Title

Targeting protein function: the expanding toolkit for conditional disruption

Authors

Amy E. Campbell[1] and Daimark Bennett[1]

Affiliation

[1], Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool, L69 7ZB, UK

Abstract

A major objective in biological research is to understand spatial and temporal requirements for any given gene, especially in dynamic processes acting over short time periods, such as catalytically-driven reactions, subcellular transport, cell division, cell rearrangement and cell migration. The interrogation of such processes requires the use of rapid and flexible methods of interfering with gene function. However, many of the most widely-used interventional approaches, such as RNAi or Crispr-CAS9, operate at the level of the gene or its transcripts, meaning that the effects of gene perturbation are exhibited over longer time-frames than the process under investigation. There has been much activity over the last few years to address this fundamental problem. Here we describe recent advances in disruption technologies acting at the level of the expressed protein, involving inducible methods of protein cleavage, (in)activation, protein sequestration or degradation. Drawing on examples from model organisms we illustrate the utility of fast-acting techniques and discuss how different components of the molecular toolkit can be employed to dissect previously intractable biochemical processes and cellular behaviours.

Introduction

Genetic manipulation, which operates at the level of the gene or its transcribed product, has proven to be indispensable for the identification of molecular components required for biological processes and understanding how these components act together to construct functional cells, tissues and organisms. There are now a myriad of tools for mutational analysis that have been accumulated for over a century, fuelling gene discovery through forward genetic screens and facilitating reverse genetics to probe gene function [1-3]. Recent developments, in particular CRISPR-based approaches and RNA interference (RNAi), promise to further transform our understanding by facilitating high-throughput reverse genetics and gene editing with nucleotide-level precision [4, 5]. Gene overexpression technologies have also become extremely advanced, featuring a high level of spatial and temporal control that has enabled a range of developmentally targeted gain-of-function studies. Heterologous gene expression systems, such as yeast GAL4-UAS, have proven to be particularly versatile in this regard [6]. Yet, quite often, the genetic tools used to perturb gene function are not able to keep pace with dynamic biological events that can act over different timescales from less than a second to many hours or days, depending on the process.

Processes acting over short times are particularly recalcitrant to genetic analysis because there is a considerable delay between perturbations at the transcriptional or post-transcriptional level and the corresponding effect on the encoded protein. Consequently, genetic approaches are typically incapable of selectively disrupting the encoded protein of interest at the time when the effects of experimental manipulation are measured. This is particularly a problem for analysing the function of proteins required at multiple points of the same process. A partial solution to this problem might be to monitor the process under investigation in real-time, so as to capture more information about the biological effects than can be revealed at a fixed time point [7]. Yet significant limitations remain, especially for the analysis of protein function *in vivo*. For instance, mutation of a gene required for development might result in early lethality making later processes impossible to analyse. Alternatively in the case of transgenic RNAi, which typically results in partial loss-of-function, it may not be possible to drive expression of the transgene early enough for knockdown to occur before the process has already taken place. This is particularly a problem early in development, where it can take a considerable time for the maternal contribution of RNA and protein to be exhausted. Another major consideration is the existence of compensatory homeostatic mechanisms that may circumvent the requirement for the protein under investigation. Although this is often a reason cited for lack of knockout phenotypes in mice, the underlying mechanisms are frequently not well described. Studies employing fast-acting methods have the potential to resolve such issues, as illustrated by a recent study of the cell-surface glycoprotein CD44. CD44 isoforms act as coreceptors for the receptor tyrosine kinases c-Met and VEGFR-2 but do not produce overt phenotypes when knocked-out in mice [8]. Using blocking antibodies, it was shown that acute disruption of CD44v6 inhibited cell proliferation and c-Met activation in wild type mice, but that ICAM-1 compensated for the CD44v6

isoform in CD44 null mice [8]. This study illustrates that rapid blockade of protein function can be a powerful way of resolving problems associated with slow-acting or constitutive methods of gene disruption. However, only a minority of proteins are currently open to pharmacological manipulation and the development of specific blocking reagents for every protein of interest based on their intrinsic properties is a long way away.

A successful strategy that has been adopted by the research community over the last few years to increase the range of targets that can be manipulated pharmacologically, has been to take well-characterised ligand-interaction domains from heterologous systems, and genetically engineer them into proteins of interest. In parallel to this chemical genetics approach, researchers have also found ways of incorporating domains responsive to other triggers, such as light, temperature and pH. This has spawned a new generation of tools that act directly at the level of the expressed protein and have the potential to provide insight into acute perturbations, give access to analysis over short times and allow reversible switching. Such tools can be broadly categorised according to their mode of action: those that, disrupt protein activity through complex (de)formation, (in)activate proteins by induced splicing or cleavage, and directly target a protein for degradation through the endogenous cell machinery. Here we focus primarily on tools for conditional control of protein function that fall into one of these three categories and are fast acting, providing examples of their application as a guide to researchers considering their use.

Complex Formation

The promotion of interactions between proteins of interest represents a powerful strategy for the conditional control of protein activity. There are a number of different mechanisms by which such interactions can be engineered to spatiotemporally regulate protein activity in response to different stimuli, in a way that is both precise and fast acting (Fig. 1, Table 1).

Chemically Induced Dimerisation (CID)

One of the first mechanisms involving engineered protein complex formation utilised small molecules, referred to as chemical inducers of dimerisation (CIDs), which simultaneously bind domains engineered into two proteins of interest, bringing them into close proximity and promoting their interaction (Fig. 1A, Table 1). Methods involving CIDs can influence protein activity by promoting (dis)aggregation, altering transcription or by changing the sub-cellular localisation of target proteins. The application of CIDs to control protein activity and study protein function spans the past two decades, with the majority of applications utilising naturally occurring CIDs, found to dimerise specific protein-ligand pairs [9-11]. For a small molecule to be successful in CID approaches it must have the ability to simultaneously bind two proteins, and therefore must have two, high affinity, highly specific protein binding domains, joined in a way that allows both target proteins to bind and interact [12].

FK1012

The first, naturally occurring CID, FK1012, was reported in 1993. FK1012 is a derivative of the immunosuppressant drug FK506, which was found in 1991 to be capable of binding calcineurin and the FK506-binding protein (FKBP12) with high affinity [13]. FK1012 is a synthetic dimer of FK506 that lacks the intrinsic biological activity of FK506 and has since been utilised as a CID to bind multiple FKBP12 domains [14], bringing target proteins together in a defined, reversible fashion and demonstrating the subtleties required for successful CID design [12]. The initial CID concept was demonstrated through a system in which addition of FK1012 activated the endogenous T-cell signalling cascade, via fusion of FKBP12 to the proximity-regulated ζ-chain of the T-cell receptor leading to receptor aggregation and subsequent activation [14]. FK1012 and other CIDs capable of dimerising a single protein domain, discovered in the years following, have since been applied to the study of many important cellular processes to, for example, induce apoptosis via aggregation of the Fas membrane signalling protein [15] or regulate transcription through ligand-dependent (dis)association of transcriptional activators with promoter regions [16].

Rapamycin

Whilst the first CIDs were only capable of homodimerisation, these approaches could, in theory, be used to generate heterodimers if two proteins of interest were tagged with the same domain. The result would however be a mixture of heterodimers and homodimers of the two individual proteins of interest. The development of methods involving naturally occurring heterodimerisers therefore followed, with the most notable heterodimeriser rapamycin dominating the field since its discovery [17]. Rapamycin is an immunosuppressant drug that selectively binds both FKBP12 and FKBP-rapamycin associated protein (FRAP/mTOR) [18]. The FKBP and FRB domains of these proteins, respectively, are sufficient for binding and retain the binding affinity of the full-length proteins [19]. A key step in the application of rapamycin as a CID was the production of rapamycin derivatives, known as “rapalogs”, which have a much lower affinity for endogenous proteins, thereby circumventing rapamycin’s immunosuppressive activity. In parallel, the rapamycin-binding regions from FKBP12 and mTOR, were remodelled to bind the rapalog at nanomolar affinity, providing an orthologous rapamycin system for CID applications [20-22]. These rapalogs have since been used to conditionally dimerise proteins to interrogate many different biological processes. Notable examples include the study of mitosis, in which rapamycin-induced binding of the endoplasmic reticulum and golgi membranes showed that these structures remain segregated during mitosis in mammalian cells [23], and also the study of phosphoinositides and their roles in endocytosis and intracellular-trafficking [24, 25].

One specific application of rapamycin-mediated control in mammalian cells is the knocksideways (KS) method (Table 1), which acutely sequesters protein activity through a change in sub-cellular localisation. The KS method is capable of rapidly re-routing target proteins containing an FKBP

domain to the mitochondria on a timescale of seconds, through rapamycin-induced binding to an FRB-containing protein with a mitochondrial targeting signal (Mito-YFP-FRB) (Fig. 1B) [26]. This method utilises the principle of mitochondrial re-routing, whereby the protein of interest accumulates on the outer mitochondrial membrane in a way that, providing the new localisation is not compatible with protein function, sequesters protein activity, but remains tolerable to cells [27]. In an initial proof-of-principle study the KS method was used to study the role of two subunits of the adaptor protein (AP) complexes of clathrin-coated vesicles AP-1 and AP-2 [26]. Robinson and colleagues [26] used rapamycin-induced re-routing of AP-1 and AP-2 to the mitochondria, in combination with siRNA knockdown of the endogenous protein, to demonstrate the requirement for both proteins in the endocytosis pathway. Whilst the phenotype of AP-2 sequestration was similar to that resulting from siRNA approaches alone, the corresponding phenotype observed for AP-1 was distinct from that of the siRNA knockdown and is accredited to more rapid depletion achieved in the KS approach [26]. The effectiveness of the KS approach for rapid changes in protein activity have since been demonstrated in a number of varied applications to fast-acting processes in mammalian cells. For example, Cheeseman and colleagues [28] used the rapamycin-mediated approach to specifically remove TACC3-ch-TOG-clathrin complexes from the mitotic spindle within a timescale of 5 minutes following rapamycin addition. By re-routing these complexes to the mitochondria and away from mitotic spindles at defined stages in mitosis they were able to deduce their role in maintaining tension in kinetochore fibres, which are essential for correct segregation of chromosomes. Again, this phenotype was distinct from that observed with siRNA alone, demonstrating the utility of the KS approach.

CIDs offer an efficient way to control the dynamics of processes reliant on oligomerisation. However, the same principles can also be used for the opposite mode of control, in which processes are inhibited by oligomerisation and activated upon addition of a ligand that dissociates the complexes. To enable this mode of control, Rollins and colleagues identified an FKBP12 mutant F36M-FKBP (F_M) with the ability to form discrete dimers that can be dissociated rapidly upon addition of ligand [22]. Using these tools, Al-Bassam and colleagues [29] were able to develop a novel pulse-chase system in which exogenous F_M tagged membrane proteins were accumulated gradually in the ER, and sequestered by the formation of aggregates. Within minutes of small molecule ligand addition, the F_M domins dissociated and the accumulated membrane proteins could be simultaneously released for synchronous continuation along the secretory pathway [29]. Through this method, Al-Bassam and colleagues were able to study proteins in a specific phase of the secretory pathway without interference from proteins in other phases of the pathway and thus overcome a major problem associated with studying this dynamic process.

In order for CID to be successful, target proteins must be considered on a case-by-case basis and prior knowledge of protein function is usually required in order to achieve thorough inactivation. With

nanomolar affinities between ligand-protein pairs, CID approaches have high specificity and high efficiency. However, this puts them at a disadvantage in terms of reversibility, as they often require an additional ligand that competes for binding to relieve protein inactivation/sequestration. Effects are often irreversible [30]. Whilst CID approaches were initially developed and demonstrated *in vitro*, these applications have since been developed to allow *in vivo* studies [31]. However, the requirement for exogenous small ligand addition and resulting potential for off-target effects somewhat limit the practicality of such applications. Also, although the KS approach demonstrates the ability for CID approaches to operate on a timescale of minutes, generally CID based methods range from minutes to hours, limited by the requirement for efficient uptake of the chemical inducer, and are consequently not among the fastest acting tools for temporal control of protein dynamics.

Light-induced dimerisation (LID)

Another way in which protein dimerisation can be induced is via light-based methods, which take advantage of naturally occurring photosensitive protein domains that dimerise upon exposure to a certain wavelength of light (Fig. 1C, Table 1). Whilst maintaining the flexibility of CID approaches in terms of the response elicited and the many ways in which protein function can be disrupted, light-based methods generally overcome many of the limitations of CID. In particular, they provide improved spatiotemporal precision, mitigate the requirement for exogenous small molecule addition and operate on a timescale of seconds. Genetically-encoded light-based (optogenetic) approaches have vastly expanded within the last decade from just a few applications to a whole toolbox of techniques with which to control protein activity [32]. Like CID, the first methods involving LID took advantage of naturally occurring photosensitive proteins, often discovered initially in plants, known as phytochromes and cryptochromes.

