About the Authors

Luca Laraia

Luca Laraia was born in Rome (Italy) and grew up in Vienna (Austria). He completed his MSc at Imperial College London (UK) before starting his PhD on the Cancer Research UK medicinal chemistry training program at the University of Cambridge (UK). Under the supervision of David Spring and Ashok Venkitaraman he is looking at disrupting protein–protein interactions relevant to cancer therapy.

David R Spring

David R Spring is currently a Professor at the University of Cambridge within the Chemistry Department and a Fellow of Trinity College (Cambridge, UK). He gained his BA (Hons) and MA in chemistry (1995) from the University of Oxford (UK), where he also achieved his DPhil (1998) for work on the proposed biosynthesis of the manzamine alkaloids under the supervision of Sir Jack E Baldwin. He then moved to Harvard University (MA, USA) to work with Stuart L Schreiber as a Wellcome Trust Postdoctoral Fellow and Fulbright Scholar (1999–2001), after which he joined the faculty at the University of Cambridge as a Biotechnology and Biological Sciences Research Council David Phillips Fellow (2001–2006), an Engineering and Physical Sciences Research Council Advanced Fellow (2006–2011) and an EPSRC Established Career Fellow (2012–2017). He was promoted to a university lectureship in 2006, to a senior lectureship in 2008, to a readership in 2011 and to a professorship in 2013. His research spans the disciplines of chemistry and biology through the synthesis of small molecules, which are applied to problems in the life sciences. In particular, he has focused on diversity-oriented synthesis, new synthetic methodologies and chemical biology in order to discover new antibiotics and anticancer drugs. He is a founder member of the Cambridge Molecular Therapeutics Programme that aims to use various approaches to find small molecules that modulate protein–protein interactions of targets relevant to cancer.

Chapter 3

Chemical library screening approaches to aid the design of protein–protein inhibitors

Luca Laraia & David R Spring

doi:10.4155/EBO.13.151

Protein–protein interaction (PPI) inhibitors discovered so far show marked differences from the majority of known drugs and especially compounds in commercial libraries. This goes some way towards explaining why high-throughput screening campaigns against PPIs are relatively unsuccessful, providing few tractable chemical starting points. PPI inhibitors tend to be larger, more hydrophobic, contain more sp^3 carbons and more stereocenters than compounds developed for more traditional targets such as kinases. This problem is beginning to be tackled by academic groups specializing in library synthesis and commercial compound vendors that are tailoring sections of their libraries for PPI inhibition and adding new compounds appropriately. In the academic environment, new strategies such as diversity-oriented synthesis provide access to screening quantities of diverse and natural product-like compounds, which possess an increased number of stereogenic centers and diversification points. Computational approaches are increasingly being used to filter commercial libraries for compounds with PPI-inhibitor

Protein–protein-interaction inhibitors tend to have larger molecular weights, increased hydrophobicity and more sp^3 carbons than enzyme inhibitors. New compound libraries that take these properties into account are needed.

properties. These combined strategies in time will hopefully deliver more promising leads for inhibiting PPIs.

Beyond the 'rule of five' for compound libraries

In the last decade, significant advancements have been made in the development of inhibitors of PPIs. The combined efforts of many laboratories have led to key insights into what may be required to identify PPI inhibitors, and the major pitfalls and difficulties. Many different approaches have been used, which include peptidomimetics, stabilized peptide therapeutics and fragment-based drug discovery, which are all discussed elsewhere in this book. However, high-throughput screening (HTS) remains one of the most widely used approaches. This typically requires large (>10,000) compound libraries, which are screened *in vitro* using a cell-free assay. One of the major problems of this approach has turned out to be the compound libraries themselves, in which putative PPI inhibitors are poorly represented, resulting in low hit rates. Traditional compound libraries, both commercial and in-house, have been biased towards historical drug targets of the last 15–20 years, which predominantly include enzymes and G-protein-coupled receptors. They have largely fallen into a set of parameters known as the 'rule of five', coined to increase the probability of obtaining an orally bioavailable molecule. Molecules designed to modulate traditional targets have not fared well against PPIs, the reasons for which are now becoming clear.

