Engineering and Characterization of RNA-binding LOV photoreceptors

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

eingereicht an der

Fakultät für Chemie, Biologie und Geowissenschaften

der Universität Bayreuth

von

Thea Ziegler

geboren in Dresden

Bayreuth, 2018

Die vorliegende Arbeit wurde in der Zeit von November 2011 bis Oktober 2018 in Bayreuth am Lehrstuhl für Biochemie, und zuvor an der Humboldt Universität zu Berlin am Institut für Biologie, unter der Betreuung von Herrn Professor Dr. Andreas Möglich angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades einer Doktorin/ eines Doktors der Naturwissenschaften (Dr. rer. nat.).

Dissertation eingereicht am: 24.10.2018

Zulassung durch die Promotionskommission: 15.11.2018

Wissenschaftliches Kolloquium: 26.04.2019

Amtierender Dekan: Prof. Dr. Stefan Peiffer

Prüfungsausschuss:

Prof. Dr. Andreas Möglich	(Gutachter)
Prof. Dr. Birte Höcker	(Gutachterin)
Prof. Dr. Frank Hahn	(Vorsitz)
Prof. Dr. Olaf Stemmann	

‚Communicate plainly what you are trying to do in science, and, who knows, you might even end up understanding it yourself.'

Stephen Hawking in the context of an interview with BBC Radio 4 [0], in which he pointed out that, in a democratic society, everyone needs to attain a basic understanding of science to insure that the changes produced by scientific and technological progress are heading in the right directions.

0. Content

1. Summary

Light-Oxygen-Voltage (LOV) photoreceptors are light-sensitive signaling proteins that provide responses to light in the ultraviolet and blue regions of the spectrum. The light stimulus is transmitted to the functional output via structural rearrangements within the light-sensing domain, which consequently modulates the activity of the signaling protein. In recent years, the compact LOV modules have become popular scaffolds for constructing new optogenetic tools. The latter enable precise spatiotemporal control over diverse biological targets in a light-dependent manner. While many photoreceptors convey control over DNA-associated processes, so far there is no candidate that directly acts on RNA. The aim of this work was therefore to provide a novel approach for the light-controlled regulation of RNA molecules, either by identifying and characterizing a naturally occurring photoreceptor that fulfils the desired property of light-regulated RNA binding, or by recombining a well-characterized LOV photosensor with a suitable RNA-binding output domain. Searching for previously uncharacterized photoreceptor candidates in the sequence databases, we discovered a promising gene entry in the gram-positive actinobacterium *Nakamurella multipartita*. The putative protein product comprises an N-terminal PAS (Per-ARNT-Sim) domain, followed by an RNA-binding ANTAR (AmiR and NasR transcription antitermination regulator) domain and a C-terminal LOV domain, accordingly referred to as 'PAL'. Based on the domain arrangement, we assumed that the RNA-binding function of the ANTAR domain may be controlled by the blue-light-responsive LOV domain. We thus amplified the PAL gene from the genomic DNA of *N. multipartita* and confirmed its sequence identity via DNA sequencing. Next, PAL was heterologously expressed in *Escherichia coli* and purified via immobilized ion affinity chromatography. The purified PAL contains a flavin chromophore and undergoes the characteristic LOV photochemistry after blue light activation. We then applied SELEX (Systematic Evolution of Ligands by EXponential enrichment) to determine specific RNA target sequences for PAL and found two different motif families defined by a common consensus sequence. The five most promising variants were analyzed with the help of electrophoretic mobility shift assays for their binding properties to PAL. The PAL photoreceptor demonstrates a blue-light-triggered binding activity for all of the thereby tested constructs with an apparent K_D of around 0.25 µM for the best-binding aptamer under light conditions, which represents an approximate 30-fold increase compared to the corresponding binding activity in the dark. By optimizing the best-binding aptamer, we achieved binding affinities in the nanomolar range (30 ± 3 nM). Regarding the structural and mechanistic investigations of the PAL photoreceptor, we succeeded in obtaining the full-length crystal structure of PAL in its dark-adapted state with a resolution of 2.75 Å. The three-dimensional structure illustrates how signal transmission can be achieved within a LOV photoreceptor with the unusual domain topology of an N-terminally positioned output domain. With a combination of SEC-MALS (Size Exclusion Chromatography

combined with Multi-Angle-Light-Scattering) and SEC (Size Exclusion Chromatography) experiments, we demonstrated that both the full-length PAL protein, as well as the isolated LOV domain, occur as a dimer in solution independently of the light conditions. Moreover, we used the architecture of PAL as a design template for the development of further light-regulated RNA-binding proteins with an altered sequence specificity, and were able to generate light-sensitive constructs by replacing the PAL ANTAR domain with that of AmiR from *Pseudomonas aeruginosa*. The results indicate a great potential of PAL for use in optogenetic applications, as it opens the possibility of generating light-dependent RNA-protein interactions with high affinity. In addition, the structural studies on PAL provide valuable mechanistic insights that will facilitate the improvement of PAL as an optogenetic tool, as well as the construction of novel PAL-oriented chimeric photoreceptor variants.

2. Zusammenfassung

Light-Oxygen-Voltage (LOV) Photorezeptoren sind licht-sensitive Proteine, die Antworten auf Licht im ultravioletten und blauen Bereich des Spektrums vermitteln. Das Lichtsignal wird über strukturelle Umlagerungen innerhalb der Sensor-Domäne zum funktionalen Output weitergeleitet, wodurch die Aktivität des Proteins moduliert wird. In den letzten Jahren sind die kompakten LOV Module zu beliebten Vorlagen für die Konstruktion neuer optogenetischer Tools geworden. Letztere ermöglichen aufgrund ihrer Lichtregulierbarkeit eine präzise raumzeitliche Kontrolle über diverse biologische Targets. Während bereits zahlreiche Photorezeptoren existieren, die Kontrolle über DNA-assoziierte Prozesse vermitteln, gibt es bisher keinen Kandidaten, der direkt mit RNA interagiert. Ziel dieser Arbeit war es daher einen neuen Ansatz für die lichtsteuerbare Regulierung von RNA-Molekülen zu entwickeln. Mögliche Strategien hierfür sind die Identifizierung und Charakterisierung eines natürlich vorkommenden Photorezeptors, der die gewünschte Eigenschaft der lichtgesteuerten RNA-Bindung erfüllt, oder die Rekombination eines gut charakterisierten LOV-Photosensors mit einer geeigneten RNA-bindenden Output-Domäne. Auf der Suche nach möglichen Photorezeptorkandidaten in den Sequenzdatenbanken entdeckten wir einen vielversprechenden Eintrag im grampositiven Aktinobakterium *Nakamurella multipartita*. Das mutmaßliche Proteinprodukt umfasst eine N-terminale PAS (Per-ARNT-Sim) Domäne, gefolgt von einer RNA-bindenden ANTAR (AmiR and NasR transcription antitermination regulator) Domäne und einer C-terminalen LOV Domäne, weshalb wir es als 'PAL' bezeichneten. Basierend auf der Domänenanordnung vermuteten wir, dass die RNA-bindende Funktion der ANTAR Domäne einer von der LOV Domäne ausgehenden Blaulicht-Kontrolle unterlegen sein könnte. Daher amplifizierten wir das PAL Gen aus der genomischen DNA von *N. multipartita* und bestätigten die Sequenzidentität anhand von DNA-Sequenzierung. Anschließend reinigten wir das PAL Protein via Affinitätschromatographie aus *Escherichia Coli* auf. Das heterolog aufgereinigte PAL enthält einen Flavin Chromophor und durchläuft nach Blaulicht-Aktivierung die charakteristische LOV Photochemie. Daraufhin wandten wir das SELEX (Systematic Evolution of Ligands by EXponential enrichment) Verfahren an um damit spezifische RNA Zielsequenzen für PAL zu ermitteln. Auf diesem Weg identifizierten wir zwei verschiedene Motiv-Familien, die ein gemeinsame Konsensus Sequenz aufweisen. Die fünf vielversprechendsten Varianten wurden anschließend mithilfe von EMSA (Electrophoretic Mobility Shift Assays) auf ihre Bindeeigenschaften zu PAL untersucht. Der PAL Photorezeptor zeigte dabei eine blaulicht-induzierte Bindeaktivität für alle getesteten Konstrukte. Das bestbindende Aptamer demonstrierte hierbei unter Blaulicht eine scheinbare K_D von etwa 0.25 µM, was einer rund 30-fachen Steigerung gegenüber der entsprechenden Bindeaktivität im Dunkeln entspricht. Durch Optimierung des bestbindenden Aptamers erreichten wir schließlich Binde-Affinitäten im nanomolaren Bereich (30 ± 3 nM).

Im Rahmen der strukturellen und mechanistischen Untersuchungen des PAL-Photorezeptors gelang es uns die Volllängen-Kristallstruktur von PAL im dunkel-adaptierten Zustand mit einer Auflösung von 2,75 Å zu erhalten. Die dreidimensionale Struktur veranschaulicht, wie die Signalübertragung innerhalb eines LOV Photorezeptors mit der ungewöhnlichen Topologie eines N-terminal positionierten funktionalen Outputs erreicht werden kann. Mit einer Kombination aus SEC-MALS (Size Exclusion Chromatography combined with Multi-Angle-Light-Scattering) und SEC (Size Exclusion Chromatography) Experimenten konnten wir zeigen, dass sowohl PAL als auch die isolierte PAL LOV Domäne unabhängig von den Lichtbedingungen als Dimer in Lösung vorkommen. Weiterhin nutzten wir die Architektur von PAL als Designvorlage für die Entwicklung weiterer lichtregulierbarer RNA-bindender Proteine mit veränderter Sequenzspezifität. Durch den Austausch der PAL ANTAR Domäne mit der des AmiR Proteins aus *Pseudomonas aeruginosa*, konnten wir licht-regulierbare Chimären erzeugen. Die bisher erbrachten Ergebnisse deuten auf ein großes Potential von PAL für den Einsatz in optogenetischen Anwendungen hin, da es die Möglichkeit eröffnet lichtabhängige RNA-Protein-Interaktionen mit hoher Affinität zu erzeugen. Die strukturellen Untersuchungen an PAL liefern darüber hinaus wertvolle mechanistische Einblicke, die die Optimierung von PAL für den Einsatz als optogenetisches Tool, sowie die Konstruktion weiterer PAL-orientierter chimärer Photorezeptoren ermöglichen.

3. Introduction

Photoreceptors are sensory proteins that mediate responses to light in all domains of life. The adaptive responses that they control are highly diverse: in animals, they convey visual perception, while in plants and microorganisms they control the orientation towards light, referred to as 'phototropism' or 'phototaxis' [1]. In photosynthetic organisms, photoreceptors promote photosynthetic efficiency, e.g. by chloroplast accumulation or stomata opening. Sensory photoreceptors further control photoperiodic feedback mechanisms that serve adaptation to day and night cycles, such as the circadian rhythm in animals or flowering periods in plants. Owing to advances in the understanding of underlying molecular functions, the application of photoreceptors in modern biology nowadays far exceeds their original role in natural contexts. Since the implementation of the light-gated cation channel 'channelrhodopsin' from the green algae *Chlamydomonas reinhardtii* as light-activatable depolarization tool in neurons, a vast number of photoreceptors have been employed as light-regulated tools for the precise spatiotemporal manipulation of diverse cellular events. The popularity of this novel method, termed 'optogenetics' [2], increased with an enormous rate over the last 10 years, yielding more than 2700 published articles in the Pubmed database during this period (as to October 2018) [3]. While the first optogenetic applications utilized natural photoreceptors as light-sensitive tools, since then many novel photoreceptors have been developed to adress new specific challenges, such as the adaptation of spectral range, light sensitivity or response kinetics, as well as the integration of new effector outputs [4–6]. This chapter will provide an overview of the key chracteristics of photoreceptors in optogenetic frameworks (Section 3.1), with a special focus on the photochemistry and signal transduction mechanisms of Light-Oxygen-Voltage (LOV) domains (Section 3.2). It will further introduce a family of RNA-binding proteins as a potential candidate for the rewiring of a light-sensing LOV module with a novel output function (Section 3.3), and summarize the most common approaches to photoreceptor engineering (Section 3.4). The sections 3.1/3.4 are based on the review article Ziegler & Möglich, 2015 [6].

3.1 Characteristics of sensory photoreceptors

Photoreceptor proteins typically incorporate a pigment, the 'chromophore', that enables them to absorb light (Section 3.1.1) [1,7]. At the appropriate wavelength, photon absorption initiates the 'photocycle', which involves a set of photochemical reactions and structural transitions within the chromophore and the adjacent protein environment (Section 3.1.2). The light-driven transformations and resulting conformational changes within the protein backbone lead to the propagation of the signal. Most photoreceptors have a modular architecture in that their sensor and output functions

are organized in different modules, called ‘domains’ (Section 3.1.3). Photoreceptor activation hence requires the forwarding of the signal from one domain to another.

3.1.1 Chromophore

Chromophores used by sensory photoreceptors typically comprise a conjugated π-electron system, such as aromatic rings or double bonds, which enables the absorption of photons within the UV to infrared (IR) range. The spectral sensitivity of the photoreceptor is determined by the chromophore identity and the surrounding protein environment, and forms the basis for the distinction of different classes (see Figure 1) [1,7]. Most chromophores derive from small metabolites, although plant UV-B receptors [8] feature intrinsic amino acid chains that form an inherent chromophore. LOV, BLUF (blue light sensors using flavins) and cryptochromes use flavin-derived chromophores sensitive to UV-B and blue light [9–11]. Flavin-derived chromophores are easily produced by mammalian cells - for optogenetic purposes the chromophore availability within the target system is an important aspect.

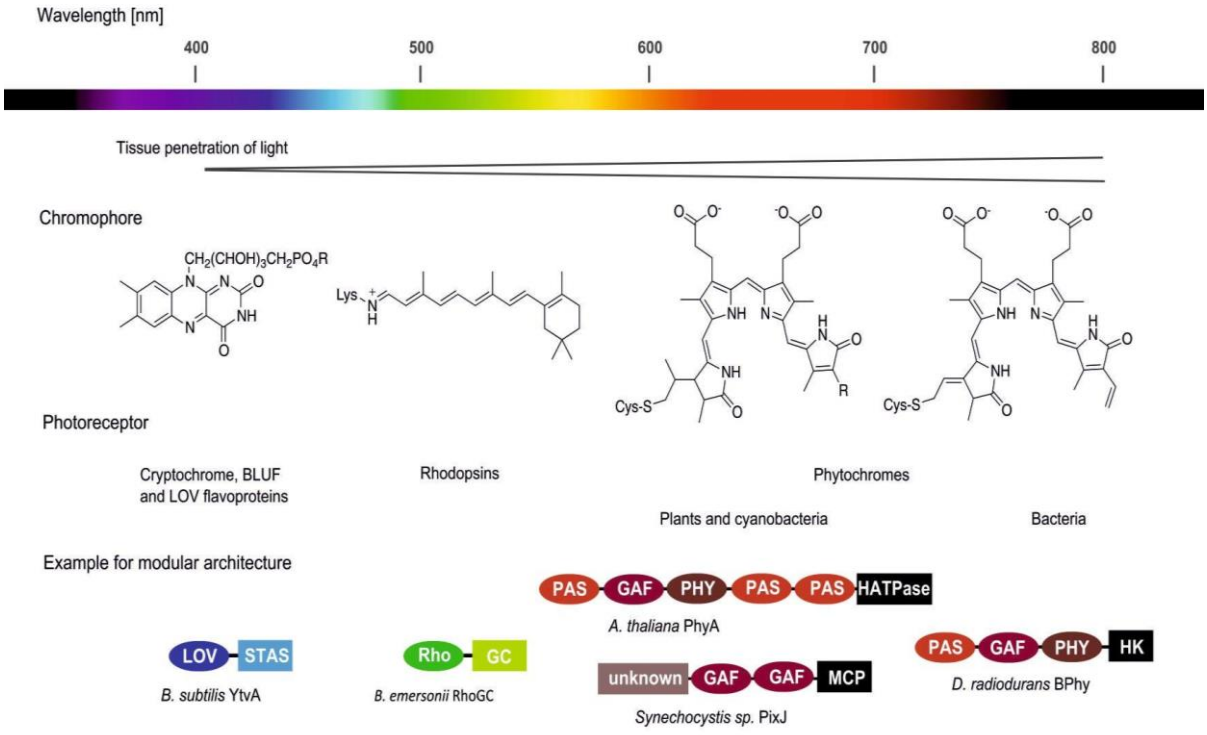


Figure 1: Properties of sensory photoreceptors; adapted from [6]. The spectral properties of a photoreceptor are determined by the identity of its chromophore, as well as the surrounding protein environment. Photoreceptors feature a modular architecture, which comprises at a minimum one sensor module that receives the light stimulus as an input signal, as well as one output or ‘effector’ module that implements a biological function in response to that stimulus.

Members of the rhodopsin family employ different retinal isomers as chromophores, thereby enabling visual perception in numerous animals and microorganisms, which can range from UV to red [13]. Phytochromes contain a simple bilin molecule as chromophore, which consists of a linear tetrapyrrole chain. Classical plant phytochromes, as well as bacterial phytochromes, perceive light in the red and far red range [14], while some algae species have extended their light sensitivity to the entire visible spectrum [15]. Cyanobacteriochromes also use linear tetrypyrroles as chromophore, but achieve an even higher spectral diversity ranging from the UV to the far red [16,17]. As part of the heme catabolism, the oxidized tetrapyrrole form biliverdin found in bacterial phytochromes is abundantly present in animal cells and tissues, while reduced tetrapyrroles such as phycocyanobilin, required by plant phytochromes and cyanobacteriochromes, are not found in higher animal tissues [6].

3.1.2 Photocycle

The term ‘photocycle’ refers to a series of photochemical reactions and structural transitions within the chromophore and the surrounding protein backbone, set off by the absorption of a photon of matching energy [7]. The principle of the photocycle is briefly described here in its simplest form, in which the photoreceptor can assume either the dark-adapted state (D) or the signaling state (S). The formation of S usually occurs within microseconds, and thus significantly faster than most cellular processes. The reaction from D to S can be described by the rate constant k(I), which depends on the light intensity I. The signaling state S might then persist from milliseconds to days before it reverts back to D in a thermal decay reaction (see Figure 2) with rate constant k_r. The probability for the formation of S determines the light sensitivity of the photoreceptor, which in turn correlates with the intrinsic quantum efficiency and absorption properties of the photoreceptor at a given wavelength. As the intrinsic quantum efficiency of most natural photoreceptors is already optimized for highly sensitive light perception, and an increase in light dose is only possible to a limited extent until it causes severe biological damage, it is difficult to influence the light sensitivity of a photoreceptor via these means. Alternatively, the experiments can be carried out under constant light conditions, which establishes an equilibrium between D and S, the so-called 'photostationary state'. Under such equilibrium conditions, the effective light sensitivity of the system no longer depends solely on the formation of S, but also on the recovery from S to D. This is of particular interest for optogenetic applications, as for some photoreceptor types, this reversion rate can be influenced by the introduction of point mutations within the chromophore region. For LOV proteins, this effect on thermal reversion can amount to several orders of magnitude [11]. Besides the spontaneous thermal decay reaction for the reversion from S to D, the return to D can be actively manipulated by

absorption of another photon of different wavelength for some photoreceptors types. These so-called 'photochromic photoreceptors' comprise the family of phytochromes, cyanobacteriochromes, and bistable rhodopsins, as well as certain engineered photoreceptors [18,19].

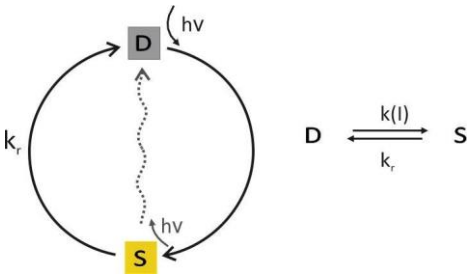


Figure 2: Simplified photocycle. A photocycle comprises at a minimum a dark-adapted state (D) and a signaling state (S) that is commonly triggered by the absorption of a photon (hv). The reaction from D to S occurs at the rate k(I) that depends on the light intensity I. The reversion from S to D occurs in a thermally driven single-step reaction with rate constant k_r. In so-called 'photochromic photoreceptors', the reversion to D can further be triggered by absorption of another photon of different energy.

3.1.3 Modularity

Like conventional signaling receptors [20], photoreceptors usually include at least one sensor module that receives an environmental stimulus as an input signal, and one output module, also called 'effector', which implements a specific function in response to that stimulus [4]. Groups of modules share recurring structural motifs, frequently organized in separate domains, as well as common principles of signal transduction. Conformational changes are mostly propagated by α-helical motifs, so-called 'linkers', which transfer structural changes from sensor to effector via thermodynamic coupling mechanisms. In homodimeric receptors, these linker elements often appear as coiled-coil bundles along the central molecule axis [21,22].

The complexity and dynamics of signal propagation within multi-modular signaling receptors render it difficult to conceive the functional features and interactions of these building blocks at once. For this reason, the decomposition of a signaling receptor into isolated building blocks with a reduced number of parameters may provide a starting point for the characterization of fundamental properties and interactions, which may lead to a higher level of understanding of the composite system. The modular nature of signaling receptors further facilitates the recombination of different sensor and effector modules in order to vary parameters of the composed system or to integrate new in- or outputs.

3.2 Light-Oxygen-Voltage photoreceptors

The responses of plants to light have been studied for many centuries [23], so that many discoveries about phototropic bending behavior were made long before the identification of the responsible photoreceptor in *Arabidopsis thaliana* [24]. Phototropins are serine/threonine protein kinases with the ability to detect light, which originates from a twofold repeated structural motif, nowadays known as 'LOV domain' [25]. Their activation requires specific UV-A or blue light (320 - 500 nm) [26–29]. Since their discovery in *Arabidopsis*, LOV proteins have been found across all kingdoms of life, mediating blue light sensitivity to more than 100 different effector types [30].

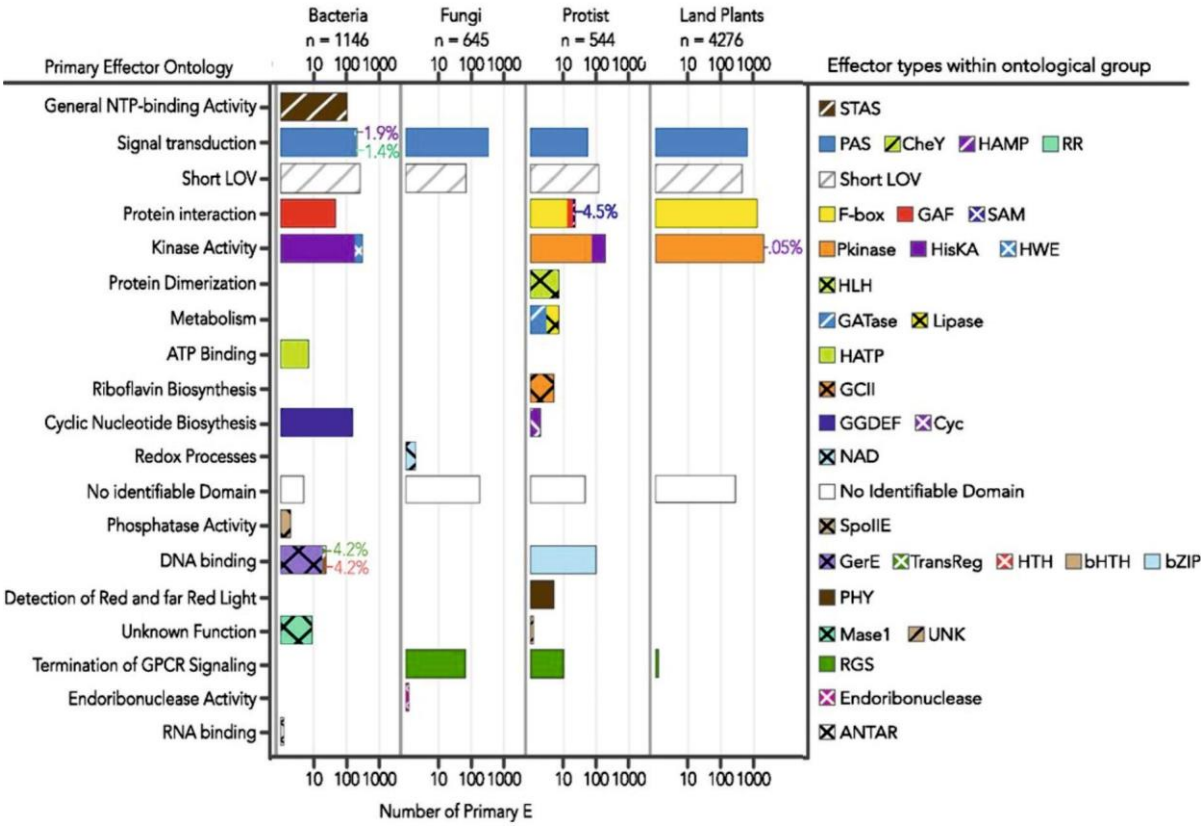


Figure 3: Diversity of LOV nearest neighboring effectors; figure extract from Glantz et al., PNAS 2016;113:E1442-E1451 [27]. The definition 'Effector' within this figure refers to the conserved domain closest to the LOV sensor based on the primary structure. The bars show the effector number of a particular functional cluster on a $\log_{10}$ scale and are hatched according to the type and proportions of the effector domains. For full names of effector abbreviations see Appendix Section 8.1. n - total number of LOV proteins found in each taxonomic rank.

3.2.1 Topological diversity

As most signaling receptors, LOV photoreceptors are modular. They are frequently connected to their respective effector modules via α-helical linker elements. The huge diversity of different physiological outputs regulated through LOV domains has been highlighted in a recent study [30], which investigated the LOV-effector topologies from a pool of two databases from more than 5,700 organisms. This led to the discovery of 119 functional clusters of effector modules. However, more than 80 % of the LOV proteins in the sample set can be assigned to five major effector categories: protein kinases, F-box domains, 'Short' LOV domains that lack a covalently attached effector module, histidine kinase phosphoacceptor domains (HisKA), and Per-ARNT-Sim (PAS) domains. Besides the conclusions about the presumed light-regulated function, the analysis of conserved domain architectures confirmed the previously reported observations [31–34] that the distribution of most effectors with respect to the LOV sensor follows a characteristic pattern, i.e. some effectors, such as the bacterial HisKA domains, occur primarily at the C-terminus, while others occur mainly N-terminally. For certain effector types, the lengths of the connecting linker elements further show a heptad-periodic pattern due to the preservation of α-helical structures within extended coiled-coil linkers [4,32,35] .

3.2.2 LOV domain structure

LOV domains belong to the PAS family, with which they share the same core domain, defined by approximately 110 amino acids that adopt the distinctive PAS fold. This characteristic tertiary motif is composed of a central five-stranded anti-parallel β-sheet and several α-helices. For the subclass of LOV proteins, one of these α-helices holds a conserved consensus motif ('GXNCRFLQ') that coordinates the photoreactive flavin chromophore (flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD)) [36,37]. Over the past few years it has been further shown that this core domain commonly features variable N- and C-terminal α-helical extensions, denoted A'α and Jα respectively (see Figure 4), that play important parts in the allosteric control of effector domains [21,38–42].

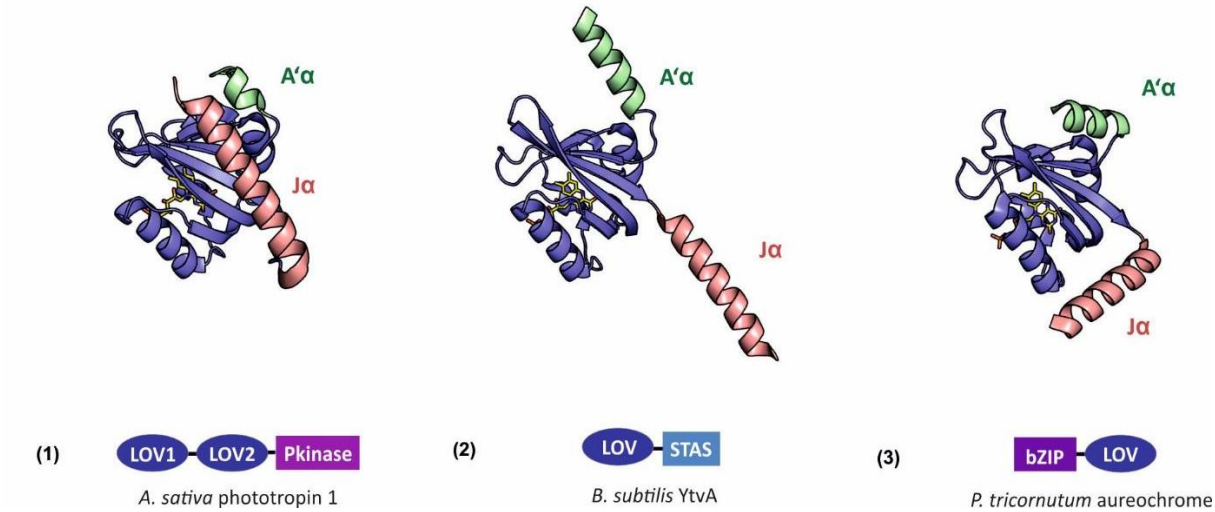


Figure 4: Structures of different LOV proteins with the associated N- terminal A'〈 and C-terminal J〈 extensions highlighted in green and red, respectively. (1) LOV2 domain from *Avena sativa* phototropin 1 (PDB code 2V0U). (2) LOV domain from the engineered histidine kinase YF1, found in *Bacillus subtilis* YtvA in its original context (PDB code 4GCZ). (3) Aureochrome 1a LOV from *Phaeodactylum tricornutum* (PDB code 5DKK).

3.2.3 LOV photochemistry

All LOV photoreceptors are defined by a joint photochemistry involving the formation of a covalent bond between the so called 'active site cysteine' and the flavin chromophore following blue light activation [43,44]. In the dark-adapted state (D447), the oxidized flavin chromophore is not covalently bound to the binding pocket and absorbs light at a maximum of 447 nm. Blue light absorption then induces the reversible formation of a covalent bond between the conserved active site cysteine and the C4a position of the flavin nucleotide. The molecular reactions of the photoactivation of LOV proteins can be summarized as follows [45–55]: blue light converts the dark-adapted D447 state to an excited singlet state, from which it decays via intersystem crossing into a triplet state ($T_1$715) within nanoseconds. The formation of the photoproduct that characterizes the signaling state then proceeds in two steps within microseconds: (i) the excited triplet state stimulates an electron and proton transfer from the active site cysteine sulfhydryl group to the N5 site of the flavin chromophore. (ii) The resulting reactive intermediate then recombines to the C4a adduct characterized by a broad absorption band around 390 nm, which defines the signaling state (S390).

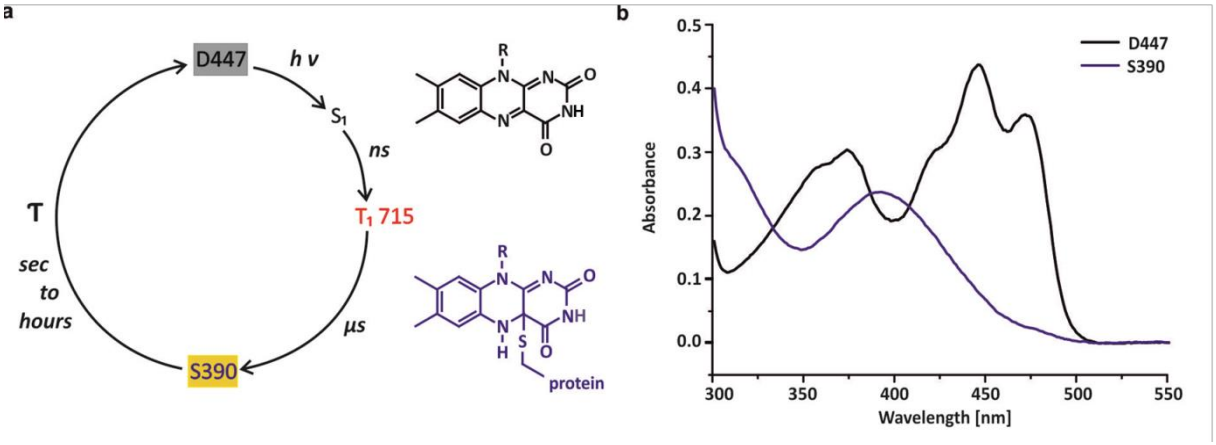


Figure 5: Photochemistry of the family of LOV receptors. **(a)** The LOV photocycle is characterized by a dark-adapted state with an absorption maximum at 447 nm (D447). Upon blue light stimulation, adduct formation proceeds through a rapidly formed triplet species with an absorption maximum at 715 nm ($T_1$715). The signaling state (S390) features a characteristic single absorption band at 390 nm. The signaling state thermally decays to D447 within seconds to hours at a time constant of $\tau = 1/k_r$ (k_r - rate constant). **(b)** UV-Vis absorbance spectra of a typical LOV protein in its dark-adapted (D447) and signaling state (S390).

The formation of the photoproduct is accompanied by an altered hydrogen bonding of the flavin chromophore (see Figure 5.a). The photoreaction is thermally reversible and reverts back to the dark-adapted state D447 within seconds to several hours [56]. In addition to the thermal decay path, it was found that near-UV light can drive the rupture of the photoproduct, thereby competing with the formation of the signaling state [52,53]. Under constant illumination, all of these three pathways combine into a photostationary equilibrium state, which determines the sensitivity to environmental changes, such as periodic or momentary oscillations in light intensities. As both the formation of the signaling state and the UV-light driven reversion to the dark-adapted state occur on a comparable time scale, the variability of LOV photoreceptors in regard to their light sensitivity rests upon the reversion rate from signaling to dark-adapted state. The reversion rate (or rate of adduct decay) is thus of large interest for the design and improvement of optogenetic tools, although its biological role is barely explored [11,56–58]. The reversion rate has been classified into three regimes in regard to the resultant sensitivity of the LOV receptor to environmental light [11]: (i) the so-called 'fast-cycling' LOV domains ($\tau < 1000$ s) afford rapid on/off-kinetics, but require high light-intensities in order to reach saturation of a respective signal. Within their natural context, they are sensitive to all environmentally occurring light intensities. In contrast, the LOV domain of the (ii) 'intermediate cycling' regime (1000 s $< \tau <$ 10000 s) become saturated at moderate light intensities (> 20 µmol photons/m^2s), but retain maximal sensitivity under low light conditions as found during dusk and

dawn. Members of the last, (iii) 'slow cycling' (10000 < τ s) group only require minimal light, enabling circadian clock photoreceptors to detect the onset of the day. However, within the scope of optogenetic applications, these slow off-kinetics often result in an effective irreversibility of the investigated biological effect. Over the past years, extensive efforts were put into the tuning of LOV photocycles, with the result that the reversion rates of diverse LOV photoreceptors can now be varied up to three orders of magnitude through the introduction of single residue mutations. The most important sites for rate altering variants were summarized in a recent report [11].

3.2.4 LOV signal transduction

While the primary photoreaction of LOV proteins occurs within the PAS core domain, defined by the central β-scaffold and α-helices that coordinate the flavin chromophore, signal propagation further involves the N- and C-terminal helical extensions that function as transmitter between sensor and effector domain [21,35,38]. While the first steps during signal propagation are highly conserved among LOV domains due to the similar photocycle, the allosteric signal transmission mediated by the terminal extensions differs strongly among individual representatives. The large number of mechanistic studies on signal transduction in LOV photoreceptors led to the identification of two essential effects that initiate the propagation of the signal upon photoproduct formation: (i) protonation of the N5 site of the flavin chromophore, causing changes of the hydrogen bonding in the surrounding protein environment, and, (ii) an increased electron density near the N5 and C4a sites of the FMN, that stabilizes the build-up of charge [11]. The influence of electron density can be illustrated by comparing the LOV1 and LOV2 domains found in the characteristic tandem motif at the N-terminus of phototropins. Phototropin LOV2 domains typically feature a conserved phenylalanine near the photoreactive site whose position is mostly occupied by a leucine in phototropin LOV1 domains. While light-activation and variation in temperature lead to strong depletion of α-helical and β-sheet content within phototropin LOV2, the content in secondary structure of phototropin LOV1 domains hardly changes upon light-activation. Several studies concluded that the diffuse electron density distribution of phenylalanine and methionine residues thus destabilizes the build-up of charge following the formation of the photoproduct, thereby promoting accelerated off-kinetics [45,60].

Owing to the large number of studies on *A. Sativa* phototropin LOV2 (AsLOV2) [38,45,55,61–64], the above described mechanisms can be brought together in the following model: protonation of the N5 site affects the hydrogen-bond contacts of Q513, a glutamine residue located within the β-core sheets conserved in most LOV photoreceptors. In this case, adduct formation is accompanied by a reduction of the β-leaf contacts as well as a strong decrease in α-helical content, which was ascribed

to the detachment of Jα from the PAS core and subsequent unfolding. The Q513 position is hence thought to function as an allosteric switch that functionally couples the C-terminal Jα helix to the photoreactive site. For the LOV domain of the stress response regulator YtvA from *Bacillus subtilis*, structural and mechanistic studies suggest that the light signal is propagated via rotational motions of the coiled-coil linker element. This linker establishes the connection to the homodimeric C-terminal effector as an elongation of the Jα helical extension [32,65,66]. The high-resolution crystal structure of the full-length engineered histidine kinase YF1, which was designed by replacing the STAS effector from YtvA by a histidine kinase function, in its dark-adapted state revealed an additional coiled-coil extension at the N-terminus of the LOV-core domains, denoted A'α. This N-terminal extension runs coaxially to the C-terminal coiled-coil linker. Mutational studies have shown a direct influence of the A'α extension on the light regulation of the effector function, since even individual mutations within this element led to attenuation or even inversion of the signal response [21]. Similar to YtvA and YF1, the vast majority of LOV photoreceptors consist of an N-terminal photosensor connected to a C-terminal effector domain. The first described exception to this 'rule' was the transcriptional regulator Aureochrome 1a from the diatom *Vaucheria frigida* (VfAur1a), which features an N-terminally coupled basic leucine zipper (bZIP) domain attached to the blue light-sensing LOV module [67]. For the Aur1a homolog from *Phaeodactylum tricornutum* (PtAur1a), the following model was proposed for signal transduction [68]: in its dark-adapted state, PtAur1a occurs as a dimer and the LOV and bZIP domain interact with each other, thereby inhibiting desoxyribonucleic acid (DNA) binding. The highly dynamic N- and C-terminal A'α and Jα extensions pack against the LOV core, until blue light activation and the resulting C4a adduct formation causes undocking of Jα from the LOV β-sheets, which in turn induces the release of A'α and subsequent dissociation of LOV and bZIP domain [69]. This permits the LOV domains to form a dimer, entailing an increased affinity of PtAur1a for its target sequence [68].

An important finding of the numerous studies on signal transmission in LOV proteins is that the comparison of LOV receptors to binary switches, which alternate between an on- and off-state, does not reflect reality. Rather, they correspond to a progressive dimmer model in which LOV proteins maintain a certain basic activity in the dark, which is only amplified by increased light intensities [11].

3.3 Photoreceptor engineering

The modular architecture of photoreceptors opens up the possibility of developing new photoreceptor candidates by recombining well characterized light-sensing domains with new cellular functions previously not controlled by light. Since the groundbreaking description of the light-

sensitive channelrhodopsin as depolarization tools in neurons, numerous other types of photoreceptors have been used for optogenetic applications, i.e. the investigation and control of biological functions with the aid of light. The success of this method is closely related to the inherent properties of sensory photoreceptors: (i) due to their protein nature, they can be genetically encoded and expressed in situ; (ii) the use of light as signal input enables a non-invasive form of control up to a certain tissue depth and high spatial-temporal resolution; and (iii), their reversibility opens up the possibility of causing transient and repeated perturbations in many applications. The first optogenetic applications were based on naturally occurring photoreceptors, but since then numerous customized photoreceptors have been developed that extend the repertoire of light-controllable cellular processes. As mentioned in Section 3.1.3, signal transduction within photoreceptors is regulated through allosteric coupling; i.e., the reception of a signal leads to conformational changes in the chromophore and the adjacent protein environment. To outline the different strategies for the development of novel photoreceptors we classified [6] the underlying signal transduction mechanisms according to whether light induction causes changes in their oligomeric state ('Associating photoreceptors') or not ('Non-associating photoreceptors').

3.3.1 Classification of photoreceptors upon underlying signaling mechanisms

Associating photoreceptors

The absorption of a photon generally leads to allosterical changes that manifest in local unfolding and/ or dissociation of the linker from the photosensory core, thereby leading to domain rearrangement and/ or modification of interaction surfaces. If the latter causes a light-induced association or dissociation of functional oligomers or heterologous binding partners, the concerned photoreceptor can be assigned to the category of associating photoreceptors. Depending on whether the light-induced change of oligomeric state is caused by a homologous or heterologous binding partner, this category can be divided into two further groups (see Figure 6). For the first group, regulation of biological activity is based on formation of the functional oligomeric state, whereas for heterologous oligomerizers biological activity is controlled through co-localization of interacting proteins, which might involve the recruitment to cellular compartments. Examples for naturally occurring homologous oligomerizers are the blue light-sensing LOV photoreceptors Vivid from *Neurospora Crassa* [70], aureochromes found in different stramenopiles [67], or the transcription factor EL222 from *Erythrobacter litoralis* [39]. The most important representatives for heterodimerizing systems are derived from *A. thaliana*, such as cryptochrome 2 and its interacting partner CIB1 [71], as well as the red to far-red sensing receptor Phytochrome B (PhyB) and its

interacting partners, called 'PIFs' (phytochrome interacting partner) [72,73]. The respective interaction mechanisms are well understood, which facilitated various optogenetic applications.

Non-associating photoreceptors

Non-associating photoreceptors form a diverse group for which signal-induced conformational changes lead to a rearrangement of subdomains via tertiary or quaternary transitions (see Figure 6). As described in Section 3.1.3, the sensor and effector domains of most photoreceptors are organized in different modules, linked together through α-helical motifs in a thermodynamically favorable manner. The physical nature of the linker is of crucial importance in that context, as linker residues have to interact with both sensor and effector sites to enable the propagation of light-induced conformational changes from the photoreactive site. Although the α-helical motif of the linker is conserved among many signaling receptors, the mechanistic details differ for individual cases: For multi-domain proteins, local unfolding [74], or torque movements [21] have been suggested as possible modes of signal transduction, often leading to rearrangements of the interfaces of neighboring modules.

The LOV2 domain from *Avena Sativa* (AsLOV2) phototropin 1 presents a particular case for photoreceptor engineering: Its C-terminal Jα helix partially unfolds upon light-perception leading to dissociation from the photosensory core [75] (Order/ disorder transitions; see Figure 6). In its natural context, this results in rearrangements of the phototropin subdomains without concomitant changes in the oligomeric state of the overall photoreceptor [76]. Thus, from a mechanistical point of view, AsLOV2 is part of the group of non-associating photoreceptors. However, several engineered AsLOV2-based receptors with obligate oligomeric effectors used AsLOV2 in association-based design approaches by fusing the Jα helix with selected effectors thereby blocking the active site in the dark-adapted state. Upon light-activation Jα unfold, which removes the steric restriction to the active site and increases biological activity [77]. Nonetheless, the mechanism of local unfolding is not limited to AsLOV2, but also appears in other light-sensitive proteins, such as Aur1a from *P. tricornutum* [78], or the LOV domain from *Rhodobacter sphaeroides* [79].

