Interface Focus

rsfs.royalsocietypublishing.org

Review

Cite this article: Sierecki E, Giles N, Polinkovsky M, Moustaqil M, Alexandrov K, Gambin Y. 2013 A cell-free approach to accelerate the study of protein–protein interactions *in vitro*. Interface Focus 3: 20130018.
http://dx.doi.org/10.1098/rsfs.2013.0018

One contribution of 6 to a Theme Issue 'Photons and biology'.

Subject Areas:
chemical biology, biophysics, biotechnology

Keywords:
protein–protein interactions, drug discovery, *in vitro* protein expression, AlphaScreen, single-molecule fluorescence

Author for correspondence:
Y. Gambin
e-mail: y.gambin@uq.edu.au

A cell-free approach to accelerate the study of protein–protein interactions *in vitro*

E. Sierecki, N. Giles, M. Polinkovsky, M. Moustaqil, K. Alexandrov and Y. Gambin

Institute for Molecular Bioscience, University of Queensland, Saint Lucia, Queensland, Australia

Protein–protein interactions are highly desirable targets in drug discovery, yet only a fraction of drugs act as binding inhibitors. Here, we review the different technologies used to find and validate protein–protein interactions. We then discuss how the novel combination of cell-free protein expression, AlphaScreen and single-molecule fluorescence spectroscopy can be used to rapidly map protein interaction networks, determine the architecture of protein complexes, and find new targets for drug discovery.

1. Introduction

For more than 20 years, targeting protein–protein interactions has been a dream in the drug discovery field. Protein complexes are ideal drug targets: they are ubiquitous, and molecules that disrupt binding would be very specific inhibitors, able to modulate protein activity in such a way that only certain pathological functions are inhibited, whereas other beneficial aspects are preserved. The majority of marketed drugs target less than 300 of the approximately 20 000 proteins encoded in the human genome, as only a minority of proteins have suitable binding sites for small molecules to modulate their activity. Typically, binding sites are restricted to grooves and clefts of enzyme catalytic sites, which are often highly conserved within protein families, compromising drug specificity. By contrast, the idea of targeting the complex and unique interfaces that characterize protein–protein interactions significantly expands the landscape of drug targets, and opens the possibility of highly specific, high-affinity binders.

Recent years have seen the discovery of a few small molecule inhibitors of interactions [1]. Of particular interest was the identification of inhibitors of the p53–mouse double minute 2 (MDM2) interaction [2–7] and its relevance to cancer [7,8]. The p53 tumour suppressor controls cell cycle arrest, and promotes DNA repair or apoptosis in response to cellular stress. Therefore, the p53 pathway is targeted in virtually every tumour type [9,10]. p53 activity is principally governed by its association with MDM2. MDM2 stops p53 action by three mechanisms. It prevents p53 from acting as a transcriptional activator by promoting the nuclear export of p53 and by binding to the p53 N-terminal transactivation domain. MDM2 also regulates p53 protein levels by ubiquitinating p53 (as its E3 ligase), targeting p53 for degradation in both the cytoplasm and the nucleus. Binding of MDM2 to a peptide derived from p53 was demonstrated in 1992 [11] and led to the development of new peptides with higher affinity and/or better stability. Vassilev *et al*. [2] reported that the Nutlin family was able to disrupt the MDM2–p53 interaction, leading to p53 pathway activation in cells and *in vivo* anti-tumour activity in a human cancer xenograft model. Phase I of the clinical trial of RG7112, a member of the Nutlin family developed by Roche, is underway for patients with advanced solid tumours and leukaemia. So far, it has demonstrated p53 stabilization and activation of p53 targets and the p53 pathway [12,13].

Inhibition of the MDM2–p53 interaction provides a new tool in the anti-cancer therapeutic arsenal and demonstrates the effectiveness of protein–protein

Table 1. Methods of validation of protein–protein interactions.

technique	high-throughput screening format	specific limits	general limits
fluorescence			
fluorescence polarization (FP)	yes	requires a size difference between the partners	fluorophore may interfere with folding or binding
Förster resonance energy transfer (FRET)	no	no study of large proteins or protein complexes requires labelling with synthetic dyes	
homogeneous time-resolved fluorescence (HTRF)	yes	no study of large proteins or protein complexes requires labelling with synthetic dyes	
complementation-based assay			
bimolecular fluorescence (BiFC)	yes	maturation time of the fluorophore	positioning and freedom of motion of the reporter fragments are crucial
protein fragment (PCA)	yes		
label-free assays			
isothermal titration calorimetry (ITC)	no	complex influence of solvent, pH	purification of the proteins is essential
surface plasmon resonance (SPR)	yes	non-specificity of surface binding ligand inflexibly bound to the surface	
impedance biosensor	yes	non-specificity of surface binding ligand inflexibly bound to the surface	

interactions as drug targets. However, such inhibitors are still rare, compared with activity modulators. Indeed, mapping protein–protein interactions presents specific challenges, mainly in the definition of the targets, existence of a suitable assay and access to proteins. We will discuss the challenges associated with the search and validation of novel protein targets. We will also present our current approach to studying protein assemblies, using a combination of cell-free protein expression (CFPE), AlphaScreen and single-molecule fluorescence spectroscopy. Finally, we will discuss the applicability of this combination of techniques to drug discovery.

The first challenge associated with inhibitors of protein–protein binding is the enormous complexity of the interaction networks present in the cell. *In vivo* immunoprecipitation, coupled to mass-spectrometric analysis and yeast-two-hybrid screens, has been used extensively for the colossal task of building a proteome-wide dataset of protein interactors. A number of databases reference potential interactions, and the interactome of *Saccharomyces cerevisiae* is almost completed.

Although genomic information is available, protein–protein interactions remain elusive targets. Indeed, validation of a protein–protein interaction is the main limitation. The experimental design used for discovery of interactors gives a number of both false-positives and false-negatives. False-positives arise from the inability to distinguish between direct and indirect interactions. This is particularly problematic when events require layers of protein–protein interactions and a specific hierarchy of binding events. False-negatives are due to a high positive signal threshold—only high-affinity complexes are targeted. However, some fundamental cell processes, such as signalling, rely on weak or transient interactions which would not be detected.

Currently, several different techniques are used to validate and quantify protein–protein interactions (table 1). They can be divided into several groups, depending on the strategy used to observe the interactions: fluorescence-based, complementation-based and label-free methods.

1.1. Fluorescence-based assays

To measure protein–protein interactions, fluorescence-based techniques rely on changes in the properties of light emitted by fluorophores bound to proteins. The fluorophores are either synthetic dyes that are bound to the molecules of interest through chemical processes or fluorescent proteins that are expressed intrinsically as part of the molecule. Synthetic dyes are usually smaller than the fluorescent proteins, but require longer and more careful handling of the sample before measurement. In either case, the concern of fluorescently tagging proteins is that the fluorescent tag may interfere with the folding or activity of the protein [14]. However, fluorescence-based methods are widely used because they provide a direct way to observe the behaviour of molecules using commonly available optical instruments.