Phytochromes

Phytochromes are photoreceptive pigments encoded by small multigene families in plants and bacteria where they monitor red/far-red wavelengths of light [33, 34]. The most thoroughly investigated phytochromes are those from *Arabidopsis thaliana*, which normally function to modulate seed germination and shade avoidance [35]. One such protein is phytochrome B (PhyB) which undergoes a conformational change upon exposure to light of visible red wavelengths (~650-670 nm) to heterodimerise with the transcription factor phytochrome interaction factor 3 (PIP3). Unlike other photosensitive proteins, this dimerisation can be reverted through exposure to longer wavelengths of light (~700-750 nm), which induces monoisomerisation of PhyB and releases PIP3, allowing for very precise control of protein activity [32, 36].

Cryptochromes

Also commonly found in plants, Cryptochromes (Cry) are plant photosensors that absorb blue light, the most well studied of which, Cry2, heterodimerises with the cryptochrome-interacting basic-helix-

loop-helix 1 (CIB1) transcription factor. Cryptochrome proteins have a C-terminal domain required for signal transduction and, like phytochromes, require a flavin adenine dinucleotide chromophore co-factor, which binds to an N-terminal DNA photolyase homology region (PHR) [37, 38]. Since the discovery of Cyr2 and its ability to heterodimerise with CIB1, this system has been adapted to circumvent the need for exogenous chromophore addition [37]. With this improved system, Kennedy and colleagues [37] induced dimerisation of Cry2-CIB1 on a sub-millisecond timescale (in under 300 μs), although the reverse process took minutes to complete. Nevertheless, this improved system has since been used in the study of a number of different cellular processes in model organisms. One field in which the Cry2-CIB1 system has been used successfully both *in vitro* [39] and *in vivo* [40] is the study of phosphoinositide signalling. This was achieved by Cry2-CIB1 mediated recruitment of a phosphoinositide phosphatase catalytic subunit responsible for the conversion of $PI(4,5)P_2$ into phosphatidylinositol-4-phosphate (PI(4)P) to the plasma membrane in a light-dependent manner. Using this approach, Guglielmi et al [40] were able to study complex morphological changes and interactions that occur within defined timescales during *Drosophila* embryogenesis. The recruitment of the catalytic subunit to the plasma membrane within seconds of blue-light illumination was sufficient for quick depletion of $PI(4,5)P_2$ which, given the role of phosphoinositides in regulating actin polymerisation, allowed control over cell contractility and facilitated the study of cell-cell interactions, force transmission and changes in tissue geometry [40].

The use of cryptochromes for conditional dimerisation has since spawned a host of methods utilising the interaction between Cry2 and binding partners such as CIB1. One such method, known as light-activated reversible inhibition by assembled trap (LARIAT), utilises light-mediated heterodimerisation to reversibly sequester target proteins into multimeric complexes in mammalian cells, by engineered interactions with multimeric proteins (Fig. 1D, Table 1) [41]. Lee and colleagues developed the LARIAT technique by fusing Ca^{2+}/calmodulin-dependent protein kinase IIα (CaMKIIα) protein, which self-assembles into a 12 subunit oligomer, to CIB1. Upon blue light stimulation, CIB1 interacts with Cry2 and forms clusters through interconnections between CIB1-conjugated CaMKIIα multimeric proteins. Through this method of optogenetic trapping, Lee *et al.* [41] were able to induce cluster formation with high spatiotemporal precision in HeLa cells within 30 seconds upon illumination, with cluster disassembly occurring within minutes of light withdrawal. Lee and colleagues also found that the extent of clustering was correlated with the intensity or number of light pulses administered, suggesting it may be possible to quantitatively control clustering simply by varying light conditions for more intricate control of protein dynamics. This approach can also be used to inactivate GFP-tagged proteins, without the need to add an additional protein tag, through the use of anti-GFP nanobodies [41]. To demonstrate this approach, Lee *et al.* trapped a number of different GFP-tagged proteins into complexes through interactions with a CIB1-conjugated anti-GFP nanobody to acutely disrupt proteins involved in fast-acting processes such as membrane retention and spindle formation.

A recent example that displays the potential of the LARIAT approach is its application to the study of intracellular membrane trafficking. Here, Nguyen and colleagues [42] developed a system whereby intracellular membranes can be rapidly and reversibly sequestered into complexes via Cry2-induced aggregation of CIB1-conjugated GTPases. Using diverse Rab GTPases as membrane markers, it was possible to access specific intracellular membrane compartments such as the golgi and endoplasmic reticulum [42]. This approach makes it possible to dissect the spatiotemporal functions of intracellular membranes in a variety of processes such as receptor transport, intracellular signalling from endosomes, protein sorting and secretion.

It is also known that many plant photosensors, including Cry2, are capable of forming aggregates upon light stimulation [43, 44]. For example, Wend and colleagues [45] demonstrated the ability of Cry2 to dimerise C-Raf and activate its kinase activity in a light-inducible manner, functionally separating C-Raf from upstream growth factor signalling, enabling a more controlled approach to study dynamic downstream effects on target protein phosphorylation and cell signalling. Interestingly Wend and colleagues also tested the ability of C-Raf-Cry2 to dimerise with CIB1-bound C-Raf and found a weaker activation of C-Raf, which they suggest may be due to a difference in stoichiometry when the larger Cry2 molecule binds to the much smaller CIB1 domain [45]. Another use of Cry2 dimerisation is in a technique called Clustering Indirectly using Cryptochrome 2 (CLICR), which involves the clustering of transmembrane receptors to activate signal transduction (Fig. 2). This is achieved by indirect clustering of Cry2 bound to a receptor-binding domain (BD); high local concentrations of the BD then serve to cluster endogenous receptors leading to signal activation [46]. An N-terminal src-homology 2 (SH2) domain, which binds tyrosine-receptor kinases and the phosphotyrosine-binding-like F3 domain from Talin, which binds β3-integrin were shown to be effective BDs [46], suggesting the method could be modified to target a wide range of transmembrane proteins. However, the selectivity of such tools need to be empirically validated for each system.

Light-oxygen-voltage domains

An alternative approach to light-induced dimerisation involves the use of light-oxygen-voltage domains (LOV) from *avena sativa* phototropin 1 (Table 1). LOV domains contain a C-terminal α-helix (Jα helix) which, upon light illumination and excitation of a flavin cofactor within the LOV domain, undergoes a large conformational change and unwinds [47]. This light-induced, structural change allows for the control of protein activity through allosteric regulation of proteins containing these LOV domains (Fig. 1E). One example of how LOV domains have aided the study of protein function is the application to the study of cell motility [48, 49]. Wu et al [48] fused Rac1 to a LOV domain, which, in its native α-helix state, blocked Rac1 interactions. This photoactivatable Rac1 (PA-Rac1) could then be reversibly and repeatedly activated in precise cellular locations by illumination with blue-light, producing localised cell ruffling and protrusions. Localised Rac1 activation was also able to promote

directed cell motility [48]. PA-Rac analogues have since been used in further *in vivo* cell migration studies, showing for instance that Rac activation is sufficient for polarisation of the border cells in *Drosophila* oogenesis and that the directionality of the subsequent migration of these cells during egg chamber development is dependent on Rac levels [50]. Whist PA-Rac has been used successfully in a number of studies, the shift between the wound and unwound Jα helix states upon illumination is less than ideal, with at best a ten-fold shift towards the unfolded state upon light irradiation [51]. Through the identification of mutations that stabilise both the wound and unwound Jα helix states, Strickland *et al.*, modified the LOV system and reduced the proportion of unwound Jα helix in the dark state to make the switch between light and dark states more defined and increase the dynamic range of the system as a whole, with up to a 70-fold shift in Jα helix state after exposure to light [52].

One example that takes advantage of the high spatial and temporal precision that can be achieved using LOV domains is the control of receptor tyrosine kinase (RTK) activation. RTKs are a family of cell surface receptors that respond to growth factor and hormone signals to regulate a variety of cell behaviours, and have previously proven difficult to study due to the rapid rates of receptor biosynthesis and degradation that can occur. Grusch and colleagues [53] used LOV-domain mediated dimerisation of mutant RTKs, insensitive to endogenous ligands, to induce transphosphorylation and therefore receptor activation on a timescale shorter than that of receptor synthesis/degradation. Through this approach they were able to mimic the cell behaviours induced by endogenous growth factors to provide control over cell signalling on the minute timescale, with diverse cellular responses in different cell types pointing to the involvement of different adapter proteins or feedback mechanisms [54].

Light-based methods for induced protein complex formation and control of protein activity offer a powerful solution to many of the drawbacks that come with chemical-based approaches whilst maintaining versatility. Although light-based methods require laser excitation to stimulate photo-activatable protein modules, the wavelengths of light used generally fall within the same range as those used for conventional fluorescence imaging, meaning cytotoxic effects are minimal and these approaches have therefore been applied successfully to both *in vivo* and *in vitro* studies [55]. The benefits of optogenetic approaches over the more traditional, well-studied small molecule approaches suggest that with further development, these tools will be invaluable in the use of complex formation for protein inactivation or sequestration in the study of fast-acting cellular processes.

Protein Cleavage/Splicing

Another common strategy for the inducible control of protein activity is to induce physical changes in protein sequence through the endogenous process of protein cleavage or splicing. As with methods for inducible protein complex formation, protein cleavage/splicing can be engineered to allow

induction via a number of different mechanisms including both small molecule based and light-based approaches (Fig. 3, Table 2). However, the mechanisms used to induce protein cleavage/splicing are often interchangeable, allowing these methods to be adapted to a wider range of systems and biological questions.

Intein-mediated protein splicing

One method that allows inducible control of protein activity uses the endogenous post-translational mechanism known as intein-mediated protein splicing (Table 2). With this method, intervening polypeptides known as inteins are used to catalyse their own removal from the flanking polypeptides, or exteins, which are subsequently joined back together. Inteins are typically removed in a four-step process involving conversion of the peptide bond linking the N-terminal extein to an ester or thioester bond and transfer of the N-extein to the C-extein by transesterification. The resulting branched ester is then resolved by asparagine cyclization followed by conversion of the newly formed ester bond linking the two exteins into an amide bond and hydrolysis of the C-terminal aminosuccinimide of the excised intein [56, 57]. Inteins are used in biotechnology for a number of different applications, including the control of protein expression or modification, post-translational processing and also protein labelling [57], but perhaps the most valuable application in terms of studying protein function is to facilitate the control of protein activity. Since inteins are extensively reviewed elsewhere [57-59] we will not discuss their use further here, except to say that they have been engineered to allow conditional protein splicing (CPS), such that the splicing process is induced by the activation of an intein through reduction or the addition of a trigger such as light, temperature, pH or the addition of a small molecule (Fig. 3A) [57, 60]. These systems have been used successfully both in cultured cells and in living animals to interrogate protein function, although they have not been widely adopted for this purpose perhaps because of their intrinsic lack of reversibility.

TEV cleavage

One common method for inducible protein cleavage as a mechanism to control protein activity exploits the ability of the *tobacco etch virus* (TEV) protease to cleave a highly specific seven amino acid recognition sequence (E-X-X-Y-X-Q-G/S) with high efficiency (Fig. 3B, Table 2) [61-63]. TEV is commonly used as a mechanism for the cleavage of fusion proteins to remove protein affinity tags prior to further protein analysis [64]; however, this system has also been applied to the control of protein activity both *in vivo* and *in vitro*. Through genetic modification, the TEV recognition sequence can be engineered into a protein of interest to allow inducible protein cleavage and inactivation when in the presence of TEV. TEV techniques have previously been demonstrated in budding yeast to provide evidence of a role for separin in anaphase initiation [65] and has since been applied to the study of proteins in both *Drosophila* cell culture and live embryos [66, 67]. For example, to show that TEV was able to effectively and specifically cleave a protein containing the recognition sequence in live *Drosophila* embryos, Harder *et al* [66] expressed the protein Megatrachea (Mega), a *Drosophila*

claudin protein localised to membrane compartments of ectodermal cells, containing an artificial TEV protease cleavage site (TEV_{pcs}) and yellow fluorescent protein (YFP) (Mega-TEV_{pcs} -YFP). Upon TEV expression using the Gal4/UAS system, Mega-TEV_{pcs}-YFP no longer showed the correct YFP localisation indicating the YFP had been cleaved from the Mega fusion protein. Harder *et al* went on to adapt this system to allow induction of TEV expression at different stages of embryo development by putting TEV under the control of the heat-shock promoter 70 (hsp70), thus generating a mechanism for the temporal control of TEV-mediated protein cleavage. Using this temporally controlled system, they were then able to induce cleavage of a Mega-TEV_{pcs} construct leading to truncation of the Mega protein and subsequent degradation of the truncated protein [66]. Clearly, a key factor in determining the timescale of TEV-mediated cleavage is the promoter from which TEV is expressed. Changes in temperature using the heat shock promoter, is capable of inducing TEV expression and protein cleavage in *Drosophila* pupae ~3 hours after a 45 minute heat shock [67], whereas rapamycin-induced expression in mammalian cells induces cleavage within 150 minutes [68]. This may make the approach unsuitable for the study of some processes that operate on a short timescale.