3.3.2 Design approaches

The design strategy for new photoreceptors largely depends on the underlying signal transduction mechanism. Figure 6 provides an overview of the design approaches described below.

Association-based

In general, effector functions that are regulated by association or dissociation in their parental context should be likewise combined with sensor domains originating from associating photoreceptors (Photoactivatable association/ dissociation). However, it is also possible to control proteins by light-induced association or dissociation, which were not regulated by oligomerization processes before, e.g. through reconstitution of split proteins. To date, the use of photoactivatable association/ dissociation is the most prevalent approach for the design of novel photoreceptor variants [5,12]. One reason for their great versatility is that the requirements on the connecting linker between sensor and effector are much less demanding than for the group of non-associating photoreceptors; i.e., it is often sufficient that the linker element provides a physical connection, for which it should be sufficiently short, flexibly and predominantly hydrophilic.

Examples of association-based approaches are provided by diverse light-induced expression systems for eukaryotes [80–82] and optically controlled receptor tyrosine kinases (RTK) [84–86]. Light-induced eukaryotic gene expression was first achieved by exploiting the red/far-red –controlled interaction of the *A. thaliana* phytochromes A and B with the interacting factor PIF3 [73]. Light-controllable eukaryotic expression was further accomplished through the use of the *N. crassa* Vivid LOV-domain, which assembles into homodimers upon blue light illumination [81,82]. For the design of light-sensitive RTKS, different aureochrome LOV domains were fused to the C-terminus of membrane-bound receptor tyrosine kinases, thereby replacing ligand-binding by blue light induction as the activating stimulus for formation of the functional RTK dimer [83]. The original sensor-effector orientation was maintained in this case, since aurreochromes also display an N-terminal effector in their natural context. Other design constructions with LOV domains of deviating sensor-effector orientations in their original contexts, showed no light-sensitivity.

Other

In the case of AsLOV2 (or likewise-behaving photoreceptors), light-triggered order/ disorder transitions allow the alteration of accessibility of active sites and epitopes, thereby regulating the activity of the protein itself and subsequent pathways. The use of photoreceptor domains that undergo order/ disorder transitions has repeatedly proven to be a promising strategy to obtain light-regulation (Peptide uncaging). In a first successful implementation of AsLOV2 Jα-unfolding, Rac1, a small GTPase involved in the regulation of cytoskeletal dynamics, was fused to the AsLOV2 Jα helix [77] and permitted the control of the active site of Rac1 through steric restriction.

For photoreceptors that transmit signals via tertiary or quaternary structural transitions, the most promising strategy is a domain replacement upon structural superposition of related (homologous)

sensor domains within similar sensor-effector architectures. For instance, GAF (cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA) domains could be exchanged by domains with homologous light-sensing GAF or derived PHY (phytochrome specific) domains: Similarly, structurally homologous LOV domains could replace related PAS domains. The availability of three-dimensional structures or homology models facilitates the identification of suitable fusion sites by enabling structural instead of sequential alignments. If no homologous relatives can be found, heterologous substitution can also lead to functional proteins, but due to the less specific alignment it is more challenging to find a suitable fusion strategy. Substitution of receptor domains linked by structured α-helical linkers (e.g. coiled-coil linkers) requires a detailed examination of the linker properties to find the correct fusion site in order to maintain the efficient transduction of the light-stimulus from sensor to effector [4,32].

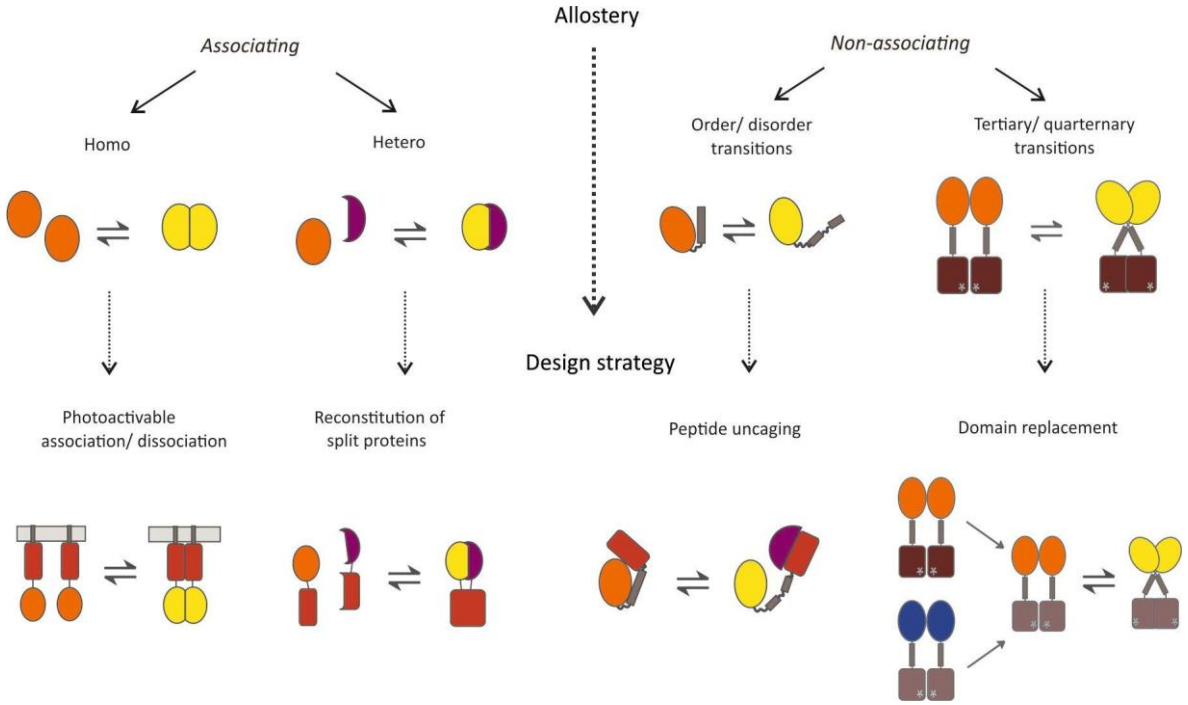


Figure 6: Allosteric mechanisms of signal transduction in natural photoreceptors and derived design strategies for the engineering of new photoreceptors. Photoreceptors that undergo light-induced changes in their oligomeric state can be employed as photoactivatable association or dissociation modules. Resulting engineered photoreceptors can be activated, e.g. by building up the functional oligomeric state, or mediating the reconstitution of cleavage proteins, which can also be used for co-localization purposes. Another popular design strategy that results from light-induced order/ disorder transitions is peptide uncaging, which allows making signal sequences accessible or inaccessible through light-controlled steric hindrance. The strategy of domain replacement is suitable for photoreceptors that do not experience any light-induced changes in their oligomeric state, and instead undergo other light-induced transitions in tertiary or quaternary structure. In this case, the conservation of the nature of the linker element is often of central importance.

3.3.3 Considerations regarding the choice of the effector

The choice of the effector domain for the design of a new photoreceptor is principally determined by the desired function. The type of the parental effector protein determines the specific activity that in turn defines the dynamic range of the perspective light-regulated receptor, as well as the option to establish an efficient high- or medium throughput assay to facilitate the screening of functional photoreceptor variants. The dynamic range of a signaling receptor is defined by the factor of catalytic activities in its higher (on) and lower (off) activity state, hence in presence or absence of the signal or vice versa. The dynamic range is primarily limited by residual dark activity. For engineered photoreceptors, the dynamic range of the parental effector domain usually limits the light-induction or -repression factor. On the other hand, it is not guaranteed that designed photoreceptors derived from high-dynamic range parent proteins will yield light-regulated derivatives with similar properties. In certain cases, the resulting dynamic range of a light-regulated system can be strongly increased by embedding the light- or dark-induced signal in cellular signaling pathways, e.g. involving second messengers [90,91] or MAP kinase-mediated signaling pathways [83].

Moreover, engineering photoreceptors often requires testing of numberous chimeric variants. A fast and simple assay that permits the detection of light-regulated activity thus presents a highly useful tool for the identification of functional photoreceptors. *In vivo* screening systems in eukaryotic or bacterial cells usually meet these criteria. In general, the desired output predefines the options for a high- or medium-throughput assay that facilitates the screening. The output should be ideally easily detectable (e.g. fluorogenic or chromogenic), orthogonal to other cellular metabolic pathways, and non-toxic to the cells. The direct or indirect coupling of the light-regulated output to the expression of a reporter gene usually results in a high screening throughput. If it is not possible to detect light-regulated activity in cellular systems, medium-throughput assays may be an option, which allow to determine a potential activity in raw cell extracts through the detection of specific metabolites or enzymatic reaction products [92].

3.4 ANTAR: extension of the optogenetic toolbox by an RNA binding domain

In addition to the critical role of messenger RNA, ribosomal RNA or transfer RNA in the conversion of genetic information to the protein level, new types of non-coding RNAs (ncRNAs) have gained increasing attention in recent years due to their implication in dynamic epigenetic regulation processes. In contrast to the coding DNA regions (genes), the proportion of non-coding DNA in the genome increases with increasing complexity of the organisms. In the human genome, this proportion is 98.8 %, although it was long referred to as 'junk DNA' until it became apparent that >

85 % are transcribed into ncRNAs [93]. The so far demonstrated functions are remarkably diverse; in eukaryotes, the diverse types of ncRNA include the class of long non-coding RNAs that hold specific expression patterns and subcellular sites, as well as many different classes of small regulatory RNAs, often involved in gene silencing [94]. In procaryotes, different types of small regulatory RNAs have been known for a long time. As posttranscriptional regulators of gene expression, they have a major influence on adaptation to different growth and environmental conditions [95]. However, for many of these RNA molecules, their natural function is not well understood, so that an RNA-binding optogenetic actuator could be of great benefit for the ongoing research in the future. A potential candidate for a suitable effector type is the family of ANTAR proteins that feature RNA–binding effector domains.

3.4.1 ANTAR domain occurrence

ANTAR proteins are involved in posttranscriptional regulation processes through antitermination of stem-loop secondary structures that pause transcription [96,97]. ANTAR domains occur in multi-modular architectures in various combinations with putative sensor domains. The Pfam protein family database [98] predicts more than 3600 occurrences of the ANTAR domain distributed across 1722 bacterial species (as of April 2018). Common domain architectures include the combination with well-known sensory domains, such as the ubiquitous GAF, PAS or nitrate and nitrite sensing domains (NIT). Representative for the latter, NasR from *Klebsiella oxytoca* positively regulates the *nasFEDCBA* operon involved in nitrogen assimilation [96]. However, most ANTAR proteins feature an N-terminal response receiver (RR) domain found in two-component systems. Two-component systems usually consist of a regulatory histidine kinase and a corresponding RR protein. The histidine kinase undergoes autophosphorylation in response to specific environmental stimuli; activation of the RR occurs upon transfer of the phosphoryl group [99]. This way, two-component systems regulate a wide range of mechanisms, such as gene expression by RR-containing transcription factors. For the so far characterized ANTAR-RR proteins, however, the regulatory action was shown to occur on a posttranscriptional level through antitermination of terminator structures within messenger RNAs [100].

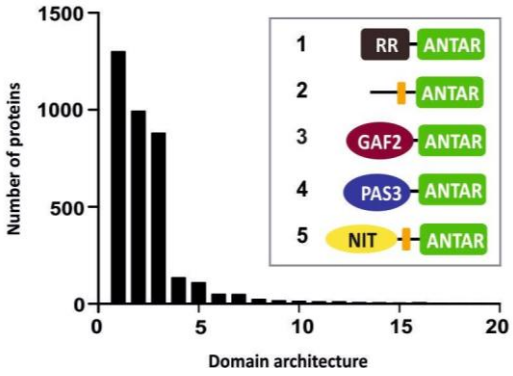


Figure 7: Domain distribution of ANTAR proteins from the Pfam protein database [98]. The five most abundant domain architectures of ANTAR proteins are shown in the box.

For example, the regulation of the ethanolamine utilization (*eut*) operon in *Enterococcus faecalis* was shown to be controlled by the RR-containing ANTAR protein EutV and the corresponding histidine kinase EutW [100]. EutW is autophosphorylated in the presence of ethanolamine, the phosphoryl group is then transferred to EutV. Phosphorylation causes EutV to dimerize thereby enabling association with its RNA target sites located upstream of the *eut* operon. The second ANTAR-containing RR actuator Rv1626 from *Mycobacterium tuberculosis* has been predicted to function in a similar manner to EutV, although the exact input and output signals are unknown [101]. AmiR from *Pseudomonas aeruginosa* represents a second class of ANTAR regulators that possess an N-terminal domain resembling the classical RR fold, but lacks the essential residues for phosphoryl group acceptance. This class of 'pseudo-RR' ANTAR proteins is expected to mediate gene-expression via physical interactions with an additional regulator protein [102]. The activity of AmiR is regulated through its negative modulator AmiC. Induction of AmiC in turn occurs by binding of small 2- to 3-carbon amide compounds, whereas butyramide acts as anti-inducer. AmiC-induction leads to release of AmiR and subsequent association with the 5' leader of its target RNA sequence, which was hypothesized to prevent the formation of a terminating stem-loop [103].

The occurrence of ANTAR proteins in domain combinations with the versatile PAS or GAF domains, which regulate diverse cellular responses to a variety of physical or chemical signals (e.g. gas molecules, redox potential, or photons), indicates that ANTAR domains constitute global bacterial regulatory modules in diverse contexts beyond nitrogen assimilation.

3.4.2 ANTAR domain structure

Only three crystal structures of ANTAR proteins are available so far: The first structure belongs to AmiR from *P. aeruginosa,* which was shown to act as a positive transcription regulator of the amidase

operon [104], the second crystal structure comes from the *M. tuberculosis* NasT homologue Rv1626 [101], and the last of the so far available structures belongs to the unusual transcriptional antiterminator NasR from *K. oxytoca* [105]. All three structures share a conserved fold within the ANTAR region composed of a three-helical bundle that comprises five highly conserved residues. These residues include three alanines, one aromatic residue, and one alanine or serine [106].

The crystal structure of AmiR was the first structure of an ANTAR protein [104]. The structure comprises the complex of AmiR together with its negative regulator AmiC, both of which occur as homodimers. AmiR itself consists of an N-terminal 'pseudo'-RR domain, i.e. a RR module displaying the distinctive α-β-α sandwich fold, but lacking the conserved residues required for the phosphoryl group acceptance. The C-terminus of this domain prolongs into a long α-helix that extends into a parallel coiled-coil with the equivalent residues of the other monomer. The C-terminus of the protein forms the three-helical bundle that defines the characteristic ANTAR fold (see Figure 8). For many ANTAR proteins a coiled-coil region along the N-terminus of the motif was predicted through bioinformatics analysis [106]. Among the crystallized ANTAR representatives, only AmiR presents a coiled-coil region centered on the N-terminal α-helix, which represents the dimerization interface for formation of the functional unit. The NIT-domain containing ANTAR regulator NasR also occurs in dimeric conformation. However, the dimerization interface does not follow the coiled-coil linker, but is predominantly formed by contacts of the large N-terminal NIT domain. The ANTAR domains are only in contact through their C-terminus, which fits into the cavity formed by the three-helical bundle (see Figure 8) of the opposite monomer. As a consequence, the relative configuration of the two ANTAR monomers of NasR greatly differs from AmiR. The ANTAR regulator Rv1216 contains a RR as its N-terminus that, in contrast to AmiR, comprises the active site residues necessary for phosphorylation, which led to the proposition of a phosphorylation-dependent mode of antitermination [101]. Rv1216 further differs from AmiR regarding the relative configurations of ANTAR and receiver domain due to kink in the linker helix between the N- and C-terminal domains of Rv1626, as well as in its oligomeric configuration being present as a monomer in solution in its inactive state [101].

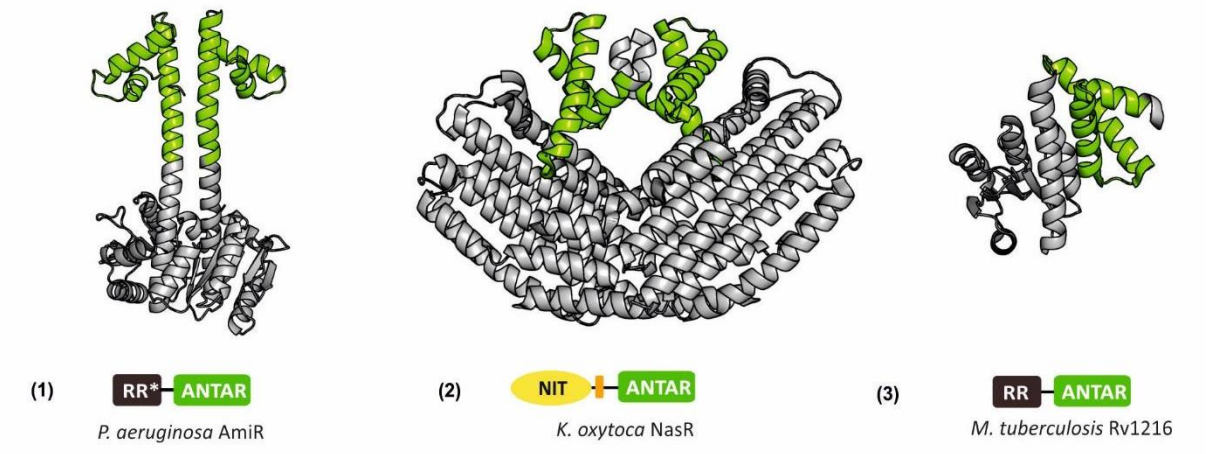


Figure 8: Crystal structures of ANTAR proteins. **(a)** Crystal structures of AmiR ((1); PDB code 1QO0 [102]), NasR ((2); PDB code 4AKK [105]) and Rv1216 ((3); PDB code 1S8N [101]). The ANTAR core domain regions featuring the characteristic three-helical bundle are shown in green.

3.4.3 Mechanism of antitermination

The first characterized and name-giving examples of the ANTAR family, AmiR and NasR, bind to RNA sequences located at the 5’ region of the transcribed AmiE and nasF operons, often called 'leader’ sequence in the literature [96,102,103,107]. Both leader sequences include two characteristic stem-loop structures, comprising an intrinsic terminator including a poly-U tail [108], as well as a proximal hairpin structure essential for antitermination (see Figure 7.c). Association of AmiR and NasR to their target sites within the leader RNAs hinders the formation of a terminator structure, thereby allowing transcription of the downstream genes. The underlying mechanism of antitermination as well as the identification of the structural and sequential features required for ANTAR-association remained unclear until investigations of the *eut* operon revealed similar intrinsic motifs upstream of four different *eut* genes. The characterization of these terminator motifs and the associated regulatory mechanism by the *eut* regulator EutV led to the proposition of common structural and sequence chracteristics for ANTAR recognition [100]. The antiterminatory motif consists of two hairpin structures, P1 and P2. The hairpin structure of P2 overlaps with the 5' end of the transcriptional terminator, which is suppressed upon association with the ANTAR regulator. However, only the functional dimer of EutV formed through signal activation is able to bind the palyndromic structures P1 and P2 simultaneously, thereby stabilizing the P2 antiterminator allowing the RNA polymerase to move on [100,109].

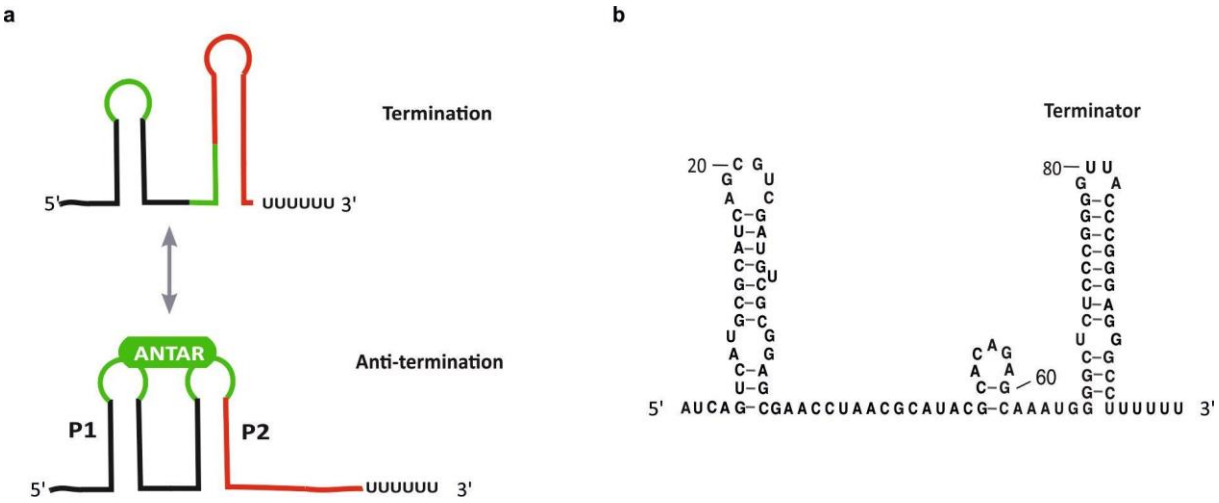


Figure 9: ANTAR mechanism of antitermination. **(a)** Mechanism of antitermination as suggested by Ramesh et al. [95] for the EutV regulator of the *eut* locus in *E. faecalis*. **(b)** Binding motif of the AmiR ANTAR regulator from *P. aeruginos*a as suggested by Wilson et al. [110].

In addition to the autoregulatory role of EutV in the expression of its own operon, the small regulatory RNA EutX was discovered recently, which influences the regulatory effect of EutV by an additional sequestration mechanism [111]. The EutX RNA contains an adenosylcobalamin (AdoCbl) - sensitive riboswitch; the presence of the cofactor AdoCbl prevents the formation of hairpin structures in EutX that provide potential binding sites for EutV. Therefore, in the presence of AdoCbl, EutX cannot bind to EutV, leaving it free to activate expression of the *eut* genes. Conversely, in the absence of AdoCbl, EutX is capable of forming the sequestering hairpin structures, resulting in down-regulation of *eut* gene expression. A similar mechanism has been described for the regulation of the *Eut* operon in *Listeria monocytogenes* [112].

3.5 Objective of research

While many of the photoreceptors used in optogenetic applications mediate control over DNA-associated processes, to my knowledge so far there is no example of a light-inducible effector that directly acts on RNA. Non-coding RNAs are involved in diverse epigenetic regulatory processes at different stages of gene expression, gene imprinting or chromatin remodeling [113]. This work aims to make such RNA-regulated cellular events optogenetically accessible through one of the following strategies: (i) the development of a novel photoreceptor by recombination of a well-characterized LOV domain with a suitable RNA-binding effector domain (Section 5.1) or (ii) the identification and characterization of a naturally occurring photoreceptor that comprises the desired building blocks for light-sensing and RNA-binding function (Section 5.2).

3.5.1 Design of an ANTAR-containing photoreceptor

The modular architecture of photoreceptors opens the possibility of engineering new light-sensitive proteins by rewiring well-characterized photoreceptor domains with interesting cell functions previously not controllable by light. In search of appealing effector functions, we noticed the family of ANTAR proteins that represent bacterial regulatory modules involved in transcriptional regulation [106,110]. The RNA-binding ANTAR modules occur in combination with diverse sensor domains, including many members of the PAS family, of which LOV domains represent a subgroup. One of the best-characterized ANTAR systems is AmiR from *P. aerigunosa*, whose sensor domain structurally resembles the classical RR domains, but lacks the essential residues for phosphoryl group reception. A first approach to subject the RNA-binding function to light control, is to replace the sensor domain of AmiR with the LOV domain from the light-regulated histidine kinase YF1, for which the underlying principles of signal transduction are already partially understood [57,87]. As the engineering process of a new photoreceptor might require the design and testing of numerous variants, a lot of effort should be put into the construction of a suitable high- or at least medium-throughput screening system. If necessary, different strategies may be applied to subject the RNA-binding function of the ANTAR domain to the control of blue light sensitive LOV domains.

3.5.2 Characterization of the natural photoreceptor PAL

Natural photoreceptors often surpass the corresponding engineered variants in terms of specific activity and dynamic range owing to the iterative optimization cycles in the course of evolution. For that reason, the gene databases should be monitored for natural signaling receptors comprising at least one ANTAR and one LOV domain, supposing that the RNA-binding function of the ANTAR

domain could be controlled by the blue light responsive LOV domain in such domain architectures. In case of success, we should confirm the sequence identity and concentrate on the characterization of the respective protein product. The detection of the presumed function, i.e. the ability to induce RNA-interactions by light, may require the identification of specific target RNA sequences. Moreover, the functional photoreceptor should be subjected to structural and mechanistic studies to investigate the underlying principles of signal transduction. This would permit to employ the novel LOV architecture as design template for the development of further light-controllable RNA-binding proteins with altered sequence specificity, e.g. via exchange of the ANTAR domain with the well-characterized archetype of AmiR from *P. aerigunosa*.

4. Materials and methods

4.1 Biological materials and chemical reagents

Unless otherwise stated, all commonly used media and buffers were purchased from Carl Roth GmbH & Co. KG or VWR International GmbH of grade 'Molecular Biology' or higher. All enzymes employed for molecular biology methods originated from Thermo Fisher Scientific Inc. All *Escherichia coli* strains and plasmids applied in molecular biology and protein expression are listed in Table 1 and Table 2. Details to antibiotics and common growth media additives are listed in Table 3.

Table 1: Genotypes of *E. coli* B strains applied in molecular biology and protein expression. DE3 entitles a lysogen that encodes T7 RNA polymerase and lacZ.

Strain	Genotype
DH10b	F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L *nup*G λ-
BL21 (DE3)	F^- *omp*T *hsd*SB(r_B^-, m_B^-) *gal dcm* (DE3)
CmpX13	F^- *omp*T *hsd*SB(r_B^-, m_B^-) *gal dcm* (DE3) *manX::ribM*
Rosetta	F^- *ompT hsdS*B(r_B^-, m_B^-) *gal dcm* (DE3) pRARE (CamR)
Arctic Express (DE3)	F^-*ompT hsdS* (r_B^-, m_B^-) *dcm*$^+$ Tet^r *gal endA* Hte [*cpn10cpn60* $Gent^r$]

Table 2: Overview of vector characteristics of the plasmids applied in molecular biology and protein expression. Kan: Kanamycin; Amp: Ampicilin; Chl: Chloramphenicol; Strep: Streptomycin.

Plasmid	Origin	Selection marker	Application	Manufacturer/ parent plasmid
pET28c	pBR322	Kan	Expression vector	Novagen
pACYC177	p15a	Amp	Cloning vector	NEB
pACYC184	p15a	Chl	Cloning vector	NEB
pCDF-Duet	CloDF13	Strep	Expression vector for mutual synthesis of two genes	Novagen
pBADM30	ColE1	Amp	Expression vector	EMBL Heidelberg
pACYC-BAD	p15a	Chl	Expression vector in reporter assay	pACYC184

Table 3: Working concentrations of antibiotics and growth media additives.

Additive	Working concentration
Ampicillin (Amp)	50 µg ml^{-1}
Kanamycin (Kan)	50 µg ml^{-1}
Gentamycin (Gen)	10 µg ml^{-1}

Chloramphenicol (Chl)	30 µg ml^{-1}
Streptomycin (Str)	100 µg ml^{-1}
Nalidixic acid (NA)	30 µg ml^{-1}
Riboflavin (Rf)	50 µM
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	1 mM
Arabinose (Ara)	0,02 % (v/w)

4.2 Molecular biology

The term 'Molecular biology' or 'cloning' refers to the process of generating recombinant desoxyribonucleic acid (DNA) sequences encoding for a target protein and its transformation into a host organism, for microbial target proteins most commonly represented by *E. coli*. Typically, the DNA is cloned downstream of a promoter in a plasmid expression vector.

During this study, DH10B cells were used for the enrichment of plasmid DNA. To achieve high efficiency transformation, stocks of chemocompetent cells were prepared according to the protocol in 4.2.1. The transfer of target genes into new plasmid vectors was either achieved by restriction cloning (Section 4.2.2) or Gibson assembly (Section 4.2.3). After a successful round of cloning, marked by the appearance of *E. coli* colonies on agar plates containing a plasmid-corresponding selection marker, the extracted plasmids were usually tested on successful insert integration by analytical restriction digests and subsequently sent to sequencing services for final verification.

4.2.1 Preparation of chemocompetent cells and transformation

For the preparation of high efficiency chemocompetent cells, all media and buffers were prepared as detailed in Table 4 and autoclaved before use. A 5 ml overnight culture was inoculated from the respective *E. coli* glycerol stock in liquid LB medium (LB Broth Miller Formulation, Amresco, VWR). Cells were grown at 37 °C and 225 rpm overnight. The main culture was started at a dilution of 1:500 in 750 ml volume of LB and grown for 2 - 4 hours to a final optical density (OD) of 0.6 to 0.8. The cell suspension was transferred to sterile centrifugation tubes and pelleted by centrifugation at 4 °C and 7000 *x* g for 7 min. Cell pellets were resuspended in 50 ml of Tfb1 buffer and chilled on ice for 10 min. After subsequent centrifugation at 7000 *x* g for 7 min and resuspension in 7.5 ml of Tfb2 of cells were aliquoted at 50 µl portions in Eppendorf tubes and immediately shock-frozen in liquid nitrogen. Competent cells were stored at - 80 °C.

Table 4: Preparation details for media and buffers employed for the generation of chemocompetent cells. KAc: Potassium acetate (CH_3COOK); MOPS: 3-(N-morpholino)propanesulfonic acid ($C_7H_{15}NO_4S$).

Solution	Components and concentration	Solvent	pH at 25°C
LB medium	10 g L^{-1} tryptone; 5 g L^{-1} yeast extract; 10 g L^{-1} NaCl	ddH_2O	7.5 (NaOH)
Tfb1	30 mM KAc; 50 mM $MnCl_2$; 100 mM KCl; 15% glycerol	ddH_2O	5.8 (acetic acid)
Tfb2	10 mM MOPS; 75 mM $CaCl_2$; 10 mM KCl; 15% glycerol	ddH_2O	7.0 (NaOH)

Transformation of DNA plasmids was performed via the heatshock method; chemocompetent DH10B cells were thawed on ice and 0.5 - 1 µl of plasmid DNA added to the cells. The cell-DNA suspension was incubated on ice for 10 min. Heat exposure at 42 °C for 60 - 90 sec was followed by 5 min incubation on ice. Then, 0.6 ml of LB medium were added to the cells followed by 1 hour incubation in a thermomixer at 37 °C and 8000 rpm. The cells were then pelleted by centrifugation at 7000 *x* g for 2 min, resuspended in 100 µl LB and plated on LB-agar plates containing an appropriate antibiotic to select for plasmid-containing *E. coli* colonies. These plates were then incubated at 37 °C overnight or for at least 12 hours to allow colonies to show up.

4.2.2 Restriction cloning

Restriction cloning refers to the more traditional way of cloning employing restriction endonucleases for the creation of DNA fragments with specific complementary flanks that can then be joined together with the help of a DNA ligase. These flanks are usually added to the target DNA (insert) by a polymerase chain reaction (PCR) amplification, whereas most cloning vectors already contain multiple cloning sites with a variety of unique restriction sites for the integration of the insert.

A common PCR protocol is shown in Table 5. The plasmid DNA and PCR products were usually digested for 5 – 30 min with the respective restriction enzymes (FastDigest, Thermo Fisher Scientific) in the according buffer (see Table 6.a). The plasmid backbone was subsequently dephosphorylated by addition of 1 µl (1U/µl) of alkaline phosphatase (FastAP Thermosensitive Alkaline Phosphatase, Thermo Fisher Scientific) into the same buffer and incubated for 10 min at 37 °C to avoid self-ligation. The digested plasmid and insert were checked via agarose-gel electrophoresis and purified by gel extraction or PCR-clean up (NucleoSpin Extract II, Machery-Nagel GmbH & Co. KG) before ligation.

The ligation of plasmids and insert was conducted at molar ratios ranging from 1:3 to 1:10 by adding a T4 DNA Ligase in the according buffer (Thermo Fisher Scientific) and incubation for at least 20 min at ambient temperature (see Table 6.b).

The reaction mixture was then transformed into chemically competent *E. coli* DH10B cells via the heat shock method, as described under 4.2.1. For plasmid preparation, 5 ml cultures were inoculated with single colonies and grown for at least 16 hours at 37 °C. The plasmid DNA was purified using the NucleoSpin Plasmid MiniPrep kit (Machery-Nagel) and checked by analytical restriction digests and sequencing (GATC Biotech AG, LGC Genomics, or Eurofins Genomics GmbH).

Table 5: Standard PCR protocol. (a) Composition of a Phusion PCR reaction mixture. (b) Thermo Cycler settings.

a

Component	Amount
Template DNA	≈60 ng
dNTPs (10 mM)	1 µl
5 x HF Buffer	10 µl
Fw Primer (10 µM)	1.5 µl
Rv Primer (10 µM)	1.5 µl
Phusion polymerase (2U/µl)	1 µl
ddH_2O	to 50 µl

b

Temperature	Time	
98°C	30 s	
98°C	30 s	35 cycles
T_m -5 °C	30 s	
72°C	20 s/ kb	
72°C	10 min	
10°C	∞	

Table 6: Fast Digestion and Ligation of DNA. (a) Compositon of a FastDigest (FD) reaction mixture. After gentle mixing the mixture was incubated at 37°C for 5 – 30 min. (b) Compositon of an ensuing Ligation reaction. After gentle mixing the mixture was incubated at room temperature for 20 - 60 min.

a

Component	Plasmid DNA	PCR product
ddH_20	15 µl	16 µl
10 x FD buffer	2 µl	3 µl
DNA	2 µl (≈0.1 µg)	10 µl (≈0.2 µg)
FD Enzyme (10 U/µl)	1 µl	1 µl
Total volume	20 µl	30 µl

b

Component	Amount
Plasmid DNA	40 – 100 ng
Insert	variable
5 x T4 Ligase buffer	2 µl
T4 DNA Ligase (1 U/µl)	1 µl
ddH_20	to 10 µl

4.2.3 Gibson cloning

Gibson cloning presents an alternative way of cloning developed by Daniel Gibson and colleagues at the Craig Venter Institute [114]. The method has the advantage that multiple overlapping DNA fragments can be joined in one single reaction called 'Gibson Assembly' (GA) without the need for specific restriction sites and the resulting 'scar' between insert and plasmid. The reaction employs three different enzymes: a T5 Exonuclease that creates single-strand DNA 3' overhangs, which allows the complementary overhang fragments to anneal; Phusion DNA Polymerase that fills in the gaps within the annealed fragments; and a Taq DNA Ligase that covalently joins the annealed fragments of DNA.

In a first step, the target genes and plasmid backbone were amplified via PCR as described above (Table 5). This step allows the addition of an overlapping region to the DNA fragments to be ligated; therefore, the oligonucleotide primers have to feature overlaps of around 20 nucleotides (nt) to the fragments to be fused, and to the insertion site within the plasmid. The PCR fragments were either directly fused via GA or, in case of multiple bands on agarose gel test, previously purified via agarose gel extraction. For the GA reaction, a master mixture was prepared as detailed in Table 7.b containing a 5 x reaction buffer (see Table 7.a) and the three key enzymes: T5 exonuclease (Epicentre), Phusion polymerase (NEB - New England Biolabs GmbH) and Taq ligase (NEB). Aliquots of 15 µl of the GA master mixture were stored at – 20 °C. To start a GA reaction, the aliquots were thawed on ice and equimolar amounts of insert and vector DNA were added to a total volume of 20 µl. In case of considerably smaller inserts, the insert was added in up to 5- fold molar excess. The reaction mixture was incubated at 50°C for 1 hour and afterwards cooled down to room temperature for 15 min. The reaction mixture was then placed on ice until transformation; for that, 5 µl o f the reaction mixture were added to chemically competent *E. coli* DH10B via the heat shock method, as described in Section 4.2.1.

Table 7: Gibson Assembly protocol. (a) Composition of the 5 x reaction buffer and (b) the GA master mixture.

a

Component	Amount
Tris-HCl pH 7.5	500 mM
$MgCl_2$	50 mM
DTT	50 mM
PEG-8000	25 %
dNTP	1 mM each
NAD	5 mM

b

Component	Amount
5 x reaction buffer	4 µl
T5 exonuclease (10 U/µl)	0.64 µl
Phusion polymerase (2 U/µl)	20 µl
Taq ligase (40 U/µl)	160 µl
ddH_2O	to 1.2 ml

4.3 Screening of engineered variants

4.3.1 Miller assay

AmiR-related chimeric constructs were functionally tested using an enzymatic reporter assay in *E. coli*. To that end, AmiR activity was coupled to a β-galactosidase (β-gal) read-out through fusion of the lacZ gene behind a transcriptional terminator targeted by AmiR [110] on a pet28c backbone. Expression of the chimeric library was put under control of the arabinose-dependent BAD-promotor on a pACYC184 backbone of compatible P15A origin. The detection is based on the 'A better Miller' protocol from 'openwetware.org' [115], which in turn was derived from a protocol of Zhang and Bremer [116]. The assay employs o-nitrophenyl-β-D-galactopyranoside (ONPG) as artificial substrate which is cleaved to yield galactose and o-nitrophenol (ONP) which has a yellow color with an absorption maximum at 420 nm. Therefore, β-gal activity can be measured in a spectrometer by the rate of appearance of yellow color, termed 'Miller Unit' (MU) in standardized amounts, defined as:

$$\mathrm{MU} = 1000\,\frac{\mathrm{OD}_{420} - 1.75 \times \mathrm{OD}_{550}}{t \times v \times \mathrm{OD}_{600}}, \qquad (1)$$

where OD_{420} measures the absorbance of the β-gal breakdown product ONP, and OD_{550} determines the scatter from cell debris when multiplied by 1.75. In the bottom of the fraction, v stands for the volume of cells used in mL at a cell density given by OD_{600} and t is the reaction time in min. To compare constructs of different nature it is advisable to eliminate the time variable within equation (1). For the interval during which the production of ONP is proportional to the concentration of β-gal, OD_{420} and t are linearly correlated (OD_{420} = slope x t), so that t can be substituted by OD_{420}/ slope.

Table 8: Preparation details for reagents required for Miller assay. Measurements in Permeabilization solution (PS) and Substrate solution (SS) are sufficient for one 96 well plate (PS requires 40 µl/ well, 96 x 40 µl = 3.84 mL; SS requires 250 µl/ well, 96 x 250 µl = 24 ml).

Solution	Components, final concentration/ quatity and solvent
Stock solutions	500 mM dibasic sodium phosphate (Na_2HPO_4), 1M NaH_2PO_4, 2 M KCl, 1 M $MgSO_4$, 1 % CTAB (hexadecyltrimethylammonium bromide), 1 % sodium deoxycholate ; all in ddH_2O
Permeabilization solution	0.8 ml 500 mM Na_2HPO_4, 20 µl 4 M KCl, 8 µL 1 M $MgSO_4$, 240 µl 1 % CTAB, 160 µl 1 % sodium deoxycholate, 21.6 µl β-mercaptoethanol; 2.75 ml ddH_2O to 4ml
Substrate solution	3.6 ml 500 mM Na_2HPO_4, 1.2 ml 1M NaH_2PO_4, 30 mg ONPG, 81 µl β-mercaptoethanol; 25.2 ml H_2O to 30 ml

All stock solutions required for the assay were previously prepared as detailed in Table 8. Permeabilization solution and substrate solution were freshly prepared the day of assay conduction.

Usually each construct was tested for dark- and light-activity, so that one plate for each condition was prepared. Cells were inoculated with 4 µL of overnight-grown precultures and grown for 3 – 4 hours at 840 rpm. For light-activity testing, the incubator was equipped with a custum-built LED-array of 10 x 8 LEDs of 470 ± 10 nm (Winger Electronics GmbH & Co. KG) at 50 µW/ cm². Light intensities were determined using a power meter (model 842-PE; Newport) and a silicon photodetector (model 918DUV- OD3; Newport). During growth, Permeabilization solution and Substrate solution were prepared as detailed in Table 8, and 40 µL aliquots of Permeabilization solution were pre-measured into a transparent 96 well microplate ('detection plate') and covered to reduce evaporation. At the end of the growth time, 300 µL of the cells were transferred into a transparent 96 well microplate ('absorbance plate') and the OD_{600} was measured immediately using the Tecan Infinite M200 PRO plate reader (Tecan Group Ltd.). After that, 10 µL of each well of the absorbance plate were transferred to the detection plate containing the Permeabilization solution and thoroughly mixed by pipetting up and down, which stabilizes the samples for several hours allowing the conduction of time-course experiments if required. A blank (solutions only) sample was always included for subtracting the background absorbance at a later point. To start the detection reaction, 250 µL of Substrate solution were added to each well of the detection plate and the development of OD_{420} and OD_{550} was followed over 10 - 90 min with the Tecan M200 PRO plate reader. Unless otherwise stated, the presented data show average values of three biological replicates ± standard deviation (SD).

4.3.2 Fluorescence-based detection

For assessment of AmiR-related chimera activity via fluorescence-based detection, the lacZ sequence of the reporter plasmid was replaced by the DsRed2 gene and reporter plasmid and chimera variants were co-transformed into competent CmpX13 cells. For functional screening, the plates were and incubated for 18 hours at 37°C and 800 rpm in the dark or under blue light (470 nm, 50 µW cm^{-2}) using custom-built arrays of 10 x 8 LEDs of 470 ± 10 nm (Winger Electronics). For analysis of the samples, the cultures were diluted 20-fold with distilled water in black-walled 96-well µClear plates (Greiner BioOne). DsRed fluorescence (excitation 554 ± 9 nm, emission 591 ± 20 nm) and OD_{600} were determined using a Tecan Infinite M200 PRO plate reader (Tecan Group Ltd).

4.4 Recombinant production of proteins

Recombinant production of proteins is one of the most essential techniques in biological science, since the generation of pure and abundant amount of a target protein provides the basis for all further *in vitro* characterization plans and intended biotechnological purposes. Once the DNA encoding for a target protein is integrated into an appropriate expression vector, the cell's protein synthesis machinery can be used to generate the target protein.

4.4.1 Expression

For all target proteins derived from *Nakamurella multipartita*, the *E. coli* strain '*Arctic Express (DE3)*' (Agilent) resulted in the highest gain of soluble and pure protein; e.g., in case of the PAL protein, 800 ml culture could yield around 1.5 mg of protein.

For protein expression in Arctic Express (DE3), 5 ml cultures were inoculated with cells harboring the respective expression plasmid the previous day and grown at 37 °C at 225 rpm in an incubator shaker (Innova 44R; New Brunswick Scientific GmbH) overnight. For growth of these precultures, the culture media had to contain both gentamycin and the respective selection marker of the expression plasmid. Next day, 800 ml media were inoculated 1 : 200 and grown at 30 °C at 225 rpm for 3 – 4 hours up to an OD_{600} of 0.7 – 0.9. At that OD, the incubator was cooled down to 12 °C and protein expression was induced with 1mM IPTG. The cells were incubated for a minimum of 48 hours or up to a weekend period of approximately 72 hours before harvest.

4.4.2 Extra precautions for work with photoreceptors

Whenever the protein construct contained a light- sensitive LOV domain, extra precautions were paid from the moment of extracting the protein from the host cell culture to avoid light exposure and excessive switching between dark and signaling state of the photoreceptor. Therefore, all work after this point was conducted under red-light conditions, at which the absorption ability of LOV domains is minimal.