Fluorescence polarization/anisotropy has the longest history and has been widely implemented [15]. This technique uses polarized excitation light to excite fluorophores whose excited state has dipoles parallel with the polarization direction. Depending on the lifetime of the excited state, the fluorophores may rotate before emitting fluorescence, and the direction of polarization of the emitted signal will be different from that of the excitation light. The rotational diffusion rate is approximately inversely proportional to the size of the molecule [16]. A typical experiment involves small ligands, with attached fluorophores, binding to larger proteins. If the ligands bind to the proteins, then the fluorophores' rotational diffusion is significantly decreased, which increases the anisotropy of the emitted light. This technique has been widely applied, owing to its versatility and relative simplicity, most notably in fluorescence polarization immunoassays [15].

Förster resonance energy transfer (FRET) has been increasingly used to study protein–protein interactions owing to its sensitivity to distances [17]. FRET involves the non-radiative transfer of energy from a photoexcited fluorophore (the donor) to another fluorophore in close proximity (the acceptor). The acceptor then emits fluorescent light at a higher

wavelength than the donor. By measuring the relative fluorescence of the two fluorophores, changes in the proximity of the molecules to which they are attached can be observed [18]. Because FRET efficiency has a $1/r^6$ distance dependence, it is well suited to distinguish interacting proteins from others in the focal volume [17]. At the same time, the extreme distance dependence prevents FRET from being useful to study interactions of large proteins, and limits the utility of intrinsically expressed fluorescent proteins as the dye pair.

FRET has also been extended, with time-resolved measurements, to a high-throughput format for drug discovery. The technique, called homogeneous time-resolved fluorescence (HTRF), uses rare-earth lanthanides in a molecular cage as donor fluorophores and compatible acceptors, both of which emit fluorescence over approximately 100 μs. By delaying the measurement of the fluorescence signal after an excitation pulse, the short-lifetime (approx. 1–10 ns) background fluorescence from the medium and impurities in the sample can be subtracted. The resulting improvement in the signal-to-noise ratio has allowed HTRF to be adapted successfully for parallelized screens in microplates. However, HTRF still suffers from several problems, including photobleaching of the fluorophores, and the other challenges of FRET, as mentioned above [19].

1.2. Complementation-based assays

In protein fragment complementation assays (PCAs), two proteins with separate, non-functional parts of a reporter bind to form a complex, subsequently making the reporter functional. Bimolecular fluorescence complementation (BiFC) is based on PCA, in which two non-fluorescent fragments derived from a fluorescent protein, such as GFP, are fused to a pair of interacting proteins. When the two partners interact, the two non-fluorescent fragments are brought into proximity and an intact fluorescent protein is reconstituted [20]. Binding between the proteins facilitates association between the fragments to produce a functional fluorophore [21]. This technique enables direct visualization of protein interactions and does not rely on measuring the activity of the protein complex. BiFC can be used to detect weak binding between proteins, as the association of the fluorophore fragments may stabilize the transient complex. One drawback to using BiFC assays is the slow maturation time of the fluorescent protein once the complex is formed. Thus, BiFC is not suitable for detecting protein complex formation in real time [22].

By analogy with BiFC, PCAs can involve a monomeric reporter enzyme separated into two inactive components that become functional when joined. Complementation assays based on several enzymes, including β-galactosidase, dihydrofolate reductase, β-lactamase and luciferases, have been used. Because the enzymatic activity of the reporter is directly reconstituted, assembly of protein complexes may be monitored in real time [23]. An important consideration in these complementation approaches is the positioning and freedom of motion of the reporter fragments on the interacting proteins, as they must be in proximity to form a functional reporter [22].

1.3. Label-free assays

Label-free methods use the intrinsic properties of the proteins of interest to detect their interactions. An advantage of these methods is in avoiding the problems of labelling the proteins and the perturbations to the system caused by the tags themselves. Surface plasmon resonance (SPR) and impedance biosensors both detect binding of molecules in solution to those bound on a surface by detecting changes in the surface properties. Isothermal titration calorimetry (ITC), on the other hand, seeks to directly measure the thermodynamic parameters of inter-molecular interactions.

The most straightforward application of ITC is to measure the strength of protein interactions. ITC can also provide the stoichiometric ratios of the participants in an interaction [24]. However, ITC data include many confounding factors, such as changes in the solvent and in proton concentration during binding, so care must be taken in interpreting the data [24,25]. While ITC can provide a range of information on the interactions of interest, it is not a high-throughput method, and requires significant amounts (at least micrograms) of the purified protein sample [24].

SPR biosensors generally use evanescent wave excitation of a thin layer of metal (i.e. approx. 50 nm gold layer) deposited on a glass substrate [26]. At a certain angle of the excitation light, called the SPR angle, the electrons in the metal layer can propagate the incident light over a wider cross section (approx. 100 nm) around the metal layer [27]. This effect is very sensitive to the local environment, and binding of molecules to the surface alters the SPR angle. Thus, SPR can be used to test the affinity of protein targets in the solution to ligands attached to the metal surface—when the protein targets bind to the ligand, the SPR angle changes. Some difficulties in using SPR include the possibility of non-specific binding of molecules to the metal surface [27] and the inflexibility of requiring the ligand to be attached to the surface.

Impedance biosensors use two electrodes, one patterned on the substrate, and another separated by the sample medium, to measure changes in electrical impedance when molecules bind and detach from the surface electrode [14]. Just as in SPR, ligands are attached to the surface electrode. When target proteins from the solution bind to the ligands on the surface, the electric impedance between the electrodes changes [28]. This change manifests itself as a different current response to a voltage applied across the electrodes [14]. In general, impedance biosensors exhibit the same issues as SPR: non-specific binding of molecules to the surface [14], and the necessity of having a specific ligand bound to the surface. Recently, attempts have been made to extend both impedance and SPR biosensors to high-throughput applications by using arrays of surface-bound ligands, similar to microarrays [14].

Protein–protein interactions can be validated by a variety of methods, but in most cases, access to the functional proteins themselves can become the limiting factor. Indeed, the assays mentioned above largely require purified proteins. But, as most therapeutic targets are human proteins, recombinant expression in *Escherichia coli* or yeast can be problematic and rate-limiting. Some valuable targets can also be lost due to known purification issues, i.e. transmembrane proteins or transcription factors binding to the genomic DNA. Finally, only a small fraction of the eukaryotic protein assemblies can be reconstituted from subunits that were expressed and purified separately, as the native assembly route relies on co-translational folding or assistance from chaperones. Simultaneous recombinant expression of multiple subunits has been achieved in insect cells and yeast systems, but this process is highly labour-intensive and scarcely used.

1.4. Cell-free expression provides rapid access to recombinant proteins *in vitro*

The development of polymerase chain reaction revolutionized biology and biotechnology by providing rapid access to any desired DNA sequence. The idea of obtaining proteins by a similar 'amplification' in a one-pot reaction is very appealing, with the goal of reducing the complexity of live cell cultures to a single reagent.

CFPE systems derived from whole cell extracts have emerged as an invaluable tool for both fundamental and applied biology mainly owing to their simplicity and efficiency. CFPE systems have several advantages over *in vivo* expression. They avoid host toxicity, are able to generate more stable and soluble proteins which would otherwise end up in inclusion bodies, enable manipulation at the genetic level and express multiple proteins simultaneously [29]. Further, they offer the option of using coupled (transcription and translation closely linked in time and space) or uncoupled reactions (transcription followed by translation) and labelling of translated proteins (biotin, ^{35}S-methionine/cysteine, ^{14}C leucine, Bodipy-FL lysine) and site-specific labelling of proteins [30,31]. To date, CFPE has contributed to the discovery of the genetic code [32], demonstrated its feasibility in protein synthesis technology [33,34], including the production of pharmaceutical proteins [35–37], and high-throughput production of protein libraries for protein evolution and structural genomics [38,39].