In general, PPI inhibitors have increased molecular masses and hydrophobicity compared with many traditional drugs; however, these properties vary depending on the specific interaction being targeted. The best examples to study these properties are those with several inhibitors available covering different scaffolds or chemotypes. These include the p53/MDM2 and the BCL-2/Bax interactions. p53 is a key tumor suppressor that can induce growth arrest and apoptosis. p53 can be inactivated by MDM2, which is over-expressed in some tumors. Therefore inactivating MDM2 by inhibiting its interaction with p53 is a viable anticancer strategy. The BCL family of proteins include several proapoptotic proteins that can be inactivated by antiapoptotic cellular factors. The inhibition of this interaction would restore proapoptotic activity and result in cancer cell death. If we compare six recently reported, potent inhibitors of the two different PPIs (e.g., p53/MDM2

Aa High-throughput screening: the use of large (typically >10,000) compound libraries in biological assays.

'Rule of five': a set of guidelines to increase the probability of finding an orally bioavailable drug.

and BCL-2/Bax; **Figure 3.1**) we can see some general trends as well as marked differences [1–6].

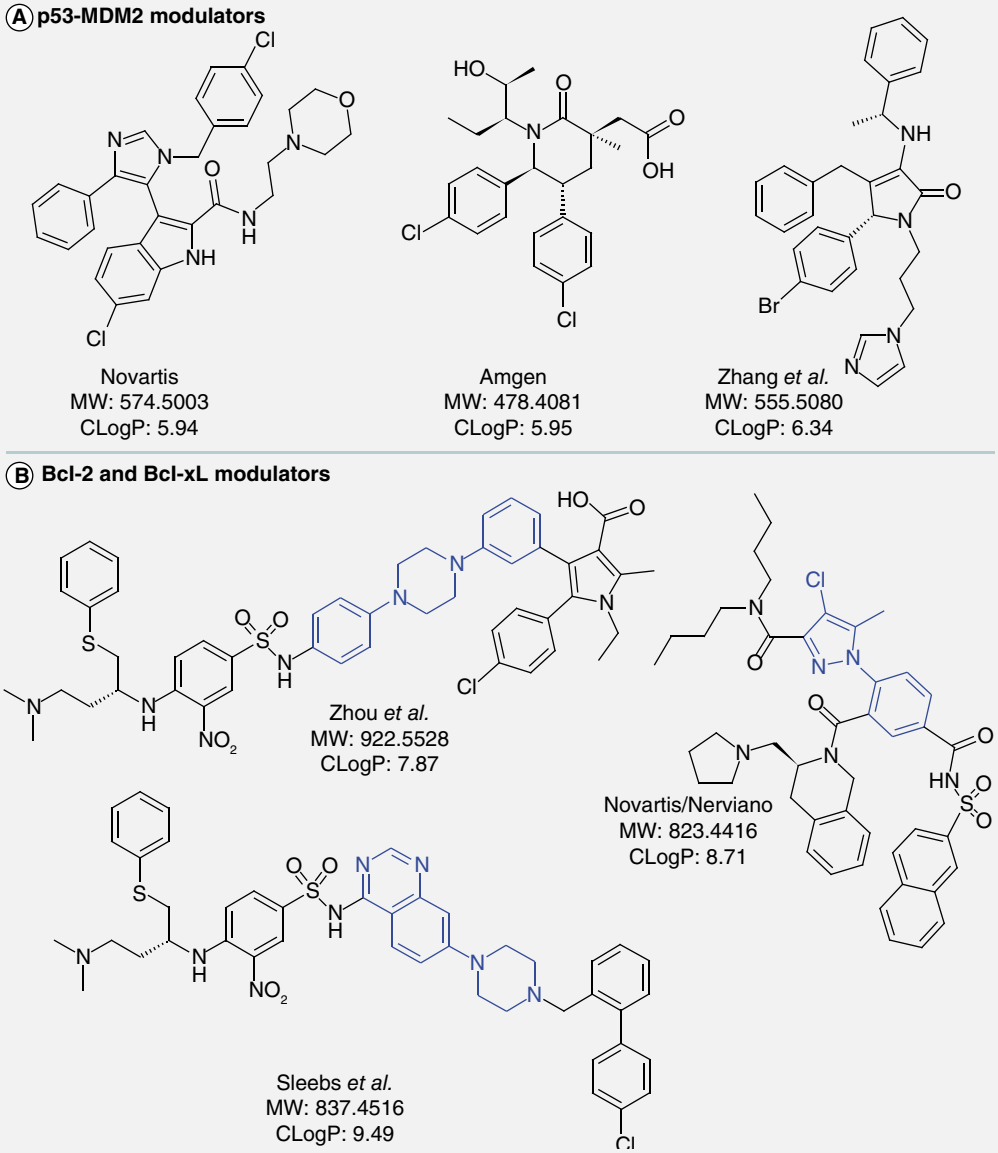
Compounds potently targeting MDM2 are smaller and more densely functionalized, while those targeting BCL-2 are significantly larger, with key functionalities also being more spread out across the molecules. These differences clearly reflect the variation in the small-molecule binding sites. MDM2 contains three adjacent hydrophobic clefts, filled by the two chlorophenyl and an ethyl substituent from lead compound identified by Amgen **(Figure 3.2A)**. Inhibitors of BCL-2, such as the one by Novartis, cover a much wider area and also contain large linker regions (highlighted in blue in **Figure 3.1**) that connect key interaction sites. The main similarity between both groups of molecules is the increased hydrophobic functionality compared with traditional drugs that target enzymes or receptors. This reflects the fact that key contacts at protein–protein interfaces are often made by burying hydrophobic residues in clefts on surfaces. Most analyzed compounds break the molecular weight (MW) and logP rules of the 'rule of five'; however, several of these compounds and derivatives are orally bioavailable and have either entered or will soon enter clinical trials [7]. Clearly, re-evaluating the structural guidelines for identifying PPI inhibitors would be prudent.