4.4.3 Purification

For protein purification, immobilized metal ion affinity chromatography (IMAC) was applied employing 1 ml Protino Ni-NTA Columns (Machery Nagel) for histidine-tagged target proteins within an Äkta Prime Fast Protein Liquid Chromatography (FPLC) System (GE Healthcare Europe GmbH). All

required buffers were prepared beforehand as detailed in Table 9 and filtered using 20 µm pore size nylon membrane filters (Merck KGaA).

Table 9: Preparation details for buffers used for protein purification and preparation.

Solution	Purpuse	Components and concentration	Solvent	pH at 25°C
Buffer A	Resuspension	50 mM Tris/HCl, 1M NaCl, 20 mM imidalzole	ddH_2O	8.0 (HCl)
Buffer B	Washing	50 mM Tris/HCL, 250 mM Arg, 5% isopropanol	ddH_2O	8.0 (HCl)
Buffer C	Elution	50 mM Tris/HCL, 200 mM NaCl, 200 mM imidazole	ddH_2O	8.0 (HCl)
Buffer D	Dialysis	1 x ICB w/ 10% glycerol	ddH_20	7.5 (NaOH)
10 x ICB	Buffer D stock	120 mM HEPES, 1350 mM KCl, 100 mM NaCl, 10 mM $MgCl_2$	ddH_20	7.7 (NaOH)

Cells were harvested by centrifugation using two beakers and spinned for 10 min at 7000 rpm in a JA10 rotor (Beckman Coulter GmbH). The pellets were resuspended in 2 x 7.5 ml of buffer A supplemented with Roche protease inhibitor mix (1 pill in 40 ml; Roche Diagnostics Germany). For cell lysis, cells were sonicated on ice alternating between 30 sec of sonification and 30 sec rest for a total of four times (total net sonification time 2 min) at power settings of 40 % – 50 % and constant duty cycle. To extract the soluble cell fraction, the cell lysate was centrifuged for 30 min at 18000 rpm using a JA25.5 rotor (Beckman Coulter). From now on, all work was done under red light conditions if the protein to be purified contained a blue-light-sensitive LOV domain. The supernatant was transferred to a falcon tube and kept on ice prior to the start of the FPLC purification process. Before applying the soluble cell fraction containing the target protein (from now on 'sample'), the FPLC system was prepared and washed with the respective buffers and the column equilibrated with 5 column volumes of buffer A. The flow rate was set to 1 ml/ min and the sample loaded onto the column through a sample tube. After load, the column was washed with buffer A until the ultraviolet (UV) signal got back to baseline level, followed by a second wash with buffer B for 10 min. Before elution, the column was equilibrated with buffer A for 5 min. Elution was initiated by setting buffer C to 100 % ; the samples were collected in 1 ml fractions for 10 min or until the UV signal got back to baseline. The sample fractions were analyzed on 15 % sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie blue [117]. Fractions of highest purity were pooled and dialyzed against 2 x 500 ml of buffer D. The protein concentration was determined via absorption spectroscopy using an extinction coefficient (ε) characteristic for LOV photosensors (ε_{447}= 12500 cm^{-1} M^{-1}).

4.5 Absorption spectroscopy

4.5.1 UV-vis spectroscopy

UV-Vis spectroscopy is based on the fact that most molecules can absorb some of light in the ultraviolet and visible (200 – 700 nm) range of the electromagnetic spectrum by promoting electrons to higher energy levels. As the energy of the absorbed light must match the energy required for the electronic transition, not all wavelengths of light are absorbed equally by a sample. This results in a characteristic absorbance spectrum that, in case of photoreceptors, matches their specific activation range, which promotes the transfer to the signaling state. Furthermore, the absorbance of a molecule varies in linear fashion with its concentration (Beer-Lambert Law), a fact that is widely used to determine the concentration of a sample in biochemistry.

During this study, all UV-vis spectra were recorded using an Agilent 8453 spectrophotometer (PDA; Agilent Technologies) and corresponding software. All spectra were corrected for the baseline signal within the range of 700 – 800 nm.

4.5.2 Dark state recovery kinetics

Dark state recovery kinetics of the LOV domain-containing variants were measured at protein concentrations of 30 -60 µM in buffer D with the same Agilent 8453 spectrophotometer at varying temperatures. The samples were illuminated with a with a blue light emitting diode (LED) (λ_{max} = 450 nm, 200 mW cm^{-2}) for 30 sec and the absorbance was followed at 447 nm. For evaluation of the data, two independent measurements were averaged and fitted with exponential functions using the GraphPad Prism 5.0 (GraphPad Software,USA). The spectra were corrected for the baseline signal between 700 and 800 nm.

4.5.3 Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy examines the absorption of circularly polarized light, which differs in the direction of its circular polarization. In optically active chiral molecules this property is differently pronounced, therefore it is widely used in biochemistry to investigate structural aspects of proteins and nucleic acids. Within proteins, the relative fractions of secondary structure content can be determined within the far-UV (190 – 250 nm), as ordered α-helices, β-sheets and β-turns all show characteristic spectra. Moreover, the peptide bonds are optically active and the ellipticity of their polarization changes in line with their local conformation. Therefore, CD spectroscopy is widely used to demonstrate conformational changes in proteins.

All shown CD spectra were recorded with a Jasco Ja-710 Spectropolarimeter (Jasco Deutschland GmbH) and respective software. Prior to the analysis, the protein samples was were dialyzed into 100 mM sodium phosphate buffer, with 150 mM NaCl, pH 7.5. For the recording of difference spectra of signaling and dark-adapted state, samples were illuminated with a blue LED (λ_{max} = 450 nm, 200 mW cm^{-2}). All spectra were corrected for the buffer-caused background signal along the measured range.

4.6 *In vitro* transcription of RNA

In vitro transcription is a standard method for laboratory synthesis of ribonucleic acid (RNA) molecules from short oligonucleotides up to several kilobases (kb). The technique requires no more than a DNA template containing a T7 promoter as well as the purified T7 RNA polymerase enzyme providing a simple way of producing up to several mg of high quality transcripts for manifold follow-up studies.

4.6.1 General precautions for work with RNA

Because of the ubiquitous presence of RNAses as well as the higher chemical instability of RNA, working with RNA is more demanding than the work with DNA. Moreover, RNAses do not require metal ion co-factors as DNAses and are highly heat-resistant even to prolonged autoclaving. For that reason the work with RNA requires special precautions, for which the most important aspects are summarized in Table 10. Diethyl pyrocarbonate (DEPC) -treated H_2O was prepared by adding DEPC up to 0.05 % and following incubation overnight at room temperature. The treated solution was then autoclaved for 30 min to remove traces of DEPC.

Table 10: General precaution guidelines for work with RNA.

Aspect	Specification
Contact / Gloves	Wear gloves all the time when working with RNA.
Working surfaces	Before work, rinse surface of bench first with ddH_2O, then with 100 % EtOH. Cover with 3 % of H_2O_2 and let soak for 10 min. Rinse again with DEPC-treated ddH_2O before use.
Equipment	Use sterile, disposable plastic ware preferred to glass ware. Bake glassware at 200 °C for at least 4 hours. Electrophoresis tanks should be cleaned with 1 % SDS and then rinsed with ddH_2O and 100 % EtOH. They should then be filled with 3 % of H_2O_2 and let soak for 10 min. Rinse again with DEPC-treated ddH_2O before use.
Reagents	Only use DEPC-treated H_20 for the preparation of buffers and reagents. Reserve separate reagents for RNA work only. Remember that TRIS buffers cannot be used with DEPC.
Manipulation	Always keep the RNA sample as well as working reagents on ice when preparing an

	experiment. Put RNA stocks back to freezer as soon as possible.
Storage	Store all RNA samples at -80 °C. Always thaw on ice!

4.6.2 High yield transcription

In preparation of the following transcription step, the template DNA had to be amplified via PCR in sufficient amounts (2 µg for transcripts < 100 nt) as described in Section 4.2.2 and Table 5. The PCR product was checked via agarose-gel electrophoresis and purified by gel extraction or PCR-clean up (NucleoSpin Extract II, Machery-Nagel) before the next step.

For the *in vitro* transcription, the TranscriptAid T7 High Yield Kit (Thermo Fisher Scientific) was used. To start the transcription reaction, all components prepared according to manufacturer's instructions and combined in a single RNase-free microfuge tube as shown in Table 11.

Table 11: Preparation details for the TranscriptAid T7 High Yield transcription reaction.

Component	Amount
DEPC-treated Water	to 20 µl
5X TranscriptAid Reaction Buffer	4 µl
ATP/CTP/GTP/UTP mix, 25 mM	6 µl
Template DNA *	2 µg
TranscriptAid Enzyme Mix (1 U/µl)	2 µl

All components were thoroughly mixed, briefly spinned and incubated at 37°C for 4 – 8 hours as recommended for short (< 100 nt) transcripts. For digestion of the DNA template, 2 units of DNase I (DNAse I, RNAse-free; Thermo Fisher Scientific) per 1 µg of template DNA were added directly to the transcription reaction mixture and incubated at 37°C for 30 min. DNAse I was removed by purification with NucleoSpin Extract II Kit (Machery-Nagel) using NucleoSpin RNA II Binding buffer (instead of Binding Buffer NTI used for DNA purification) and DEPC-treated H_2O for elution. The RNA concentration was then determined spectroscopically in the Tecan Spark 10M Reader (Tecan Group Ltd). As the RNA probes were usually subsequently treated with site-specific terminal labeling methods, dephosphorylation of the transcripts was ensued directly employing FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific) in the appropriate buffer (10 x FastAP Buffer; Thermo Fisher Scientific) at 37 °C for 15 min and subsequent heat inactivation at 75°C for 5 min.

4.7 Electrophoretic Mobility Shift Assay

Electrophoretic Mobility Shift Assays (EMSAs) is a method for detecting interactions between proteins and nucleic acids. It takes advantage of the different mobility of free nucleic acids compared to nucleic acids bound in protein complexes in electrophoretic gels [118]. Radioactive labels are frequently used for detection, as their high sensitivity allows the detection of lowest nucleic acid or protein concentrations.

4.7.1 Radiolabeling of RNA

Isotope-labeling of the RNA-probes was generally done by 5'-end labeling employing [γ-^{33}P] –ATP (Hartmann Analytic GmbH, Braunschweig, Germany). To start the labeling reaction, all components were combined on ice as shown in Table 12 in a single RNase-free microfuge tube and incubated at 37 °C for 1 hour.

Table 12: Preparation details for the 5'-end labeling reaction employing [γ-^{33}P] –ATP.

Component	Amount
DEPC-treated Water	to 20 µl
Dephosphorylated RNA probe	20 pmol
[γ-^{33}P] –ATP	5 µl (20 pmol)
10 x T4 PNK Buffer A	2 µl
T4 Polynucleotide Kinase (10 U/ml)	1 µl

The reaction mixture was purified purification by spin-column chromatography using Ambion NucAway Spin Columns (Thermo Fisher Scientific) to remove the free radiolabeled nucleotides.

4.7.2 EMSA assay

All EMSA assays in this study were conducted in 1 x Intracellular buffer (ICB) (12 mM HEPES, 135 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 10% glycerol, pH 7.5; 'buffer D', see Table 9) suitable for loading on polyacrylamide gels. The radioactively labeled RNA stock was first diluted to a concentration of 30 nM, then several protein-RNA mixtures were prepared with different proportions of the purified protein variants and a final RNA concentration of 0.5 nM in a total volume of 30 µl each. The protein-RNA mixtures were incubated for 20 min at ambient temperature in the dark, after which a first fraction of the protein-RNA preparations was loaded onto a 6% polyacrylamide tris-borate-EDTA

(TBE) gel, made of Acrylamide/Bis 37.5:1, 40% (w/v) solution (Merck) and 1 x TBE run buffer, supplemented with 2 mM $MgCl_2$. Gel runs were performed in the dark with a precooled (4 °C) 0.5 x TBE running buffer (later also supplemented with 2 mM $MgCl_2$) for 30 min (15 V cm-1). Meanwhile, the remaining protein-RNA mixtures were illuminated for 20 minutes with a blue light array (≈ 100 mW cm^{-2} at 447 nM) and subsequently likewise separated on a 6% polyacrylamide gel under continuous blue light illumination. Gels were then dried on Whatman blotting paper (GE Healthcare) in a slab gel dryer (GD2000; Hoefer) connected to a vacuum system and exposed to a film (BioMax MR; Kodak) for at least 12 hours. Autoradiographies were digitized using an FLA-7000 phosphorimager (Fuji Film Europe).

4.8 Systematic Evolution of Ligands by Exponential Enrichment

Systematic Evolution of Ligands by Exponential Enrichment (SELEX) is a method for identifying specific DNA or RNA aptamers for different types of target ligands, including proteins [119]. The process uses a library of random oligonucleotide sequences, flanked by constant 5'- and 3'-termini that serve as primer annealing sites. The library sequences are incubated together with the target ligand, and the none-binding or only weakly aptamers are removed by a separation or washing step, for which the stringency increases with an increasing number of selection cycles. In a last step, the bound sequences are eluted and amplified by either PCR or reverse transcription (RT) PCR for DNA or RNA aptamers, respectively, to prepare for subsequent selection cycles. Within this study, the procedure was intended for the identification of RNA aptamers, so that special precaution was paid to avoid RNA degradation (as detailed in Table 10) during the whole proceeding. The SELEX experiments were performed in collaboration with the Günter Mayer group of the University of Bonn (Bonn, Germany) and special help of Anna Maria Pyka.

4.8.1 Preparation

All buffer and reagents, as well as the master mixtures for RT-PCR and transcription reaction were prepared in advance according to the preparation details shown in Table 13 and Table 14.

Table 13: Preparation details for media and buffers employed for SELEX.

Solution	Components and concentration	Solvent	pH at 25 °C
10 x intracellular buffer (ICB)	120 mM HEPES, 1350 mM KCl; 100 mM NaCl; 10 mM $MgCl_2$	DEPC-tr. ddH_2O	7.7 (NaOH)

Quenching buffer	200 mM Tris	DEPC-tr. dd H_2O	7.9 (HCl)
$MgCl_2$ stock	100 mM $MgCl_2$	DEPC-tr. ddH_2O	/
1 x ICB w/ glycerol	10 % of 10 x ICB, 10 % glycerol	DEPC-tr. ddH_2O	auto-adjusts to
1.25 x ICB w/ glycerol	12.5 % of 10 x ICB, 12.5 % glycerol	DEPC-tr. dd H_2O	7.5 from 10 x ICB

Table 14: Preparation details for master mixtures for 30 reactions (rcts) for (a) RT-PCR and (b) transcription.

a

Reagent	Amount for 30 rcts
5x Colorless GoTaq Flexi Buffer (Promega)	600 µl
5 x first strand buffer (Invitrogen)	120 µl
100 mM DTT	60 µl
100 µM Fw primer	30 µl
100 µM Rv primer	30 µl
25 mM MgCl2	180 µl
25 mM (each) dNTPs	36 µl
Total volume	1056 µl
Single aliquot volume	35.2 µl

b

Reagent	Amount for 30 rcts
5x Transcription buffer (200 mMTris, pH 7.9)	600 µl
100 mM DTT	150 µl
NTP-Mix (25 mM each)	300 µl
100 mM MgCl2	450 µl
MilliQ	1065 µl
Total volume	2565 µl
Single aliquot volume	85.5 µl

a

5’-GGGGGAATTCTAATACGACTCACTATAGGGAGGACGATGCGG - 40 N - CAGACGACTCGCTGAGGATCCGAGA-3’

3’- CCCCCTTAAGATTATGCTGAGTGATATCCCTCCTGCTACGCC - 40 N - GTCTGCTGAGCGACTCCTAGGCTCT-5’

b

Fw primer: 5’-GGGGGAATTCTAATACGACTCACTATAGGGAGGACGATGCGG-3’

Rv primer: 5’-TCTCGGATCCTCAGCGAGTCGTC-3’

c

5’-GGGAGGACGAUGCGG – 40 N - CAGACGACUCGCUGAGGAUCCGAGA-3’

Figure 10: Sequence details of the Sul I pool showing **(a)** the DNA template comprising 40 random nucleotides (’40 N’) and constant 5’- and 3’-termini, **(b)** the primer sequences for transcription and RT-PCR of the pool and **(c)** the resulting ssRNA transcript. The red colored Section marks the T7 promoter site for transcription initiation with T7 RNA polymerase

The random oligonucleotide pool 'Sul I' comprising 40 random nucleotides ('40 N') and constant 5'- and 3'-termini (see Figure 10.a) was used as template for transcription of the first SELEX round. Transcription of the template pool with the primer sequences shown in Figure 10.b results in a single-strand RNA transcript of a total of 80 nt length (see Figure 10.c).

For convenience, the PAL protein was immobilized to streptavidin-coated microplates (Streptavidin Coated HBC Black 96-Well Plates, Thermo Fisher Scientific) via biotinylation of primary amino groups with a commercial Sulfo-NHS biotinylation reagent (EZ-Link Sulfo-NHS-LC-Biotin, Thermo Fisher Scientific). Prior to the start of the first SELEX cycle, sufficient amounts of the protein should be biotinylated and aliquoted to guarantee consistent testing conditions over the course of the SELEX.

4.8.2 Biotinylation

For one coupling reaction, 10 nmol of the protein solution were mixed with 4 µL of Sulfo-NHS-LC-Biotin in a total volume of 100 µl and incubated for 2 hours at 4 °C. The remaining NHS-Biotin was quenched with 20 µL of 200 mM Tris-HCl, pH 7.9, for 10 min on ice. The reaction mixture was then purified using Zebra Spin Desalting Columns (Thermo Fisher Scientific) allowing purification of up to 700 µl sample volume following the manufacturer's instructions. Biotinylation rate and potential sample loss during purification were controlled with the help of Dot Blot analysis and SDS-PAGE with subsequent Coomassie staining [117]. The purified biotinylated protein was fractionated into aliquots and stored at -80 °C.

4.8.3 Protein immobilization

From here on, attention was payed to conduct all work under dark or red light conditions. For irreversible coupling of the protein to the streptavidin-coated microplates, the wells were first prepared by washing each well three times with 200 µl 1 x ICB . After that, 100 µl of the biotinylated protein were added to each well and incubated over night at 4 °C with slow shaking. To remove the non-bound protein, each well was washed three times with 200 µl 1x ICB. The protein-coupled microplates are stable at 4°C for several hours during preparation and testing of the transcripts for start of the next SELEX cycle.

4.8.4 Transcription

For the 1st cycle, 10 µl of the Sul I-pool DNA were added to the previously prepared transcription master mixture. In the following rounds, the Sul I-pool template was replaced by 10 µl of the product

from the RT-PCR reaction of the previous SELEX round. In addition, 1.24 µl RNasin (Recombinant RNasin Ribonuclease Inhibitor, Promega GmbH) and 3.3 µl T7 RNA Polymerase (30 U/µl, homemade) were added to the reaction mixture. A negative control containing ddH_2O instead of the template was sustained during all SELEX cycles. The reaction was performed for 20 min by transferring the mixture to the thermomixer preheated to 37°C. Samples were checked on a 2.5 % agarose gel. The transcription product was then directly used for the next SELEX cycle, and the remnant stored at -80 °C.

4.8.5 SELEX selection

For the first selection cycle, 1 nmol of the purified Sul I RNA Mix were mixed with 160 µl 1.25 x ICB, 28 µl water and 2µl of 100 mM $MgCl_2$ to a final volume of 200 µl. The addition of $MgCl_2$ in this step results in a final $MgCl_2$ concentration of 1 mM; attention was payed to maintain this concentration constant for the entire SELEX procedure, as even small changes in $MgCl_2$ concentration can have large impacts on aptamer secondary structure. In the following rounds, 20 µl of the transcription product were combined with 40 µl water and 240 µl of 1 x ICB up to a final volume of 300 µl. As the diluted transcription product already displays a $MgCl_2$ concentration of 1 mM, no further $MgCl_2$ is required at this point. The selection mixture was added to the immobilized protein and incubated under blue light irradiation for 30 min at 25 °C in a thermomixer and slow shaking (300 rpm). After that, washing was carried out under blue light irradiation at 25 °C as indicated in Table 15. After washing, the RNA was eluted by incubating the sample for 30 min at 25 °C in the dark. The supernatant was transferred to a new vial. The eluate was then used for the RT-PCR in the next step.

Table 15: Instructions for SELEX washing protocol. **(a)** Overview of a complete washing cycle; step 4 – 7 embody one subcycle. **(b)** The number of subcycles increases in the following rounds by one more subcycle at every second SELEX round up to a maximum of 4 washing subcycles.

a	Single steps of one washing cycle
	(1) Remove supernatant.
	(2) Resuspend wells with 200 µl 1x ICB.
	(3) Incubate sample for 30 seconds.
	(4) Remove and discard the supernatant.
	(5) Resuspend wells with 200µl 1x ICB.
	(6) Incubate sample for 3 minutes.
	(7) Remove and discard the supernatant.

b	Number of subcycles for increasing rounds
1st Round:	One washing subcycle.
3rd Round:	Two washing subcycles.
5th Round:	Three washing subcycles.
7th Round:	Four washing subcycles
9th Round:	Four washing subcycles.

4.8.6 RT-PCR

For RT of the eluated transcripts, 62.8 µl of the SELEX eluate or 62.8 µl water as negative control were combined in 0.5ml PCR tubes and incubated at 65 °C for 5 min. After cooling down to 4 °C, 1 µl Superscript II Reverse Transcriptase (Thermo Fisher Scientific) and 1 µl GoTaq Flexi DNA polymerase (Promega) were added and the RT reaction initiated according to the protocol described in Table 16.

Table 16: PCR-program for Reverse transcription (RT) reaction.

Description	Temperatur	Time	Number of cycles
RT step	54 °C	10 min	
Initial denaturation	95 °C	1 min	1st round: 4 cycles
Denaturation	95°C	30 sec	> Given the strong decrease of
Annealing	60°C	30 sec	RNA during each round, the
Elongation	72°C	30 sec	number of cycles has to be
Final Elongation	72°C	1 min	strongly increased (check band
Description	4°C	∞	intensity on agarose gel!)

The 5 µl product of the RT-PCR reaction + 1 µL 6 x DNA loading dye were separated on a 2.5 % agarose gel for validation. If the band intensity was too weak (less intense than the 50 bp band), the RT-PCR was re-run for an appropriate number of cycles (avoid over-amplification) and re-checked on a 2.5 % agarose gel.

The Sections 4.8.3 to 4.8.6 form an entire SELEX cycle, which were repeated 9 times in the first SELEX experiment. After the 9th cycle, a filter binding assay was performed by Anna-Maria Pyka (University of Bonn) to validate the enrichment of the selected transcripts; after positive validation, a fraction of 45 randomly picked clones of the RT-transcribed pool from the 9th cycle was analyzed via DNA sequenzing.

4.9 Ribogreen Assay

The Ribogreen Assay presents an alternative way to detect protein-nucleic acid interactions. As during SELEX, the protein is immobilized to streptavidin-coated microplates (Streptavidin Coated HBC Black 96-Well Plates, Thermo Scientific) via biotinylation of primary amino groups with a commercial Sulfo-NHS biotinylation reagent (EZ-Link Sulfo-NHS-LC-Biotin, Thermo Scientific). The RNA target is added in varying concentrations. Unbound fractions can be separated by washing, whereas the

bound fractions are detected employing the fluorescent Ribogreen RNA Reagent (Quant-iT Ribogreen RNA Reagent , Thermo Scientific).

As for all RNA works, special precaution was payed to avoid RNA degradation, as described in Table 10. The method was adopted from the Mayer group of the University of Bonn (Bonn, Germany). All buffer and reagents were prepared in advance according to the preparation details shown in Table 17.

Table 17: Preparation details for reagents and buffers required for the proceeding of the Ribogreen assay.

Solution	Components and concentration	Solvent	pH at 25 °C
10 x intracellular buffer (ICB)	120 mM HEPES, 1350 mM KCl; 100 mM NaCl; 10 mM MgCl2	DEPC-tr. ddH_2O	7.7 (NaOH)
10 x TE	100 mM TRIS; 10 mM EDTA	DEPC-tr. dd H_2O	7.5 (HCl)
1 x ICB w/ glycerol	10 % of 10 x ICB, 10 % glycerol	DEPC-tr. ddH_2O	auto-adjusts to
2 x ICB w/ glycerol	20 % of 10 x ICB, 20 % glycerol	DEPC-tr. dd H_2O	7.5 from 10x ICB

4.9.1 Protein immobilization

Attention was payed to conduct all work under dark or red light conditions. For irreversible coupling of the protein to the streptavidin-coated microplates, the wells were treated as previously described in Section 4.8.3; each well was washed three times with 200 µl 1 x ICB. After that, 100 µl of the biotinylated protein were added to each well and incubated over night or for a minimum of 2 hours at 4 °C with slow shaking. To remove the non-bound protein, each well was washed three times with 200 µl 1x ICB.

Before proceeding with the Ribogreen assay 50 µl of 2 x ICB-Mix were added to the immobilized protein to avoid drying of the protein.

4.9.2 RNA serial dilution and Ribogreen detection

In the next step, the target RNA is added in varying concentrations (e.g., 3160, 1000, 316, 100, 31.6, 10, 3.16, 1, 0.316, 0 nM) to the immobilized protein. The highest RNA concentration was calculated and prepared as 2 x stock (6320 nM) of the final RNA concentration (3160 nM) in ddH_2O, as starting point for serial dilution, which will yield the final concentration by adding it to the 2 x ICB buffer. The serial dilution was set up in a sterile deep-well mixing block (e.g. 2.0 mL Deep-Well Mixing Block, 96 wells per Block, Matrix Technologies GmbH) by stepwise diluting the stock RNA 1:3.16 into ddH_2O.

Before start of the incubation, the Ribogreen incubation buffer for final detection was prepared with a 1 : 500 dilution of Ribogreen in 1 x TE buffer made from 10 x TE buffer stock. After that, 50 µl of each RNA concentration were added to the immobilized protein in 2 x ICB and incubated under blue light for 30 min at 25 °C. (For the dark control plates, the incubation was started at the end of the light incubation period to avoid trouble during elongated washing.) Washing also occurred under steady blue light following the instruction in Table 18.

Table 18: Ribogreen assay instructions for washing.

(1)	Remove supernatant.
(2)	Resuspend wells with 200 µl 1x ICB.
(3)	Incubate sample for 30 seconds.
(4)	Remove and discard the supernatant.
(5)	Resuspend wells with 200µl 1x ICB.
(6)	Incubate sample for 3 minutes.
(7)	Remove and discard the supernatant.
(8)	Repeat step (5) to (7) 3 x .

After washing, 150 µl of the Ribogreen incubation buffer were added to each well and incubated for 1 hour at room temperature in the dark. The supernatant (125 µl) was then transferred into a fresh black 96 well plate (Greiner BioOne) and Ribogreen fluorescence was quantified using a Tecan Infinite M200 PRO plate reader (Tecan Group Ltd.) at an extinction of 485 nm and emission of 530 nm.

4.10 Size exclusion chromatography

Size exclusion chromatography (SEC) is a technique used to separate molecules in solution by their size. As macromolecules, such as proteins, may adapt different structural conformations, the molecule size does not correlate necessarily with the molar mass of the molecule. In contrast, the combination of SEC and Multi-Angle Light Scattering (SEC-MALS) offers an absolute technique for determining the molar mass and root mean square (rms) radius in solution [120].

For SEC-MALS analysis, a MALS detector (Dawn EOS, Wyatt Technology Europe GmbH) connected to a high performance liquid chromatography system (HPLC) (1100 Series, Agilent Technologies

Deutschland GmbH) with a variable UV detector was used at a detection wavelength of 280 nm. The system was combined with a refractive index detector (Shodex RI-71; Showa Denko Europe GmbH). For data analysis the ASTRA 6 software (Wyatt Technology) was used, which yielded estimates for molecular weight (MW). The SEC-MALS system belongs to the Scheibel Laboratory of the University of Bayreuth (Bayreuth, Germany), during use I was kindly supported by Dr. Martin Humenik from the group of Prof. Scheibel. Prior to the start of the run, the system coupled to a Superose 6 10/300 GL column (GE Healthcare) was equilibrated with 1x ICB (12 mM HEPES, 135 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 10 % glycerol) buffer. Per run, 50 µl of sample with a protein concentration of 100 µM were applied to the system. Before loading, the measured variants were preincubated at around 22 °C (ambient temperature) in the dark for 2 min, or in case of the light experiments, illuminated with a blue LED (λ_{max} = 450 nm, 200 mW cm^{-2}) for another minute. For the dark investigations, the column was protected from light during gel filtration.

Subsequent SEC experiments were repeated at 4 ° C using a Äkta Prime FPLC system (GE Healthcare) without MALS detector, as at this temperature the reversion kinetics are slowed. The system coupled to a Superdex 75 10/300 GL column (GE Healthcare) was previously equilibrated in 1x ICB buffer. Again, 50 µl sample of 100 µM were applied for each run. Before loading, the measured variants were illuminated with a blue LED (λ_{max} = 450 nm, 200 mW cm^{-2}) for 1 min in case of the light experiments, all variants were preincubated on ice prior to the start. For the dark investigations, the column awas protected dditionally from light during gel filtration. The data were evaluated using the UNICORN control software (GE Healthcare).

4.11 X-ray crystallography

Protein crystallography enables the visualization of protein structures at an atomic level only accessible with electro-magnetic radiation in the nanometer range, called ‘X-rays’. As the diffraction of a single protein molecule would be too weak to be measurable, the protein has to be brought into an ordered three-dimensional array form (‘crystal’), a process that is called ‘crystallization’. The x-rays are diffracted by the electrons in the crystal, which results in a three-dimensional map of the crystal structure [121].

4.11.1 PAL crystallization

Determination of crystallization conditions

Prior to all crystallization experiments, the protein fractions were centrifuged at 55.000 rpm for 60 min at 4 °C (Optima-MAX-TL; Beckman Coulter GmbH, Krefeld, Germany) in the laboratory of Prof.

Dirk Schüler. For the screening of crystallization conditions, vapor diffusion sitting drop 96-well MRC2-plates (Molecular Dimensions) were set up with commercial sparse matrix screens (Qiagen) at 9 mg/ ml and 4.5 mg/ ml protein concentration using a Phoenix liquid handling system (35 µl reservoir, 0.2/0.2 µl drops; Art Robbins Instruments). To prepare protein solutions of the desired concentration, 1 x ICB buffer (12 mM HEPES, 135 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 10 % glycerol, pH 7.5) was used. The plates were monitored by eye using a Nikon SMZ1500 stereomicroscope (Nikon GmbH) in combination with a red-light filter (B+W 58mm 090 5x F-Pro; Nikon GmbH).

Optimization grid screens were performed manually under same temperature and light conditions in 24-well plates (VDXm plates; Hampton Research) with the 'Hanging drop' method. To examine whether the crystals were composed of protein or salt, small amounts of JBS True Blue (Jena NakaBioscience) were added to the drop and the crystal color evaluated after 10 min. In a further attempt of optimization, the 'Microseeding' method was applied, the microseed stock was generated following the instructions from Douglas Instruments ('*Procedure for making the microseed stock*') and added to the protein sample in a 1: 100 dilution before combining the crystallization drops.

Crystal mounting and data collection

Under optimal conditions (0.1 M bicine, pH 9.2; 15 % PEG 20000; 2 % dioxane; 0.8 M imidazole 1 : 1 with 4.5 mg/ ml protein solution), needle-shaped crystals emerged after one day and were harvested after 3 - 7 days. Crystals were cryo-protected by soaking for at least 30 sec in small drops of crystallization solution with 25 % (v/v) glycerol. Cryo-protected crystals were mounted on a Cryo-Loop (Hampton research) and immediately flash-frozen in liquid nitrogen. Monochromatic oscillation X-ray diffraction data sets were collected at BESSY II (Helmholtz-Zentrum Berlin) at beamlines 14.1 at 100 K and processed using XDSAPP [122].

Phasing, model building and presentation

As molecular replacement in Phaser could not solve the phase problem from the native PAL diffraction data, selenomethionine (SeMet)-substituted PAL was grown and crystallized as described below. Experimental phasing was performed by single-wavelength anomalous dispersion (SAD) from a SeMet-substituted crystal with the help of the SHELXC/D/E pipeline [123]. The experimental phasing, as well as large parts of the model building were done by Dr. Sébastien Moniot. Initial automatic rebuilding of the PAL structure was performed with AutoBuild [124] within Phenix and Buccaneer [125]; model corrections and completion were then continued by several cycles of refinement with phenix.refine [126]. The model figures shown in this work were created with PyMOL [127]. The final model was deposited to the Protein Data Bank under the accession code 6HMJ.

Selenomethionine-substituted crystal growth

For selenomethionine (SeMet) -substituted PAL expression, 50 ml minimal medium (M9 Broth, see Table 19) containing 0.4 & of glucose and antibiotics (gentamycin and kanamycin) were inoculated with Arctic Express (DE3) harboring the respective expression plasmid and grown overnight at 37 °C at 225 rpm in an incubator shaker (Innova 44R, New Brunswick Scientific). All media and reagents were prepared as detailed in Table 19.

Table 19: Preparation details for reagents and buffers required for SeMet-substituted crystal growth.

Solution	Components, final concentration and solvent (optional)
M9 Broth (Sigma-Aldrich)	10 g/ L in ddH_2O
20 % D-glucose	0.2 g/ ml, 0.4 %
Additives (10 ml)	0.85 M NaCl, 8.5 mM; 0.01 M $CaCl_2$, 0.1 mM; 0.2 M $MgSO_4$, 2 mM; 0.1 mg/ ml thiamine, 1 µg/ ml
Cocktail of 6 amino acids	L-isoleucine, L-leucin, L-lysin, L-phenylalanine, L-threonine, L-valine; each to 100 mg/ L
L-selenomethionine (SeMet)	60 mg/ L

Next day, cells were harvested, the supernatant discarded and the cells resuspended in 10 ml M9 Broth (Sigma-Aldrich Chemie GmbH, Munich, Germany). The concentrated cell suspension was then used to inoculate 800 ml M9 Broth (Sigma-Aldrich) plus 0.4 % glucose, additives and antibiotics. The culture was grown at 30 °C at 225 rpm for 5 hours up to an OD_{600} of 0.7 – 0.9. At that OD, the incubator was cooled down to 12 °C, the cocktail of 6 amino acids (lacking methionine) and SeMet were added and protein expression was induced with 1 mM IPTG. The cells were incubated for 72 hours at 12°C at 225 rpm before harvest. Protein expression and crystallization were then realized as described in section 4.4.3 and 4.11.1.

4.12 Cultivation of *Nakamurella multipartita*

The *N. multipartita* stock culture was ordered from DSMZ (DSM-44233; Leibniz-Institut DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) in freeze-dried form. For cultivation, 5 ml cultures of TSB media (Tryptic Soy Broth, Sigma-Aldrich) were inoculated from the freeze dried stock under sterile conditions and grown at 28 °C at 225 rpm in an incubator shaker (Innova 44R, New Brunswick Scientific). Usually it took from several days up to more than a week until first signs of turbidity could be perceived in the culture media. After reaching saturation,

inoculation of fresh culture medium with 1/100 of the saturated culture diluted to an OD_{600} of 1.0 led to faster growth reaching optical density close to saturation (OD_{600} = 1.5) after approximately 48 hours, after which they were harvested for the RNA-co-immunoprecipitation experiments. To assure exclusive growth of *N. multipartita*, either gentamycin or nalidixic acid may be added to the culture media, however leading to reduced growth rates.

4.13 RNA-co-immunoprecipitation

While SELEX procedure from Section 4.8 provided a way of determining a randomly generated artificial RNA binding sequence to our PAL target, the identified SELEX targets do not necessarily match the natural binding sequences and features. RNA-co-immunoprecipitation (RIP) presents an alternative method to map natural protein-RNA binding sites by isolationg the protein of interest together with its associated RNA from its native environment, and subsequent mapping of the RNA via microarrays or next-generation-sequencing methods.

4.13.1 Synthesis and testing of an PAL-specific polyclonal anti-rabbit antibody

The production of a polyclonal anti-rabbit high-affinity antibody against PAL was assigned to to an external service company (Davids Biotechnologie GmbH). Around 1.5 mg of purified PAL were sent to Davids to serve as antigen for immunization and subsequent purification of the antibody. Polyclonal antibodies recognize multiple epitopes of the antigen, which involves the risk of a broader background signal. Therefore, affinity purification was chosen for concentration of the antibody. Before application in RIP experiments, the specificity and suitability for IP use of the PAL-Ab were confirmed in Western blot experiments.

Western blotting

For Western blotting, proteins were typically separated via SDS-PAGE on 12 – 15 % polyacrylamide gels, made of Acrylamide/Bis 37.5:1, 40% (w/v) solution (Merck), in a first step. All media and reagents were prepared as detailed in Table 20.

Table 20: Preparation details for reagents and buffers required for Western blot.

Solution	Components	Solvent	Final pH
10 x TBS (Tris-buffered saline)	200 mM TRIS, 1.5 M NaCl	ddH_2O	7.5 (HCl)

1 X TBST (TBS + Tween)	20 mM TRIS, 150 mM NaCl, 0.1 % Tween 20	dd H_2O	7.5 (HCl)
Transfer buffer	25 mM TRIS, 192 mM Glycin, 20% (v/v) Methanol	dd H_2O	8.3 (HCl)
Detection buffer	100 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$	dd H_2O	9.0 (HCl)

For transfer of the proteins to the detection membrane, a semi-dry protocol was employed using Polyvinylidene difluoride (PDVF) membranes (Immobilon-P, 0.45 µm, Merck) and a Trans-Blot Semi-Dry Transfer Cell (BioRad). After running the SDS-PAGE gel, the gel was equilibrated in Transfer buffer for 5 min to remove electrophoresis buffer salts and detergents. PDVD membrane and 4 blotting papers were prepared according to the dimension of the gel (max. 6 x 9 cm), the membrane was then activated with methanol; both blot papers and membrane were soaked in Transfer buffer before transferring them to the transfer cell. For assembly of the transfer unit, the protein-containing gel, PDVD membrane and blotting papers were stacked according to the manufacturer's instructions and run at 400 mA for 25 min.

For detection of the protein, first unspecific interactions between membrane and antibodies were blocked through incubation of the membrane in 10 ml blocking buffer (1 x TBST + 5 % (m/V) skim milk powder) for 1 hour or overnight at 4 °C under light shaking. After blocking, the membrane was washed with 1 x TBST and the primary antibody (PAL-Ab, rabbit, Davids Biotechnologie GmbH) was added in a 1 : 1000 dilution within 10 ml 1 x TBST + 0.5 % (m/V) skim milk powder and incubated for 1 hour at RT. Before applying the secondary antibody, the membrane was washed three times for 5 min with 1 x TBST. The alkaline phophatase (AP)-coupled secondary antibody (Goat anti-Rabbit IgG (H+L) Secondary Antibody, AP, Thermo Fischer Scientific) was added in a dilution of 1 : 50000 in 10 ml 1 x TBST and incubated for 1 hour at room temprerature, followed by three more washing steps à 5 min in 1 x TBST. For final exposure, 1.5 mg BCIP and 4.24 µg NBT were added to 10 ml Detection buffer, after 2 to 10 min the reaction was stopped by removing the detection mixture.

4.13.2 RIP procedure

Highest measures of precautions were paid during the entire RIP procedure to avoid RNase contamination, as the overall quantities of cell-isolated RNA are distinctively lower then during work with RNA originating from *in vitro* processes.

The RIP assay requires the cultivation of larger culture volumes. Therefore, two *N. multipartita* pre-cultures were grown up to saturation as described in Section 4.12 and subsequently used to

inoculate 2 x 800 mL of TSB media 1 : 200. The cultures were then grown for about 48 hours at 28 °C at 225 rpm under constant blue light illumination (50 µW/ cm² at 450 nm) to an OD_{600} of 1.2 – 1.4 before harvest. The cells were harvested by centrifugation at 5000 x g for 10 min and immediately resuspended in 50 ml sterile TSB media and 100 ml RNAprotect (RNAprotect Bacteria reagent, Quiagen GmbH, Hilden, Germany) designed to stabilize cellular RNA transcripts prior to isolation procedures. The cell suspension was incubated for 5 min at room temperature and again centrifuged for 10 min at 5000 x g. The cells were resuspended in 15 ml 1 x ICB (12 mM HEPES, 135 mM KCl; 10 mM NaCl; 1 mM $MgCl_2$, pH 7.5) and subjected to mechanical disruption using the microfluidizer instrument (Modell 110; Microfluidics). The cell lysate was then centrifuged in 2 ml fractions for 3 min at 20000 x g to separate soluble and insoluble fraction. The supernatant (soluble fraction) was then employed for the RNA co-immunoprecipitation in the next step.

For the third RIP approach ('Pure RNA'), the RIP experiment was conducted with previously purified RNA from *N. multipartita* cells, for which the RNeasyProtect Bacteria Mini Kit (Quiagen) was used. The below protocol was designed according to the instructions from the 'RNAprotect Bacteria Reagent Handbook 01/2015'. 50 mL of TSB media were inoculated and subsequently grown as described above; the cells were then equally collected by centrifugation at 5000 x g for 10 min and immediately resuspended in 15 ml sterile TSB media and 30 ml RNAprotect (Quiagen). After 5 min incubation at room temperature, the supernatant was removed and the cells were resuspended in 16 ml of RLT buffer via vigorous vortexing for 10 sec. Disruption of the cells was realized using the fluidizer (Modell 110; Microfluidics), after which the lysate fraction was centrifuged in 8 ml fractions for 3 min at at maximum speed. The supernatant was transferred to a fresh falcon, and mixed with an equal volume of ethanol (70%). After that, the protocol followed the instructions from Protocol 7 'RNAprotect Bacteria Reagent Handbook 01/2015', using 22 Mini spin column to process the prepared cell lysate. The purified RNA was then employed for the RNA co-immunoprecipitation in the next step.

For the immunoprecipitation, Dynabeads magnetic beads (Dynabeads Protein A, Thermo Scientific) were employed. The protocol below was hence adopted from the respective manufacturer's instructions. For target antigen- (Ag) immunoprecipitation, 2 mL of the *N. multipartita* supernatant were pre-icubated with 10 µl (15 µg) of the PAL-Ab (except for the negative control, for which this pre-incubation step was left out) for 30 min at room temperature prior to bead capture; as recommended for the work with protein/ nucleic acid complexes. The supernatant-Ab suspension was then added to 50 µl of the Dynabeads, previously prepared as detailed in the manufacturer's guidelines, thoroughly mixed and incubated for another 10 min at room temperature. After that, the supernatant-Ab suspension was removed and the Dynabeads-Ab-Ag complex washed three times

using 200 µL TBST (see Table 20) for each wash. For separation of the Dynabeads-Ab-Ag suspension, the magnet was used to remove the supernatant between each wash. After the last washing step, the Dynabeads-Ab-Ag complex was resuspended in 100 µl TBST buffer and transferred to a fresh tube to avoid co-elution of proteins bound to the tube wall. For target antigen elution, TRIzol reagent (Thermo Scientific) was used which supports a denaturing elution and further facilitates the RNA isolation procedure in the coming step. For that, the supernatant was removed from the Dynabeads and 200 µl of TRIzol were added and thoroughly mixed by pipetting to resuspend the Dynabeads-Ab-Ag complex. The homogenized sample was stored at -80 °C before proceeding with the RNA isolation the next day.