Currently, the more common CFPE extracts are derived from *E. coli*, rabbit reticulocytes (RRL), wheat germ (WGE) and insect cells, *Spodoptera frugiperda* (ICE), all of which are commercially available. The prokaryotic *E. coli* system is the most well-characterized and widely used. It is simple to use, the cheapest to manufacture, and can produce high protein yield of 500 $\mu g\ ml^{-1}$ to 1 $mg\ ml^{-1}$ in a 2 h batch reaction. However, its major limitation is the lack of eukaryotic chaperone proteins and co- and post-translational modification machinery required for producing functional eukaryotic proteins. The eukaryotic systems—RRL, ICE and WGE—are able to overcome these obstacles, but they are laborious and expensive to prepare, have significant variations between batches, are difficult to scale up reliably, are more difficult to manipulate genetically and yield a lower quantity of protein of 1–100 $\mu g\ ml^{-1}$ in a 2 h batch reaction. Beyond the eukaryotic CFPE systems mentioned above, other systems such as one based on yeast [40] have been developed, and human cell extract made from HeLa cells is commercially available.

We have recently exploited the unique gene expression machinery of *Leishmania* protozoa to create a new cell-free system termed LTE [41]. While *Leishmania tarentolae* has the advantage of being a eukaryotic system, it is also a fast-growing, fermentable organism that is relatively easy to genetically modify and process. A volume of 50 ml of extract can be produced from a 10 l batch culture, and the protein yield is comparable with commercially available systems, with 0.1 $mg\ ml^{-1}$ in a 2 h batch reaction and up to 0.5 $mg\ ml^{-1}$ in an overnight bi-layer reaction. In addition, *L. tarentolae* has become a more versatile and attractive system owing to the development of an expression vector containing species-independent translational sequences to enable initiation of cap-independent translation in almost all known cell-free expression systems.

Cell-free systems can provide easy and rapid access to proteins, including difficult targets such as membrane-embedded or membrane-binding proteins. These systems also have an enormous potential for rapid co-expression of proteins: more than one DNA sequence can be introduced into the system, an unlimited number of genes can be co-expressed and the relative gene dosage can be precisely controlled.

The main limitation of CFPE systems is the relatively low concentration of the expressed proteins, and the high background of the protein machinery required for translation and protein production. CFPE systems are crowded environments, and label-free techniques such as SPR and ITC would be overloaded with non-specific signals. In order to tackle this issue, we rely on extremely sensitive techniques such as AlphaScreen, coupled with single-molecule fluorescence methods for direct validation of hits and detailed mechanistic studies.

1.5. Interaction platform based on AlphaScreen technology and single-molecule fluorescence

Our team is developing a protein–protein interaction and drug-screening platform based on the cell-free expression system for rapid access to proteins, AlphaScreen assay for sensitive detection of interactions and single-molecule fluorescence to explore protein oligomerization and stoichiometry of interactions.

Amplified luminescent proximity homogeneous assay (ALPHA) screen is a technology developed in the mid-1990s and commercialized by Perkin-Elmer. It involves coating proteins on beads to locally increase concentrations giving this assay a wide detection range, from pM to mM affinities.

In short, a donor bead, coupled to protein A, contains a photosensitizer, phthalocyanine, which is excited by a laser pulse at 680 nm and emits singlet oxygen (figure 1*a*). An acceptor bead, coupled to protein B, is filled with thioxene derivatives that react with singlet oxygen and emit luminescence between 520 and 620 nm. In solution, a singlet oxygen can travel only up to 200 nm before returning to its stable triplet state. This means that a signal can be detected only if the beads are less than 200 nm away, that is, if protein A and protein B interact.

A key characteristic of ALPHA screen is the bell-shape response to protein concentration (figure 1*b*). The signal increases with the protein concentration up to a maximum. This behaviour is due to the loading of the beads; as more proteins are added, more beads are able to interact and the signal increases. However, increasing the protein concentration beyond the optimum results in a loss of signal. At this point, the excess proteins in solution compete with the proteins bound to the beads, resulting in a loss of interaction between beads and a decrease in signal.

This particularity can be exploited to confirm that hits are true positives. A competition assay consists of adding an excess of one of the interactors (preferably without an interacting tag, to prevent it from binding to the beads). If the signal is due to interaction between two proteins, the system is in the 'overloading' configuration and the signal is lost. When the interaction between the beads is non-specific, further addition of proteins will have no effect.

Practical considerations also make AlphaScreen attractive. AlphaScreen beads that interact with a variety of peptidic, non-peptidic and Fc-tags are available. We can use proteins directly in the expression mix, without additional purification. The assay is run in a 384-well format and is amenable to automation.

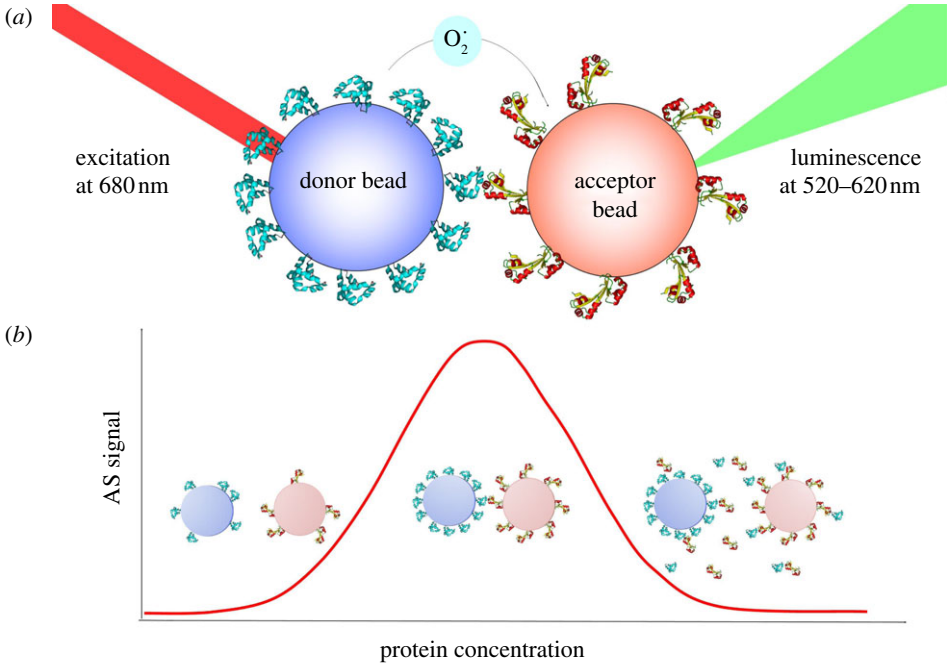

Figure 1. Principle of the ALPHA screen. (*a*) A donor bead, coated with protein A is excited by a laser pulse at 680 nm. This triggers the release of singlet oxygen, which has a half-life of 4 µs and can diffuse over 200 nm. If an acceptor bead, coated with protein B, is brought into proximity of the donor bead by the interaction between protein A and protein B, luminescence between 520 and 620 nm is emitted as a result of a reaction between singlet oxygen and thioxene derivatives encapsulated in the acceptor bead. (*b*) The 'hooking' effect is a signature of the AlphaScreen (AS) signal and is linked to bead loading. When the concentration of proteins is too low, the signal is low as only few interactions occur, keeping the beads far from each other. At optimal protein concentration, all the beads are covered with proteins and bead–bead contacts are maximized, leading to a high signal. When the system is overloaded with proteins, free-floating proteins can compete with bead-immobilized proteins for interaction partners, diminishing the number of bead–bead interactions and the resulting signal.