A more widespread analysis of all PPI inhibitors to date (conducted and stored in the 2P2I database) [8] has prompted researchers to formulate new guidelines to increase the probability of finding PPI inhibitors in screening libraries [9]. These are termed the 'rule of four' and include a large MW (>400), a higher logP (>4), an increased number of rings (more than four) and hydrogen-bond acceptors (more than four). It is important to note that this analysis was conducted on heavily optimized compounds, and may not necessarily reflect the requirements for initial leads. In addition, it does not take into account the fact that several PPI inhibitors were initially identified from fragment screens.

Observing other PPI targets with fewer validated compounds shows that while increasing hydrophobicity is generally required, a large molecular weight may not be **(Figure 3.3)**. Several recently discovered PPI inhibitors with significantly different structures maintain a smaller size, despite having undergone some degree of optimization.

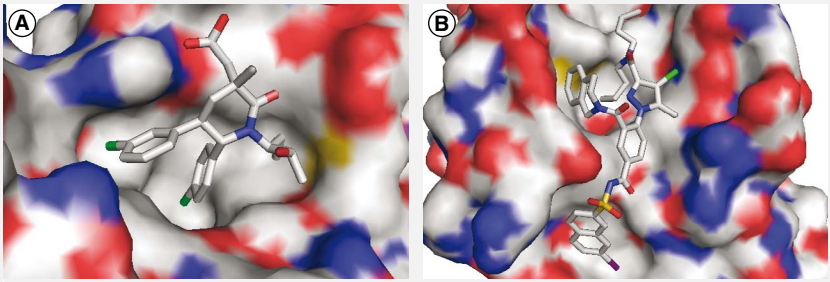
MI-2–2, an inhibitor of menin and the mixed lineage leukemia fusion protein that occurs in some types of leukemia, is a promising anticancer therapy [10], while GS-B, an inhibitor of the interaction between HIV integrase and the human protein LEDGF, could potentially provide a new treatment for HIV [11]. Both compounds have a MW of below 500 and CLogP

Figure 3.1. Small molecules targeting different protein–protein interactions display varying sizes and ClogPs.



Recently reported compounds targeting **(A)** the p53–MDM2 interaction and **(B)** the BCL-2 protein family. Calculated properties to exemplify noncompliance with 'rule of five' are displayed. Regions that do not contribute significantly to binding to BCL-2 are highlighted in blue.

