shi, T. Purohit, A. G. Muntean, R. J. Sorenson, H. D. Showalter, M. J. Murali, A. M. Belcher, T. Hartley, J. L. Hess and T. Cierpicki, Nat. Chem. Biol., 2012, 8, 277–284.
17 G. Zimmermann, B. Papke, S. Ismail, N. Vartak, A. Chandra, M. Hoffmann, S. A. Hahn, G. Triola, A. Wittinghofer, P. I. H. Bastiaens and H. Waldmann, Nature, 2013, 497, 638–642.
18 J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells and K. M. Shokat, Nature, 2013, 503, 548–551.
19 T. D. Heightman, J. F. Callahan, E. Chiarpin, J. E. Coyle, C. Griffiths-Jones, A. S. Lakdawala, R. McMenamin, P. N. Mortenson, D. Norton, T. M. Peakman, S. J. Rich, C. Richardson, W. L. Runsey, Y. Sanchez, G. Saxty, H. M. G. Willems, L. Wolfe, A. J. A. Woolford, Z. Wu, H. Yan, J. K. Kerns and T. G. Davies, J. Med. Chem., 2019, 62, 4683–4702.
20 W. McCoull, R. D. Abrams, E. Anderson, K. Blades, D. Barton, M. Box, J. Burgess, K. Byth, Q. Cao, J. Chuaqui, R. J. Carbajo, T. Cheung, E. Code, A. D. Ferguson, S. Fillery, N. O. Fuller, E. Gangl, N. Gao, M. Grist, D. Hargreaves, M. R. Howard.
