7 Putting genetics into optogenetics: knocking out proteins with light

7.1 Introduction

The disruption of gene function has been a central method for understanding biological processes for over a century. However, there are three major weaknesses to classical genetic analysis: pleiotropic mutations, strain lethals, and cell lethals. First, many genetic knockouts affect multiple tissues, which can complicate the interpretation of phenotypes for any particular cell. These pleiotropies may arise from non-autonomous effects from other tissues or broad nonspecific defects, such as general ill health. Second, some genes are mutated to lethality; that is, the mutants die during development, making it difficult to study the role of a gene at later time points. Third, some genes are required for the viability of a cell; for example, genes involved in basic cellular functions, like transcription and translation. A cell with such a mutation will die even as a mosaic in an otherwise wild-type animal. The most common methods to circumvent these drawbacks of traditional genetic analysis are mRNA destabilization by RNA interference, or inducible DNA mutations; for example, by using Cre/Lox recombination. These methods allow the study of gene function later in development in specific tissues. However, the temporal resolution of these methods is limited by the half-life of the protein of interest. Ideally, an experimenter would be able to instantaneously remove a protein from any cell of choice. Here we review methods for the induction of protein degradation, and speculate about the potential use of light to stimulate protein degradation.

7.2 Protein degradation

Methods for inducing protein degradation should satisfy multiple criteria to maximize their utility. First, the method should be rapidly inducible so that acute phenotypes can be studied. Second, protein levels should be reduced as close as possible to that of a genetic null. Third, the technique should be applicable to a broad range of protein substrates and model organisms. This list of requirements is ambitious and current technologies for inducible protein degradation are not yet ideal.

Perhaps the simplest and most direct method for disrupting the function of a protein would be to have a protease that is highly specific for the protein of interest, to simply cut the protein and inactivate it. The classic examples for such proteases are the clostridial neurotoxins [1, 2, 3]; these toxins are made by bacteria from the genus Clostridium, and are among the deadliest toxins known. These toxins are highly specific proteases for particular SNARE proteins that function in the release of neu-
rotransmitters at the synapse. They bind with extremely high specificity and cleave the target protein (Figure 7.1A). However, nature has not been so generous to provide a broad-range of specific proteases, so their use is limited. One could imagine engineering such proteases for new targets, but so far this approach has not been pursued.

![Diagram](image)

**Figure 7.1:** Sequence-specific protease destruction of proteins. (a) Tetanus toxin. Clostridial neurotoxins are made by bacteria that cleave specific proteins that function in neurotransmitter release. The tetanus toxin cleaves the SNARE protein synaptobrevin. (b) TEV protease cleavage. Tobacco Etch Virus (TEV) proteases cleave at a specific seven amino acid sequence. This sequence can be inserted into sequence and cleave the protein to disrupt function.

An alternative to developing target-specific proteases is to modify the target to be sensitive to a pre-existing protease (Figure 7.1B). For example, one could insert the seven amino target sequence for the Tobacco Etch Virus (TEV) protease into a gene. Transgenics that express this modified protein can then be targeted by expression of the TEV protease [4]. A problem with this method is that it does not generate a knockout, but rather two fragments of the target protein. These fragments can bind other proteins acting in the pathway and eliminate their function as well, thereby generating a more severe phenotype than the simple loss of the target protein. Alternatively, the fragments may have novel targets and act in a ‘neomorphic’ fashion on other molecular pathways.

A more interpretable result can be achieved not just by cleaving the protein, but instead by fully degrading it. Inducible protein degradation systems (or degrons) use the conserved ubiquitin–proteasome pathway to quickly degrade specific proteins. The ubiquitin pathway begins with the transfer of ubiquitin from the E1 ubiquitin-activating enzyme to the E2 ubiquitin-conjugating enzyme (Figure 7.2A). The E3 ubiquitin ligase recognizes both the target protein and the E2 conjugating enzyme and transfers the ubiquitin to a lysine in the target protein. At least three more ubiquitins must be added to the first ubiquitin modification to generate a poly-ubiquitinated substrate that is tar-
geted for degradation by the proteasome [5]. E3 ubiquitin ligases can be divided into two molecular classes based on conserved domains: HECT and RING [6, 7].

