

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/104917/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Halliwell, Lisa M., Jathoul, Amit P., Bate, Jack P., Worthy, Harley L., Anderson, James C., Jones, D. Dafydd and Murray, James Augustus Henry 2018. Δ Flucs: brighter photinus pyralis firefly luciferases identified by surveying consecutive single amino acid deletion mutations in a thermostable variant. *Biotechnology and Bioengineering* 115 (1), pp. 50-59. 10.1002/bit.26451

Publishers page: <http://dx.doi.org/10.1002/bit.26451>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



ΔFlucs: brighter *Photinus pyralis* firefly luciferases identified by surveying consecutive single amino acid deletion mutations in a thermostable variant[†]

Lisa M. Halliwell*¹, Amit P. Jathoul*¹, Jack P. Bate¹, Harley L. Worthy¹, James C. Anderson², Dafydd D. Jones¹ and James A. H. Murray¹.

¹School of Biosciences, University of Cardiff, Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX, UK.

²Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK.

**Authors contributed equally.*

E-mail: JathoulA@Cardiff.ac.uk

[†]This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/bit.26451]

Additional Supporting Information may be found in the online version of this article.

**This article is protected by copyright. All rights reserved
Received July 6, 2017; Revision Received September 8, 2017; Accepted September 11, 2017**

This article is protected by copyright. All rights reserved

Abstract

The bright bioluminescence catalysed by *Photinus pyralis* firefly luciferase (Fluc) enables a vast array of life science research such as bioimaging in live animals and sensitive in vitro diagnostics. The effectiveness of such applications is improved using engineered enzymes that to date have been constructed using amino acid substitutions. We describe Δ Flucs: consecutive single amino acid deletion mutants within 6 loop structures of the bright and thermostable x11 Fluc. Deletion mutations are a promising avenue to explore new sequence and functional space and isolate novel mutant phenotypes. However, this method is often overlooked and to date there have been no surveys of the effects of consecutive single amino acid deletions in Fluc. We constructed a large semi-rational Δ Fluc library and isolated significantly brighter enzymes after finding x11 Fluc activity was largely tolerant to deletions. Targeting an ‘omega-loop’ motif (T352-G360) significantly enhanced activity, altered kinetics, reduced K_m for D-luciferin, altered emission colours and altered substrate specificity for redshifted analogue DL-infraluciferin. Experimental and in silico analyses suggested remodelling of the Ω -loop impacts on active site hydrophobicity to increase light yields. This work demonstrates the further potential of deletion mutations, which can generate useful Fluc mutants and broaden the palette of the biomedical and biotechnological bioluminescence enzyme toolbox. This article is protected by copyright. All rights reserved

Introduction

Beetle luciferases, such as the bright *Photinus pyralis* (*Ppy*) firefly luciferase (Fluc) catalyse a 2-step reaction of D-luciferin (D-LH₂) with adenosine triphosphate (Mg²⁺.ATP) and oxygen resulting in bright bioluminescence color tunable from green to red with native substrate D-LH₂ (White et al., 1971; Viviani et al., 2001). Fluc is a useful tool in a wide range of applications, for example, biomedical imaging small animal models (Zinn et al., 2008) including of cancer, stem cells and gene therapies, and sensitive fieldable molecular diagnostics (Gandelman et al., 2010). Extensive protein engineering of Fluc has been performed to study the structure-function relationship and improve activities (Jathoul et al., 2012) and/ or quantum yields (Branchini et al., 2014), catalytic efficiencies, emission kinetics, alter emission colors (Branchini et al., 2010) and improve stabilities, including spectral stability (Tisi et al., 2002), pH tolerance and resistance to thermal inactivation (Law et al., 2006; Branchini et al., 2007; Hall et al., 1999). Enhanced engineered luciferases lead to new or improved applications (Baggett et al., 2004) with luciferin or analogues (Reddy et al., 2010; Kuchimaru et al., 2016), such as promising redshifted dual color bioimaging analogue infraluciferin (iLH₂) (Jathoul et al., 2014; Anderson et al., 2017). By combining beneficial thermostabilising and activity improving changes from different studies in *Ppy* Fluc (Tisi et al., 2002b; Law et al., 2006; White et al., 1996; Prebble et al., 2001; Willey et al., 2001) one the most thermostable *Ppy* derivatives was previously developed and is being used for bioluminescence imaging (BLI): x11 Fluc (Jathoul et al., 2012) (F14R, L35Q, A105V, V182K, T214C, I232K, D234G, E354R, D357Y, S420T, F465R) is stable up to c.55°C and has c.50% K_{cat} of WT *Ppy* Fluc. Since BLI in small animals is carried out at 37°C the high resistance of x11 Fluc to thermal inactivation can be sacrificed for enhanced activity as long as other properties, such as its high spectrally stability and pH tolerance are also preserved.

Accepted Preprint

To date, Fluc engineering has been largely achieved by substitution and insertional mutagenesis (Tafreshi et al., 2007) or domain replacements. Insertion/deletion (Indel) mutations of single amino acids occur in proteins in nature, but less commonly than substitutions (Taylor et al., 2004). Indels however alter the structure of the protein backbone, resulting in both short and long-range structural changes and the formation of new conformational space that cannot be sampled by substitutional mutagenesis (Shortle and Sondek, 1995). It is harder to model the effects of Indels compared to side-chain substitutions but the sampling space for deletions is much more limited (to n , where n is the number of residues in the target protein) than natural amino acid substitutions (20^n) or insertions ($20^{(n-1)}$). Thus, in surveying single consecutive deletions one can fully explore their impact in target regions and improve our fundamental understanding about the impact of deletions on structure-function relationships of Fluc. Deletions are generally thought to compromise the structural integrity of proteins and have until now been used to study function but not to improve the practical usefulness of Fluc by altering its various properties. Deletions of conserved functionally important residues, those closer to the active site, and of those disrupting buried hydrophobic core packing are generally deleterious (Simm et al., 2007). Protein termini and loops are more tolerant to single amino acid deletions than secondary structures such as α -helices and β -sheets, though helical and strand termini deletions are often tolerated and can impact positively on function (Jones, 2005). We hoped surveying loop deletions in x11 Fluc would lead to the isolation of novel phenotypes, as demonstrated with other proteins: a deletion in TEM-1 β -lactamase enhances its activity 64-fold towards ceftazidime as a substrate (Simm et al., 2007). Six amino acids deletions in *Staphylococcal* nuclease promote a more compacted state and increase catalytic activity and stability (Baldisseri et al., 1991). Random single amino acid deletion sampling in enhanced green fluorescent protein (EGFP) has shown deletions can be tolerated throughout the β -barrel, particularly in loops, leading to the identification of a better expressed and rapid folding EGFP

Accepted Preprint

deletion mutant, with higher in vivo fluorescence due deletion $\Delta G4$ within an α -helix. This caused a helical registry shift and the formation of a new polar interaction network that putatively stabilised a beneficial cis proline peptide bond (Arpino et al., 2014; Arpino et al., 2014b). Using the stable x11 Fluc scaffold, we hypothesised that we might observe potentially beneficially deletions which may also be destabilising and thus not ordinarily observed in less stable Flucs, but on to which further functions could be engineered (Tokuriki et al., 2008).

