Quarterly Reviews of Biophysics, Page 1 of 44. © Cambridge University Press 2012
doi:10.1017/S0033583512000108 Printed in the United States of America

Biophysical and computational fragment-based approaches to targeting protein–protein interactions: applications in structure-guided drug discovery

Anja Winter[1], Alicia P. Higueruelo[1], May Marsh[1], Anna Sigurdardottir[1], Will R Pitt[1,2] and Tom L. Blundell[1]*

[1] Department of Biochemistry, University of Cambridge, Cambridge CB1 2GA, UK
[2] Department of Medicinal Chemistry, UCB Pharma, Slough SL1 3WE, UK

Abstract. Drug discovery has classically targeted the active sites of enzymes or ligand-binding sites of receptors and ion channels. In an attempt to improve selectivity of drug candidates, modulation of protein–protein interfaces (PPIs) of multiprotein complexes that mediate conformation or colocation of components of cell-regulatory pathways has become a focus of interest. However, PPIs in multiprotein systems continue to pose significant challenges, as they are generally large, flat and poor in distinguishing features, making the design of small molecule antagonists a difficult task. Nevertheless, encouragement has come from the recognition that a few amino acids – so-called hotspots – may contribute the majority of interaction-free energy. The challenges posed by protein–protein interactions have led to a wellspring of creative approaches, including proteomimetics, stapled α-helical peptides and a plethora of antibody inspired molecular designs. Here, we review a more generic approach: fragment-based drug discovery. Fragments allow novel areas of chemical space to be explored more efficiently, but the initial hits have low affinity. This means that they will not normally disrupt PPIs, unless they are tethered, an approach that has been pioneered by Wells and co-workers. An alternative fragment-based approach is to stabilise the uncomplexed components of the multiprotein system in solution and employ conventional fragment-based screening. Here, we describe the current knowledge of the structures and properties of protein–protein interactions and the small molecules that can modulate them. We then describe the use of sensitive biophysical methods – nuclear magnetic resonance, X-ray crystallography, surface plasmon resonance, differential scanning fluorimetry or isothermal calorimetry – to screen and validate fragment binding. Fragment hits can subsequently be evolved into larger molecules with higher affinity and potency. These may provide new leads for drug candidates that target protein–protein interactions and have therapeutic value.

* Author for correspondence: Tom Blundell, The University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK.
Tel: +44(0)1223 333628; Email: tlb20@cam.ac.uk

1. Background

Targeting regulatory sites of enzymes, receptors and ion channels has become a new focus in drug discovery. Most activity has been in the design of small molecule allosteric activators or inhibitors that stabilise particular conformers (Lee & Craik, 2009). An alternative and increasingly popular approach is to target the protein–protein interfaces (PPIs) of multiprotein complexes that mediate conformation or colocation of regulatory components (Wells & McClendon, 2007). Both approaches target sites other than enzyme active sites or receptor ligand-binding sites; we therefore classify them as 'allo-targeting'. As allosteric sites and PPIs tend to be less conserved than active sites, they may offer greater opportunities for selectivity, but at the same time may pose greater challenges in drug design as they often have less well-defined binding sites. In this review, we focus on targeting PPIs.

Multiprotein assemblies provide an important contribution to the molecular regulation of most biochemical pathways involved in cell signalling, growth and survival, and their dysregulation is

often causal of disease. Biochemical assays and high-throughput interactomics studies have indicated the presence of ~130 000 possible interactions between proteins in humans, excluding splice variant complexity (Venkatesan *et al.* 2009), and many may constitute useful therapeutic targets. Genes that share similar or the same disease phenotypes often encode proteins that interact with each other (Lage *et al.* 2007). PPIs are also important therapeutic targets, where diseases are the result of protein misfolding and aggregation, as exemplified by rare genetic diseases such as Huntington's disease (Young, 2003), Duchenne muscular dystrophy (Khurana & Davies, 2003) and Alzheimer's disease (Bonda *et al.* 2010).

Although modulation of PPIs for therapeutic intervention is becoming of increasing interest, multiprotein systems continue to pose significant challenges. The inter-protein surfaces are generally large (~1500–3000 Å^2), flat and lack distinguishing features (Blundell *et al.* 2000), making the design of small molecule antagonists a difficult task (Blundell *et al.* 2006; Hopkins & Groom, 2002). Many consider them 'undruggable'. However, encouragement has come from the recognition that a few amino acids – so-called hotspots – may contribute the majority of interaction free energy. These amino acids are usually found in the centre of the interface, surrounded by residues that have a lesser effect on stability (Bogan & Thorn, 1998). Hotspots often have flexible aromatic side chains, which can allow conformational change leading to binding sites that can accommodate small molecules (Wells & McClendon, 2007).

The challenges posed by PPIs have led to a wellspring of creative approaches. Proteomimetics mimic elements of surface structure that occur at protein interfaces, including porphyrins and α-helical mimetics (Yin & Hamilton, 2004); for a good review see Fletcher & Hamilton (2005). Stapled peptides target α-helical peptide molecular recognition sites. They are stabilised by carefully optimised chemistry to cross-link adjacent residues in the helix and are therefore resistant to proteolysis, which occurs when helices are unfolded. Targets include Bcl-2 homology domain 3 : myeloid cell leukaemia sequence 1 (BH3:Mcl-1) (Stewart *et al.* 2010), p53:MDM2/X (protein 53:murine double minute) (Bernal *et al.* 2010) and mastermind-like protein 1 (MAML-1):Notch (Moellering *et al.* 2009), all of which involve helices central to the protein–protein interactions. Foldamers are fully synthetic mimics of protein secondary structure elements (for a review, see Wilson, 2009), peptide aptamers recognise specific protein domains (Buerger & Groner, 2003; Choi *et al.* 2009) and a plethora of antibody inspired constructs are designed to prevent protein binding and/or activation (Traczewski & Rudnicka, 2011). Other inventors have taken natural products as their inspiration, not least because of the extent of the natural product derived pharmacopoeia. Many of these molecules are highly complex in structure. One of the best known therapeutics that modulates a PPI is the macrocycle rapamycin (Chen *et al.* 1995; Choi *et al.* 1996).

A more generic approach is to use fragment-based drug discovery. Fragments allow novel areas of chemical space to be explored more efficiently, but the initial hits have low affinity. This means that they will not normally disrupt PPIs, unless they are tethered, an approach that has been pioneered by Wells and co-workers, with some impressive successes (Wells & McClendon, 2007). An alternative fragment-based approach is to stabilise the uncomplexed components of the multiprotein system in solution and employ conventional fragment-based screening (Blundell *et al.* 2002; Hajduk & Greer, 2007; Murray & Blundell, 2010; Shuker *et al.* 1996). This approach requires sensitive biophysical methods – nuclear magnetic resonance (NMR), X-ray crystallography, surface plasmon resonance (SPR), differential scanning fluorimetry (DSF) or isothermal calorimetry (ITC) – to screen and validate fragment binding. Fragment hits can

subsequently be evolved into larger lead- and drug-like molecules with higher affinity and potency.

This review focuses on two aspects of targeting PPIs. The first concerns current knowledge of the structures and properties of PPIs and the small molecules that can modulate them. The second relates to the use of biophysical and structural approaches to investigating binding of chemical fragments, their growth or cross-linking with the aim of providing new leads for drug candidates that target PPIs and have a therapeutic value.

2. PPIs and their chemical modulators

2.1 Protein–protein interactions

Multiprotein assemblies may be characterised as obligate, where the subunits are not observed independently *in vivo*, or non-obligate where they are independently observed, often being synthesised in different cells or tissues. Components and complexes of non-obligate assemblies are in dynamic equilibrium, albeit with a range of binding affinities and on- and off-rates (Nooren & Thornton, 2003). Those with high affinities are often referred to as 'permanent', and those with lower affinities as 'transient'. Further descriptors relate to whether the interaction region of a subunit is a discontinuous epitope, where strands of polypeptide from different regions of the sequence contribute, or a continuous epitope in which a continuous region of polypeptide, often comprising a single secondary structure element (α-helix or β-strand), makes up the interaction interface. Many of the latter group involve flexible polypeptides or disordered regions of polypeptide chains that assemble with a more classical globular protein to give a globular complex (Wright & Dyson, 1999, 2009). Schwyzer proposed this idea for polypeptide hormones in the 1970s (Schwyzer *et al.* 1979). It was exemplified by glucagon where biophysical and X-ray analyses (Sasaki *et al.* 1975) together with NMR studies in a lipid–water interface (Braun *et al.* 1983) suggested a disorder-to-order transition of the polypeptide on receptor binding (Blundell, 1979; Blundell & Wood, 1982). Such transient PPIs potentially offer additional distinct sub-sites for targeting using small molecule inhibitors (Blundell *et al.* 2006) as discussed below for the RAD51–breast-cancer-associated gene 2 (BRCA2) interaction (Pellegrini *et al.* 2002).

Several databases provide access to the results of published high-throughput protein interaction studies. The Database of Interacting Proteins (DIP; Xenarios *et al.* 2000), containing 23 201 proteins from 372 organisms totalling 71 276 high quality and manually verified data regarding experimentally determined protein–protein interactions. IntAct (Hermjakob *et al.* 2004) is an open source molecular interaction database and software suite from the European Bioinformatics Institute (EBI; http://www.ebi.ac.uk/intact). The Molecular Interaction Database (MINT; Chatr-Aryamontri *et al.* 2007) primarily focuses on PPIs from mammalian genomes. Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), a database of known and predicted direct physical and indirect functional PPIs, contains 2 590 259 proteins from 630 species (Szklarczyk *et al.* 2011). The International Molecular Exchange Consortium (Orchard *et al.* 2007), an international group of data resources, facilitates exchange of data and avoidance of duplication. However, there are still outstanding issues to be solved, in particular, the lack of standardised approaches in data created before the formation of the International Molecular Exchange Consortium (http://www.imexconsortium.org/; a non-redundant set of protein–protein interaction data from a broad taxonomic range of organisms), which prevents data from different sources from being consolidated (Turinsky *et al.* 2011).

Table 1. *Databases and resources for structurally characterised protein–protein interactions, adapted from Bickerton (2009) PhD Thesis, University of Cambridge and Lee* et al. *(2009).*

Database and access	Description
3DID (Stein *et al.* 2011) http://3did.irbbarcelona.org/	Intra- and inter-molecular interactions between Pfam domains from high-resolution crystal structures. Includes GO-based functional annotations and predicted similarity of interactions across families
DAPID (Chen *et al.* 2006) http://gemdock.life.nctu.edu.tw/dapid/	Domain-annotated protein interactions Basis for prediction of novel interactions through 3D-domain interologs
DIMA (Luo *et al.* 2011) http://webclu.bio.wzw.tum.de/dima/	Known and predicted protein domain interactions
DOCKGROUND (Gao *et al.* 2007) http://dockground.bioinformatics.ku.edu/	Dynamic generation of non-redundant bound-bound datasets for docking
DOMINE (Yellaboina *et al.* 2011) http://domine.utdallas.edu/cgi-bin/Domine	Known and predicted protein domain (domain–domain) interactions
ICBS (Dou *et al.* 2004) http://icbs.ics.uci.edu/	Interactions mediated by inter-chain β-sheet formation
NEGATOME (Smialowski *et al.* 2010) http://mips.helmholtz-muenchen.de/proj/ppi/negatome	Experimentally supported non-interacting protein pairs
PIBASE (Davis & Sali, 2005) http://pibase.janelia.org/pibase2010/queries.html	Classification of polar and non-polar surface area using SCOP and CATH domain definitions
PICCOLO (Bickerton *et al.* 2011) http://www-cryst.bioc.cam.ac.uk/piccolo	Classification of interfacial atomic contacts in PDB ASU and PISA-predicted quaternary assemblies
PRISM (Ogmen *et al.* 2005) http://prism.ccbb.ku.edu.tr/	Clustering interface using sequence-order independent method
ProtBuD (Xu *et al.* 2006) http://dunbrack.fccc.edu/ProtBuD/index.php	Comparison of asymmetric and biological units from PDB and PQS
PsiBase (Gong *et al.* 2005) http://psibase.kobic.re.kr/	Networks of SCOP family/superfamily relationships
SCOPPI (Winter *et al.* 2006) http://www.scoppi.org/	Interface classification according to geometry of domain associations
SCOWLP (Teyra *et al.* 2008) http://www.scowlp.org/	Protein-binding regions classified by SCOP families and interfaces defined by interacting partner: protein, peptides, nucleic acids and saccharides
SNAPPI-D (Jefferson *et al.* 2007) http://www.compbio.dundee.ac.uk/SNAPPI/	Cross-linking among Pfam, SWISSPROT, InterPro, GO terms, secondary structures and multiple structure alignments

2.2 Structurally characterised PPIs

Numerous databases describe PPIs in terms of their three-dimensional (3D) protein structures, and these are summarised in Table 1 (Bickerton, 2009; Lee *et al.* 2009). 3D Interaction Domains (3DID; (Stein *et al.* 2011)) describe intra- and inter-molecular interactions between protein families (Pfam) domains from high-resolution crystal structures, and includes functional annotations based on Gene Ontology and predicted similarity of interactions across families. Another database of 3D structures of domains is DAPID (Domain Annotated Protein–Protein Interaction Database; Chen *et al.* 2006).

PICCOLO (Bickerton, 2009; Bickerton *et al.* 2011; http://www-cryst.bioc.cam.ac.uk/piccolo) is a comprehensive database of structurally characterised PPIs derived from quaternary assemblies, generated by the EBI Protein Interfaces Surfaces and Assemblies (PISA) resource (Krissinel & Henrick, 2007). It identifies interfaces clustered to reflect interface groups sharing equivalent residue patches and relative geometry. The interactions are described at the level of interacting pairs of atoms, residues and polypeptide chains. PICCOLO distinguishes 12 different interaction types using distance and angle terms, which are used to analyse physico-chemical properties of PPIs including residue propensity, hydropathy, polarity, residue contact preference and sequence entropy.

Nussinov and co-workers have prepared a very useful review of the databases and tools for studying PPIs (Tuncbag *et al.* 2009). Table 1 summarises databases and resources for structurally characterised protein–protein interactions.

2.3 Small molecule binding to PPIs

Less effort has been focused on databases that allow comparison of small molecule binding to PPIs. One of the first was TIMBAL, a hand-curated database holding current small molecules inhibiting multi-protein complexes (Higueruelo *et al.* 2009). The database gives insights into which types of molecules are favoured by protein interfaces. Comparisons of these small molecules with drugs from the MDDR database (MDL Information Systems Inc., San Leandro, CA, USA) [MDL® drug data report], classical screening compounds from commercial vendors and drug-like molecules in the Protein Data Bank (PDB) show that inhibitors of PPIs are on average bigger and more lipophilic, and have fewer hydrogen bonding features. Furthermore, for the subset of molecules in the PDB, TIMBAL molecules engage in more hydrophobic contacts and fewer hydrogen bonds than drug-like molecules, consistent with their molecular property profile (Higueruelo *et al.* 2009).

2P2IDB (http://2p2idb.cnrs-mrs.fr) is a hand-curated database of protein–protein complexes with known inhibitors (Bourgeas *et al.* 2010). The authors have compared protein–protein and protein–inhibitor interfaces in terms of a number of parameters including: geometry, atom and residue properties, buried and accessible surface areas, and protein–protein dissociation constants (K_D). Most complexes found in this database have accessible surface areas that are significantly smaller than the average, and these authors observed no major conformational changes between protein in the complex, bound to an inhibitor and the unbound states. Interfaces represented in this database are also more hydrophobic and with less charged residues and more non-polar atoms than protein-interaction interfaces that have no known inhibitor.

An example of a protein–protein interaction site that has been targeted for therapeutic intervention is human RAD51 (Fig. 1), a recombinase with homology to bacterial RecA and archeal RadA. RAD51 self-associates using a conserved FxxA sequence motif to form an ordered helical nucleoprotein filament on DNA substrates, and catalyses the strand exchange reaction that is essential for DNA repair by homologous recombination. In humans, the interaction between RAD51 and the BRCA2 is required for this activity. Interaction with BRCA2 disrupts the self-association of RAD51 by mimicking the self-association FxxA motif with its conserved, ca. 40 amino acid motifs known as the BRC repeats. These repeats interact not only with the FxxA binding on RAD51 but also with an adjacent hydrophobic patch to the FxxA site as revealed by the crystal structure of a complex between RAD51 and the fourth BRC repeat, BRC4 (Pellegrini *et al.* 2002). Expression of BRC4 on its own in human cells competes with

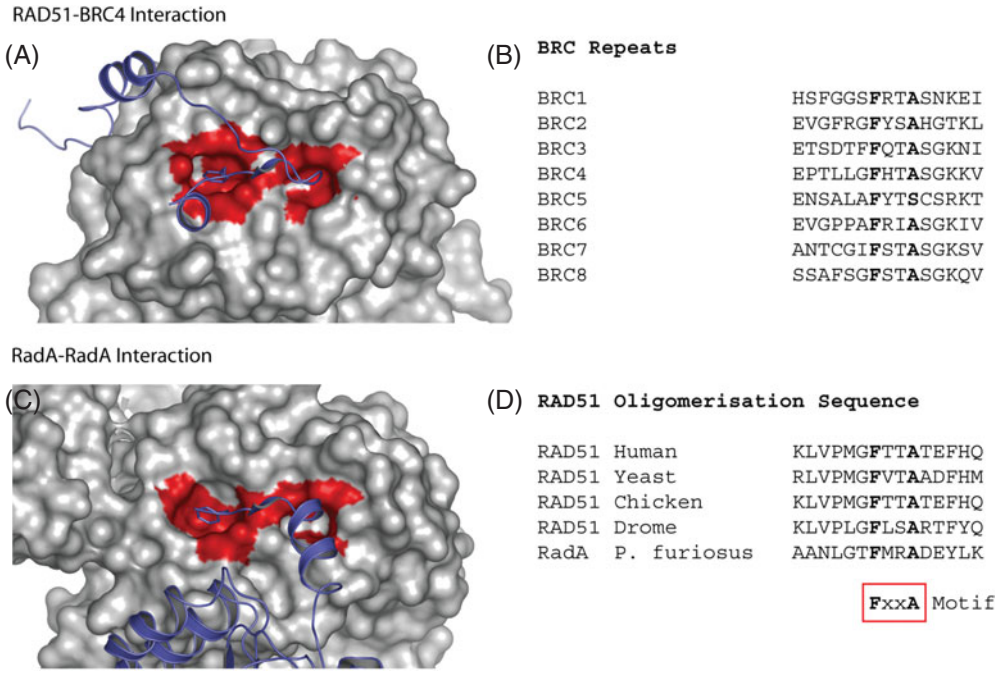

Fig. 1. PPI involving RAD51 and its homologues. (A) The structure of RAD51 complexed with the region BRC4 of BRCA2 (Pellegrini *et al.* 2002) demonstrating the existence of two well-defined pockets on RAD51 that are occupied by side chains of the conserved FxxA motif. RAD51 is shown as grey van-der-Waals surface. BRC4 is shown in purple cartoon form. Residues within a 4 Å radius of the FxxA motif are highlighted in red on the RAD51 surface. (B) Sequences of homologous repeats in human BRCA2. (C) An equivalent view of *P. Furiosus* RadA in a protein oligomeric filament showing the similarity of the interface with that of the RAD51 BC4 complex. The interacting oligomerisation region of the adjacent RadA protomer is shown as a purple cartoon. (D) Oligomerisation sequences of RAD51 orthologues and RadA.

endogenous BRCA2 for RAD51 and renders cells sensitive to DNA-damaging agents. BRC4 is unstructured in solution and undergoes concerted folding and binding on interacting with RAD51.

2.4 Binding-site druggability

Two factors conventionally define a 'druggable' protein target, first whether its modulation has a therapeutic effect and second whether it is able to bind to a small drug-like molecule (Hopkins & Groom, 2002). Druggability predictors usually refer to the latter which is sometimes also called 'ligandability' (Edfeldt *et al.* 2011).