ALPHA screen appears in most publications as a validation technique for protein–protein interactions [42]. It has also been used as a tool to identify interaction domains [43] and to study the effect of mutations on protein association [44]. In a few papers, ALPHA screen technology is used as a drug-screening platform [45–47] and as a protein–protein interaction detection platform [48,49], notably in cells.

In all assays, the question of aggregation and functionality of expressed proteins has to be considered. How can we distinguish between weakly bound monomers and a mixture of tight binders and non-functional oligomers or aggregates? How can we make sure that the interactions detected are specific and not owing to coprecipitation of misfolded proteins? This is extremely difficult to assess with conventional techniques, such as size exclusion chromatography and ultracentrifugation, which require protein purification and amounts of sample that are not compatible with CFPE.

1.6. Monomer, defined oligomer or random aggregate? Single-molecule fluorescence provides a unique protein quality control

During the past two decades, novel fluorescence methods, especially single-molecule fluorescence [50], have emerged with the promise of revealing hidden information about protein folding and interactions [51].

In conventional ensemble experiments, large numbers of proteins are interrogated simultaneously, and their averaged properties are observed. By contrast, single-molecule experiments can provide an incredible wealth of information by directly measuring the distributions of molecular properties. Single-molecule fluorescence methods are especially attractive and powerful for the study of complex biological processes as they enable direct quantification of folded versus unfolded proteins and can reveal rare populations (oligomers or aggregates) which would remain buried in the averaged values.

The principles and advantages of single-molecule fluorescence can be easily explained from a known technique, fluorescence correlation spectroscopy (FCS). The method is based on a simple idea [52]: on a confocal microscope, an excitation laser creates a very small observation volume (approx. 1 fl). The proteins are tagged with fluorophores, either genetically encoded (GFP and RFP), or bright organic dyes (such as the Alexa or the Cy series). Proteins diffuse freely by Brownian motion, and they enter and exit the confocal volume constantly.

FCS is a 'small ensemble' technique: at all times, a few proteins—ideally between 20 and 50—are present in the focal volume. The total measured fluorescence fluctuates around an average value—every time a protein exits or enters the detection volume, the overall fluorescence drops or increases accordingly (figure 2*a*). The fluctuations of intensity are computed, and the autocorrelation of the intensity over time leads to a calculation of the diffusion time, the typical time it takes for a protein to diffuse through the focal volume. Binding between proteins or the formation of aggregates can be detected, as the physical size, and consequently the diffusion time, will increase upon complex formation.

Single-molecule detection would require dropping the concentration lower, typically less than 100 pM, in order to observe single proteins [53]. In that case, the measured fluorescence trace is mostly background, with bright bursts signalling when a molecule enters the focal volume (figure 2*b*). It is possible to analyse the brightness of the

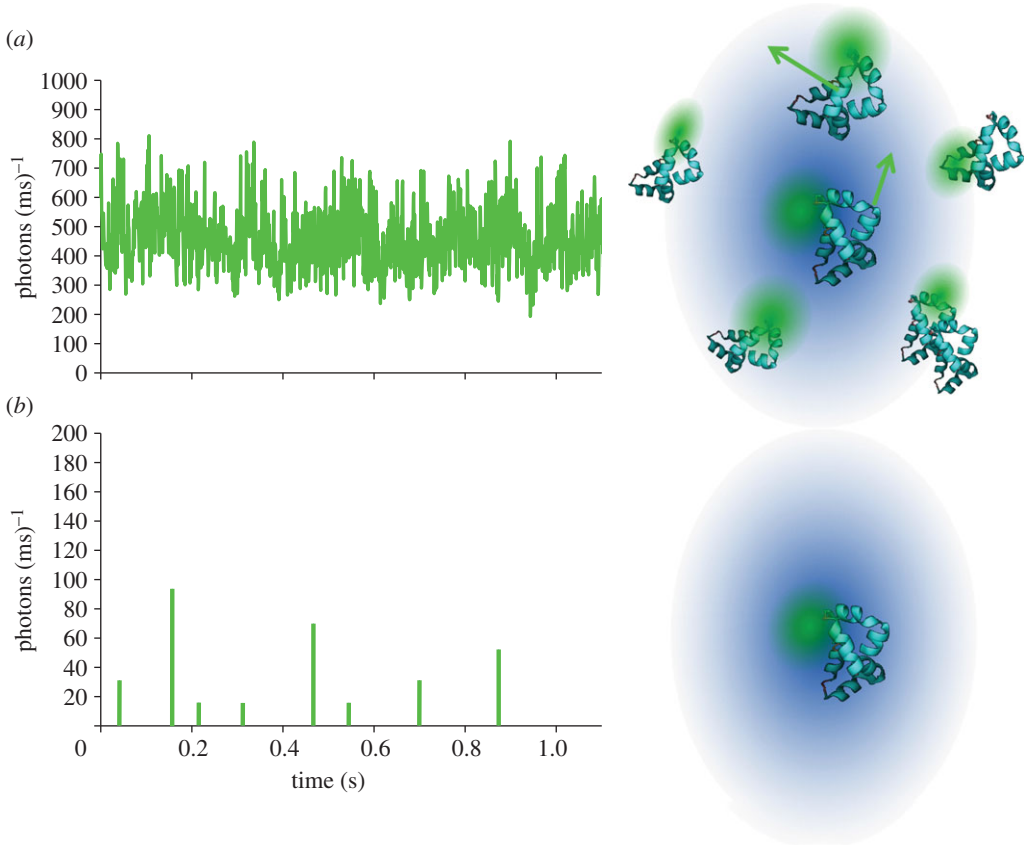

Figure 2. Fluorescent traces typically obtained for (*a*) FCS and (*b*) single-molecule experiments. Both experiments are conducted on a confocal microscope, with a small detection volume. The FCS experiments are conducted at protein concentrations in the nanomolar range; fluctuations of the average fluorescence intensity are due to movements of the fluorescent proteins, and the signal is analysed to obtain residence times in the detection volume. Single-molecule experiments require picomolar concentrations of proteins to detect single fluorescent 'events', corresponding to single proteins transiting the focal volume.

burst (the number of photons) to estimate the number of fluorophores present, as we expect a dimer to be twice as bright as a monomer.

The evident advantage of such single-molecule 'counting' is the direct access to distributions of oligomeric states; each burst of fluorescence (each protein) can theoretically be categorized as monomer, dimer or oligomer. The technique enables initial testing of the proteins' 'folding': most unfolded proteins should have hydrophobic residues exposed and form aggregates. In the case of aggregates, one would expect random and disperse sizes to be observed, whereas defined oligomers of smaller, more homogeneous size could reveal that some proteins are natively in a functional oligomeric form.