Results and Discussion

i. Rational construction of x11 Fluc deletion mutant (Δ Fluc) library and deletion tolerance

Loops are a rich source of functional and structural diversity and are likely to accommodate backbone deletions without considerably impacting on more ordered secondary structures. Therefore, we focused our survey on loop targets for a higher chance of isolating functional mutants with new properties. Since the crystal structure of x11 Fluc had not been solved, we constructed a homology model of x11 Fluc using wild-type (WT) *Photinus pyralis* (*Ppy*) luciferase (4G36.pdb) as a template and relaxed it in solvent to help us select loops. The model contained a few differences to the WT Fluc crystal structure including alteration of a β -sheet (A361-V365) so that turns/ loops are now predicted between A361-G363 (Supplementary Information Figs. S1. a. and b.) and these changes were in agreement with program YASPIN (Lin et al., 2005). The final 6 loop targets were (1-6) M1-G10, L172-T191, T352-F369, D375-R387, D520-L526 and K543-L550 (Fig. 1.), all of which are solvent exposed. Of 43 x11 Fluc deletion mutants (Δ Flucs) obtained from the library, 41 (95%) retained bioluminescent activity, compared to 60% of loop deletions sampled in EGFP by Arpino et al. Our loop selection criteria and considerations (Supplementary Information Table 1) were as follows: (i). Target disordered regions which could improve enzyme stability or folding and thus activity and/ or

expression (Oldfield et al., 2005): PrDOS indicated that Met1-Ala4 of the N-terminus and the final 15 residues (Ile535 –Leu550) of the C-terminus were relatively disordered. Seven amino acid deletions of the N-terminus (Sung and Kang, 1998; Wang et al., 2002) and 8-12 deletions of the C-terminus largely inactivate the enzyme, but the effect of sequential single amino acid deletions have not been assessed, so the N- (including small beta-sheet K8 and K9) and C-termini loops were selected as targets (1 and 6, respectively) and deletions in both were well tolerated. N-terminal deletions had little effect on activity, while C-terminal deletions caused a moderate reduction in activity (Fig. 1). Loop target 2 was located at the top of the N-domain and contains 6 residues (P173-Y179) with a high predicted probability of disorder, and high B-factors associated with residues P174 and G175 in *Ppy* Fluc structure 4G36.pdb. However, deletions in this region markedly reduced activity. Deletions in localised areas of enzymes can have distinct effects (Simm et al., 2007) but this loop is relatively far from the active site, lowly conserved (Supplementary Information Fig. S2), and has a high flexibility. Perhaps the dynamics in the region is important for Fluc function rather than its absolute sequence identity.

(ii). We also targeted loops both distal and proximal (within 5-10Å) (Kaiser, 1988) to the active site: six loops are within 5Å of the active site of Fluc, (Supplementary Information Table 2) the longest of which is an omega (Ω)-loop motif (T352-G360) at the junction of subdomains B and C, in which substitutions (White et al., 1996; Willey et al., 2001) and insertions (Tafreshi et al., 2007; Moradi et al., 2009) give rise to a wide range of phenotypes including improved resistance to thermal inactivation (contains x11 Fluc mutations E354R and D356Y) and altered emission colors (e.g. E354I and D357Y redshift emission) (Willey et al., 2001). Ω -loops are extended turns that resemble an Ω shape and play roles in molecular recognition and substrate specificity of enzymes (Fetrow et al., 1995; Guntas et al., 2012; Stojanoski et al., 2015). The Ω -loop of Fluc has low B factors is a part of a larger solvent exposed structure including a solvent-exposed beta-sheet (A361-V365) in *Ppy* Fluc, of which residue V362 side-chain is

within 5Å of the active site. Since deletions are often tolerated at ends of beta-sheets (Arpino et al., 2014b), our model predicted this entire region as a large loop, and given its functional importance, we tested the effect of deletions both in the Ω-loop and the region beyond it (T352-F369 - target 3) and observed significant enhancements of in vivo activity with ΔP359 and ΔG363. This region was relatively tolerant to deletions in terms of activity and color with the exception of ΔT352, which was redshifted and displayed relatively low activity. An interesting a hypsochromic shift of emission color from orange to yellow-green was observed over time in colonies expressing x11 Fluc ΔP353 (Supplementary Information Fig. S3). With the exception of loop L172-T191 in subdomain A, deletions in loops distal to the active site affected activity less than those relatively proximal to it, which is similar to the results reported of the effect of deletions more proximal to the chromophore of EGFP, which abrogate fluorescence or reduce maturation rates. The 4th loop target (D520-L526) was within the active site and deletions in this region markedly reduced activity. (iii). We also targeted less conserved loops in beetle luciferases: with the exception of loops S82-Q87, S198-P211, V240-H245, G316-S320, D375-R387, L411-W417, K445-G446, H461-D466, P473-E479, S504-G515 and V519-D531 (11 out of a total of 42 loops or turns in x11 Fluc) loops are lowly conserved, particularly in the N-domain. A relatively conserved loop target 5 (D375-R387) was thus chosen to balance the level of conservation between targets. However, there was no apparent correlation of deletion tolerance to the level of conservation, which is in contrast to the situation with substitutions.

The emission color of x11 Fluc ($\lambda_{\max} = 555\text{nm}$) was remarkably unaltered by the majority of deletions since only 5 mutants had red-shifted λ_{\max} of bioluminescence spectra (ΔT352: 605nm, ΔG363: 571nm, ΔD520: 571nm, ΔE521: 572nm and ΔP523: 571nm). One double deletion mutant (ΔF176 + ΔN177) was obtained and one deletion with an additional

substitution (Δ P183 + E184D). Whereas the double deletion mutant displayed bioluminescence activity ($\lambda_{\text{max}} = 555\text{nm}$), the deletion and substitution was a knock-out. Only 2 single Δ Flucs (Δ V384 and Δ R387) exhibited a near-complete loss of activity, demonstrating that loop regions of x11 Fluc are largely tolerant to single amino acid deletions in loops. Ordinarily deletions in proteins of residues with low residue solvent accessibility (RSA) such as Gly and Pro are better tolerated than larger residues, or those that make fewer contacts (Jackson et al., 2017). Aside from in the N-terminal, deletions of Pro (P353 and P359), Gly (G360 and G363), Ala (A361), Val (V362), Leu (L376) and Asp (D377) were best tolerated or enhanced brightness in vivo. Structural predictors of deletion tolerance residue solvent accessibility (RSA) and weighted contact number (WCN) (Fig. 2.a.) were calculated for each residue targeted (Supplementary Information Table 1). A slight negative correlation was seen between deletion target RSA value and in vivo activity of mutants (Spearman's $r = -0.099$), and a positive correlation in the case of WCN (Spearman's $r = 0.255$ and 0.207 for carbon atom and side-chain, respectively) (Supplementary Information Fig. S4). Deletion of residues with lower WCN are predicted to be better tolerated, yet aside from small areas such as loop D375-R387 (Spearman's r values of loop D375-R387 structural predictor correlations to activity: RSA = 0.406, WCN carbon atom = -0.483 and WCN side-chain = -0.003), overall we observed a positive correlation of the WCN of a deleted residue and in vivo activity.

ii. Identification of novel properties from the deletion library

Assaying library cell lysates in a 96-well plate format revealed a 'fingerprint' pattern of loop deletion tolerance of x11 Fluc with which we tested the effect of conditions (Figs. 2.a. and b.). Mutant x2 Fluc, x11 Fluc Δ P353 and a few other mutants (Δ P173, Δ D187, Δ T191 and Δ D356) displayed higher activity at lower luciferin concentrations, however, since the K_m of pure x2