One disadvantage of TEV is that it readily cleaves itself at a specific site to yield a truncated enzyme with greatly reduced activity [69, 70]. There have therefore been a number of iterative changes made to TEV protease to adapt the protease for more diverse applications, for example a TEV mutant has recently been designed specifically to be active in the secretory pathway [71]; various other TEV mutants offer the same recognition site cleavage, but an increased stability and reduced auto-cleavage activity [69]. Whilst TEV offers highly specific and efficient cleavage, *a priori* knowledge about protein composition is required to choose a position where the TEV recognition site can tolerably be inserted that will inactivate the target protein whilst reducing the possibility that the resulting protein fragments will retain function or even have novel functions of their own. There is also an optimum level of TEV protease expression at which cleavage occurs but background activity is minimalised; this level is likely to depend on both the variation of TEV protease and the system in which it is applied so would need to be considered during experimental design [68].

The timescales and spatiotemporal resolution of both intein-mediated protein splicing and TEV cleavage are dependent on the engineered mechanism of induction. Whilst the ability to customise these techniques allows them to be applied to a wide range of systems and cellular processes, the fastest acting methods of protein cleavage/splicing with the highest spatiotemporal resolution are again those induced by light.

CALI/FALI

One method for light-inducible protein cleavage applied to the study of protein function is chromophore-assisted light inactivation (CALI), Table 2. Chromophores are photosensitive groups,

often responsible for the colour of organic molecules, which produce highly reactive free radicals such as reactive oxygen species (ROS) upon illumination with a specific wavelength of light. Using the CALI approach, a chromophore-tagged protein of interest is inactivated through mild illumination for a period of time sufficient to induce generation of reactive oxygen species and induce protein cleavage of proximate proteins, but short enough to ensure the reactive oxygen species act within a defined radius (30-40 Å) to minimise off target effects (Fig. 3C) [72]. The specificity of CALI approaches is determined by the short half-life of the free radical species, which ensures that only proteins within a radius of 1.5-6 nm relative to the chromophore are affected [72-74]. After free radical generation, proteins are typically inactivated within 1 second, which, combined with laser irradiation of micrometer accuracies, allows for high spatial resolution and highly controlled protein inactivation [74].

Originally, CALI approaches used an antibody-based mechanism to attach a chromophore, such as the dye malachite green, to a protein of interest (Fig. 3C). Fluorophores such as fluorescin isothiocyanate, which are more efficient at ROS production, were later employed in a similar approach called fluorophore-assisted light inactivation (FALI) [75]. However, the need for microinjection of a dye-labelled, non-function blocking antibody specific to the protein of interest limited the widespread application of these approaches. Subsequent methods made use of genetic modification to label proteins with a generic tag that can then be fluorescently-labelled through extracellular addition of a specific reagent (Fig. 3C). Proteins tagged with 1-2 small tetracysteine (TC) motifs will specifically bind to the membrane permeable biarsenical dye resorufin-based arsenical hairpin binder (ReAsH) or the fluorescein-based arsenical hairpin binder (FlAsH). For example, Marek and colleagues [76] used FlAsH labelling to visualise Synaptotagmin I (Syt I) at the neuromuscular junction (NMJ) in late stage, live *Drosophila* larvae. Through photo-inactivation they were able to inactivate Syt I within seconds and provide supporting evidence for a model, previously based on genetic data alone, in which Syt I plays a role post-vesicle docking to mediate vesicle-fusion and calcium-dependent transmitter release [76].

Whilst dyes are added extracellularly in ReAsH/FlAsH based CALI/FALI approaches, the application of these techniques *in vivo* is limited by the difficult task of achieving sufficient uptake in live animals and also the inability to spatially control the production of ROS and limit it to particular cells or sub-cellular compartments [77]. There is also the problem of non-specific binding of the membrane-permeable dyes. For CALI/FALI to become a more widely used technique there was therefore a need for a system that sidestepped the requirement for exogenous addition and could be encoded completely through genetic manipulation. There has been some limited success using eGFP, a tag commonly used to study protein localisation and function. For example, CALI illumination of GFP-myosin II was shown to result in unequal-size daughter cells during asymmetric cell divisions in a *C. elegans* Q neuroblast cell lineage [78]. It is however believed that the chromophore within the GFP

structure is protected by the outer shell meaning the generation of free-radical species upon illumination is restricted and therefore GFP holds a low phototoxicity [77, 79].

KillerRed

The first example of a successful genetically-encoded CALI reagent, with a 1,000-fold increase in phototoxicity compared to GFP, is the GFP-related red fluorescent protein KillerRed, developed from the nonfluorescent red chromoprotein anm2CP of *Hydrozoa* jellyfish (Fig. 3C) [79]. Whilst it was originally unclear why there was such an improvement in phototoxicity, a study into the structure-function relationship of killerRed by Pletnev and colleagues provided crystallographic data revealing unique structural features that may facilitate ROS generation [80]. KillerRed has been used in a number of applications in biological research, such as the control of protein activity in the study of neuronal development in model organisms [81] and has been used to induce cell specific killing of eukaryotic cells in culture via mitochondrial or membrane-targeted KillerRed [79]. In this regard KillerRed has also proven to have uses in medicine; one recent and exciting application of KillerRed was to the experimental cancer treatment known as photodynamic therapy (PDT), which aims to use photosensitisers to selectively kill tumour cells through reactive oxygen species generation upon laser illumination [53, 82]. Following subcutaneaous injection of KillerRed expressing *Escherichia coli* (KR-*E.coli*) into mouse xenograft models of human carcinoma cell lines the Terekawa lab [82] monitored the intensity and spread of fluorescence through the tumour cells. After 24 hours the KR-*E.coli* spread throughout the whole tumour and were subsequently irradiated with orange light (540-580 nm) to induce ROS production. The generation of ROS led to necrosis and tumours gradually disappeared to leave healed skin after just 1 week demonstrating the ability of KR-*E.coli* to kill cancer cells originating from humans. Whilst there are many questions still to be answered before this technique is applicable to humans, these results provide an insight into the capabilities of genetically-encoded CALI approaches.

KillerRed has proven to be an exciting solution to the difficulties associated with previous CALI approaches. However, one limitation is its tendency to homodimerise, which can potentially interfere with protein function [77, 79]. There have been a number of other novel photosensitisers discovered since, such as SuperNova, a monomeric form of KillerRed [83], and the fluorescent flavoprotein mini Singlet Oxygen Generator (miniSOG), which can also be used to generate an insoluble deposit of singlet oxygen species that can be stained for visualisation using electron microscopy [84]. More recently, the toolkit of phototoxic proteins was expanded further with the addition of KillerOrange, an orange mutant version of KillerRed that results in ROS formation upon illumination with either blue or green light, meaning it can be used in combination with KillerRed or other photosensitisers activated by different wavelengths of light [85].

CALI/FALI approaches enable specific protein inactivation through phototoxicity and can act with high spatial resolution through the ability to express tagged proteins in specific cells and trigger inactivation at a sub-cellular level (Table 2). The expansion of the phototoxic protein toolbox should also allow for the creation of more intricately controlled systems in which different proteins can be inactivated at different timepoints or in different cell populations, which will form a useful tool both for fundamental research and for potential medical applications. However, CALI/FALI-based methods suffer from a number of limitations, including the requirement for either exogenous ligand addition or the inclusion of a fairly large protein tag that may interfere with protein function. Importantly, there is also potential for off-target effects on proteins in close vicinity to the ROS generator, challenging the specificity of these approaches. For example Guo and colleagues [86] found that the inhibition of calcium ion currents, mediated by the class C G-protein coupled receptor (GPCR) mGluR8a, was greatly attenuated following FALI inactivation. Whilst initial results were consistent with acute inactivation of mGluR8a, Guo and colleagues also reported collateral damage to proximal proteins with no overt link to pathways of GPCR signalling [86]. These factors have limited the adoption of such methods to study protein function and cannot be overlooked when using techniques involving phototoxicity for protein inactivation.

Protein Stability/Degradation

Whilst protein cleavage or splicing can often lead indirectly to protein instability and degradation, it is likely that this will occur after some delay depending on the half-life of the protein. It is also possible for the resulting protein fragments to retain function or bind other proteins and perform independent functions of their own, potentially generating a more severe phenotype than simple protein knockout alone. It therefore follows that a more thorough and interpretable approach for complete removal of proteins from cells is to target them directly for degradation by the endogenous cellular degradation pathways (Fig. 4, Table 3). Although the direct degradation of target proteins means these methods are technically irreversible and somewhat limited in their application, in most cases protein levels return to normal following relief of the degradation stimulus and thus these methods can still provide a useful tool for studying protein function [9, 10, 87]. Methods for inducible protein degradation generally involve the proteasome pathway for protein degradation, in which ubiquitin is transferred from the E1 ubiquitin-activating enzyme to the E2 ubiquitin-conjugating enzyme and subsequently to a lysine residue within the target protein in a transfer facilitated by the E3 ubiquitin ligase [88]. This process is repeated to polyubiquitinate the target protein until it has a sufficient number of ubiquitin molecules for targeted degradation by the proteasome [88].

N-end Degron

A common strategy for the direct induction of protein degradation utilises the UBR1 E3 ligase pathway and the N-end rule (Table 3), which states that the half-life of a protein is determined by both the accessibility of lysine residues, for ubiquitination, and the identity of the amino acid at the N-

terminus [89, 90]. Varshavsky and colleagues [90] demonstrated this principle through cleavage of ubiquitin, via a yeast deubiquitinating protease, from a fusion protein expressed in yeast, containing the 5' end of a lacI linker followed by β-galactosidase (β-gal), resulting in the exposure of a new N-terminal amino acid. The half-life of β-gal following cleavage could then be vastly altered by simply changing the exposed amino acid residue. For example, β-gal with an N-terminal Arginine or Phenylalanine residue had a half-life of ~3 minutes, whereas Methionine or Valine resulted in a half-life greater than 20 hours [89]. This method of degradation is conserved from bacteria to higher eukaryotes, and means proteins tagged with the unstable lacI degron can be targeted for degradation within minutes [90]. However, this strategy is not inherently inducible and therefore requires modification for the conditional control of protein degradation.

One of the first examples of an inducible N-end degron involved the use of a temperature sensitive dihydrofolate reductase variant (tsDHFR) where an N-terminal destabilising arginine was only exposed at non-permissive temperatures [91]. By fusing the tsDHFR to the N-terminus of a target protein, degradation can be induced by a switch to the non-permissive temperature of 35°C, exposing the N-terminus at which point the N-end rule takes effect (Fig. 4A). For example, this system has been used successfully in *Drosophila*, to inducibly polyubiquitinate an eGFP reporter protein at the neuromuscular synapse following a 30-minute heat shock at 35°C, in order to track the degradation of polyubiquitinated proteins [92]. By tracking the degradation of the eGFP, Speese and colleagues [92] showed that ubiquitinated presynaptic proteins are not removed from the synaptic terminal but rather undergo local proteasome-mediated degradation at pre-synaptic sites. In addition to examples from *Drosophila*, this technique has also been used successfully to characterise many essential proteins in budding yeast [93]. However, such techniques are generally limited to systems that can survive the required temperature changes and also to proteins that retain function with the required N-terminal modification. Despite this, the approach has since been used successfully in chicken DT40 cells, in which the method was first tested using a tsDHFR degron fused to eGFP. Upon transfer of the cells to the non-permissive temperature of 42°C, the protein was rapidly depleted to ~10% of initial levels within 90 minutes leading to undetectable levels after 120 minutes [94]. Moving cells back to the permissive temperature of 35°C resulted in an efficient recovery to the pre-depletion level within 150 minutes. Su and colleagues [94] then used the approach to deplete RAD51, finding that RAD51, which plays an important role in homologous DNA recombination (HDR), does not stop DNA synthesis but causes cell cycle arrest in G2, suggesting HDR becomes important at G2. This, along with the many other applications of tsDHFR-based approaches, show that whilst this method is limited in its potential applications, tsDHFR can still provide a useful tool in the study of protein function.