Most current prediction methods depend on identification and scoring of pockets in the surface for their likelihood to accommodate a small molecule. However, contact areas between proteins exhibit some flexibility of side chains and loops creating new cavities that can bind small molecules. These cavities cannot always be observed from static structures of either the free protein target or the protein–protein complex (Wells & McClendon, 2007) and are therefore difficult to predict. Nonetheless, an increased effort has been made over the last decade to predict druggable PPIs.

Currently available predictors for classical receptor and enzyme targets are based on geometrical scanning (PocketFinder, based on Ligsite; Hendlich *et al.* 1997) and PocketDepth

(Kalidas & Chandra, 2008), favourable interaction energy between the protein and a van-der-Waals probe (Q-SiteFinder; Laurie & Jackson, 2005) and i-SITE (Morita *et al.* 2008), Voronoi tessellation and alpha spheres (Fpocket; Le Guilloux *et al.* 2009), and random forest classifiers using residue-based properties (SitePredict; Bordner, 2009). Schmidtke and co-workers evaluated and compared Finder, Fpocket, PocketFinder and SiteMap using a test set containing 5416 protein–ligand complexes and 9900 apo-forms and found that they are all around 95% successful in identifying binding sites (Schmidtke *et al.* 2010). A comprehensive review in pocket detection and druggability has been recently published (Perot *et al.* 2010). An open source repository for druggable and un-druggable proteins is maintained to help improve druggability scores (Schmidtke & Barril, 2010). Analysis of these structures reveals that in addition to hydrophobicity of the cavities, polar groups have an important role in molecular recognition and should be considered in the druggability predictions.

However, this is a retrospective approach, mostly through assessment of the 3D structures of proteins that have proved successful, and therefore of limited value in assessing a new class of targets, such as PPIs, and certainly not those where pockets are induced by ligand binding. Indeed Fuller and co-workers showed that pocket-finding algorithms predict marked differences between the binding pockets that define PPIs and those that define protein–ligand interactions (PLIs) of currently marketed drugs. They suggest that approaches that simultaneously target several small pockets at the PPI are likely to be more successful (Fuller *et al.* 2009).

Using a different approach, ANCHOR (Meireles *et al.* 2010) is a web server to find the anchor residues at PPIs. Anchor residues are preformed recognition motifs and correspond to surface side chains that bury large solvent accessible areas upon binding (Rajamani *et al.* 2004). The ANCHOR web service calculates the change in buried area and its estimated contribution to the energy of interaction, as well as displaying stereo and chemical properties of the proposed sites for drug binding. It also holds calculated anchor residues for the whole PDB. Druggable Protein–Protein Interaction Assessment System (Dr PIAS; http://asp.gridasp.net/drpias/pias_top.php) is another service to assess druggability for PPIs (Sugaya & Furuya, 2011; Sugaya & Ikeda, 2009). In this case, in addition to the structural properties of the pockets at the interface, the authors also include in the model small molecule chemical information and functional data of the proteins involved in the complex. Running Q-SiteFinder, ANCHOR and Dr PIAS on the RAD51/BRC4 structure (Pellegrini *et al.* 2002) (Fig. 1), we found that all three predictors identify the pocket that binds phenylalanine of FxxA; however; only ANCHOR predicts the small alanine pocket as druggable as well.

Another approach to assist the search for small molecules binding to PPIs is to overlay homologous proteins that bind to other proteins and to small molecules (Davis & Sali, 2010) in order to find small molecules capable of substituting one of the binding partners of the protein complex. In a related approach, Jochim & Arora (2009, 2010) identified examples of the binding of small molecules to grooves, which accommodate interactions with helices in PPIs. Common features of the binding subsites that make them tractable for small molecules were deduced and potential candidates for synthetic ligands in other proteins were identified. These were subsequently used to rank potentially druggable sites.

Complementing the computational approaches mentioned above, Hajduk and co-workers showed that the hit rate from an NMR-based fragment screen is a good indicator of the druggability of a target. In their study, fragment hits were found for almost all targets (which included several protein–protein binding sites), but a hit rate of $\geqslant 0.1\%$ was found to indicate druggability (Hajduk *et al.* 2005). In a recent study, Edfeldt *et al.* (2011) incorporated hit rate, best

affinity and hit diversity into a traffic light ligandability score. They showed that fragment-screening-based ligandability was a very good indicator of success in high-throughput screening (HTS) and likelihood of hit-to-lead project progression for 36 AstraZeneca targets. Almost all of the 36 targets were enzymes or receptors. No mention is made of protein–protein targets explicitly in the paper, but 'interleukin' and 'G-protein-coupled receptor (GPCR) extracellular domain' were both confirmed as being protein/peptide–protein interaction targets (Edfeldt, F. N. B., personal communication). Both targets had low ligandability, failed HTS and were not taken forward to hit-to-lead optimisation.

It is clear that a redefinition of druggability or ligandability of PPIs is needed. We require less focus on large cavities and more on defining shallow grooves and small pockets that might provide anchors, as shown by the pocket analysis of Rad51 described above. Such computational approaches can be advantageously combined with experimental approaches to defining ligandability, especially by observing experimentally the frequency of fragment binding by X-ray or NMR approaches (see section 3.2.2 below).

2.5 Hotspots

Pioneering work in 1995 by Clackson and Wells on human growth hormone and its receptor demonstrated that a small subset of residues involved in the PPI contributes most of the binding free energy (Clackson & Wells, 1995) and more extensive alanine-scanning mutagenesis studies revealed that this was a general phenomenon (Bogan & Thorn, 1998). Although the data, which have been organised in databases, most notably Alanine Scanning Energetics database (ASEdb; Thorn & Bogan, 2001), are neither extensive nor diverse, they are used most often to train models to predict hotspots. A systematic attempt to extend the data by mining the literature is recorded in the Binding Interface Database (BID; Fischer *et al.* 2003).

Bogan & Thorn (1998) found that hotspots often include tyrosine, arginine or tryptophan, but rarely leucine, serine, threonine or valine residues. Tryptophan and tyrosine residues can contribute to aromatic pi-interactions with cations and CH groups, establish weak H-bonds through aromatic hydrogens with carbonyls and form hydrogen bonds through the indole nitrogen of tryptophan and the phenolic hydroxyl of tyrosine groups; they also have large hydrophobic surfaces that can protect hydrogen bonds from water (Chakrabarti & Bhattacharyya, 2007). Arginine can also form a range of similar favourable interactions, in addition to ion pairs that force it to the periphery of an interacting interface. Aspartate and asparagine are favoured over glutamate and glutamine, presumably due to differences in side-chain conformational entropy. Energetically less important residues surrounding centrally located hotspots most likely occlude bulk solvent, leading to an O-ring model (Bogan & Thorn, 1998). Hotspots coincide with hubs (highly connected residues) at the PPIs and can be understood in terms of small-world networks where highly connected nodes keep the whole network connected (del Sol & O'Meara, 2005). Nussinov and co-workers identified 'coupling' and 'hot regions', where highly packed residues make it easier to remove water upon binding in agreement with the O-ring theory and the small-world network idea (Keskin *et al.* 2005). The 'double water exclusion' (DWE) hypothesis provides an alternative view point, in which the cluster of residues in the hotspot is described by a biclique subgraph where a vertex is used to represent a residue, an edge to indicate a close distance between two residues, and a bipartite graph to represent a pair of interacting proteins (Li & Liu, 2009).

Kortemme & Baker (2002) proposed the first useful physical model (Robetta) for detecting hotspots at PPIs by analysing energies of packing interactions, hydrogen bonds and solvation. Their 'computational hotspots' are well correlated with 'experimental hotspots' in ASEdb. Presaging Critical Residues in Protein Interfaces Database (PCRPi-DB) is a public repository that archives computationally annotated hotspots in protein complexes. The hotspot predictor combines a set of seven different measures that account for energetic, structural and evolutionary information into a common probabilistic framework by using Bayesian Networks based on training sets from ASEdb and BID (Segura & Fernandez-Fuentes, 2011). More recent predictions have been based on support vector machine (SVM) methods (Lise *et al.* 2011) where the structure of a complex is used to predict hotspot residues.

Hotspots can be seen as ligand-efficient chemical 'foot holds' from which established fragments can be evolved into lead-like molecules. Thus, a further definition of hotspots in PPIs is emerging from fragment screening, in which the location of binding sites for multiple small molecules defines the hotspot. Organic solvent mapping studies using X-ray crystallography (Fitzpatrick *et al.* 1993; Mattos & Ringe, 1996) and NMR (Liepinsh & Otting, 1997; Mattos & Ringe, 1996) have shown that small probe molecules such as isopropanol have a strong tendency to bind to particular pockets on the surface of proteins. This phenomenon is also observed for larger fragments using NMR methods (Hajduk *et al.* 2005). Ciulli *et al.* (2006) dissected a ligand into its constituent fragments to examine their uneven contribution to binding affinity.

Fragment-based approaches targeting the protein–protein interaction between the human recombinase RAD51 and the hub protein BRCA2 have shown that the small pocket that accommodates phenylalanine in RAD51, where the conserved FxxA motifs in BRCA2 bind (see Fig. 1), accumulates multiple fragments (Scott *et al.* 2012). This pocket can therefore be seen as a 'hotspot' in RAD51.

Wells and co-workers state that protein–protein interaction hotspots mean that fragments can find ligand-efficient 'footholds', but higher affinity is often achieved by binding 'cryptic' binding sites within the binding interface (Thanos *et al.* 2006; Wells & McClendon, 2007). It is clear that, where protein–protein hotspots and ligand-binding hotspots coincide, competitive small molecule inhibitors can be developed.

3. Fragment-based approaches to screening PPIs

The previous section has shown that PPIs present difficult targets in terms of current definitions of druggability. Furthermore, describing the binding of small molecules to PPIs can be complicated by conformational changes that sometimes occur during binding. On the positive side, targeting PPIs offers increased opportunities for identifying selective ligands for members of large protein superfamilies such as protein kinases or proteases and there is a clear need to find drugs that target those interfaces. Furthermore, as we have seen above in Section 2.5, the existence of hotspots encourages the view that small molecules may be able to modulate these interactions effectively.

A key question is how we can achieve this. HTS using drug-like molecules of 300–500 Da may be less effective with PPIs than for many targets, as libraries have been developed largely to explore the chemical space of GPCRs, proteases and kinases, and do not have representative sets that are proven to bind protein interfaces. Here, we discuss the use of fragment-based drug discovery, where small molecules of a molecular weight of less than 300 Da are used to probe the

chemical space of a protein target. As fragments are small, they are likely to bind to a more diverse set of protein targets and the library size does not need to be large. Efficient detection of low affinity binding has relied on the development of a range of assays, predominantly based on biophysical methods. As the size of chemical space grows exponentially with the number of atoms, a library, of say 2000 fragments, can cover a relatively large area of chemical space. This fact combined with the lower complexity as compared with larger molecules means that the chances of finding fragment hits are relatively high (Hann *et al.* 2001). Once fragment hits have been identified, characterising their binding mode and chemical evolution into hit compounds becomes the focus.

3.1 Fragment libraries

A library of fragments with molecular weights of 100–300 Da is used to explore chemical space of the target protein efficiently. A 'rule of three' for fragments (Congreve *et al.* 2003), equivalent to the 'rule of five' for drug-like molecules (Lipinski *et al.* 2001), defines an optimal molecular weight of less than 300 Da, less than three H-bond donors and acceptors, less than three rotatable bonds and *c*log *P* less than 3 (Congreve *et al.* 2003). This results in small and relatively rigid molecules that present good starting points for evolution into drug candidates. A major advantage of fragment-based drug discovery is the small size of the library used; many fragment libraries contain less than 1500 members *versus* drug-like compound libraries containing 100 000 to 2 million members (Congreve *et al.* 2008; Hajduk & Greer, 2007). Chemical companies have developed various fragment libraries that are rule-of-three compliant. For example, compounds in Maybridge libraries (Thermo Fisher Scientific: http://www.maybridge.com/) are selected from over 30 000 compounds and assembled into diverse libraries of up to 1500 members that comply with the rule of three. Some libraries contain specific bromo- and fluoro-fragment collections for NMR. Maybridge (www.maybridge.com) supply both fluorine (5300 members) and bromine-containing (1500 members) commercial libraries. KeyOrganics also supply fluorine-containing fragment library (1201 members: www.keyorganics.com). Ligand-based approach (LBA) targeted libraries and receptor-based approach (RBA) targeted libraries, for example against kinases, proteases, GPCRs and ion channels, are available (see for example http://homepage.mac.com/swain/Sites/CMC/DDResources/Hit_iden/frag_collection.html), but have often been kept in house by commercial companies. In a recent fragment-based approach against GPCRs, Heptares used a focused ZoBio (www.zobio.com) fragment library with molecular weights of 134–194 Da (Congreve *et al.* 2011). A fragments blog that is very up to date on the fragment libraries and development and which includes a list of other suppliers is to be found at http://practicalfragments.blogspot.com/search/label/fragment%20libraries.

Recently, with the objective of increasing the chemical space explored by the current fragment libraries, diversity-oriented synthesis (DOS) has been used to generate more complex, 3D-rich expanded set of fragments (Hung *et al.* 2011).

Fragments that bind are generally ligand efficient (Congreve *et al.* 2005), making several interactions per atom with the target protein, which is difficult to achieve with larger molecules. Kuntz *et al.* (1999) found that a binding affinity increase of up to 1.5 kcal per mol per heavy (non-hydrogen) atom is achievable for fragments up to 12 heavy atoms, but after this point the maximum gain decreases dramatically. However, as the fragment interactions have to compensate for loss of rotational and translational entropy and fewer interactions mean a lower enthalpic contribution, affinities tend to be low, often between 0.1 and 10 mM.

Little has been reported in the literature on targeting PPIs with fragment libraries. In our hands the Maybridge library has proved useful, giving multiple hits with a diverse set of PPIs, as we describe below (Scott *et al.* 2012; D. Coyne, personal communication), but a more chiral set of molecules would be helpful.

3.2 Strategies for screening fragments

3.2.1 Tethering

Wells and co-workers have used dynamic combinatorial chemistry (DCC) combined with fragment-based lead discovery to develop the 'tethering' approach. Reversible disulphide bonds between a cysteine residue in a protein and a thiol-containing fragment are used to capture otherwise weak-binding fragments and identify them by mass spectrometry (Erlanson *et al.* 2000). Cysteines are often engineered into proteins local to the target site. This method was later developed into 'tethering with dynamic extenders' where a small-molecule 'extender' irreversibly alkylates the cysteine residue of a protein and also provides a thiol group to capture fragments (Erlanson *et al.* 2003). However, the chemistry used in this approach is difficult and modification of the protein is a prerequisite for fragment screening. Furthermore, specific libraries composed of thiol-containing fragments are needed.

3.2.2 Screening fragments using biophysical methods

If tethering is not used, the low affinities often observed for fragment binding necessitate the use of relatively sensitive assays, in particular biophysical assays that can detect binding to one partner, where inhibition might not be unambiguously detectable in a bioassay. To attain sufficiently high-throughput, an initial coarse-grained screen is often used. Depending on the nature of the assay, this screen may provide some quantification of binding affinity or may serve simply to identify hits with a response above a given threshold. Hits can then be validated and characterised in greater detail but lower throughput, either using the same techniques with different experimental design or using additional techniques, to acquire a wider set of thermodynamic, kinetic and structural data to inform compound progression (see Fig. 2 for a general outline of the approach). We have found that this approach works well when targeting PPIs.

The choice of technique to use at each stage of this screening cascade is determined by the physical properties of the target protein, the level of throughput required and the equipment available to the researcher. For example, in the initial stage of screening, if soakable crystals can be obtained, a lot of information on fragment binding can be obtained by X-ray crystallography in the context of a high-throughput screen. If one binding partner is relatively robust so that it can be regenerated on-chip, SPR can be used to screen many compounds with a very small amount of protein. Ligand-observed NMR is very sensitive and can be used to screen cocktails of compounds without the need for deconvolution. The fluorescence-based thermal shift assay is a relatively inexpensive and rapid technique that requires only a target protein with an experimentally accessible thermal denaturation transition.

In the later stages of screening, involved in validating and characterising the hits (Fig. 2), more precise measurements of affinity can be obtained using ITC and SPR. Both techniques can be used to perform competition experiments with a known ligand to confirm that compounds compete for binding at a specific site. In addition, ITC directly measures the change of enthalpy

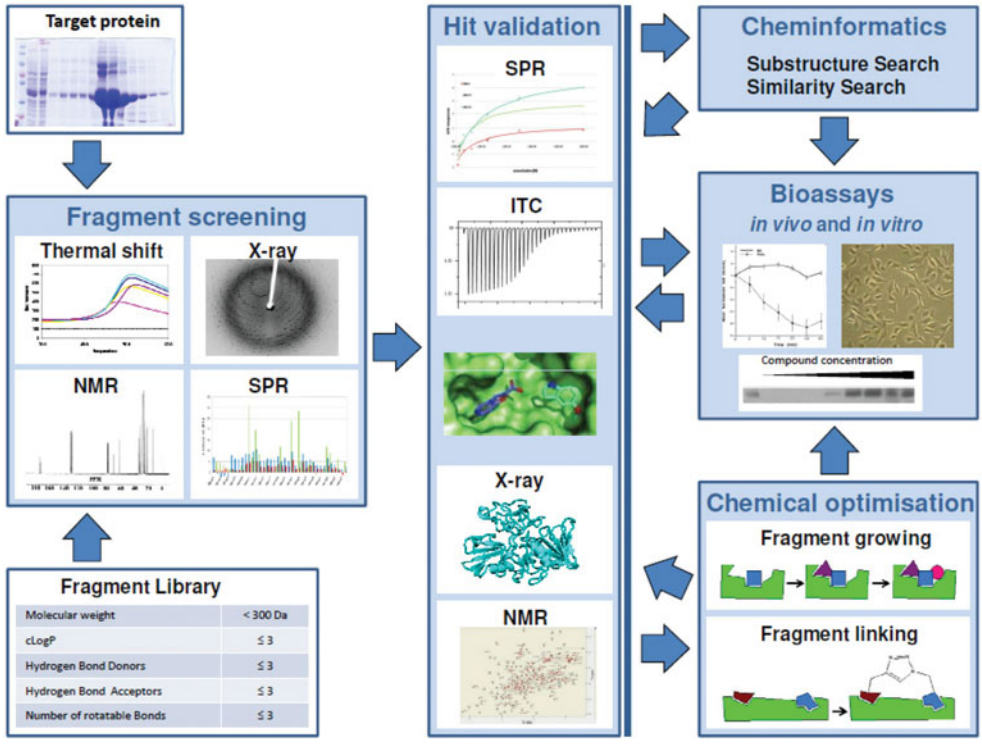

Fig. 2. Strategy for a fragment-based screening campaign: methods for initial high-throughput fragment screening include thermal shift (DSF), SPR, ligand-based NMR and X-ray crystallography. Validation of potential fragment hits can be achieved by nanospray mass spectrometry, the kinetic parameters such as association and dissociation rates defined using SPR, and the major thermodynamic parameters derived using ITC. X-ray crystallography and NMR are used to define binding modes of the fragment and to guide fragment optimisation or purchase of related compounds (analogue search). Assessment of biological activity assists in the chemical optimisation of the fragments on their way to a drug lead.

upon binding from which the change of entropy upon binding can be calculated, and SPR directly measures the rate constants for association and dissociation of the protein–fragment complex. At this point in a fragment-screening campaign, structural data obtained by X-ray crystallography or multi-dimensional heteronucleus-edited NMR spectroscopy can give invaluable insight into compound orientation and binding mode. Structure–activity relationships (SAR) can be rationalised at an atomic level and used to guide chemical synthesis of new compounds. These techniques are the most commonly used for detection and quantification of fragment binding and are described in detail later in this review. Many other biophysical techniques can be used at each stage of screening with appropriate experimental design. Examples include fluorescence resonance energy transfer (FRET) and the related amplified luminescent proximity homogeneous assay (ALPHA), fluorescence polarisation (FP) or fluorescence anisotropy (FA), and even single-molecule techniques such as fluorescence correlation spectroscopy.