Although single-molecule technology has become more popular in the past decade, three critical obstacles have prevented its wider spread and its application to protein–protein interactions and protein complexes. First, its throughput is still far from adequate for screening assays, as it takes 10–15 min to collect a sufficient number of protein events for a single interaction. Second, owing to the low-nanomolar concentrations required for single-molecule detection, the methods cannot be applied to the analysis of low affinity interactions. Most biologically relevant complexes have typical dissociation constants (K_d) in the high mM range and dissociate quickly upon dilution [54]. Finally, genetically encoded fluorophores such as GFP can be used for FCS, but are not sufficiently bright for single-molecule detection of freely diffusing proteins. In general, single-molecule experiments are not yet connected with CFPE. Even if the experiments require only a few thousand proteins to obtain sufficient statistics, the proteins are generally expressed on a large scale, purified, selectively labelled with organic dyes and repurified before measurement.

An interesting option is a hybrid approach between single molecule and FCS, in an attempt to remove the 'averaging' of multiple diffusion times in complex systems. The measure relies on genetically encoded fluorophores, and can be directly performed in cell-free extracts, in order to eliminate purification steps that would disrupt interactions and complexes.

Figure 3 demonstrates the various analyses that can be conducted, comparing the behaviour of a correctly folded monomeric protein (blue) and a misbehaving protein that has a tendency to form aggregates (red).

This hybrid technique is perfectly suited to detect the formation of protein complexes and oligomers in real time. We routinely use these approaches to perform the essential quality control of the proteins expressed in the cell-free system and detect false-positives.

Using single-molecule fluorescence, we can obtain a more detailed picture of the interactions, study stoichiometries and binding mechanisms, and directly validate the mode of action of inhibitors.

These experiments require labelling of the two proteins with two different fluorophores (such as Alexa488 and Alexa594), for two-colour simultaneous detection or single-molecule 'coincidence' experiments. In these experiments, two lasers are turned on continuously, exciting the same focal volume. The fluorescence from the two fluorophores (proteins) is collected separately in two channels, and we analyse the coincidence of the 'bursts' in the two channels.

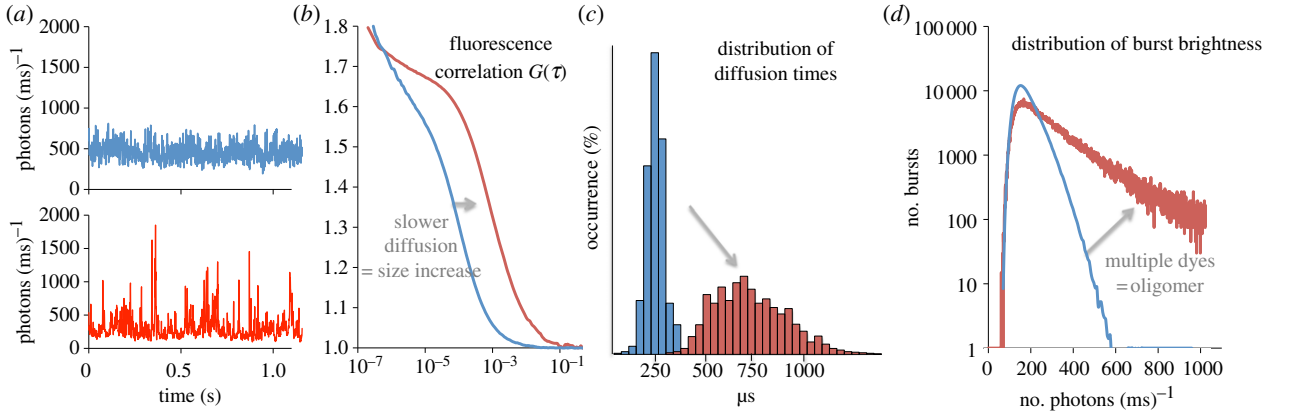

Figure 3. Principle of oligomer detection using FCS or single-molecule analysis. (*a*) By dropping protein concentration to 1 nM, one can no longer distinguish individual well-separated bursts, but the intensity would show bursts of large amplitude in the case of aggregates. (*b*) FCS autocorrelation curves obtained for the monomer and the aggregate show different diffusion times. These curves have been obtained by continuous measurement for 30 s. If the number of aggregates is small, then it would become difficult to fit the FCS curve and observe this subpopulation, as it would remain buried in the average with thousands of monomers. (*c*) Acquisition of very short FCS curves, and plot of the distributions of diffusion times obtained. Measures can be as short as 3 s, and more weight is given to bright events and slow diffusion (aggregates of many fluorophores). In the case of a well-behaved system, the distribution of times (sizes) will be narrow, but as more oligomers or aggregates are present, the distribution will widen towards higher values; typically, if even one aggregate is detected in every 3 s window, the whole distribution will shift considerably. (*d*) Analysis of the intensity traces based on brightness. Indeed, the two traces in (*a*) show that the most evident properties of oligomers are the high-intensity bursts they create. The distribution of burst brightness can be obtained quickly (5 s), and reveals small deviations from the expected monomeric distribution.