In addition to a temperature-sensitive DHFR variant, a small molecule-controlled version has also been engineered, for which the drug methotrexate (MTX) regulates stability of the N-terminus. Whilst

the presence of MTX fails to inhibit recognition and therefore polyubiquitination of tsDHFR by E3 ligase, the stable, high-affinity interaction between MTX and DHFR impedes protein unfolding and prevents degradation by the proteasome [95, 96]. This system was demonstrated in both yeast and mammalian cells in culture, although the occupancy of the proteome by the MTX-tsDHFR complex is likely to inhibit degradation of other cellular proteins leading to off-target effects [97]. Also, DHFR is required for the production of tetrahydrofolate, which is subsequently required for the synthesis of purines, thymidylate and several amino acids [98]. The inhibition of DHFR by MTX therefore interferes with the synthesis of DNA, RNA and even proteins, meaning MTX is undesirable as a regulatory small molecule for exogenous addition.

TIPI (TEV protease mediated induction of protein instability)

An alternative way in which the N-end degron system can be made inducible, and more widely applicable, is through the use of TEV in a technique called TEV protease-mediated induction of protein instability (TIPI), Table 3. This technique combines TEV with the N-end rule, whereby TEV cleaves a recognition sequence engineered into a cryptic N-degron, attached to the N-terminus of a target protein, to reveal an unstable N-end amino acid (Fig. 4B) [99]. According to the N-end rule, this unstable amino acid targets the protein for polyubiquitination and degradation via the UBR1 E3 ligase pathway [89, 90]. TEV protease has previously been shown to allow degeneracy within its recognition sequence and is particularly flexible to changes at position 7, the amino acid residue that forms the N-end following TEV cleavage [63]. TEV protease can therefore cope with the incorporation of an amino acid that induces degradation following cleavage via the N-end rule. Taxis et al [99] first developed the TIPI approach by designing a construct containing a reporter followed by a TEV protease recognition site, N-degron and SF3b155$^{381\text{-}424}$ termed Reporter-TDegX-tag, where X represents the amino acid at position 7 which becomes the new N-terminal amino acid upon cleavage. The inclusion of a relatively short SF3b155$^{381\text{-}424}$ domain allowed for more efficient cleavage as its binding to a mutant version of the human spliceosome subunit 14 (p14*) recruited p14*-TEV to the recognition site (Fig. 4B). Taxis and colleagues initially demonstrated this system in yeast using a GFP-TDegX-Don1p fusion protein, with p14*-TEV expression driven by the Gal1 promoter, monitoring cleavage via the release of GFP and testing the effect of the amino acid at position X on TEV-cleavage efficiency and protein half-life. Phenylalanine or asparagine were found to provide optimal conditions for both TEV cleavage and rapid degradation of the target protein, Don1p. The effectiveness of TIPI was shown by using the approach to deplete several different proteins in yeast, obtaining phenotypes correlating to those observed via genetic knockdowns [99].

The potential of TIPI as an approach for conditional degradation is yet to be realised, however its power and versatility has recently been demonstrated in some alternative applications. For instance, TIPI has been further modified (mTIPI) to facilitate production of recombinant proteins; it does this by blocking endocytosis in yeast and combating a common problem whereby highly active endocytosis

in protein expression systems reduces the overall protein yield [100]. For conditional control this method simply requires expression of the p14*-TEV fusion protein, which could be induced via the same methods as for TEV cleavage, including temperature, pH, small molecule addition or the Gal4/UAS system, making it a versatile tool for inducible degradation and the study of protein function both *in vivo* and *in vitro*.

Auxin-induced degron (AID)

Another way in which proteins can be directly targeted for degradation by the ubiquitin-proteasome machinery is via F-box proteins, which bind target proteins and recruit the Cullin-RING complex, also called the SCF complex consisting of Skp1, Cullin and F-box, to ubiquitinate the target protein [101]. Eukaryotes contain multiple forms of SCF, whereby the F-box protein conveys specificity towards different target proteins. One F-box particularly suited to small molecule-induced control of protein degradation is the transport inhibitor response 1 (TIR1) protein, which binds target proteins in the presence of the plant hormone auxin and has a highly conserved interaction with the E3 ligase protein Skp1 (Fig. 4C) [102, 103]. This is commonly referred to as an auxin-inducible degron (AID) system (Table 3), and was initially shown to be applicable to most eukaryotes (excluding plants), including budding yeast and cell lines derived from human, mouse, hamster, monkey and chicken [104]. The AID degron consists of IAA17, also known as AXR3, from *Arabidopsis thaliana* and when expressed at either the C- or N-terminus of GFP in budding yeast also expressing AtTIR1, under the control of the galactose-inducible *GAL* promoter, the SCF-TIR1 complex was able to assemble and degrade GFP to less than 3% of initial levels within 30 minutes of auxin addition [104]. This approach was successfully used to degrade several essential nuclear or cytoplasmic proteins in yeast. However, to apply the system to mammalian cells it was first necessary to modify TIR1 to convey a higher thermostability and allow use at 37°C. This was achieved by sourcing TIR1 from the rice plant *Oryza sativa* (osTIR1), which also provided an improvement over atTIR1 for use in yeast [104].

This method has since been used successfully to degrade both nuclear and cytoplasmic proteins in *C.elegans* [105] and mammalian cells [106] to help identify the function of several different target proteins at specific time points of the cell cycle. Holland and colleagues [106] tested five differentially localised proteins, some of which were known to be incorporated into protein complexes. Four out of the five AID-YFP tagged proteins expressed under doxycline control (Plk4, CENP-A, TFR2 and cyclin B1), showed quantitative protein degradation within 80 minutes of auxin addition. Degradation of the other protein studied, H2B, occurred more slowly within 3 hours [106]. These results show that AID is capable of rapidly depleting proteins involved in stable complexes with relatively long half-lives, however, the time taken for depletion following induction can vary. Holland and colleagues demonstrated the ability of AID to induce proteolysis with the same or very similar degradation kinetics at all phases of the cell cycle and also found that proteins re-appeared almost immediately upon removal of auxin stimulus. To demonstrate the ability of AID to study protein function, Holland

and colleagues [106] achieved rapid functional inactivation of BubR1, an essential component of the mitotic checkpoint, by depleting endogenous BubR1 protein with siRNA, replacing it with siRNA-resistant GFP-AID-BubR1 and inducing mitotic arrest through nocodazole addition. GFP-AID-BubR1 rescued the function of the depleted endogenous BubR1 and this rescue could be rapidly reverted through auxin-induced degradation of the GFP-AID-BubR1 fusion protein to produce a more complete null phenotype than mRNA depletion alone [106].

The AID approach has since been developed further to increase versatility through minimisation of the degron size and the inclusion of a series of epitope tags to allow detection using fluorescence microscopy or commercially-available antibodies [107]. Morawska and colleagues [107] developed a series of vectors for PCR-based genomic tagging strategies containing different iterations of the AID degron with epitope tag, allowing for both C- or N-terminal tagging and providing a range of selection markers which they then demonstrated through application to a series of different yeast proteins. Whilst these vectors increase the versatility and facilitate use of the AID approach, individual proteins must still be considered on a case by case basis to design the most effective degron; it may even be necessary to test multiple iterations to ensure proteins retain function.

PROTACS

Another chemical-based method for the specific degradation of target proteins by the endogenous ubiquitin-proteasome machinery is through the use of heterobifunctional small molecules known as Proteolysis Targeting Chimeras (PROTACs), Table 3. PROTACs consist of one moiety that binds the target protein linked to an E3 ligase to directly recruit the protein for proteasome-mediated degradation. Initially developed to target disease-causing proteins for destruction, the first generation of PROTACs were based on large peptide motifs derived from known ubiquitin ligase substrates [108]. However, these were limited by high molecular weight, poor cellular uptake and potential metabolic instability [109, 110]. Following a switch to small molecule-based PROTACs, the past decade has seen a series of improvements to PROTAC technology, aided by the development of small ligands for a number of E3 ligases, including MDM2, cIAP1, CRBN and VHL (as reviewed by [110]). Whist these improved small molecule PROTACs were able to successfully degrade target proteins, the overall uptake of this technique for the conditional control of protein degradation has been limited by a number of uncertainties, including PROTAC stability and E3 ligase binding affinity [109, 110]. However, more recent advances in the field have provided a new generation of highly specific, high affinity, low molecular weight PROTACs with the potential to expand the use of PROTAC technology [111-114].

For example, Bondeson and colleagues [111] used structure-guided approaches to develop low molecular weight (~450 Da), high affinity ligands for the Cullin-RING ligase 2 VHL E3 complex ($CRL2^{VHL}$). Linked to small molecules that bind specific cellular targets, Bondeson and colleagues

were able to efficiently degrade specific proteins in cultured cell lines, including the serine-threonine kinase RIPK2, which is involved in innate immune signalling, and the estrogen-related receptor alpha (ERRα), which is implicated in the regulation of various cellular metabolism pathways, with dose dependent degradation and maximal degradation levels of >95% and 86%, respectively. Bondeson and colleagues also demonstrated this approach *in vivo* using a PROTAC targeting ERRα, reducing its levels by ~50% and significantly reducing mouse heart and kidney tumours by >40% [111]. Using a similar approach, Zengerle and colleagues [114] successfully designed potent PROTACs using optimised, drug-like VHL ligands [115] and BET bromodomain protein ligands to selectively degrade certain members of the Bromo- and Extra-terminal (BET) protein family, including the epigenetic regulator BRD4 previously identified as a potential therapeutic target for acute myeloid leukaemia and ovarian cancer [114].

In addition to targeting proteins to the CRL2^{VHL} E3 complex, potent PROTACs have also been developed to utilise the interaction between immunomodulatory drugs (IMiDs), such as thalidomide, and the CRL4^{CRBN} E3 ligase. CRL4^{CRBN}, together with an IMiD, forms a tertiary complex with the transcription factor Ikaros, resulting in its ubiquitination and degradation. This approach has since been used for the efficient and specific degradation of BRD2, BRD3 and BRD4 by attaching BET bromodomain protein ligands to an IMiD [112, 113].

Whilst the principle behind PROTAC technology is not novel, there has been a recent surge of developments to generate a newer generation of more potent PROTACs, which address the limitations of previous iterations. These newer PROTACs offer greatly increased potency whilst retaining high specificity to their target proteins both *in vitro* and *in vivo*. It is also possible to modify this specificity through manipulation of the linker between the two PROTAC moieties [111, 114]. The diversity of recently described examples shows how the PROTAC approach can work on different protein targets in a number of different systems. However, compared with small molecules used in other approaches to conditionally control protein dynamics, PROTACS are larger and more complex molecules and so may suffer limitations with respect to their pharmacokinetic properties.

deGradFP

The F-box/SCF complex-based approach has also been utilised in the deGradFP method to specifically degrade GFP-tagged fusion proteins via an anti-GFP nanobody/F-box chimera (Fig. 4D, Table 3). A method to allow specific degradation of proteins tagged with GFP is desirable as GFP tagged constructs already exist for many proteins and degradation can be easily monitored by the loss of fluorescence. The degrade GFP (deGradFP) method involves the engineered F-box fusion protein NSlmb-vhhGFP4, consisting of an F-box domain derived from the *Drosophila* protein Slmb and the single-domain anti-GFP antibody fragment vhhGFP4 which recognises GFP and its close derivatives (Fig. 4D) [116]. This method was initially demonstrated in *Drosophila*, where, with NSlmb-

vhhGFP4 expression restricted to the posterior of early stage embryos by the Gal4/UAS system, an EYFP tagged histone H2A variant (His2Av-EYFP) was rapidly depleted by the deGradFP system [116]. Caussinus and colleagues used the *engrailed-Gal4* driver to express both NSlmb-vhhGFP4 and nuclear mCherry in embryos ubiquitously expressing His2Av-EYFP. Using mCherry levels as a reporter for the expression of NSlmb-vhhGFP4, His2Av-EYFP started to be degraded after ~30 minutes following NSlmb-vhhGFP4 expression, with less than 10% of the maximum EGFP intensity remaining after ~ 3 hours [116]. Caussinus and colleagues then went on to show the versatility of the deGradFP approach through the successful depletion of the cytoplasmic protein Spaghetti squash (Sqh), nuclear protein Apterous (Ap) and the transmembrane protein Crumbs (Crb) all of which were tagged with GFP, expressed in a null background and degraded upon induction of NSlmb-vhhGFP4 expression via the Gal4/UAS system. There were, however, a couple of cases in which the deGradFP was not effective against GFP-tagged target proteins. For example, E-Cadherin/Shotgun (Shg) could not be degraded using this method, possibly as it exists in a large protein complex which may mean the GFP tag is not accessible to the vhhGFP4 antibody [116]. Also, NSlmb-vhhGFP4 was unable to induce degradation of GFP alone, perhaps as the small size of GFP prevents exposure to the SCF-recruited E2 enzyme and thus prevents poly-ubiquitination. It was however possible to degrade GFP containing a small nuclear localisation signal via this method, so whilst it is possible that a minimum size limit exists, below which degradation does not occur, this limit must be very close to the size of GFP alone and should not greatly limit the versatility of the approach [116].

deGradFP has been proven to be a useful approach for the induced degradation of target proteins particularly in combination with RNAi knockdowns in order to generate a more effective depletion of protein levels [117, 118]. More recently, a similar approach involving the modification of the E3 ubiquitin ligase adapter protein SPOP to alter target protein specificity was proposed [119]. By fusing an anti-GFP nanobody directly to a truncated SPOP adapter protein completely lacking its substrate-binding domain, Schin and colleagues [119] claim to have developed an approach that is more efficient than deGradFP, which simply involves an NSlmb deletion mutant for which the binding domain has been modified. This remains to be proven in terms of biological applications, but may offer an alternative in cases where deGradFP is not effective. deGradFP, and related methods, can in theory be adapted to allow knockdown of many endogenous proteins, providing high-affinity antibodies are available for target proteins [87].