Figure 3.2. Different small-molecule properties reflect differences in the protein surfaces they target.

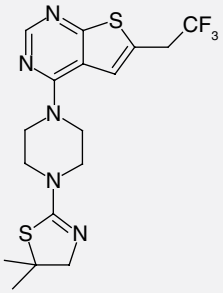


Inhibitors of **(A)** p53/MDM2 and **(B)** BCL-2/Bax highlight the difference in binding clefts. MDM2 contains three adjacent hydrophobic hotspots targeted by hydrophobic portions of the inhibitor. BCL-2 contains two larger hydrophobic regions that are targeted by lipophilic groups connected by a long linker.

values between 3.8 and 5.5; very reasonable values for drug leads. Therefore, rather than excluding compounds that fall below the 'rule of four' *a priori*, the authors propose a more inclusive approach that is less restrictive. For traditional HTS, only compounds classed as fragments would be removed in the first instance (or screened with the appropriate techniques), while larger compounds would remain. Those that remain would be subjected to strict parameters for removal of compounds containing undesirable functionalities. It has become particularly important to pay attention to structural features of molecules that may render them promiscuous and more likely to be false positives **(Figure 3.4)**, particularly when searching for PPI inhibitors, as will be outlined below.

Screens for novel PPI inhibitors often give fewer hits with lower potency, due to the lack of appropriate libraries. As a result researchers are often faced with the difficult decision to progress compounds that may otherwise have been discarded. To reduce the chance of a false positive it is important to be aware of common issues encountered in typical *in vitro* assays. These include nonspecific binding, aggregation and compound solubility problems, which can all result in a positive assay result. In addition,

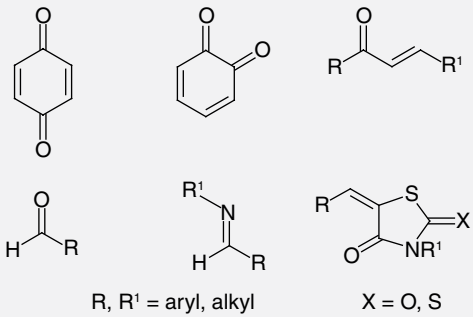
Figure 3.3. Recently identified PPI inhibitors with properties more similar to published guidelines for small-molecule orally bioavailable therapeutics.



MI-2
Menin/MLL inhibitor
MW: 415.50
CLogP: 3.88

GS-B
LEDGF/integrase inhibitor
MW: 437.20
CLogP: 5.49

Figure 3.4. Potentially reactive functionalities that may cause false positives in screening campaigns.



caution should be taken when using redox and otherwise reactive compounds, as there are examples in the literature where such compounds have resulted in wrong conclusions being drawn about PPIs. To counteract this, guidelines have emerged on how to remove or at least be aware of such compounds [12]. Researchers setting up screening libraries from scratch may also wish to use the appropriate computational filters before purchasing or synthesizing compounds containing questionable functional groups. It must be noted that like all such general 'rules' these filters are not absolute, and there are compounds with such properties in marketed drugs. If compounds with redox or reactive properties do appear as hits in screening campaigns, the appropriate control experiments to validate these compounds should be undertaken. These include confirming the hit in an orthogonal primary assay as well as a secondary biophysical assay, such as isothermal titration calorimetry, surface plasmon resonance or NMR. Where possible data should be complemented with credible structure–activity relationships and structural information on protein binding achieved through x-ray crystallography or NMR.

Increasing molecular diversity & complexity

It is now accepted that to increase the probability of identifying PPI inhibitors our screening collections must contain molecules with greater size, diversity and complexity. Interfaces between proteins can vary in size, shape and structure, and therefore it is unlikely that one single class or structural motif in a small molecule will be able to provide hits or leads against a whole range of PPIs. To combat this problem, several approaches to increase diversity and biological relevance have been outlined and, in some cases, successfully implemented. One of these is termed diversity-oriented synthesis (DOS) [13]. The general aim of DOS is to construct libraries of great structural diversity in a few easily executable synthetic steps to deliver milligram quantities of compounds for screening. There are four main classifications for molecular diversity: skeletal, appendage, functional group and stereochemical. Most DOS approaches focus on maximizing skeletal diversity, as variation in the molecular

Aa Diversity-oriented synthesis: the use of efficient reactions to produce a large and diverse set of small molecules for screening.