As TRIzol reagent (Thermo Scientific) was employed for RNA isolation, the further proceeding followed the manufacturer's instructions. Phase separation was done as detailed in the directions; for RNA-precipitation, 1 µl glycogen (Glycogen, RNA-grade, Thermo Scientific) was added to the sample to increase the RNA yield. The washing protocol followed the manufacturer's directions; for RNA resuspension, 15 µl of RNase-free H_2O were applied to the RNA pellet.

For further analysis, 5 µl of the resuspended RNA from all samples were subjected to dephosphorylation via FastAP (FastAP Thermosensitive Alkaline Phosphatase, Thermo Scientific) and subsequent 5'-end radiolabeling as described in Section 4.7.1. The radiolabeled fractions were then loaded onto a 5 % polyacrylamide TBE gel, made of Acrylamide/Bis 37.5:1, 40% (w/v) solution (Merck) and 1 x TBE buffer, and run in 0.5 x precooled TBE buffer (4 °C) for 30 min (15 V cm^{-1}). The gel was then dried on Whatman blotting papers (GE Healthcare) in a slab gel dryer (GD2000, Hoefer) connected to a vacuum system and exposed to a film (BioMax MR, Kodak) for at least 12 hours. Autoradiographies were digitized using an FLA-7000 phosphorimager (Fuji Film Europe, Düsseldorf).

5. Results

5.1 Engineering of an ANTAR-containing photoreceptor

The successful engineering of chimeric photoreceptors proved that sensor and effector modules of different origin can be functionally rewired [4]. The starting point for the plan of engineering an ANTAR-containing photoreceptor was the question of whether the signal transduction mechanism of common LOV protein architectures can be transferred to different effector types. For the light-regulated histidine kinase YF1 [21,87], some essential aspects of the underlying signal transduction mechanisms were already known at the beginning of this work. YF1 is composed of an N-terminal blue-light-sensitive LOV domain and a C-terminal histidine kinase. This effector function is the most abundant output in PAS proteins and occurs in almost half of all multi-domain PAS proteins [128]. During the search for less frequent PAS effectors, we came across effectors involved in transcriptional or posttranscriptional regulation, such as the family of helix-turn-helix, helix-loop-helix, or ANTAR proteins. Due to the great potential of a prospective RNA-binding optogenetic tool, e.g. for the elucidation of non-coding RNA involving processes, we focused on the ANTAR effector funtion. As discussed in Section 3.3, a common design strategy is based on the sequential and structural superposition of sensor and effector domain and subsequent domain replacement within a signal receptor (see Figure 6 in Section 3.3). The following Section (5.1.1) outlines the outcome of this approach. Since the development of a new photoreceptor often requires the screening of many chimeric protein variants, the receptor activity to an easily detectable reporter readout was an important intermediate objective. Section 5.1.2 summarizes the outcome of another design strategy based on the use of LOV domains that act as photoactivatable dimerization modules (see Figure 6). In Section 5.1.3, a naturally occurring photoreceptor was used as a design template, the sequence of which was found in the gene databases during this work.

5.1.1 Domain replacement based on structural superposition

At the beginning of the project only three crystal structures of ANTAR proteins were available, belonging to AmiR from *P. aerigunosa* [104], Rv1626 from *M. tubercolosis* [101], and NasR from *K. oxytoca* [105]. All three structures (see Figure 8) share a three-helical bundle, but only AmiR presents a coiled-coil region centered on the first long α-helix. Whereas Rv1626 was crystallized as a monomer, NasR also occurs in a dimeric conformation. However, here the dimer interface is mainly formed by the adjacent NIT effector giving rise to a coiled-coil-lacking fold that is well conserved among NasR orthologues. Since the light-sensing LOV domain in YF1 forms a tight dimer through a hydrophobic interface and connects to its effector via a coiled-coil linker, AmiR was the most suitable

candidate for structural superposition with YF1. In addition, the target RNA sequence for AmiR was already identified in a previous study, which would facilitate subsequent functional tests [110].

Computer modeling of the YF1-LOV-AmiR chimera LOVA

The recombination of different protein domains requires the fusion of the linker elements from the parental proteins. The sensor-effector connecting linker of homodimeric photoreceptors frequently occurs as structured coil-coiled motif [21,129], which not only provides spatial proximity but also significantly contributes to the transfer of the activation signal (see Section 3.3). The physical properties and maintenance of the three-dimensional structure of the linker element thus plays an essential role for an intact signal transmission of the engineered chimeras, rendering the choice of a suitable fusion site one of the most error-prone steps in the design process [22,87].

For the design of a blue-light-activated ANTAR chimera ('LOVA'), the linker regions of the crystal structures from YF1 (PDB code: 4GCZ) and AmiR (PDB code: 1QO0) were superposed by least-squares minimization in PyMOL. Only the coiled-coil segments that present a relatively high sequence similarity (> 40 %) were used for the alignment. The overlay produced by the 'align' algorithm was similar to the best alignment match for the residues (resi) 122 - 151 from AmiR and resi 125 - 151 from YF1 found by the pairwise sequence alignment tool 'Needle'[130] with 13.8 % identity, 44.8 % similarity, and score 17 (see Figure 11.a). The overlap resulting from the linker alignment of the two parental proteins represents the window for possible fusion sites for the assembly of the LOVA chimera. An additional consideration was that the YF1 LOV domain originally derives from the functional stress response regulator YtvA from *B. subtilis* whose coiled-coil linker region differs to that of YF1 on a sequence level. To compare the coiled-coil helical pattern of the three different linker sequences, a multiple sequence alignment of the three linker regions was performed using ClustalX2 [131]. The hydrophobicity pattern was better maintained, and the positions of the YF1 and the AmiR linkers were shifted by three residues compared to the structure-based alignment (see Figure 11.b). Further comparison of the two pairs (YtvA:AmiR vs. YF1:AmiR) revealed a better agreement between the YF1:AmiR linker sequences regarding the hydrophobicity pattern. Therefore, the fusion site of the initial LOVA chimera was derived from the overlay of YF1 and AmiR. For the original YF1 system, the light-dependence of the kinase function can be influenced by changes in linker length [87]. For this reason, a test of several linker variants was pursued, which also contains the chimeric variant proposed from the structure-based alignment (later named '+3-LOVA'; see 'LOVA linker variants').

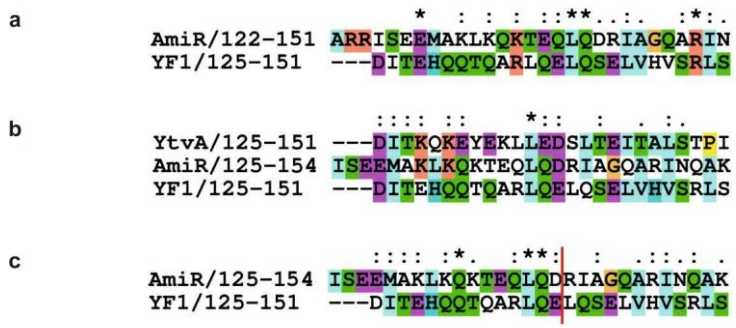


Figure 11*:* Linker region alignments for planning of the YF1-LOV-AmiR chimera 'LOVA'. **(a)** Structural alignment of the linker regions of the AmiR and YF1 sequences. **(b)** Sequence- based alignment of the linker regions of YtvA, AmiR and YF1. **(c)** Choice of the fusion site (red line) for the initial YF1-LOV-AmiR chimera 'LOVA' derived from the window of YF1 and AmiR in (b).

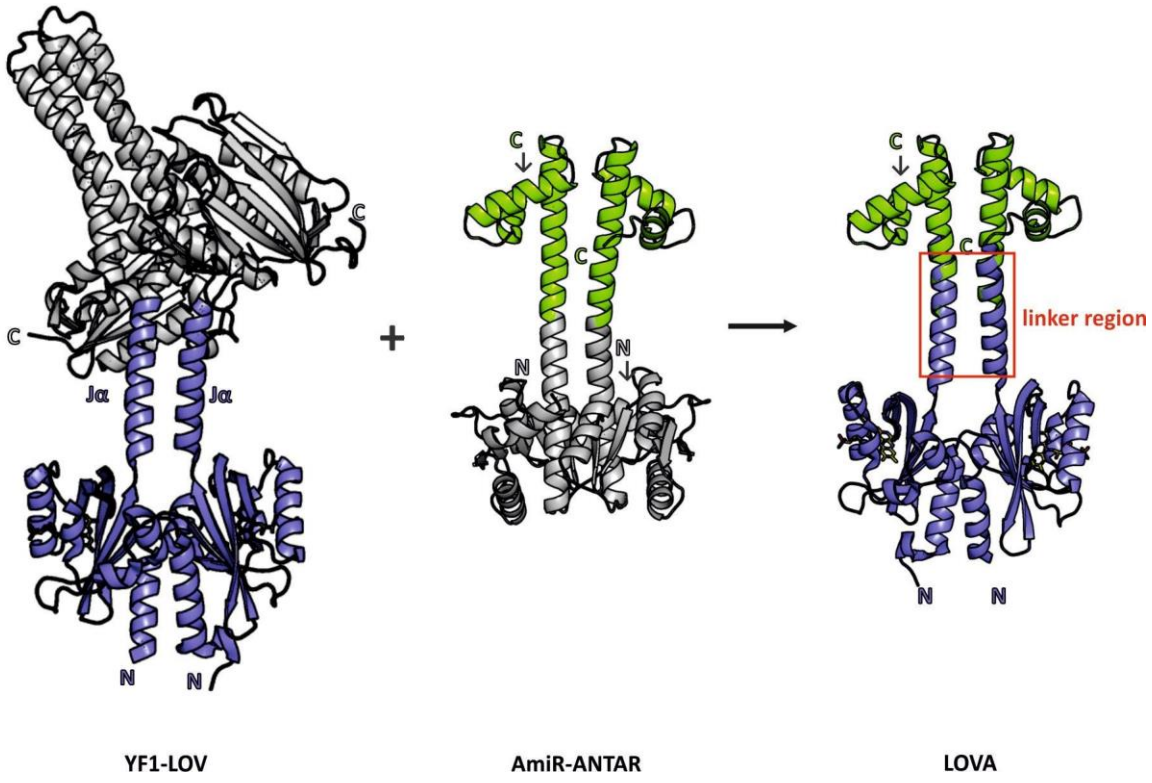


Figure 12: In-silico planning of the YF1-LOV-AmiR chimera LOVA. The crystal structures of the YF1 (left, PDB code 4GCZ) and AmiR (middle, PDB code 1QO0) signaling receptors both form parallel hommodimers featuring extended α-helical coiled-coil linker elements with an elevated structural homogeneity. The structural model of the LOVA chimera (right) combines the N-terminal LOV sensor module of YF1 (left) with the C-terminal regulatory ANTAR effector domain from AmiR. The exact fusion site was determined according to the considerations that resulted from the structural superposition of the coiled-coil linker elements (red box) of the two signal receptors.

For the choice of the definitive fusion site within the aligned region the following general design principles [132] were considered: (i) conservation of tertiary structure of sensor and effector domains (or all involved domains); (ii) avoidance of disruption of quaternary structure, especially interfaces between single domains and the central helical backbone; and, as discussed above, (iii) maintenance of the hydrophobicity pattern within the coiled-coil linker sequence, as helical linkers often play important roles in signal transduction. The initial fusion construct of YF1-LOV and AmiR ('LOVA') was thereby defined by resi 1 - 138 from YF1 and 142 - 196 from AmiR (see Figure 11.c and Figure 12).

Functional testing of LOVA

For experimental testing of the initial LOVA chimera, the fusion of the gene sequences defined above (see Figure 11.c) was cloned into the pBAD-30 vector between SpeI and SalI sites of the multiple cloning site. A test of expression gave best results for overnight expression (≈ 16 hours) at 20 °C employing the CmpX13 *E. coli* strain. For purification, the protein expression was conducted at larger scale with 800 mL culture volume. Ni-NTA affinity purification employing buffer B1 for washing resulted in satisfactory amounts (≈ 2 mg or 2.5 mg/L) of pure and soluble target protein. For the envisaged activity assay, the protein was dialyzed against fresh sodium phosphate buffer (50 mM sodium phosphate, 200 mM NaCl, 10 % glycerol, pH 7.5). To inspect photochemical functionality of the LOV-sensor, the absorption spectra of the UV and visible (UV-Vis) range were recorded and the ability to switch from dark-adapted to signaling state was checked. The absorption spectra of the dark-adapted LOVA chimera showed the typical fine structure of the oxidized flavin mononucleotide (FMN) chromophore with a main absorption band at 448 nm and two side peaks (see Figure 13). After light exposure these absorption bands decreased and instead a maximum at 390 nm emerged, indicating the formation of the covalent bond between the flavin chromophore and the reactive cysteine. Reversion to the dark-adapted state occurred within the range of several hours, comparably to the recovery kinetics within the YF1 context (not shown).

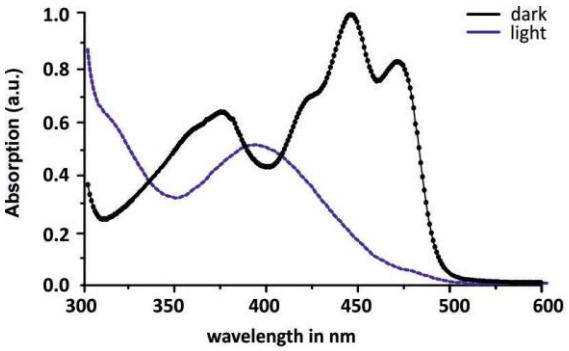


Figure 13: UV-Vis absorption spectra of the purified LOVA chimera. The dark-adapted ground state spectrum is shown in black with the characteristic absorption maximum at 448 nm, blue light induction triggers the appearance of the typical signature of the LOV signaling state with an absorption maximum at 390 nm. a.u. – arbitrary units.

To validate preservation of the ANTAR effector functions, an Electrophoretic mobility shift assay (EMSA) employing the target RNA sequence for AmiR identified by Wilson and colleagues [110] was conducted. This target RNA sequence derives from the leader region of the *P. aeruginosa* amidase operon, henceforth referred to as 'ami-lead'. The DNA sequence of the ami-lead segment comprising the original pE promoter was cloned between the BglII and NdeI sites of a pet28c vector. To generate a template for *in vitro* synthesis of the target RNA, the ami-lead sequence was amplified via PCR with a T7 promoter-containing forward primer ranging from the putative start of transcription to the terminus of the transcription terminator. The purified template was then employed for the RNA *in vitro* synthesis reaction, yielding a target RNA transcript of ≈ 120 nt length (see Figure 14). The EMSA was subsequently performed with constant concentrations of the radiolabeled transcript (≈ 50 pM) and varying amounts of the LOVA chimera. As positive control, a further gel shift assay was performed in parallel with the ANTAR domain of the parental AmiR protein ('ANTAR-AmiR'; resi 139 to 196/end) (see Figure 15, (1)). To investigate the effect of light induction on the RNA binding properties of LOVA chimera, the assay was performed under dark and blue light conditions (see Figure 15, (2-3)). The quantitative evaluation of the gel radiographs with the help of a Hill-fit determined the values for the dissociation constant (K_d), that indicates the effective protein concentration at 50% RNA occupation. These are in a similar range for LOVA chimera and the ANTAR-AmiR positive control ($K_d^{ANTAR\text{-}AmiR}$= 0.56 ± 0.13 µM vs. $K_d^{LOVA\text{-}dark}$= 0.51 ± 0.12 µM). For the differently treated LOVA-samples, no relevant differences between the dark and blue light treated samples could be detected ($K_d^{LOVA\text{-}dark}$= 0.51 ± 0.12 µM vs. $K_d^{LOVA\text{-}light}$= 0.54 ± 0.12 µM). To assess the specificity of the LOVA-RNA interaction, a further round of the assay was carried out with a 1000-fold excess (50 nM) of unlabeled yeast tRNA (see Figure 15. (4)). The computed binding affinity for LOVA within the competition gel shift experiment was similar to the previous EMSA runs ($K_d^{LOVA\text{-}illum}$= 0.54 ± 0.12 µM vs. $K_d^{LOVA\text{-}unspec.}$= 0.59 ± 0.07 µM), showing no signs of competitive binding through unspecific RNA interactions. These results demonstrated that the RNA binding capacity of the LOVA chimera is

sequence-specific, thereby providing the basis for a bacterial reporter assay that will facilitate the investigation of further chimeras.

5'- GGCGCCGGCGCCAUCAGGUCAUGCGCAUCAGCGUCGAUGUCGCGGGACCGAACCUAACGCAUACGCACAGAGCAAAUGGCU CUCCCGGGGUUACCCGGGAGGGCCUUUUUUUCGUCCCAAAAAAUAAC -3'

Figure 14: Transcript of the ami-lead segment for use as target RNA within the EMSA assay. The putative transcription terminator region is highlighted in grey.

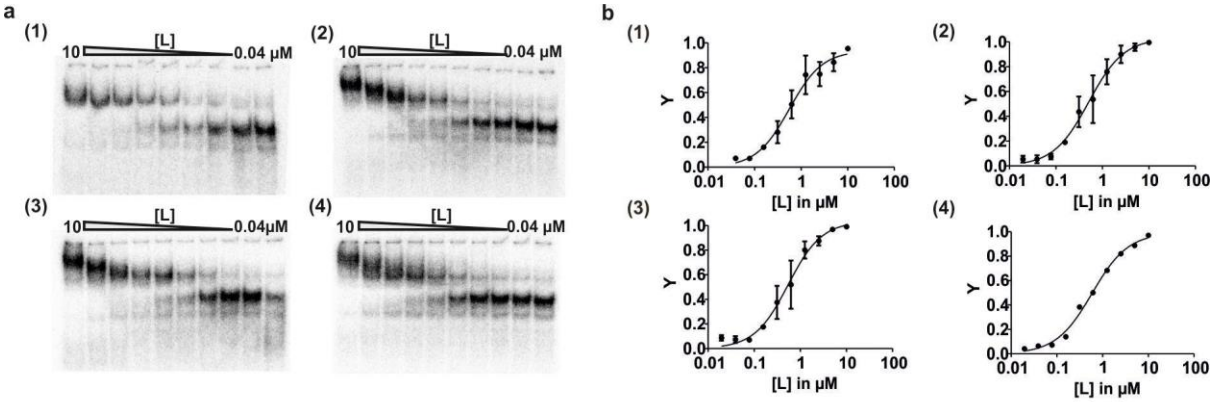


Figure 15: Determination of the binding affinity of LOVA to its RNA target **(a)** EMSAs of the positive control (1) and LOVA (2-4) in the presence of ≈ 50 pM target RNA. The protein sample was added in dilutions of 1:2 starting with 10 µM final concentration. Before separation on a 6 % - TBE gel, the LOVA-RNA mixtures were incubated 20 min at ambient temperature in the dark [(2) → 'LOVA dark'],and then for another 20 min under continuous blue light illumination [(3) → 'LOVA illum']. The same light conditions were maintained during gel electrophoresis. The AmiR positive control was treated similarily to the LOVA dark samples. An additional gel shift assay was conducted, for which 50 nM of yeast tRNAs were included to the LOVA-RNA mixtures; the samples were then directly subjected to 20 min blue light illumination before electrophoretic separation [(4) → 'LOVA + unspec.']. **(b)** RNA dissociation curves for AmiR (1) and LOVA (2-4) obtained by quantification of both free and ligand-bound RNA of the EMSAs shown in (a)**.** For (1), (2) and (3) a reproduction (n = 2) was performed and included in the evaluation. To determine the respective affinity constants, the fractions of bound RNA (Y= $[R{\cdot}L]/[R]_{total}$) were plotted against the amount of ligand [L]. The thereby computed values for K_d, hill coefficient (h) and Y_{max} are shown below in Table 21. [L] – ligand (ANTAR-AmiR control or LOVA), Y – fraction of bound RNA.

Table 21: Computed fitting values of the RNA affinity curves shown in Figure 15.b. K_d – dissociation constant, h - hill coefficient, Y_{max} – computed maximal value for the fraction of bound RNA (Y; Y = [R·L]/[R]total).

Best-fit values	(1) ANTAR-AmiR	(2) LOVA dark	(3) LOVA light	(4) LOVA + unspec.
Y_{max}	0.94 ± 0.07	1.03 ± 0.09	1.06 ± 0.08	0.99 ± 0.04
h	1.22 ± 0.26	1.15 ± 0.23	1.19 ± 0.23	1.1 ± 0.1
K_d [µM]	0.56 ± 0.13	0.51 ± 0.12	0.54 ± 0.12	0.59 ± 0.07

Reporter assay for AmiR-derived chimeras

As the engineering of novel photoreceptors may well require the testing of a substantial number of candidates, we constructed a beneficial high-throughput screening assay for the detection of light-regulated chimeras. In order to track the activity of the parental AmiR effector, Wilson et al. [110] previously established an *E. coli* screening system. The synthesis of an amidase reporter readout, provided on a first plasmid, was regulated through AmiR provision with the help of a second plasmid. A third plasmid capable of just expressing the leader mRNA was further able to abolish the effect of AmiR antitermination. Since the RNA-binding activity of the LOVA chimeras originates from the AmiR ANTAR domain, the design of the initial screening system was oriented on their approach. However, on grounds of the inherent advantages of fluorescence-based reporter systems, we decided to replace the prior amidase readout by the red fluorescent protein DsRed2 (see Figure 16.a). For the construction of the initial reporter construct ('pE_DsRed'), the gene sequence of DsRed2 was inserted between the BglII and NdeI sites behind the 'ami-lead' segment on a pet28c plasmid. The 'ami-lead' sequence comprises the original promoter (pE) as well as the transcriptional terminator targeted by the AmiR ANTAR regulator (see Figure 9.b). The expression of the AmiR-derived library constructs was put under control of the arabinose-dependent BAD promotor on a pACYC184 backbone of compatible origin. For setting up the assay, the ANTAR-part of AmiR (from L139 to A196/end) was used as positive control, as well as a negative control composed of the empty library_pACYC184 plasmid. For functional screening, AmiR-chimeras and the initial pE_DsRed reporter construct were co-transformed into competent CmpX13 cells and grown for 18 hours at 37 °C. Fluorescence (excitation 554 ± 9 nm, emission 591 ± 20 nm) and optical density at 600 nm were measured using the Tecan M-200 plate reader.

Even for high fluorescence gain settings (up to 200) the measurements revealed only very weak signals, lacking any significant difference between negative and positive control (not shown). Therefore, we looked out for possible problems or alterable parameters in the gene expression of the reporter construct: Within its original context the pE promoter is adapted to an enzymatic readout. The full amplitude of a biological signal is usually achieved through the action of various factors since the regulation of gene expression takes place at different levels, whereby the processes of transcription and translation are particularly important. During transcription, an RNA polymerase produces a messenger RNA (mRNA) molecule from the gene sequences encoded in the DNA, involving many other factors. The RNA polymerase binds to specific promoter sequences, i.e. conserved nucleotide sequences, for procaryotes usually located in the -35 and -10 regions, whose name refers to the position of the transcription start site. The nature of these regions significantly determines the transcription rate of a gene by the RNA polymerase. During translation, the resulting mRNA is translated into a sequence of amino acids requiring a multi-component ribosome complex

that binds to a conserved ribosome binding site (RBS) at the 5' end of the mRNA. The type and completeness of the RBS sequence determines the affinity of the interaction of the ribosome complex to the respective mRNA, which in turn represents an important aspect for the rate of translation. In addition, several other factors play a role, e.g. the process of untwisting the supercoiled structure of the plasmid, or the number of available initial templates that depends on the copy number of a plasmid, which in turn is governed by the origin of replication.

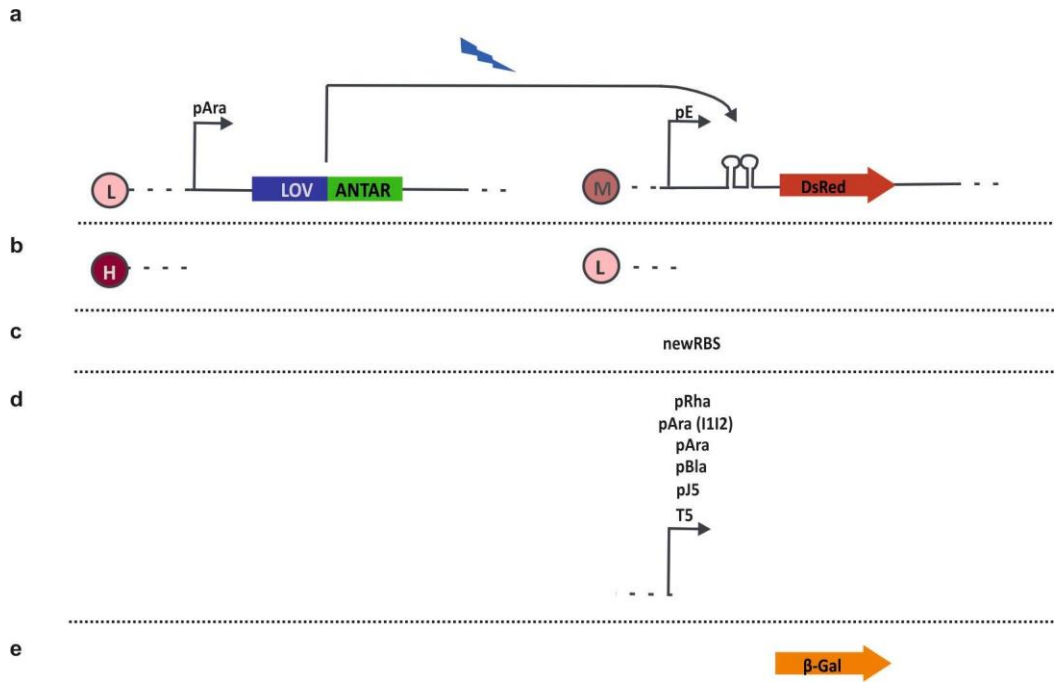


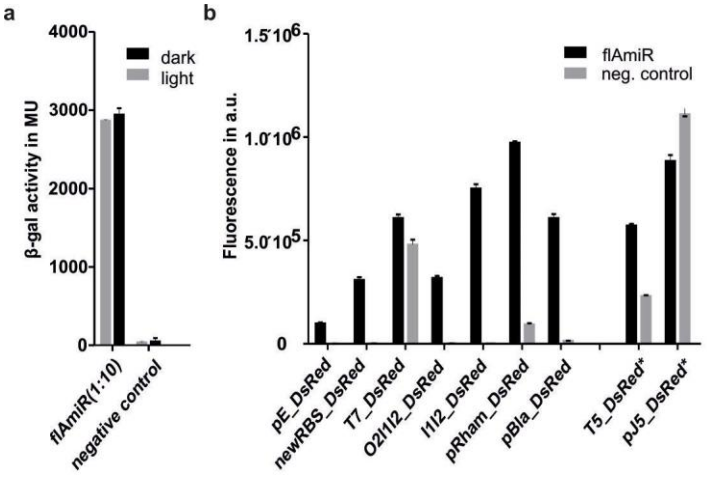
Figure 16: Schematic illustration of the reporter assay for AmiR-derived chimeras. **(a)** Design of the reporter assay: The AmiR-derived library constructs were placed on a low copy (L) pACYC184 backbone under the control the arabinose-dependent promoter pAra. For construction of the reporter plasmid, the leader sequence of the Ami-operon, comprising a weak *E. coli* promoter (pE), as well as the transcriptional terminator sequence, was placed in front of the DsRed gene on a pet28c backbone featuring a medium copy number (M). In order to optimize the output signal, the influence of different parameters was assessed by **(b**) varying the copy number of both library and reporter plasmid from low to high (H) and medium to low (L); **(c)** replacement of the ribosome binding site (RBS) of the reporter plasmid; **(d)** substitution of the pE promoter by various other promoters [134,135] of different strength; and finally, **(e)** the replacement of the DsRed reporter gene by the lacZ gene that encodes the β-galactosidase enzyme (β-Gal).

In order to increase the signal strength of the pE_DsRed reporter construct within the screening assay, we tested the influence of various parameters: in a first step (see Figure 16.b), we investigated the influence of the copy number of the plasmids by exchanging the medium-copy-number pet28c

backbone of the pE_DsRed reporter for that of the pACYC184 plasmid with a low copy number. Simultaneously, the pACYC184 backbone of the library plasmid was replaced by the backbone of the high copy plasmid pBAD30, so that the compatibility of the origins of replication was maintained. In a second step (see Figure 16.c) the RBS of the pE_DsRed reporter was replaced by the original pet28c RBS. None of these measures showed a significant influence on the strength of the fluorescence signal. The pE promoter was then substituted with a diverse set of well-characterized promoters of disparate strength [133] (see Figure 16.d; a complete list of all of the created constructs can be found in Table A3 within Appendix Section 8.2). Three of the novel promoter constructs (pJ5_DsRed, T5_DsRed, and T7_DsRed) led to significant improvement of the fluorescent signal, but still almost no variation between positive and negative control resulted, as the background noise seemed to increase proportional to the signal. Wilson et al. [110] previously observed a leakiness of the transcriptional terminator in conjunction with stronger promoters variants for their screening system, so that we concluded that these results might be due to similar reasons.

Since the above approaches did not succeed, we decided to substitute the fluorescent readout with an enzymatic reporter and chose β-galactosidase (β-Gal) as specific actor, whose action provides a classic approach to quantify levels of reporter expression in a simple and reliable colorimetric assay. Therefore, the DsRed2 gene of the reporter plasmid was replaced by the *lacZ* gene (see Figure 16.e). For the screening, the reporter plasmid and chimera variants were then co-transformed into competent Novagen TUNER (DE3) cells; which are *lacYZ* deletion mutants of the BL21 *E. coli* strain. The *lacY* mutation facilitates the homogeneous entry of IPTG into the cell, while the *lacZ* gene codes for the cell- own β-galactosidase. For more details regarding the implementation of the assay, see Section 4.3.1. First activity tests of the LOVA chimeras along with the controls yielded activities in the basal range with Miller Units (MUs) < 100 for any of the constructs, including the positive control (not shown). MUs are defined such that the fully induced lac-operon amounts to an activity of 1000 MUs, and to ≈ 1 MU in its non-induced state [115] providing a clear indication that even the positive control only achieved small levels of induction so far. Beyond that, it is likely that the maximal achievable range of MUs will be much higher in our system, as the number of β-gal actuators should relate to the copy number of the reporter plasmid, which is around 20 per cell. This implied that the positive control used until this point (the ANTAR domain of the AmiR protein from resi 139 - 196/end) might not be functional. For that reason, the full-length AmiR (flAmiR) sequence was cloned into the pKT-library vector to substitute the initial positive control. This new flAmiR control yielded Miller units within a range of 25000 to 35000 MUs in fresh rounds of the assay (see Figure 17.a; note that the cell volume of the positive control sample was diluted 1 : 10 for realization of the assay).

The observation that the reporter signal for the previously employed ANTAR-AmiR positive control is at the level of the negative control, is in apparent contradiction to the results of the initial EMSA experiments (see Figure 15). However, ligand-receptor associations are concentration-dependent equilibrium processes and the determination of the binding constant K_d does not provide an absolute measure such as 'binding' or 'non-binding' to assess the functionality of a receptor. The physiological concentrations under which the receptor function is optimally pronounced are often unknown. Since we observed a binding of the RNA target sequence regardless of the light conditions, we had classified the previously tested constructs as constitutively active and invested a lot of time in optimizing parameters of the reporter plasmid.



Construct	Fluo	norm. Fluo	Gain
pE_DsRed	98872	1,00	180
newRBS-pE_DsRed	307045	3,11	180
T7_DsRed	128726	1,30	180
O2I1I2_DsRed	316837	3,20	180
I1I2_DsRed	748625	7,57	180
pRham_DsRed	878639	8,89	180
pBla_DsRed	594131	6,01	180
T5_DsRed	340965	3,45	159
pJ5_DsRed	-226395	-2,29	159

Figure 17: Establishment of a reporter screening system. **(a)** The enzymatic Miller screening assay employing the pE_β-gal construct yields a strong signal in the range of approximately 30,000 MUs for the positive flAmiR control. Bars represent means from three measurements. **(b)** Evaluation of the fluorescence assay of different promoter variants of the DsRed2 reporter construct in combination with either positive control (flAmiR) or negative control. The stars behind the two last constructs indicate the application of a different gain setting, as indicated in (c). Again, bars represent means from three measurements. **(c)** Table indicating the dynamic range of the fluorescent signal (Fluo) and normalized Fluo (with reference to the original pE_DsRed reporter), as well as the applied gain for the fluorescence measurement in (b). The fluorescent signal (or dynamic range) was calculated from the difference in signal of the flAmiR construct and the negative control. MU – Miller unit; a.u. – arbitrary unit.

To complete the experiments on the fluorescence-based reporter system, the different promoter versions of the DsRed-reporter constructs were tested once again with the new functional positive control. For that purpose, all promoter and RBS variants of the DsRed-reporter constructs were co-transformed with the new flAmiR positive control into CmpX13 cells. In combination with the new

functional positive control, all of the promoter variants displayed an increased fluorescent signal compared to the pE-DsRed reporter construct, and only the T7_DsRed and pJ5_DsRed promoter variants continued to cause a leakiness of the transcriptional terminator. The sensitivity and utility of a reporter screening system is largely determined by the size of the dynamic range. The pRham_DsRed, I1I2_DsRed and pBla_DsRed reporter constructs (for which the original pE promoter was replaced by regions of the rhamnose, arabinose, or β-lactamase promoter, respectively [133,134]) show highly increased dynamic ranges compared to the original pE_DsRed reporter construct (8.89-,7.52-, and 6.01-fold increase, respectively) measured at a gain of 180. For the pJ5 and T5 promoter constructs, the detected signal was too strong at that gain, so the measurement was repeated at a lower gain of 159 (Figure 17.b). These two last reporter constructs exhibit a strong leakiness, which manifests in a high basal activity in the dark-adapted state (resulting in a negative value for the dynamic range of the pJ5_DsRed contruct). The successful implementation of a screening system for AmiR-derived chimeras can be regarded as an important milestone that could be useful for the screening of further engineered photoreceptor variants at a later stage.

Linker variants of the LOVA chimera

For the parental photoreceptor YF1, Möglich et al. have shown that modulations of the sensor- and effector-connecting linker element significantly influence the activity level as well as the light response [87]. By varying the length of the coiled-coil linker element, Möglich and colleagues were able to set the YF1 functionality in such a way that the kinase activity could be adjusted to either light- or dark-dependency. Hence, the linker properties of a signaling receptor seems to play an important role for functional signal transmission. Therefore, we analyzed the linker lengths of naturally occurring PAS-ANTAR proteins. The linker regions of native PAS-ANTAR signaling receptors were analyzed with the help of a script provided by Andreas Möglich, which facilitated the following steps: Using the Pfam database [135] all proteins harboring either a PAS (Pfam clan CL0183) or an ANTAR domain (Pfam family PF03861) were identified. Protein sequences that appear in both data sets were subsequently retrieved from the Uniprot database [136]. For better comparison to the LOVA chimera, further criteria, such as a defined order where the PAS domain resides N-terminally from the ANTAR domain and a maximal distance of 100 amino acids (aa) between them, were introduced, narrowing the protein sequences within the data sets down to approximately two thirds of the original fraction. With these remaining PAS-ANTAR candidates, a multiple sequence alignment was performed, whose analysis revealed one predominantly occurring length of 30 aa between the conserved parts of sensor and effector, defined by the aspartate of the ‘DIT’ motif for the PAS modules and the glutamine of the ‘Q’AKG motif for the ANTAR modules. With only 24 residues within the defined range, the linker length of the initial LOVA chimera differs from the predominant length

of other PAS-ANTAR sensors. The LOVA protein demonstrates activity in both dark and light-activated state, which correlates to the '-3' or '+4' position of the YF1-linker contructs if one assumes similar signaling mechanisms as for YF1 [137]. However, no obvious heptad periodicity could be detected, as previously reported for YF1 and further PAS-HisKA signaling receptors [32], which would be a hint for a shared signaling mechanism. Nonetheless, the lack of an observed heptad periodicity might still be due to the limited number of available PAS-ANTAR protein sequences of the defined architectures (n < 150), which restricts the statistical analysis. The investigation of additional linker variants comprising varied linker lengths seemed a good way to investigate this question. The established reporter assay (see previous Section, '*Reporter assay for AmiR-derived chimeras'*) considerably facilitated the implementation of these tests.

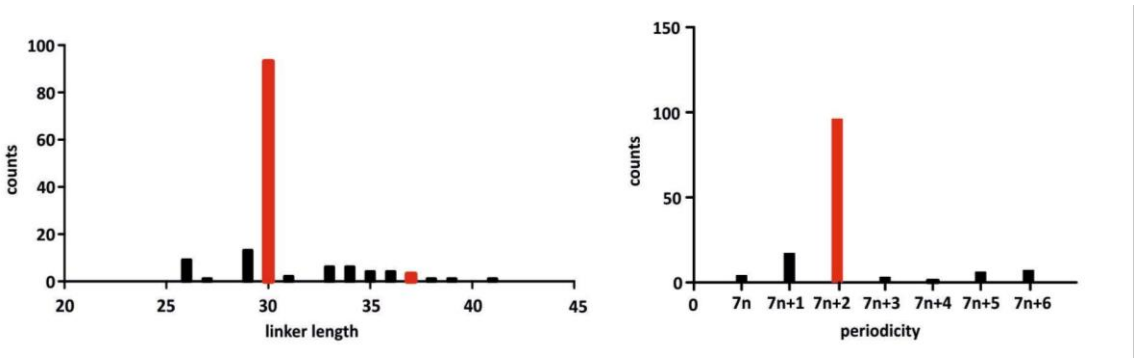


Figure 18: Linker length distribution for naturally occurring PAS-ANTAR signaling receptors comprising at least one of each domains from the Pfam PAS clan CL0183 and the family of ANTAR proteins PF03861**.** For the left figure, the x axis indicates the number of residues between the conserved parts of sensor and effector, defined by the aspartate of the 'DIT' motif at the C-terminus of the PAS modules and the alanine of the 'Q'AKG motif within the ANTAR core domains. The right figure displays the linker length distribution according to the heptad periodicity. The predominant linker lengths and thereof resulting heptad period are marked in red. However, only one predominant linker length can be observed, from which no heptad periodocity can be derived. A possible reason for this is the restricted number of available PAS-ANTAR protein sequences (n < 150).

Table 22: Overview of the AmiR-derived LOVA chimera designed upon domain replacement with the YF1 LOV domain, as well as the tested truncated variants.

Name	Composition
LOVA	YF1 resi 1 - 138 + AmiR resi 142 - 197
-4 - LOVA	YF1 resi 1 - 138 + AmiR resi 146 - 197
-3 - LOVA	YF1 resi 1 - 138 + AmiR resi 145 - 197
-2 - LOVA	YF1 resi 1 - 138 + AmiR resi 144 - 197
-1 - LOVA	YF1 resi 1 - 138 + AmiR resi 143 - 197
+1 - LOVA	YF1 resi 1 - 138 + AmiR resi 141 - 197

+2 - LOVA	YF1 resi 1 - 138 + AmiR resi 140 - 197
+3 - LOVA	YF1 resi 1 - 138 + AmiR resi 139 - 197
+4 - LOVA	YF1 resi 1 - 138 + AmiR resi 138 - 197
+5 - LOVA	YF1 resi 1 - 138 + AmiR resi 137 - 197
+6 - LOVA	YF1 resi 1 - 138 + AmiR resi 136 - 197
+7 - LOVA	YF1 resi 1 - 138 + AmiR resi 135 - 197
+8 - LOVA	YF1 resi 1 - 138 + AmiR resi 134 - 197
+9 - LOVA	YF1 resi 1 - 138 + AmiR resi 133 - 197
+10 - LOVA	YF1 resi 1 - 138 + AmiR resi 132 - 197
+3-ccANTAR-AmiR*	AmiR resi 128 - 197
+10-ccANTAR-AmiR	AmiR resi 121 - 197
+17-ccANTAR-AmiR	AmiR resi 114 - 197
+24-ccANTAR-AmiR	AmiR resi 107 - 197
+31-ccANTAR-AmiR	AmiR resi 100 - 197
flAmiR	AmiR resi 1 - 197

Therefore, linker variants shifted in both senses were designed and cloned (see Table 22), covering the range of position -4 to +10 in respect to the linker length of the initial LOVA-construct that refers to positon ‘zero’. The activity of all so far designed LOVA variants, as well as of the positive control, lies within the basal level (see Figure 19.a). The so far designed chimeras cover the range of position -4 to +10 in respect to the linker length of the initial LOVA-construct, which refers to positon ‘zero’. For EutV, an ANTAR protein comprising an N-terminal RR domain, Ramesh et al. previously reported that the RNA binding affinities for different variants of the protein varied strongly depending on the included receptor regions. For full-length EutV, the determined binding affinity was within a micromolar range (comparable to the determined affinities of the LOVA chimera and the ANTAR-AmiR control). In addition, two truncated variants of EutV were investigated: one variant that comprised ANTAR domain plus the coiled-coil region (ccANTAR-EutV), and another variant comprising only the C-terminal ANTAR domain (ANTAR-EutV; see Figure 41 in Section 6.1.1). The binding affinity determined in EMSA experiments was around 100-fold higher for ccANTAR-EutV compared to ANTAR-EutV alone, indicating that the coiled-coil region is important for functional EutV-RNA interactions. Therefore we decided to test if elongating the N-terminal coiled-coil region of the AmiR ANTAR domain would similarly improve its RNA-binding affinity, and similarly created varied truncated versions of AmiR comprising the ANTAR domain and increasing parts of the coiled coil (ccANTAR-AmiR, see Figure 19).

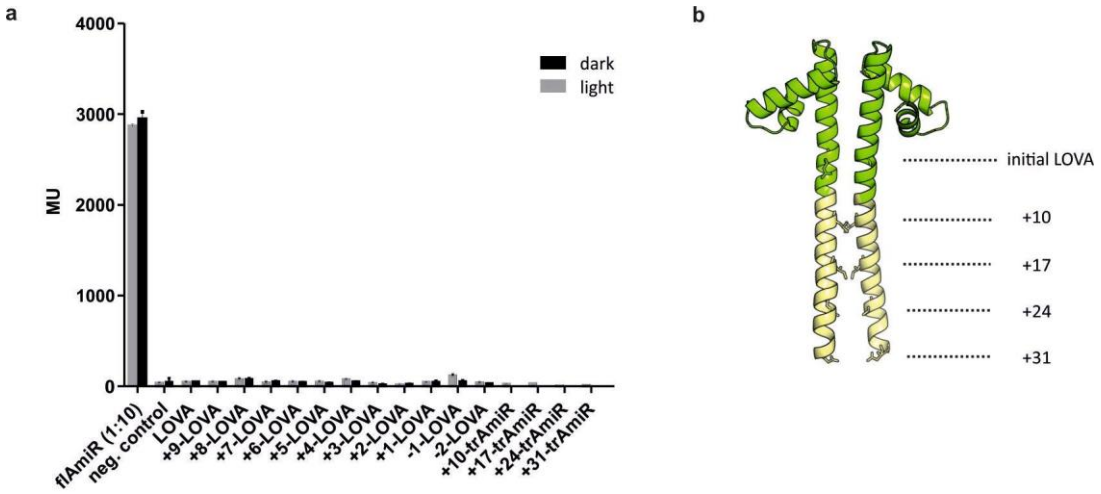


Figure 19: Activity test of LOVA variants. **(a)** Evaluation of the Miller assay of LOVA linker variants and truncated AmiR versions. **(b)** AmiR ANTAR domain (green) and elongated coiled coil (pale green) region for illustration of the different versions of the truncated ccANTAR-AmiR constructs (+10-, +17-, +24 and +31-ccANTAR-AmiR) tested in (a). The numbering of the added residues corresponds to the fusion site of the initial LOVA chimera and goes from C- towards the N-terminus, so that the +10-, +17-, +24 and +31 positions correspond to L132, I125, V118, and D111 respectively, shown in stick optics.