Affinities (K_D) measured by any of the techniques mentioned above can be used to calculate the free energy of binding and so to calculate ligand efficiency (LE), which is a useful tool to monitor the effectiveness of compound optimisation. LE is defined as the free energy (ΔG) of binding of a ligand for a specific protein averaged for each heavy atom (non-hydrogen atom) (Hopkins *et al.* 2004). Recently, group efficiencies (GE) have been used in fragment growing and

optimisation campaigns (Saxty *et al.* 2007). GE allow the estimation of an individual group's contribution towards the overall free energy of binding (ΔG), giving a quick and simple insight into how efficient one modification is over another.

Fragment hits are evolved into larger compounds by optimising and adding groups to pick up additional PLIs that result in improved affinity for the target. During the optimisation process, it is often useful to evaluate the compounds not only by their physicochemical parameters but also by their biochemical and biological activity. Enzyme assays and cell-based assays are now widely used to assess a compound's activity at an early stage of compound optimisation.

Most of the experience of fragment-based approaches has been gained from conventional targets, including kinases and proteases. Here, we review attempts to target PPIs based on these now well-established methods.

3.3 Production of soluble proteins using protein engineering

Since many biophysical assays, for example X-ray crystallography, NMR and SPR, require high concentrations of proteins to screen for weak-binding affinities, fragment-based approaches require significant quantities of pure, soluble recombinant protein. It is important that the binding site of interest is free from endogenous ligands and, in the case of X-ray screening, is not occluded by crystal contacts allowing it to bind fragments.

PPIs are frequently extensive (>2000 Å^2), and exposure of the hydrophobic regions of the binding site may compromise stability, solubility and monodispersity of the protein sample. These characteristics mitigate against HTS. Any tendency of the target protein to aggregate when isolated from the multi-protein complex should be minimised. This can sometimes be achieved by engineering the protein target.

For targets that self-associate or form very large polydisperse macromolecular complexes, complexity or heterogeneity in the order of association can complicate biophysical analysis. In such cases, the production of a simplified surrogate system (e.g. monomerised protein) may simplify the measurement of binding constants and other thermodynamic parameters.

Some screening techniques such as FP, FRET and NMR can be facilitated by recombinant expression of target proteins containing fluorescent, isotope-labelled or spin-labelled non-natural amino acid analogues (Liu & Schultz, 2010; Wang & Schultz, 2004; Young & Schultz, 2010). Such an expansion of the genetic code for protein expression offers the chance to place reporters at specific sites within interaction interfaces thereby analysing specificity as well as affinity of binding.

Hyvönen and co-workers (Scott *et al.* 2012) have created an extensive toolbox for fragment-based drug discovery against the RAD51–BRCA2 interaction. Human RAD51 cannot be expressed in a stable, monomeric, unliganded form. RadA, the archaeal homologue of RAD51, is therefore used as a surrogate protein, and a number of mutants of RadA have been engineered that are unable to self-associate, are free of protein or peptide ligands, and have humanised interaction surfaces around the binding site. With respect to creating new crystallographic forms for screening, crystal packing interaction sites have been targeted for mutagenesis, successfully encouraging other packing arrangements with binding sites open for soaking. By screening a fragment library against these humanised proteins and validating the fragment hits biophysically and structurally, Hyvönen and co-workers have shown that this protein–protein interaction site is amenable to small molecule intervention.

3.4 HTS of fragment libraries

As described above, fragments are often initially screened using a fluorescence-based thermal shift assay (also known as DSF), SPR, ligand-based NMR or X-ray crystallographic soaking of cocktails. Here, we describe the different approaches.

3.4.1 Fluorescence-based thermal shift assays

The fluorescence-based thermal shift assay is a rapid and sensitive tool for monitoring thermostability of proteins. Temperature-induced protein denaturation is monitored using an environmentally sensitive dye such as Sypro Orange (Invitrogen) that fluoresces in the environment of the hydrophobic core as the protein unfolds (Pantoliano *et al.* 2001; Holdgate & Ward, 2005; Leung *et al.* 1996). The interaction between ligands and proteins usually induces changes in protein thermal stability with modifications in the midpoint denaturation temperature, enthalpy of unfolding and heat capacity. These modifications are due to the coupling of unfolding with binding equilibrium and can be attained by changes in protein structure and conformational flexibility induced by ligand interaction. Hence, an increased thermal stability of the protein–ligand complex corresponds to a reduced conformational flexibility of the protein or a stabilisation. These effects were studied in detail by Celej *et al.* using bovine serum albumin (BSA) interacting with three different anilinonaphthalene sulphonate (ANS) derivatives and employing different biophysical techniques (Celej *et al.* 2003). The increase and decrease of protein thermal stability upon ligand binding and influence of ligand concentration and affinity were recently described in a quantitative model by Cimmperman *et al.* (2008).

Using β-site amyloid precursor protein-cleaving enzyme 1 (BACE1) as a test system, Lo *et al.* tested 13 compounds and determined their unfolding temperature (ΔT_m) (Lo *et al.* 2004). Using the shift of the unfolding temperature obtained in the presence of ligands relative to that obtained in the absence of ligands, they assessed the ligand-binding affinity at T_m and compared it with values obtained from ITC, which proved to correlate well. However, one should be cautious in extracting further information such as the potency ranking order and the magnitude of binding constants from the ΔT_m values because the same binding affinity can give rise to different values of ΔT_m (and *vice versa*) depending on the enthalpy, entropy and change in heat capacity of binding and protein folding (Brandts & Lin, 1990; Holdgate & Ward, 2005). Furthermore, this study demonstrates that positive hits can be identified only if thermal shift assays are conducted at appropriate compound concentrations relative to their dissociation constants, i.e. at least 2–3 times K_D of the compound. Vedadi and co-workers determined the thermal stability of 61 recombinant proteins in the presence and absence of a range of physiologically relevant compounds with the aim of finding a stabilising compound that would aid purification and crystallography. No T_m could be determined for 17 proteins. They found that T_m shifts of more than 4 K translate into values for IC_{50} of better than 1 μM (Holdgate & Ward, 2005; Vedadi *et al.* 2006; Waldron & Murphy, 2003).

The fluorescence-based thermal shift assay is unsuitable for proteins that do not show a well-defined sigmoidal unfolding curve in an unliganded state. Proteins with hydrophobic binding pockets or cavities accessible to the dye tend to display an atypically high initial fluorescence. Other difficulties might arise from poorly folded proteins, proteins with high thermal stability or from multiple, independently unfolding domains (Vedadi *et al.* 2006). Additionally, since aggregation of a denatured protein buries hydrophobic surface area, it causes the

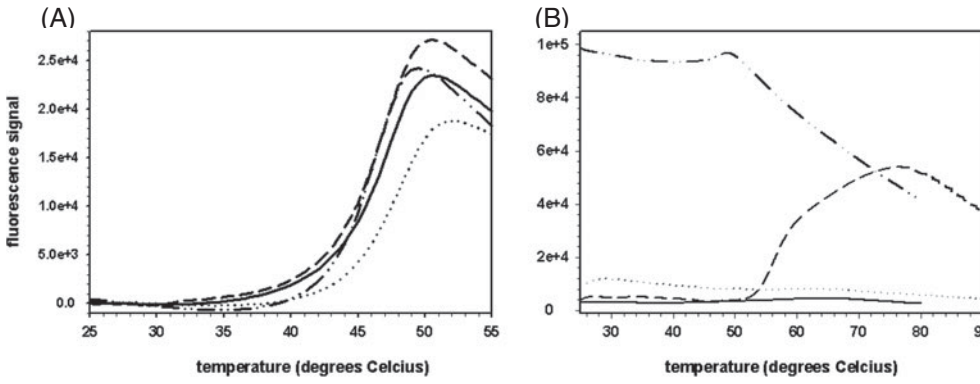

Fig. 3. Thermal shift curves of Met and NK1. (A) Fragments interacting with Met cause a shift in melting temperatures. Compounds 1 (— — —) and 2 (— ·· —) cause a slight destabilisation of the protein as compared with control (——), whereas compound 3 (·····) stabilises the protein. (B) Melting curves of NK1 obtained using different buffers at varying pH (30 mM phosphate, pH 7, 150 mM NaCl (——), 50 mM Bis Tris, pH 6.7 (·····), CAPS, pH 10.5 (— — —) indicate complex melting processes, possibly related to the presence of two domains and disulphide bridges, as well as a tendency to dimerise through domain swapping. One of the curves (— ·· —) shows a fragment that interacts with the fluorescent dye.

fluorescent signal to decrease; where this occurs data must be excluded from the analysis. In some cases, compounds themselves can have an influence on the outcome of this assay. Highly fluorescent compounds or compounds that bind both the native and unfolded states of the protein are not suitable for this assay.

Figure 3 shows unfolding curves for two proteins that interact with each other, an example of targeting protein–protein interactions: the Met receptor shows a well-defined apo-unfolding curve and was used for fragment screening whereas NK1 (comprising the first two domains of HGF) does not show a suitable unfolding curve, probably because it has two independent domains, one of which has many disulphides, as well as a tendency to dimerise through domain swapping (Sigurdardottir, 2012).

The fluorescence-based thermal shift assay is the cheapest of those described in this review – both in initial equipment cost and in consumables used per experiment. However, relatively high protein and fragment consumption can be limiting factors in some projects. Furthermore, this technique is the least sensitive of the techniques presented here confirming only 8–23% of NMR hits against three kinase targets (for examples, see Hubbard & Murray, 2011) indicating that binders in millimolar range may be difficult to detect.

However, the simplicity and general applicability of the thermal shift assay still make it attractive for identification of fragment hits. Thus, for targeting the PPI in the Notch ankyrin domain Abdel-Rahman *et al.* used a library of 1201 rule-of-three-compliant fragments, mainly from the Maybridge library that was not target-tailored but was reasonably diverse (Abdel-Rahman *et al.* 2011). The fluorescent-based thermal shift assay identified 36 positive hits with a shift in thermal unfolding temperature (T_m) of 0.5 K or more, an impressive outcome when the protein interface is relatively featureless and flat.

3.4.2 Surface plasmon resonance (SPR)

SPR is a label-free optical biosensing technique that measures adsorption of material onto planar metal (typically gold) surfaces. The SPR method is based on optical measurement

of refractive index changes associated with the binding of analyte molecules to molecules immobilised on the SPR sensor surface (Jonsson *et al.* 1991). Its low protein and compound (analyte) consumption, and high sensitivity make it invaluable for fragment-based drug design. Case studies demonstrate the successful identification of selective low molecular weight inhibitors for pharmacologically relevant drug targets through the SPR screening of fragment libraries (Giannetti, 2011). Fragment hits as weak as 5 mM can be detected in short time frames, with screening of a 5000-membered library and the first stage of hit validation possible in one month. Fragment concentrations used in a screening campaign usually range from 100 μM to 1 mM (Giannetti, 2011). In addition, most SPR instruments are multiplexed allowing multiple targets to be screened simultaneously, thereby increasing turnaround speed. Proteins and other biomolecules can be immobilised on a variety of chip surfaces leaving room for individual requirements in assay design. However, great care must be taken in data correction, scaling and normalisation (Giannetti, 2011). Additionally, solvent effects need to be taken into account, especially if dimethyl sulphoxide (DMSO) is used to solubilise the fragments in the fragment library (Navratilova & Myszka, 2006).

It is helpful to use a known binder as means of identifying over-stoichiometric binders often caused by fragment stacking or other self-interactions. High compound concentrations can result in signals arising from a number of secondary effects, for example non-specific binding (i.e., poorly defined interactions to a single site) and methodological limitations (e.g. carryover, insufficient correction for buffer-related signals and drifting baselines) (Geitmann *et al.* 2011). Screening against different proteins (e.g. mutants with disabled active sites) enables identification of promiscuous binders and ensures specificity against the target protein (Hamalainen *et al.* 2008). A stepwise hit selection and validation protocol is useful for compensating for these secondary effects. For example, a primary screen with relatively generous response cut-off criteria could be used in the first instance to detect binders. This screen should, if possible, highlight promiscuous binders by using one or more suitable reference proteins. These preliminary hits could then be applied to a second screen where competition against a known binder is assessed or the K_D is determined. Using a series of screens, promiscuous binders, fragment stacking or false positives can be easily identified.

The literature provides several examples where SPR has successfully been used as the screening method (Elinder *et al.* 2011; Giannetti, 2011; Perspicace *et al.* 2009; Xiang *et al.* 2011). In our hands, SPR is more sensitive than the thermal shift assay for identifying hits in millimolar range. The affinity range that can be reliably determined is limited primarily by the aqueous solubility of the fragment. Kobayashi and co-workers compared SPR with solution ligand-observed NMR in fragment screening. They found that SPR detection is most suited to ligands that bind with K_D better than 1 mM. However, they concluded that the good correlation between SPR and potency in a bioassay makes this a good method for hit validation and characterisation studies (Kobayashi *et al.* 2010).

Summarising different studies at Vernalis, Hubbard and Murray reported that 84–100% of NMR hits of five different targets were confirmed in SPR with affinities ranging from 0.1 to 1000 μM (Hubbard & Murray, 2011). Given the screening speed, sensitivity and ability to drive SAR down to biochemical potency, SPR methods represent a robust approach to initiating and supporting a fragment-based lead identification, but additional experiments such as protein-observed NMR or X-ray crystallography need to be performed to determine structural information about the protein–fragment complex.

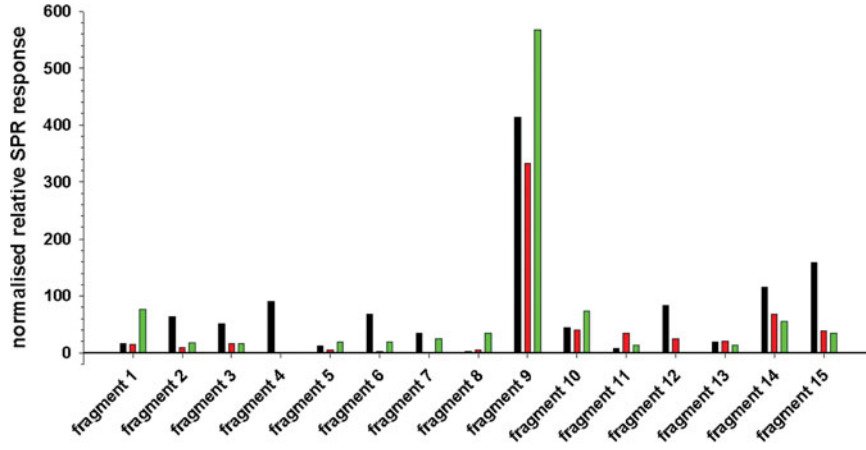

Fig. 4. SPR screening. Binding levels of 15 fragments from fragment screening showing hits, over-stoichiometric binders and promiscuous binders to NK1 (green), Met (black) and SPH domain (red).

We have used the approach to target successfully protein–protein interactions with the aim of interrupting the binding between hepatocyte growth factor (HGF)/scatter factor (SF) and Met receptor (Jubb *et al.* 2012; Sigurdardottir, 2012). Protein consumption could be reduced 5000-fold using SPR as compared with using thermal shift assay for fragment screening. A fragment library of ∼1300 members was screened against NK1, SPH-domain and Met receptor at concentrations of 0.5 or 1 mM. Promiscuous binders were easily identified by comparing binding levels of each fragment against each target (see Fig. 4). Promiscuous binders show binding to all three proteins and over-stoichiometric binders or aggregating fragments can be identified by extremely high responses. In Figure 4, fragments 1 and 8 are specific for NK1, fragments 2, 3, 4, 6, 12 and 15 are specific for Met and fragment 11 is specific for SPH. Fragment 9 is probably aggregating on the chip surface and might be omitted from the library in future. Fragments 5, 7, 10, 13 and 14 bind to all proteins equally well or do not bind at all. These fragments are probably promiscuous binders, which need to be excluded from further experiments. Hit rates for each protein were around 5–10%, which is consistent with results from other screening campaigns against more conventional targets (Giannetti, 2011; Hubbard & Murray, 2011; Perspicace *et al.* 2009). Only fragment hits that showed clear selectivity for one target were taken forward for hit validation. After employing this selectivity filter, hit rates for each target protein were 0.1% (SPH), 5% (NK1) and 8% (Met). These data are encouraging and indicate that SPR could find wide applicability for screening PPIs.

3.4.3 NMR (ligand-observed)

NMR spectroscopy has been very popular as an initial screening method and was the first experimental method used for screening fragments (Shuker *et al.* 1996). NMR methods for fragment screening are classified as either ligand- or receptor-observed methods, depending on whether ligand binding is monitored through ligand or receptor resonances. Generally ligand-based methods are used for primary screening and receptor-based methods for mapping the binding site and further development. Therefore, receptor-based methods are discussed in detail in Section 3.5.4.

Ligand-based methods detect changes in the characteristics of the ligand (such as relaxation and diffusion) that occur when it transiently binds to the receptor (Lepre, 2011). The typical ligand-observed method compares the NMR spectrum of a mixture of ligands in the presence and absence of a protein. This approach renders the molecular weight of the protein irrelevant. Additionally, there is no need to produce isotope-labelled protein. Currently, the most widely used ligand-based methods are saturation transfer difference (STD; Mayer & Meyer, 1999) and water-ligand observed via gradient spectroscopy (WaterLOGSY; Dalvit *et al.* 2000). STD relies on intermolecular magnetisation transfer directly via the protein, whereas WaterLOGSY experiments utilise the transfer via the bulk water for the detection of molecules interacting with the receptor. These methods consume very little protein and are capable of detecting ligands with affinities ranging from 10^{-8} to 10^{-3} M (Lepre, 2011). Another useful screening technique is 1D ^{1}H relaxation experiments (Hajduk *et al.* 1997), which detect small molecule binding by comparing line shapes in the presence or absence of receptor. Binding-induced ligand transverse autorelaxation rate ($R_2 = 1/T_2$) enhancements may be visible as simple broadening of proton resonance lines upon the addition of receptor (Lepre *et al.* 2004). Initial ligand-based NMR screening is usually carried out with cocktails containing a mixture of fragments thereby decreasing the amount of protein needed and data collection time. If changes in the ligand spectrum are detected, single fragments from that cocktail are then analysed separately (so-called deconvolution). Cocktails usually contain 8–10 fragments and are designed to avoid chemical reactions between the fragments and overlapping resonance spectra. A competition-binding step during NMR screening has been used to test the specificity of an identified ligand and for extracting the dissociation-binding constant using titration experiments.

The clear advantage of ligand-based NMR methods is that they require only soluble protein; there is no need to crystallise, immobilise or tag the target protein. It is possible to detect fragment binding over a very broad affinity range making this technique suitable for weak binders up to a K_D of 1 mM (Kobayashi *et al.* 2010). However, this technique is susceptible to false positives arising from fragments that aggregate in solution, but this can be detected by running control experiments. NMR also consumes relatively large amounts of reagents compared with other methods. To screen a library of 1200 fragments, about 30–40 mg of protein is required which is considerably more than needed for SPR. Additionally, NMR has lower throughput than biochemical HTS. However, when fragment cocktails are used for screening, the throughput can be increased substantially.

Many examples have been reported for classical targets. For example, Stockman *et al.* (2009) conducted a study to identify small, ligand-efficient fragments that might inhibit the 3-phosphoinositide-dependent kinase-1 (PDK1) in the phosphoinositide 3-kinase (PI3K) signalling cascade. Using STD NMR experiments, 372 fragment hits were identified from a fragment library containing 10 237 compounds. In a second round of experiments, fragments were investigated for binding either the ATP (adenosine-5′-triphosphate) site or the allosteric PDK1 interacting fragment (PIF) pocket of PDK-1. In a second example, WaterLOGSY experiments were performed to screen a 2000-compound general fragment library against the aspartic protease β-secretase (BACE1), which is a well-established target for the treatment of Alzheimer's disease. Hits were further characterised using SPR, and a series of six-substituted isocytosines could be identified as a novel scaffold for BACE-1 inhibitors (Geschwindner *et al.* 2007).