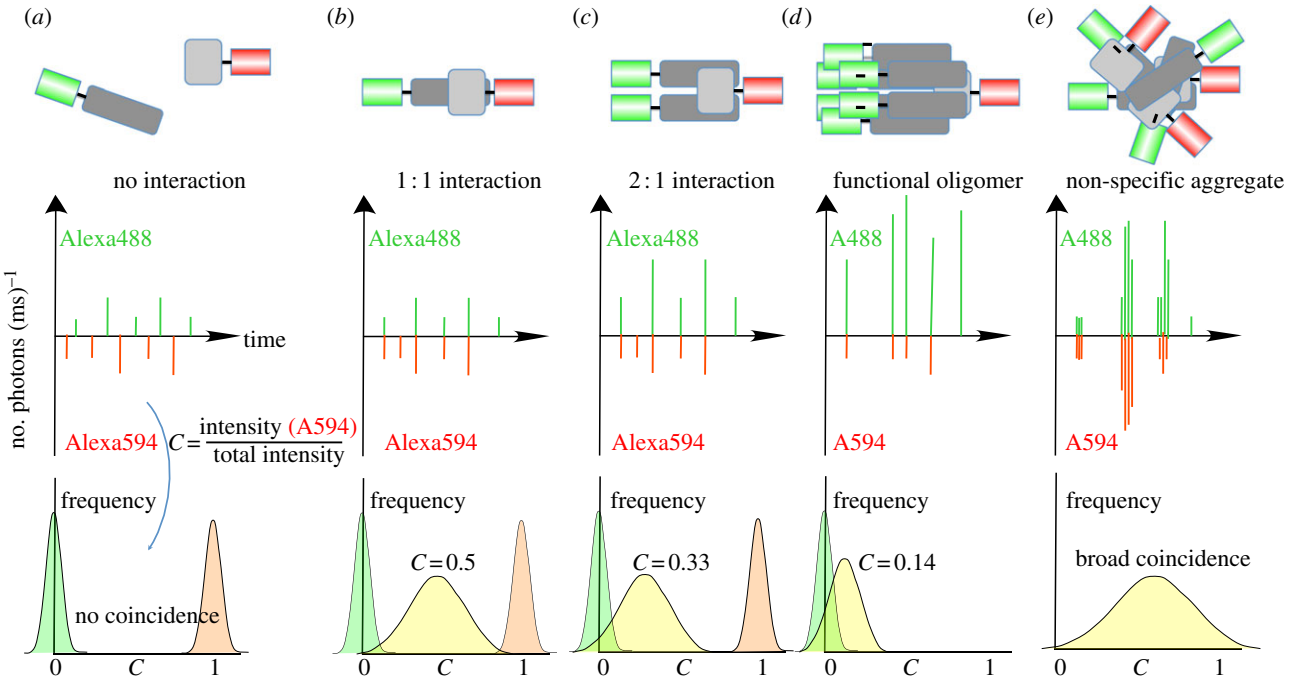

Figure 4. Two-colour coincidence experiments and signals expected in various scenarios. (*a*) The case of no interaction and (*b*) the case of a 1 : 1 interaction between two proteins labelled with Alexa488 and Alexa594. In the absence of interaction, the bursts of fluorescence in the two channels will be decorrelated. We define a coincidence ratio, *C*, as the intensity in the A594 channel divided by the total fluorescence. If only the A594-tagged protein is present, $C = 1$, and if only the A488-tagged protein is present, $C = 0$. When two proteins bind, the bursts of fluorescence in both channels will correspond, and we will measure coincidence ratios centred around $C = 0.5$. We plot single-molecule histograms of *C*-values for each burst of fluorescence to detect binding. (*c*,*d*) The brightness of the bursts and the position of the coincidence peak will shift according to stoichiometries, and we can detect the presence of functional oligomers. (*e*) In the case of non-functional aggregates, we typically observe long fluorescent bursts of large, slow-diffusing objects and very broad coincidence as random numbers of proteins are involved.

As described in figure 4, this technique can precisely quantify the fraction of bound proteins and differentiate between different stoichiometries.

2. Perspectives

Using the combination of cell-free expression system, AlphaScreen and single-molecule spectroscopy, we are setting up a platform to identify valuable protein–protein interaction targets and develop drug-screening campaigns (figure 5).

For each protein of interest, potential interactors, such as possible activators and repressors of a given transcription factor, are selected from the literature. We are fortunate to have access to the human ORFeome, a library of 20 000 open reading frames (ORFs). The relevant ORFs can be selected and cloned into a series of gateway vectors compatible with a

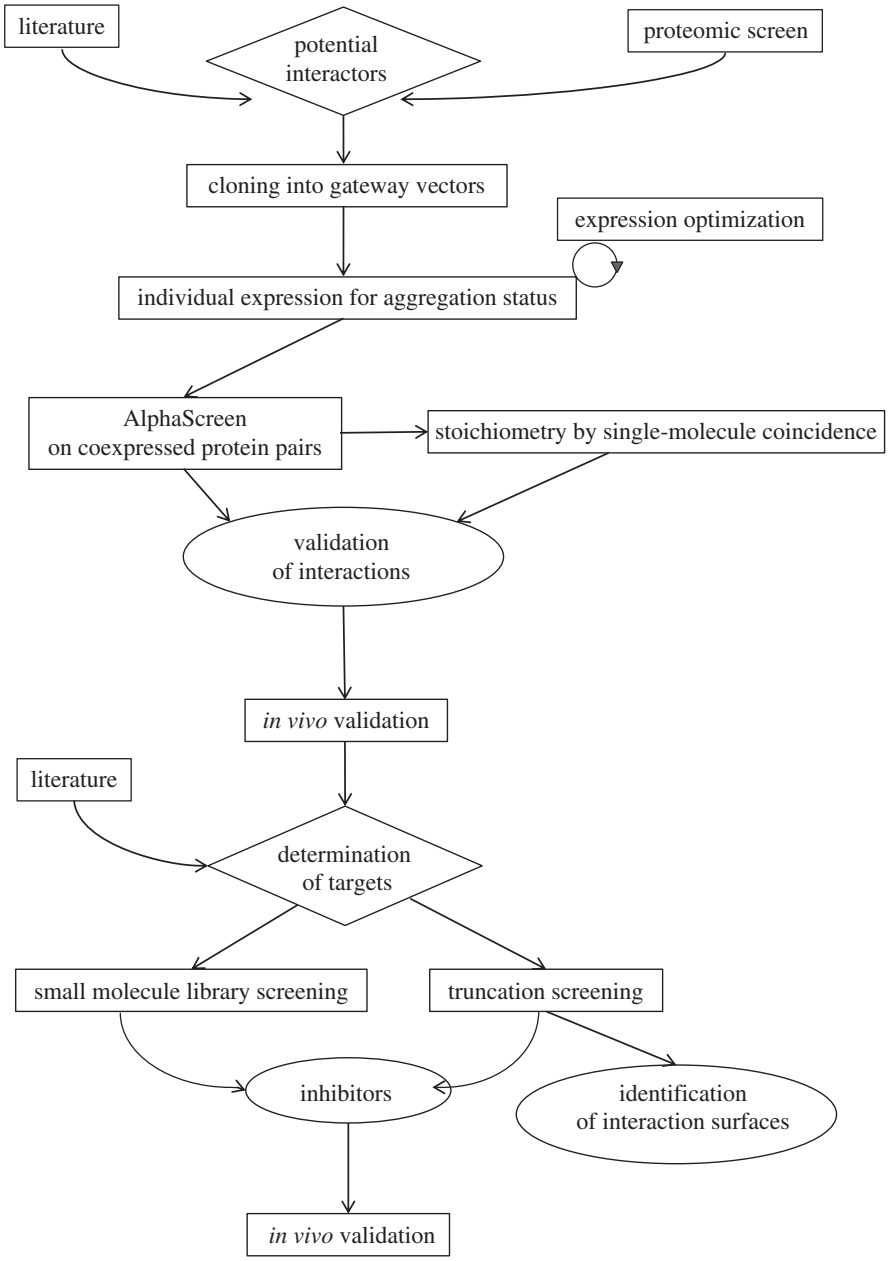

Figure 5. Typical flowchart for the validation of protein interactome and screening for inhibitors of protein–protein interactions.

cell-free expression system from *L. tarentolae*. Fusion proteins bearing a fluorescent tag (GFP or Cherry) can then be obtained.

The proteins are first expressed independently and their aggregation status is assessed by single-molecule 'counting'. The parameters of the cell-free system can be tuned to avoid excessive aggregation (lower lysate concentration, shorter expression times, etc.). Pairwise interactions are then tested using AlphaScreen technology. To minimize the possibility of false-positives, all the interactions are tested in two configurations with N-terminal or C-terminal tagging of the proteins. With this experimental set-up, we expect to detect cases where the fluorescent tag interferes in the interaction. AlphaScreen provides rapid access to interaction heat maps as we can routinely analyse more than a hundred pairs per day.

The positive hits are further characterized by single-molecule coincidence. The stoichiometry of the interactions provides valuable structural information and is difficult to obtain by other techniques. Indeed, a binding event between two proteins may be impossible to understand if the stoichiometry data are missing, as oligomers can create interaction surfaces that do not exist in the monomeric form.