Ubiquitin-independent

Ubiquitin-mediated protein degradation is by far the most common strategy for control of protein degradation, however it has previously been shown that localisation to the proteasome is sufficient for degradation [120] and so it is worth mentioning here that there are also a handful of methods for the conditional control of ubiquitin-independent protein degradation. The most common of these is the C-degron; consisting of a 36 amino acid sequence from Ornithine decarboxylase (ODC) this

forms a bridged association to the proteasome, acting as both the recognition and degradation initiation signal [87]. An exciting use of this technique allows for light-induced protein degradation via the use of LOV2 domains [121]. Renicke and colleagues designed a system in which the C-degron, fused to the C-terminus of the LOV domain, can be masked by the Jα-helix under dark conditions, but exposed upon illumination with blue-light via Jα-helix unfolding, leading to ubiquitin-independent protein degradation. Whilst this provides an exciting alternative, ubiquitin-dependent methods remain the most widely used and well-studied methods for inducible protein degradation.

Concluding comments and future perspectives

Manipulation of genes, at the level of DNA or RNA, has proven to be a specific and immensely powerful way of understanding the roles of encoded proteins in their native cellular environment. However, distinguishing between the initial and steady-state consequences of gene disruption, especially *in vivo*, is often problematic. The new generation of tools and methods that are emerging to meet this challenge address the issue by offering both rapid and specific control of protein function. Different sets of tools can be employed across a range of time scales to challenge biological processes operating at the subcellular, cellular and multicellular level. Methods for conditional control of protein complex formation (Table 1), in particular, open the door to *in vivo* analysis of biochemical processes operating over short times (sec/min), such as intracellular signaling cascades, which are initiated within seconds of receiving the stimulus. Furthermore, the ability to reversibly switch activities on and off enables the systematic perturbation of biochemical pathways, thereby revealing how information is processed from upstream stimulus to downstream effectors at each step and providing insights into rate limiting components and feedback control [122]. Repetitive perturbation at different times has particular value in the dissection of biological systems where frequency variation, including oscillatory behaviours, encodes information [123, 124]. Other techniques, including those that conditionally control protein splicing/cleavage or degradation, typically operate over min-h timescales (see Tables 2-3) but have proven utility in studying the molecular mechanisms of downstream events, such as changes in cell proliferation, differentiation, migration or adhesion, which operate over longer times.

It is important to note, as discussed in the sections above, that current methodologies for conditional perturbation of protein function are not without their technical limitations and researchers must weigh up whether the available tools offer the appropriate flexibility and precision for the desired experiment. One of the main considerations for researchers wishing to utilise the methods we have described is to decide which inducer they should use. Chemically-induced methods have been the mainstay of the field for many years, but the use of chemical inducers is often restricted because of their promiscuous binding profiles, which can lead to off-target effects and cytotoxicity. As we have discussed, higher affinity ligands might mitigate these effects, because the compounds can be used at lower concentrations, typically suffer from not being reversible. More fundamentally however, the

inability to target some chemicals to specific subcellular localisations, combined with their relatively slow uptake in cells, with effects occurring in minutes to hours, make these methods unsuitable for the study of certain biological processes and *in vivo* models.

Recently, there has been a surge of activity to develop methods of induction using light. One of the key attractions of this approach is that perturbations can be both rapid and reversible, in a spatially-resolved manner. Consequently, with light-dependent systems, even should an activated protein diffuse out of the area in which it received the activating input, it will then switch off, preventing phenotypic outputs from losing their spatial resolution [55]. Correspondingly, there has been a great deal of focus on improving the properties of the light-sensitive domains used in these approaches, much in the same way that there have been iterative improvements made to fluorescent proteins for use in cell imaging. This will make it possible to tailor the perturbation dynamics. For instance, in the case of light-induced complex formation, a derivative of the Cry2 domain (Cry2-olig) with altered off kinetics may make it more suitable for sequestration and inhibition of protein activity [125]. A challenge to future efforts for improvement to such domains will be not just to identify variants that confer beneficial properties in isolation, but ones that retain additive effects of multiple genetic changes; this will require screening procedures that simultaneously optimise constructs against multiple parameters.

The ultimate goal of the approaches we have described is to understand the role of molecules in biological phenomena with quantitative precision. Quantitative *in vivo* biochemistry however, requires measurement not just of the effects of perturbation on a given process, but also the magnitude of the perturbation to the target protein in time and space. Whilst reliable molecular readouts might be available for the former, the latter may be somewhat harder to measure. Fluorescently-labelled target proteins offer an attractive solution for methods relying on protein degradation, since fluorescence intensity can provide a measure of protein concentration with spatiotemporal resolution [126]. However, whilst quantitation of other perturbations, e.g. protein cleavage or activation, may be straightforward in cell populations at fixed time points, measurements with single cell resolution or in real time will be much more challenging to achieve. A number of advanced cell imaging approaches may make this possible, but these techniques are themselves technically demanding and may not be widely available to researchers who do not have access to specialised equipment. For instance, measurements of protein dynamics (e.g. with Fluorescence cross-correlation spectroscopy, or Raster image correlation spectroscopy) and protein proximity (e.g. with Förster resonance energy transfer - FRET, and Fluorescence-lifetime imaging) can be employed to determine effects on protein complex formation [127], whilst specialist (e.g. FRET-based) reporters can measure e.g. the activity of enzymes responsible for post-translational modifications [128]. Future developments will therefore have to consider appropriate strategies, not just to perturb protein function, but to simultaneously measure the extent of that perturbation along with the biological effect(s).

Concurrently, an ongoing challenge for technology developers is to make the techniques universally appealing and easy to employ. Ultimately, whether a technique is proven to be robust and fit for purpose will depend on uptake and testing by user communities. From a practical perspective, it may still take a significant investment of time to tailor an approach to the experimental system under investigation, despite efforts to make the available tools universally applicable. Prior knowledge of protein structure/function relationships may be required, for example, to guide the production of fusion proteins that retain normal activity and respond effectively to regulatory stimuli. These issues have inevitably limited the uptake of many of these techniques. Methods that utilise tools that are already widely-employed, such as GFP-tagged proteins, are likely to be amongst the most popular in the short-term because the methodologies can be rapidly deployed. For instance, deGradFP will no doubt be of particular interest to the *Drosophila* and zebrafish communities, which are creating transgenic libraries in which endogenous genes have been tagged with GFP [129, 130]. Indeed, such collections may be the starting point for well-designed, temporally-controlled screens, to identify novel genes involved in developmentally-regulated biological processes.

As the field continues to mature, an area of future development is likely to be how multiple techniques might be combined for improved spatiotemporal control of a single protein, or for the induction of more than one protein, which has application in the engineering and study of artificial networks. Importantly, conditional control of protein function is not exclusive of genetic manipulation. Ultimately therefore, as gene-editing “knock-in” strategies mature [131], it may become routine to incorporate any number of different conditional domains on a genome-wide scale to facilitate such studies.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/J014516/1) and a RCUK block grant to the University of Liverpool for open access charges.

References

1 Venken, K. J. and Bellen, H. J. (2014) Chemical mutagens, transposons, and transgenes to interrogate gene function in Drosophila melanogaster. Methods. **68**, 15-28

2 Lin, C. Y., Chiang, C. Y. and Tsai, H. J. (2016) Zebrafish and Medaka: new model organisms for modern biomedical research. J Biomed. Sci. **23**, 19

3 Boulin, T. and Hobert, O. (2012) From genes to function: the C. elegans genetic toolbox. Wiley Interdiscip. Rev. Dev. Biol. **1**, 114-137

4 Boettcher, M. and McManus, M. T. (2015) Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. Mol. Cell. **58**, 575-585

5 Sternberg, S. H. and Doudna, J. A. (2015) Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol. Cell. **58**, 568-574

6 Duffy, J. B. (2002) GAL4 system in Drosophila: a fly geneticist's Swiss army knife. Genesis. **34**, 1-15

7 Neumann, B., Walter, T., Heriche, J. K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sieckmann, F., Pau, G., Kabbe, R., Wunsche, A., Satagopam, V., Schmitz, M. H., Chapuis, C., Gerlich, D. W., Schneider, R., Eils, R., Huber, W., Peters, J. M., Hyman, A. A., Durbin, R., Pepperkok, R. and Ellenberg, J. (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature. **464**, 721-727

8 Olaku, V., Matzke, A., Mitchell, C., Hasenauer, S., Sakkaravarthi, A., Pace, G., Ponta, H. and Orian-Rousseau, V. (2011) c-Met recruits ICAM-1 as a coreceptor to compensate for the loss of CD44 in Cd44 null mice. Mol. Biol. Cell. **22**, 2777-2786

9 Fegan, A., White, B., Carlson, J. C. T. and Wagner, C. R. (2010) Chemically controlled protein assembly: Techniques and applications. Chem. Rev. **110**, 3315-3336

10 Rakhit, R., Navarro, R. and Wandless, T. J. (2014) Chemical Biology Strategies for post-translational control of protein function. Chem. Biol. **21**, 1238-1252

11 Wijdeven, R. H., Neefjes, J. and Ovaa, H. (2014) How chemistry supports cell biology: The chemical toolbox at your service. Trends Cell Biol. **24**, 751-760

12 Schultz, L. W. and Clardy, J. (1998) Chemical Inducers of Dimerization: The Atomic Structure of FKBP12-FK1012A-FKBP12. Bioorg. Med. Chem. Lett. **8**, 1-6

13 Liu, J., Farmer, J. D., Lane, W. S., Friedman, J., Weissman, I. and Schreiber, S. L. (1991) Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell. **66**, 807-815

14 Spencer, D. M., Wandless, T. J., Schreiber, S. L. and Crabtree, G. R. (1993) Controlling signal transduction with synthetic ligands. Science. **262**, 1019-1024

15 Belshaw, P. J., Spencer, D. M., Crabtree, G. R. and Schreiber, S. L. (1996) Controlling programmed cell death with a cyclophilin-cyclosporin-based chemical inducer of dimerization. Chem. Biol. **3**, 731-738

16 Ho, S. N., Biggar, S. R., Spencer, D. M., Schreiber, S. L. and Crabtree, G. R. (1996) Dimeric ligands define a role for transcriptional activation domains in reinitiation. In Nature ed.)^eds.). pp. 822-826

17 Putyrski, M. and Schultz, C. (2012) Protein translocation as a tool: The current rapamycin story. FEBS Lett. **586**, 2097-2105

18 Brown, E. J., Albers, M. W., Shin, T. B., Ichikawa, K., Keith, C. T., Lane, W. S. and Schreiber, S. L. (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature. **369**, 756-758

19 Chen, J., Zheng, X. F., Brown, E. J. and 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. Proc. Natl. Acad. Sci. USA. **92**, 4947-4951

20 Choi, J., Chen, J., Schreiber, S. L. and Clardy, J. (1996) Structure of the KKBP12-Rapamycin complex Interactign with the Binding Domain of Human FRAP. Science. **273**, 239-242

21 Liberles, S. D., Diver, S. T., Austin, D. J. and Schreiber, S. L. (1997) Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. Proc. Natl. Acad. Sci. USA. **94**, 7825-7830