Fluc for D-LH₂ is higher than x11 Fluc (Jathoul et al., 2012) we did not think this to be indicative of lower K_m values for D-LH₂. The library was assayed in the presence of inorganic pyrophosphate (PPi), which can act both as an inhibitor and activator (e.g. in the presence of inorganic pyrophosphatase - PPiase) of the Fluc reaction (Fontes et al., 2008). We observed inhibition of most Flucs by 200μM PPi in lysates and x11 Fluc was slightly more affected than WT or x2 Flucs. Since x11 Fluc ΔP359 and ΔG363 are brighter than other enzymes, they displayed similar activity in the presence of PPi as other enzymes did in the absence of PPi, however, they were also inhibited by PPi similarly to WT Fluc. In contrast, x11 Fluc ΔL376 and ΔD377 displayed improvements in activity in the presence of 200μM PPi, suggesting this region may be involved in the inhibition effect. Next, the library was assayed for the effect of thermal inactivation on activity. After incubation at 42°C for 15min WT Fluc was nearly inactivated, whereas x2, x11 and most x11 ΔFlucs had higher activity. Surprisingly, the majority of the deletion mutations did not compromise the activity of x11 Fluc at 42°C since most had similar levels of activity as x11 Fluc. ΔFlucs ΔP359, and ΔG363 displayed markedly higher activities (more than double) than parental x11 Fluc at 42°C and ΔA361, ΔV362, ΔL376, ΔD377 and ΔG379 also showed similar, but smaller enhancements. After incubation at 60°C WT and x2 Flucs were almost inactivated, however, x11 Fluc displayed 74% activity. ΔFlucs ΔK8, Δ187, ΔP353, ΔG355, ΔD356, ΔK358, ΔP359, ΔG360, ΔV362, ΔG363 and ΔV365 all had reduced activity after 60°C incubation. Mutants ΔK5 and ΔD377 has similar activities at RT and after incubation at 60°C, however, ΔL376 appeared to have higher activity after incubation at 60°C than at RT. Overall, our approach of sequential deletion mutagenesis has led to the observation of a variety of effects which appear to be region specific and deletion mutagenesis shows promise in its potential for the engineering of x11 Fluc properties and obtain novel mutant phenotypes. A distinct effect on substrate specificity was also seen with

DL-iLH₂ (Supplementary Fig. S5). Δ P353 had consistently lower activity than x11 Fluc with the DL-iLH₂, whereas Δ P359 had enhanced activity. Therefore, as seen with TEM β -lactamase, Ω -loop deletions can affect Fluc substrate specificity. The bright Ω -loop mutants could be advantageous for in vivo reporters in combination with DL-iLH₂.

iii. Purification and characterisation of Ω -loop region Δ Flucs

The Ω -loop mutants (Fig. 3. a.) were purified and all mutants aside from Δ P353 displayed higher specific activities than x11 Fluc, with Δ P359 and Δ G363 giving 234% and 188% of x11 Fluc, respectively. These increases were in-line with the levels of activity observed in vivo at 42°C, showing that the deletions directly improve the activity of x11 Fluc and give a higher integrated light yield per unit protein. In particular, Δ G363 increased in its bioluminescence activity following initial flash decay (Fig. 3.b.). Hydrogen bonds from the Ω -loop to adjacent residues (e.g. in β -sheet region H431-F433) are thought to stabilise the active site near the 6'-hydroxyl group of luciferin and remodelling of the Ω -loop impacts on the polarity of emitter site (Tafreshi et al., 2007). Backbone deletions clearly have similar effects, and mutants might be brighter due to a remodelling effect which increases the active site hydrophobicity.

Michaelis-Menten kinetic parameters were determined for D-LH₂ and ATP (Table 1). x11 Fluc Δ P359 (2 μ M) and Δ G360 (3.3 μ M) had slightly lower K_ms for D-LH₂, suggesting that the conformation of the Ω -loop is important for the affinity for D-LH₂. In contrast, mutants Δ P353, Δ A361, Δ V362 and Δ G363 had similar or slightly higher K_ms than x11 Fluc (6 μ M - 11 μ M). The K_{cat}/K_m of Δ P359 of D-LH₂ was nearly doubled compared to x11 Fluc, a further potential advantage for in vivo imaging. The K_m for ATP of x11 Fluc was 76 μ M, which was similar to mutants Δ P353 and Δ G360 (67 μ M and 58 μ M, respectively). The K_m for ATP of Δ P359 was slightly higher (100 μ M) and the other mutants Δ A361 (117 μ M), Δ V362 (150 μ M), Δ G363

(188 μ M) displayed increasing K_m s apparently positively correlated with the distance of the deletion from the N-terminal end of the Ω -loop. Increases in the K_m for ATP by deletions in the Ω -loop were more frequent than for D-LH₂.

Next the pH dependence of bioluminescence spectra and activity of purified mutants were compared to x11 Fluc (Fig. 4. i.-iii. and Supplementary Information Fig. S6 i.-iv.). Between pH values of 6.3-8.8, the Ω -loop variants Δ P353, Δ P359, Δ G360, Δ A361 had very similar λ_{max} and spectral shapes to x11 Fluc (555nm) showing that deletions had little effect on the high spectral stability of x11 Fluc. At pH 8.8 x11 Fluc Δ G363 displayed an increase in bandwidth toward the red, indicating that the deletion has an effect on the structure of the emitter in alkali conditions, which is interesting since bathochromic shift of WT Fluc is normally observed under acidic conditions (McElroy and Seliger, 1961). The color of x11 Fluc Δ P353 was green, independent of pH or time, thus the hypsochromic shift effect was only observed in vivo. This mutant had lower activity at acidic pH compared to x11 Fluc and conversely mutant Δ G360 was less active at higher pH. Clearly the identity or conformation of this loop region is important for the properties of pH dependence of activity and color. When the two brighter Δ Flucs (particularly Δ G363) were assayed after incubation on ice or room temperature for different periods they displayed significant variation in activity, suggesting an effect on stability. Since these two mutants could make promising reporters for in vivo bioluminescence imaging we assessed their resistances to thermal inactivation. WT Fluc activity was completely inactivated at 40°C, whereas x11 Fluc was inactivated at 60°C (Fig. 5.a.). Δ Flucs Δ P353 and Δ P359 are significantly brighter than x11 Fluc between c.30-40°C, however, they are inactivated at 50°C and thus are significantly less resistant to thermal inactivation than x11 Fluc. This shows that the Ω -loop conformation is also important for resistance to thermal inactivation and the effect of thermostabilising mutations E354R/ D357Y in x11 Fluc is likely disrupted by deletions. Mutant Δ G363 proved to be very unstable and was

also completely inactivated at 50°C. However, mutants $\Delta P353$ and $\Delta P359$ could be advantageous for bioimaging and it is interesting that so many properties of x11 Fluc are modulated by these two opposing prolines positioned either side of enclosing turns of the Ω -motif of Fluc. Deletion of P353 deletes two intramolecular H-bonds, one between G355 and the side chain hydroxyl of Y357, and another between D356 and P353, and, $\Delta P359$ deletes a beta turn (Fig. 3.a.). We used circular dichroism (CD) spectroscopy to examine if Ω -loop deletions have an effect on the secondary structure of x11 Fluc (Fig. 5.b.) (Atae and Hossienkhani, 2015), although we expected the effect to be minimal. Indeed, the Ω -loop Pro deletion mutants showed little difference compared to x11 Fluc. To see if deletions effected the hydrophobicity of the active site which could explain the increased brightness observed, fluorescence spectra were acquired with 8-anilinonaphthalene-1-sulfonic acid (ANS). This showed that the two brighter enzymes $\Delta P359$ and $\Delta G363$ had more hydrophobic active sites than x11 Fluc, giving more intense and blueshifted ANS spectra, which could explain their relative brightness.

iv. Conclusions

Our survey of sequential single amino acid deletions in a solvent-exposed loops of x11 Fluc has revealed Fluc is remarkably tolerant to loop deletions and that they can invoke a variety of interesting phenotypes particularly in regions proximal to the active site, like the Ω loop, which has high functional plasticity. Ω -loop Δ Flucs markedly enhanced the brightness enzymes, altered K_m s for both D-LH₂ and ATP, emission kinetics, pH dependence of activity and color, and, resistances to thermal inactivation. Mutants could be of benefit to bioimaging with iLH₂ at physiological pH and temperature in mammals. Other potentially thermostable mutants ($\Delta L376$ and $\Delta D377$) warrant further investigation for their potential for application in

isothermal LAMP-BART⁴ nucleic acid diagnostics. These new Δ Fluc mutants also provide valuable sequence landscapes on to which to engineer further properties.