![Diagram of the ubiquitin-proteasome pathway](image)

**Figure 7.2:** Ubiquitin-proteasome protein degradation. (a) The canonical ubiquitin-proteasome pathway. The E1-ubiquitin activating enzyme transfers ubiquitin to the E2 ubiquitin-conjugating enzyme. There are many E3 ubiquitin ligases and they add ubiquitin to specific substrates. Polyubiquitinated proteins are degraded by the proteasome. (b) Heat-inducible degron. The sequence for a temperature-sensitive dihydrofolate reductase (DHFR) sequence can be attached to the N-terminus of a protein. This sequence has an arginine at the N-terminus, which destabilizes proteins, but in this case the arginine is buried in the folded DHFR structure. A shift to 35 °C induces a conformational change in the tsDHFR, which exposes the N-terminal Arg. The E3 ubiquitin ligase, UBR1, binds the Arg-DHFR, recruits E2 Ub-conjugase, and promotes the poly-ubiquitination and degradation of the DHFR and the attached protein of interest. (c) Small molecule inducible degron. The FKBP domain undergoes a conformational change upon small molecule binding that exposes a degron. We presume that this results in binding by chaperones (Hsp) and recognition by the E3 ubiquitin ligase CHIP. The E3 ligase recruits E2 Ub-conjugase, which attaches a ubiquitin chain on the degradation domain, resulting in degradation of this domain and the protein of interest.
The RING finger E3 ligase family has been used to destabilize specific proteins [8]. These proteasome-dependent degrons differ in their method for inducing interactions between the E3 ubiquitin ligase complex and the target protein. The protein can be destabilized by: amino acids at the N-terminus of the protein that promote degradation; a sequence that can be unfolded by temperature; a sequence that can be unfolded by a small molecule; or by binding the target protein to F-box proteins in the E3 ligase. We discuss each of these approaches below.

The UBR1 E3 ligase acts on proteins with destabilizing N-terminal amino acids according to the N-end rule [9]. Varshavsky and coworkers observed that the half-life of a protein is determined by the identity of its N-terminal amino acid and accessible lysine residues, now known as the N-end rule [10]. They placed the coding sequence for ubiquitin at the 5’ end of a lacI linker sequence followed by the β-galactosidase (β-gal) gene. A yeast deubiquitinating protease rapidly cleaved off the ubiquitin, and a new N-terminal amino acid was exposed. The half-life of β-gal radically changed, depending on the identity of the N-terminal residue. For example, Met and Val N-termini resulted in a β-gal half-life of greater than 20 hours, whereas Arg and Phe N-termini showed a half-life of approximately 3 minutes [11]. Varshavsky and coworkers found that degradation was dependent on lysines found in the lacI linker that had been placed at the N-terminus of β-gal [12]. β-gal degradation relies on the UBR1 protein, which is an E3 ubiquitin ligase that recognizes destabilizing N-terminal residues and recruits an E2 ubiquitin conjugase to add poly-ubiquitin chains to the protein. These poly-ubiquitin chains mark the β-gal protein for degradation by the proteasome.

This unstable LacI sequence can be attached to a protein of interest to generate an unstable version of the protein [13]. Proteins tagged with this sequence are unstable and are degraded in about three minutes. However, this degron is not inducible. To generate a heat-inducible degron, Dohmen et al. screened for a variant of dihydrofolate reductase in which the N-terminus is only exposed at high temperatures (tsDHFR) [14]. The N-terminal acid in this sequence is a destabilizing arginine (Figure 7.2B). When the tsDHFR degron was fused to a protein of interest, the target protein was rapidly degraded at the non-permissive temperature [14]. The tsDHFR degron has been used to degrade proteins in yeast [7, 15] and at the *Drosophila* neuromuscular junction [16]. There are two limitations to this technology. First, the degron only destabilizes proteins at 35 °C, which limits its utility to organisms that can survive at this temperature. Second, the degron only works as an N-terminal fusion, which restricts its application to proteins that are functional with large N-terminal extensions.