22 Rollins, C. T., Rivera, V. M., Woolfson, D. N., Keenan, T., Hatada, M., Adams, S. E., Andrade, L. J., Yaeger, D., van Schravendijk, M. R., Holt, D. A., Gilman, M. and Clackson, T. (2000) A ligand-reversible dimerization system for controlling protein-protein interactions. Proc. Natl. Acad. Sci. USA. **97**, 7096-7101

23 Pecot, M. Y. and Malhotra, V. (2004) Golgi membranes remain segregated from the endoplasmic reticulum during mitosis in mammalian cells. Cell. **116**, 99-107

24 Zoncu, R., Perera, R. M., Balkin, D. M., Pirruccello, M., Toomre, D. and De Camilli, P. (2009) A phosphoinositide switch controls the maturation and signaling properties of APPL endosomes. Cell. **136**, 1110-1121

25 Zoncu, R., Perera, R. M., Sebastian, R., Nakatsu, F., Chen, H., Balla, T., Ayala, G., Toomre, D. and De Camilli, P. V. (2007) Loss of endocytic clathrin-coated pits upon acute depletion of phosphatidylinositol 4,5-bisphosphate. Proc. Natl. Acad. Sci. USA. **104**, 3793-3798

26 Robinson, M. S., Sahlender, D. A. and Foster, S. D. (2010) Rapid Inactivation of Proteins by Rapamycin-Induced Rerouting to Mitochondria. Dev. Cell. **18**, 324-331

27 Bear, J. E., Loureiro, J. J., Libova, I., Fassler, R., Wehland, J. and Gertler, F. B. (2000) Negative regulation of fibroblast motility by Ena/VASP proteins. Cell. **101**, 717-728

28 Cheeseman, L. P., Harry, E. F., McAinsh, A. D., Prior, I. A. and Royle, S. J. (2013) Specific removal of TACC3-ch-TOG-clathrin at metaphase deregulates kinetochore fiber tension. J Cell Sci. **126**, 2102-2113

29 Al-Bassam, S., Xu, M., Wandless, T. J. and Arnold, D. B. (2012) Differential trafficking of transport vesicles contributes to the localization of dendritic proteins. Cell Rep. **2**, 89-100

30 Haruki, H., Nishikawa, J. and Laemmli, U. K. (2008) The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol. Cell. **31**, 925-932

31 Broermann, A., Winderlich, M., Block, H., Frye, M., Rossaint, J., Zarbock, A., Cagna, G., Linnepe, R., Schulte, D., Nottebaum, A. F. and Vestweber, D. (2011) Dissociation of VE-PTP from VE-cadherin is required for leukocyte extravasation and for VEGF-induced vascular permeability in vivo. J Exp. Med. **208**, 2393-2401

32 Weitzman, M. and Hahn, K. M. (2013) Optogenetic approaches to cell migration and beyond. Curr. Opin. Cell Biol. **18**, 1199-1216

33 Hughes, J., Lamparter, T., Mittmann, F., Hartmann, E., Gärtner, W., Wilde, A. and Börner, T. (1997) A prokaryotic phytochrome. In Nature ed.)^eds.). p. 663

34 Vierstra, R. D. and Davis, S. J. (2000) Bacteriophytochromes: new tools for understanding phytochrome signal transduction. Sem. Cell Dev. Biol. **11**, 511-521

35 Quail, P. H. (2002) Phytochrome Photosensory Signalling Networks. Nat. Rev. Mol. Cell Biol. **3**, 85-93

36 Kendrick, R. E. and Kronenberg, G. H. M. (1994) Photomorphogenesis in plants. Kluwer Academic Publishers, Dordrecht-Boston-London

37 Kennedy, M. J., Hughes, R. M., Peteya, L. A., Schwartz, J. W., Ehlers, M. D. and Tucker, C. L. (2010) Rapid blue-light-mediated induction of protein interactions in living cells. Nat. Methods. **7**, 973-975

38 Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D. and Lin, C. (2008) Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. Science. **322**, 1535-1539

39 Idevall-Hagren, O., Dickson, E. J., Hille, B., Toomre, D. K. and De Camilli, P. (2012) Optogenetic control of phosphoinositide metabolism. Proc. Natl. Acad. Sci. USA. **109**, E2316-2323

40 Guglielmi, G., Barry, J. D., Huber, W. and DeRenzis, S. (2014) An Optogenetic Method to Modulate Cell Contractility during Tissue Morphogenesis. Dev. Cell. **35**, 646-660

41 Lee, S., Park, H., Kyung, T., Kim, N. Y., Kim, S., Kim, J. and Heo, W. D. (2014) Reversible protein inactivation by optogenetic trapping in cells. Nat. Methods. **11**, 633-636

42 Nguyen, M. K., Kim, C. Y., Kim, J. M., Park, B. O., Lee, S., Park, H. and Heo, W. D. (2016) Optogenetic oligomerization of Rab GTPases regulates intracellular membrane trafficking. Nat. Chem. Biol.

43 Bugaj, L. J., Choksi, A. T., Mesuda, C. K., Kane, R. S. and Schaffer, D. V. (2013) Optogenetic protein clustering and signaling activation in mammalian cells. Nat. Methods. **10**, 249-252

44 Rosenfeldt, G., Viana, R. M., Mootz, H. D., von Arnim, A. G. and Batschauer, A. (2008) Chemically induced and light-independent cryptochrome photoreceptor activation. Mol. Plant. **1**, 4-14

45 Wend, S., Wagner, H. J., Muller, K., Zurbriggen, M. D., Weber, W. and Radziwill, G. (2014) Optogenetic control of protein kinase activity in mammalian cells. ACS Synth. Biol. **3**, 280-285

46 Bugaj, L. J., Spelke, D. P., Mesuda, C. K., Varedi, M., Kane, R. S. and Schaffer, D. V. (2015) Regulation of endogenous transmembrane receptors through optogenetic Cry2 clustering. Nat. Commun. **6**, 6898

47 Hahn, K. M. and Kuhlman, B. (2010) Hold me tightly LOV. Nat. Methods. **7**, 595, 597

48 Wu, Y. I., Frey, D., Lungu, O. I., Jaehrig, A., Schlichting, I., Kuhlman, B. and Hahn, K. M. (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. Nature. **461**, 104-108

49 Wu, Y. I., Wang, X., He, L., Montell, D. and Hahn, K. M. (2011) Spatiotemporal Control of Small GTPases with Light Using the LOV domain. Methods Enzymol. **497**, 1-14

50 Wang, X., He, L., Wu, Y. I., Hahn, K. M. and Montell, D. J. (2010) Light-mediated activation reveals a key role for Rac in collective guidance of cell movement in vivo. Nat. Cell Biol. **12**, 591-597

51 Yao, X., Rosen, M. K. and Gardner, K. H. (2008) Estimation of the available free energy in a LOV2-J alpha photoswitch. Nat. Chem. Biol. **4**, 491-497

52 Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H. and Sosnick, T. R. (2010) Rationally improving LOV domain-based photoswitches. Nat. Methods. **7**, 623-626

53 Dolmans, D. E., Fukumura, D. and Jain, R. K. (2003) Photodynamic therapy for cancer. Nat. Rev. Cancer. **3**, 380-387

54 Grusch, M., Schelch, K., Riedler, R., Reichhart, E., Differ, C., Berger, W., Ingles-Prieto, A. and Janovjak, H. (2014) Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. EMBO J. **33**, 1713-1726

55 Toettcher, J. E., Voigt, C., Weiner, O. D. and Lim, W. A. (2011) The promise of optogenetics in cell biology: interrogating molecular circuits in space and time. Nat. Methods. **8**, 35-38

56 Paulus, H. (2000) Protein splicing and related forms of protein autoprocessing. Annu. Rev. Biochem. **69**, 447-496

57 Topilina, N. I. and Mills, K. V. (2014) Recent advances in in vivo applications of intein-mediated protein splicing. Mobile DNA. **5**, 5

58 Cheriyan, M. and Perler, F. B. (2009) Protein splicing: A versatile tool for drug discovery. Adv. Drug Deliv. Rev. **61**, 899-907

59 Starokadomskyy, P. L. (2007) Protein splicing. Mol. Biol. **41**, 278-293

60 Vila-Perello, M. and Muir, T. W. (2010) Biological Applications of Protein Splicing. Cell. **143**, 191-200

61 Carrington, J. C. and Dougherty, W. G. (1988) A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proc. Natl. Acad. Sci. USA. **85**, 3391-3395

62 Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F. and Fletterickt, R. (1989) Characterization of the Catalytic Residues of the Tobacco Etch Virus 49-kDa Proteinase. Virology. **172**, 302-310

63 Kapust, R. B., Tözsér, J., Copeland, T. D. and Waugh, D. S. (2002) The P1' specificity of tobacco etch virus protease. Biochem. Biophys. Res. Commun. **294**, 949-955

64 Kapust, R. B. and Waugh, D. S. (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Express. Purif. **19**, 312-318

65 Uhlmann, F., Wernic, D., Poupart, M. A., Koonin, E. V. and Nasmyth, K. (2000) Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. Cell. **103**, 375-386

66 Harder, B., Schomburg, A., Pflanz, R., Küstner, K. M., Gerlach, N. and Schuh, R. (2008) TEV protease-mediated cleavage in Drosophila as a tool to analyze protein functions in living organisms. BioTechniques. **44**, 765-772

67 Pauli, A., Althoff, F., Oliveira, R. A., Heidmann, S., Schuldiner, O., Lehner, C. F., Dickson, B. J. and Nasmyth, K. (2008) Cell-Type-Specific TEV Protease Cleavage Reveals Cohesin Functions in Drosophila Neurons. Dev. Cell. **14**, 239-251

68 Williams, D. J., Puhl, H. L. and Ikeda, S. R. (2009) Rapid modification of proteins using a rapamycin-inducible tobacco etch virus protease system. PLoS One. **4**

69 Fang, J., Chen, L., Cheng, B. and Fan, J. (2013) Engineering soluble tobacco etch virus protease accompanies the loss of stability. Protein Express. Purif. **92**, 29-35

70 Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. and Waugh, D. S. (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. Protein Eng. **14**, 993-1000

71 Cesaratto, F., Lopez-Requena, A., Burrone, O. R. and Petris, G. (2015) Engineered tobacco etch virus (TEV) protease active in the secretory pathway of mammalian cells. J. Biotechnol. **212**, 159-166

72 Jacobson, K., Rajfur, Z., Vitriol, E. and Hahn, K. (2008) Chromophore-assisted laser inactivation in cell biology. Trends Cell Biol. **18**, 443-450

73 Jay, D. G. (1988) Selective destruction of protein function by chromophore-assisted laser inactivation. Proc. Natl. Acad. Sci. USA. **85**, 5454-5458

74 Li, W., Stuurman, N. and Ou, G. (2012) Chromophore-assisted laser inactivation in neural development. Neurosci Bull. **28**, 333-341

75 Beck, S., Sakurai, T., Eustace, B. K., Beste, G., Schier, R., Rudert, F. and Jay, D. G. (2002) Fluorophore-assisted light inactivation: A high-throughput tool for direct target validation of proteins. Proteomics. **2**, 247-255

76 Marek, K. W. and Davis, G. W. (2002) Transgenically encoded protein photoinactivation (FlAsH-FALI): Acute inactivation of synaptotagmin I. Neuron. **36**, 805-813

77 Boulina, M. E., Lukyanov, K. A., Britanova, O. V., Onichtchouk, D., Lukyanov, S. and Chudakov, D. M. (2006) Chromophore-assisted light inactivation (CALI) using the phototoxic fluorescent protein KillerRed. Nat. Protoc. **1**, 947-953

78 Ou, G., Stuurman, N., D'Ambrosio, M. and Vale, R. D. (2010) Polarized myosin produces unequal-size daughters during asymmetric cell division. Science. **330**, 677-680

79 Boulina, M. E., Chudakov, D. M., Britanova, O. V., Yanushevich, Y. G., Staroverov, D. B., Chepurnykh, T. V., Merzlyak, E. M., Shkrob, M. A., Lukyanov, S. and Lukyanov, K. A. (2006) A genetically encoded photosensitizer. Nat. Biotechnol. **24**, 95-99

80 Pletnev, S., Gurskaya, N. G., Pletneva, N. V., Lukyanov, K. A., Chudakov, D. M., Martynov, V. I., Popov, V. O., Kovalchuk, M. V., Wlodawer, A., Dauter, Z. and Pletnev, V. (2009) Structural basis for phototoxicity of the genetically encoded photosensitizer KillerRed. J Biol. Chem. **284**, 32028-32039

81 Kobayashi, J., Shidara, H., Morisawa, Y., Kawakami, M., Tanahashi, Y., Hotta, K. and Oka, K. (2013) A method for selective ablation of neurons in C. elegans using the phototoxic fluorescent protein, KillerRed. Neurosci. Lett. **548**, 261-264