Materials and Methods

Plasmids, bacterial strains, reagents and buffers. pET16b (EMD Millipore, Watford, UK) plasmids with WT and x11 Fluc genes and x2 Fluc in pDEST17 (Thermo Fisher Scientific, MA, USA) were previously constructed (White et al., 1996; Jathoul et al., 2012). x2 Fluc was amplified from pDEST17 and subcloned into pET16b. D-LH₂ potassium salt was purchased from Europa Bioproducts, Ely, Cambridgeshire, UK and ATP from Roche Diagnostics, IN, USA. 'TEM' buffer contains 100mM Tris-acetate, 2mM ethylenediaminetetraacetic acid (EDTA) and 10mM magnesium sulphate, pH 7.8.

Construction of single amino acid loop deletion library of x11 Fluc. Phusion High Fidelity DNA polymerase (1 U) (NEB, MA, USA) was used to amplify DNA according to the Quikchange (Agilent Technologies, CA, USA) protocol to delete residues in x11 Fluc. Primers are listed in Supplementary Information Table 3. DpnI digested DNA was transformed into high efficiency DH5 α . The next day DNA prepared from bulk transformants was transformed into *E. coli* B121 (DE3) pLysS (Agilent Technologies, CA, USA).

Colony and 96-well format lysate screening of the mutant library. Colony screening was carried out as previously described (Wood and Deluca, 1987). For plate screening, single colonies of Δ Flucs were used to inoculate 5ml LB broth with 100 μ g/ml carbenicillin and grown overnight at 37°C. Optical density normalised cultures were induced with IPTG (1mM) for 3

hours at RT and 100 μ l pipetted into a 96-well plates, which were centrifuged at 4500rpm for 20min. Any remaining LB broth was removed by aspiration and pellets were resuspended in 0.1% Triton-X100 in TEM buffer with 10% glycerol. Assays were carried out in a Fluoroskan Ascent luminometer (Thermo Scientific, MA, USA).

Over-expression, purification of mutants, spectroscopic and luminometric methods. Cloning, over-expression, purification by nickel-nitrilotriacetic acid (NTA) resin affinity chromatography and desalting of enzymes was as previously described (Law et al., 2006). *i. Measurement of specific activities, kinetics and bioluminescence spectra.* Specific activities of the enzymes were measured with 500 μ M LH₂ and 1mM ATP. 50 μ l of substrates were injected onto 50 μ l of 0.5 μ M Fluc and light integrated over 250s. Michaelis-Menten kinetic parameters were determined using the flash height method and derived using the Hanes-Woolf plot (Cornish-Bowden, 1999): 50 μ l varying concentrations of substrate were injected onto 0.5 μ M Fluc with 50 μ l of the remaining constituent. Bioluminescent spectra at different pH were acquired in a Varian Cary Eclipse fluorimeter (Agilent Technologies, CA, USA) and corrected for varying photomultiplier tube spectral sensitivity using lucifer yellow-CH (Sigma-Aldrich, MO, USA). *ii. Resistance to thermal inactivation.* 500 μ l aliquots of 0.5 μ M Luc solutions (pH 7.8) were incubated at different temperatures and 50 μ l aliquots were placed on ice before assaying with 50 μ l of 500 μ M D-LH₂ and 1mM ATP in a BMG Fluorstar injecting luminometer (BMG labtech, Offenburg, Germany). *Extrinsic 8-anilinonaphthalene-1-sulfonic acid (ANS) fluorescence spectroscopy.* Fluorescence spectra of ANS were measured in the absence or presence of Fluc proteins to probe the active site. 5 μ l of 10 μ M ANS was added to 50 μ l 0.5 μ M enzymes and fluorescence emission spectra were recorded between 400 and 600nm with an excitation wavelength of 350nm. *CD spectroscopy.* Far-UV circular dichroism (CD) spectra of

Fluc proteins were acquired using 4 μ M proteins analysed at 25°C in Applied Photophysics Chirscan (Surrey, UK) with a Quantum Northwest (WA, USA) temperature controller.

In silico analyses and modelling. A x11 Fluc homology model was prepared using Swiss Model (Bordoli et al., 2009) and deletion mutant models were constructed via manual removal of the individual residues with the PyMOL (Version 1.8 Schrödinger, LLC). All 4 models (x11 Fluc, Δ P353, Δ P359 and Δ G363) were subjected to an automated processing script using Gromacs (Berendsen et al., 1995) and Maya Chem tools. This involved resetting the amino acid residue numbering to start at 1, filling in missing atoms of residues, while ensuring correct hydrogen placement. For energy minimisation and molecular dynamics the models were charge balanced to 0 by adding Na and Cl ions and to mimic in vitro conditions models were solvated with water ions to fill a virtual box around the protein. The protein was then relaxed to an energy minimum using Gromacs via a steepest descent method. Molecular dynamics was run over 1 ns under constant simulation temperature and pressure allowing the model protein's structure to relax. Models were viewed and manipulated on PyMOL. *Residue solvent accessibility and weighted contact number determination.* For residue solvent accessibility (RSA) calculations, the x11 Fluc model was submitted to the CMBI XSSP server to generate DSSP values. These were then inputted into a python script calc_RSA.py to generate a list of RSA values. Weighted contact numbers (WCN) calculated as described (Jackson et al., 2017) using script calc_WCN.py. Scripts are made free to use on GitHub page.

vi. Acknowledgements.

Dr. Halliwell was funded by a KESSII PhD studentship sponsored by 3M (Bridgend, Wales, UK); Dr. Jathoul was part funded by the Cardiff synthetic biology initiative (CBSI) and part

by a Ser Cymru II Fellowship. We acknowledge Dr. Robert Mart (Department of Chemistry, Cardiff University) for assistance measuring CD spectra. Harley Worthy was funded by a BBSRC-facing Cardiff University Studentship. DDJ would like to thank BBSRC (BB/H003746/1, BB/M000249/1 and BB/FOF/263), EPSRC (EP/J015318/1) and Cardiff SynBio Initiative/SynBioCite for supporting this work. JAHM would like to thank the BBSRC (BB/L022346/1). The authors declare no conflict of interest.

References

Anderson JC, Grounds H, Jathoul AP, Murray JAH, Pacman SJ and Tisi LC. 2017. Convergent synthesis and optical properties of near-infrared emitting bioluminescent infra-luciferins. *RSC Adv.* 7: 3975-82.

Arpino JA, Reddington SC, Halliwell LM, Rizkallah PJ, Jones DD. 2014. Random single amino acid deletion sampling unveils structural tolerance and the benefits of helical registry shift on GFP folding and structure. *Structure* 22:889-98.

Arpino, J.A.J., Rizkallah, P.J. and Jones, D.D. 2014. Structural and dynamic changes associated with beneficial engineered single-amino-acid deletion mutations in enhanced green fluorescent protein. *Acta Crystallographica Section D* 70, 2152-62.

Ataei F and Hosseinkhani S. 2015. Impact of trifluoroethanol-induced structural changes on luciferase cleavage sites. *J. Photochem. Photobiol. B: Biol.* 144: 1-7.

Baggett B, Roy R, Momen S, Morgan S, Tisi L, Morse D and Gilles RJ. 2004. Thermostability of firefly luciferases affects efficiency of detection by in vivo bioluminescence. *Mol. Imaging* 3: 324-32.

Baldisseri D, Torchia D, Poole L and Gerlt J. 1991. Deletion of the Omega-Loop in the Active Site of Staphylococcal Nuclease. 2. Effects on Protein Structure and Dynamics. *Biochemistry* 30: 3628-33.