These new constructs were similarly tested with the help of the Miller assay, in which the increased binding affinities would be translated in higher MUs. The numbering of the added residues corresponds to the fusion site of the initial LOVA chimera and runs from the C- towards the N-terminus. The initial positive control comprising just the ANTAR part of AmiR corresponds to the +3 position of ccANTAR-AmiR within this numeration. From here the coiled-coil was elongated in heptades, resulting in four further constructs: +10-, +17-, +24- and +31-ccANTAR-AmiR (see Table 22 and Figure 19.b). However, even the longest of these variants ('+31-ccANTAR-AmiR'), which contained the full extent of the AmiR coiled-coil did not display any increase activity within the Miller assay. The efforts of designing further fusion constructs based on superposition of the AmiR ANTAR domain and the YF1 LOV domain were therefore abandoned at that stage.

5.1.2 Association-/ dissociation-based design approach

In a recent literature study of previous engineering photoreceptors [6], we identified association-based design approaches as a particularly promising for the development of new photoreceptors. This is explained by the frequent occurrence of oligomerization processes in signaling cascades, as well as by lower demands on the connecting linker element of sensor and effector domain. For association-based design approaches, it is often sufficient if the linkers is short, flexible and

hydrophilic. Moreover, this approach is even suitable for proteins that were not originally regulated by oligomerization processes. As the attempt to achieve light-control over the AmiR ANTAR domain through structure-based design approaches failed, the substitution of the YF1 LOV domain by light-inducible associating LOV domains seemed to be an attractive alternative. Several LOV domains exist, for which light-inducible association of the functional homodimer has been successfully implemented, such as the aureochrome LOV domains found in diverse Stramenopiles, or the LOV domain from *N. crassa* Vivid [138] (see Section 3.3.2 for more details).

On grounds of these successful examples, both of the above mentioned LOV domains were selected for fusion with the ANTAR module of AmiR. The Vivid LOV sequence was derived from resi 37–186 of *N. crassa* Vivid ('VVD-LOV'; Uniprot entry Q9C3Y6 with N56K C71V substitution). Among the different available aureochrome LOV domains, the one from *N. gaditana* hypothetical protein NGA_0015702 ('NgAur-LOV'; resi 87–228 of Uniprot entry K8Z861) was selected. For fusion with the ANTAR domain, two different fusion sites were chosen: A112, which marks the end of the elongated coiled-coil, and S90, which includes the first β-strand as well as the first α-helix of the AmiR RR domain (see scheme in Figure 20.a and Table 23).

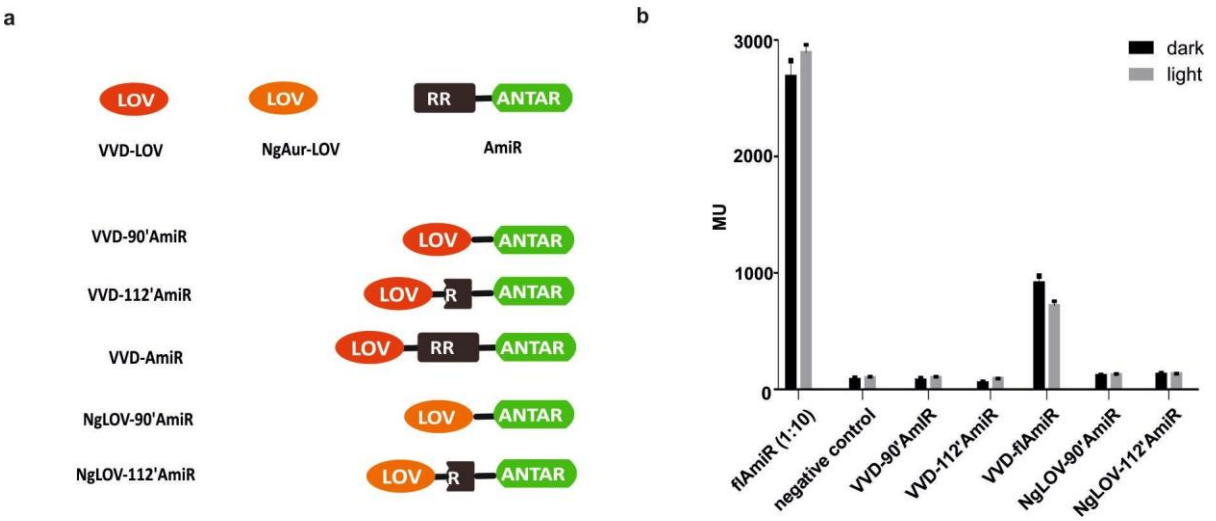


Figure 20: Association-based design approach. **(a)** Scheme of the associating LOV-AmiR chimeras designed for this approach. The VVD LOV sequence was derived from *N. crassa* vivid (VVD-LOV, residue 37–186 of Uniprot entry Q9C3Y6 with N56K and C71V substitution), whereas the aureochrome LOV domain was derived the from *N. gaditana* hypothetical protein NGA_0015702 sequence (NgAur-LOV, residue 87–228 of Uniprot entry K8Z861. Within the AmiR ANTAR effector, two different fusion sites were chosen: A112, which marks the end of the elongated coiled-coil, or S90, which includes the first β-strand as well as the first α-helix of the AmiR RR. The two LOV domains were fused N-terminally to these sites via a short flexible linker composed of ten hydrophilic residues ('DSAGSAGSAG'), resulting in four novel LOV-AmiR chimeras, as well as an additional control for which the VVD LOV domain to the flAmiR protein via the same flexible linker. **(b)** Miller screening assay results for the AssLOV constructs shown in (a).

Table 23: Overview of the AmiR-derived chimera designed employing associating LOV modules.

Name	Composition
VVD-90‘AmiR	VVD resi 37 - 186 + linker + AmiR resi 90 - 197
VVD-112‘AmIR	VVD resi 37 - 186 + linker + AmiR resi 112 - 197
VVD- AmiR	VVD resi 37 - 186 + linker + AmiR resi 1 - 197
NgLOV-90‘AmiR	NgAur resi 87 - 228 + linker + AmiR resi 90 - 197
NgLOV-112‘AmiR	NgAur resi 87 - 228 + linker + AmiR resi 112 - 197
flAmiR / positive control	AmiR resi 1-197

The two LOV domains were fused N-terminally to these sites via a 10 residue long short flexible linker (‘DSAGSAGSAG’), resulting in four novel LOV-AmiR chimeras. In addition, a fifth construct was constructed as additional control, since none of the so far designed YF1 LOV-AmiR fusion constructs led to a detectable signal within the Miller screening system. For that purpose, the VVD-LOV domain was added to the flAmiR protein via the same flexible linker (‘VVD-AmiR’; see Figure 20). In the following screening via Miller assay, only that last construct (‘VVD-flAmiR’) resulted in a detectable signal demonstrating constitutive activity under both dark and light conditions. Compared to the positive control (flAmiR) the activity of the VVD-flAmiR chimera is nonetheless reduced by a factor of almost 30. A possible explanation might be a reduced protein expression of the VVD-flAmiR chimera due the N-terminal attachment of the VVD LOV domain. The activity for the remaining four LOV-AmiR chimeras was in the range of the negative control, which made it difficult to extract any further directions for additional engineering approaches out of these experiments. As we had further just discovered a natural photoreceptor in the sequence databases with the desired domain combination of a LOV and an ANTAR domain (see below/ Section 5.1.3), the attempt to design further chimeric variants of the AmiR ANTAR module in combination with associating LOV domains was aborted.

5.1.3 Discovery of a new design template

The strategies for the construction of an RNA-binding photoreceptor described in Section 5.1.1 and 5.1.2 did not lead to any major success with regard to the activity and light-sensitivity of the engineered photoreceptor chimeras. The design of the LOV-AmiR chimeras from Section 5.1.1 and 5.1.2 was oriented on the more common PAS-ANTAR architectures within which the effector is C-terminally coupled to the PAS sensor module. The discovery of a gene bank entry from the Gram-positive bacterium *Nakamurella multipartita*, whose protein product comprises a PAS domain followed by an ANTAR, and a LOV domain at the C-terminus (hence termed 'PAL'; see Section 5.2), therefore strongly attracted our attention. Based on the domain architecture PAL, we suspected that

the RNA-binding function of the ANTAR domain might be subjected to the control of the blue-light-sensitive LOV domain. The discovery of PAL thus opened up a completely new design approach for the engineering of a light-sensitive ANTAR receptor, based on domain replacement of the PAL ANTAR domain by the ANTAR module of AmiR.

The pairwise alignment of the two ANTAR sequences (PAL resi 125– 181, Uniprot ID C8XJT7; AmiR resi 146 – 196, Uniprot ID P10932) with the tool 'EMBOSS Needle' [130] results in an alignment score of 29, 14.5 % identity, and 33.9 % similarity (see Figure 21.a). For the planning of the new PAL-AmiR chimera, the domain boundaries of the ANTAR domains of AmiR and PAL were derived from the multiple sequence alignments of the overall family of ANTAR proteins generated by Pfam [98], as multiple sequence alignmenst are much more reliable for the identification of domain boundaries than pairwise alignments that easily ignore eventual insertions within the core domain. The fusion site was set behind the last C-terminal residue of AmiR (A196). As each fusion site increases the risk of disrupting the functionality of the chimera (e.g., by destroying secondary structure motifs often resulting in protein folding problems), we decided to replace all residues N-terminal to the fusion site with the complete AmiR protein including the N-terminal RR instead of substituting only the isolated ANTAR core domains. This first fusion construct hence contains resi 1 – 196 of AmiR and resi 187 of PAL and was termed ‚AmiLOV' (see Figure 21.a and Table 24). Furthermore, we decided to explore the effect of another domain replacement with the LOV domain of the *Nannochloropsis gaditana* Aureochrome ('NgAur'; *N. gaditana* NGA_0015702; Uniprot ID K8Z861). Similar to PAL, the bZIP effector within the Aureochrome is N-terminally attached to the LOV sensor. For that reason, we wanted to investigate if the signal propagation from sensor to effector within the Aureochrome could be based on similar signaling mechanisms as in PAL. The fusion construct, termed 'AmiNgLOV', was designed based on a multiple sequence alignment of several LOV domains and comprises the N-terminal resi 1 - 246 from AmiLOV, as well as resi 91 – 228 from NgAur (see Figure 21.b and Table 24).

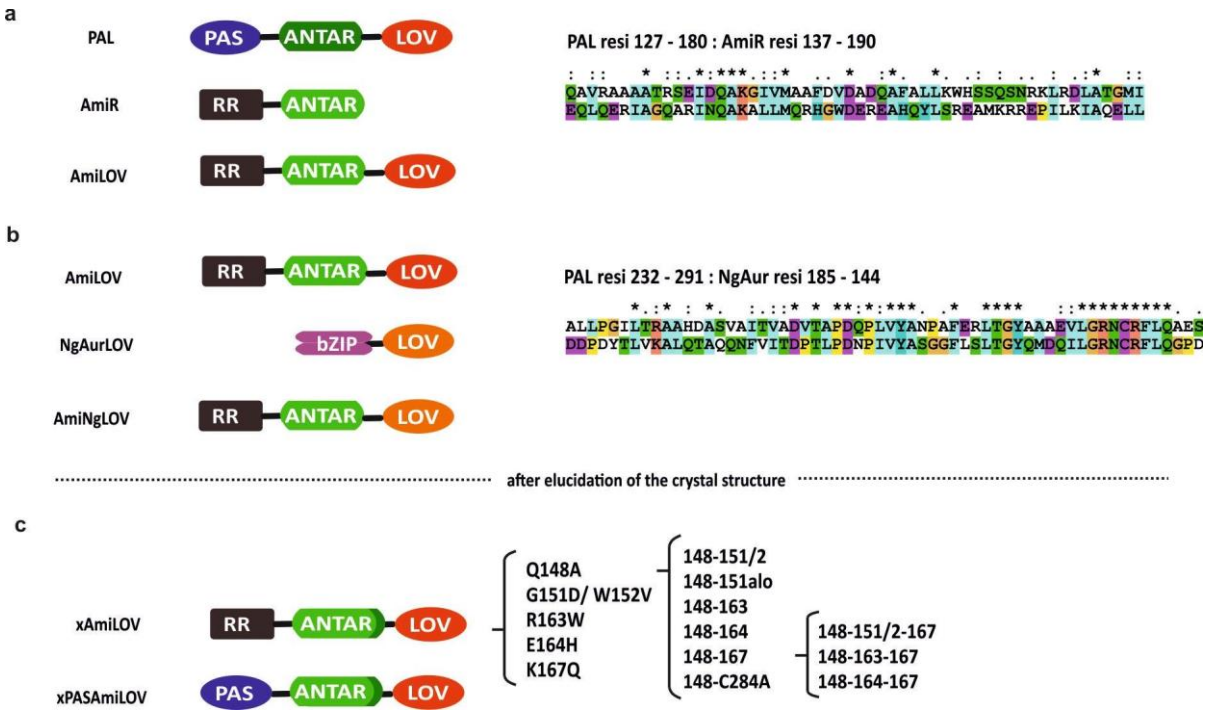


Figure 21: Alternation of the sequence specificity via exchange of the ANTAR domain **(a-c)** Schematic overview of the designed chimera variants: **(a)** The initial AmiLOV chimera was derived from a sequence alignment of the ANTAR domains from AmiR and PAL; since no further domains or residues follow C-terminally downstream of AmiR, the fusion site was set behind the last AmiR residue (A196), thereby replacing all PAL residues N-terminally to the fusion site with the complete AmiR protein comprising the N-terminal RR domain. **(b)** The AmiNgLOV cimera was subsequently derived from a domain exchange of the PAL LOV domain with the LOV domain of the Aureochrome from *N. gaditana* . **(c)** The elucidation of the crystal structure enabled the adjustment of the fusion sites with the help of a model resulting from structural superposition of the AmiR and PAL ANTAR domains. The structural model was then employed to identify single residues, whose substitution might lead to an improvement of the chimera activity.

Table 24: Overview of the chimera designed upon the discovery of PAL.

Name	Composition
AmiLOV	AmiR resi 1 – 196 + PAL resi 187 - 365
AmiNgLOV	AmiLOV resi 1 - 247 + NgAur resi 91 - 228
35-shlinkAmiLOV	AmiR resi 1 – 196 + PAL resi 199 - 365
21-shlinkAmiLOV	AmiR resi 1 – 196 + PAL resi 213 - 365
7-shlinkAmiLOV	AmiR resi 1 – 196 + PAL resi 227 - 365
7-shlinkAmiNgLOV	AmiR resi 1 – 196 + NgAur resi 80 - 228
xAmiLOV	AmiR resi 1 – 187 + PAL 178 - 365
xPASAmiLOV	PAL resi 1 - 134 + AmiR resi 145 – 187 + PAL resi 178 -365
flAmiR / positive control	AmiR resi 1-197

A positive outcome from the engineering efforts described in Section 5.1.1 was the construction of a functional reporter system sensitive to the effector activity of AmiR, which enables the fast and

simple screening of chimeric variants without the laborious steps of protein expression and purification. The gene sequence of AmiLOV and all following chimeras was therefore cloned into the same pACYC184 backbone as employed for the testing of the AmiR-derived variants from Section 5.1.1, which is compatible to the pE_lacZ and pE_DsRed reporter plasmids (see Figure 16 and Table A3

in Section 8.2). A positive control comprising the flAmiR protein, as well as a negative control composed of the empty library_pKT plasmid, were employed as a standard throughout the tests using the reporter assay. For functional screening, the AmiR chimera variants and the pE_DsRed reporter construct were co-transformed into competent CmpX13 cells. The assay was then conducted as described in Section 4.3.1. The AmiLOV and the AmiNgLOV chimera demonstrated significantly higher activity levels within the reporter assay than all other AmiR-derived chimeras engineered so far. The activity levels of the initial AmiLOV and AmiNgLOV chimeras are approximately 10 % and 20 % of the positive control. However, neither of the two chimera variants displayed any difference between light and dark activity throughout the initial tests within the bacterial reporter assay (Figure 23.a). In order to assess whether a variation of the linker length would affect the light sensitivity of the constructs, we created chimeric variants with shortened linkers (‘shlink' variants, designated according to length of remaining linker residues between the ANTAR and LOV core domain boundaries; e.g. 7/ 21/ 35-shlink-AmiLOV; see Table 24). While linker length variations of the AmiLOV variants resulted in no further influence on the chimera activity, linker length shortening of the AmiNgLOV construct (‘7shlink-AmiNgLOV’) led to first light-induced differences in chimeric activity (ratio of light over dark activity ‘L/D’ = 1.75; see Figure 23.a) within the tests using the reporter assay.

In the meantime, the elucidation of the PAL crystal structure (see Section 5.2.6) led to the provision of a model that allowed the structural superposition of the two ANTAR domains, which permitted to inspect the fusion sites previously chosen upon determinants derived from the sequence alignment (see Figure 22). For that purpose, the crystal structures from PAL and the AmiR ANTAR domain (PDB code: 1QO0, resi 137-190) were superposed in PyMOL. We consequently designed and cloned another PAL-AmiR chimera (‚xAmiLOV') based on the insights from the PAL crystal structure. Motivated by the now realizable visualization of the fusion interfaces, we designed another chimera in which only the ANTAR core domains were exchanged ('xPASAmiLOV'), which led to an additional fusion site. The latter displayed an activity level comparable to the negative control, i.e. no detectable activity within the activity tests using the reporter assay. In contrast, the xAmiLOV chimera showed a significant light-induced increase in activity after the displacement of the fusion interface (light activity $(L)^{xAmiLOV}$ = 1980 ± 74; dark activity $(D)^{xAmiLOV}$ = 1070 ± 40; $L/D^{xAmiLOV}$ = 1.85).

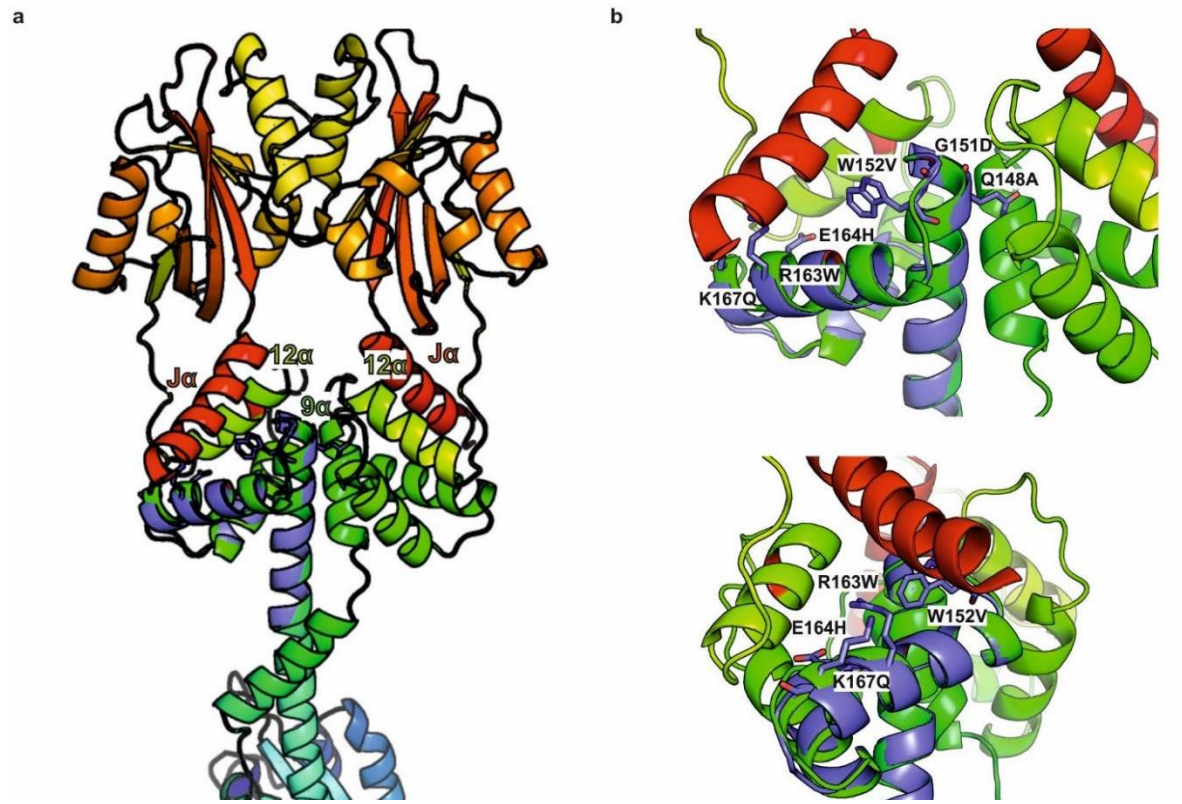
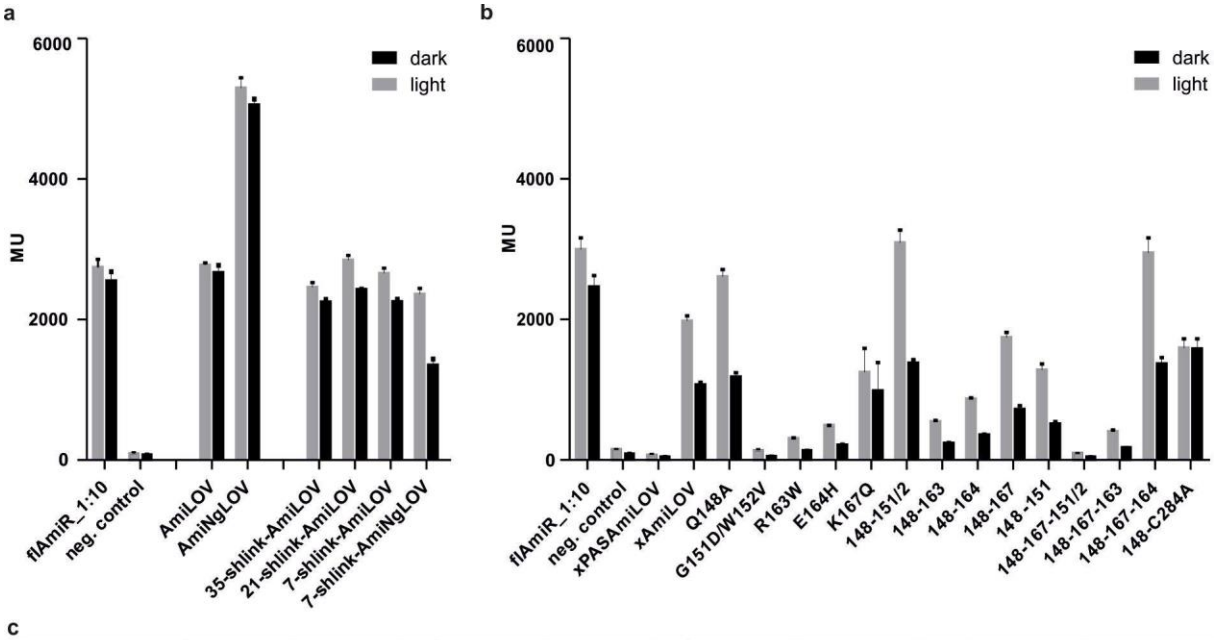


Figure 22: Structural model for the design of the xAmiLOV variants. **(a)** Structural overlay of the crystal structures from PAL and AmiR ANTAR domains (highlighted in blue**). (b)** Zoomed-in view of the mutated residues of the xAmiLOV chimera produced within this study. The Q148A point mutation results in the clearest effect on the activity and light-sensitivity of the chimera. In the structural model, the glutamine residue Q148 previously interfered with the coiled-coil interface of the PAL 9〈 helices.

We now attempted to achieve further improvements of the xAmiLOV chimera through the introduction of point mutations with the help of the structural overlay. For a start, I concentrated on: (i) residues, which could interfere with the coiled-coil interface between the 9α helices of PAL (Q148A); (ii) bulky hydrophobic residues which might interfere with the Jα interface (W152V; R163W); or, (iii) the conservation of charged residues which might be involved in signaling based on their proximity to other charged residues (G151D, E164H). I also mutated a further lysine residue (K167Q) in proximity to E164, as its charge might interfere with the introduced E164H mutation. All single point mutations were created in multiple combinations and tested as a pool with the help of the reporter assay (see Figure 23.b), the naming of mutants based on the introduced mutations always refers to the xAmiLOV chimera.

Among the single point mutations, only Q148A results in a modest improvement of overall activity and L/D ratio compared to the original xAmiLOV chimera (L^{Q148A} = 2600 ± 101; D^{Q148A} = 1180 ± 61; L/D^{Q148A} = 2.21). All other introduced point mutations led to sharp reductions of the overall activity (see Figure 23.b/c). Interestingly, in many cases, combining the Q148A mutation with the remaining single mutations results in suppression of this negative effect. For xAmiLOV-Q148A-G151D-W152V (abbreviated '148-151/52'), the elimination of the bulky tryptophan residue (W152V) also leads to a significant increase in light activity accompanied by a slight reduction in the L/D ratio ($L^{148-151}$ = 1280 ± 75, $D^{148-151}$ = 510 ± 33, $L/D^{148-151}$ = 2.49; vs. $L^{148-151/52}$ = 3100 ± 182; $D^{148-151/52}$ = 1380 ± 50, $L/D^{148-151/52}$ = 2.24). The K167Q mutation was based on the consideration that the positive charge of lysine could interfere with the negative charge introduced by E164. In the activity tests, the combination of Q148A-E164H -K167Q mutations (148-167-164) also led to a marked increase in light activity, accompanied by a slight decrease in the L / D ratio ($L^{148-164}$ = 870 ± 20, $D^{148-164}$ = 360 ± 12, $L/D^{148-164}$ = 2.42; vs. $L^{148-164-167}$ = 2900 ± 217, $D^{148-164-167}$ = 1370 ± 89; $L/D^{148-164-167}$ = 2.15). As an additional control for the reliability of the reporter assay, the active site cysteine of xAmiLOV-Q148A was mutated to alanine. According to expectations, the resulting mutant (xAmiLOV-Q148A-C284A) displayed a complete loss of light-sensitivity, but an intermediate activity level between dark and light-activated state of the inititial xAmiLOV-Q148A variant ($L^{148-C284A}$ = 1600 ± 134; $D^{148-C284A}$ = 1600 ± 120; $L/D^{148-C284A}$ = 1.01).



c

construct	L (mean)	u_L (SD)	N	D (mean)	u_D (SD)	N	L/D	$u_{L/D} = ?(A^2+B^2)$	→	A = (1/D*u_L)²	B = (-L/D²*u_D)²
flAmiR_1:10	2744,06	113,00	3	2548,69	225,54	3	1,08	0,00929		0,00197	0,00908
neg. control	94,29	16,56	3	79,81	19,06	3	1,18	0,09052		0,04306	0,07962
AmiLOV	2776,68	28,87	3	2670,36	173,84	3	1,04	0,00458		0,00012	0,00458
AmiNgLOV	5295,11	145,18	3	5053,94	146,91	3	1,05	0,00124		0,00083	0,00093
35-shlink-AmiLOV	2457,69	67,71	3	2249,90	63,84	3	1,09	0,00132		0,00091	0,00096
21-shlink-AmiLOV	2844,63	67,42	3	2429,67	2,91	3	1,17	0,00077		0,00077	0,00000
7-shlink-AmiLOV	2657,72	75,94	3	2254,40	61,77	3	1,18	0,00154		0,00113	0,00104
7-shlink-AmiNgLOV	2361,18	81,81	3	1350,72	150,35	3	1,75	0,03804		0,00367	0,03786
flAmiR_1:10	2996,72	162,61	3	2464,89	159,93	3	1,22	0,00759		0,00435	0,00622
neg. control	155,11	3,64	3	93,98	9,92	3	1,65	0,03042		0,00150	0,03038
xPASAmiLOV	78,06	9,88	3	52,71	6,02	3	1,48	0,04532		0,03513	0,02863
xAmiLOV	1977,75	74,83	3	1068,95	39,74	3	1,85	0,00681		0,00490	0,00473
Q148A	2608,37	100,98	3	1181,26	61,39	3	2,21	0,01506		0,00731	0,01317
G151D/W152V	139,54	16,13	3	61,68	3,65	3	2,26	0,07068		0,06837	0,01792
R163W	306,93	8,15	3	141,35	5,55	3	2,17	0,00800		0,00332	0,00727
E164H	492,18	2,29	3	211,48	13,62	3	2,33	0,02245		0,00012	0,02245
K167Q	1248,07	340,62	3	981,72	406,97	3	1,27	0,30272		0,12039	0,27776
148-151/2	3088,78	182,04	3	1378,44	50,27	3	2,24	0,01868		0,01744	0,00668
148-163	548,16	16,59	3	238,32	10,43	3	2,30	0,01124		0,00485	0,01014
148-164	866,22	19,85	3	358,10	12,41	3	2,42	0,00767		0,00307	0,00703
148-167	1743,71	74,60	3	720,00	54,10	3	2,42	0,03481		0,01074	0,03312
148-151	1278,66	86,63	3	513,35	32,58	3	2,49	0,03789		0,02848	0,02499
148-167-151/2	99,19	2,37	3	50,13	7,02	3	1,98	0,07687		0,00224	0,07684
148-167-163	404,34	21,43	3	181,49	4,30	3	2,23	0,01422		0,01395	0,00279
148-167-164	2944,54	216,78	3	1367,63	88,55	3	2,15	0,03176		0,02513	0,01943
148-C284A	1592,27	134,04	3	1580,75	144,46	3	1,01	0,01111		0,00719	0,00847

Figure 23: Miller screening assay results for the PAL-based design constructs shown in Figure 21. The depicted bars are the mean value of three measurements; the error bar indicates the SD. **(a)** Activities of the initial AmiLOV and AmiNgLOV chimeras, as well as the shortened linker ('shlink') variants, which were named after the length of remaining linker residues between ANTAR and LOV domain (e.g. 7-shlink-AmiLOV). **(b)** Activities of the second chimera generation created upon superposition of the AmiR and PAL ANTAR domain crystal structures ('xAmiLOV' and 'xPASAmiLOV'). The xAmiLOV chimera displays a significant light-induced increase in activity after the structure-based displacement of the fusion interface. Motivated by this initial success, we subsequently used the structural model to introduce targeted point mutations for further optimization of the light-sensitivity of the xAmiLOV chimera. **(c)** Overview of determined light (L) and dark (D) activities along with the corresponding standard deviation for all of the in (a) and (b) tested constructs. As a measure of light-sensitivity, the ratio of light over dark activity (L/D) was determined, along with the associated error ($u_{L/D}$) to assess the accuracy of the results. $u_{L/D}$ was calculated through Gaussian error propagation from the root of the squared sum of the individual uncertainties A and B that result from the errors for L and D. N – number of measured samples.

With a number of light-sensitive constructs, this last approach, oriented on a natural template, is the most successful attempt to construct a photosensitive ANTAR receptor so far. By replacing the ANTAR domain of the natural photoreceptor PAL with that from AmiR, these new PAL-based chimeras should exhibit an altered sequence specificity. The results of the reporter assay indicate that the chimeras recognize the ami-lead sequence upstream of the reporter gene leading to antitermination. However, for most chimeras the reporter signal is also induced under dark conditions, and only enhanced by light activation. Even for the best-performing chimeras in terms of the resulting signal-strength (e.g. xAmiLOV 148-151/52 or 167-164), the reporter signal is significantly reduced (≈ 10-fold) compared to the AmiR positive control. This suggests that the ability to bind the RNA target sequence was negatively affected by the fusion to the PAL-LOV domain.

5.2 Characterization of the natural photoreceptor PAL

Parallel to the efforts of engineering an ANTAR-containing photoreceptor, we monitored the gene databases for novel emerging LOV photoreceptors in regular intervals. In order to identify potentially interesting LOV-effector combinations, a Python script provided by Andreas Möglich was used to identify distinct LOV proteins among the PAS Clan CL0183 in Pfam 26.0/ 27.0 [135]. The resulting output list was manually examined for promising novel LOV effector occurrences that led to the discovery of a predicted protein with the desired building blocks of a LOV and ANTAR domain (uniprot entry C8XJT7). This so far unique combination of domains was found in the genome of *Nakamurella multipartita*, a species of the monogeneric family Nakamurellaceae within the actinobacterial suborder Frankineae. Further analysis of the gene sequence revealed an additional PAS domain at the N-terminus, turning the complete domain architecture into PAS-ANTAR-LOV, which is why the resulting protein product was named 'PAL'. Based on the domain arrangement of PAL, we suspected that the RNA-binding function of the ANTAR domain could be controlled by the LOV domain. For this reason, we regarded the PAL protein as a promising target for further investigations.

5.2.1 Preliminary proof of function

To ensure that the underlying gene sequence of PAL is not the result of a sequencing artifact, and to exclude covert stop codons within the locus, the nucleotide sequence was extracted from the *N. multipartita* strain Y-104 (DSMZ 44233) via PCR and verified through DNA sequencing ahead of the project start. The sequence was cloned between NcoI and XbaI of the pet28c vector. The testing of various expression conditions provided best results for the Arctic Express (DE3) *E. coli* strain at 12°C for a prolonged period of around 40 hours. For purification, the expression was conducted with 800 mL cultures; Ni-NTA affinity purification, employing buffer B2 (see Table 9 in Section 4.4.3) for washing, resulted in ≈ 1.5 mg of pure and soluble target protein of the expected size of ≈ 40 kDa (see Figure 24.b). For the intended activity assays, the protein was dialyzed against fresh buffer D (see Table 9 in Section 4.4.3), which was found to be the most suitable recipe for ensuring maximum protein stability between all buffer substances and additives tested (in buffers without additional divalent cations and glycerol as co-solvent, protein stability was found to be severely impaired).

As it might be of advantage to separate light-sensing and effector functions for subsequent investigations (such the determination of specific RNA targets or mechanistic investigations), the C-terminal LOV domain ('PAL-LOV'; comprising residues 220 to 365 of the PAL sequence, Uniprot C8XJT7; see Figure 24.a) and the N-terminal PAS-ANTAR segment (PAL-[PA]; comprising residues 1 to 198 of the PAL sequence, Uniprot C8XJT7; see Figure 24.a) were likewise cloned into pet28c and

expressed and purified under similar conditions as the full-length PAL photoreceptor. Expression and purification of the isolated LOV domain yielded soluble target protein with the expected size of ≈ 15 kDa. In contrast, the PAL-[PA] construct showed a strongly reduced expression and poor solubility, resulting in several weak bands, including one with the expected size of ≈ 25 kDa (see Figure 24.b).

The ANTAR domain should exhibit a preference for single stranded RNA molecules among different types of nucleotides. To provide initial evidence of the suspected function of PAL, an EMSA with different kinds of unspecific nucleotide sequences, i.e. single stranded (ss) or double stranded (ds) RNA or DNA, was performed (see Figure 24.c). The ribonucleotide sequence of the utilized oligo was 5'-GUGAUCCAACCGACGCGACAAGCUAAUGCAAGA-3' for ssRNA; the DNA oligo was of similar sequence except for desoxyribonucleotides with 'T's (thymin) instead of 'U's (uracil). For the experiments with ds oligos, a reverse complementary strand was added and annealed with the forward strand before conduction of the EMSA. Qualitative comparison of the four different gels revealed a binding preference for ssRNA (see Figure 24.c (4)), providing a first proof of proper ANTAR domain functioning. To see if the association to the RNA is influenced by different light conditions, another EMSA experiment employing only the ssRNA oligo was conducted under dark and under continuous blue light illumination conditions, which did not reveal any differences in binding behavior. We reasoned that this could be due to the use of an unspecific RNA sequence and that the detection of light-regulated binding might require the identification of more specific RNA targets.

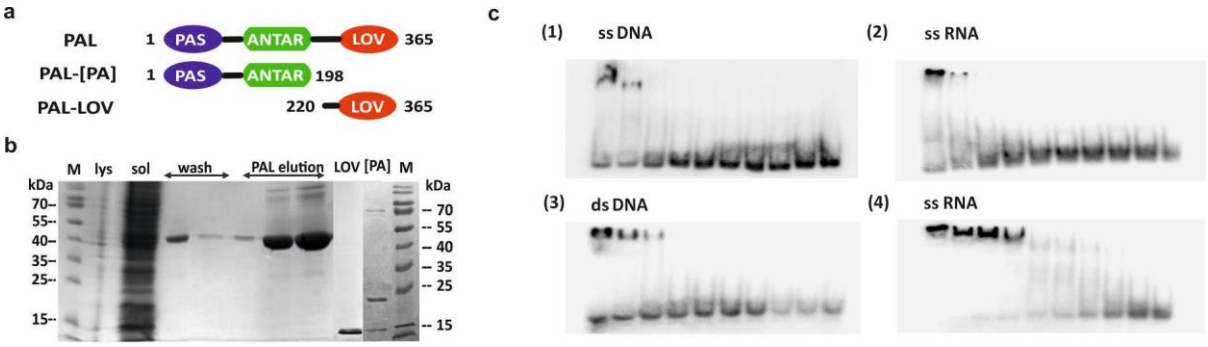


Figure 24: Recombinant protein expression of PAL and it subdomains and preliminary proof of predicted ANTAR domain function. **(a)** Schematic depiction of the investigated PAL constructs. The nucleotide sequence was derived from the *N. multipartita* strain Y-104 genome and corresponds to Uniprot entry C8XJT7. **(b)** SDS-Page analysis (Coomassie-stained) of the different PAL-constructs purified from Arctic Express (DE3). The shown elution fractions from PAL/ PAL-LOV and PAL-[PA] originate from separate analyses. The full-length PAL protein appears as strong band at the expected size of ≈ 40 kDa, while the isolated PAL LOV domain appears at ≈ 15 kDa. The PAL-[PA] construct showed only weak expression and solubility. M: protein ladder; lys: lysate fraction; sol: soluble fraction; wash: washing fractions; elution: elution fractions; [PA]: PAL-[PA] elution; LOV: PAL-LOV elution. **(c)** EMSA with varying concentrations of PAL (16/8/4/2/1/0.5/0.25/0.125/0.0625/0 µM) in presence of ≈ 50 pM of either single stranded (ss) or double stranded (ds) unspecific DNA and RNA oligos of the same sequence: (1) ssDNA; (2) ssRNA; (3) dsDNA; (4) ssRNA. Qualitative comparison of the four different gels reveals a binding preference for ssRNA (4) as expected for an ANTAR-domain containing signaling receptor.

5.2.2 Photochemical characterization

To assess the functionality of the light-sensing LOV module, UV-Vis spectra were recorded for both the dark-adapted and signaling state. The UV-Vis absorption spectrum of the dark-adapted PAL protein showed the characteristic fine structure of the oxidized flavin chromophore with a main absorption band at 447 nm and two minor peaks. Upon blue light illumination, these absorption bands decreased to a 390 nm maximum characteristic of the signal state, indicating the formation of the covalent bond of the photoproduct (see Figure 5 in Section 3.2.3). The return to the dark-adapted state took less than 30 minutes at 22 °C; more detailed investigations of the dark reversion kinetics of PAL and the isolated PAL-LOV domain resulted in data curves that could be fitted by a monoexponential function (see Figure 25.b). For the dark recovery of fl PAL, a time constant (τ) of 1270 ± 100 s was determined at 25 °C, whereas the isolated PAL LOV domain recovered in 470 ± 40 s from S to D at the same temperature, indicating that the presence of adjacent PAS and ANTAR domains decelerates the dark recovery rate of the isolated PAL LOV domain. To analyze the effect of temperature, the recovery rates for PAL and PAL-LOV and were recorded at 10 °C, 15 °C, 20 °C, 25 °C and 30 °C, and monoexponentially fitted to determine the values for τ , as well as of its reciprocal, which yields the rate constant k. To determine the activation energy (E_a) for PAL and PAL-LOV, the logarithm of k was then plotted against the inverse temperature (1/T in K^{-1}) in an Arrhenius plot (see Figure 25.c). The thereby obtained values of E_a are 60.6 ± 3.3 kJ/ mol for PAL and 85.2 ± 8.5 kJ/ mol for PAL-LOV.

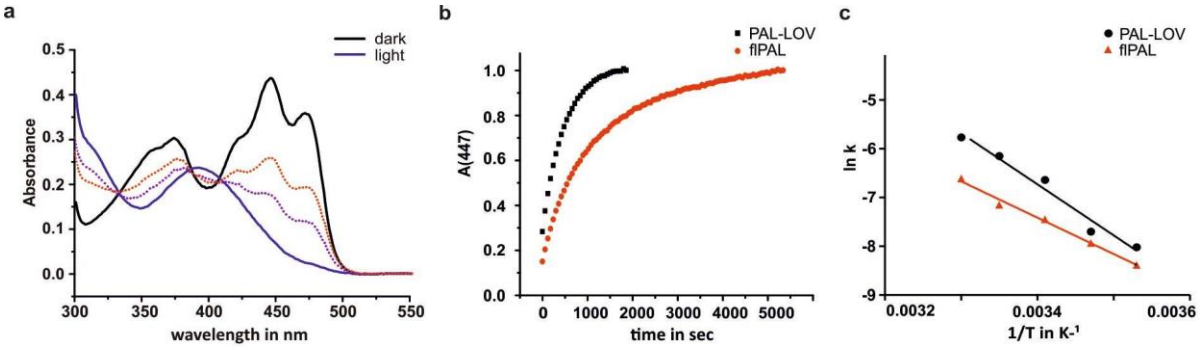


Figure 25: Photochemical characterization of PAL and PAL-LOV. **(a)**UV-Vis Spectra of PAL. In the dark (black line), the spectrum adapts the typical fine structure of an oxidized flavin chromophore with main absorption at 447 nm. Light activation induces formation of a new maximum at 390 nm (blue line), characteristic for the photoproduct of the signaling state. Spectra of intermediate states during dark reversion are shown in dashed lines. **(b)** Recovery kinetics of PAL and PAL-LOV by measuring the absorbance at 447 nm A(447) after blue light-activation at 25 °C. Exponential fitting of the time traces yields time constants (τ) of 1270 ± 100 s for fl PAL and 470 ± 40 s for PAL-LOV. The time constants were calculatesd from the mean of two independent measurements. **(c)** Arrhenius plot for determination of the activation energies (E_a) for PAL and PAL-LOV, in which the logarithm of the rate constant k was plotted against the inverse temperature (1/T in K^{-1}). The thereby obtained values are 60.6 ± 3.3 kJ/ mol for PAL and 85.2 ± 8.5 kJ/ mol for PAL-LOV.