The most biologically relevant interactions can then be subjected to drug screening. Once again, the throughput of the AlphaScreen assays is exploited. Two sources of inhibitors are available. The traditional option is to screen established libraries of small molecules, either from synthetic or natural origins. Interestingly, we verified that most of the compounds in the LOPAC library do not affect the protein transcription in our cell-free expression system, allowing the targeting of interactions during coexpression. Another source of inhibitors is peptides obtained from truncations of one of the partners. This option has the added bonus of providing a wealth of structural information on the binding surface. It is particularly appealing when targeting two different interactions, on two different binding sites. The identified hits can then be coupled to cell-penetrating peptides for cell delivery or serve as leads for the development of peptidomimetics.

Using this combination of techniques, we are able to validate interaction partners, characterize the complexes, define the important residues and identify potential inhibitors. The single-molecule coincidence experiments can validate the

hits and help define the mode of action of the inhibitor. This is especially interesting in complex systems, when one of the proteins forms dimers or oligomers, as we can detect whether the inhibitor disrupts the oligomer or blocks binding to the oligomer. The novel combination of techniques enables us to perform these steps in a reasonable amount of time, before the longer and more difficult validation of inhibitors inside a cell or an organism.

3. Conclusion

Unravelling the enormous complexity of protein interaction networks represents one of the greatest challenges in biological research. With the availability of genome sequences, a detailed molecular dissection of the main hubs of structural, metabolic and signalling networks in cells is long overdue. Protein complexes with more than four non-identical subunits represent more than 20% of the eukaryotic proteome, yet they account for less than 1% of biophysically characterized proteins [55]. We believe that using cell-free protein co-expression coupled to AlphaScreen and single-molecule fluorescence should accelerate the study of multi-subunit protein complexes and facilitate drug discovery on protein–protein interactions.

Acknowledgements. K.A. and Y.G. are supported by an Australian Research Council (ARC) Future Fellowship (FT0991611 and FT110100478).

Funding statement. This research is supported by ARC Discovery Project (grant nos. DP130102396 and DP120101423), by a National Health and Medical Research Council Project (grant no. APP1025082) and an NHMRC Programme (grant nos. 511005 and APP1037320).

References

1. Fletcher S, Hamilton AD. 2007 Protein–protein interaction inhibitors: small molecules from screening techniques. *Curr. Top. Med. Chem.* **7**, 922–927. (doi:10.2174/156802607780906735)
2. Vassilev LT *et al*. 2004 *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–848. (doi:10.1126/science.1092472)
3. Galatin PS, Abraham DJ. 2004 A nonpeptidic sulfonamide inhibits the p53-MDM2 interaction and activates p53-dependent transcription in mdm2-overexpressing cells. *J. Med. Chem.* **47**, 4163–4165. (doi:10.1021/jm034182u)
4. Shangary S *et al*. 2008 Reactivation of p53 by a specific MDM2 antagonist (MI-43) leads to p21-mediated cell cycle arrest and selective cell death in colon cancer. *Mol. Cancer Ther.* **7**, 1533–1542. (doi:10.1158/1535-7163.MCT-08-0140)
5. Ding J, Zhang Y, Zheng J. 2008 Reactivation of p53 by inhibiting Mdm2 E3 ligase: a novel antitumor approach. *Curr. Cancer Drug Targets* **11**, 987–994.
6. Zhuang C *et al*. 2012 Discovery, synthesis, and biological evaluation of orally active pyrrolidone derivatives as novel inhibitors of p53-MDM2 protein–protein interaction. *J. Med. Chem.* **55**, 9630–9642. (doi:10.1021/jm300969t)
7. Millard M, Pathania D, Grande F, Xu S, Neamati N. 2011 Small-molecule inhibitors of p53-MDM2 interaction: the 2006–2010 update. *Curr. Pharm. Des.* **17**, 536–559. (doi:10.2174/138161211795222649)
8. Vu BT, Vassilev L. 2011 Small-molecule inhibitors of the p53-MDM2 interaction. *Curr. Top. Microbiol. Immunol.* **348**, 151–172. (doi:10.1007/82_2010_110)
9. Muller PA, Vousden KH. 2013 p53 mutations in cancer. *Nat. Cell Biol.* **15**, 2–8. (doi:10.1038/ncb2641)
10. Miliani de Marval PL, Zhang Y. 2011 The RP-Mdm2-p53 pathway and tumorigenesis. *Oncotarget* **2**, 234–238.
11. Momand J, Zambetti GP, Olson DC, George D, Levine AJ. 1992 The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69**, 1237–1245. (doi:10.1016/0092-8674(92)90644-R)
12. Andreeff M *et al*. 2011 A multi-center, open-label, phase I study of single agent RG7112, a first in class p53-MDM2 antagonist, in patients with relapsed/refractory acute myeloid and lymphoid leukemias (AML/ALL) and refractory chronic lymphocytic leukemia/small cell lymphocytic lymphomas (CLL/SCLL). In *53rd ASH Annual Meeting and Exposition, New Orleans, LA*.
13. Andreeff M *et al*. 2012 Results of the phase 1 trial of RG7112, a small-molecule MDM2 antagonist, in acute leukemia. In *54th ASH Annual Meeting and Exposition, Atlanta, GA*.
14. Daniels JS, Pourmand N. 2007 Label-free impedance biosensors: opportunities and challenges. *Electroanalysis* **19**, 1239–1257. (doi:10.1002/elan.200603855)
15. Jameson DM, Ross JA. 2010 Fluorescence polarization/anisotropy in diagnostics and imaging. *Chem. Rev.* **110**, 2685–2708. (doi:10.1021/cr900267p)
16. Gradinaru CC, Marushchak DO, Samim M, Krull UJ. 2010 Fluorescence anisotropy: from single molecules to live cells. *Analyst* **135**, 452–459. (doi:10.1039/b920242k)
17. Kenworthy AK. 2001 Imaging protein–protein interactions using fluorescence resonance energy transfer microscopy. *Methods* **24**, 289–296. (doi:10.1006/meth.2001.1189)
18. Yan Y, Marriott G. 2003 Analysis of protein interactions using fluorescence technologies. *Curr. Opin. Chem. Biol.* **7**, 635–640. (doi:10.1016/j.cbpa.2003.08.017)
19. Degorce F, Card A, Soh S, Trinquet E, Knapik GP, Xie B. 2009 HTRF: a technology tailored for drug discovery: a review of theoretical aspects and recent applications. *Curr. Chem. Genomics* **3**, 22–32. (doi:10.2174/1875397300903010022)
20. Barnard E, McFerran N, Trudgett A, Nelson J, Timson DJ. 2008 Development and implementation of split-GFP-based bimolecular fluorescence complementation (BiFC) assays in yeast. *Biochem. Soc. Trans.* **36**(Pt 3), 479–482. (doi:10.1042/BST0360479)
21. Hu C-D, Kerppola TK. 2003 Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotech.* **21**, 539–545. (doi:10.1038/nbt816)
22. Kerppola TK. 2008 Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.* **37**, 465–487. (doi:10.1146/annurev.biophys.37.032807.125842)
23. Michnick SW, Ear PH, Manderson EN, Remy I, Stefan E. 2007 Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat. Rev. Drug Discov.* **6**, 569–582. (doi:10.1038/nrd2311)
24. Ghai R, Falconer RJ, Collins BM. 2012 Applications of isothermal titration calorimetry in pure and applied research: survey of the literature from 2010. *J. Mol. Recognit.* **25**, 32–52. (doi:10.1002/jmr.1167)
25. Ladbury JE. 2010 Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Biochem. Soc. Trans.* **38**, 888–893. (doi:10.1042/BST0380888)
26. Jason-Moller L, Murphy M, Bruno J. 2001 Overview of Biacore systems and their applications. In *Current protocols in protein science*. London, UK: John Wiley & Sons Inc. (doi:10.1002/0471140864.ps1913s45)
27. Nedelkov D, Nelson RW. 2003 Surface plasmon resonance mass spectrometry: recent progress and outlooks. *Trends Biotechnol.* **21**, 301–305. (doi:10.1016/S0167-7799(03)00141-0)
28. McGuinness R. 2007 Impedance-based cellular assay technologies: recent advances, future promise. *Curr. Opin. Pharmacol.* **7**, 535–540. (doi:10.1016/j.coph.2007.08.004)
29. Ohuchi S, Nakano H, Yamane T. 1998 *In vitro* method for the generation of protein libraries using PCR amplification of a single DNA molecule and