Ligand-based NMR has also been used successfully for targeting protein–protein interactions, for example with B-cell leukaemia/lymphoma X_L (Bcl-X_L), which inhibits apoptosis by binding a 16-residue α-helical portion of the pro-apoptotic molecule Bcl-2 homologous antagonist/killer

(BAK) or a 26-residue α-helical portion of another pro-apoptotic molecule Bcl-2-associated death (BAD). Two different fragments were shown to bind simultaneously to Bcl-X_L and were linked to a single molecule (Oltersdorf *et al.* 2005). We describe this successful campaign at targeting PPIs in Section 5.

3.4.4 X-ray crystallography (Screening)

X-ray crystallography is an information-rich screening technique, which is widely used in fragment-based drug discovery (Blundell *et al.* 2002; Hartshorn *et al.* 2005; Hubbard *et al.* 2007; Nienaber *et al.* 2000; Verlinde *et al.* 1997). In addition to allowing the detection of binding hits, it provides information about both the site and mode of binding in a single experiment. It is now an established part of several drug development platforms for example the Pyramid at Astex Therapeutics (Davies *et al.* 2006), SeeDs approach at Vernalis (Hubbard *et al.* 2007) and at Johnson and Johnson (Hubbard *et al.* 2007; Spurlino, 2011).

The main technical challenge of the method is to obtain crystal structures with different compounds using either crystal soaking or co-crystallisation. At the screening stage, soaking is preferable as it provides higher throughput and enables a higher concentration of compound to be used. Indeed, a high compound concentration allows the characterisation of even weak-binding fragments (Carr & Jhoti, 2002). In practice, the maximal fragment concentration is limited only by solubility and the tolerance of the crystals for fragment and co-solvent. The most common co-solvent used to maintain fragment solubility in the soaking solution is DMSO, so the crystal form used must be stable in the presence of DMSO.

In the simplest case, successful soaking requires a crystal form where the binding site is unoccupied by ligand and unobstructed by crystal contacts. Targeting PPIs offers new challenges to the use of X-ray crystallographic screening. The fact that PPIs evolved to bind to proteins suggests that they are likely to form crystal contacts if their natural partner is absent. This along with the additional difficulties in obtaining soluble unliganded protein for PPI targets, as described in Section 3.3 of this review, makes obtaining a suitable crystal form for screening more difficult. A technique that we have found particularly useful for targeting the RAD51–BRCA2 interaction is microseed matrix screening combined with cross seeding between different mutants (D'Arcy *et al.* 2007). This has enabled us to both crystallise previously un-crystallisable mutants and to engineer crystal forms with unobstructed binding sites to be optimal for soaking. The crystals must be readily reproducible in order to obtain optimal results because screening a library of fragments will require 100–1000 s of soaks depending on whether fragment cocktails are used and the number of fragments in one cocktail.

At Astex Therapeutics, an average of four fragments is used per soak with the fragments selected to be as chemically diverse as possible within a particular cocktail (Davies *et al.* 2006; Hartshorn *et al.* 2005). This both minimises the chance of more than one fragment binding per soak and eases de-convolution if more than one fragment binds. A different approach, adopted by Johnson and Johnson (Spurlino, 2011), is to assemble fragments into groups of five similarly shaped compounds. They maintain that this approach increases the chances of multiple fragment hits in one cocktail and reinforces areas of density where multiple fragments are bound. In this case, the hits are de-convoluted by carrying out multiple rounds of refinement with each different fragment from the cocktail as a starting point. However, in our experience, it is desirable to carry out single-molecule soaks for all cocktails where binding is observed.

Screening a large number of crystals requires the acquisition and processing of a concomitantly large number of datasets. In the past, this would have presented a significant obstacle. However, technical improvements have expedited all stages from crystallisation to structure determination. Advances in laboratory X-ray sources now allow efficient in-house data collection. Our in-house X-ray diffraction system (X8 PROTEUM from Bruker (www.bruker.com)) routinely enables datasets to be collected in less than 2 h and up to a maximum resolution of 1.5 Å. Improvements in automated crystal mounting in synchrotron sources not only allow data to be collected remotely (Gabadinho *et al.* 2008), but can improve the efficiency of data collection (Muchmore *et al.* 2000), especially by using simple scripting methods to increase automation.

Typical hit rates observed in fragment screening against conventional targets range from 0.5 to 10%. Thus, although the fragment libraries contain only 500—1000 compounds, they are able to sample significant chemical space (Hartshorn *et al.* 2005). However, X-ray crystallography has the disadvantage of screening fragments in a protein environment that may be partially constrained by crystal contacts, whereas NMR retains flexibility of protein side chains and conformational changes upon binding of the fragment.

In practice, to increase efficiency of the initial screening process, a pre-screening step is often implemented before screening potential hits with X-ray crystallography (Murray & Blundell, 2010). However, screening all fragments of the library with X-ray crystallography offers the advantage of evaluating fragment hits based solely on their binding mode and not only on potency. For example, a weaker potency fragment whose orientation offers more opportunity for fragment growing could be preferred over a tighter binding fragment that would be more difficult to develop.

Fragment-based approaches are proving successful in several protein–protein interaction campaigns. An example has been the interface between the human recombinase RAD51 and the hub protein BRCA2 (see Section 2.3 above). Figure 1 shows that the interface on the RAD51 is fairly flat, but there are distinct pockets for the phenylalanine and alanine side chains that bind the FxxA motif. Eighty-one of 1338 fragments screened using thermal shift showed induced shifts greater than 0.8 K and were selected for further analysis (Scott *et al.* 2012). Of these, 54 bound almost entirely in the pocket that binds Phe. Figure 5 shows that the chemistry and interactions of the fragments are rather more varied than one might assume from the selectivity of binding of the Phe in the BRCA2 sequence (see sequences of binding motifs in Fig. 1). The fragments exploit both the lipophilic nature of the pocket as well as opportunities to form well-defined H-bonds, particularly with the main chain carbonyl function of Leu 214 and the nearby side chain of Gln 217. The affinities are in the range of 400–1000 μM. This shows that small side chain pockets in PPIs may act as hotspots that undergo little conformational rearrangement on ligand binding.

3.5 Validation of fragment binding

Once high-throughput screens have been completed, then the fragment binding needs to be validated and characterised by careful analysis of the thermodynamics, kinetics and structure.

3.5.1 Thermodynamics of fragment binding: isothermal titration calorimetry

ITC measures the heat evolved or absorbed by all chemical processes that occur when one component of a reaction mixture is titrated into a solution of another component at constant

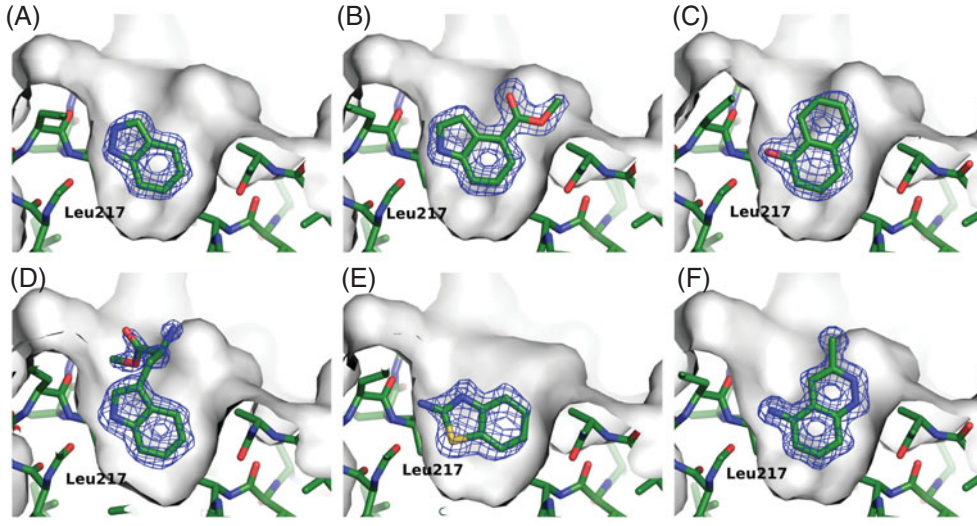

Fig. 5. Fragments binding to the Phe pocket of the RAD51 interface with the BRC4 peptide of BRCA2. The fragments bind with the following K_Ds (A) 540 μM (B) 1000 μM (C) 600 μM (D) 730 μM (E) 430 μM (F) 460 μM (from Scott *et al.* 2012).

temperature (Wiseman *et al.* 1989). It can be used to monitor any process that has an associated change in enthalpy (ΔH). In the case of macromolecular association, the enthalpy of binding, which can be exothermic or endothermic, is measured as one partner is titrated against another. When the heat of dilution has been accounted for, the heat of binding per mole of injectant decreases in magnitude as the fractional saturation of the complex increases. For the simplest case of a 1:1 association, the given curve can be fitted to a modified Langmuir isotherm (Hill equation) to obtain a value for the equilibrium dissociation constant K_D and the stoichiometric ratio (n). The entropy of binding (ΔS) can be calculated from the Gibbs free energy of binding (ΔG, calculated from the K_D) and the measured enthalpy of binding (Leavitt & Freire, 2001; Pierce *et al.* 1999). Accurate fitting of ITC data is dependent on careful experimental design. Both the concentration of the binding partner in the ITC cell and the molar excess of the titrant in the syringe can be manipulated to give data that can be reliably interpreted for a wide range of dissociation constants (nM–mM) and binding enthalpies (Turnbull & Daranas, 2003; Wiseman *et al.* 1989). Although traditionally regarded as low throughput, data acquisition has been expedited by the introduction of commercial small volume ITC instruments that are amenable to automation. In the future, the development of arrays of nano-calorimeters could allow the collection of ITC data in a high throughput screen (Torres *et al.* 2010).

The concentration of protein used typically ranges from 5 to 100 μM in a cell volume of 200–1800 μl, depending on the enthalpy and stoichiometry of binding. The titrant is typically at a 10-fold molar excess for tight binding (nM to low μM), with a greater excess being required for association stoichiometries greater than 1:1 or for weaker binding titrants.

In line with other binding techniques, ITC may be used to monitor competition experiments where one ligand displaces another from a binding site on a protein. A general solution for the fitting of competitive binding of two ligands (Wang, 1995) has been modified to fit ITC data (Sigurskjold, 2000). These experiments ensure that a putative competitive inhibitor is binding to the expected site on a target protein and also can be of particular use measuring extremely tight

association (low K_D) for which simple binding data could not be reliably fitted (Velazquez-Campoy *et al.* 2000, 2001).

One of the earliest reported applications of ITC to fragment-based drug discovery was to probe hot spots at the co-factor binding site of the *E. coli* ketopantoate reductase using fragments derived from the natural ligand, nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) (Ciulli *et al.* 2006). The binding affinities and thermodynamics of the fragments were measured by ITC and the fragment binding sites were located by a combination of site-directed mutagenesis and ITC. In a second example, ITC in competition with a natural phosphopeptide ligand was used to characterise the binding affinity and thermodynamics of fragments found by virtual screening and ligand-directed NMR to bind to the phosphotyrosine-binding pocket of the viral oncoprotein viral-sarcoma (v-Src) (Taylor *et al.* 2007). Recently, Edink *et al.* (2011) optimised fragment hits toward high-affinity lead compounds that bind to the acetylcholine-binding protein (AChBP). Analysis of fragment binding by ITC contributed to structure-based optimisation of the fragments. Comparison of ligand efficiencies and thermodynamic data from different compounds and protein mutants revealed interactions with a ligand-induced AChBP subpocket. Induced changes in the pocket were shown to be important for selectivity and may be of use in the design of subtype-selective ligands for human nicotinic receptors. In a similar study, Hung *et al.* (2009) used a fragment-growing approach of an initial fragment hit, 5-methoxyindole, to enhance binding to pantothenate synthetase (PS) for development of new tuberculosis therapeutics. 5-methoxyindole was found to bind with low affinity of 1.1 mM and binding could be enhanced 1000-fold by elaborating the original fragment scaffold.

ITC data can also be used in a more general fashion to inform and direct compound progression. Since the value of ΔH is directly measured in an ITC experiment, and the value of ΔS is calculated using equilibrium constant that is fitted independent of the enthalpy, those parameters are less prone to correlated errors than estimates from van't Hoff analysis of binding monitored by other techniques. Therefore, ITC-derived thermodynamic data are better suited to statistical analysis and correlation with the physico-chemical properties of ligands and binding sites. The Structure/Calorimetry of Reported Protein Interactions Online (SCORPIO) database collated by the Ladbury group holds structural and calorimetric data on nearly 100 unique protein ligand complexes (Olsson *et al.* 2008). Analysis of these data reveals striking differences in the range of affinities and dominant thermodynamic contributions to binding energy for biological ligands compared with compounds from medicinal chemistry. Further analysis allows assessment of the relative importance of burial of polar and apolar surface area in determining ligand selectivity and affinity respectively, and undermines confidence in commonly assumed physicochemical explanations for increased affinity upon increasing compound lipophilicity.

ITC has been used extensively as part of the RAD51–BRC4 screening effort to monitor the thermodynamics of fragment and compound binding. Figure 6 shows examples of binding isotherms for a fragment, a peptide and a peptide fragment chimaera to humanised RadA.

3.5.2 Kinetics of fragment binding (SPR)

As we have seen in Section 3.4.2, SPR has become a valuable tool for screening fragment binding to drug targets. Here, we focus on the use of SPR to carry out kinetic analysis by determining the on and off rates and enabling calculation of the binding affinity constant (K_D) of the compound. Additionally, the method is also useful in competition studies to map the binding locus.

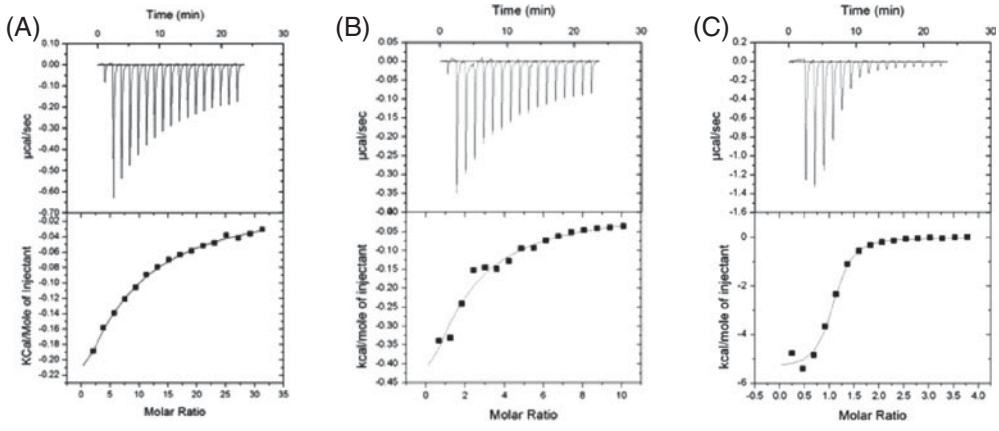

Fig. 6. Example binding isotherms of a fragment, a peptide and a peptide fragment chimaera to humanised RadA. (A) Example fragment titration with $K_D = 1.5$ mM. (B) Example peptide titration with $K_D = 280\ \mu$M. (C) Example peptide fragment chimaera $K_D = 3\ \mu$M (Scott *et al.* 2012).

There exists a wealth of studies described in the literature where SPR had been used to determine binding affinities for compounds and to monitor compound optimisation (Cole *et al.* 2010; Geitmann *et al.* 2011; Nordin *et al.* 2005). Most fragments show a squared response curve due to fast on and off rates with little or no curvature, and a 1:1 Langmuir-binding model cannot be fitted. In these cases, kinetic experiments are carried out using different concentrations of the compound and determining their respective equilibrium-binding levels. Response levels *versus* concentration are therefore fitted using a global fit that is able to determine steady-state affinity constants (Giannetti, 2011). In our experience, K_D values of a protein–protein-binding event obtained using both fitting methods vary slightly but are generally in good agreement. However, K_D values obtained by other methods such as ITC can differ because the protein in SPR is immobilised on a surface possibly introducing restrictions in flexibility or induced conformational changes upon ligand binding.

Kinetic studies have been used in our laboratory to determine the binding affinity of fragments targeting a PPI as a means of hit validation (Sigurdardottir, 2012; Jubb *et al.* 2012). After data correction, 13% of our initial fragment hits selective for Met were found to bind with an affinity of 100 μM or lower, and 34% were found to bind with affinities lower than 500 μM (see Fig. 7*a*). This amounts to 3% of the whole fragment library used. However, nearly 20% of the initial fragment hits did not show the expected binding curves but instead either a linear concentration dependency or no binding; they were therefore rejected from further analysis. This behaviour might be due to low solubility, aggregation or unwanted interaction with the chip surface, and these fragments are not considered hits. Figure 7*b* shows four examples of binding curves of fragments with different binding affinities.

SPR can also be used to characterise drug–drug interactions by evaluating binding of one fragment, while keeping the concentration of the other fragment constant. This type of competition assay can give valuable information about whether or not fragments bind to the same locus on a protein; this is especially useful in the absence of structural information. Results can be used to focus on medicinal chemistry efforts to either combine fragments that target neighbouring sites or to evolve two different fragment scaffolds independently.

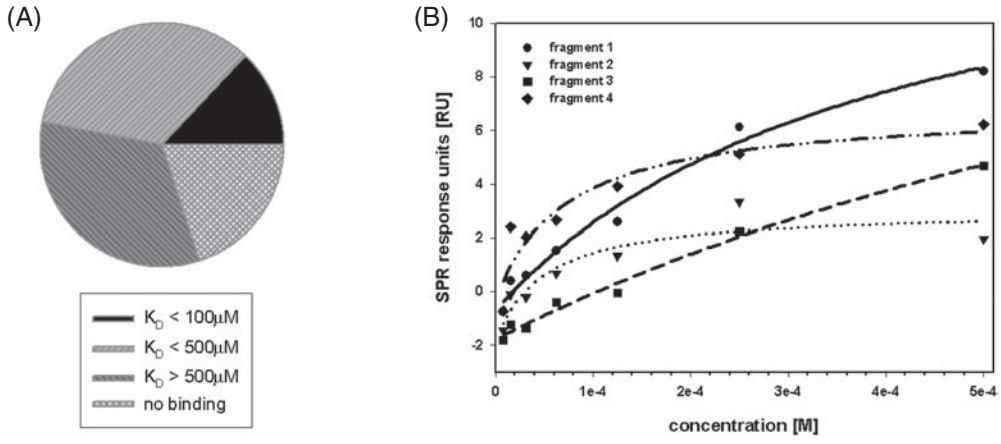

Fig. 7. Evaluation of fragment binding using SPR. (A) Distribution of steady-state binding constants (K_D) among fragment hits for Met. (B) Concentration-dependent binding of selected fragments to immobilised Met.

3.5.3 Structure (NMR: protein based)

NMR can be used to identify the binding site of a fragment hit to its protein target. When a ligand binds to a protein, the chemical shifts of both ligand and protein proton resonances are affected, but mostly nuclei located within the binding site (Meyer & Peters, 2003). Once the protein has been enriched with ^{15}N (or ^{13}C) and the chemical shift assignments are known, the protein is titrated with a ligand and chemical shift changes are monitored. 2D ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectroscopy is the most common method to monitor backbone chemical shifts. Only the protein resonances are visible in such spectra and, therefore, the ligand can be included at a high concentration (Homans, 2004). By comparing the HSQC spectra in the absence of a ligand with a similar spectrum acquired in the presence of a ligand, information on where the ligand is binding can be obtained. Additional information on binding constants and stoichiometry can be obtained by titrating the ligand into the protein sample and observing chemical shifts of selected residues.

HSQC measurements have been used to detect binding of some common buffer molecules to NK1 (Sigurdardottir, 2012). NK1 is a natural splice variant of HGF/SF and contains the high affinity binding site for Met. Buffers and some other common molecules are often found in crystal structures. The buffer molecules occupy the binding pocket itself, making it possible to use this information, combined with fragment information, to aid the drug discovery process. Figure 8*a* shows overlaid HSQC spectra of NK1 in the absence and presence of 4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid (HEPES) at one concentration. The dissociation constant K_D was calculated by plotting the weighted average ^{1}H and ^{15}N chemical shifts of the residues experiencing significant variations as a function of ligand concentration added during the titration and considering the one site binding model (Garrett *et al.* 1997). A K_D of 3.3 mM was calculated for HEPES.