Since a PAL-LOV variant with a slowed-down adduct decay could be of advantage for further experiments (such as the intended investigation of the functional oligomeric state of the photoreceptor), we looked for suitable residue mutations that have slowed down the off-kinetics of other LOV proteins. For this purpose, I created an alignment with further LOV proteins, which had already been tested for the effect of individual residual mutations in former studies (see Figure 26). I focused on the T250 position (or T32 in the isolated PAL LOV domain) whose mutation to more bulky, hydrophobic residues had a strong decelerating effect on the off-kinetics in other LOV photoreceptors [139]. Initially, three different variants were designed and expressed (T32V-/ T32I-/ T250L-PAL-LOV), but only the T32V PAL-LOV variant resulted in functional expression of the holoprotein. The T32V PAL-LOV protein was subsequently characterized via UV-Vis spectroscopy. Evaluation of the recovery kinetics after light-activation, measured at an absorbance of 447 nm at 20°C, yielded a time constant of 1980± 10 s at 20 °C for the T32V PAL-LOV mutant (see Figure 26.b). The reversion time of T32V PAL-LOV is hence decelerated about 2.5-fold compared to wild-type PAL-LOV at 20° C ($\tau^{\text{PAL-LOV}} = 765 \pm 5$ s), and the effect of the T32V mutation is thus in line with expectations.

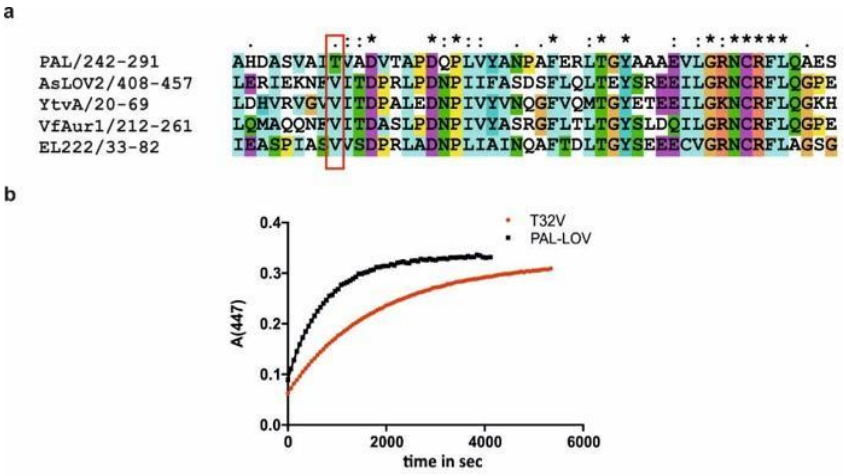


Figure 26: Design and characterization of a decelerated dark reversion variant. **(a)** Alignment of different LOV proteins. The mutated T250 residue is framed in red. **(b)** Recovery kinetics of PAL-LOV and T32V PAL-LOV by measuring the absorbance at 447 nm (A(447)) after blue - activation at 20 °C. Exponential fitting of the time traces yields reversion rates () of 765 ± 5 s for PAL-LOV and 1980 ± 10 s for T32V PAL-LOV.

5.2.3 Identification of specific RNA-targets via SELEX

To determine specific RNA target sequences for PAL, we applied Systematic Evolution of Ligands by EXponential enrichment (SELEX). For the conduction of the SELEX experiments we collaborated with the laboratory of Günter Mayer at the University of Bonn. The facilities of the Mayer laboratory were not adapted to the work with light-sensitive proteins at that time. For that reason, we initially considered to detach the RNA-binding function from the light-sensing domain to facilitate the

conduction of the experiments. The N-terminal PAS-ANTAR domains were thus separately expressed and purified (see Figure 24), but the yield of expression of the LOV domain-lacking PAL-[PA] construct as well as its stability were strongly reduced compared to the full-length PAL receptor. As both factors are important prerequisites for functional studies, and also the binding behavior of a truncated construct might not fully be conserved, we abandoned the idea of utilizing the PAL-[PA] construct for the SELEX experiments and used the full-length protein instead.

The protocol for the SELEX experiments with PAL was established with the help of Anna-Maria Pyka. A typical SELEX cycle of this protocol included the following steps: (i) protein immobilization; (ii) transcription of the oligonucleotide pool; a multistep selection (iii - v): (iii) coincubation of the RNA pool sequences with the immobilized target protein, (iv) a wash protocol whose stringency increases with increasing number of SELEX cycles, and (v) elution of the remaining RNA sequences; and (vi) amplification of the eluted sequences via RT-PCR. Attention was paid to conduct all work under dark or red light conditions.

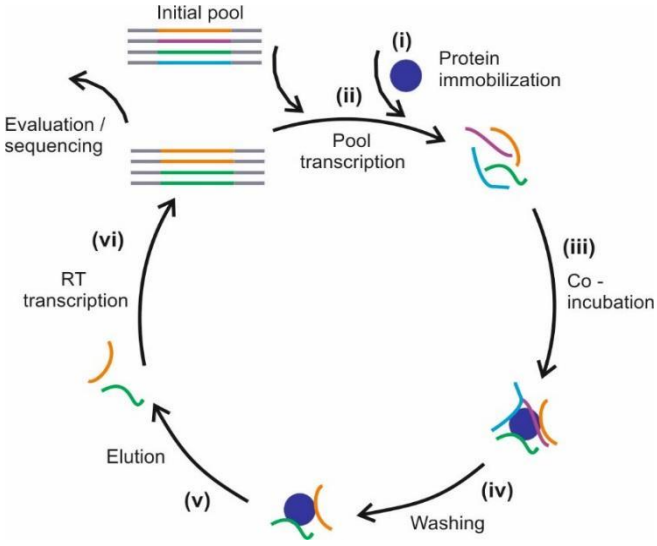


Figure 27: Overview of the SELEX procedure. One SELEX cycle of the established protocoll included the following steps: (i) immobilization of the target protein; (ii) transcription of the RNA pool; (iii) co-incubation of the pool sequences with the immobilized target protein, (iv) washing, (v) elution of the remaining RNA sequences; and (vi) RT-PCR.

The required number of PCR cycles for the reverse transcription of retained RNA sequences at the end of the selection cycle provides a first hint about alterations in the specificity of the pool sequences. The more specific the pool, the more sequences will bind to the target protein and resist the washing steps and thus remain as a starting template for the RT-PCR, resulting in a reduction of the required number of RT-PCR cycles. After the 9th round, the number of RT-PCR cycles was reduced

from 10 to 7 while the number of washing cycles was increased from 1 to 10, so that we decided to start the sequence analysis of the RNA pools. For that purpose, the pool sequences of SELEX round 1, 6, 8, 9 were subjected to a filter-binding assay (realized by Anna-Maria Pyka) to compare the respective affinities to our target protein. The results of the filter-binding assay confirmed that the RNA-protein binding that corresponds to the pool's affinity increased with every SELEX round (see Figure 28.a). Within the last SELEX rounds the increase in overall binding was only slight, so that the RNA-protein binding at the applied conditions seemed close to saturation. Therefore, we decided to terminate the first sequence of SELEX experiments at this point. The reverse transcribed sequences of the 9th SELEX round were then cloned into a sequencing vector and around 50 clones was sent to sequencing (realized by Anna-Maria Pyka). The evolution of the best-binding pool sequences during the SELEX process typically leads to the emergence of conserved sequence motifs that promote the specific interaction with the target. In order to identify related binding motifs within the sequenced aptamers, the sequencing results were scanned using the MEME suite [140] which offers different analysis services for conserved motifs within nucleotide and protein sequences. With the help of the MEME Motif Discovery tool, ninety percent of the sequences could be assigned to two motif families, whose consensus sequences are shown in Figure 28.b.

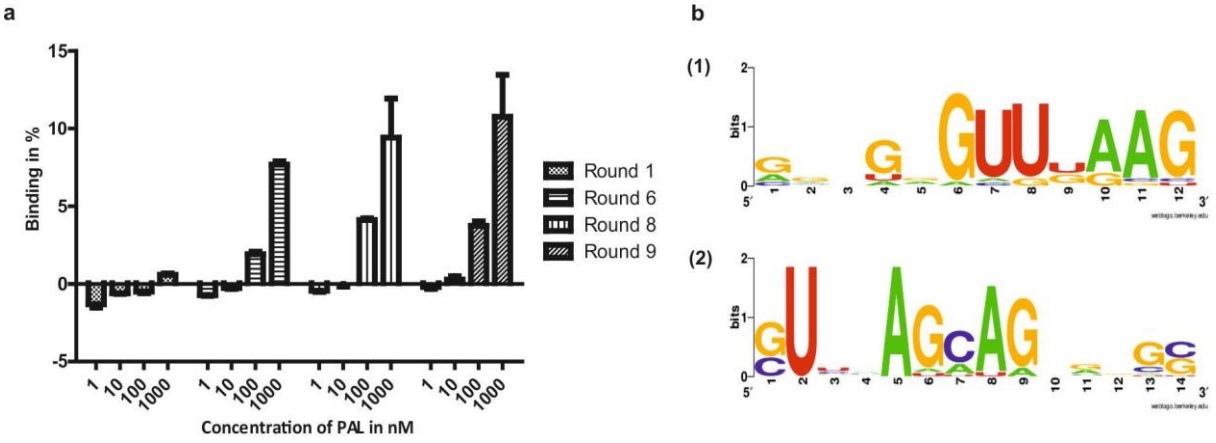


Figure 28: Evaluation of SELEX selection cycles. **(a)** Comparison of the binding affinity of the RNA-pool sequences to PAL at different selection stages. The overall binding affinity of the RNA-pool to PAL increases with an increasing number of SELEX selection cycles. **(b)** Consensus sequence of identified conserved target motifs. The analyzed sequences could be allocated to two motif families using the MEME Motif Discovery tool.

The individual members of each of two motif families were then compared by Anna-Maria Pyka via filter-binding assay to identify the most promising aptamers in terms of specific binding behavior ('apatamer' is a commonly used term for nucleic acid species that have been optimized for binding to a molecular target by repeated rounds of *in vitro* selection). Within motif family (1), the SELEX clones 04, 36 and 46 were identified as best candidates for further characterization, while within motif family (2) the clones 10 and 40 performed best. These five aptamers were then further characterized via EMSA. All initial tests for determination of appropriate experimental parameters were performed with SELEX clone 40. The RNA probe for the EMSA experiment was synthesized via *in vitro* transcription and subsequent 5'-end radiolabeling employing [γ^{33}P]-ATP (see Section 4.7.1). For acquisition of a complete binding curve the concentration range of 0.02 - 5.12 µM ligand [L] on a binary log scale was determined as appropriate. To verify the selectivity of PAL for the aptamer, an excess (10 µg / µl final concentration) of unlabeled tRNAs was added to the sample mixtures. The assay was then performed under safe red light or under continuous blue light illumination to assess effect of light on the binding affinity of PAL to the aptamer. The development of the EMSA gel traces gave immediate evidence that blue light-treatment significantly increases the affinity of PAL for its target RNA. Under dark conditions, almost no binding could be observed within the tested concentration range, whereas blue light treatment resulted in the appearance of strong PAL-RNA bands (see Figure 29.a). Quantitative evaluation of the data revealed a Kd value of (Kd40+unsp.RNA = 825 ± 30 nM) for the blue light-treated samples.

For comparison of the binding parameters of the different RNA aptamers identified via SELEX, the experiment was repeated without the unlabeled tRNAs under both dark and blue light conditions (see Figure 29.b). Within the dark-treated fraction of the samples, only very faint RNA-binding of PAL could be observed for any of the five aptamers within the tested concentration range of 0.02 - 5.12 µM. In contrast, blue light illumination of the remaining sample fractions resulted in the appearance of strong PAL-RNA bands for which quantitative data evaluation yielded Kd values between 0.24 to 1.3 µM (see Figure 29.c and Table 25). Among the five tested aptamers, the SELEX clone 04 displayed the lowest Kd constant (Kd04 = 240 ± 20 nM). For the dark-treated samples, the employed concentration range was not sufficient to record a complete RNA-binding curve. However, for the most affine binding 04 clone, the recorded points were sufficient for a computational estimation of a fitting value for the K_d constant (K_d^{04dark} = 9000 ± 900 nM) which indicates that blue light stimulation induces a more than 30-fold enhancement of the PAL-04-aptamer association. This is in line with the visual assessment of the K_d constant on the basis of the EMSA gel, which is possible because the K_d value corresponds to the effective concentration with a binding behaviour of 50 %. For the blue light-treated samples of the 04 aptamer, this point lies somewhere between 160 nM and 320 nM [PAL] whereas for the dark-treated samples it is almost reached at 5120 nM, supporting the outcome of

the computational estimation that blue light causes a more than 30-fold enhancement of the PAL-04-aptamer association.

Comparison of the predicted minimum fold energy (MFE) secondary structures of the examined SELEX clones [141,142] shows that all clones from motif family 1 (04, 36 and 46) form a conserved hairpin structure with the 3'-end primer region (see Figure 30.a). The hairpin structure contains the conserved consensus motif, which emerges just at the end of the 40 nt random region for the identified aptamers from motif family 1. This hairpin structure is completely congruent for the 04 and the 36 aptamer; in both cases the unpaired terminal loop (T-loop) residues form a hepta-loop rich in A and G residues. This observation is in line with the K_d values determined for the two aptamers, which are in a similar range (K_d^{04} = 240 ± 20 nM and K_d^{36} = 500 ± 250 nM). The hairpin structure of the 46 aptamer differs from the above structural motif within the T-loop region in that it contains 9 unpaired residues instead of 7. However, its association affinity for PAL does not differ significantly from the magnitude determined for the 04 and 36 aptamers (K_d^{46} = 410 ± 60 nM). The two aptamers from motif family 2 (40 and 10) do not share any common secondary structural features. At the sequence level, the region of the identified consensus motif also differs for most of the identified clones of that motif family. The K_d value of the 40 aptamer is in a similar range as the characterized aptamers from motif family 1 (K_d^{40} = 390 ± 20 nM), whereas clone 10 demonstrates a comparably poor association affinity (K_d^{10} = 1300 ± 100 nM). The MFE secondary structure prediction of the 40 aptamer includes three different hairpin-like features, the one closest to the 5' end resembles the hairpin of the 04 and 36 clones in its pronounced stem region, as well as in the number of unpaired residues of the terminal loop (7 nt) and the type of nucleotides (mainly A and G) that build the T-loop region (see Figure 30.b). For the second hairpin, the T-loop region holds a 'GCA' motif that also appears in the 5'-proximal hairpin of the 10 clone. Nevertheless, there are no more pronounced similarities between the two clones of family two, which makes it hard to select any structural feature for further optimization.

Therefore, we decided to focus on the candidates of motif family 1 for further characterization and optimization of the PAL-aptamer association which (i) comprise the SELEX clone 04 with the highest affinity to PAL, and (ii) all feature the same structural hairpin motif that contains the conserved consensus sequence.

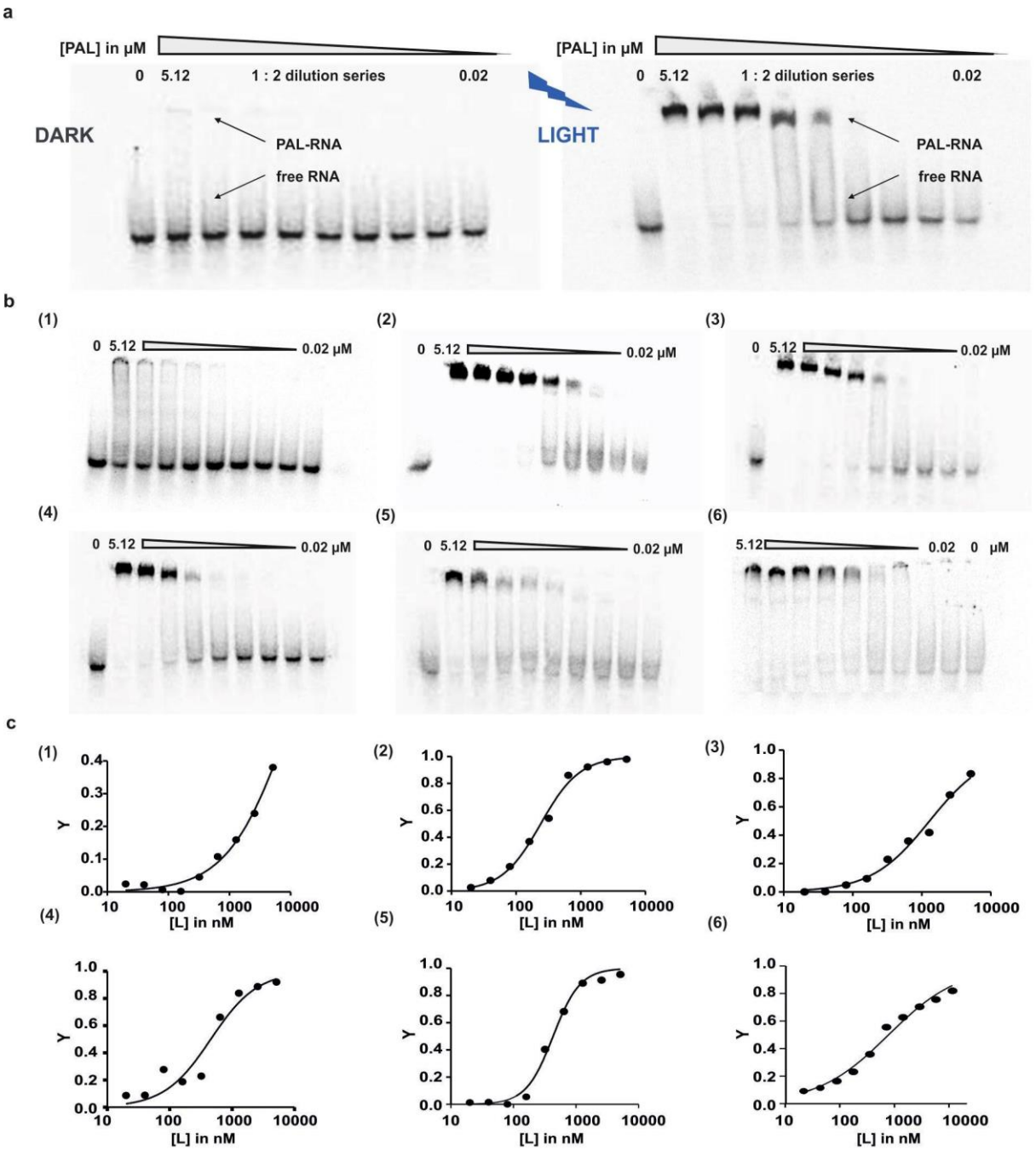


Figure 29: Characterization of SELEX clones via EMSA. The EMSAs were conducted with varying concentrations of PAL (5.12/2.56/1.28/0.64/0.32/0.16/0.08/0.04/0.02 µM) in presence of ≈50 pM of the radiolabeled aptamer. **(a)** Blue light induction significantly increases the affinity of PAL for its target RNA**.** The depicted EMSA demonstrates the interaction of the 40 SELEX clone with PAL under dark (left) and blue light (right) conditions. An excess of unlabeled tRNAs was added to the reaction mixture to validate the selectivity for the aptamer. **(b)** EMSAs of PAL with SELEX clone 04 under dark conditions (1), as well as SELEX clones 04 (2), 10 (3), 36 (4), 40 (5) and 46 (6) under blue light conditions. **(c)** RNA-dissociation curves of PAL of the EMSAs depicted in (a).

Table 25: Computed fitting values of the RNA affinity curves shown in Figure 29.c. K_d - dissociation constant, h - hill coefficient, Y_{max} – computed maximal value for the fraction of bound RNA (Y; Y = $[R{\cdot}L]/[R]_{total}$). Y_{max} was set = 1 for calculation of the association curve of 04 under dark conditions.

Best-fit values	(1) 04 dark	(2) 04 light	(3) 10 light	(4) 36 light	(5) 40 light	(6) 46 light
Y_{max}	= 1.00	1.00 ± 0.03	1.17 ± 0.2	1.0 ± 0.2	0.94 ± 0.2	0.86 ± 0.04
h	0.9 ± 0.1	1.4 ± 0.1	1.0 ± 0.1	1.1 ± 0.4	2.4 ± 0.3	0.86 ± 0.08
K_d [nM]	9000 ± 910	240 ± 20	1300 ± 100	500 ± 250	390 ± 20	410 ± 60

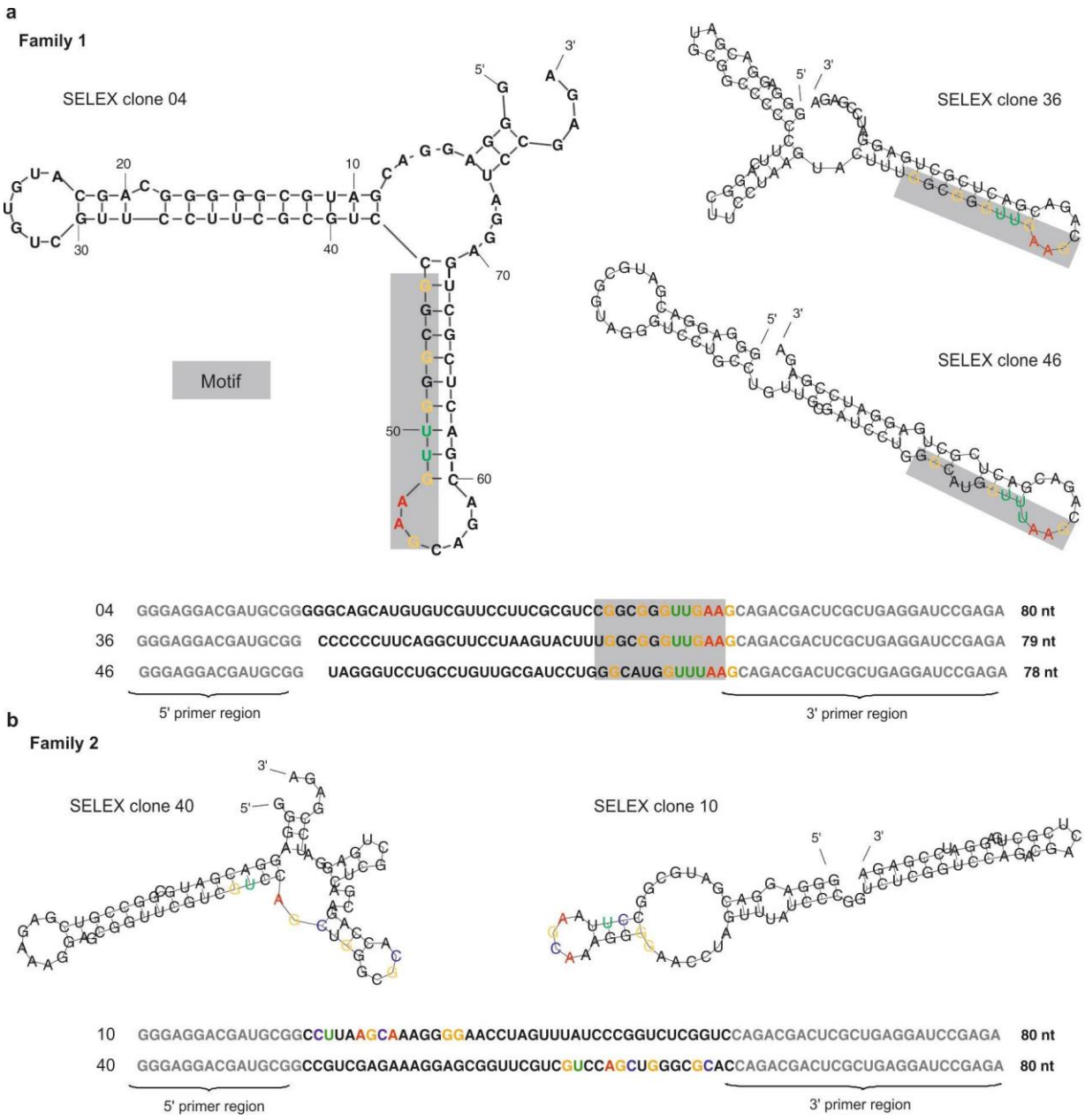


Figure 30: Predicted secondary structure of the RNA aptamers identified via SELEX. **(a)** The SELEX clones from motif family 1 (04, 36 and 46) form a conserved hairpin structure within the 3' primer region that contains the conserved consensus motif. **(b)** The two aptamers from motif family 2 (40 and 10) do not share any common secondary structure features; the region of the identified consensus motif also differs on a sequence level for most of the identified clones of that motif family. The RNAfold [141], as well as the Mfold [142] web servers were used for computation of the hybridization and folding models.

5.2.4 Investigation of aptamer requirements for PAL-RNA association

The identification of specific RNA aptamers that are light-dependently bound by PAL opens up many new possibilities for further characterization and application of the photoreceptor. However, for many experimental investigations (e.g. fluorescence anisotropy for kinetic measurements of the PAL-

aptamer association/ dissociation; co-crystallization for elucidation of structural determinants) or application purposes a shortened, defined binding motif would be of advantage. Therefore, we decided to investigate the minimal required binding region of the shared 3'hairpin motif of SELEX clone 04 and 36 from motif family 1. For that purpose, three different shortened variants were designed and designated according to their length (37 nt → 04.37, etc.; see Figure 31.a).

In parallel, further experiments (see Section 5.2.5) have shown that the PAL receptor assumes a dimeric conformation within its signaling state, suggesting that the combination of two binding motifs may favour the PAL-aptamer association. The so far characterized ANTAR proteins were also shown to bind a tandem hairpin motif [100,110]. We thus designed three further tandem variants of the 04 hairpin motif, connected by a linker element of varying length (7 nt, 10 nt or 14 nt; designated as 04di7, 04di10 and 04di14; see Figure 31.b).

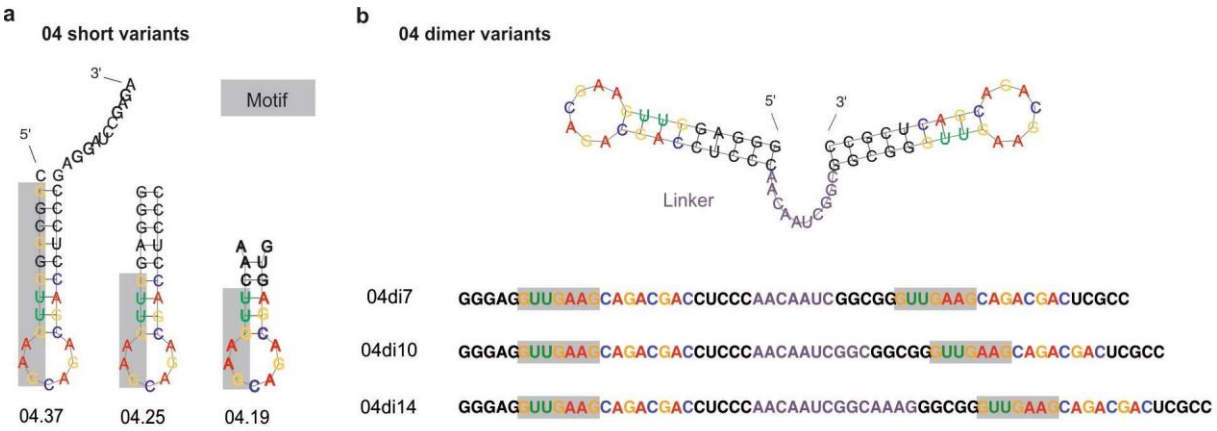


Figure 31: Overview of characterized variants of the 04 aptamer. **(a)** Schematic secondary structure representation of shortened 04 SELEX clone variants, named according to their length (04.37 □ 37 nt; etc.). **(b)** Secondary structure scheme of the dimer variants, for whose design the 04 hairpin motif was connected via linker elements of varying length (7 nt, 10 nt and 14 nt for 04di7, 04di10 and 04di14, respectively).

For conduction of the initial tests, the longest 04.37 variant was ordered and synthesized by IDT (Integrated DNA Technologies, BVBA; Leuven, Belgium). This variant was intended for later use in fluorescence anisotropy measurements as part of the bachelor thesis project of Franziska Wilhelm. Therefore, an overhang of unpaired residues at the 3' end of the hairpin stem was left to assure that an eventually attached fluorophore would not interfere with the PAL-aptamer association. EMSA experiments with radiolabeled sequences of the original 04 aptamer ('04-ori') and the shortened

04.37 variant indicated a comparable affinity of the two aptamers for PAL ($K_d^{04.37}$ = 350 ± 70 nM vs. K_d^{04} = 200 ± 90 nM; see Appendix Figure A1, sectors (1) and (2)). However, experimental problems with the EMSA assay led to rather large deviations between the individual experimental reproductions at that stage, which impeded an exact assessment of the association affinity. Most of the bands occurred smeared and streaky, and even the protein-free control band appeared much less defined than during the initial experiments. Since the labeled RNA probe was freshly transcribed each time before the radiolabeling procedure, storage-related decomposition of the RNA template is unlikely. RNA molecules are highly susceptible to degradation by nucleases, so that the most likely reason for this observation is nucleic acid degradation through ribonucleases. For that reason, all extracts, buffers and reagents were freshly prepared and treated with DEPC, and the amount of RNAse inhibitor was increased after the occurrence of the problems. Although these measures resulted in a clear improvement, the quality of the autoradiographs remained lower than at the beginning of the EMSA experiments. To get to the bottom of the problem, I followed the instructions of various EMSA troubleshooting protocols, e.g. reduction of gel running times, minimization of time between gel loading and electrophoresis, use of more concentrated polyacrylamide gels, use of smaller sample volumes, securing of a constantly low temperature during the electrophoresis run, addition of glycerol as stabilizing solutes in the reaction buffer, or variation of the salt concentration in reaction and electrophoresis buffers to stabilize electrostatic interactions [143]. Despite all these measures, the quality of the EMSAs could not be restored to the status of the first tests.

The initial tests of the different tandem variants did not reveal any improvement in respect to the binding affinity for PAL. The computed fitting values for the K_d constant, obtained by quantification of both free and ligand-bound RNA of the EMSAs, were in a similar range as for the original 04 clone (K_d^{04di10} = 490 ± 60 nM; see Appendix Figure A1 (3)). The examination of the tandem variants was conducted in parallel by me via EMSA and Anna-Maria Pyka in Bonn with a new assay, which employs a fluorescent dye for detection of the RNA (‘Ribogreen assay’). The Ribogreen assay offers some advantages; yet, only longer (> 100 nt) aptamers are suitable for this procedure because the fluoroscent dye used for detection is based on RNA intercalation. For this reason, a comparison between EMSA and Ribogreen assay is only possible for the longer 04di tandem constructs.

The preliminary Ribogreen assay results indicated a considerable improvement of the association affinity of the 04 tandem variants to PAL with K_d values in a range of 10 – 30 nM, and thus about a magnitude of 10 lower than the values obtained for the original 04 aptamer via EMSA analysis. The persistent experimental problems alone seemed not sufficient to explain the discrepancies of the determined K_d values between the EMSA assay and the Ribogreen assay. For this reason, in addition to the ongoing troubleshooting efforts, I started to run buffer optimization experiments, in which the

so far used running buffer (0.5 x TBE), as well as two further buffer variants, were additionally supplied with 2 mM $MgCl_2$. The optimization experiments of the electrophoresis running buffer composition resulted in a considerable improvement of the association affinity of PAL to the original 04 and the 04di10 tandem construct ($K_d^{04\text{-}ori}$ = 130 ± 30 nM; K_d^{04di10} = 30 ± 3 nM; see Figure 32 and Table 26). Protein-nucleic acid interactions are extremely sensitive to the concentration of monovalent and divalent salt ions [143] and were thus presumably destabilized by the differing conditions of the running buffer prior to the adjustment. The K_D values achieved with the help of the optimized electrophoresis running buffer are in good agreement with the dimensions of the K_D values determined by the Ribogreen assay from our cooperation partners in Bonn.

The K_d value of the 04di10 tandem construct is thus within the same range as determined by the Ribogreen assay under the improved conditions. Further tests of the shortened 19 nt-long variant of the 04 construct revealed an affinity of 80 ± 8 nM to PAL under the same conditions ($K_d^{04.19}$ = 80 ± 8 nM; see Figure 32 and Table 26).

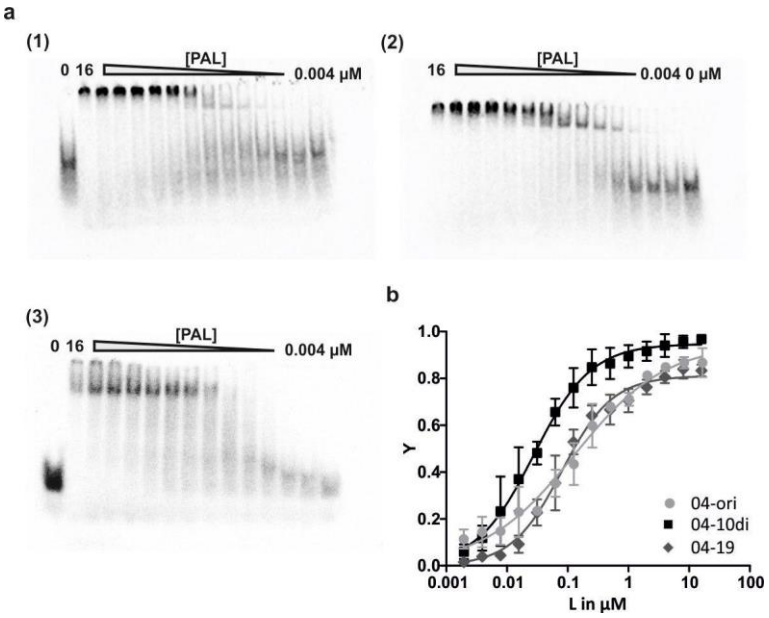


Figure 32: Improved EMSA assay for which the electrophoresis running buffer was supplied with 2 mM MgCl2. The EMSAs were conducted with varying concentrations of PAL (16/8/4/2/1/0.5/0.25/0.125/0.062/0.031/0.016/0.008/0.004 µM) in presence of 50 pM of the radiolabeled aptamer. **(a)** EMSA analysis of PAL plus the original 04 clone ('04-ori') (1), the 04-10di (2) tandem construct, or the shortened 04 19mer ('04-19') (3) under constant blue light illumination **(b)** RNA-dissociation curves obtained by quantification of both free and ligand-bound RNA of the EMSAs shown in (a) including the data for an additional reproduction of the EMSA analysis (n = 2).

Table 26: Computed fitting values of the RNA affinity curves shown in Figure 32.b. K_d – dissociation constant, h - hill coefficient, Y_{max} – computed maximal value for the fraction of bound RNA (Y; $Y = [R{\cdot}L]/[R]_{total}$).

Best-fit values	(1) 04-ori	(2) 04di10	(3) 04-19
Y_{max}	0.94 ± 0.04	0.95 ± 0.02	0.81 ± 0.02
h	0.59 ± 0.06	0.92 ± 0.09	1.04 ± 0.09
K_d [nM]	130 ± 30	30 ± 3	80 ± 8

5.2.5 RNA co-immunoprecipitation

In Section 5.2.3 we were able to identify first specific RNA target sequences for PAL. However, these sequences were derived from a pool of randomly generated RNA sequences that do not necessarily match the natural binding sequence and structural features. The RNA co-immunoprecipitation (RIP) procedure presents an alternative method for the mapping of natural protein-RNA binding sites via the joint immunoprecipitation (IP) of a protein of interest together with associated RNA molecules and subsequent determination of the RNA sequences, e.g. via RNA-sequencing.

A successful RIP experiment usually requires the availability of a suitable antibody (Ab); therefore, a polyclonal anti-rabbit high-affinity Ab against PAL (PAL-Ab) was produced and tested for suitability within IP use and specificity prior to the start of the RIP experiments. The test of specificity should assure that the PAL-Ab does not display any cross-reactivity to other *N. multipartita* (*Nm*) proteins or cell components, from which the RNA-PAL complex will be isolated. Therefore, the PAL-Ab was tested on different dilutions of soluble *N. multipartita* cell lysate ('*Nm* lysate') within Western blot experiments. The cultivation of *N. multipartita* was achieved as described in Section 4.12. Cell lysis of the required small volumes was done through three freeze-thawing cycles with liquid nitrogen. Since it is not known under what conditions and to what extent PAL is expressed in its original host, the lack of a signal would give only little indication of the quality of our antibody. Therefore, in a second sample, around 1.3 µg of purified recombinant PAL were added to 100 µl of the *Nm* cell lysate ('*Nm* + PAL'). In addition, a third sample of pure recombinant PAL ('PAL purif') was loaded onto the gel as positive control. For all three samples, 10 µl of three different dilutions were then separated on a 12 % SDS gel and analyzed via Western blot (see Figure 33.a, (1)). Western blot detection revealed no detectable band within the mere *Nm* lysate fractions. Within the second sample (*Nm* lysate + PAL), a sharp single band can be perceived at the level of 40 kDa within the undiluted and 1 : 10 diluted fractions, wich is at the same level as the PAL protein from the positive control (PAL purif).

To evaluate the suitability of the PAL-Ab for IP use, the antibody was tested in a purification trial. For this purpose, in a further experiment, 6.5 µg of purified recombinant PAL were added to 2 ml of a *N. multipartita* fraction, which was previously lysed and centrifuged to remove the insoluble cell fraction ('IP *Nm*+PAL'). In addition, the IP separation was conducted with mere *Nm* lysate ('IP *Nm*') and with mere buffer ('IP buffer') to assign eventual bands stemming from the Ab. For the IP separation, the PAL-Ab was coupled to Dynabeads magnetic beads (Dynabeads Protein A, Thermo Scientific) and purified as specified in the manufacturer's instructions.The eluted fractions were separated on a 12 % SDS gel and analyzed via Western blot (see Figure 33.a, (2)). Western blot detection discloses three strong bands at the level of > 100 kDa, 55 – 70 kDa, and 40 – 55 kDa for all of the IP-treated samples, including the 'IP buffer' control, so that these bands can be assigned to Ab

fragments. For the 'IP *Nm*+PAL' fraction, an additional band at the 40 kDA level, distinctive for PAL, could be detected, proving the suitability of the PAL-Ab in IP applications. Yet, even after concentration via IP-treatment, no signal could be perceived for the native PAL protein within the *Nm* cell lysate. Yet, the detection sensitivity of co-immunoprecipitated RNA might be higher in RNA sequencing experiments than the detection sensitivity reached via radiolabeling methods, so that we decided to move on with the RIP for a first comparative attempt with three different approaches: In a first approach, the RIP was conducted under native conditions, using only the *N. multipartita* supernatant for co-incubation with the PAL-Ab ('Native'). In a second approach, the RIP was carried out with an additional portion (150 pmol) of recombinant PAL, which was added to the *N. multipartita* supernatant prior to the IP. In a third approach, the RIP experiment was conducted with previously purified RNA from *N. multipartita* cells ('Pure RNA'), for which the preparation of the RNA was carried out separately (see Section 4.13.2).

As negative control experiment, the RIP was conducted using Dynabeads without the coupled PAL-Ab, applying the 'native' fraction. As a positive control, 150 pmol of the *in vitro* transcribed 04 RNA were added to the 'native' fraction prior to the IP. The *N. multipartita* cultures were grown for ≈ 48 hours at 28 °C at 225 rpm under constant blue light illumination (50 µW/ cm² at 450 nm) up to an OD_{600} of 1.2 – 1.4 before harvest. For stabilization of cellular RNA transcripts, RNAprotect Bacteria reagent was used (see Section 4.13.2 for details of the experimental conduction). For evaluation of the RIP experiments, small fractions of the pulled-down RNA from each of the above approaches and controls were subjected to dephosphorylation via FastAP and subsequent 5'-end radiolabeling as described in Section 4.7.1.

Evaluation of the autoradiography (see Figure 33.b) reveals strong signals for the 'Pure RNA' (Figure 33.b, (3)) and the 'Native+PAL' (see Figure 33.b, (4)) approaches. For the mere 'Native' approach (Figure 33.b, (5)), only a very weak signal could be detected. Unfavorably, the positive control lane is lacking any detectable signal. The RNA derived from the pulldown experiments of all three approaches, as well as the negative control, was nevertheless send to our cooperation partners Gerhard Lehmann and Gunter Meister from the University of Regensburg for sequencing. However, the evaluation (which was still in progress at the end of my time in Bayreuth, so that it is not included in this work) showed no significant differences between the data sets.

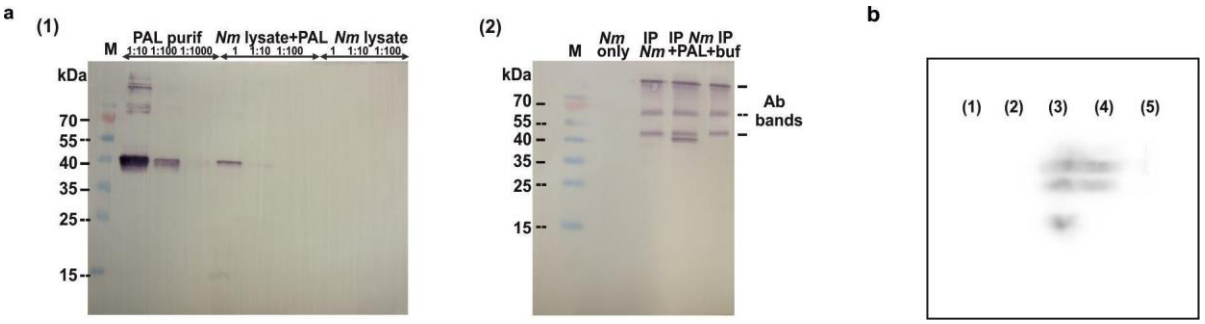


Figure 33: Preliminary tests and evaluation of the RIP experiments. **(a)** Testing of the generated PAL-Ab. For all samples, volumes of 10 µl were loaded and separated on a 12 % SDS gel and analyzed via Western blot. (1) Test of specificity. 'M': Protein ladder; 'PAL purif': positive control, comprising different dilutions (1:10, 1:100, 1:1000) of the purified recombinant PAL protein (≈ 33 µM); '*Nm* lysate+PAL': contains the lysate of *N. multipartita* (*Nm*) cells plus 1.3 µg of recombinant PAL in different dilution (1, 1:10, 1:100); '*Nm* lysate': contains different dilutions (1, 1:10, 1:100) of the mere lysate of *Nm* cells. (2) Test of IP suitability. 'M': Protein ladder; '*Nm* only': *Nm* lysate without IP treatment; 'IP *Nm*': IP with *Nm* cell lysate; '*Nm* +PAL': IP with *Nm* cell lysate plus 6.5 µg of recombinant PAL; 'IP + buf': IP with mere buffer. **(b)** Evaluation of RIP pull-down products from different approaches. (1) Positive control: RIP with *Nm* supernatant containing 150 pmol of recombinant PAL plus 150 pmol of the *in vitro* transcribed 04 RNA. (2) Negative control: RIP using Dynabeads without the coupled PAL-Ab with *Nm* supernatant. (3) 'Pure RNA': RIP with previously purified *Nm* RNA plus 150 pmol of purified recombinant PAL. (4) 'Native+PAL': RIP with *Nm* supernatant containing 150 pmol of recombinant PAL. (5) 'Native': RIP with *Nm* supernatant without any further components.