Berendsen HJC, van der Spoel D and van Drunen R. 1995. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* 91: 43-56.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. and Schwede, T. 2009. Protein structure homology modelling using SWISS-MODEL Workspace. *Nat. Protocols* 4: 1.

Branchini BR, Ablamsky DM, Murtiashaw MH, Uzasci L, Fraga H, Southworth TL. 2007. Thermostable red and green light-producing firefly luciferase mutants for bioluminescent reporter applications. *Anal. Biochem.* 361: 253-62.

Branchini BR, Ablamsky DM, Davis AL, Southworth TL, Butler B, Fan F, Jathoul AP, Pule MA. 2010. Red-emitting luciferases for bioluminescence reporter and imaging applications. *Anal. Biochem.* 396: 290-7.

Branchini BR, Southworth TL, Fontaine DM, Davis AL, Behney CE, Murtiashaw MH. 2014. A *Photinus pyralis* and *Luciola italica* chimeric firefly luciferase produces enhanced bioluminescence. *Biochemistry* 53: 6287-9.

Cornish-Bowden A. 1999. *Fundamentals of enzyme kinetics*. 2nd Edition. Portland Press, London. Pgs. 19-39 and 106-108.

da Silva LP and Esteves da Silva JCG. 2011. Study on the Effects of Intermolecular Interactions on Firefly Multicolor Bioluminescence. *Chemphyschem* 12: 3002-3008.

da Silva LP and Esteves da Silva JCG. 2012. Firefly Chemiluminescence and Bioluminescence: Efficient Generation of Excited States. *Chemphyschem* 13: 2257-2262.

Fetrow JS. 1995. Omega loop: non regular secondary structure significant in protein function and stability. *FASEB J.* 9: 708-17.

Fontes R, Fernandes D, Peralta F, Fraga H, Maio I, Esteves da Silva JC. 2008. Pyrophosphate and tripolyphosphate affect firefly luciferase luminescence because they act as substrates and not as allosteric effectors. *FEBS J.* 275: 1500-9.

Gandelman OA, Church VL, Moore CA, Kiddle G, Carne CA, Parmar S, Jalal H, Tisi LC, Murray JA. 2010. Novel bioluminescent quantitative detection of nucleic acid amplification in real-time. *PLoS One* 5: e14155.

Guntas G, Kanwar M, Ostermeier M. 2012. Circular permutation in the Ω -loop of TEM-1 β -lactamase results in improved activity and altered substrate specificity. *PLoS One* 7: e35998.

Hall MP, Gruber MG, Hannah RR, Jennens-Clough ML and Wood KV (1999). Stabilisation of firefly luciferase using directed evolution. In *Biolum. Chemilum.: perspectives for the 21st century*. Editors: Roda A, Pazzagli M, Kricka LJ and Stanley PE. Wiley, Chichester. Pgs. 392-5.

Jackson, E.L., Spielman, S.J. and Wilke, C.O. 2017. Computational prediction of the tolerance to amino-acid deletion in green-fluorescent protein. *PLoS One.* 12: e0164905.

Jathoul A, Law E, Gandelman O, Pule M, Tisi L, Murray J. 2012. Development of a pH-tolerant thermostable *Photinus Pyralis* luciferase for brighter in vivo imaging. In *Biolumin.—Recent Adv. Ocean. Meas. Lab. Appl.* (D. Lapota). InTech: DOI: 10.5772/37170.

Jathoul AP, Grounds H, Anderson JC, Pule MA. 2014. A dual-color far-red to near-infrared firefly luciferin analogue designed for multiparametric bioluminescence imaging. *Angew. Chem.* 53: 13059-63.

Jones DD. 2005. Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 β -lactamase to an amino acid deletion. *Nucleic Acids Res.* 33: e80.

Kaiser ET. 1988. Catalytic Activity of Enzymes Altered at Their Active Sites. *Angew. Chem.* 27: 913-22.

Khalifeh K, Ranjbar B, Alipour BS, Hosseinkhani S. 2011. The effect of surface charge balance on thermodynamic stability and kinetics of refolding of firefly luciferase. *BMB Rep.* 44: 102-6.

Kheirabadi M, Sharafian Z, Naderi-Manesh H, Heineman U, Gohlke U and Hosseinkhani S. 2013. Crystal structure of native and a mutant of *Lampyris turkestanicus* luciferase implicate in bioluminescence color shift. *Biochim. et Biophys. Acta* 1834: 2729-2735.

Kuchimaru T, Iwano S, Kiyama M, Mitsumata S, Kadonosono T, Niwa H, Maki S, Kizaka-Kondoh S. 2016. A luciferin analogue generating near-infrared bioluminescence achieves highly sensitive deep-tissue imaging. *Nat Commun.* 7: 11856.

Law GHE, Gandelman OA, Tisi LC, Lowe CR and Murray JAH. 2006. Mutagenesis of solvent exposed amino acids in *Photinus pyralis* luciferase improves thermostability and pH-tolerance. *Biochem. J.* 397: 305-312.

Marques SM and Esteves da Silva JCG. 2009. Firefly Bioluminescence: A Mechanistic Approach of Luciferase Catalyzed Reactions. *Iubmb Life* 61: 6-17.

Tisi LC, Law EL, Gandelman O, Lowe CR and Murray JAH (2002). The basis of the bathochromic shift in the luciferase from *Photinus pyralis*. In *Biolum. Chemilum.: Progress and Current Applications*. Edited by Stanley PE and Kricka LJ. Pgs. 57-60. World Scientific, Singapore.

Tisi LC, White PJ, Squirrell DJ, Murphy MJ, Lowe CR and Murray JAH (2002b). Development of a thermostable firefly luciferase. *Anal. Chim. Acta* 457: 115-23.

Lin K, Simossis VA, Taylor WR, Heringa J. 2005. A simple and fast secondary structure prediction method using hidden neural networks. *Bioinformatics* 21: 152-9.

McElroy WD and Seliger HH (1961). Mechanisms of bioluminescent reactions. In McElroy WD and Glass B. *Light and Life*. The Johns Hopkins University Press, Baltimore. Pgs 219-257.

Moradi A, Hosseinkhani S, Naderi-Manesh H, Sadeghizadeh M and Alipour B. 2009. Effect of charge distribution in a flexible loop on the bioluminescence colour of firefly luciferases. *Biochemistry* 48: 575-82.

Mortazavi M, Hosseinkhani S. 2011. Design of thermostable luciferases through arginine saturation in solvent-exposed loops. *Protein Eng. Des. Sel.* 2011 24: 893-903.

Mortazavi M, Hosseinkhani S. 2017. Surface charge modification increases firefly luciferase rigidity without alteration in bioluminescence spectra. *Enzyme Microb. Technol.* 96: 47-59.

Oldfield CJ, Ulrich EL, Cheng Y, Dunker AK and Markley JL. 2005. Addressing the intrinsic disorder bottleneck in structural proteomics. *Proteins* 59: 444-53.

Prebble SE, Price RL, Lingard B, Tisi LC and White PJ (2001). Protein engineering and molecular modelling of firefly luciferase. In *Proc. 11th Int. Symp. Biolum. Chemilum.* Editors: Case JF, Herring PJ, Robinson BH, Haddock SHD, Kricka LJ and Stanley PE. World Scientific, Singapore. Pgs. 181-4

Reddy GR, Thompson WC, Miller SC. 2010. Robust light emission from cyclic alkylaminoluciferin substrates for firefly luciferase. *J Am Chem Soc.* 132: 13586-7.

Said Alipour B, Hosseinkhani S, Ardestani SK, Moradi A. 2009. The effective role of positive charge saturation in bioluminescence color and thermostability of firefly luciferase. *Photochem. Photobiol. Sci.* 8: 847-55.