An alternative to induction of the degron by elevated temperature is induction by small molecules or drugs (Figure 7.2C). The Wandless lab screened for DHFR and FK506-binding protein (FKBP) domains that are stabilized or destabilized in the presence of a small molecule [17, 18]. The destabilized domains can be fused to either end of a target protein, and addition of the small molecule either stabilizes the protein, or leads to target protein degradation. However, degrons of this type suffer from one major deficiency – they are slow. The time to degradation is reported to be greater than 4 hr
Putting genetics into optogenetics: knocking out proteins with light

This is most likely due to the mechanism of degradation. The destabilizing domain becomes unfolded when the molecule is removed [18, 19, 21, 22], or even more usefully becomes unfolded when the small molecule is added [17]. The unfolded protein is then likely to be recognized by chaperone proteins, which attempt to stabilize and refold it. In this model, the chaperone and target protein complex are bound by CHIP, which is a U-box E3 ligase, structurally related to the RING finger ligases [23]. The CHIP E3 ligase recruits the E2 Ub-conjugase, which poly-ubiquitinates the target protein and results in its degradation. CHIP recruitment is likely the limiting step; the unfolded protein can go through multiple rounds of attempted refolding by the chaperone before it is finally degraded [24]. If essential proteins are degraded slowly, the cell will undergo a prolonged and perhaps highly pleiotropic cell death. If researchers wish to study essential genes, a degron must be fast – a strength of the tsDHFR system, which can be achieved by more direct coupling to proteasome degradation.

The Cullin-RING complex ubiquitin ligases are particularly adaptable for the direct targeting of the ubiquitin-proteasome machinery to a protein substrate [8, 23, 25]. F-box proteins from E3 ligase complexes can be modified to act as protein-specific degrons [8]. The F-box domain protein binds the target protein and recruits the Cullin-RING complex, also called the SCF complex for its conserved constituents Skp1, Cullin, and F-box, which ubiquitinates the target (Figure 7.3A) [26]. Importantly, the F-box recruits all the required machinery for degradation; targeting a protein of interest to an F-box results in degradation within ~1 hr in yeast and mammalian cells [26]. For example, a 35aa segment from the human papilloma virus (HPV) binds the oncogenic retinoblastoma RB protein; this segment was fused to a Cullin-RING complex F-box protein. The HPV-F-box fusion protein caused the degradation of the retinoblastoma protein. In another example, an F-box was designed to specifically target and degrade β-catenin in trans [27]. The APC protein from the canonical Wnt signaling pathway, binds to β-catenin. Su et al. fused this binding domain to an F-box to make a chimeric F-box protein. They expressed this chimera in colorectal cancer cells, which targeted free β-catenin for degradation [27].

These methods are specific to a single target and are not inducible, but demonstrate that researchers need only recruit the F-box to a protein target to induce degradation. A more broadly applicable method is to use antibody fragments fused to the F-box protein to recruit the target to the Cullin-RING complex. For example, the Affolter group fused the coding sequence for a single-domain antibody targeted to GFP to an F-box [28]. When expressed in cells containing a GFP-tagged target protein, the anti-GFP domain binds the GFP moiety of the target protein and the Cullin-RING complex induces degradation of the protein (Figure 7.3A). This design is convenient for multiple reasons: first, GFP fusion constructions already exist for many proteins, second, the degradation is measurable by loss of fluorescence, and third, tissue-specific degradation is possible by expressing the antibody-F-box chimera in specific cells. Importantly, it has been shown that this degron works in vivo in Drosophila melanogaster larvae [28]. The method, however, is not yet inducible.
Figure 7.3: Cullin-RING ubiquitin ligases. Protein degradation is often mediated by interactions between F-box containing proteins and their specific substrates. (a) Antibody-driven recruitment. F box-antiGFP fusion protein binds to the GFP-tagged target protein. The F-box protein recruits the bound protein to the E3 ubiquitin ligase. The target protein is polyubiquitinated and degraded by the 26S proteasome. (b) Auxin-mediated recruitment. TIR1 normally targets transcriptional repressors for degradation via the Cullin-RING ubiquitin ligase. This interaction depends on the plant hormone auxin. The auxin-inducible degron (AID) can be fused to a target protein. When auxin is added TIR1 binds the domain and targets the protein of interest for degradation.