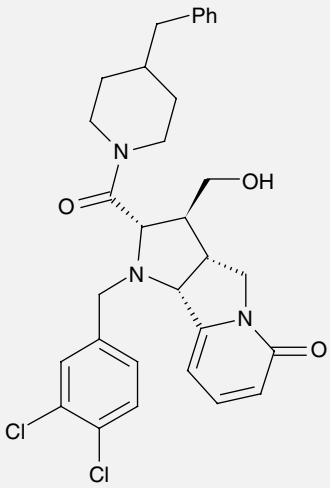
More emphasis needs to be set on functional group filters during library design to avoid false positives in high-throughput screening campaigns.

Diversity-oriented synthesis and the basic input/output system provide new strategies to address the diversity and natural product likeness issues in current libraries.

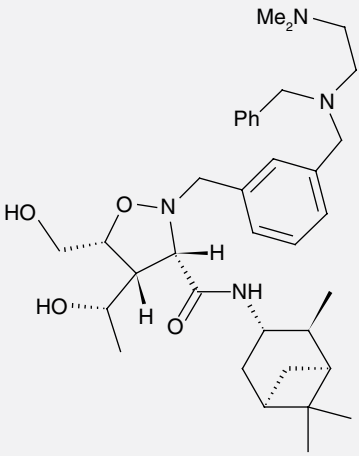
skeleton greatly affects the overall compound structure and results in the most diverse scaffolds, which will hopefully be bioactive in different ways. Several approaches to achieve this goal have been described. The reagent-based approach utilizes different reagents but common, cheap starting materials to assemble complex scaffolds. Conversely, the substrate-based approach uses different substrates assembled in few synthetic steps, which can be subjected to similar reaction conditions, leading to different scaffolds. The final molecular construction approach is termed the 'build/couple/pair' strategy, which can be thought of as a combination of the two previous strategies. As these concepts have been extensively reviewed in the literature, we will only discuss the successes as a result of this strategy and the future directions for this approach in PPI-inhibitor discovery.

Several small-molecule inhibitors of BCL-2 have been identified using DOS strategies **(Figure 3.5)**. Marcaurelle and coworkers have identified selective BCL-2 inhibitors from a cytidine-inspired DOS library [14]. Given their complexity, these compounds were prepared in few synthetic steps following a large-scale synthesis of key precursors. The compounds obtained are low micromolar inhibitors of BCL-2 and represent a useful starting point for medicinal chemistry. In line with the trends previously identified, these compounds have a large MW (>500), and relatively large CLogP values. In addition, they also display an increased number of stereocenters, a trait usually associated with natural products (NPs). These are predicted to confer selectivity for a specific target and are often included as a priority in compound libraries obtained from DOS. To assess the importance of stereogenic centers and sp^3 carbons, Schreiber *et al.* screened 15,000 compounds from diverse sources such as DOS, NPs and commercially available ones, on 100 unrelated protein targets [15]. The results confirmed that an increase in sp^3 content and stereogenic centers gave increased specificity and a better overall hit rate across the targets. Conversely, compounds with the least sp^3 carbons and stereocenters gave the highest degree of promiscuity. This correlates with the observation that stereocenter count increases during drug development [16]. Combined, this information confirms the importance of complexity in compound libraries and validates DOS as a strategy to address this. For compounds obtained by DOS to fulfil their potential, it is crucial that they are screened in as many PPI (and other) assays as possible. The strength of these libraries is the diverse bioactivity that the compounds display, which can require

Figure 3.5. Small-molecule inhibitors of BCL-2 identified by diversity-oriented synthesis.