Shortle D and Sondak J. 1995. The Emerging Role of Insertions and Deletions in Protein Engineering. *Curr. Op. Biotech.* 4: 387-393.

Simm A, Baldwin A, Busse K and Jones D. 2007. Investigating the Protein Structural Plasticity by Surveying the Consequence of an Amino Acid Deletion from TEM-1 β -lactamase. *FEBS Lett.* 581: 3904-8.

Stojanoski V, Chow DC, Hu L, Sankaran B, Gilbert HF, Prasad BV, Palzkill T. 2015. A triple mutant in the Ω -loop of TEM-1 β -lactamase changes the substrate profile via a large conformational change and an altered general base for catalysis. *J. Biol. Chem.* 290: 10382-94.

Sung D and Kang H. 1998. The N-terminal amino acid sequences of the firefly luciferase are important for the stability of the enzyme. *Photochem Photobiol.* 68: 749-53.

Tafreshi NK, Hosseinkhani S, Sadeghizadeh M, Sadeghi M, Ranjbar B and Naderi-Manesh H. 2007. The influence of insertion of a critical residue (Arg356) in structure and bioluminescence spectra of firefly luciferase. *J. Biol. Chem.* 282: 8641-7.

Taylor MS, Ponting CP and Copley RR. 2004. Occurrence and Consequences of Coding Sequence Insertions and Deletions in Mammalian genomes. *Genome Res.* 14, 555–66.

Tokuriki, N., Stricher, F., Serrano, L., Tawfik, D.S. 2008. How protein stability and new functions trade off. *PLoS Comput Biol.* 4: e1000002.

Ugarova NN. 1989. Luciferase of *Luciola mingrelica* fireflies. Kinetics and regulation mechanism. *J. Biolum. Chemilum.* 4: 406-18.

Viviani VR, Uchida A, Suenaga N, Ryufuku M and Ohmiya Y (2001). Thr 226 is a key residue for bioluminescence spectra determination in beetle luciferases. *Biochem. Biophys. Res. Comm.* 280: 1286-91.

Wang XC, Yang J, Huang W, He L, Yu JT, Lin QS, Li W, Zhou HM. 2002. Effects of removal of the N-terminal amino acid residues on the activity and conformation of firefly luciferase. *Int. J. Biochem. Cell Biol.* 34: 982-91.

White PJ, Squirrell DJ, Arnaud P, Lowe CR and Murray JAH (1996). Improved thermostability of the North American firefly luciferase: saturation mutagenesis at position 354. *Biochem. J.* 319: 343-50.

Willey TL, Squirrell DJ and White PJ (2001). Design and selection of firefly luciferases with novel in vivo and in vitro properties. In *Proc. 11th Int. Symp. Biolum. Chemilum.* Editors: Case JF, Herring PJ, Robinson BH, Haddock SHD, Kricka LJ and Stanley PE. World Scientific, Singapore. Pgs 201-4.

Wood KV and DeLuca M (1987). Photographic detection of luminescence in *Escherichia coli* containing the gene for firefly luciferase. *Anal. Biochem.* 161: 501-7.

White EH, Rapaport E, Seliger HH and Hopkins TA (1971). The chemi- and bioluminescence of firefly luciferin: An efficient chemical production of electronically excited states. *Bioorg. Chem.* 1: 92-122.

Zinn KR, Chaudhuri TR, Szafran AA, O'Quinn D, Weaver C, Dugger K, Lamar D, Kesterson RA, Wang X, and Frank SJ. 2008. Noninvasive Bioluminescence Imaging in Small Animals. *ILAR J.* 49(1): 103–115.

List of figures and table captions

Figure 1. Location and effect of deletions in x11 Fluc. a. Protein sequence of x11 Fluc with thermostabilising mutations (blue boxes: F14R, L35Q, A105V, V182K, T214C, I232K, D234G, D357Y, E354R, S420T, F465R), secondary structure elements (from a homology model of x11 Fluc constructed using Swiss modeller and relaxed in solvent using Gromacs), regions of conservation (multiple alignment with clustal omega: beetle species and luciferase accession numbers in Supplementary Information Fig. S2.), common insertions (red arrows), deletions (red boxes), gold box: deletion in *Lampyrus noctiluca* (AAR207941), purple tube: α -helix, light purple tube: 3-10 helix, yellow arrow: β -sheet, grey box: loop, cyan box: turn, dark blue box: bridge, and, room temperature in vivo activity results (Fig. 2) shown by coloured hatch markings. Target loops are underlined (M1-G10, L172-T191, T352-F369, D375-R387, D520-L526 and K543-L550). b. 3D model of x11 Fluc showing mutations (pink balls), 3 visible target loops (blue) and 5'-O-[(N-dehydroluciferyl)-sulphamoyl]-adenosine (DLSA).

Figure 2. 'Fingerprint' pattern of x11 Fluc loop deletion mutant activity. a. *E. coli* BL21 expressing WT, x2, x11 Flucs and deletion mutant cell densities normalised by OD₆₀₀ and induced for 6 hours at RT with 1mM IPTG. Bioluminescence activity measured for 100ms after 5 min settle time after injection of 500 μ M D-LH₂ and 1mM ATP in a Fluoroskan Ascent luminometer (Thermo Fisher Scientific, MA, USA). Low luciferin concentration measured with 20 μ M D-LH₂ and 1mM ATP; Plus PPi measured with 200 μ M PPi; thermal inactivation measured after incubation at 42°C and 60°C for 15 min. b. Inset picture: light emission from colonies expressing x11 Fluc Ω -loop region mutants taken over 30s using a Nikon D700 camera. Overlaid lines indicate the RSA and side-chain WCN for each residue. Errors are standard error of the mean (SEM).

Figure 3. Molecular models of Ω -loop (T352-G360), emission kinetics of x11 Fluc and mutants $\Delta P353$, $\Delta P359$ and $\Delta G363$ with D-LH₂. a. Graphics of Ω -loop in models of x11 (i) and Δ Flucs $\Delta P353$, $\Delta P359$ and $\Delta G363$ (ii-iv, respectively). Structures were constructed by homology modelling and molecular dynamics relaxation of structures using Swiss-model and Gromacs, respectively. Displayed using Pymol¹. Beta-turn in which all C α within 7Å (i: K358, i + 1: P359, i + 2: G360 and i + 3: A361) shown by dotted line. b. Light emission was captured in the BMG luminometer after injection of saturating D-LH₂ and ATP onto 0.5 μ M enzymes kept on ice and assayed immediately.

Table 1. Mutant and control specific activities and kinetic parameters. i. Proteins kept on ice until assay by injecting saturating D-LH₂ and ATP onto enzymes diluted in cold TEM buffer and integrating bioluminescence between 30s-250s using a BMG Fluostar Optima (BMG Labtech, Ortenberg, Germany). ii. Flash-height derived (Ugarova, 1989) K_{cat} values: since the initial flash heights of enzymes were similar to x11 Fluc the K_{cat} values were similar.

Figure 4. pH dependence of colour and activity. pH dependence of bioluminescence spectra of x11 Fluc, $\Delta P359$ and $\Delta G363$ (i. – iii.) acquired using a Varian Cary Eclipse fluorimeter after incubating saturating conditions of D-LH₂ and ATP with enzymes in TEM buffer at different pH for a period of 1 min. iv. pH dependence of activity. These values were determined by integrating the area under bioluminescence spectra acquired at different pH values. Errors are standard errors of the mean (SEM).