One F-box in particular is well-suited for small molecule induction of protein degradation. The F-box protein TIR1 (transport inhibitor response-1) degrades particular plant transcriptional repressor proteins in response to binding the plant hormone auxin [29, 30, 31, 32]. TIR1 only binds its target protein when the auxin is present, but on the other hand, it is not species specific, since it can interact with the E3 ligase protein Skp1 from yeast to humans (Figure 7.3B). Nishimura et al. fused the TIR1 targeting domain of the plant transcriptional repressors to a protein of interest and expressed it in transfected cells. They then expressed the F-box TIR1 in these cells and found that when they applied auxin that the target protein was degraded within an hour [33]. Importantly, auxin hormones are biologically silent in non-plant cells so that they are unlikely to have off-target effects. While this tool has been effectively used in yeast and cell culture from many organisms [34, 35, 36], it remains to be seen if this degron will be useful in more complex organisms.

The methods discussed so far are temporally regulated by either elevation of temperature (for example, tsDHFR is stable at 23 °C, but unstable at 35 °C), or small molecule induction. Most organisms cannot easily tolerate such temperature shifts, or cannot easily take up large concentrations of small molecules. Thus, these methods are largely limited to microorganisms or cell culture. Nor are these methods fast or
limited to specific tissues. Light-inducible methods of protein degradation could increase induction speed and degrade proteins locally.

### 7.3 Light stimulation

Investigators can deactivate proteins with light using photo-reactive organic dyes. This technique, chromophore-assisted light inactivation (CALI), uses dyes such as malachite green or fluorescein (FALI) [37, 38]. Stimulation of the chromophore generates short-lived reactive oxygen species, which can oxidize and inactivate proteins within a limited diameter. More recently, the fluorophore KillerRed provides an example of how light can be used to destroy proteins (Figure 7.4A). KillerRed is a red fluorescent GFP-like protein that releases reactive oxygen species when stimulated by green light. KillerRed can be fused to a protein of interest and then stimulation by light will generate reactive oxygen species, which destroy the protein, with spatial and temporal control [39]. A problem with protein inactivation using reactive oxygen species is that it is not specific and will damage proteins complexed with, or near the target protein. The radius of damage has been estimated to range from 3 nm to 50 nm [40]; to put these values into perspective, GFP has a diameter of 2.5 nm. In addition, KillerRed is cytotoxic and can also be used to induce cell death. Experimenters using KillerRed to target specific proteins should interpret results with caution due to nonspecificity and cytotoxicity. A more precise method for disrupting protein function is needed; light controlled protein-protein interactions may be the solution.

There are a number of strategies for light-induced protein-protein interactions [41]. One could imagine generating a light-inducible degron using light-activated dimerization domains, in this case a gift – once again – from plants. The cryptochrome 2 (CRY2) protein from *Arabidopsis* dimerizes to the CIB1 protein when exposed to blue light [42] and this interaction occurs rapidly (within 1 second) even in mammalian cells [43]. One could fuse CRY2 to an F-box protein and fuse CIB1 to the target protein. Stimulation with blue light would cause these proteins to dimerize, and the target protein would be polyubiquitinated and degraded (Figure 7.4B). This technique may require optimization of the proteins in eukaryotes living at temperatures below 30 °C. At these temperatures, the CRY2 and CIB1 irreversibly dimerize after light stimulation [44]. However, any optimization may be worth the effort; the sub-second dimerization time of CRY2–CIB1 is orders of magnitude faster than heat shock or small molecule methods of induction.