MW: 566.5180
CLogP: 4.03
Four stereocenters



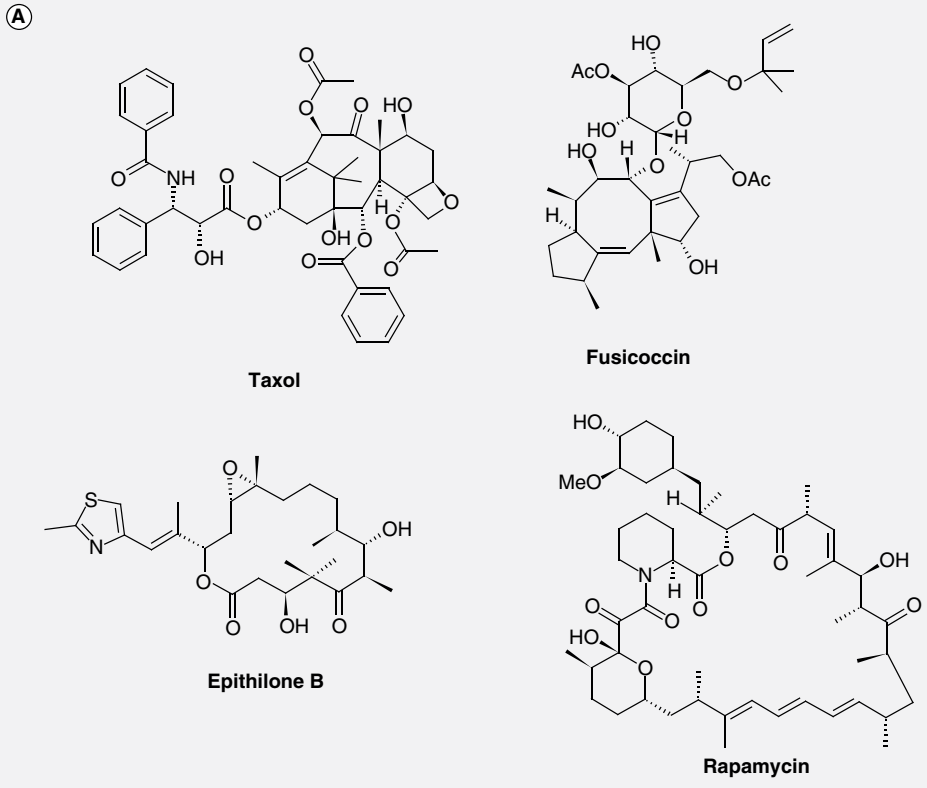
MW: 606.8384
CLogP: 5.60
Eight stereocenters

many screening campaigns to be identified. With new PPI targets being identified and screened for on a regular basis, the potential of DOS for inhibiting PPIs will continue to grow.

Given the importance of diversity in screening collections, it is imperative to note that this can also be achieved by screening NP libraries. NPs have long been thought to target key cell-survival mechanisms, which has recently also been quantified computationally [17]. This suggests that they may be most suitable for targeting pathogens or signaling pathways in diseases such as cancer. Despite well-known limitations, such as availability and synthetic intractability, NPs have been a vast source of drugs, and have provided some of the most well-known PPI modulators **(Figure 3.6)** [18].

Modulators of microtubule polymerization, such as the stabilizers taxol and epithilone B, as well the destabilizers colchicine and vinblastine, are all NPs. In addition, colchicine and vinblastine bind to different sites on tubulin, which are also distinct from the taxol binding site, suggesting that nature can offer different ways to modulate the same PPIs. New NP modulators of microtubules are still being discovered and approved for clinical use (e.g., eribulin, derived from the NP halichondrin). Other PPI

Figure 3.6. Natural products provide a wealth of protein–protein interaction modulators.



(A) Natural product stabilizers and **(B)** inhibitors of PPIs.

stabilizing compounds include the immunosuppressive NP rapamycin. It exerts its activity by binding to FKBP12, and this complex can bind and inhibit mTOR complex 1. Analogs of rapamycin have also found use as anticancer therapy due to several tumors having deregulated mTOR signaling. Another interesting example of a NP PPI stabilizer is fusicoccin. This compound stabilizes the interaction between a 14-3-3 protein and the plant proton pump PMA2, leading to the opening of the gas-exchange pores on its leaves [19].