Figure 5. Resistance to thermal inactivation, secondary structure effect and active site hydrophobicity. a. Resistance to thermal inactivation: 50 μ l 200 μ M D-LH₂ was added to 50 μ l 0.5 μ M Fluc after incubation at temperatures. 50 μ l 6mM ATP then injected to initiate light emission, captured using BMG Fluostar (PMT gain 2500v). Very small error bars are SEM. b. Average CD spectra of 4 μ M proteins acquired between 200-400nm at 25°C in triplicate as detailed in the Materials and Methods. The blue coloured spectrum of Δ P353 underlies x11 Fluc spectrum. c. Fluorescence emission spectra of ANS +/- Flucs to probe the active site. 5 μ l 1.95mM ANS was added to 50 μ l 3.9 μ M enzymes and fluorescence emission spectra were recorded between 380 and 700nm in a Tecan Infinite M200 fluorometer (Männedorf, Switzerland). D. Normalised fluorescence spectra of mutants with ANS.

Enzyme	Specific activity ⁱ (RLU/mg) (x 10 ⁷)	Percentage of WT Fluc (%)	D-luciferin			ATP		
			K _m (μM)	K _{cat} (RLU/s) (x 10 ¹⁶) ⁱⁱ	K _{cat} /K _m (RLU/s/μM) (x 10 ¹⁵)	K _m (μM)	K _{cat} (RLU/s) (x 10 ¹⁶) ⁱ	K _{cat} /K _m (RLU/s/μM) (x 10 ¹⁵)
x11 Fluc	1.7	100	5.0±0.0	2.0±0.0	3.3±0.0	75.8±7.6	2.5±0.0	0.33±0.01
x11 ΔP353	0.9	54	6.0±1.5	2.0±0.0	3.3±0.6	66.6±8.3	2.0±0.0	0.30±0.02
x11 ΔP359	3.9	234	2.3±0.6	2.2±0.1	1.3±0.3	100.0±0.0	2.0±0.0	0.20±0.00
x11 ΔG360	2.5	150	3.3±0.8	2.0±0.0	6.0±0.5	58.3±6.3	2.0±0.0	0.34±0.01
x11 ΔA361	2.6	158	11.1±1.5	2.1±0.2	1.9±0.1	116.7±9.9	3.3±0.0	0.3±0.02
x11 ΔV362	3.0	183	6.7±0.6	2.5±0.0	3.9±0.4	150.0±1.5	1.6±0.9	0.10±0.03
x11 ΔG363	3.1	188	6.0±0.0	1.0±0.0	1.7±0.0	188.4±23.5	1.0±0.0	0.05±0.01

Table 1. Mutant and control specific activities and kinetic parameters. i. Proteins kept on ice until assay by injecting saturating D-LH₂ and ATP onto enzymes diluted in cold TEM buffer and integrating bioluminescence between 30s-250s using a BMG Fluostar Optima (BMG Labtech, Ortenberg, Germany). ii. Flash-height derived (Ugarova, 1989) K_{cat} values: since the initial flash heights of enzymes were similar to x11 Fluc the K_{cat} values were similar. Errors of constants are the standard error of the mean.