Another example of the applicability of the HSQC spectrum to identify and validate binding modes of fragments for protein–protein targets is the perturbation of the ZipA–FtsZ complex (Tsao *et al.* 2006). Interaction of membrane-anchored zinc-regulated and ion-regulated transporter proteins (ZipA) with the cytosolic filamenting temperature-sensitive mutant Z (FtsZ)

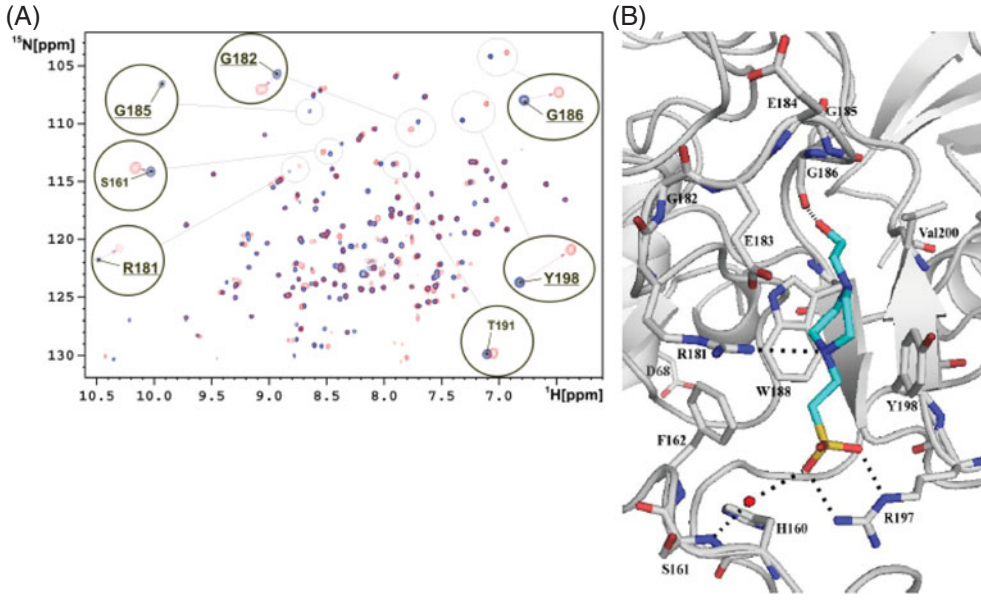

Fig. 8. Evaluation of HEPES binding to the lysine-like pocket of NK1. (A) Overlaid 2D ^{1}H-^{15}N HSQC spectra of 15N enriched NK1 in the absence (blue) and presence (red) of a ligand. (B) X-ray structure of the same compound obtained by soaking.

protein is required for bacterial cell division. Inhibitors of this complex might be a new antibiotic class that could help to overcome bacterial drug resistance. Less than 1000 fragments (with molecular weights of 200–250 Da with $c\log P < 2.5$) were screened in cocktails of six or fewer compounds by ^{1}H-^{15}N HSQC spectrum of ZipA. In total, 16 fragments were found to bind to ZipA, of which seven bound to the FtsZ-binding pocket.

Although protein-based NMR is not high-throughput it is an appealing method to validate binding modes of ligands because it is simple, easy to use and affordable. It is very sensitive and particularly valuable for screening weakly binding ligands that otherwise might be missed. It also gives information about where the ligands bind. In contrast, X-ray crystallography has the advantage of providing information on not only where but also how the ligand binds.

3.5.4 X-ray crystallography

Used as a validation tool, X-ray crystallography affords the opportunity to visualise compound binding in atomic detail. This information can then be used to monitor the progression of the compound at every iterative round of the optimisation process. The ability to use structural information dramatically increases the success of the fragment-based drug design process (Blundell *et al.* 2002; Hajduk & Greer, 2007; Hartshorn *et al.* 2005). The lower throughput requirements of a validation technique mean that the disadvantages of co-crystallisation, i.e. screening for new crystallisation conditions with each different compound are less problematic and therefore sustainable in a drug discovery campaign.

Fragment-based approaches are proving successful in targeting protein–protein interactions. An example has been the interface between the human recombinase RAD51 and the hub protein BRCA2. In this case, the analysis has progressed by two routes: the exploration of short peptides

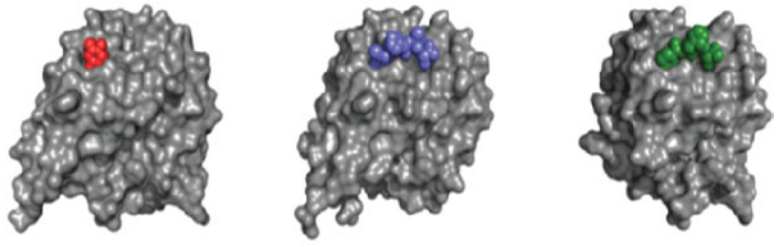

Fig. 9. Structures of humanised RadA bound to a fragment and peptides. X-ray crystal structures of humanised monomeric RadA bound to a fragment (red), peptide (blue) and a fragment/peptide chimaera (green). Protein (grey) and compounds are all shown as van-der-Waals spheres (Scott *et al.* 2012).

based on the natural ligand and fragment screening. Around 20 peptide–protein structures have been studied by X-ray crystallography and these demonstrated the subtle changes of conformation that allow peptide H-bonds to be optimised as the sequence changes. These structures of the peptide–protein complexes along with structures of protein–fragment complexes were used to design fragment-peptide chimaeric molecules, and these in turn served as templates for the design of other molecules (see Figure 9).

An example of how a potential difficulty can be utilised to give useful binding information is seen from our own work on the HGF/SF splice variant NK1. This protein was crystallised by chance with the buffering component HEPES present in a pocket thought to be important in receptor binding. This gave valuable information on how ligands bind in this pocket and how it could be exploited for the development of inhibitors of the interaction with the Met receptor (see Fig. 8*b*). In an effort to investigate this further, seven buffer compounds have now been successfully soaked into NK1 crystals (Sigurdardottir, 2012). This information, in combination with results from fragment screening, will be used to evolve fragments into a lead-like molecule.

The propensity of PPIs to form crystal contacts can be exploited in the discovery of new binding sites as exemplified by recent work (Sledz *et al.* 2010). Through crystal packing analysis of 19 different crystal forms of the polo-box domain (PBD) of polo-like kinase 1 (Plk1), Sledz *et al.* observed a previously unreported binding site formed by rearrangement of surface residues involved in crystal packing. This binding site was shown to participate in binding to a peptide derived from a biologically relevant ligand by using biophysical techniques, mutagenesis and X-ray crystallography.

3.5.5 Bioassays

Biochemical assays exploiting the enzymatic activity of the target protein have the advantage of being high throughput and low cost, so there has been interest in utilising such assays in fragment-based drug discovery. However, bioassays require inhibition of interactions, whereas most biophysical methods detect binding to a protein interface of one component rather than disrupting the interaction. Nevertheless, Geitmann *et al.* (2011) employed a primary SPR-based fragment screen followed by an interaction and an enzyme inhibition assay for stepwise hit selection and validation of a novel human immunodeficiency virus type 1 reverse transcriptase (HIV-1RT) inhibitor. For non-enzyme targets, a variety of biochemical, biophysical and cell-based biological assays may be useful, such as ALPHA-screen (PerkinElmer), assays

assessing enzymatic activity of down-stream effectors such as the phosphorylation assay (Ferraris *et al.* 2009) or migration assays (Boyden, 1962).

In some fragment-based screening campaigns, biological assays have proved remarkably helpful. For example, apo-crystals of the Met receptor were not available for compound soaking, so instead a variety of biological and biophysical assays were explored to assess the biological activity of fragments. Phosphorylation assays proved to be a fast and easy way to probe for inhibition of downstream signalling [extracellular-signal-regulated kinase (Erk) and Akt phosphorylation] caused by the fragments and early compounds, in the presence of stimulating HGF/SF. In this assay, Vero cells were grown to confluency, starved and treated with 0.1 nM HGF/SF in the presence of compound. Several fragments and compounds found from initial fragment hits using substructure and similarity searches were found to prevent Akt and Erk phosphorylation indicating biological activity (A. Winter and E. Gherardi, unpublished results).

However, most fragments are too small to exhibit biological activity and several rounds of fragment growth or evolution are usually required before cell-based assays can be utilised.

4. Growing and cross-linking

Once fragment binding has been identified and characterised fragment hits are elaborated into larger compounds with higher affinity and potency. Most fragments in commercially available fragment libraries present 'handles' for conjugation and hit evolution and are usually pharmacophore-rich, but not too complex (see Section 3.1). When the structure of a protein–ligand complex has been determined by NMR or X-ray crystallography, fragment evolution can be monitored directly and the SAR estimated for every optimisation step (SAR by NMR or SAR by X-ray crystallography, respectively). In cases where no structural information is available, other avenues for evolving the fragment hits need to be explored such as finding related compounds using substructure searches. These initial compounds can be tested for their activity in biochemical and biological assays. However, structural information of the protein in complex with the compound will eventually be highly beneficial. Indeed, the availability of tight binders can sometimes decrease the conformational heterogeneity and encourage the formation of crystals, which can be used later in the campaign.

With structural information in hand, two different strategies can be employed to construct larger compounds from one or more fragment starting points: fragment growing and fragment linking. The former has proved to be more straightforward and less problematic, since only one starting point is required (Congreve *et al.* 2005; Howard *et al.* 2006). A starting building block can be grown into a novel ligand with high affinity by the addition of polar and lipophilic groups at appropriate positions. The chemical repertoire available for fragment growing strategies is much larger because the binding site is explored consecutively and less favourable geometries are discovered immediately. Therefore, fragment growing is becoming more popular compared with fragment linking among pharmaceutical and academic groups, and several successful examples can be found in the literature (Carr *et al.* 2005; Edink *et al.* 2011; Saxty *et al.* 2007).

In contrast, fragment linking is often hampered by linkers that perturb the optimal binding geometries of the fragments or introduce excessive conformational flexibility (Chung *et al.* 2009). In these cases, linking chemistry and linker length as well as linker optimisation is crucial for finding a high-affinity binder. It is often impossible to predict, even with high-resolution structures in hand, how a linker will affect binding of two fragments. Often, linkers do not interact well with the target and therefore decrease LE. Thus, hidden strain and other energetic penalties

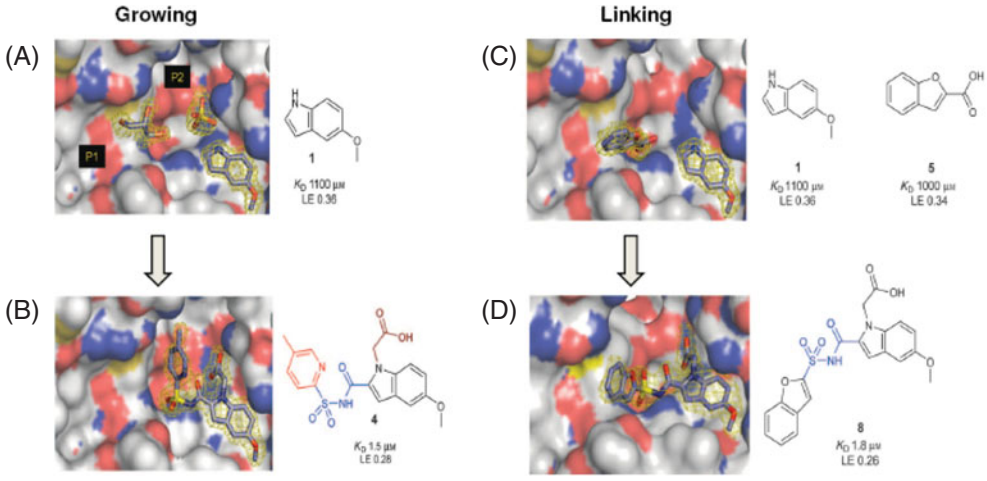

Fig. 10. Fragment-growing *versus* fragment-linking strategy designed for *M. tuberculosis* pantothenate synthetase (PS). (A) The crystal structure of 1 bound together with sulphate and glycerol molecules reveals fragment-growing opportunities at both the C2 and N1 positions of the indole. Possible growth pockets are P1 (pantoate-binding site) and P2 (pyrophosphate-/b-alanine-binding site. (B) After two rounds of optimisation the potent inhibitor 4 was generated. The initial 5-methoxyindole fragment maintains its original position throughout the elaboration process; thus, this group is a suitable anchor core for growth. (C) Fragments 1 and 5, soaked as a cocktail into crystals of PS, bind simultaneously at the active site. Fragment 5 occupies the P1 pocket. (D) Crystal structure of inhibitor 8, which contains a more conformationally constrained acyl sulphonamide linker leading to the desired additivity of the binding energies of the initial fragments. The 2Fo_Fc electron-density maps superimposed around each ligand are shown in yellow and contoured at 1σ. The ligands are shown as sticks with carbon atoms in blue, nitrogen atoms in darker blue, oxygen atoms in red and sulphur atoms in yellow. Adapted from Hung *et al.* (2009).

can only be discovered by iterative optimisation and binding measurements (Rohrig *et al.* 2007). These allow binding affinities and LE to be determined, both of which are a useful criteria for evaluating the quality of the starting fragment hits and for assessing intermediate inhibitors en route to the final lead compound (Hajduk, 2006).

In a recent study, Barelier *et al.* (2010) deconstructed nine evolved inhibitors of the Bcl-X_L into 22 fragments. They find that almost half of the generated fragments do not bind alone to the protein and the ones that bind do not always bind in the same position or pocket as they occupy in the mature molecule. This work highlights the complexity of fragment growing/linking and, as mentioned before, the difficulty to predict *a priori* how the chemical modification is going to affect the fragment-binding mode.

An example for fragment linking versus fragment growing is shown in Fig. 10 (Hung *et al.* 2009). An initial fragment hit 1 displayed good shape complementary and specificity upon fragment binding and was selected for a fragment growing strategy (Fig. 10*a*). After several rounds of optimisation, binding of the original fragment 1 could be improved from $K_D = 0.5$ mM, LE = 0.32 to K_D value of 1.5 μM and an LE value of 0.28 for compound 4 (Fig. 10*b*). In addition, two other fragments were identified that bind 3.1 Å apart from each other in adjacent pockets, and a fragment-linking strategy was employed to combine both fragments (Fig. 10*c*). The linked compound was found to bind an order of magnitude more tightly than either fragment ($K_D = 75$ μM, LE = 0.20); however, the binding of the linked compound was found to be much weaker than would be expected from complete additivity of the binding energies of the

respective fragments (Hung *et al.* 2009). Optimising the linker led to compound 8 with improved binding affinity of 1.8 μM, 1000-fold better than the original fragments (Fig. 10*d*). However, a fragment-growing strategy rather than fragment linking would have permitted a multistep optimisation of LE while retaining a small compound size to enable a more flexible exploration of the binding site. The LE values of compounds derived from fragment growing are often higher than those derived from fragment linking. This phenomenon was observed in several studies, which have led to the view that growing one fragment into an adjacent pocket where another fragment was found to bind might be the better strategy (Hajduk & Greer, 2007; Howard *et al.* 2006; Hung *et al.* 2009). Other successful examples for fragment linking are inhibitors of thrombin (Howard *et al.* 2006). We describe the approach for a PPI, Bcl-X_L (Oltersdorf *et al.* 2005; Petros *et al.* 2006) in Section 5.

Where no crystal or NMR structure of the target protein is available, other routes for fragment growing and optimisation can be explored. Substructure and more general similarity searches, e.g. using the Tversky similarity index (Tversky, 1977), seek to find larger compounds that are structurally related to the hits identified in the fragment-screening process. Compounds can be purchased from commercial sources avoiding labour-intensive synthetic chemistry at that stage, so that the optimisation process can be started. Databases such as ZINC (Irwin & Shoichet, 2005) contain commercially available compounds and are searchable for different physicochemical properties such as molecular weight, H-bond donors and acceptors or log *P* values. The fragment-based campaign on the Met receptor is an example of its application to a PPI where no unliganded structure was available. Based on substructure searches, several compounds were purchased and shown to have biological activity in phosphorylation assays (A. Winter, T. L. Blundell and E. Gherardi, unpublished results). Comparative analysis of structures of active compounds is now being used to guide synthesis or purchase of further compounds with potentially improved affinity and biological activity.

A variety of *in silico* tools has been developed to support the different phases of fragment-based drug discovery, most of which have yet to be applied to PPIs. Indeed the very different nature of PPI hotspots from classical binding sites may show that they are not useful in such cases. Recent reviews cover the computational aspects of designing leads from fragments; see for example (Mauser & Guba, 2008; Pitt *et al.* 2010; Schneider & Fechner, 2005). Fragment merging comprises a third computational strategy that may augment fragment growing and linking. An example is the program LEA3D (Douguet *et al.* 2005), recently released as web server (Douguet, 2010), where fragments derived from known drugs and biological molecules are merged using a multiparametrical fitness function. Most of the programs used to evolve fragments into more complex molecules can be classified as *de novo* strategies. Indeed, the concept of *de novo* design is to achieve with a computer what the scientists do in the lab. Examples of algorithms for fragment growing are SkelGen (Dean *et al.* 2006), AutoGrow (Durrant *et al.* 2009), CombiSMoG (Grzybowski *et al.* 2002) and SYNOPSIS (Vinkers *et al.* 2003). Regarding fragment linking, GANDI (Dey & Caflisch, 2008) links pre-docked fragments using similarity to known inhibitors. CAVEAT (Lauri & Bartlett, 1994) and CONFIRM (Thompson *et al.* 2008) store linkers from known structures or binding modes to facilitate the design of bridged fragments. Computational methods such as homology modelling and *in silico* docking have successfully been utilised in a recent study to evaluate the binding mode of a compound series to PI3K (Kim *et al.* 2011). *In silico* docking is particularly successful for enzymes where a defined binding pocket/active site is present.

An early example, introduced in Section 3.5.3, is the ZipA–FtsZ complex, where inhibitors have been sought to impede bacterial cell division, thereby generating a new class of antibiotics

that would help to overcome bacterial drug resistance. Researchers at Wyeth used an arsenal of techniques to successfully identify small molecules that disrupt ZipA–FtsZ interaction, such as HTS using FP assay (Kenny *et al.* 2003), structure-based drug design (Jennings *et al.* 2004*a*, *b*), virtual screening (Rush *et al.* 2005) and NMR-based fragment screening (Tsao *et al.* 2006). Tsao *et al.* used structural information from three different fragments (Sutherland *et al.* 2003) and merged them into a bigger molecule 6-fold more potent, but 20% less ligand efficient. This case is an elegant example of the integration of structural, computational and biological tools to find inhibitors of protein–protein interactions. However, the intrinsic hydrophobic nature and flatness of the binding interface yielded molecules that did not succeed as therapeutic agents as their cell penetration, solubility and specificity required was not matched to their ability to bind to ZipA. Nevertheless, the integration of these techniques allowed the discovery team to reach the key decision point of the project minimising time and resources (Tsao *et al.* 2006).

Growth of fragments targeting the protein–protein interaction between the human recombinase RAD51 and the hub protein BRCA2 has advanced using peptides as guides, taking advantage of about 20 structures of peptide RAD51 complexes. Structural characterisation of peptide binding has been combined with fragment screening information to design chimaeric peptide/fragment molecules, which have increased affinity when compared with the peptide progenitors. This is further described in the following section (Scott *et al.* 2012).

5. State-of-the-art

Few fragment-based campaigns of PPIs have been reported in the literature, let alone advanced into the clinic.