5.2.6 Structural characterization

Identification of suitable crystallization conditions

As a first step of the crystallization experiments, a sitting drop MRC2 96-well plate was set up with the JCSG+ commercial sparse matrix screen (Qiagen) at concentrations of 27 mg/ ml and 13.5 mg/ ml PAL in buffer D (see Table 9 in Section 4.4.3) using the Phoenix liquid handling system. The examination of this first plate revealed heavy precipitations at these concentrations, so that the following plates were set up with lower protein concentrations. To then determine the optimal ambient parameters , four JCSG+ sparse matrix screen plates with protein concentrations of 9 mg/ ml and 4.5 mg/ ml were incubated at 4° C or 20° C under either dark or blue light (50 µW/ cm^2) conditions. A fifth, additional plate was monitored with the help of a Rockimager 1000 (Formulatrix) at room temperature, with the resulting disadvantage of rather undefined lighting conditions. In case of the illuminated plates and of the Rockimager-monitored plate, the undesirable phenomena of high protein loss or phase separation occurred within short time, in the majority of the wells. Under dark conditions, however, a significantly higher number of wells preserved a 'promising appearance' (i.e., little or no protein aggregation, uniform grain size, no phase separation) after several days at both 4° C and 20° C. However, after 7 days, undesirable strong protein aggregations occurred in more

than 50 % of the wells of the 20° C plate, so that we decided to continue the screening with the ambient parameters set to 'dark' and '4 °C'. For these tests, further commercial sparse matrix screens (JCSG Core Suite I, II, III & IV and AmSO4 Suite; Qiagen) were prepared with 0.2 µL drops at the same protein concentrations as before (9 mg/ ml and 4.5 mg/ ml) mixed 1 : 1 with the reservoir solution. These new screening plates were monitored together with the initial JCSG+ plates on day 1, 2, 3, 6 and then once a week for four months. Potential protein crystallization hits were usually first manually replicated and then tested with the help of JBS True Blue, a dye helping to discriminate protein crystals from salt crystals or amorphous objects. Most of these potential crystals resulted to be salt or non-protein material, except for one single condition (C10 of the JCSG+/ 4°C/ dark plate: 0.1 M bicine, pH 9.0; 10 % PEG 20000; 2 % dioxane 1 : 1 with 4.5 mg/ ml protein solution), within which a slim needle appeared after almost two months. Optimization trials were performed under the same temperature and lighting conditions in 24-well plates with the 'hanging drop' method. For optimization, 24-well plates were set up manually with varying pH and PEG concentrations around the initial crystallization condition and three different protein to reservoir ratios (2 : 1, 1 : 1, 1 : 2) at drop volumes of 2 µl. The thereby determined best condition (0.1 M bicine, pH 9.2; 15 % PEG 20000; 2 % dioxane 1 : 1 with 4.5 mg/ ml protein solution) was further optimized using the 'Additive Screen' (Hampton Research) in a 96-well MRC2-plate. The second drop of each well was hereby supplied with 25 mM NTPs, which resulted in much larger protein needles in most cases (but still too small for diffraction tests). In further attempts of optimization, the six most promising additive conditions were tested at varying concentrations or combined with NTP in a 24-well plate format. In addition, the effect of imidazole acetate was assessed (motivated by rather intuitive reasons, as imidazole usually has a strong impact on the photocycle of LOV proteins and enhances the dark recovery via a base-catalyzed mechanism [144]), which resulted in the strongest improvements observed so far at 0.8 M (see Figure 34). Under this condition (0.1 M bicine, pH 9.2; 15 % PEG 20000; 2 % dioxane; 0.8 M imidazole acetate, at a ratio of 1 : 2 with 4.5 mg/ ml protein solution at drop volumes of 2 µl), first crystals could be obtained that exhibited diffraction patterns up to 2.75 Å on a synchrotron source (see Figure 35). A first dataset was collected at the BESSY II synchrotron (HZ, Berlin) on beamline 14.1 at a wavelength of 0.9181Å. Processing the dataset using with XDSAPP [122] indicated that the crystals belong to space group $P2_12_12_1$; a cell content analysis, considering the molecular weight of PAL (39304.6 Da) and the unit cell dimensions, suggested the presence of 4 molecules in the asymmetric unit. The statistics of the high-resolution diffraction dataset are shown in Table 27.

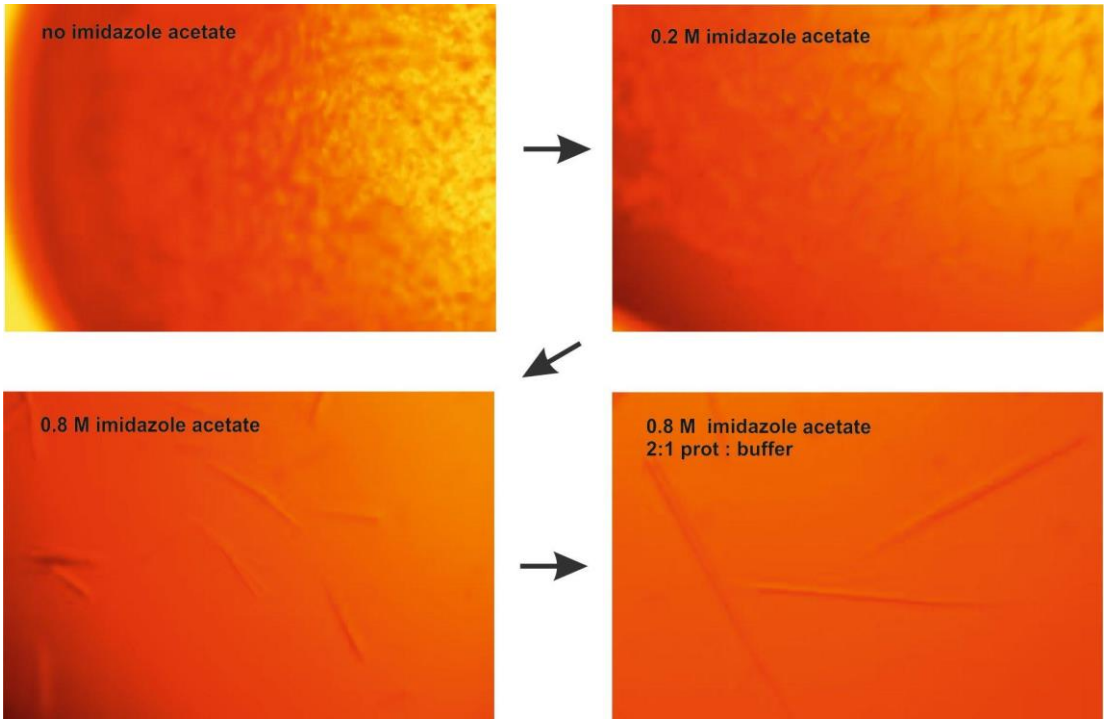


Figure 34: Optimization of the initial crystallization conditions (0.1 M bicine, pH 9.2; 15 % PEG 20000; 2 % dioxane; 1 : 1 with 4.5 mg/ ml PAL) with imidazole acetate as additive. The addition of 0.8 M imidazole acetate led to a striking improvement of the original clusters of tiny needle-shaped crystals. Further variation of the protein to reversoir solution ratio (prot : buffer) to 2 : 1, led to the production of crystals that allowed the collection of first datasets at Bessy II synchrotron (HZ, Berlin).

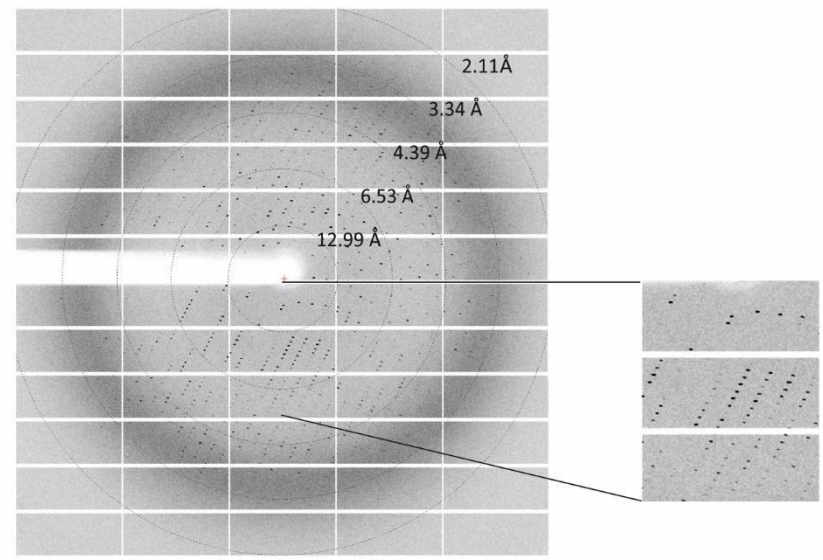


Figure 35: Diffraction pattern for the native PAL crystal in overall and zoomed view. The crystal pattern exhibited diffraction spots up to 2.75 Å and was indexed in space group $P2_12_12_1$.

Yet, the phase problem could not be solved by molecular replacement, due to the poor coverage of the protein sequence by molecular replacement models (only the predicted LOV- (resi246 to 352) and ANTAR domains (resi 127 to 180) were available). The moderate resolution of the dataset is further reducing the probability of a successful molecular replacement. Neither combined models, nor automatic, brute-force, molecular replacement programs, such as Phenix MRage, or MrBump and BALBES of the CCP4 suite, succeeded in solving the structure by molecular replacement, so that experimental phasing was required to elucidate the structure of PAL.

Experimental phasing and model building

In order to apply experimental phasing to solve the PAL structure, selenomethionine (SeMet)-incorporated PAL was produced in *E. coli* following the protocol described in Section 4.11.1. SeMet incorporation into proteins through recombinant production in *E. coli* is a well-established strategy for phasing [145]. For this purpose, a minimal medium without methionine, but SeMet instead, is used for cultivation. The addition of high amounts of isoleucine, lysine and threonine inhibits the pathway for methionine biosynthesis, thereby forcing *E. coli* to use SeMet for protein expression [146]. Purification was performed, as described for native PAL, *via* IMAC (see Section 4.4.3) and resulted in sufficient protein quantities for crystallization. With the purified SeMet-PAL sample, I set up grid screens with the same conditions as for native PAL crystallization (0.1 M bicine, pH 9.2; 15 % PEG 20000; 2 % dioxane; 0.8 M imidazole 1 : 1 with 4.5 mg/ ml protein solution). After one night, needle-shaped crystals with a similar appearance as for native PAL appeared, of which several were harvested after 5 days. A diffraction data of one of the SeMet-PAL crystals was collected at BESSY II synchrotron (Helmholtz-Zentrum, Berlin) on beamline 14.1 at a wavelength of 0.979656 Å. The best of the tested SeMet crystals diffracted up to 3.68 Å resolution and, using the anomalous signal from incorporated SeMet residues by single wavelength anomalous dispersion (SAD) and the SHELXC/D/E pipeline, produced a satisfying set of initial phases, which permitted to solve the structure[145]. The statistics of the merged data sets are shown in Table 27.

The experimental phasing of the structure, as well as large parts of the model building were done under the supervision and with huge support of Dr. Sébastien Moniot. Automatic model building was done with Phenix Autobuild [124] and Buccaneer [145], for manual reconstruction of the missing segments, Coot [147,148] was used. The partial models originating from both software were subsequently combined and completed through iterative cycles of manual corrections and refinement using phenix.refine [126] to R_{work}/ R_{free} values of 0.2316 / 0.2529, respectively. The refinement statistics of the data that led to the structural model used in this work are shown in Table 27. This model comprises the amino acids 10 to 365; the N-terminus, as well as residues 218 - 221 of

monomer A and residues 216 - 223 of monomer B were not modeled as the electron density is not defined (chain C and D also contain some undefined sections, but only chain A and B were used within this work), probably due to a high intrinsic flexibility in this region. The overall structural model features four molecules of PAL within each asymmetric unit that assemble into two dimers (see Figure 36). The Matthews coefficient was calculated to 3.52, which corresponds to a solvent content of 65.1 %.

Table 27: Crystallographic data collection and refinement statistics. Values for the highest resolution shell are shown in parentheses. SeMet-PAL: data for selenomethionine-substituted PAL crystal, PAL: data for. For Native data, anomalous Friedel pairs were considered independently for the SeMet-PAL dataset. RMS: Root-mean-square deviation.

Data collection	**SeMet-PAL**	**PAL**
Synchrotron beamline	BESSY II, MX 14.1	BESSY II, MX 14.1
Wavelength [Å]	0.979656	0.9181
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell dimensions [Å /°]	a = 67.01 b = 150.41 c= 220.54	a = 67.00 b = 150.36 c = 219.83
	α= 90.00 β = 90.00 γ = 90.000	α = 90.00 β = 90.00 γ = 90.000
Resolution [Å]	45.64 - 3.68 (3.81 - 3.68)	46.97 - 2.75 (2.85 - 2.75)
R-meas	0.2497 (1.011)	0.1524 (1.155)
Mean I/σ	6.73 (1.75)	9.83 (1.28)
Wilson B-factor [Å²]	92.57	51.80
Completeness [%]	99.56 (97.28)	98.67 (99.36)
Unique reflections	46562 (4504)	57953 (5739)
Redundancy	5.8 (5.7)	4.5 (4.3)
$CC_{1/2}$	0.989 (0.681)	0.995 (0.597)
Refinement		
R_{work}/ R_{free}		0.2316 / 0.2529
Number of atoms		10961
Protein		10706
Ligand/ ion		136
Water		119
Protein residues		1418
Average B-factor [Å²]		66.25
Protein		66.60
Ligand/ ion		53.34
Water		49.46
RMS bond lengths [Å]		0.008
RMS bond angles [°]		1.10
Ramachandran favored [%]		97.50
Ramachandran outliers [%]		0.29

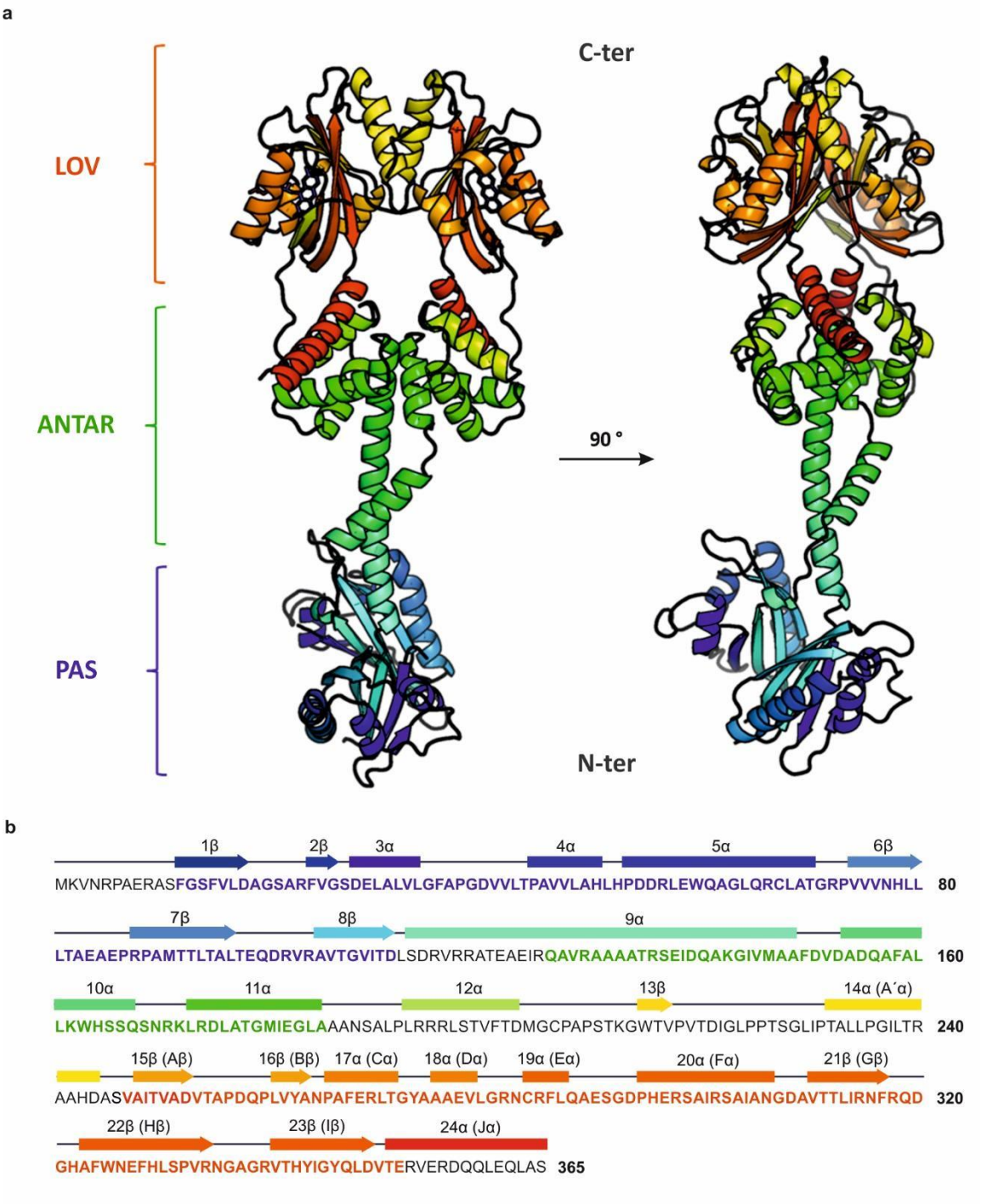


Figure 36: Crystal structure of the dark-adapted PAL photoreceptor. **(a)** The crystal structure of dark-adapted PAL. FMN cofactors within the LOV domains are shown in stick representation; domain coloring in chainbow colors with the N-terminal PAS domain (resi 10 - 73) shown in blue tones, the ANTAR domain (resi 127 - 202) shown in green, and the LOV domain in tones ranging from yellow (depicting the N-terminal A'α extension; resi 231 - 244) over orange (LOV-core region; resi 246 - 352) to red (depicting the C-terminal Jα helical extension; resi 351 – 365). Notably, the ANTAR domain comprises an extra helix (12α) in addition to the characteristic three-helical bundle of the ANTAR core region (aa 127 - 180). **(b)** Overview of secondary structure elements of the PAL monomer. Residues of the domain core regions are represented in bold optic (PAS → blue; ANTAR → green; LOV → orange). The individual structural elements (α-helices and β-strands) were consecutively numbered; for the LOV domain, the classical annotation for structural motifs is shown in parentheses.

PAL crystal structure

The interface of the dimer is formed by an α-helical spine that extends from the C-terminus of the PAS core modules as coiled-coil α-helical linker between PAS (resi 10 – 111) and ANTAR (resi 127 – 180) domains; a second coiled-coil segment is formed by the N-terminal α-helical extensions of the LOV core domains (resi 246 – 350), denoted A'α (resi 231 – 245). Interestingly, the two LOV domains, connected with the ANTAR domains via a long (28 aa) flexible linker, are rotated 180° along their horizontal axis, permitting the C-terminal α- helical extensions, denoted Jα (resi 351 – 365), to form an interface with the adjacent helical surface of the ANTAR domain. For the LOV-ANTAR connecting flexible linker region, the electron density is not defined for all residues due to the high flexibility in this region. Within the PAS-ANTAR connecting coiled-coil region (resi 113 – 149) of chain B, a pronounced kink leads to disruption of the α- helical fold within resi 116 – 121 and resi 136 – 139, probably caused by intermolecular packing within the PAL crystal lattice. As a consequence, the PAS domains fall out of the symmetry along the axis of the helical spine, otherwise maintained in the rest of the molecule.

Both LOV domains comprise one FMN chromophore as a ligand; the absence of a covalent bond between the FMN and the active site cysteine (resi 284) confirms that the PAL photoreceptor assumes its fully dark-adapted state (see Figure 37). The FMN chromophore is stabilized through a network of hydrogen bonds by the residues N283, R285, Q288, R301, N316, and N326. The active site cysteine is embedded within the conserved GXNCRFLQ motif (resi 281 – 288; here X corresponds to R282), its thiol group is oriented towards the dimethyl-benzene ring of the FMN chromophore, and away from the C4a position. Residue T250, whose substitution significantly affects the dark recovery kinetics in other LOV photoreceptors, forms a hydrogen-bond with Q347. Notably, the PAL LOV domains contain a sixth β sheet (13β, resi 213 – 215; see Figure 36.b) in addition to the five-stranded antiparallel β sheet that define the characteristic PAS-fold (Aβ, Bβ, Gβ, Hβ, Iβ in common LOV annotation; here 15β, 16β, 21β, 22β, 23β as part of the full-length structure; see Figure 36.b) [149]. The additional 13β sheet arises within the flexible linker region that connects ANTAR and LOV domain and aligns antiparallel with Aβ (15β; resi 248 - 253) via polar interaction through T214 and T312. The N-terminal A'α helices (resi 231 – 245) assemble into a coiled coil at the LOV-domain interface and are stabilized by interactions among the aliphatic residues L234 and L238, as well as through intermolecular hydrophobic contacts with the β sheets of the adjacent LOV domains via I237 (see Figure 38.b). Besides these hydrophobic, nondirectional interactions, A'α further builds up polar intermolecular contacts with the β sheets of the opposite LOV monomer via D244 and Y346/ S331 or A242 and N265 at its C-terminus (see Figure 38.a/b). Particularly, the β-sheet cores of the two LOV modules make no direct contact with each other. The C-terminal Jα helical extensions (resi 351 – 365) extend from the LOV core module to the C-terminus of the ANTAR domain; the junction

between the LOV core and the C-terminal Jα extension is provided by the conserved residues D349, V350, and T351 (DIT motif; [87]) that form a pair of hydrogen-bonds with each other (see Figure 38.a). Further polar contacts of the Jα helix arise from R353 with E326 of the β sheets of the LOV core; or Q358 with the backbone of S209, which is part of the flexible linker region that connects LOV and ANTAR domain; as well as Q359 and H164 of the ANTAR module. In addition to these polar interactions, Jα is stabilized by interactions among the aliphatic residues V354 and W325 of Hβ (22β; resi 323 – 334), or L363 and L160 of helix 10α (resi 154 – 164) of the ANTAR domain (see Figure 38.c). Interestingly, the C-terminal His-tag ('LEHHHHHH'; resi 366- 374) elongates the Jα helix in a continuous α-helical fold.

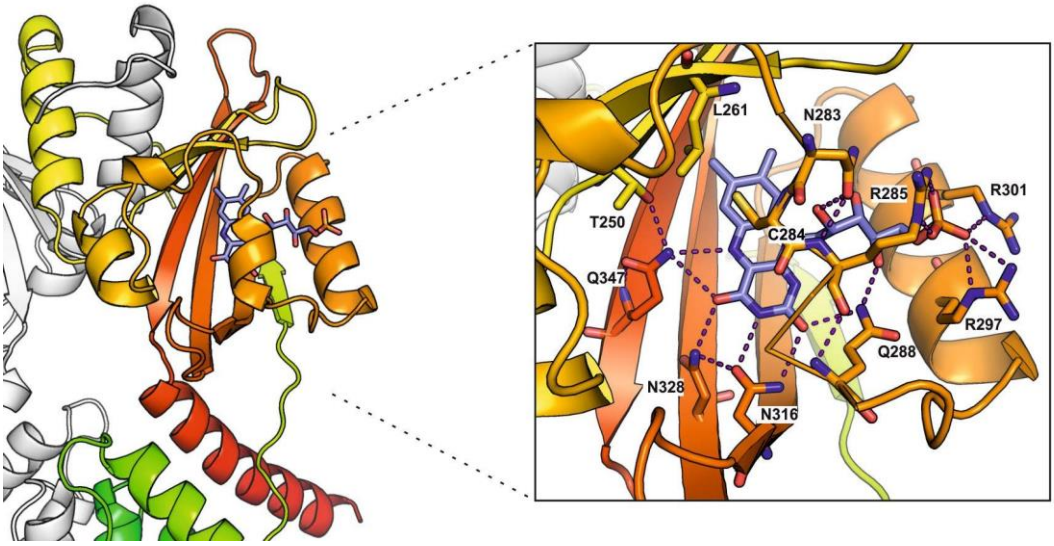


Figure 37: FMN binding pocket of the PAL LOV-domain. The FMN chromophore is shown in stick representation highlighted in blue; residues involved in commented interactions are also shown in stick representation. The detail view shows polar interactions and steric contacts of the FMN chromophore with the surrounding active site residues. For the Eα (19α) helix, the cartoon was replaced by the ribbon representation for a better view of the chromophore interactions. The residues N283, R285, Q288, R301, N316, and N328 are involved in polar interactions with the FMN chromophore. Residue T250 whose substitution significantly influences the dark recovery kinetics in other LOV proteins, forms a hydrogen-bond with Q347. L261 is located in direct proximity to C284, placing its two methyl groups only 4.9/ 5.0 Å away from Cβ of C284.

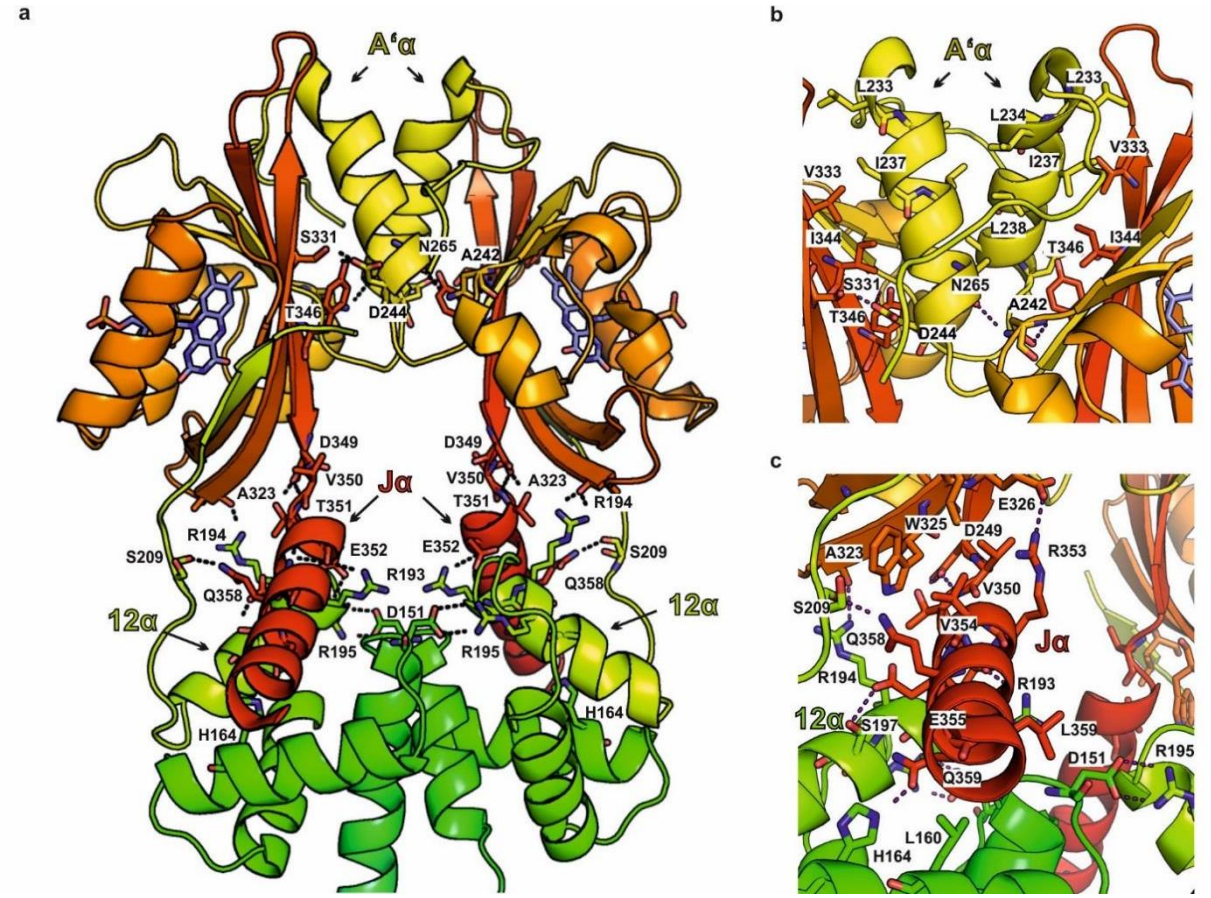


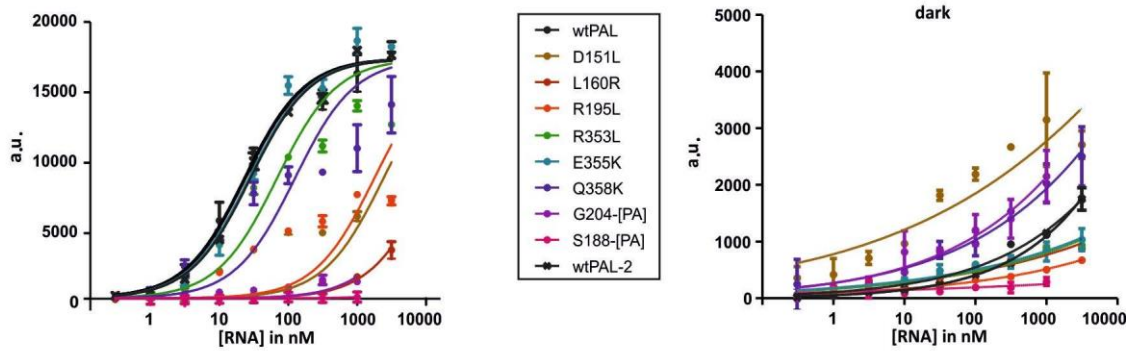
Figure 38: Detail view of the LOV-ANTAR interface. Residues involved in commented interactions are shown in stick representation, the FMN chromophore is shown in stick optic highlighted in blue. **(a)** Overview of polar contacts of the A'α and Jα helical extensions with their surrounding protein environment. A'α is stabilized by polar intermolecular contacts with the β-sheets of the opposite LOV monomer via D244 and Y346/ S331 or A242 and N265 at its C-terminus. The ANTAR domain comprises an extra helix (12α), which is engaged in several polar contacts with Jα (e.g. R193 with E352 and S197 with E355) and the LOV core region (R194 with A323). Further polar contacts of Jα arise from R353 and E326 of the LOV β sheets, Q359 and H164 of the ANTAR domain, or Q358 with the backbone of S209, which is part of the flexible linker region that connects LOV and ANTAR domain. Within helix 12α of the ANTAR domain, R195 forms an intermolecular salt bridge with D151 of helix 9α of the opposite monomer. **(b)** Overall interactions of the A'α helical extension. In addition to the polar contacts at its C-terminus commented in (a), A'α is stabilized through interactions among the aliphatic residues L234 and L238, as well as through intermolecular hydrophobic contacts with the β sheets of the adjacent LOV domains via I237. **(c)** Overall interactions of the Jα helical extension. In addition to the polar interactions commented in (a), Jα is stabilized by contacts between the hydrophobic residues V354 and W325 of Hβ, or L363 and L160 of helix 10 of the ANTAR domain. Residues belonging to the C-ter His-tag are not shown.

Influence of residue mutations

In order to identify some first residues involved in signaling, we examined the effect of mutations within residues of the LOV-ANTAR interface. The Ribogreen assay presents a fast and robust alternative to the so far employed EMSA to assess the RNA binding activity of potential mutants. Here, the target RNA is added in varying concentrations, while the PAL protein is immobilized to streptavidin-coated microplates. Unbound RNA fractions are eliminated by washing, and the final bound fractions are detected with the help of the fluorescent Ribogreen RNA reagent.

We decided to start with a number of ten mutants, all aiming at a variation of the RNA-binding activity: D151L-, L160R-, R195L-, T351V-, R353L-, E355K-, and Q358K-PAL (see Figure 38.a/c), as well as one further mutant lacking the Jα helical extension (dJα-PAL; resi 1 - 352). In addition, two variants lacking the blue-light-sensitive LOV domain were investigated: S188-[PA] (resi 1 – 188), ending at the terminus of α11, and G204-[PA] (resi 1 – 204), ending at the C-terminus of the additional helix 12α. The PAL-[PA] variant designed prior to the elucidation of the crystal structure ended at resi 198 and thus in the middle of helix α11, which might be the reason for the earlier problems with protein expression and stability. Like the EMSA experiments, the Ribogreen assay requires the expression and purification of all PAL variants to be tested. The yield of expression differed strongly among the different mutant variants; for the dJα- and T351V-PAL constructs, the expression and solubility were to poor to include them into the tests. For the remaining PAL mutants, the Ribogreen assay was conducted with two replica measurements per variant. Fitting of the obtained data points led to strong variations of B_{max}, which corresponds to the total number of available receptors (here the 04-10di RNA aptamer) that, in theory, should actually remain constant uned the given experimental conditions. In contrast to Y, which is normalized by definition and thus adopts a value between 0 and 1, B_{max} occurs in arbitrary units. Remember that Y is given by the ratio $[RL]/[R_t]$. $[R_t]$ depicts the total amount of receptors in the sample and divides into two populations, i.e. the fraction [R] of free receptor, and [RL], the fraction of bound receptor within the receptor/ligand complex ($[R_t] = [R] + [RL]$) [150]. Hence, B_{max} should remain constant under the given experimental conditions, as it corresponds to the total number of available receptors (the 04-10di RNA template). For wild-type PAL, which was placed as a control on each plate, the amplitude of B_{max} stayed in a similar range for the independent measurements (see Figure 39.a). The strong variations in the mutants could be an indication of increased degradation of the RNA substrate. To compare the mutant data with wtPAL, B_{max} was set to 17500 (which is in between the two values determined for wtPAL) as a fitting constraint for the evaluation of the mutant data obtained by means of the Ribogreen assay (also, for the light curves, h was set to 1; see Figure 39.b).

a



b

light	B_{max} [a.u.]	SE	h	SE	K_D [µM]	SE	dark	B_{max} [a.u.]	h	SE	K_D [µM]	SE
wtPAL	17150	550	0,92	0,08	0,024	0,002	wtPAL	=17500	0,36	0,03	1,334	0,755
D151L	=17500	-	= 1	-	2,410	0,692	D151L	=17500	0,20	0,03	3,189	4,307
L160R	=17500	-	= 1	-	11,655	0,970	L160R	=17500	0,22	0,02	1032,141	1216,745
R195L	=17500	-	= 1	-	1,773	0,527	R195L	=17500	0,25	0,01	1257,162	917,404
R353L	=17500	-	= 1	-	0,069	0,017	R353L	=17500	0,24	0,02	327,320	371,225
E355K	=17500	-	= 1	-	0,027	0,003	E355K	=17500	0,23	0,03	484,865	742,509
Q358K	=17500	-	= 1	-	0,125	0,041	Q358K	=17500	0,30	0,03	0,921	0,630
G204-[PA]	=17500	-	= 1	-	9,910	2,439	G204-[PA]	=17500	0,32	0,04	0,440	0,426
S188-[PA]	=17500	-	= 1	-	187,133	350,099	S188-[PA]	=17500	0,11	0,01	1,3E+13	2,5E+14
wtPAL_2	17700	450	1,00	0,07	0,028	0,002	wtPAL_2	=17500	0,45	0,02	0,379	0,089

Figure 39: Test of mutant variants of PAL by means of the Ribogreen assay. **(a)** RNA dissociation curves of the different PAL mutants. For performance of the assay, the 04-10di aptamer was added in varying concentrations ([RNA])to the samples, while the PAL protein was immobilized to the plates. The depicted curves were obtained by evaluation of two replica measurements. **(b)** Fitting values for the fitted curves shown in (a). Comparison of the determined K_d values for for the light-adapted state shows that most of the introduced mutations led to a removal of the light state RNA binding activity. Only the E355K-, R353L- and Q358K-mutant behave comparably to wild-type PAL (wtPAL), whereas all other mutant variants display constitutive inactivity, i.e. display light state K_d values larger than the ones obtained for wtPAL in its dark state. B: total numbers of available receptors; here the 04-10di RNA aptamer. a.u. – arbitrary units; K_d – dissociation constant; h - hill coefficient; B_{max} – total number of available receptors; SE – standard error.

The thereby obtained RNA dissociation curves of the light-activated state reveal K_d values in a similar range as within the EMSA experiments for wtPAL (K_d^{wtPAL} = 24 ± 2 nM or $K_d^{wtPAL-2}$= 28 ± 2 nM; see Figure 39.b). Precise comparison of the dark state values of the obtained K_d is more critical, as the incomplete dissociation curves result in inaccurate fitting, causing high standard errors (SE) among different experiments (e.g., K_d^{wtPAL} = 1334 ± 755 nM or $K_d^{wtPAL-2}$= 379 ± 89 µM under dark conditions, resulting in mean and SD of 856.5 nM and 422 nM, respectively). Therefore, the comparison of the dark state activity was restricted to orders of magnitude (i.e., factors of ten). Most of the introduced mutations led to a removal of the light state RNA binding activity. For D151L-, L160R-, and R195L-PAL, the RNA binding affinity is smaller than for dark state wtPAL under both dark and blue light conditions. Unexpectedly, even the two [PA]-variants turned out to be constitutively inactive, i.e. demonstrate K_d values larger than the dark-adapted wtPAL. The E355K-PAL variant displays a

comparable binding affinity for the 0410di substrate in its light state (K_d^{E355K}= 27± 3 nM) as wtPAL, and demonstrates an even lower binding affinity in the dark than wtPAL. R353L and Q358K behave comparably to wtPAL, but display slightly reduced binding affinities in their light-activated states (K_d^{R353L} = 69 ± 17 nM; K_d^{Q358K} = 125 ± 41 nM) compared to wtPAL. In summary, the E355K-, R353L- and Q358K-mutants behave comparably to wtPAL, whereas all other mutant variants display constitutive inactivity. The large impact of the D151L-, L160R-, and R195L point mutations hence indicates an important role of these residues within the transmission of the light signal.

Structural changes upon light-activation

The crystal structure of the PAL photoreceptor in its dark-adapted state allowed the identification of important key residues involved in the communication between the light-sensing LOV domain and the RNA-binding ANTAR module. However, the dynamic structural transitions that occur upon light-activation remain to be elucidated. To assess eventual transitions of secondary structure elements upon light activation, CD spectra for both PAL and PAL-LOV were recorded under dark and blue light conditions (see Figure 40.a). Irrespective of the light condition, both variants feature negative bands at 208 nm and 222 nm, characteristic for high α-helical content of proteins (Holzwarth, 1965). For the isolated PAL-LOV domain, light-treatment induces a slight reduction (≈ 7 %) of the CD signal at 208 nm, revealing signs for a structural transition of an α-helical element upon light-activation. For the full-length PAL receptor (flPAL), no apparent changes can be detected within the near UV region (200 – 250 nm) upon light-treatment. Larger structural order/ disorder transitions or unfolding of secondary structure elements within the PAS and ANTAR domains are hence unlikely.

To investigate the oligomeric state in solution of PAL and PAL-LOV, SEC-MALS experiments were performed under dark and light conditions. The SEC-MALS set-up could only be used at ambient temperature, at which the reversion times for PAL and PAL-LOV from signaling to dark-adapted state are too fast to perform a complete SEC run (see Figure 25). As continuous illumination of the protein sample is not possible during the measurements, we decided to conduct additional size-exclusion chromatography (SEC) experiments at 4°C, at which the theoretical reversion time for both flPAL and the isolated PAL LOV domain is > 3 hours, to assess the effect of blue light on the oligomeric state. For flPAL, the MALS-derived molecular weight (MW) signal resulted in peak tailing at around 120 kDa and a continuous decrease of the MW signal, which suggests a concentration dependent equilibrium between trimers, dimers and monomers. For PAL-LOV, peak tailing of the MW occurred at around 30 kD. Quantification of the MALS-derived average MW of PAL and PAL-LOV (Figure 22.b) in their dark states yielded values of 88.2 kDa and 28.1 kDa, respectively, which are close to the theoretically expected MW of dimers (78.6 kDa for PAL and 31.9 kDa for PAL-LOV; predicted by ExPASy [151]. The

repetition of the SEC experiments at 4°C revealed a slight illumination-induced retardation of the elution peak for flPAL, which could be due to conformational changes during the transition from dark to signaling state. The different characteristics of the peak absorption maxima are due to the different absorption properties of the LOV domain between dark-adapted and blue light induced signaling state. For the isolated PAL-LOV domain, no apparent change can be detected in the elution peak upon light-induction.

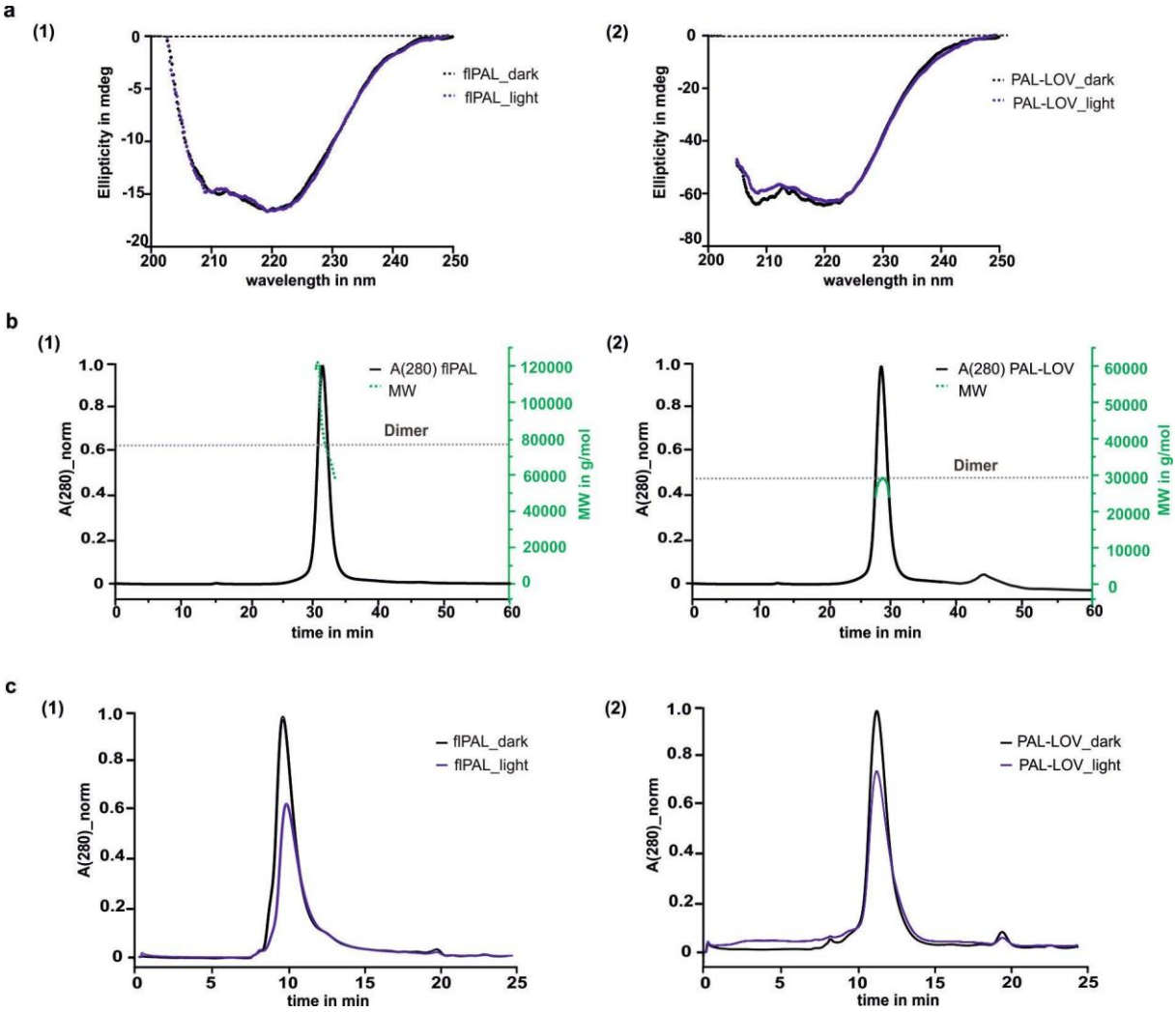


Figure 40: Investigation of structural changes upon light-activation. **(a)** CD measurements of full-length PAL (flPAL) (1) and PAL-LOV (2) in the dark (black traces) and after blue light illumination (blue traces). **(b)** Normalized MALS detection of fl PAL (1), and PAL-LOV (2) in the dark, fractionated by SEC using a Superose 6 10/300 GL column (1) or a Superdex 75 column (2) at a flowrate of 0.5 ml/min, measured at ambient temperature. The MALS-derived signals are depicted in green. Determination of the average molecular weight (MW) yielded values of 88.2 kDa for PAL and 28.1 kDa for PAL-LOV, which are close to the theoretically expected molar masses of dimers. **(c)** SEC of fl PAL (1) and PAL-LOV (2) in the dark (black traces) and after blue light illumination (blue traces), using a Superdex 75 column at a flowrate of 0.5 ml/min at 4 °C. A(280)_norm – absorbance at 280 nm (normalized).