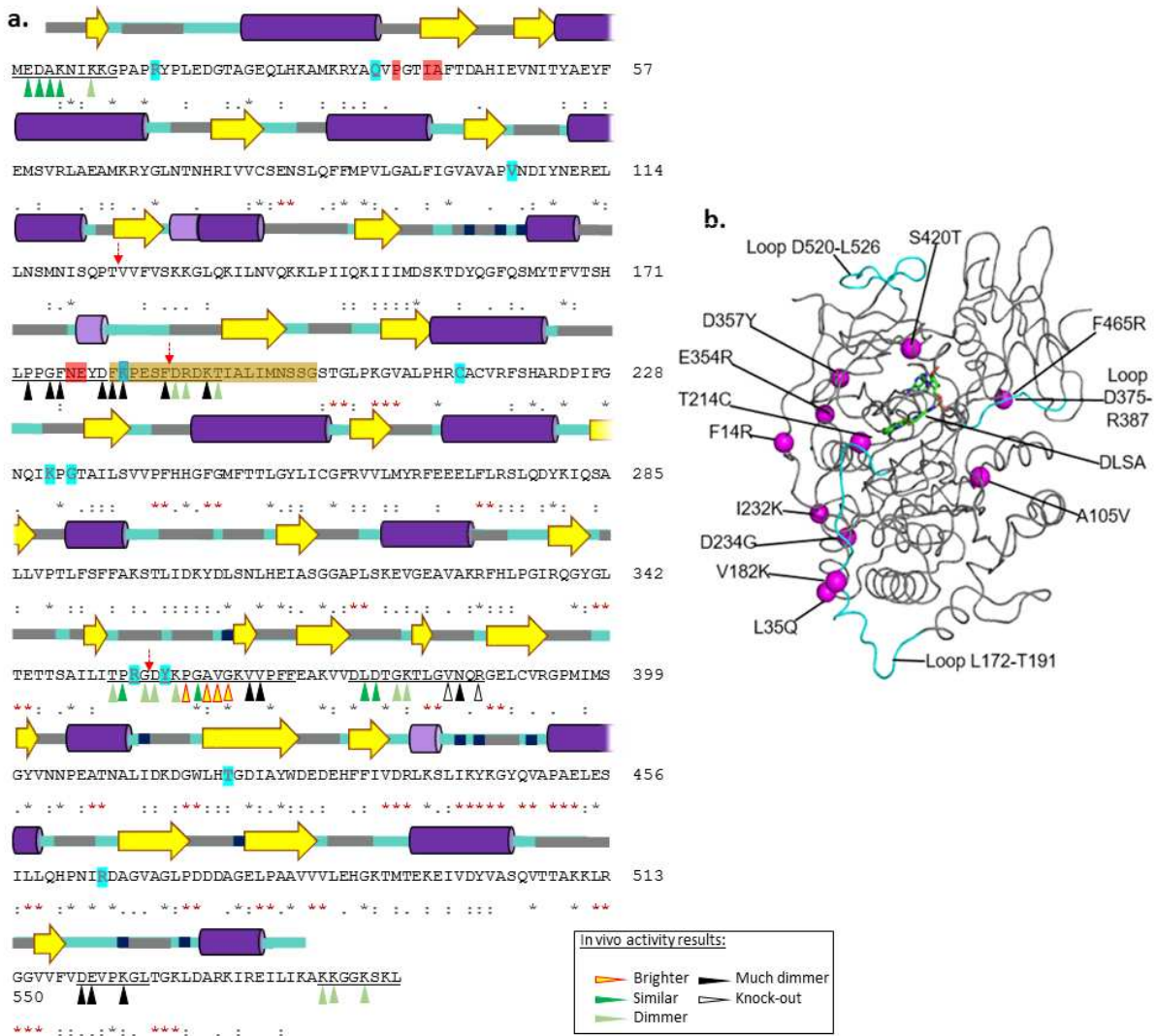


Figure 1

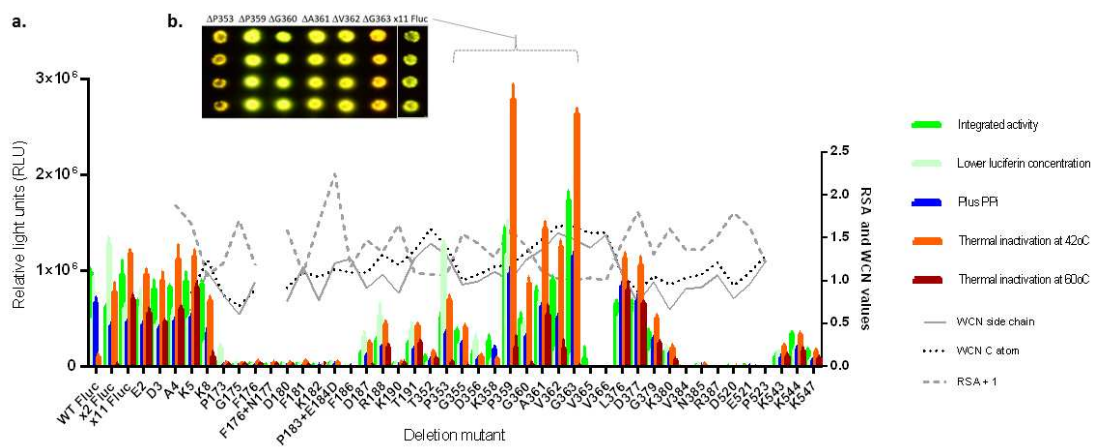


Figure 2

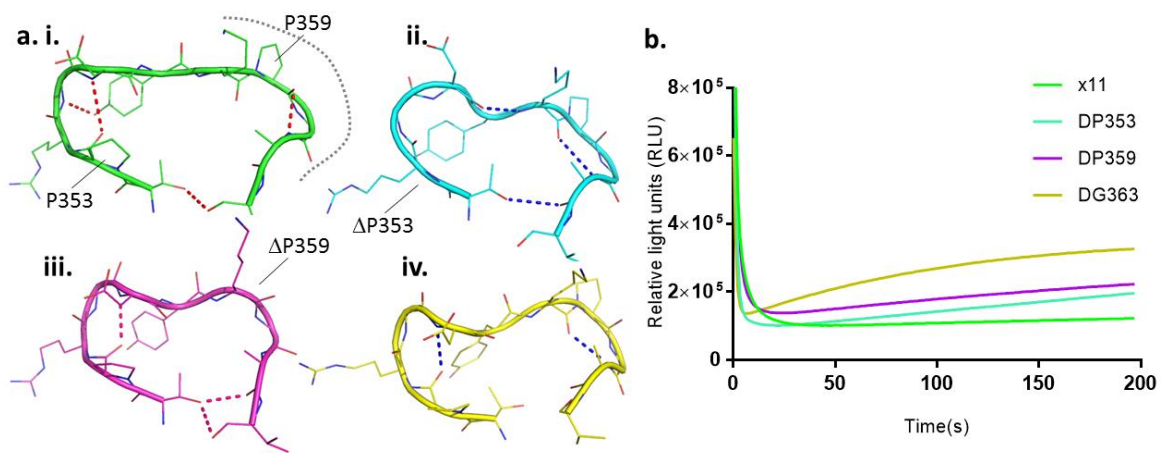


Figure 3

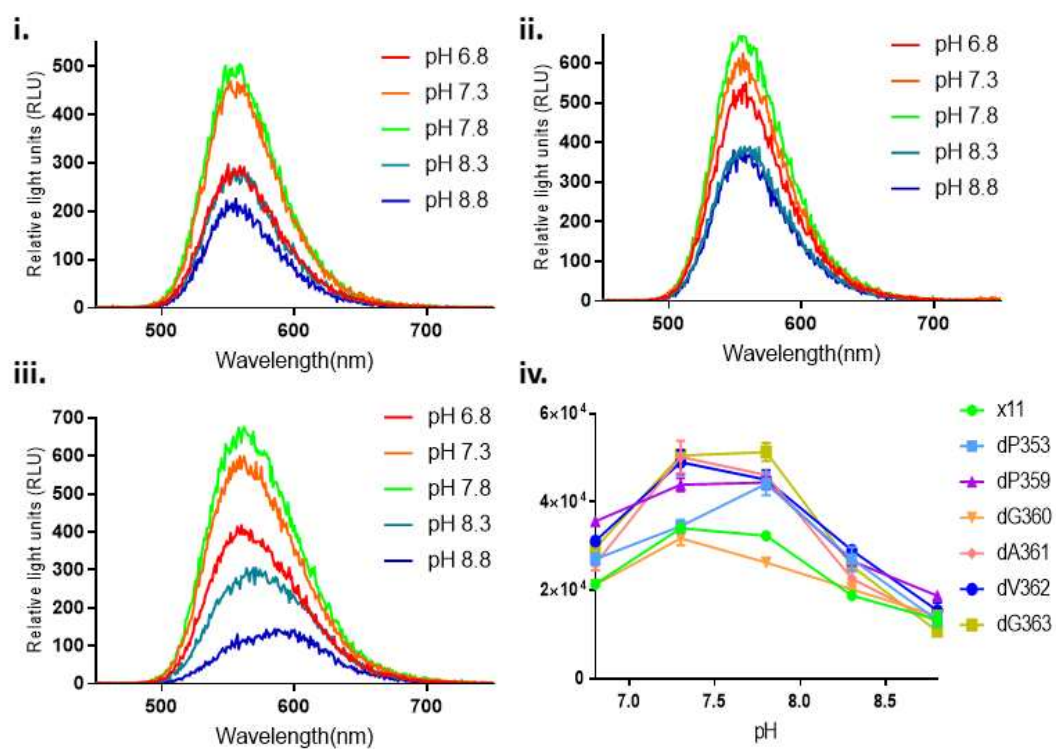


Figure 4

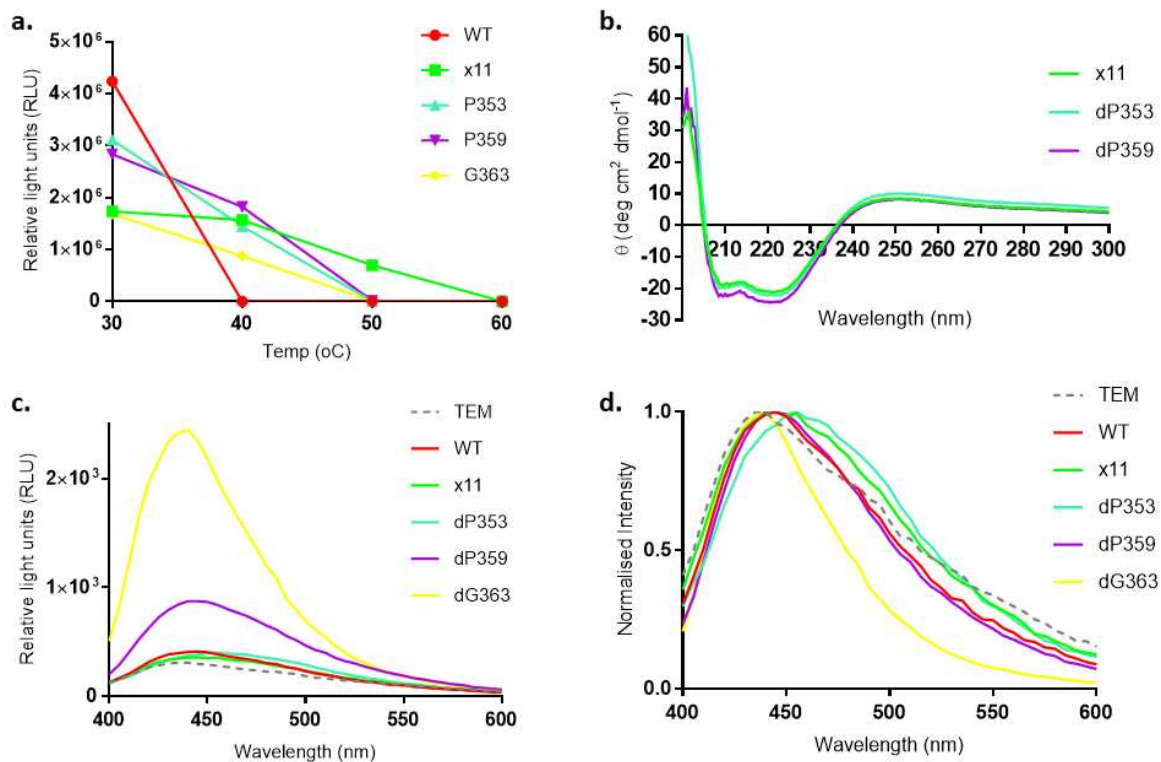


Figure 5