All of the NPs discussed display large MWs, a high sp^3 carbon content and a greater number of stereocenters. This further supports the need for new guidelines for PPI-inhibitor design and the need for diverse and NP-like chemical libraries. To address this challenge, the strategy of biology-inspired synthesis has also been developed [20]. This involves synthesizing compounds based on known bioactive scaffolds, such as NPs, and evaluating them in several screening campaigns. This strategy has been successful in several target areas, and it will be interesting to see this concept applied to PPI screening efforts.

Commercial efforts for PPI screening libraries

HTS capabilities have moved from being exclusively conducted in industrial settings, to being available in larger academic institutions. This includes the possibility to set up screening libraries aimed towards identifying PPI inhibitors. To address bias in commercial libraries towards traditional targets, several companies have embarked on synthetic and acquisition efforts to enhance diversity in their libraries specifically enriching in PPI-inhibitor chemical space. Suppliers such as ChemDiv (CA, USA) [101] and Asinex (Moscow, Russia) [102] have both released targeted PPI libraries, which comprise a mixture of peptidomimetic scaffolds with a higher degree of diversity. The synthesis of compounds inspired by peptide or protein secondary structure is another area of active research in the PPI-inhibitor field. Compounds that can mimic secondary structure motifs at the PPI binding interface have proven to be useful inhibitors. In addition, these mimics can be useful across several PPIs with the same secondary structure, therefore significantly reducing the need to design specific scaffolds for individual PPIs. Obtaining such compounds commercially may enhance the hit rates of HTS campaigns against PPIs. Pharmaceutical companies are also interested in the potential wealth of new targets in PPI space, and have invested heavily in companies that provide novel libraries or new approaches for targeting them. For example, Forma Therapeutics (MA, USA), has developed a proprietary library containing compounds derived from DOS, and has a strong focus on PPIs as a target class. Another example

is Ensemble Therapeutics (MA, USA), which has used its patented DNA-templated chemistry to construct very large (>1 million) compound libraries based exclusively on macrocycles. Other companies using macrocycles specifically to target PPIs include TranzymePharma (Now Ocera Therapeutics, CA, USA) and Polyphor Ltd (Allschwil, Switzerland). Macrocycles have been described as particularly efficacious against PPIs, as their larger size is more suitable for targeting the large interfaces between proteins; however, the full potential for this target class remains to be discovered [21]. It is likely that specialized laboratories or dedicated companies such as those outlined, will be required to address the intrinsic difficulties with macrocycle library synthesis and modulation of the often-problematic pharmacokinetic and pharmacodynamic properties. The anticipated results in this area should be announced in the near future and will be indicative of whether this particular approach can be fruitful for providing clinically relevant PPI inhibitors.

Conclusion

Several approaches exist for creating or acquiring libraries for targeting PPIs, although the debate on the exact composition of such libraries continues. However, a general consensus has emerged that suggests that current libraries are inadequate due to their historical target bias, and that PPI inhibitors tend to be larger and more complex than inhibitors or other biological targets (e.g., enzymes and G-protein-coupled receptors). Current rules for filtering and prioritizing compounds such as the 'rule of five' are looking increasingly inadequate for PPIs. It is likely that being less restrictive on size and hydrophobicity, but more stringent on functional groups that are known to be problematic in cell-free assays is the best way forward. Proprietary libraries that address these factors are being produced, and several are commercially available for those who are not equipped to construct their own using techniques such as DOS. The latter are areas in which academic expertise in complex molecule synthesis can be particularly useful, and provide new avenues for such complex targets.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Summary

- Traditional compound libraries have been unsuccessful in screens for protein–protein interaction (PPI) inhibitors.
- PPI inhibitors are often larger and more complex than inhibitors of more traditional targets.
- Synthetic and computational strategies to tailor libraries for PPI inhibitors are being adopted.
- Commercial organizations and academics are contributing towards the establishment of PPI specific libraries.

References

1 Miyazaki M, Kawato H, Naito H *et al.* Discovery of novel dihydroimidazothiazole derivatives as p53/MDM2 protein–protein interaction inhibitors: synthesis, biological evaluation and structure–activity relationships. *Bioorg. Med. Chem. Lett*. 22(20), 6338–6342 (2012).

2 Perez HL, Banafi P, Bertrand J *et al.* Identification of a phenylacylsulfonamide series of dual BCL-2/BCL-xL antagonists. *Bioorg. Med. Chem. Lett*. 22(12), 3946–3950 (2012).