82 Yan, L., Kanada, M., Zhang, J., Okazaki, S. and Terakawa, S. (2015) Photodynamic treatment of tumor with bacteria expressing killerred. PLoS One. **10**, 1-15

83 Takemoto, K., Matsuda, T., Sakai, N., Fu, D., Noda, M., Uchiyama, S., Kotera, I., Arai, Y., Horiuchi, M., Fukui, K., Ayabe, T., Inagaki, F., Suzuki, H. and Nagai, T. (2013) SuperNova, a monomeric photosensitizing fluorescent protein for chromophore-assisted light inactivation. Sci. Rep. **3**, 2629

84 Shu, X., Lev-Ram, V., Deerinck, T. J., Qi, Y., Ramko, E. B., Davidson, M. W., Jin, Y., Ellisman, M. H. and Tsien, R. Y. (2011) A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. PLoS Biol. **9**, e1001041

85 Sarkisyan, K. S., Zlobovskaya, O. A., Gorbachev, D. A., Bozhanova, N. G., Sharonov, G. V., Staroverov, D. B., Egorov, E. S., Ryabova, A. V., Solntsev, K. M., Mishin, A. S. and Lukyanov, K. A. (2015) KillerOrange, a Genetically Encoded Photosensitizer Activated by Blue and Green Light. PloS One. **10**, e0145287

86 Guo, J., Chen, H., Puhl, H. L. and Ikeda, S. R. (2006) Fluorophore-assisted light inactivation produces both targeted and collateral effects on N-type calcium channel modulation in rat sympathetic neurons. The Journal of physiology. **576**, 477-492

87 Yu, G., Rosenberg, J. N., Betenbaugh, M. J. and Oyler, G. A. (2015) Pac-Man for biotechnology: co-opting degrons for targeted protein degradation to control and alter cell function. Curr. Opin. Biotechnol. **36**, 199-204

88 Jackson, P. K., Eldridge, A. G., Freed, E., Furstenthal, L., Hsu, J. Y., Kaiser, B. K. and Reimann, J. D. R. (2000) The lore of the RINGs: Substrate recognition and catalysis by ubiquitin ligases. Trends Cell Biol. **10**, 429-439

89 Bachmair, A., Finley, D. and Varshavsky, A. (1986) In vivo half-life of a protein is a function of its amino-terminal residue. Science. **234**, 179-186

90 Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. Protein Sci. **20**, 1298-1345

91 Dohmen, R. J., Wu, P. and Varshavsky, A. (1994) Heat-inducible degron: a method for constructing temperature-sensitive mutants. Science. **263**, 1273-1276

92 Speese, S. D., Trotta, N., Rodesch, C. K., Aravamudan, B. and Broadie, K. (2003) The Ubiquitin Proteasome System Acutely Regulates Presynaptic Protein Turnover and Synaptic Efficacy. Med. Eng. Phys. **13**, 899-910

93 Kanemaki, M., Sanchez-Diaz, A., Gambus, A. and Labib, K. (2003) Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. Nature. **423**, 720-724

94 Su, X., Bernal, J. A. and Venkitaraman, A. R. (2008) Cell-cycle coordination between DNA replication and recombination revealed by a vertebrate N-end rule degron-Rad51. Nat. Struct. Mol. Biol. **15**, 1049-1058

95 Johnston, J. A., Johnson, E. S., Waller, P. R. H. and Varshavsky, A. (1995) Methotrexate inhibits proteolysis of dihydrofolate reductase by the N-end rule pathway. In J Biol. Chem. ed.)^eds.). pp. 8172-8178

96 Lévy, F., Johnston, J. A. and Varshavsky, A. (1999) Analysis of a conditional degradation signal in yeast and mammalian cells. Eur. J Biochem. **259**, 244-252

97 Bence, N. F., Sampat, R. M. and Kopito, R. R. (2001) Impairment of the ubiquitin-proteasome system by protein aggregation. Science. **292**, 1552-1555

98 Rajagopalan, P. T. R., Zhang, Z., McCourt, L., Dwyer, M., Benkovic, S. J. and Hammes, G. G. (2002) Interaction of dihydrofolate reductase with methotrexate: ensemble and single-molecule kinetics. Proc. Natl. Acad. Sci. USA. **99**, 13481-13486

99 Taxis, C., Stier, G., Spadaccini, R. and Knop, M. (2009) Efficient protein depletion by genetically controlled deprotection of a dormant N-degron. Mol. Syst. Biol. **5**, 1-7

100 Rodríguez-Limas, W. A., Tannenbaum, V. and Tyo, K. E. J. (2015) Blocking endocytotic mechanisms to improve heterologous protein titers in Saccharomyces cerevisiae. Biotechnol. Bioeng. **112**, 376-385

101 Zhou, P., Bogacki, R., McReynolds, L. and Howley, P. M. (2000) Harnessing the Ubiquitination Machinery to Target the Degradation of Specific Cellular Proteins. Mol. Cell. **6**, 751-756

102 Gray, W. M., Pozo, J. C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W. L., Yang, M., Ma, H. and Estelle, M. (1999) Identification of an SCF ubiquitin–ligase complex required for auxin response in Arabidopsis thaliana. Genes Dev. **13**, 1678-1691

103 Ruegger, M., Dewey, E., Gray, W. M., Hobbie, L., Turner, J. and Estelle, M. (1998) The TIR1 protein of Arabidopsis functions in auxin response and is related to human SKP2 and yeast Grr1p. Genes Dev. **12**, 198-207

104 Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. and Kanemaki, M. (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods. **6**, 917-922

105 Zhang, L., Ward, J. D., Cheng, Z. and Dernburg, A. F. (2015) The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. Development. **142**, 4374-4384

106 Holland, A. J., Fachinetti, D., Han, J. S. and Cleveland, D. W. (2012) Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc. Natl. Acad. Sci. USA. **109**, E3350-3357

107 Morawska, M. and Ulrich, H. D. (2013) An expanded tool kit for the auxin-inducible degron system in budding yeast. Yeast. **30**, 341-351

108 Sakamoto, K. M., Kim, K. B., Kumagai, A., Mercurio, F., Crews, C. M. and Deshaies, R. J. (2001) Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. Proc. Natl. Acad. Sci. USA. **98**, 8554-8559

109 Deshaies, R. J. (2015) Protein degradation: Prime time for PROTACs. Nat. Chem. Biol. **11**, 634-635

110 Toure, M. and Crews, C. M. (2016) Small-Molecule PROTACS: New Approaches to Protein Degradation. Angew Chem. Int. Ed. Engl. **55**, 1966-1973

111 Bondeson, D. P., Mares, A., Smith, I. E., Ko, E., Campos, S., Miah, A. H., Mulholland, K. E., Routly, N., Buckley, D. L., Gustafson, J. L., Zinn, N., Grandi, P., Shimamura, S., Bergamini, G., Faelth-Savitski, M., Bantscheff, M., Cox, C., Gordon, D. A., Willard, R. R., Flanagan, J. J., Casillas, L. N., Votta, B. J., den Besten, W., Famm, K., Kruidenier, L., Carter, P. S., Harling, J. D., Churcher, I.

and Crews, C. M. (2015) Catalytic in vivo protein knockdown by small-molecule PROTACs. Nat. Chem. Biol. **11**, 611-617

112 Lu, J., Qian, Y., Altieri, M., Dong, H., Wang, J., Raina, K., Hines, J., Winkler, J. D., Crew, A. P., Coleman, K. and Crews, C. M. (2015) Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. Chem. Biol. **22**, 755-763

113 Winter, G. E., Buckley, D. L., Paulk, J., Roberts, J. M., Souza, A., Dhe-Paganon, S. and Bradner, J. E. (2015) DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. Science. **348**, 1376-1381

114 Zengerle, M., Chan, K. H. and Ciulli, A. (2015) Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. ACS Chem. Biol. **10**, 1770-1777

115 Galdeano, C., Gadd, M. S., Soares, P., Scaffidi, S., Van Molle, I., Birced, I., Hewitt, S., Dias, D. M. and Ciulli, A. (2014) Structure-guided design and optimization of small molecules targeting the protein-protein interaction between the von Hippel-Lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. J Med. Chem. **57**, 8657-8663

116 Caussinus, E., Kanca, O. and Affolter, M. (2011) Fluorescent fusion protein knockout mediated by anti-GFP nanobody. Nat. Struct. Mol. Biol. **19**, 117-121

117 Dunst, S., Kazimiers, T., von Zadow, F., Jambor, H., Sagner, A., Brankatschk, B., Mahmoud, A., Spannl, S., Tomancak, P., Eaton, S. and Brankatschk, M. (2015) Endogenously Tagged Rab Proteins: A Resource to Study Membrane Trafficking in Drosophila. Dev. Cell. **33**, 351-365

118 Brankatschk, M., Dunst, S., Nemetschke, L. and Eaton, S. (2014) Delivery of circulating lipoproteins to specific neurons in the Drosophila brain regulates systemic insulin signaling. Elife. **3**

119 Shin, Y. J., Park, S. K., Jung, Y. J., Kim, Y. N., Kim, K. S., Park, O. K., Kwon, S. H., Jeon, S. H., Trinh le, A., Fraser, S. E., Kee, Y. and Hwang, B. J. (2015) Nanobody-targeted E3-ubiquitin ligase complex degrades nuclear proteins. Sci. Rep. **5**, 14269

120 Janse, D. M., Crosas, B., Finley, D. and Church, G. M. (2004) Localization to the proteasome is sufficient for degradation. J Biol. Chem. **279**, 21415-21420

121 Renicke, C., Schuster, D., Usherenko, S., Essen, L. O. and Taxis, C. (2013) A LOV2 domain-based optogenetic tool to control protein degradation and cellular function. Chem. Biol. **20**, 619-626

122 Mettetal, J. T., Muzzey, D., Gomez-Uribe, C. and van Oudenaarden, A. (2008) The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae. Science. **319**, 482-484

123 Paszek, P., Jackson, D. A. and White, M. R. (2010) Oscillatory control of signalling molecules. Current opinion in genetics & development. **20**, 670-676

124 Tsien, R. W. and Tsien, R. Y. (1990) Calcium channels, stores, and oscillations. Annual review of cell biology. **6**, 715-760

125 Taslimi, A., Vrana, J. D., Chen, D., Borinskaya, S., Mayer, B. J., Kennedy, M. J. and Tucker, C. L. (2014) An optimized optogenetic clustering tool for probing protein interaction and function. Nat. Commun. **5**, 4925

126 Cluzel, P., Surette, M. and Leibler, S. (2000) An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. Science. **287**, 1652-1655

127 Diekmann, S. and Hoischen, C. (2014) Biomolecular dynamics and binding studies in the living cell. Physics of life reviews. **11**, 1-30

128 Zhou, X., Herbst-Robinson, K. J. and Zhang, J. (2012) Visualizing dynamic activities of signaling enzymes using genetically encodable FRET-based biosensors from designs to applications. Methods Enzymol. **504**, 317-340

129 Kawakami, K., Abe, G., Asada, T., Asakawa, K., Fukuda, R., Ito, A., Lal, P., Mouri, N., Muto, A., Suster, M. L., Takakubo, H., Urasaki, A., Wada, H. and Yoshida, M. (2010) zTrap: zebrafish gene trap and enhancer trap database. BMC Dev. Biol. **10**, 105

130 Nagarkar-Jaiswal, S., Lee, P. T., Campbell, M. E., Chen, K., Anguiano-Zarate, S., Gutierrez, M. C., Busby, T., Lin, W. W., He, Y., Schulze, K. L., Booth, B. W., Evans-Holm, M., Venken, K. J., Levis, R. W., Spradling, A. C., Hoskins, R. A. and Bellen, H. J. (2015) A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. Elife. **4**

131 Pinder, J., Salsman, J. and Dellaire, G. (2015) Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. Nucleic Acids Res. **43**, 9379-9392

Figure 1. Schematic representation of methods for the control of protein activity by induced complex formation. Methods described here can be broadly categorised by their method of induction. (A) Generalised mechanism for small-molecule mediated approaches involving small molecule, chemical inducers of dimerisation (CID) to sequester protein activity by promoting (dis)aggregation, altering transcription or by changing the sub-cellular localisation of target proteins. (B) A specific example of this approach is illustrated by the Knocksideways method, which involves the rapid sequestration of target proteins to the mitochondria through rapamycin mediated interaction between the FKBP tagged target protein and the mitochondrially localised mito-RFP-FRB trap. (C) The generalised mechanism for light-based approaches induced by exposure to a specific wavelength of light. More specific examples of light-based methods are also illustrated in D-E. (D) Schematic of the light-activated reversible inhibition by assembled trap (LARIAT) method, whereby target proteins tagged with CRY2 are sequestered into large protein complexes upon a light-induced conformational change to produce the photoactivated form of CRY2. This interacts with CIB1-bound multimeric complexes to aggregate the target protein and sequester activity. (E) Representation of LOV-domain-based approaches, which involve a light-induced conformational change leading to unravelling of the jα-helix and the loss of its interaction with the LOV domain leaving the target protein free to bind interactors and perform its function.