A novel method to block protein function by light stimulation uses the conformational changes in a fluorescent protein to activate or inactivate the protein [45]. Specifically, Michael Lin’s lab used the dimerization properties of the fluorescent protein Dronpa. Dronpa is fluorescently active as a dimer. Intense stimulation with 500 nm light causes Dronpa to dissociates into monomers that are in a dark state, that is they are no longer fluorescently active. The monomers will redimerize when stimulated
by 400 nm light. Lin and coworkers fused Dronpa monomers to each terminus of the CDC42 GEF domain of intersectin (Figure 7.4C). These monomers could form a dimer by bringing the N- and C- termini together. This ‘caged’ conformation of the protein is inactive. Stimulation using 500 nm light stimulation caused the Dronpa dimer dissociate, and thereby uncaged the GEF. (d) Light-activated protease. Dronpa monomers were fused to both termini of the Hepatitis C Virus protease. The wild-type Dronpa monomers form tetramers that occlude the activity of the protease. 500 nm light stimulation causes monomerization and uncages the protease, which can now cleave the target sequence. In this imaginary example the protease cleavage sequence is inserted into a protein and cleavage generates an inactive protein. (e) Light-activated F-box-antibody degron. Dronpa monomers, fused to an antiGFP F-box, dimerize and cage the antibody recognition domain. When 500 nm light is applied, the Dronpa dimer falls apart into monomers, exposing the antibody binding surface. antiGFP then recruits the GFP-tagged target protein to the E3 ligase. In this specific example, the antiGFP antibody would need to be specific for the jellyfish-derived GFP, and not crossreact with the coral-derived Dronpa. The target protein is then polyubiquitinated and degraded by the 26S proteasome.

Lin and colleagues generalized the process by using Dronpa dimerization to inactivate a protease (Figure 7.4D) [45]. They used the hepatitis C protease since it exhibits no toxicity in mammalian cells. The protease target sequence was placed between a membrane tether and mCherry. Stimulation with 490 nm light switched Dronpa fluorescence off and caused the fluorescent proteins to adopt a monomeric configuration and activated the protease. Cleavage was monitored by measuring cytosolic mCherry, which peaked 60 minutes after activation. This approach was not designed to produce a protein knockout but rather to demonstrate that a protease could be activated. However, one could easily imagine combining this approach with that described above for the TEV protease. In short, the TEV protease target sequence, or the hepatitis C target sequence, could be engineered into a protein, such that cleavage would yield two nonfunctional targets and permanently inactivate the protein. On
the other hand, this method would share some of the same disadvantages described above; specifically, cleavage would not mimic the loss of a protein but rather generate dominant negative protein fragments.

Alternatively, Dronpa dimerization could be combined with degron technology. For example, the anti-GFP F-box (see Figure 7.4E) can be fused to Dronpa monomers on either terminus. In the dark state, Dronpa dimers could either shield the anti-GFP antibody from recognizing its epitope, or Dronpa dimerization may prevent the F-box domain from binding the E3 ligase complex. Application of 500 nm light would dissociate Dronpa into monomers and restore an ‘open’ conformation of the anti-GFP F-box chimera. Anti-GFP could then bind the GFP-tagged substrate and the F-box could recruit the E3 ligase complex and degrade the target protein. Additionally, one could apply 400 nm light to re-dimerize the Dronpa monomers, and inactivate the degron.
and reverse the phenotype. This method is worth developing because it has many advantages over current systems. First, light is the fastest potential method for the induction of changes in protein function. Second, light can be non-invasively applied to many organisms, with perhaps the exception of deep tissue in mammals.

Acute protein knockouts could allow the study of embryonic lethal and pleiotropic genes. Currently, degrons are limited by their speed of induction and applicability across systems. Light induction may be the solution. Optical control may allow rapid and cell-specific degradation of proteins. These methods could bypass some of the inherent weaknesses of the DNA mutations used in genetic analysis and disrupt proteins directly and on a rapid time scale. These methods could lead to a revolutionary tool for perturbation studies in biology.

References


[44] Tucker CL. Personal communication.