3 Rew Y *et al.* Structure-based design of novel inhibitors of the MDM2–p53 interaction. *J. Med. Chem.* 55(11), 4936–4954 (2012).

4 Zhou H, Aguilar A, Chen J *et al.* Structure-based design of potent BCL-2/BCL-xL inhibitors with strong *in vivo* antitumor activity. *J. Med. Chem*. 55(13), 6149–6161 (2012).

5 Zhuang C, Miao Z, Zhu L *et al.* Discovery, synthesis, and biological evaluation of orally active pyrrolidone derivatives as novel inhibitors of p53-MDM2 protein–protein interaction. *J. Med. Chem*. 55(22), 9630–9642 (2012).

6 Furet P, Chene P, De Pover A *et al.* The central valine concept provides an entry in a new class of non peptide inhibitors of the p53–MDM2 interaction. *Bioorg. Med. Chem. Lett.* 22(10), 3498–3502 (2012).

7 Gandhi L, Camidge DR, Ribeiro de Oliveira M *et al.* Phase I study of navitoclax (ABT-263), a novel BCL-2 family inhibitor, in patients with small-cell lung cancer and other solid tumors. *J. Clin. Onc.* 29(7), 909–916 (2011).

8 Bourgeas RL, Basse MJ, Morelli X, Roche P. Atomic analysis of protein–protein interfaces with known inhibitors: the 2P2I database. *PLoS ONE*, 5(3), e9598 (2010).

9 Morelli X, Bourgeas RL, Roche P. Chemical and structural lessons from recent successes in protein–protein interaction inhibition (2P2I). *Curr. Opin. Chem. Biol.* 15(4), 475–481 (2011).

10 Shi A, Murai MJ, He S *et al.* Structural insights into inhibition of the bivalent menin–MLL interaction by small molecules in leukemia. *Blood* 120(23), 4461–4469 (2012).

11 Tsiang M, Jones GS, Niedziela-Majka A *et al.* New class of HIV-1 integrase (IN) inhibitors with a dual mode of action. *J. Biol. Chem.* 287(25), 21189–21203 (2012).

12 Baell JB, Holloway GA. New substructure filters for removal of Pan Assay Interference Compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* 53(7), 2719–2740 (2010).

13 O' Connor CJ, Beckmann HSG, Spring DR. Diversity-oriented synthesis: producing chemical tools for dissecting biology. *Chem. Soc. Rev.* 41(12), 4444–4456 (2012).

14 Marcaurelle LA, Johannes C, Yohannes D, Tillotson BP, Mann D. Diversity-oriented synthesis of a cytisine-inspired pyridone library leading to the discovery of novel inhibitors of BCL-2. *Bioorg. Med. Chem. Lett.* 19(9), 2500–2503 (2009).

15 Clemons PA, Bodycombe NE, Carrinski HA *et al.* Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *Proc. Natl Acad. Sci. USA* 107(44), 18787–18792 (2010).

16 Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation

as an approach to improving clinical success. *J. Med. Chem.* 52(21), 6752–6756 (2009).

17 Dancik V, Seiler KP, Young DW *et al.* Distinct biological network properties between the targets of natural products and disease genes. *J. Am. Chem. Soc.* 132(27), 9259–9261 (2010).

18 Koehn FE, Carter GT. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* 4(3), 206–220 (2005).

19 Rose R, Erdmann S, Bovens S *et al.* Identification and structure of small-molecule stabilizers of 14-3-3 protein–protein interactions. *Angew. Chem. Int. Ed.* 49(24), 4129–4132 (2011).

20 Wetzel S, Robin S Bon, Kumar K *et al.* Biology-oriented synthesis. *Angew. Chem. Int. Ed.* 50(46), 10800–10826 (2011).

21 Marsault E, Peterson ML. Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.* 54(7), 1961–2004 (2011).

Websites

101 ChemDiv. Fragment library. http://us.chemdiv.com/index.php/discovery-tools-chemistry/screening-libraries/drug-like/special-edition

102 Asinex. Protein–protein interactions. www.asinex.com/PPI_Library.html