One exception has been the Abbott Laboratories' targeting of Bcl-X_L, which inhibits apoptosis by binding a 16-residue α-helical portion of the pro-apoptotic molecule BAK or a 26-residue α-helical portion of another pro-apoptotic molecule BAD. Small molecules that mimic the key α-helix involved in this interaction have been designed with high affinities of up to 5 nM. Abbott Laboratories used fragment-based NMR to improve the SAR (by NMR), combined with NMR-structure-guided medicinal chemistry (Bruncko *et al.* 2007), as described in Section 3 above. Figure 11 shows the published progression of the design process (Oltersdorf *et al.* 2005). Initially two different fragments that could be bound simultaneously to Bcl-X_L (Fig. 11 *a*) were evolved into a single molecule (Fig. 11 *b*), by 'structure-guided iterative library approach' (Petros *et al.* 2006). Several rounds of optimisation led to a high-affinity organic compound that bound to the hydrophobic helical domain of Bcl-X_L, Bcl-2 and Bcl-W (Fig. 11 *c*). With their smaller contact regions with the protein, the LE is almost twofold higher than that of a small helix. This is a particular challenging example with impressive success. However, optimised molecules are highly lipophilic and do not mimic available specific polar contacts that BAD engages (Fig. 11 *d*).

Fragment-based approaches targeting the protein–protein interaction between the human recombinase RAD51 and the hub protein BRCA2 has progressed by two routes: the exploration of short peptides based on the natural ligand and fragment screening (Scott *et al.* 2012). Fragment screening was initially carried out using thermal shift with hits validated using both ligand observed NMR and X-ray crystallography. ITC and X-ray crystallography were used to explore the binding of various peptides designed around the natural ligand to the target protein. NMR and FP detected competition assays have been exploited to verify binding at the expected sites. As detailed in Sections 3.3 and 3.4.4, a number of different humanised monomeric mutants have been designed to be amenable for use with the different screening methods. The use of different

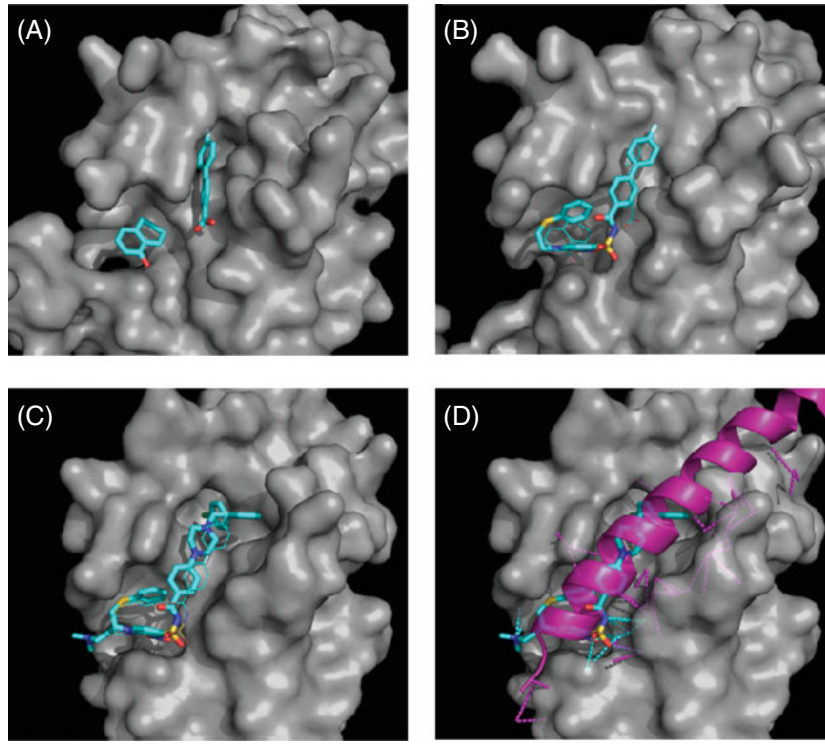

Fig. 11. Bcl-X_L is represented as grey surface. (A) Two different fragments in cyan co-crystallised with Bcl-X_L (1YSG). (B) Molecule in cyan sticks (1YSI) generated from merging and optimising previous fragments (from 1YSG), both are superposed using the protein chain as a guide and represented by fine lines. (C) Optimised compound ABT-737 in cyan sticks bound to Bcl-X_L (2YXJ), molecule from 1YSI is superposed using the protein and represented by fine lines. (D) Compound ABT-737 in cyan sticks bound to Bcl-X_L (2YXJ), BAD peptide in magenta is superposed using the protein and represented by magenta helix. Dotted lines represent polar interactions between the protein and BAD in magenta and between protein and small molecules in cyan. Figures generated with the program PyMol (www.pymol.org).

crystal seeding techniques between these mutants has been used to crystallise previously un-crystallisable mutants and to engineer crystal forms with unobstructed binding sites to be optimal for soaking. Information from both the fragment screening and peptide SAR has been used to design chimaeras of peptides and fragments with increased potency. Future challenges involve targeting further sites on different parts of the same protein to give opportunities for fragment linking to increase potency.

An NMR-based screen has been successfully used to find fragments that bind to the I-domain allosteric site (IDAS) of the inflammatory disorder target leukocyte function-associated antigen-1 (LFA-1) (Liu *et al.* 2001). Compounds, such as those discovered in this study, inhibit the binding of intercellular adhesion molecule-1 (ICAM-1) to LFA-1. This protein–protein interaction is part of the mechanism of immune cell adhesion and tissue invasion.

An interesting target has been the CC chemokine receptor type 5 (CCR5), which is the target of Maraviroc (Dorr *et al.* 2005), an approved drug for the treatment of HIV-1 infection. This inhibitor reduces cellular penetration of HIV by inhibiting the binding of viral gp120 to the chemokine receptor CCR5. Although it is derived from a HTS hit, fragments that bind to the same site on the receptor have recently been discovered using SPR (Navratilova *et al.* 2011).

A further example is the targeting of PPIs by tethering fragments to interleukin-2 (IL-2). The small molecule SP4206, which binds IL-2, is not an accurate mimic of the receptor but traps a conformation of IL-2 in which a groove is present for small-molecule binding and in which a loop has been repositioned to embrace the furanoic acid moiety at one end of the small molecule (Thanos *et al.* 2006). It is worth highlighting that the first IL-2 binder was initially developed at Hoffmann–La Roche more than a decade ago (Tilley *et al.* 1997). It was designed to mimic IL-2 and disturb the interaction with its IL-2 receptor α (IL-2Rα). However, when the interaction was characterised it showed that the molecule was binding to IL-2 itself, underlining the difficulty of designing molecules to mimic interfaces and the power of structural data.

An example of a molecule that modulates a PPI and is also an allosteric inhibitor is MK-2206, a non-ATP-competitive Akt inhibitor developed by Merck (Tolcher *et al.* 2007). Akt1 is a serine/threonine AGC protein kinase family member, involved in cell regulation, growth and survival, dysregulation of which is associated with many cancers. MK-2206 depends on the pleckstrin-homology (PH) domain for binding. Interestingly, Wu *et al.* (2010) have determined the crystal structures of an inhibitor, which appears to have a similar mechanism and demonstrates interactions between the PH and kinase domains. The authors describe an 'intricate balance in the enzymatic regulation, where the PH domain appears to lock the AKT kinase in an inactive conformation and the kinase domain disrupts the phospholipid binding site of the PH domain'. This structure provides a basis for understanding the mechanism of this group of inhibitors and demonstrates the close relationships that allosteric inhibitors may have with those that modulate protein–protein interactions. This year, a phase I study was conducted by Merck looking at administering MK-2206 alongside chemotherapy or erlotinib for solid tumours that have continued to grow despite having other treatment with encouraging results (Hirai *et al.* 2010).

6. Future prospects

The numbers of candidate compounds at the pre-clinical stage or in the clinic that have derived from targeting PPIs are still too small to allow generally applicable conclusions to be drawn. Indeed Maraviroc (Dorr *et al.* 2005), an approved drug for the treatment of HIV-1 infection and developed by conventional methods, is said to be the only protein–protein inhibitor on the market to date (Wilson, 2009).

With respect to fragment-based approaches, early efforts using tethering, arising mainly from the beautiful work of Jim Wells and colleagues at Sunesis, are impressive and have given confidence that PPIs can be targeted. The application of SAR by NMR by Steve Fesik and colleagues at Abbott has demonstrated that the method could be used especially on PPIs involving helical epitopes. In the past 4 years, improvements in fragment-screening technologies, experience with fragment growth and optimisation and new developments in computational methods have been pushing the boundaries of what is druggable – a notoriously ill-defined concept in general – and what is not.

6.1 Further developments in methodology for fragment-based drug discovery

As we have described, fragment binders can be discovered for almost any protein target, including intrinsically disordered examples.

We are beginning to gain useful knowledge by computational analysis of the increasing numbers of reports of compounds that modify protein–protein complexes. Much could be

achieved by enriching fragment libraries with chemical structures that have been successful in targeting PPIs as well as by providing a greater variety of 3D shapes (Hung *et al.* 2011). We must build on the idea of using the fragment-hit rate as an indicator of druggability (Edfeldt *et al.* 2011; Hajduk *et al.* 2005).

Improving specific contacts in the earlier phase of the discovery, instead of optimising purely for binding affinity, should decrease the tendency to increase lipophilicity and non-specific hydrophobic interactions (Keseru & Makara, 2009). Indeed, binding thermodynamic data show a negative correlation with maximal enthalpy and ligand size (Ferenczy & Keserü, 2010). This means that there is a better chance of introducing specific interactions in the very early stages of ligand optimisation and that this is especially achievable for fragment hits.

A consequence of the activity in fragment-based screening has been increased interest in biophysical screening assays. As a result, extremely sensitive assays for the detection of protein binders have become more accessible. In turn, this has led to the increased use of techniques such as SPR to determine which binders also inhibit protein–protein interactions. Nevertheless, progress is often hampered by experimental noise at affinities close to the limit of assay sensitivity and ligand concentrations close to the limit of solubility. For this reason, we must continue to develop multiple complementary assay systems to confirm activity. Even so, painstaking fragment analogising and growth may become becalmed within a sea of flat SAR, indicating that it is time to seek a new approach or a different target.

6.2 Classes of PPIs

Success in targeting PPIs clearly depends on target type. We have commented on the characteristically different interfaces that involve proteins that undergo concerted folding and binding at the 'receptor' protein (Blundell *et al.* 2000; Boehr *et al.* 2009; Dyson & Wright, 2002; Wright & Dyson, 1999). These tend to have better defined binding sites, either as grooves to accommodate helices, for example Bcl, or small pockets that bind side chains but can be exploited to bind fragments, for example RAD51. Globular partners that bind flexible/disordered peptides tend to have relatively little induced conformational change on fragment or even larger ligand binding. They constitute many of the most promising targets for fragment-based approaches.

On the other hand, interactions between globular proteins tend to involve large, flat and featureless interfaces. The greatest success here has been with tethering methods that allow the fragments to induce pockets at hotspots. Nevertheless, the evidence from the Notch receptor ankyrin domain indicates that untethered fragments can be observed to bind on such surfaces. However, at this time, little progress has been made to link or grow them.

6.3 Co-operative and allosteric effects

Although co-operative and allosteric inhibitors are not the focus of this review, trapping conformational states of proteins involved in protein complexes or oligomers may be fruitful in the context of interrupting protein–protein interactions.

Homomultimers are an interesting example of where co-operative ligand effects can occur. For example, the dimer interfaces of homodimeric enzymes have been suggested as potentially favourable targets (Cardinale *et al.* 2010). Binding of the same ligand to multiple identical binding sites can occur in homo-oligomers, potentially leading to cooperative effects. An example of a

dimer interface where two copies of a fragment are known to bind was described by workers at Abbott on human survivin (Wendt *et al.* 2007).

A related approach is the stabilisation of benign multimeric forms of proteins, which in other complexes can be disease causing (Lawrence *et al.* 2008). An example of a fragment-sized compound that stabilises the native oligomeric state of a protein is Tafamidis, currently in clinical trials for the treatment of transthyretin familial amyloid polyneuropathy (TTR-FAP) (Razavi *et al.* 2003). Two copies of this compound, which has a molecular weight of 308 Da, stabilise the native transthyretin tetramer, preventing amyloidosis.

The cAMP-response-element-binding protein (CREB) and its binding partner CREB-binding protein (CBP) (Best *et al.* 2004; Sugase *et al.* 2007) are examples of concerted folding and binding; when the KID domain of CREB interacts with the KIX domain of CBP, the KID domain forms two helices. Best and co-workers show that a shallow hydrophobic groove on the surface of KIX that accommodates an amphipathic helix binds several compounds. One of these, KG-501, disrupts the interaction and attenuates the function when added to live cells. These observations remind us that future drug discovery will need to be aware of the importance of protein dynamics (Lee & Craik, 2009), which clearly play important roles in drug discovery.

6.4 Concluding thought

Wells and McClendon were right to encourage us to reach for the figuratively 'high-hanging fruit' in 2007 (Wells & McClendon, 2007). Surely the future of many selective therapeutic approaches will be in targeting protein–protein interactions, which regulate many of the important pathways in cells and are often dysregulated in disease. Fragment-based approaches offer a realistic approach in succeeding in this endeavour.

7. Acknowledgements

We thank the Wellcome Trust, Bill and Melinda Gates Foundation, BBSRC, European Union and UCB for financial support. We are grateful to Dr Anthony Coyne, Dr John Skidmore and Dr Chris Murray for reading and providing very useful comments on our manuscript and to Professor Chris Abell, Dr Ermanno Gherardi and Dr Marko Hyvonen for helpful discussions.

8. References

Abdel-Rahman, N., Martinez-Arias, A. & Blundell, T. L. (2011). Probing the druggability of protein–protein interactions: targeting the Notch1 receptor ankyrin domain using a fragment-based approach. *Biochemical Society Transactions* **39**, 1327–1333.

Barelier, S., Pon, J., Marcillat, O., Lancelin, J. M. & Krimm, I. (2010). Fragment-based deconstruction of Bcl-X_L inhibitors. *Journal of Medical Chemistry* **53**, 2577–2588.

Bernal, F., Wade, M., Godes, M., Davis, T. N., Whitehead, D. G., Kung, A. L., Wahl, G. M. & Walensky, L. D. (2010). A stapled p53 helix overcomes HDMX-mediated suppression of p53. *Cancer Cell* **18**, 411–422.

Best, J. L., Amezcua, C. A., Mayr, B., Flechner, L., Murawsky, C. M., Emerson, B., Zor, T., Gardner, K. H. & Montminy, M. (2004). Identification of small-molecule antagonists that inhibit an activator:coactivator interaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17622–17627.

Bickerton, G. R. (2009). Molecular characterization and evolutionary plasticity of protein–protein interfaces, University of Cambridge.

Bickerton, G. R., Higueruelo, A. P. & Blundell, T. L. (2011). Comprehensive, atomic-level characterization of structurally characterized protein–protein interactions: the PICCOLO database. *BMC Bioinformatics* **12**, 313.

Blundell, T. (1979). Conformation and molecular biology of polypeptide hormones II. Glucagon. *Trends in Biochemical Sciences* **4**, 80–83.

Blundell, T. & Wood, S. (1982). The conformation, flexibility and dynamics of polypeptide hormones. *Annual Review of Biochemistry* **51**, 123–154.

Blundell, T. L., Burke, D. F., Chirgadze, D., Dhanaraj, V., Hyvonen, M., Innis, C. A., Parisini, E., Pellegrini, L., Sayed, M. & Sibanda, B. L. (2000). Protein–protein interactions in receptor activation and intracellular signalling. *Biological Chemistry* **381**, 955–959.

Blundell, T. L., Jhoti, H. & Abell, C. (2002). High-throughput crystallography for lead discovery in drug design. *Nature Reviews Drug Discovery* **1**, 45–54.

Blundell, T. L., Sibanda, B. L., Montalvao, R. W., Brewerton, S., Chelliah, V., Worth, C. L., Harmer, N. J., Davies, O. & Burke, D. (2006). Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **361**, 413–423.

Boehr, D. D., Nussinov, R. & Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology* **5**, 789–796.

Bogan, A. A. & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* **280**, 1–9.

Bonda, D. J., Lee, H. P., Lee, H. G., Friedlich, A. L., Perry, G., Zhu, X. & Smith, M. A. (2010). Novel therapeutics for Alzheimer's disease: an update. *Current Opinion in Drug Discovery and Development* **13**, 235–246.

Bordner, A. J. (2009). Predicting protein–protein binding sites in membrane proteins. *BMC Bioinformatics* **10**, 312.

Bourgeas, R., Basse, M. J., Morelli, X. & Roche, P. (2010). Atomic analysis of protein–protein interfaces with known inhibitors: the 2P2I database. *PLoS ONE* **5**, e9598.

Boyden, S. (1962). The chemotactic effect of mixtures of antibody and antigen on polymorphonuclear leucocytes. *Journal of Experimental Medicine* **115**, 453–466.

Brandts, J. F. & Lin, L. N. (1990). Study of strong to ultratight protein interactions using differential scanning calorimetry. *Biochemistry* **29**, 6927–6940.

Braun, W., Wider, G., Lee, K. H. & Wuthrich, K. (1983). Conformation of glucagon in a lipid–water interphase by ^{1}H nuclear magnetic resonance. *Journal of Molecular Biology* **169**, 921–948.

Bruncko, M., Oost, T. K., Belli, B. A., Ding, H., Joseph, M. K., Kunzer, A., Martineau, D., Mcclellan, W. J., Mitten, M., Ng, S. C., Nimmer, P. M., Oltersdorf, T., Park, C. M., Petros, A. M., Shoemaker, A. R., Song, X., Wang, X., Wendt, M. D., Zhang, H., Fesik, S. W., Rosenberg, S. H. & Elmore, S. W. (2007). Studies leading to potent, dual inhibitors of Bcl-2 and Bcl-X_L. *Journal of Medical Chemistry* **50**, 641–662.

Buerger, C. & Groner, B. (2003). Bifunctional recombinant proteins in cancer therapy: cell penetrating peptide aptamers as inhibitors of growth factor signaling. *Journal of Cancer Research and Clinical Oncology* **129**, 669–675.

Cardinale, D., Salo-Ahen, O. M., Ferrari, S., Ponterini, G., Cruciani, G., Carosati, E., Tochowicz, A. M., Mangani, S., Wade, R. C. & Costi, M. P. (2010). Homodimeric enzymes as drug targets. *Current Medicinal Chemistry* **17**, 826–846.

Carr, R. & Jhoti, H. (2002). Structure-based screening of low-affinity compounds. *Drug Discovery Today* **7**, 522–527.

Carr, R. A., Congreve, M., Murray, C. W. & Rees, D. C. (2005). Fragment-based lead discovery: leads by design. *Drug Discovery Today* **10**, 987–992.

Celej, M. S., Montich, G. G. & Fidelio, G. D. (2003). Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein Science: A Publication of the Protein Society* **12**, 1496–1506.

Chakrabarti, P. & Bhattacharyya, R. (2007). Geometry of nonbonded interactions involving planar groups in proteins. *Progress in Biophysics and Molecular Biology* **95**, 83–137.

Chatr-Aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L. & Cesareni, G. (2007). MINT: the Molecular INTeraction database. *Nucleic Acids Research* **35**, D572–D574.

Chen, J., Zheng, X. F., Brown, E. J. & Schreiber, S. L. (1995). Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 4947–4951.

Chen, Y. C., Chen, H. C. & Yang, J. M. (2006). DAPID: a 3D-domain annotated protein–protein interaction database. *Genome Informatics* **17**, 206–215.

Choi, J., Chen, J., Schreiber, S. L. & Clardy, J. (1996). Structure of the FKBP12–rapamycin complex interacting with the binding domain of human FRAP. *Science* **273**, 239–242.

Choi, Y. S., Hur, J., Lee, H. K. & Jeong, S. (2009). The RNA aptamer disrupts protein–protein interaction between beta-catenin and nuclear factor-kappaB p50 and regulates the expression of C-reactive protein. *FEBS Letters* **583**, 1415–1421.

Chung, S., Parker, J. B., Bianchet, M., Amzel, L. M. & Stivers, J. T. (2009). Impact of linker strain and flexibility in the design of a fragment-based inhibitor. *Nature Chemical Biology* **5**, 407–413.

Cimmperman, P., Baranauskiene, L., Jachimoviciute, S., Jachno, J., Torresan, J., Michailoviene, V., Matuliene, J., Sereikaite, J., Bumelis, V. & Matulis, D. (2008). A quantitative model of thermal stabilization

and destabilization of proteins by ligands. *Biophysics Journal* **95**, 3222–3231.