Figure 2. Clustering Indirectly Using Cryptochrome 2 (CLICR). In the dark, un-induced state, monomers of Cry2 fused to a receptor targeting binding domain (BD) (Cry2-BD) exist in an unclustered state and therefore have a weak affinity for the target receptor. Upon blue light stimulation, Cry2-BD molecules oligomerise, increasing the local concentration of BD and conferring a high avidity for the target receptor. These oligomers undergo membrane translocation and cluster target transmembrane receptors.

Figure 3. Illustration of methods for conditional protein splicing or protein cleavage to (in)activate target proteins. (A) Illustration of conditional protein *cis*-splicing induced by activation of an intein through a change in redox state or via a trigger, which may be addition of a small molecule, as depicted here, a change in pH, temperature or irradiation with a specific wavelength of light. Protein *trans*-splicing is also possible, whereby dimerisation domains can be used to reassociate split intein fragments upon addition of a trigger. (B) Illustration of TEV protease mediated cleavage of a TEV recognition site engineered within a protein of interest leading to protein inactivation upon induction of TEV protease expression. (C) Schematic representation of chromophore/fluorophore-assisted light inactivation (CALI/FALI) induced by addition of a dye conjugated ligand or antibody, or via genetically-encoded methods involving photosensitisers such as KillerRed, eGFP, miniSOG or SuperNova. Upon irradiation with a specific wavelength of light these produce reactive oxygen species ($^1O_2$) leading to inactivation of proteins in the immediate vicinity.

Figure 4. Illustration of methods for inducible protein degradation divided into those involving the N-end rule or F-box based pathways (A) N-end degron method, involving the exposure of an unstable N-terminal amino acid which can be induced by a number of mechanisms including temperature or small molecule based mechanisms. This unstable end is subsequently targeted for polyubiquitination by an E3 ligase, such as Ubr1P in yeast, and degraded via the proteasome pathway. (B) TEV protease mediated induction of protein instability (TIPI) which utilises TEV protease mediated cleavage of a seven amino acid TEV recognition site to reveal an unstable N-terminal amino acid, subsequently targeted for proteasome mediated degradation as for the N-end degron approach. The efficiency of TEV cleavage is increased by the inclusion of a short SF3b155$^{381-424}$ domain downstream of the TEV recognition site, which binds to a mutant version of the human spliceosome subunit 14 (p14*) bringing p14*-TEV to the recognition site. (C) Illustration of the mechanism for auxin-induced degradation of target proteins tagged with an auxin-inducible degron (AID). Upon addition of auxin, the AID tagged protein is recruited to an engineered E3 ubiquitin ligase SCF complex, containing the transport inhibitor response 1 (TIR1) F-box protein from plants, which binds target proteins in the presence of the plant hormone auxin. Expressing TIR1 in non-plant cells is enough to result in formation of the SCF complex which then binds AID tagged proteins in an auxin-dependent manner, leading to poly-ubiquitination and degradation via the proteasome. (D) Schematic of the deGradFP method for inducible protein degradation. The GFP tagged target protein is recruited to an engineered SCF complex containing the F-box protein NSlmb conjugated to an anti-GFP single chain antibody (vhhGFP4). Target protein is polyubiquitinated via the recruited E2 ubiquitin ligase and subsequently degraded via the endogenous proteasome machinery.

Table 1. Summary of methods for conditional control of protein complex formation

Technique	Protein Disruption	Timescale	Can be induced by …			
			Small molecule /hormone	Light	Temperature	pH
Chemically induced dimerisation (CID)	Sequestration/ (in)activation	min/h	✓ e.g. FK1012, Rapamycin	✗	✗	✗
Knocksideways (KS)	Sequestration	min	✓ Rapamycin	✗	✗	✗
Light-activated dimerisation	Sequestration/ (In)activation	sec/min	✗	✓	✗	✗
Light-Activated Reversible Inhibition by Assembled Trap (LARIAT)	(In)activation	sec	✗	✓ Blue light: 450-495 nm	✗	✗
Light-Oxygen-Voltage (LOV) domains	(In)activation	sec/min	✗	✓ Blue light: 450-495 nm	✗	✗

Table 2. Summary of methods for conditional control of protein splicing/cleavage

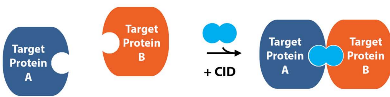
Technique	Protein Disruption	Timescale	Can be induced by …			
			Small molecule /hormone	Light	Temperature	pH
Intein mediated protein splicing	Inactivation via protein splicing	h	✓	✓	✓	✓
Tobacco Etch Virus (TEV)-protease cleavage	Inactivation via protein cleavage	min	Promoter dependent			
Chromophore-Assisted Light Inactivation (CALI)	Inactivation by ROS	Often <1 sec	✗	✓ e.g. Malachite green: 616-624 nm KillerRed: 540-580 nm	✗	✗
Fluorophore-Assisted Light Inactivation (FALI)	Inactivation by ROS	<10 min	✗	✓ e.g. Fluorescin isothiocyanate 493-518 nm	✗	✗

Table 3. Summary of methods for conditional control of protein degradation

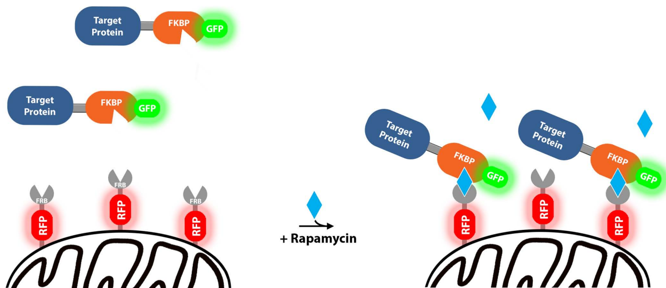
Technique	Protein Disruption	Timescale	Can be induced by …			
			Small molecule /hormone	Light	Temperature	pH
N-end Degron	Unstable N-end amino acid ubiquitination	min-h, sometimes <30 min	✓ e.g. Methotrexate	✗	✓ e.g. 37-42°C	✗
TEV protease mediated induction of protein instability (TIPI)	Unstable N-end amino acid ubiquitination	min/h	Promoter dependent			
Auxin-induced degradation (AID)	F-box induced ubiquitination	<30 min	✓ Auxin	✗	✗	✗
Proteolysis Targeting Chimeras (PROTACS)	Direct targeting to E3 ligase complexes	min/h	✓ PROTAC	✗	✗	✗
deGradFP	F-box induced ubiquitination	min/h	Promoter dependent			

Chemically Induced Dimerisation

A.

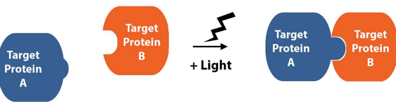


B. Knocksideways (KS) Method

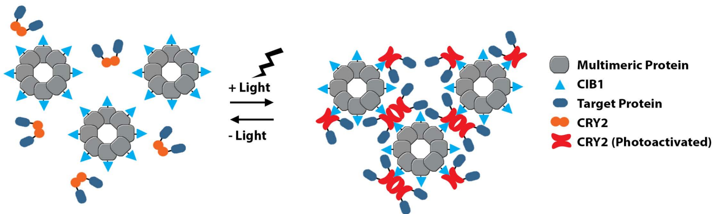


Light Induced Dimerisation

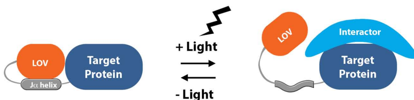
C.

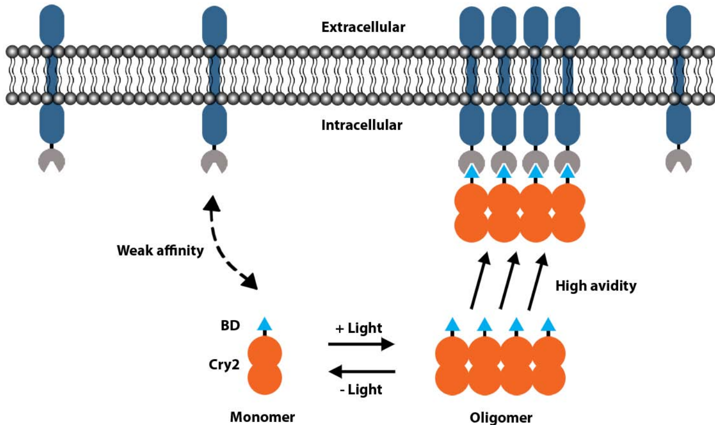


D. LARIAT



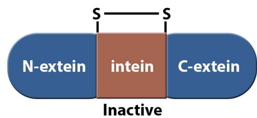
E. LOV domains



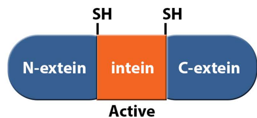

Extracellular
Intracellular
Weak affinity
High avidity
BD
Cry2
+ Light
- Light
Monomer
Oligomer

Protein Splicing

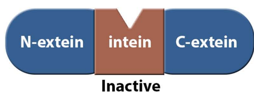
A. Redox state



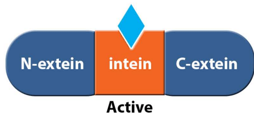
Reduction

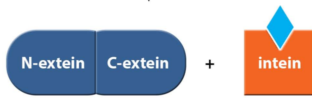


Triggered Splicing



Trigger

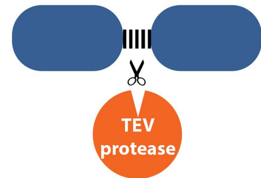


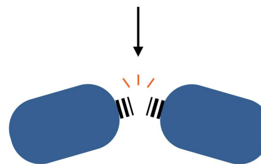


Protein Cleavage

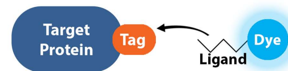
B. TEV-protease

Engineered TEV recognition site

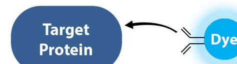




C. CALI/FALI

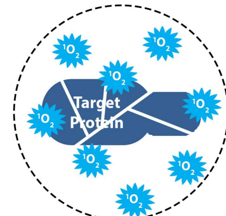


OR



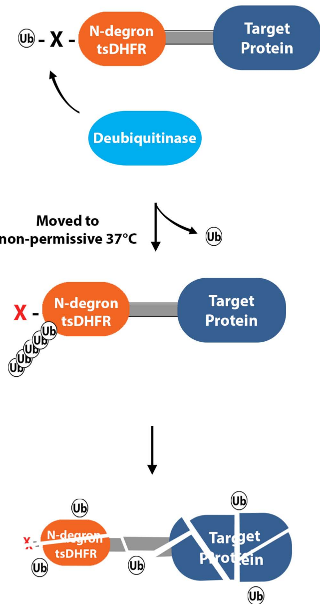
OR

Light irradiation

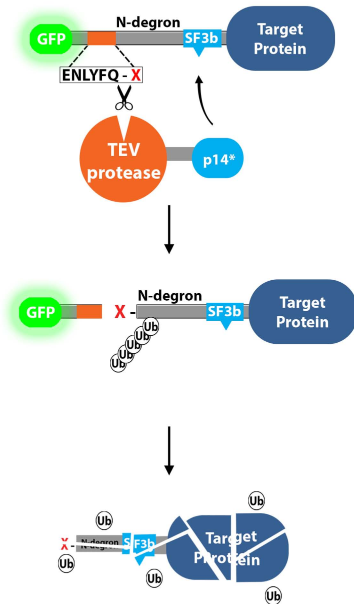


N-end Rule Degradation

A. N-degron

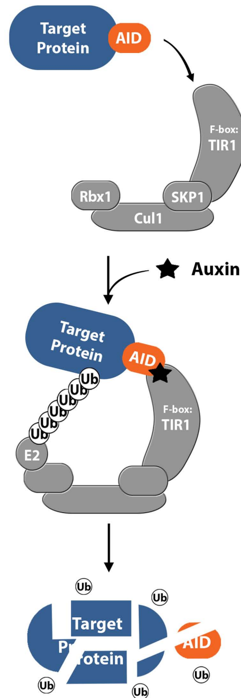


B. TIPI



F - box Degradation

C. Auxin Induced



D. DeGradFP