coupled transcription/translation. *Nucleic Acids Res.* **26**, 4339–4346. (doi:10.1093/nar/26.19.4339)

30. Goerke AR, Swartz JR. 2009 High-level cell-free synthesis yields of proteins containing site-specific non-natural amino acids. *Biotechnol. Bioeng.* **102**, 400–416. (doi:10.1002/bit.22070)
31. Saraogi I, Zhang D, Chandrasekaran S, Shan S. 2011 Site-specific fluorescent labeling of nascent proteins on the translating ribosome. *J. Am. Chem. Soc.* **133**, 14 936–14 939. (doi:10.1021/ja206626g)
32. Nirenberg MW, Matthaei JH. 1961 The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl Acad. Sci. USA* **47**, 1588–1602. (doi:10.1073/pnas.47.10.1588)
33. Katzen F, Chang G, Kudlicki W. 2005 The past, present and future of cell-free protein synthesis. *Trends Biotechnol.* **23**, 150–156. (doi:10.1016/j.tibtech.2005.01.003)
34. Swartz J. 2006 Developing cell-free biology for industrial applications. *J. Ind. Microbiol. Biotechnol.* **33**, 476–485. (doi:10.1007/s10295-006-0127-y)
35. Goerke AR, Swartz JR. 2008 Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol. Bioeng.* **99**, 351–367. (doi:10.1002/bit.21567)
36. Kanter G, Yang J, Voloshin A, Levy S, Swartz JR, Levy R. 2007 Cell-free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. *Blood* **109**, 3393–3399. (doi:10.1182/blood-2006-07-030593)
37. Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ. 2011 Microscale to manufacturing scale-up of cell-free cytokine production: a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**, 1570–1578. (doi:10.1002/bit.23103)
38. Goshima N *et al*. 2008 Human protein factory for converting the transcriptome into an *in vitro*-expressed proteome. *Nat. Methods* **5**, 1011–1017. (doi:10.1038/nmeth.1273)
39. Griffiths AD, Tawfik DS. 2003 Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J.* **22**, 24–35. (doi:10.1093/emboj/cdg014)
40. Wang X *et al*. 2008 An optimized yeast cell-free system: sufficient for translation of human papillomavirus 58 L1 mRNA and assembly of virus-like particles. *J. Biosci. Bioeng.* **106**, 8–15. (doi:10.1263/jbb.106.8)
41. Mureev S, Kovtun O, Nguyen UTT, Alexandrov K. 2009 Species-independent translational leaders facilitate cell-free expression. *Nat. Biotech.* **27**, 747–752. (doi:10.1038/nbt.1556)
42. Eglen RM, Reisine T, Roby P, Rouleau N, Illy C, Bossé R, Bielefeld M. 2008 The use of AlphaScreen technology in HTS: current status. *Curr. Chem. Genomics* **1**, 2–10. (doi:10.2174/1875397300801010002)
43. Takahashi H, Takahashi C, Moreland NJ, Chang Y-T, Sawasaki T, Ryo A, Vasudevan SG, Suzuki Y, Yamamoto N. 2012 Establishment of a robust dengue virus NS3–NS5 binding assay for identification of protein–protein interaction inhibitors. *Antiviral Res.* **96**, 305–314. (doi:10.1016/j.antiviral.2012.09.023)
44. Kaur G, Delluc-Clavieres A, Poon IKH, Forwood JK, Glover DJ, Jans DA. 2010 Calmodulin-dependent nuclear import of HMG-box family nuclear factors: importance of the role of SRY in sex reversal. *Biochem. J.* **430**, 39–48. (doi:10.1042/BJ20091758)
45. Rayter S *et al*. 2007 A chemical inhibitor of PPM1D that selectively kills cells overexpressing PPM1D. *Oncogene* **27**, 1036–1044. (doi:10.1038/sj.onc.1210729)
46. Hu G, Li X, Li Y, Sun X, Liu G, Li W, Huang J, Shen X, Tang Y. 2012 Inhibitors of HIV-1 integrase-human LEDGF/p75 interaction identified from natural products via virtual screening. *Chin. J. Chem.* **30**, 2752–2758. (doi:10.1002/cjoc.201200897)
47. Demeulemeester J, Tintori C, Botta M, Debyser Z, Christ F. 2012 Development of an AlphaScreen-based HIV-1 integrase dimerization assay for discovery of novel allosteric inhibitors. *J. Biomol. Screen.* **17**, 618–628. (doi:10.1177/1087057111436343)
48. Madono M, Sawasaki T, Morishita R, Endo Y. 2011 Wheat germ cell-free protein production system for post-genomic research. *New Biotechnol.* **28**, 211–217. (doi:10.1016/j.nbt.2010.08.009)
49. Mohamed MR *et al*. 2009 Proteomic screening of variola virus reveals a unique NF-κB inhibitor that is highly conserved among pathogenic orthopoxviruses. *Proc. Natl Acad. Sci. USA* **106**, 9045–9050. (doi:10.1073/pnas.0900452106)
50. Moerner WE, Orrit M. 1999 Illuminating single molecules in condensed matter. *Science* **283**, 1670–1676. (doi:10.1126/science.283.5408.1670)
51. Gambin Y, Deniz AA. 2010 Multicolor single-molecule FRET to explore protein folding and binding. *Mol. Biosyst.* **6**, 1540–1547. (doi:10.1039/c003024d)
52. Ferreon AC, Gambin Y, Lemke EA, Deniz AA. 2009 Interplay of alpha-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proc. Natl Acad. Sci. USA* **106**, 5645–5650. (doi:10.1073/pnas.0809232106)
53. Gambin Y, Schug A, Lemke EA, Lavinder JJ, Ferreon ACM, Magliery TJ, Onuchic JN, Deniz AA. 2009 Direct single-molecule observation of a protein living in two opposed native structures. *Proc. Natl Acad. Sci. USA* **106**, 10 153–10 158. (doi:10.1073/pnas.0904461106)
54. Gambin Y, VanDelinder V, Ferreon ACM, Lemke EA, Groisman A, Deniz AA. 2011 Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing. *Nat. Methods* **8**, 239–241. (doi:10.1038/nmeth.1568)
55. Gavin AC *et al*. 2002 Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147. (doi:10.1038/415141a)