Ciulli, A., Williams, G., Smith, A. G., Blundell, T. L. & Abell, C. (2006). Probing hot spots at protein-ligand binding sites: a fragment-based approach using biophysical methods. *Journal of Medical Chemistry* **49**, 4992–5000.

Clackson, T. & Wells, J. A. (1995). A hot spot of binding energy in a hormone–receptor interface. *Science* **267**, 383–386.

Cole, D. C., Olland, A. M., Jacob, J., Brooks, J., Bursavich, M. G., Czerwinski, R., Declercq, C., Johnson, M., Joseph-Mccarthy, D., Ellingboe, J. W., Lin, L., Nowak, P., Presman, E., Strand, J., Tam, A., Williams, C. M., Yao, S., Tsao, D. H. & Fitz, L. J. (2010). Identification and characterization of acidic mammalian chitinase inhibitors. *Journal of Medical Chemistry* **53**, 6122–6128.

Congreve, M., Carr, R., Murray, C. & Jhoti, H. (2003). A 'rule of three' for fragment-based lead discovery? *Drug Discovery Today* **8**, 876–877.

Congreve, M., Chessari, G., Tisi, D. & Woodhead, A. J. (2008). Recent developments in fragment-based drug discovery. *Journal of Medical Chemistry* **51**, 3661–3680.

Congreve, M., Murray, C. W. & Blundell, T. L. (2005). Structural biology and drug discovery. *Drug Discovery Today* **10**, 895–907.

Congreve, M., Rich, R. L., Myszka, D. G., Figaroa, F., Siegal, G. & Marshall, F. H. (2011). Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. *Methods in Enzymology* **493**, 115–136.

D'arcy, A., Villard, F. & Marsh, M. (2007). An automated microseed matrix-screening method for protein crystallization. *Acta Crystallographica, Section D, Biological Crystallography* **63**(Pt 4), 550–554.

Dalvit, C., Pevarello, P., Tato, M., Veronesi, M., Vulpetti, A. & Sundstrom, M. (2000). Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *Journal of Biomolecular NMR* **18**, 65–68.

Davies, T. G., van Montfort, R. L. M., Williams, G. & Jhoti, H. (2006). Pyramid: an integrated platform for fragment-based drug discovery. In *Fragment-based Approaches in Drug Discovery*, pp. 193–214. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Davis, F. P. & Sali, A. (2005). PIBASE: a comprehensive database of structurally defined protein interfaces. *Bioinformatics* **21**, 1901–1907.

Davis, F. P. & Sali, A. (2010). The overlap of small molecule and protein binding sites within families of protein structures. *PLoS Computational Biology* **6**, e1000668.

Dean, P. M., Firth-Clark, S., Harris, W., Kirton, S. B. & Todorov, N. P. (2006). SkelGen: a general tool for structure-based *de novo* ligand design. *Expert Opinion on Drug Discovery* **1**, 179–189.

Del Sol, A. & O'MEARA, P. (2005). Small-world network approach to identify key residues in protein–protein interaction. *Proteins* **58**, 672–682.

Dey, F. & Caflisch, A. (2008). Fragment-based de novo ligand design by multiobjective evolutionary optimization. *Journal of Chemical Information Modeling* **48**, 679–690.

Dorr, P., Westby, M., Dobbs, S., Griffin, P., Irvine, B., Macartney, M., Mori, J., Rickett, G., Smith-Burchnell, C., Napier, C., Webster, R., Armour, D., Price, D., Stammen, B., Wood, A. & Perros, M. (2005). Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrobial Agents and Chemotherapy* **49**, 4721–4732.

Dou, Y., Baisnee, P. F., Pollastri, G., Pecout, Y., Nowick, J. & Baldi, P. (2004). ICBS: a database of interactions between protein chains mediated by beta-sheet formation. *Bioinformatics* **20**, 2767–2777.

Douguet, D. (2010). e-LEA3D: a computational-aided drug design web server. *Nucleic Acids Research* **38**(Web Server issue), W615–W621.

Douguet, D., Munier-Lehmann, H., Labesse, G. & Pochet, S. (2005). LEA3D: a computer-aided ligand design for structure-based drug design. *Journal of Medical Chemistry* **48**, 2457–2468.

Durrant, J., Amaro, R. & Mccammon, J. (2009). AutoGrow: a novel algorithm for protein inhibitor design. *Chemical Biology & Drug Design* **73**, 168–178.

Dyson, H. J. & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology* **12**, 54–60.

Edfeldt, F. N., Folmer, R. H. & Breeze, A. L. (2011). Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today* **16**, 284–287.

Edink, E., Rucktooa, P., Retra, K., Akdemir, A., Nahar, T., Zuiderveld, O., Van Elk, R., Janssen, E., Van Nierop, P., Van Muijlwijk-Koezen, J., Smit, A. B., Sixma, T. K., Leurs, R. & De Esch, I. J. (2011). Fragment growing induces conformational changes in acetylcholine-binding protein: a structural and thermodynamic analysis. *Journal of the American Chemical Society* **133**, 5363–5371.

Elinder, M., Geitmann, M., Gossas, T., Kallblad, P., Winquist, J., Nordstrom, H., Hamalainen, M. & Danielson, U. H. (2011). Experimental validation of a fragment library for lead discovery using SPR biosensor technology. *Journal of Biomolecular Screening: The Official Journal of the Society for Biomolecular Screening* **16**, 15–25.

Erlanson, D. A., Braisted, A. C., Raphael, D. R., Randal, M., Stroud, R. M., Gordon, E. M. & Wells, J. A. (2000). Site-directed ligand discovery. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9367–9372.

Erlanson, D. A., Lam, J. W., Wiesmann, C., Luong, T. N., Simmons, R. L., Delano, W. L., Choong, I. C., Burdett, M. T., Flanagan, W. M., Lee, D., Gordon, E. M. & O'Brien, T. (2003). *In situ* assembly of enzyme inhibitors using extended tethering. *Nature Biotechnology* **21** 308–314.

Ferenczy, G. G. & Keserü, G. M. (2010). Enthalpic efficiency of ligand binding. *Journal of Chemical Information Model* **50**, 1536–1541.

Ferraris, D. M., Gherardi, E., Di, Y., Heinz, D. W. & Niemann, H. H. (2009). Ligand-mediated dimerization of the Met receptor tyrosine kinase by the bacterial invasion protein InlB. *Journal of Molecular Biology* **395**, 522–532.

Fischer, T. B., Arunachalam, K. V., Bailey, D., Mangual, V., Bakhru, S., Russo, R., Huang, D., Paczkowski, M., Lalchandani, V., Ramachandra, C., Ellison, B., Galer, S., Shapley, J., Fuentes, E. & Tsai, J. (2003). The binding interface database (BID): a compilation of amino acid hot spots in protein interfaces. *Bioinformatics* **19**, 1453–1454.

Fitzpatrick, P. A., Steinmetz, A. C., Ringe, D. & Klibanov, A. M. (1993). Enzyme crystal structure in a neat organic solvent. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8653–8657.

Fletcher, S. & Hamilton, A. D. (2005). Protein surface recognition and proteomimetics: mimics of protein surface structure and function. *Current Opinion in Chemical Biology* **9**, 632–638.

Fuller, J. C., Burgoyne, N. J. & Jackson, R. M. (2009). Predicting druggable binding sites at the protein–protein interface. *Drug Discovery Today* **14**, 155–161.

Gabadinho, J., Hall, D., Leonard, G., Gordon, E., Monaco, S. & Thibault, X. (2008). Remote access experiments at the macromolecular crystallography beamlines of the ESRF. *Synchrotron Radiation News* **21**, 24–29.

Gao, Y., Douguet, D., Tovchigrechko, A. & Vakser, I. A. (2007). DOCKGROUND system of databases for protein recognition studies: unbound structures for docking. *Proteins* **69**, 845–851.

Garrett, D. S., Seok, Y. J., Peterkofsky, A., Clore, G. M. & Gronenborn, A. M. (1997). Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. *Biochemistry* **36**, 4393–4398.

Geitmann, M., Elinder, M., Seeger, C., Brandt, P., De Esch, I. J. & Danielson, U. H. (2011). Identification of a novel scaffold for allosteric inhibition of wild type and drug resistant HIV-1 reverse transcriptase by fragment library screening. *Journal of Medicinal Chemistry* **53**, 699–708.

Geschwindner, S., Olsson, L. L., Albert, J. S., Deinum, J., Edwards, P. D., De Beer, T. & Folmer, R. H. (2007). Discovery of a novel warhead against beta-secretase through fragment-based lead generation. *Journal of Medical Chemistry* **50**, 5903–5911.

Giannetti, A. M. (2011). From experimental design to validated hits a comprehensive walk-through of fragment lead identification using surface plasmon resonance. *Methods in Enzymology* **493**, 169–218.

Gong, S., Yoon, G., Jang, I., Bolser, D., Dafas, P., Schroeder, M., Choi, H., Cho, Y., Han, K., Lee, S., Lappe, M., Holm, L., Kim, S., Oh, D. & Bhak, J. (2005). PSIbase: a database of protein structural interactome map (PSIMAP). *Bioinformatics* **21**, 2541–2543.

Grzybowski, B. A., Ishchenko, A. V., Shimada, J. & Shakhnovich, E. I. (2002). From knowledge-based potentials to combinatorial lead design in silico. *Accounts of Chemical Research* **35**, 261–269.

Hajduk, P. J. (2006). Fragment-based drug design: how big is too big? *Journal of Medical Chemistry* **49**, 6972–6976.

Hajduk, P. J. & Greer, J. (2007). A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Reviews. Drug Discovery* **6**, 211–219.

Hajduk, P. J., Huth, J. R. & Fesik, S. W. (2005). Druggability indices for protein targets derived from NMR-based screening data. *Journal of Medical Chemistry* **48**, 2518–2525.

Hajduk, P. J., Olejniczak, E. T. & Fesik, S. W. (1997). One-dimensional relaxation and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *Journal of the American Chemical Society* **119**, 12257–11226.

Hamalainen, M. D., Zhukov, A., Ivarsson, M., Fex, T., Gottfries, J., Karlsson, R. & Bjorsne, M. (2008). Label-free primary screening and affinity ranking of fragment libraries using parallel analysis of protein panels. *Journal of Biomolecular Screening: The Official Journal of the Society for Biomolecular Screening* **13**, 202–209.

Hann, M. M., Leach, A. R. & Harper, G. (2001). Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences* **41**, 856–864.

Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J. & Jhoti, H. (2005). Fragment-based lead discovery using X-ray crystallography. *Journal of Medical Chemistry* **48**, 403–413.

Hendlich, M., Rippmann, F. & Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of Molecular Graphics & Modeling* **15**, 359–363, 389.

Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., Vingron, M., Roechert, B., Roepstorff, P., Valencia, A., Margalit, H., Armstrong, J., Bairoch, A., Cesareni, G., Sherman, D. & Apweller, R. (2004). IntAct: an open source molecular interaction database. *Nucleic Acids Research* **32**, D452–D455.

Higueruelo, A. P., Schreyer, A., Bickerton, G. R., Pitt, W. R., Groom, C. R. & Blundell, T. L. (2009). Atomic

interactions and profile of small molecules disrupting protein–protein interfaces: the TIMBAL database. *Chemistry Biology and Drug Design* **74**, 457–467.

Hirai, H., Sootome, H., Nakatsuru, Y., Miyama, K., Taguchi, S., Tsujioka, K., Ueno, Y., Hatch, H., Majumder, P. K., Pan, B. S. & Kotani, H. (2010). MK-2206, an allosteric Akt inhibitor, enhances antitumor efficacy by standard chemotherapeutic agents or molecular targeted drugs in vitro and *in vivo*. *Molecular Cancer and Therapeutics* **9**, 1956–1967.

Holdgate, G. A. & Ward, W. H. (2005). Measurements of binding thermodynamics in drug discovery. *Drug Discovery Today* **10**, 1543–1550.

Homans, S. W. (2004). NMR spectroscopy tools for structure-aided drug design. *Angewandte Chemie International Edition* **43**, 290–300.

Hopkins, A. L. & Groom, C. R. (2002). The druggable genome. *Nature Reviews Drug Discovery* **1**, 727–730.

Hopkins, A. L., Groom, C. R. & Alex, A. (2004). Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **9**, 430–431.

Howard, N., Abell, C., Blakemore, W., Chessari, G., Congreve, M., Howard, S., Jhoti, H., Murray, C. W., Seavers, L. C. & Van Montfort, R. L. (2006). Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *Journal of Medical Chemistry* **49**, 1346–1355.

Hubbard, R. E., Davis, B., Chen, I. & Drysdale, M. J. (2007). The SeeDs approach: integrating fragments into drug discovery. *Current Topics in Medicinal Chemistry* **7**, 1568–1581.

Hubbard, R. E. & Murray, J. B. (2011). Experiences in fragment-based lead discovery. *Methods in Enzymology* **493**, 509–531.

Hung, A. W., Ramek, A., Wang, Y., Kaya, T., Wilson, J. A., Clemons, P. A. & Young, D. W. (2011). Organic synthesis toward small-molecule probes and drugs special feature: route to three-dimensional fragments using diversity-oriented synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6799–6804.

Hung, A. W., Silvestre, H. L., Wen, S., Ciulli, A., Blundell, T. L. & Abell, C. (2009). Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angewandte Chemie International Edition* **48**, 8452–8456.

Irwin, J. J. & Shoichet, B. K. (2005). ZINC – a free database of commercially available compounds for virtual screening. *Journal of Chemical Information Model* **45**, 177–182.

Jefferson, E. R., Walsh, T. P., Roberts, T. J. & Barton, G. J. (2007). SNAPPI-DB: a database and API of structures, iNterfaces and alignments for protein–protein interactions. *Nucleic Acids Research* **35** (Database issue), D580–D589.

Jennings, L. D., Foreman, K. W., Rush, III, T. S., Tsao, D. H., Mosyak, L., Kincaid, S. L., Sukhdeo, M. N., Sutherland, A. G., Ding, W., Kenny, C. H., Sabus, C. L., Liu, H., Dushin, E. G., Moghazeh, S. L., Labthavikul, P., Petersen, P. J., Tuckman, M., Haney, S. A. & Ruzin, A. V. (2004a). Combinatorial synthesis of substituted 3-(2-indolyl)piperidines and 2-phenyl indoles as inhibitors of ZipA–FtsZ interaction. *Bioorganic and Medicinal Chemistry* **12**, 5115–5131.

Jennings, L. D., Foreman, K. W., Rush, III, T. S., Tsao, D. H., Mosyak, L., Li, Y., Sukhdeo, M. N., Ding, W., Dushin, E. G., Kenny, C. H., Moghazeh, S. L., Petersen, P. J., Ruzin, A. V., Tuckman, M. & Sutherland, A. G. (2004b). Design and synthesis of indolo[2,3-a]quinolizin-7-one inhibitors of the ZipA–FtsZ interaction. *Bioorganic and Medicinal Chemistry Lett* **14**, 1427–1431.

Jochim, A. L. & Arora, P. S. (2009). Assessment of helical interfaces in protein–protein interactions. *Molecular Biosystems* **5**, 924–926.

Jochim, A. L. & Arora, P. S. (2010). Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. *ACS Chemical Biology* **5**, 919–923.

Jonsson, U., Fagerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Lofas, S., Persson, B., Roos, H., Ronnberg, I. *et al.* (1991). Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques* **11**, 620–627.

Jubb, H., Higueruelo, A., Winter, A. & Blundell, T. (2012). Structural biology and drug discovery for protein–protein interactions. *Trends in Pharmacological Sciences* **33**, 241–248.

Kalidas, Y. & Chandra, N. (2008). PocketDepth: a new depth based algorithm for identification of ligand binding sites in proteins. *Journal of Structurall Biology* **161**, 31–42.

Kenny, C. H., Ding, W., Kelleher, K., Benard, S., Dushin, E. G., Sutherland, A. G., Mosyak, L., Kriz, R. & Ellestad, G. (2003). Development of a fluorescence polarization assay to screen for inhibitors of the FtsZ/ZipA interaction. *Analytical Biochemistry* **323**, 224–233.

Keseru, G. M. & Makara, G. M. (2009). The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews. Drug Discovery* **8**, 203–212.

Keskin, O., Ma, B., Rogale, K., Gunasekaran, K. & Nussinov, R. (2005). Protein-protein interactions: organization, cooperativity and mapping in a bottom-up systems biology approach. *Physical Biology* **2**, S24–35.

Khurana, T. S. & Davies, K. E. (2003). Pharmacological strategies for muscular dystrophy. *Nature Reviews Drug Discovery* **2**, 379–390.

Kim, O., Jeong, Y., Lee, H., Hong, S. S. & Hong, S. (2011). Design and synthesis of imidazopyridine analogues as inhibitors of phosphoinositide 3-kinase

signaling and angiogenesis. *Journal of Medical Chemistry* **54**, 2455–2466.

Kobayashi, M., Retra, K., Figaroa, F., Hollander, J. G., Ab, E., Heetebrij, R. J., Irth, H. & Siegal, G. (2010). Target immobilization as a strategy for NMR-based fragment screening: comparison of TINS, STD, and SPR for fragment hit identification. *Journal of Biomolecular Screening: The Official Journal of the Society for Biomolecular Screening* **15**, 978–989.

Kortemme, T. & Baker, D. (2002). A simple physical model for binding energy hot spots in protein–protein complexes. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14116–14121.

Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**, 774–797.

Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. (1999). The maximal affinity of ligands. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9997–10002.

Lage, K., Karlberg, E. O., Storling, Z. M., Olason, P. I., Pedersen, A. G., Rigina, O., Hinsby, A. M., Tumer, Z., Pociot, F., Tommerup, N., Moreau, Y. & Brunak, S. (2007). A human phenome–interactome network of protein complexes implicated in genetic disorders. *Nature Biotechnology* **25**, 309–316.

Lauri, G. & Bartlett, P. A. (1994). CAVEAT: a program to facilitate the design of organic molecules. *Journal of Computer-Aided Molecular Design* **8**, 51–66.

Laurie, A. T. & Jackson, R. M. (2005). Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics* **21**, 1908–1916.

Lawrence, S. H., Ramirez, U. D., Tang, L., Fazliyez, F., Kundrat, L., Markham, G. D. & Jaffe, E. K. (2008). Shape shifting leads to small-molecule allosteric drug discovery. *Chemistry & Biology* **15**, 586–596.

Le Guilloux, V., Schmidtke, P. & Tuffery, P. (2009). Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **10**, 168.

Leavitt, S. & Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current Opinion in Structural Biology* **11**, 560–566.

Lee, G. M. & Craik, C. S. (2009). Trapping moving targets with small molecules. *Science* **324**, 213–215.

Lee, S., Brown, A., Pitt, W., Higueruelo, A., Gong, S., Bickerton, G., Schreyer, A., Tanramluk, D., Baylay, A. & Blundell, T. (2009). Structural interactomics: informatics approaches to aid the interpretation of genetic variation and the development of novel therapeutics. *Molecular Biosystems.* **5**, 1456–1472.

Lepre, C. A. (2011). Practical aspects of NMR-based fragment screening. *Methods in Enzymology* **493**, 219–239.

Lepre, C. A., Moore, J. M. & Peng, J. W. (2004). Theory and applications of NMR-based screening in pharmaceutical research. *Chemical Reviews* **104**, 3641–3675.

Leung, S. M., Senisterra, G., Ritchie, K. P., Sadis, S. E., Lepock, J. R. & Hightower, L. E. (1996). Thermal activation of the bovine Hsc70 molecular chaperone at physiological temperatures: physical evidence of a molecular thermometer. *Cell Stress Chaperones* **1**, 78–89.

Li, J. & Liu, Q. (2009). 'Double water exclusion': a hypothesis refining the O-ring theory for the hot spots at protein interfaces. *Bioinformatics* **25**, 743–750.

Liepinsh, E. & Otting, G. (1997). Organic solvents identify specific ligand binding sites on protein surfaces. *Nature Biotechnology* **15**, 264–268.

Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **46**, 3–26.

Lise, S., Buchan, D., Pontil, M. & Jones, D. T. (2011). Predictions of hot spot residues at protein–protein interfaces using support vector machines. *PLoS ONE* **6**, e16774.

Liu, C. C. & Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annual Review of Biochemistry,* **79**, 413–444.

Liu, G., Huth, J. R., Olejniczak, E. T., Mendoza, R., Devries, P., Leitza, S., Reilly, E. B., Okasinski, G. F., Fesik, S. W. & Von Geldern, T. W. (2001). Novel p-arylthio cinnamides as antagonists of leukocyte function-associated antigen-1/intracellular adhesion molecule-1 interaction. 2. Mechanism of inhibition and structure-based improvement of pharmaceutical properties. *Journal of Medical Chemistry* **44**, 1202–1210.

Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M. & Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical Biochemistry* **332**, 153–159.

Luo, Q., Pagel, P., Vilne, B. & Frishman, D. (2011). DIMA 3.0: domain interaction map. *Nucleic Acids Research* **39** (Database issue), D724–D729.

Mattos, C. & Ringe, D. (1996). Locating and characterizing binding sites on proteins. *Nature Biotechnology* **14**, 595–599.

Mauser, H. & Guba, W. (2008). Recent developments in de novo design and scaffold hopping. *Current Opinion in Drug Discovery and Development* **11**, 365–374.

Mayer, M. & Meyer, B. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie International Edition* **38**, 1784–1788.

Meireles, L. M., Domling, A. S. & Camacho, C. J. (2010). ANCHOR: a web server and database for analysis of protein–protein interaction binding pockets for drug discovery. *Nucleic Acids Research* **38**(Web Server issue), W407–W411.

Meyer, B. & Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to

protein receptors. *Angewandte Chemie International Edition* **42**, 864–890.

Moellering, R. E., Cornejo, M., Davis, T. N., Del Bianco, C., Aster, J. C., Blacklow, S. C., Kung, A. L., Gilliland, D. G., Verdine, G. L. & Bradner, J. E. (2009). Direct inhibition of the NOTCH transcription factor complex. *Nature* **462**, 182–188.

Morita, M., Nakamura, S. & Shimizu, K. (2008). Highly accurate method for ligand-binding site prediction in unbound state (apo) protein structures. *Proteins* **73**, 468–479.

Muchmore, S. W., Olson, J., Jones, R., Pan, J., Blum, M., Greer, J., Merrick, S. M., Magdalinos, P. & Nienaber, V. L. (2000). Automated crystal mounting and data collection for protein crystallography. *Structure* **8**, R243–R246.

Murray, C. W. & Blundell, T. L. (2010). Structural biology in fragment-based drug design. *Current Opinion in Structural Biology* **20**, 497–507.

Navratilova, I., Besnard, J. & Hopkins, A. L. (2011). Screening for GPCR ligands using surface plasmon resonance. *ACS Medicinal Chemistry Letters* **2**, 549–554.

Navratilova, I. & Myszka, D. (2006). *Investigating Biomolecular Interactions and Binding Properties using SPR Biosensors*. New York: Springer.

Nienaber, V. L., Richardson, P. L., Klighofer, V., Bouska, J. J., Giranda, V. L. & Greer, J. (2000). Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnology* **18**, 1105–1108.

Nooren, I. M. & Thornton, J. M. (2003). Structural characterisation and functional significance of transient protein–protein interactions. *Journal of Molecular Biology* **325**, 991–1018.

Nordin, H., Jungnelius, M., Karlsson, R. & Karlsson, O. P. (2005). Kinetic studies of small molecule interactions with protein kinases using biosensor technology. *Analytical Biochemistry* **340**, 359–368.

Ogmen, U., Keskin, O., Aytuna, A. S., Nussinov, R. & Gursoy, A. (2005). PRISM: protein interactions by structural matching. *Nucleic Acids Research* **33** (Web Server issue), W331–W336.

Olsson, T. S., Williams, M. A., Pitt, W. R. & Ladbury, J. E. (2008). The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. *Journal of Molecular Biology* **384**, 1002–1017.

Oltersdorf, T., Elmore, S. W., Shoemaker, A. R., Armstrong, R. C., Augeri, D. J., Belli, B. A., Bruncko, M., Deckwerth, T. L., Dinges, J., Hajduk, P. J., Joseph, M. K., Kitada, S., Korsmeyer, S. J., Kunzer, A. R., Letai, A., Li, C., Mitten, M. J., Nettesheim, D. G., Ng, S., Nimmer, P. M., O'CONNOR, J. M., Oleksijew, A., Petros, A. M., Reed, J. C., Shen, W., Tahir, S. K., Thompson, C. B., Tomaselli, K. J., Wang, B., Wendt, M. D., Zhang, H., Fesik, S. W. & Rosenberg, S. H. (2005). An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435**, 677–681.

Orchard, S., Kerrien, S., Jones, P., Ceol, A., Chatr-Aryamontri, A., Salwinski, L., Nerothin, J. & Hermjakob, H. (2007). Submit your interaction data the IMEx way: a step by step guide to trouble-free deposition. *Proteomics* **7** (Suppl. 1), 28–34.

Pantolino, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P. & Salemme, F. R. (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of Biomolecular Screening* **6**, 429–440.

Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L. & Venkitaraman, A. R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287–293.

Perot, S., Sperandio, O., Miteva, M. A., Camproux, A. C. & Villoutreix, B. O. (2010). Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discovery Today* **15**, 656–667.

Perspicace, S., Banner, D., Benz, J., Muller, F., Schlatter, D. & Huber, W. (2009). Fragment-based screening using surface plasmon resonance technology. *Journal of Biomolecular Screening: The Official Journal of the Society for Biomolecular Screening* **14**, 337–349.

Petros, A. M., Dinges, J., Augeri, D. J., Baumeister, S. A., Betebenner, D. A., Bures, M. G., Elmore, S. W., Hajduk, P. J., Joseph, M. K., Landis, S. K., Nettesheim, D. G., Rosenberg, S. H., Shen, W., Thomas, S., Wang, X., Zanze, I., Zhang, H. & Fesik, S. W. (2006). Discovery of a potent inhibitor of the antiapoptotic protein Bcl-X_L from NMR and parallel synthesis. *Journal of Medicinal Chemistry* **49**, 656–663.

Pierce, M. M., Raman, C. S. & Nall, B. T. (1999). Isothermal titration calorimetry of protein–protein interactions. *Methods* **19**, 213–221.

Pitt, W., Higueruelo, A. & Todorov, N. (2010). Fragment-based methods for lead discovery. In *In Silico Lead Discovery* (ed. M. Miteva), pp. 83–116. Paris: Bentham Science Publishers.

Rajamani, D., Thiel, S., Vajda, S. & Camacho, C. J. (2004). Anchor residues in protein–protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11287–11292.

Razavi, H., Palaninathan, S. K., Powers, E. T., Wiseman, R. L., Purkey, H. E., Mohamedmohaideen, N. N., Deechongkit, S., Chiang, K. P., Dendle, M. T., Sacchettini, J. C. & Kelly, J. W. (2003). Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angewandte Chemie International Edition* **42**, 2758–2761.

Rohrig, C. H., Loch, C., Guan, J. Y., Siegal, G. & Overhand, M. (2007). Fragment-based synthesis and SAR of modified FKBP ligands: influence of different

linking on binding affinity. *Chem Med Chem* **2**, 1054–1070.

Rush, T. S., 3rd, Grant, J. A., Mosyak, L. & Nicholls, A. (2005). A shape-based 3-D scaffold hopping method and its application to a bacterial protein–protein interaction. *Journal of Medical Chemistry* **48**, 1489–1495.

Sasaki, K., Dockerill, S., Adamiak, D. A., Tickle, I. J. & Blundell, T. (1975). X-ray analysis of glucagon and its relationship to receptor binding. *Nature* **257**, 751–757.

Saxty, G., Woodhead, S. J., Berdini, V., Davies, T. G., Verdonk, M. L., Wyatt, P. G., Boyle, R. G., Barford, D., Downham, R., Garrett, M. D. & Carr, R. A. (2007). Identification of inhibitors of protein kinase B using fragment-based lead discovery. *Journal of Medical Chemistry* **50**, 2293–2296.

Schmidtke, P. & Barril, X. (2010). Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *Journal of Medical Chemistry* **53**, 5858–5867.

Schmidtke, P., Souaille, C., Estienne, F., Baurin, N. & Kroemer, R. T. (2010). Large-scale comparison of four binding site detection algorithms. *Journal of Chemical Information Model* **50**, 2191–2200.

Schneider, G. & Fechner, U. (2005). Computer-based de novo design of drug-like molecules. *Nature Reviews Drug Discovery* **4**, 649–663.

Schwyzer, R., Kriwaczek, V. M., Baumann, K., Haller, H.-R., Wider, G. & Wiltzius, A. P. (1979). Hormone-receptor interactions: a study of the binding of hormone-substituted tobacco mosaic virus to membrane vesicles by dynamic light scattering and by transient electric bi-refringence. *Pure and Applied Chemistry* **51**, 831–835.

Scott, D. E., Ehebauer, M. T., Pukala, T., Marsh, M., Blundell, T. L., Venkitaraman, A. R., Abell, C. & Hyvönen, M. (2012). Using a fragment-based approach to target protein–protein interactions. Manuscript in Preparation.

Segura, J. & Fernandez-Fuentes, N. (2011). PCRPi-DB: a database of computationally annotated hot spots in protein interfaces. *Nucleic Acids Research* **39** (Database issue), D755–D760.

Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534.

Sigurdardottir, A. (2012). *Targeting hepatocyte growth factor/scatter factor for drug discovery using a fragment-based approach*. PhD thesis, University of Cambridge.

Sigurskjold, B. W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Analytical Biochemistry* **277**, 260–266.

Sledz, P., Zheng, H., Murzyn, K., Chruszcz, M., Zimmerman, M. D., Chordia, M. D., Joachimiak, A. & Minor, W. (2010). New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Science: A Publication of the Protein Society* **19**, 1395–1404.

Smialowski, P., Pagel, P., Wong, P., Brauner, B., Dunger, I., Fobo, G., Frishman, G., Montrone, C., Rattei, T., Frishman, D. & Ruepp, A. (2010). The Negatome database: a reference set of non-interacting protein pairs. *Nucleic Acids Research* **38** (Database issue), D540–D544.

Spurlino, J. C. (2011). Fragment screening purely with protein crystallography. *Methods in Enzymology* **493**, 321–356.

Stein, A., Ceol, A. & Aloy, P. (2011). 3did: identification and classification of domain-based interactions of known three-dimensional structure. *Nucleic Acids Research* **39** (Database issue), D718–D723.

Stewart, M. L., Fire, E., Keating, A. E. & Walensky, L. D. (2010). The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nature Chemical Biology* **6**, 595–601.

Stockman, B. J., Kothe, M., Kohls, D., Weibley, L., Connolly, B. J., Sheils, A. L., Cao, Q., Cheng, A. C., Yang, L., Kamath, A. V., Ding, Y. H. & Charlton, M. E. (2009). Identification of allosteric PIF-pocket ligands for PDK1 using NMR-based fragment screening and ^{1}H-^{15}N TROSY experiments. *Chemical Biology and Drug Design* **73**, 179–188.

Sugase, K., Dyson, H. J. & Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025.

Sugaya, N. & Furuya, T. (2011). Dr. PIAS: an integrative system for assessing the druggability of protein–protein interactions. *BMC Bioinformatics* **12**, 50.

Sugaya, N. & Ikeda, K. (2009). Assessing the druggability of protein–protein interactions by a supervised machine-learning method. *BMC Bioinformatics* **10**, 263.

Sutherland, A. G., Alvarez, J., Ding, W., Foreman, K. W., Kenny, C. H., Labthavikul, P., Mosyak, L., Petersen, P. J., Rush, T. S., 3rd, Ruzin, A., Tsao, D. H. & Wheless, K. L. (2003). Structure-based design of carboxybiphenylindole inhibitors of the ZipA–FtsZ interaction. *Organic and Biomolecular Chemistry* **1**, 4138–4140.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L. J. & Von Mering, C. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research* **39** (Database issue), D561–D568.

Taylor, J. D., Gilbert, P. J., Williams, M. A., Pitt, W. R. & Ladbury, J. E. (2007). Identification of novel fragment compounds targeted against the pY pocket of v-Src SH2 by computational and NMR screening and thermodynamic evaluation. *Proteins* **67**, 981–990.

Teyra, J., Paszkowski-Rogacz, M., Anders, G. & Pisabarro, M. T. (2008). SCOWLP classification:

structural comparison and analysis of protein binding regions. *BMC Bioinformatics* **9**, 9.

Thanos, C. D., Delano, W. L. & Wells, J. A. (2006). Hot-spot mimicry of a cytokine receptor by a small molecule. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15422–15427.

Thompson, D., Denny, A., Nilakantan, R., Humblet, C., Joseph-Mccarthy, D. & Feyfant, E. (2008). CONFIRM: connecting fragments found in receptor molecules. *Journal of Computer-Aided Molecular Design* **22**, 761–772.

Thorn, K. S. & Bogan, A. A. (2001). ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics* **17**, 284–285.

Tilley, J. W., Chen, L., Fry, D. C., Emerson, S. D., Powers, G. D., Biondi, D., Varnell, T., Trilles, R., Guthrie, R., Mennona, F., Kaplan, G., Lemahieu, R. A., Carson, M., Han, R.-J., Liu, C.-M., Palermo, R. & Ju, G. (1997). Identification of a small molecule inhibitor of the IL-2/IL-2Rα receptor interaction which binds to IL-2. *Journal of the American Chemical Society* **119**, 7589–7590.

Tolcher, A., Yap, T., Fearen, I., Taylor, A., Carpenter, C. & Brunetto, A. (2007). A phase I study of MK-2206, an oral potent allosteric Akt inhibitor, in patients with advanced solid tumor. *Journal of Clinical Oncology* **27**, 3503.

Torres, F. E., Recht, M. I., Coyle, J. E., Bruce, R. H. & Williams, G. (2010). Higher throughput calorimetry: opportunities, approaches and challenges. *Current Opinion in Structural Biology* **20**, 598–605.

Traczewski, P. & Rudnicka, L. (2011). Treatment of systemic lupus erythematosus with epratuzumab. *British Journal of Clinical Pharmacology* **71**, 175–182.

Tsao, D. H., Sutherland, A. G., Jennings, L. D., Li, Y., Rush, III, T. S., Alvarez, J. C., Ding, W., Dushin, E. G., Dushin, R. G., Haney, S. A., Kenny, C. H., Malakian, A. K., Nilakantan, R. & Mosyak, L. (2006). Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorganic and Medicinal Chemistry* **14**, 7953–7961.

Tuncbag, N., Kar, G., Keskin, O., Gursoy, A. & Nussinov, R. (2009). A survey of available tools and web servers for analysis of protein–protein interactions and interfaces. *Briefings in Bioinformatics* **10**, 217–232.

Turinsky, A. L., Razick, S., Turner, B., Donaldson, I. M. & Wodak, S. J. (2011). Interaction databases on the same page. *Nature Biotechnology* **29**, 391–393.

Turnbull, W. B. & Daranas, A. H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *Journal of the American Chemical Society* **125**, 14859–14866.

Tversky, A. (1977). Features of Similarity. *Psychological Reviews* **84**, 327–352.

Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, Jr., P. J., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R. & Marsden, B. D. (2006). Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15835–15840.

Velazquez-Campoy, A., Kiso, Y. & Freire, E. (2001). The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. *Archives of Biochemistry and Biophysics* **390**, 169–175.

Velazquez-Campoy, A., Todd, M. J. & Freire, E. (2000). HIV-1 protease inhibitors: enthalpic *versus* entropic optimization of the binding affinity. *Biochemistry* **39**, 2201–2207.

Venkatesan, K., Rual, J. F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K. I., Yildirim, M. A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J. M., Cevik, S., Simon, C., De Smet, A. S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R. R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M. E., Roth, F. P., Hill, D. E., Tavernier, J., Wanker, E. E., Barabasi, A. L. & Vidal, M. (2009). An empirical framework for binary interactome mapping. *Nature Methods* **6**, 83–90.

Verlinde, C. L. M. J., Kim, H. B. B., Mande, S. C. & Wgj, H. (1997). *Anti-trypanosomiasis Drug Development Based on Structures of Glycolytic Enzymes*. New York: Marcel Dekker; Inc.

Vinkers, H. M., De Jonge, M. R., Daeyaert, F. F. D., Heeres, J., Koymans, L. M. H., Van Lenthe, J. H., Lewi, P. J., Timmerman, H., Van Aken, K. & Janssen, P. A. J. (2003). SYNOPSIS: SYNthesize and OPtimize System in Silico. *Journal of Medicinal Chemistry* **46**, 2765–2773.

Waldron, T. T. & Murphy, K. P. (2003). Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. *Biochemistry* **42**, 5058–5064.

Wang, L. & Schultz, P. G. (2004). Expanding the genetic code. *Angewandte Chemie International Edition* **44**, 34–66.

Wang, Z. X. (1995). An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Letters* **360**, 111–114.

Wells, J. & Mcclendon, C. (2007). Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009.

Wendt, M. D., Sun, C., Kunzer, A., Sauer, D., Sarris, K., Hoff, E., Yu, L., Nettesheim, D. G., Chen, J., Jin, S., Comess, K. M., Fan, Y., Anderson, S. N., Isaac, B., Olejniczak, E. T., Hajduk, P. J., Rosenberg, S. H. & Elmore, S. W. (2007). Discovery of a novel small

molecule binding site of human survivin. *Bioorganic and Medicinal Chemistry Lett* **17**, 3122–3129.

Wilson, A. J. (2009). Inhibition of protein–protein interactions using designed molecules. *Chemical Society Reviews* **38**, 3289–3300.

Winter, C., Henschel, A., Kim, W. K. & Schroeder, M. (2006). SCOPPI: a structural classification of protein–-protein interfaces. *Nucleic Acids Research* **34**(Database issue), D310–D314.

Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* **179**, 131–137.

Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *Journal of Molecular Biology* **293**, 321–331.

Wright, P. E. & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology* **19**, 31–38.

Wu, W. I., Voegtli, W. C., Sturgis, H. L., Dizon, F. P., Vigers, G. P. & Brandhuber, B. J. (2010). Crystal structure of human AKT1 with an allosteric inhibitor reveals a new mode of kinase inhibition. *PLoS ONE* **5**, e12913.

Xenarios, I., Rice, D. W., Salwinski, L., Baron, M. K., Marcotte, E. M. & Eisenberg, D. (2000). DIP: the database of interacting proteins. *Nucleic Acids Research* **28**, 289–291.

Xiang, Y., Hirth, B., Asmussen, G., Biemann, H. P., Bishop, K. A., Good, A., Fitzgerald, M., Gladysheva, T., Jain, A., Jancsics, K., Liu, J., Metz, M., Papoulis, A., Skerlj, R., Stepp, J. D. & Wei, R. R. (2011). The discovery of novel benzofuran-2-carboxylic acids as potent Pim-1 inhibitors. *Bioorganic and Medicinal Chemistry Lett* **21**, 3050–3056.

Xu, Q., Canutescu, A., Obradovic, Z. & Dunbrack, R. L., Jr. (2006). ProtBuD: a database of biological unit structures of protein families and superfamilies. *Bioinformatics* **22**, 2876–2882.

Yellaboina, S., Tasneem, A., Zaykin, D. V., Raghavachari, B. & Jothi, R. (2011). DOMINE: a comprehensive collection of known and predicted domain–domain interactions. *Nucleic Acids Research* **39**(Database issue), D730–D735.

Yin, H. & Hamilton, A. D. (2004). Terephthalamide derivatives as mimetics of the helical region of Bak peptide target Bcl-X_L protein. *Bioorganic and Medicinal Chemistry Lett* **14**, 1375–1379.

Young, A. B. (2003). Huntingtin in health and disease. *Journal of Clinical Investigation* **111**, 299–302.

Young, T. S. & Schultz, P. G. (2010). Beyond the canonical 20 amino acids: expanding the genetic lexicon. *The Journal of Biological Chemistry* **285**, 11039–11044.