6. Discussion

6.1 Engineering of an ANTAR photoreceptor

The design strategies of Section 5.1.1 and 5.1.2 that aimed at the development of a light-regulated ANTAR photoreceptor did not produce any light-sensitive chimeras, and showed either no or only very weak signs of RNA binding activity. The approaches of the two applied design strategies were fundamentally different; while the engineering approach of Section 5.1.1 was based on a domain replacement upon structural superposition of the α-helical linker region, the design strategy in Section 5.1.2. made use of 'associating' LOV domains, for which light activation causes a change in the oligomeric state (see Section 3.3.2). The discovery of the PAL gene sequence later on offered a new design template on the basis of which the more successful ANTAR chimeras from Section 5.1.3 were developed.

6.1.1 Design approaches prior to the discovery of PAL

The starting point for the design of the first LOV-ANTAR chimera 'LOVA' was the question, whether the underlying signaling mechanisms of the light-regulated histidine kinase YF1 [21,87] could be applied to effector domains of different nature. Among the potential effector domains, we chose the family of ANTAR proteins because the potential acquisition of an optogenetic tool with RNA-binding output function was an attractive side effect in addition to the mechanistic traits. ANTAR proteins naturally occur in diverse combinations with putative sensor modules, including the family of PAS sensors that represent the superfamily of our blue light-sensing LOV modules. The use of a PAS-ANTAR signaling receptor as a design template and subsequent domain replacement of its sensor module by a LOV domain was therefore an obvious consideration. However, none of the PAS-sensor-comprising ANTAR candidates had been characterized up to this time, which would have complicated the choice of an appropriate fusion site as well as the conduction of functional assays due to the lack of structural insights and known RNA target sequences. Among the formerly characterized ANTAR proteins, only AmiR featured a homodimeric conformation and a coiled-coil linker element suitable for structural superposition with the YF1 linker region. This led to the first LOVA chimera developed on the basis of sequence and structural analysis of the two parental proteins (see Section 5.1.1, Figure 11).

Benefits and limitations of bacterial reporter assays

The design of synthetic photoreceptors often requires the testing of high numbers of chimeras. Coupling fluorescence or other easily detectable readouts to chimeric protein activity in a bacterial

reporter assay enables high or medium throughput screening, allowing the simultaneous investigation of multiple protein variants without the time-consuming process of protein purification. Despite these advantages, the use of a bacterial reporter assay also has disadvantages: In case of the LOVA chimera, the use of the β-gal reporter system indicated that the activity level of the variants tested so far, including the initially employed positive control (AmiR resi L139 to A196/end; hereinafter termed 'ANTAR-AmiR'), was at the same level as the negative control. The 'Miller unit' was defined such that the fully induced lac-operon would amount to an activity of 1000 MUs, and the non-induced level would yield around 1 MU [152]. Upon the discovery that the induction rate of ANTAR-AmiR was insufficient to generate a positive reporter signal within the in-cell context, we decided to employ the full-length AmiR receptor as positive control for the screening assay. This resulted in a strong signal ranging from 25000 to 35000 MUs in new rounds of the assay. The reporter plasmid features a pBR322 origin of replication; its copy numer is at around 20 per cell, which accounts for the elevated number of MUs. The availability of a defined measurement unit offers thus an advantage compared to fluorescent outputs, as fluorescence measurements depend largely on the experimental setup. Normalization by a control from a different system is possible, but since each system requires a specific 'gain' setting (see following Section, 'Flourescence-based detection'), we long assumed in the experiments on the LOVA chimeras, for which we had observed binding activity in the EMSAs, that the detection of this activity requires a higher gain than, for example, the Dusk and Dawn system [153].

However, the quantitative translation of the enzymatic activity of the employed β-gal reporter to the output signal depends on several factors, as protein expression and functional testing occur in parallel. Since both processes are sensitive to experimental parameters, such as incubation temperature and time, light intensity, inducer concentration, the plasmid backbone(s), or their simultaneous combination, the reporter assay system is rather complex, making it difficult to distinguish between non-functional protein design and inadequate protein expression. This is a major problem for the use of a reporter assay in combination with rational design approaches, which tend to involve only limited numbers of fusion proteins into their planning and testing. In case of non-functional variants with a complete loss of activity, only little information can be derived by means of a reporter assay. Moreover, the activity curves of signaling receptors within cellular networks are often non-linear due to the cooperative behavior of many signaling proteins [154]. Therefore, small alterations of an input trigger may lead to large changes of the output signal and *vice versa*. The generation of an output signal therefore importantly depends on the adaption to the signaling network, so that dynamic ranges assessed with the help of a reporter screening system are only comparable for protein variants within the same experimental setup. While *in vitro* approaches provide well-defined systems, for which single parameters can be easily controlled, the extraction of

information from a complex cellular system, as given within a bacterial reporter assay, is more difficult. As a consequence, the determination of absolute protein activities, which are a big advantage of *in vitro* characterization, is usually not possible with this method. Yet, in case of successful design experiments in frame of optogenetic application contexts, the direct testing of the functional light response of engineered chimeras within the target system can be more valuable than the determination of absolute protein activities. The removal of a protein from its natural context often leads to unforeseen problems during the implementation of *in vitro* assays; e.g., the requirement for additional cofactors for formation of the functional (dimeric or oligomeric) conformation, or the creation of a suitable ionic environment, which often demand extensive and time-consuming optimization.

Fluorescence-based detection

Fluorescence-based reporter systems are frequently the method of choice, as they do not require any further exogenous substrates for detection owing to their intrinsic chromophore structures. This enables direct monitoring over different growth phases of the cell and a high screening throughput in combination with flow cytometry techniques. [12]. The use of the pE_β-gal reporter instead of a fluorescent reporter system (such as the I1I2_DsRed construct that demonstrated the highest dynamic range among the different promoter-DsRed constructs) is primarily due to the chronological order of the experiments. This means that the improved positive control (consisting of flAmiR) was first successful in combination with the β-gal-reporter, with which the moderate number of chimeric variants within the applied rational design approach could be easily mastered. Nevertheless, the use of the fluorescent DsRed reporter would have clear advantages for the testing of a higher number of variants, e.g. when using directed evolution- or library-based design approaches [155]. Such approaches usually permit to exctract valuable information from the reporter-based detection and selection of successful chimera variants [153], e.g. through the creation of fusion-libraries of two different protein modules that resulted in the creation of chimeras with different linker properties [88].

Nonetheless, it is important to remember that the details of the measurement device (i.e. the sensitivity of the photomultiplier tube) and the setup parameters, such as excitation and emission wavelength and gain, strongly influence detected fluorescence intensity [154]. Fluorescence measurements are particularly susceptible to errors in ranges of weak gene expression, where autofluorescence often causes a high associated error [154]. Autofluorescence derives both from biological structures, as well as many non-biological materials, such as organic plastic polymers. Therefore, at low expression levels, autofluorescence variations often cause high associated errors,

as experienced during the initial attempts to construct the pE_DsRed Reporter System with the LOVA chimeras and the erroneous positive control.

Domain replacement guided by structural superposition

The initial LOVA chimera exhibited good expression in *E. coli*. The qualitatively conserved spectral and photochemical properties (see Figure 13 in Section 5.1.1) indicate that the structure of the LOV module within the LOVA chimera has been preserved. Within the EMSA, the LOVA receptor demonstrated a binding affinity (500 – 600 nM) to the ami-lead transcript comparable to the initial positive control, i.e. the isolated ANTAR domain of AmiR, but no apparent difference under dark and light conditions. However, the β-gal reporter assay revealed that the activity level of the LOVA chimera, as well as the ANTAR-AmiR positive control employed so far, was at the same level as the negative control. The positive control used up to this point only consisted of the C-terminal ANTAR module of AmiR (resi L139 to A196/end). Since protein domains are commonly considered as an autonomous unit of organization that can fold independently into a stable structure and exist and function independently of the rest of the polypeptide chain [156], we assumed that this would be sufficient to maintain the RNA-binding function of the AmiR receptor and thus represent an appropriate positive control. Moreover, the via EMSA determined binding affinity of ANTAR-AmiR was in a similar range as for EutV, another recently characterized representative of the ANTAR family from *E. faecalis* [100]. Since no further comparable experimental data on RNA binding activity was available for AmiR, and the binding affinity of the EutV protein was in the same range as within our EMSA tests, we initially assumed that the binding affinity of ANTAR- AmiR corresponded to a binding level of the active receptor state. In addition, ligand receptor associations are usually concentration-dependent equilibrium processes. Since the physiological concentrations, under which the receptor function is optimally expressed, are often unknown, the determination of the dissociation constant K_d unfortunately does not offer an absolute measure such as ‘binding’ or ‘non-binding’ in order to evaluate the functionality of a receptor [150].

The strongly reduced or terminated reporter activity using the AmiR-ANTAR construct, compared to the full length protein, suggests that the truncation of the ANTAR domain leads to a reduction in RNA binding activity. This is in line with the report of Ramesh et al. [100] on their experiences with the ANTAR signaling receptor EutV. In this study, the RNA binding activities of three different EutV variants were examined: the full-length EutV protein (flEutV) which comprised the N-terminal receiver domain, a coiled-coil region and the C-terminal ANTAR domain; a second construct that included the coiled-coil region (ccANTAR-EutV) and the ANTAR domain; and, just the C-terminal ANTAR domain (ANTAR-EutV) (see Figure 41). The data suggested that the target RNA is bound with

decreasing affinity by ccANTAR-EutV, flEutV and ANTAR-EutV, respectively. Compared to ANTAR-EutV, the binding affinity of ccANTAR-EutV to the RNA target substrate determined via EMSA binding assays was about 100-fold higher. However, the RNA binding activity of ccANTAR-EutV was in a similar range as for the isolated AmiR ANTAR domain within our EMSA tests ($K_d^{ccANTAR\text{-}EutV} \approx 0.7$ µM). Whereas flAmiR shows activities in the fully induced range in the β-gal reporter assay, the binding activity of the flEutV is significantly lower ($K_d^{flEutV} \approx 10$ µM) than for ccANTAR-EutV within the EMSA experiments of Ramesh et al [100]. This fact might be explained by the different activation mechanisms of the two ANTAR proteins: the activation of EutV occurs by means of signal-induced phosphorylation of the N-terminal response receiver (RR) domain, which stimulates dimerization of the EutV monomers thereby enabling the association with its target RNA. The signal transfer is mediated by the corresponding sensor histidine kinase EutW that undergoes autophosphorylation in response to ethanolamine and consequently acts as positive regulator. The activation of AmiR, on the other hand, is carried out with the aid of the negative regulator protein AmiC [102]. The induction of AmiC by small-chain amides leads to release of AmiR allowing its association with its target RNA sequence. Consequently, AmiR displays a constitutive activity in absence of AmiC, whereas in case of EutV, the RR domain seems to attenuate the RNA binding activity of the ccANTAR in its unphosphorylated state in absence of EutW. However, a clear explanation for the reduced binding affinity of ANTAR-AmiR compared to the full-length receptor is not given by previous studies. One possibility is that the RR-part AmiRs is required for the production of the functional dimer unit, but the RR might also (somehow) be required for the antitermination observed in our assay.

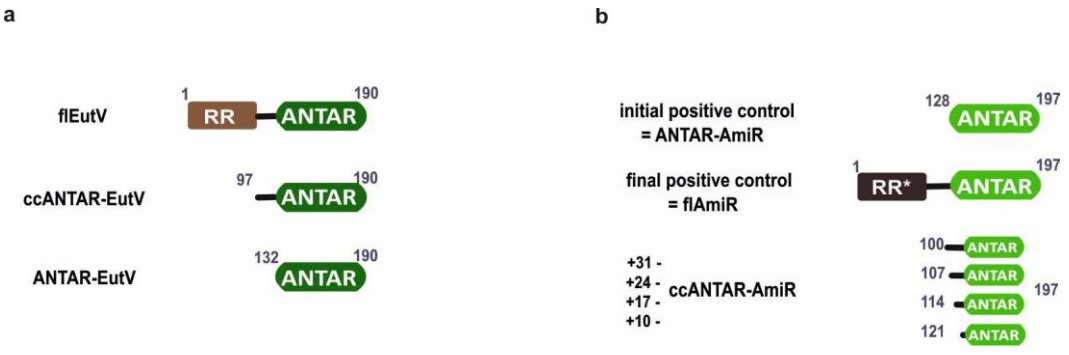


Figure 41: Different activation mechanisms of EutV and AmiR. Scheme for illustration of the conclusions drawn in the previous Section. Even in absence of its positive regulator AmiC, AmiR displays a constitutive RNA binding activity C, whereas in case of EutV, the RR domain of EutV seems to attenuate the RNA binding activity of the ccANTAR module. **(a)** Overview of truncated contructs from the study of Ramesh et al. [95] together with the via EMSA assessed affinity constants (K_d) to their target sites. **(b**) Overview of the truncated AmiR contructs tested within this study. RR – common response regulator receiver domain; RR* - pseudo RR, i.e., residues essential for phosphoryl acceptance lacking.

Effect of linker-length variation

As an insufficient length of the coiled-coil linker might prevent the formation of the functional dimer, we decided to assess if an elongation of the N-terminal coiled-coil region of the AmiR ANTAR domain would improve its RNA-binding affinity. For that purpose, additional Miller assays with extended versions of the initial positive control (ANTAR-AmiR), comprising the ANTAR core module and increasing parts of the coiled-coil (see Figure 19 in Section 5.1.1), were performed to assess if these changes would result in higher reporter activity. However, even the longest of these variants ('+31-ccANTAR-AmiR'; see Table A3 in Section 8.2) that included the full extent of the AmiR coiled-coil region did not result in an increase of the output signal. We hence assumed that further shifting of the fusion site towards an elongated AmiR linker would most likely not improve the binding characteristics. The potential reasons for the persistent inactivity of the LOVA chimeras are diverse: first of all, the light-induced structural signal elicited by the LOV domains might not be compatible with the activation mechanisms of the AmiR ANTAR effector. Another potential reason might be an unfavorable choice of the selected fusion site. Whereas PAS/LOV signaling receptors feature the highly conserved DIT motif that marks the C-terminus of the PAS/LOV core domain, the AmiR coiled-coil linker region that connects the N-terminal (pseudo-) RR domain with the C-terminal ANTAR domain does not feature such a clear boundary. Moreover, the C-terminus of the coiled-coil α-helix features interactions with the three-helical bundle that defines the ANTAR core domain. To maintain these interactions, the fusion site had to be selected further down the coiled-coil linker region. For YF1, it was shown that some of the N-terminal residues of the Jα coiled-coil linker element are essential for the transmission of the activating light stimulus, therefore the fusion site had to be set inside the α-helical motif of the linker region for both sensor and effector. This process is error-prone and can easily lead to irregularities within the linker helix, which would subsequently impair the transduction of the intramolecular signal. Further investigations would be required to clarify which of the discussed explanations is actually valid. Considering that the activity of AmiR within its natural context is controlled by the positive regulator AmiC, differences between the signaling mechanisms are conceivable.

Besides, the linker length analysis for the family of PAS-ANTAR signaling receptors did not reveal any conserved pattern (see Figure 18 in Section 5.1.1), as found for the family for PAS-HisKA. The linker regions of the PAS-HisKA family display a significant heptad periodicity regarding their lengths, which is a result of the preservation of the hydropathy pattern characteristic for coiled-coil elements [35,157]. For YF1, the maintenance of this heptad-periodicity was found to be crucial to establish light-dependent kinase activity. In contrast, in case of the LOVA chimeras, the alteration of the linker length between sensor and effector module, which led to light-regulated constructs in other photoreceptor-engineering approaches, failed to improve the RNA-binding function and light-

sensitivity of the chimeras. The linker variants of the tested LOVA chimeras encompass linker lengths variations of 14 residues, which corresponds to the insertion of two helical turns between the - 4 and +10 –LOVA contructs (see Table A3 in Section 8.2). For the shortened linker variants of the initial LOVA chimera (-1 to -4 – LOVA; see Table A3), the remaining residues of the AmiR linker sequence C-terminal to the start of the ANTAR core domain were consecutively reduced until the beginning of the conserved three-helical bundle. The alternative reduction of N-terminal residues that belong to linker residues of YF1 would have eliminated essential residues, for which single point mutations (e.g. for the residues D21 or V27) were found to have severe effects on signal transduction within YF1 [21,155]. For that reason, a further reduction of the linker length seemed not beneficial. On the other hand, the successive elongation of the linker by 10 residues did not result in any change in activity or light-sensitivity, so that we decided to cease the efforts of designing a light-regulated ANTAR chimera based on the superposition of YF1 and AmiR at this point. Still, further systematic analysis would be required to assess if the lack of activity of the tested chimeras is due to mechanistic differences or to an unfavorable design. One notable further option could be to shift the original fusion site more towards the C-terminus of the linker sequence without moving the overlay grid further, e.g. from E138 to L143. This would be particularly interesting, since Gleichmann et al. [155] observed that point mutations of the YF1 linker residues E142 and L143 led to constitutive kinase activity irrespective of the light conditions. However, unlike the constitutively active YF1 mutants, the so far tested LOVA chimeras have shown no signs of activity within the reporter assay up to now and, and only low affinity to their target RNA in the case of the initial LOVA chimera tested via EMSA. For this reason, we decided to adopt a new strategy instead.

Use of associating LOV modules

As outlined in Section 3.3.2, the use of associating photoreceptors has been proven quite successful for optogenetic engineering attempts in the past, even for effector types that were not regulated by association- or dissociation-based processes before. For that reason, we decided to employ another design approach for building a light-responsive ANTAR protein, by using a LOV domain type for which the transmission of the light signal was shown to occur upon association-based mechanisms.

With (i) the LOV domain from *N. crassa* Vivid ('VVD-LOV'), and (ii) the LOV domain from the *N. gaditana* Aureochrome ('NgAur-LOV'), two different LOV modules were fused at the N-terminus of two different fusion sites of the ccANTAR module from *P. aeruginosa* AmiR via a short and flexible linker, resulting in four distinct LOV-AmiR chimeras. As additional control, the VVD LOV domain was attached to the N-terminus of the full-length AmiR protein via the same linker ('VVD-flAmiR') in order to investigate if this would interfere with the functionality of the ANTAR effector. Only this last

construct resulted in a detectable signal, even though the activity level was sharply reduced compared to the positive control consisting of flAmiR alone. The signal of the remaining associating LOV-AmiR variants was at the level of the negative control, regardless of the light conditions. Only little information can be drawn from these results. The lack of a signal within the reporter assay can have several reasons, as detailed in the previous Section (*'Benefits and limitations of a reporter assay'*), e.g., problems with folding of the protein structure, consequently leading to a reduced yield of protein expression. This is not unlikely, especially for the S90 variants whose fusion point lies within the RR unit. However, other scenarios are conceivable, e.g. the AmiR-RR might be required for the formation of the functional dimer of AmiR, which in turn is important for activation of the RNA binding function [102]. Also, the VVD-flAmiR control construct, whose RNA binding function should be hardly affected by the loosely attached VVD-LOV at its N-terminus, exhibited an about 30-fold reduced activity within the bacterial screening assay compared to the flAmiR positive control. Possible causes are, again, problems with protein folding and expression, which could be caused by an insufficient linker length. An alteration of the length of the connecting linker element would thus be a conceivable option for an optimization attempt. Although the requirements for the sensor-effector connecting linker are comparably low for the design of associating photoreceptors, an insufficient length could cause problems with the formation of the functional dimer, whereas an excessively long linker may reduce the regulatory effect on the respective effector module.

However, only two associating LOV candidates have been tested so far. The different LOV modules have different inert properties: despite several advantages of VVD-LOV, such as its small size or its superior light-sensitivity, some of its properties entail certain limitations for the use within optogenetic applications. In its natural context, light-activation induces the formation of rapidly exchanging dimers that exhibit a rather low dimerization affinity (5.5 - 13 µM) [40]. Even though the employed variant comprises two point mutations (N56K and C71V) that were shown to improve the affinity of the VVD monomers for one another [81], this might not be sufficient to effectively drive the formation of the functional effector conformation in context of the bacterial screening system, within which numerous other protein candidates may interfere with the monomer interactions. In contrast to the vast majority of LOV domains, the NgAur LOV domain features a N-terminally attached bZIP effector. It is possible that the particular sensor-effector orientation plays a crucial role for the functional transduction of the signal from sensor to effector, as the original domain orientation was similarly maintained in the far only successful implementation of NgAur-LOV in an engineering context [83]. Conceivable candidates for further testing would be, e.g., the LOV domain of EL322, a bacterial transcription factor that dimerizes under blue light thereby activating its DNA-binding capacity [39]; or the LOV domain of *R. sphaeroides* [79], which dissociates into monomers after blue light absorption, thus showing the opposite signal polarity as the LOV systems mentioned

above, as well as other Aureochrome proteins from *V. frigida* or *P. tricornutum*. However, since only little information can be extracted from the use of photoactivatable dimerizers from a mechanistic point of view, and we had already discovered the sequence of the natural PAL photoreceptor at the time of these experiments, we ceased the design efforts using associating LOV modules at that point.

6.1.2 Design approaches after the discovery of PAL

The demonstration of light-regulated function of PAL opened up a new perspective for the design of novel RNA-binding photoreceptors, based on the use of the uncommon domain architecture of PAL as a design template. The replacement of the PAL ANTAR module with different ANTAR domains could importantly contribute to the understanding of the underlying signal transduction mechanism. Moreover, the new PAL-derived chimeras may feature divergent properties from the parental PAL protein, such as a differing binding affinity, or the degree of light regulation.

Initial constructs

The already characterized *P. aeruginosa* AmiR and *E. faecalis* EutV systems, for which the target RNA recognition motifs are known, were particularly suitable for domain replacement experiments of the PAL ANTAR domain. Since we already held a working screening assay for the detection of the AmiR function, we decided to start the new approach once again with the AmiR-ANTAR module. We were also interested in investigating potential similarities among the signal transduction mechanisms of PAL-LOV and NgAur-LOV, which are suggested by the parallels of the particular domain orientation of sensor and effector. The fusion site of the initial 'AmiLOV' chimera was based on a sequence alignment of the ANTAR domains of PAL and AmiR, a second chimera ('AmiNgLOV') was then subsequently generated by substitution of the PAL-LOV domain with NgAur-LOV based on a sequence alignment of the two LOV domains (see Figure 39 in Section 5.1.3). This exchange benefits from the presence of a highly conserved DIT consensus sequence motif at the C-terminus of most PAS and LOV domains [35], which delineates the domain boundary.

To investigate whether the PAL-derived chimera variants could exert a regulatory effect on RNA binding function, the two protein constructs were tested within the reporter assay. Both the AmiLOV and the AmiNgLOV variant showed a significantly higher level of activity than any other AmiR chimera produced so far. During the first functional tests with the help of the pE_lacZ reporter assay, neither the AmiLOV nor the AmiNgLOV chimera showed differences between the light and dark activity levels. The shortening of the sensor-effector connector element had also no effect on the activity of the shortened linker ('shlink') variants of the AmiLOV starting construct, whereas the

7shlink-AmiNgLOV chimera showed first light-induced changes between light and dark activity (L/D =1.75) in the reporter screening tests after truncation of the linker element. More experiments would be required to explain the role of the NgAur-LOV domain on the observed light effect, such as tests of different linker length within ANTAR and LOV domain. However, after solving the PAL crystal structure, I focused on optimizing the AmiLOV variant, for which a structural model was available through the combination of the AmiR and PAL crystal structures. Still, a continuation of the AmiNgLOV experiments would be particularly interesting to clarify whether the implementation of the NgAur LOV domain in its original orientation allows to exercise its regulatory potential on this new effector domain.

Structure-aided design

The elucidation of the crystal structure permitted the adaptation of fusion sites with the help of a model that resulted from the structural superposition of the ANTAR domains of AmiR and PAL. During the investigation of the structural model, it turned out that the previous fusion site of the AmiLOV chimera falls within the PAL-own 12α helix (see Figure 22.a in Section 5.1.3), which is adjacent to the C-terminus of the ANTAR core motif. This is potentially problematic because the fusion site could intervene with the helical secondary structure motif. The newly designed xAmiLOV construct displayed first signs of light-induced activity (L/D ≈ 1.85; see Figure 23.b/c in Section 5.1.3), following the displacement of the fusion interface, while the xPASAmiLOV chimera demonstrated an activity level comparable to the negative control in tests with the reporter assay. Since the introduction of the additional fusion interface involves the risk of disrupting protein folding and stability in various ways, the loss of activity is most likely due to a defective design. The structural model was further used to identify single residues within xAmiLOV (see Figure 22.b), whose substitution could lead to further improvements of the chimeric activity. The Q148A point mutation resulted in an improvement in activity and L/D ratio (L/D ≈ 2.21; see Figure 23.c) compared to the original xAmiLOV chimera. In the structural model, the glutamine residue Q148 had previously disrupted the coiled-coil interface of the PAL 9α helices (see Figure 22.a); the Q148A point mutation shows the intended effect of resolving this problem. All of the otherwise introduced point mutations resulted in a reduction of the overall activity. However, the combination of the Q148A point mutation seems to suppress this negative effect in many cases, such as in the case of xAmiLOV-Q148A-G151D-W152V and xAmiLOV-Q148A-K167Q-E164H (see Figure 23). Also, the conservation of charged residues in proximity to one another within the AmiR structure (e.g., G151D, E164H) was expected to have a positive effect on the internal signal transduction, which appears to be the case in combination with the Q148A mutation.

However, as pointed out before (see previous Section, '*Benefits and limitations of a reporter assay*'), the output signal of the reporter assay is influenced by many different factors, so that there is a risk that the differences between the light- and dark-activity levels might be due to artifacts. Nevertheless, the low associated errors (uL/D), calculated by Gaussian error propagation of the individual uncertainties from the light and dark activity measurements, are a good indicator for the accuracy of the results. The complete loss of photosensitivity of the mutated active site cystidine xAmiLOV-148-C284A control is a further positive sign that the observed light effect in the reporter screening system is not just an artifact. Still, caution is advised with this type of control, as it was shown that the signaling state can still be achieved at higher light intensities despite this mutation in other LOV photoreceptors. Although the C → A mutation of the active site cysteine prevents the formation of the covalent bond, it promotes the photoreduction of the flavin cofactor to the neutral semichinone at high light doses. Like the cysteinyladduct, the semichinone radical is protonated at the N5 position, so that the signal can propagate similar to wild-type LOV proteins via a cascade of changes in hydrogen bonds [158]. Under the given experimental conditions, however, we observe the expected effect of the removed photosensitivity. In contrast, the shift of the fusion site of the initial AmiLOV chimera (designed before crystal structure availability) had a positive effect on photosensitivity, so that there is a clear correlation between changes in the regulating photoreceptor chimera and the measured reporter signal.

In conclusion, this last design approach, which was closely oriented on a natural design template, is by far the most successful compared to the previous engineering attempts in the scope of this work. The exchange of the ANTAR domain has created a new chimeric PAL-variant with an altered sequence specificity, which has a good potential to extend the range of available optogenetic tools in the future. Moreover, these results demonstrate that the availability of structural details, provided by the crystal structure of PAL in its dark-adapted state, as well as by the SEC and SEC-MALS data, largely facilitate the generation of light-regulated RNA-binding photoreceptor proteins. The experiments further confirm the conclusions already drawn in an earlier review [6] that design approaches become more difficult the further one deviates from the natural (or at least provably functional) system, and benefit greatly from prior existing mechanistic knowledge. For a quantitative comparison of the RNA binding function with that of the original PAL photoreceptor that served as a design template, subsequent *in vitro* experiments would be required. However, the establishment of such experiments, including the optimization of protein expression, purification and stability, is often very time-consuming. Therefore, I gave up the project after one test run as my laboratory time was often differ from the conditions required in *in vitro* experiments, so that the benefit should be thoroughly weighed up.

Outlook

The construction of a AmiR-derived chimera based on the blueprint of PAL enables the employment of the already established pE_β-gal or DsRed-based reporter assays for the testing of mutant variants, which provides a great means to deepen the knowledge of the internal signal transduction mechanisms in PAL and its derivatives. The use of a fluorescent reporter, e.g. the I1I2_DsRed plasmid, under the control of a PAL-derivative would further permit the construction of a larger library, e.g. by random mutagenesis as previously described [155,159] and subsequent selection of variants based on detection of the fluorescent reporter, e.g. via fluorescecnce-activated cell sorting (FACS). This would permit the analysis and isolation of variants featuring a desired property, such as an improved dynamic range or an inverted signal polarity.

6.2 Characterization of the natural photoreceptor PAL

Nature has yielded several photoreceptors with immediate optogenetic utility in the past [160], such as different photoactivated cyclases from several organisms [90,161–163], or the bacterial transcription factor EL222, which was used to construct a light-inducible eukaryotic gene expression system [39,164]. These natural actuators often surpass the corresponding engineered variants in terms of specific activity and dynamic range [6]. After the discovery of the PAL gene sequence within the data bases, we thus focused on the functional and mechanistic characterization of the respective protein product that comprises the desired building blocks for light-sensing and RNA-binding function. Our studies on PAL provide structural and functional insights into allosteric signaling of a multi-domain LOV protein featuring a domain architecture opposite to the sensor-effector topologies found in most other LOV proteins, for which the effector domain is situated C-terminally to the LOV sensor.

6.2.1 Functional characterization

In order to identify some first specific RNA target sequences, we opted for the SELEX method [165]. This led to the unveiling of two different families within the SELEX pool sequences of the last selection cycle, defined by a common consensus motif. The RNA aptamers identified via SELEX were first tested by Anna-Maria Pyka in Bonn for their binding behavior to PAL under dark and blue light conditions via a nitrocellulose filter binding assay. The implementation of the filter binding assay is straightforward, but quantitative analyses are often hampered due to filter-retention properties [166]. For that reason, the five most promising variants were subsequently analyzed with the help of EMSA assays in order to assess their apparent affinity to PAL under dark and blue light conditions.

The EMSA method offers a number of advantages [143], e.g. the assay is easy to perform in principle (although, depending on the type of oligos used, complications may occur) and compatible with a variety of binding conditions. In classical experiments the electrophoretic mobility of the nucleic acid oligo is followed, which can be done with a broad range of sizes and structures. The oligo can be labeled radioactively, or with fluorophores [167] or biotin [168]. Using radioactively labeled nucleic acids, the method is sensitive enough to be performed with low (up to picomolar) concentrations and small sample volumes. Within this study, the RNA substrates were prepared through *in vitro* transcription and radiolabeled with ^{33}P isotopes. The EMSAs were then performed at increasing PAL concentrations and evaluated by means of autoradiographic monitoring. PAL demonstrated a blue light-dependent binding activity for all of the tested constructs within the EMSA experiments. Prior to further optimization, the best-binding 04 aptamer demonstrated an apparent K_D of around 130 ± 30 nM under light conditions, which represents an approximate 30-fold enhancement compared to the corresponding binding activity in the dark. Further optimization by combining the sequences of the minimal required region of the 04 substrates resulted in an improved affinity of 30 ± 3 nM. In order to make more detailed statements on the essential properties of the aptamers for efficient association with PAL, further investigations on sequential and structural requirements of the target RNAs through mutations of individual residues are undertaken by our cooperation partners at the University of Bonn.

Comparison of the identified RNA targets with known ANTAR binding motifs

Using the SELEX method, we succeeded in identifying the first RNA target sequences that are light-dependently bound by the PAL photoreceptor. However, the determined aptamer sequences are artificial substrates, and the contained consensus sequences are too short to assign them to specific genomic target regions within *N. multipartita*. The via SELEX determined 40 bp-long RNA target sequences are also significantly shorter than the conserved double loop structural motifs of the previously characterized ANTAR representatives, so that it could resemble at most a sub-region of this motif. A further extension of the random nucleotide sequence of the SELEX pool would have been difficult, as the number of possible sequences in the library is 4^n ; e.g., $1{,}2 \times 10^{24}$ for n = 40 and $1{,}46 \times 10^{48}$ for n = 80, so that only a marginal fraction of all possible sequence combinations would have come into contact with the PAL target per incubation cycle.

For the 'bestperforming' aptamer, the 04di10 tandem variant, a binding affinity to PAL of around 30 nM was determined in the EMSA experiments. For the original 04 clone, the determined K_d value is about twice as high (around 80 nM), which could be due to measurement inaccuracies due to the impaired quality of the EMSAs. For the measurements using the Ribogreen Assay, the K_d values for

the 04di10 tandem construct are in a similar range as within the EMSA experiments. However, the binding affinity of PAL to the original 04 aptamer could not be assessed with this method, as it is only suitable for longer (< 100 nt) aptamers. There is no comparable data for the binding affinity of other ANTAR receptors to their RNA target sequences. Even though for Eut, several EMSA experiments were carried out in which an RNA binding affinity of around 10 µM was determined for the full-length protein, it was not possible to reproduce these experiments in sufficient quality together with the activating phosphodonor to determine the RNA binding affitity in the activated state [100].

Nevertheless, it is interesting to check whether the aptamers identified via SELEX show similarities to the natural target sequences of already characterized ANTAR proteins. Baker and Perego [169] found several regulatory sequences in the Eut operon of *E. faecalis* that share a common 13 nt-long sequence ('AGCAANGRRGCUY'; however, within the EutP 5'-leader region I could only identify a modified 12 nt-long form: 'RGCAANGRRGCY'). This consensus sequence overlaps with a 5'-proximal region of the intrinsic terminator element, suggesting that this novel RNA motif might constitute a specific antiterminator structure. In another study, Ramesh and colleagues demonstrated that this motif contains all major determinants of EutV-ANTAR domain recognition [100]. A smilar RNA motif was found within the previously characterized AmiR and NasR ANTAR regulators, indicating that the mechanism of ANTAR-mediated control is based on a common recognition element. According to the model of Ramesh et al. [100], this structural motif consists of a pair of stem loops (P1 and P2), of which the 5'-proximal P2 stem loop conmprises the 13 nt-long consensus sequence (see Figure 43.a). For EutV, it was shown that the recognition of the target sequence relies on a combination of sequence and structural determinants. Regarding the sequence determinants, the residues 1 and 4 of the hexaloop region ('A' and 'G') were found to be essential for efficient recognition by EutV, which were also the highest conserved positions withinthe consensus sequence [100]. The same residues were previously identified as crucial components for successful antitermination by the NasR ANTAR regulator [96]. The 13 nt-long consensus sequence does not appear within the via SELEX identified sequences, which is not particularly surprising since the SELEX-based sequence optimization is based on a synthetic starting template. On a structural basis, the shared core motif of the SELEX clones from family 1 (see Figure 28 in Section 5.2.3, or Figure 42) also differs from the structural P2 motif found within the lead sequences of previously characterized ANTAR regulators within the number of nucleotides of the loop-region that contains 7 nt instead of 6 nt (see Figure 43). However, the hairpin region of the 04 clone does similarly contain mainly A and G as bases. Once natural target sequences of PAL are identified, it will be interesting to compare them in terms of sequential and secondary structural features, as well as in terms of affinity and dynamics. The gain of knowledge of structural and sequential binding determinants may be useful for the construction of

new aptamer templates for future applications, or for the construction of a PAL-based reporter gene assay that exploits the transcriptional antitermination (or different) mechanism.

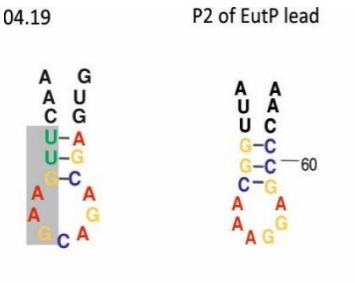


Figure 42: Comparison of the via SELEX identified core motif to known ANTAR binding motifs. The common core motif of the SELEX clones from family 1, here exemplified by the shortened 04 19mer, differs from the structural P2 motif found within the previously characterized leader sequences of ANTAR regulators in the number of nucleotides of the loop-region, that comprises 7 nt instead of 6 nt, but features a similarly high content of A and G bases.

In order to assess the validity of the model from Ramesh et al. [100] for an unrelated species, I tried to identify the said consensus sequence in the AmiE lead sequence from *P. aeruginosa* (see Figure 43.b). The mentioned 13-nt consensus sequence can be easily identified within the stemloop region 5'-proximal to the intrinsic terminator (again slightly different: 'AGCAANURRGCUY'), but is missing in the first of the two stem loops (P1). In an earlier study that investigated the transcriptional antitermination mechnism of AmiR, Wilson and colleagues [110] were able to identify two loci within the AmiE lead sequence (CCGAAC and CACAGAGCA, starting 36 and 54 positions downstream of the transcription start site). The first of these two regions is located at the 5'-end of the first hairpin within the AmiE lead sequence (see Figure 43.b). If the antitermination mechanism described by Ramesh et al. (2012) for EutV can be transferred to AmiR, the AmiR interaction with the structural motif may form the P1 motif shown in Figure 43.b, which does not contain the complete 13 nt-long consensus sequence, but at least the highly conserved 1 and 4 positions within the hexaloop region of the stem loop. Wilson et al. [110] found that point mutations within these two regions abolished sequence recognition by AmiR. The 12 bases between these regions were insensitive to miss-mutations; however, insertions or deletions in this intermediate segment led to a reduced efficiency of antitermination. This supports the idea that the three-dimensional structure of the two regions is important for an efficient interaction with AmiR. If the model of the structural change resulting from the interaction of AmiR with the AmiE lead transcript applies, (see Figure 43.b) valuable hints could be drawn for the design of new ANTAR-recognition motifs. Since our established reporter assay is only compatible with AmiR-derived constructs, our engineering efforts to date have been limited to

working with this particular ANTAR domain. Even though we had once considered using the well-characterized EutV receptor for our engineering efforts, we soon refrained from doing so because the in *E. faecalis* established assay [100], could not easily be transferred to *E. coli*. The assay is based on EutV interaction with the EutP lead sequence, that would cause the expression of a reporter gene through antitermination, but the leakiness of the terminator was too high in *E. Coli* to detect any diffences in signal upon EutV expression. If one compares the two structural motifs shown in Figure 43 (a) and (b), it it obvious that the P2 terminator within the EutP lead sequence is significantly more pronounced in terms of stem length than the P2 hairpin within the AmiE lead sequence. Through shortening the EutP lead terminator, e.g. by combining the two motifs (see Figure 43.c), it might be possible to create an *E. coli* - compatible reporter assay for the screening of EutV-derived chimera variants.

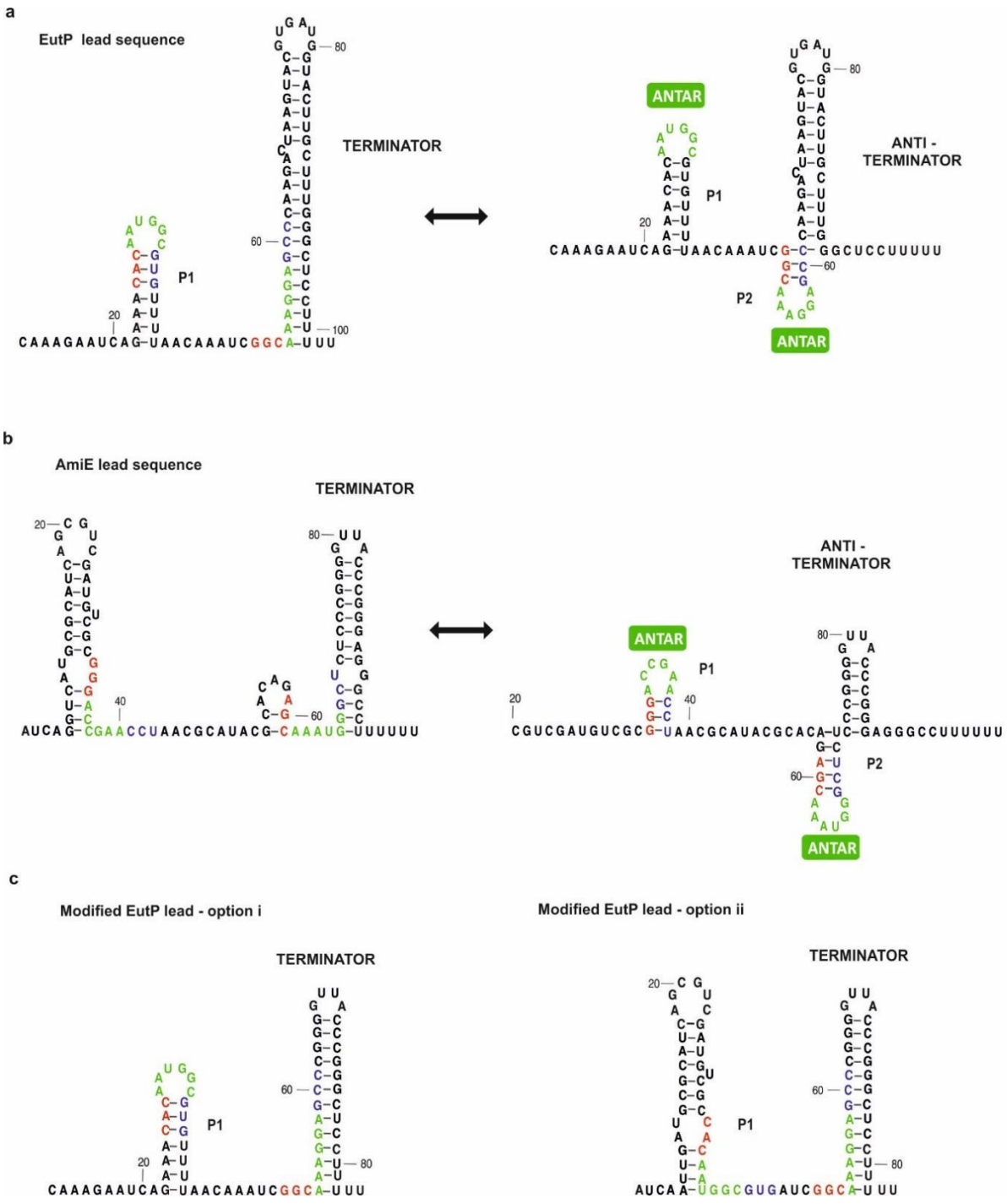


Figure 43 (at left): Proposed mechanism of antitermination. **(a)** Predicted secondary structure of the EutP lead sequence from the Eut operon of *E. faecalis*, with the consensus motif that constitutes the antiterminator structure highlighted in colors. For EutV and NasR, the positions A1 and A4 of the hexamerix loop region were shown to be essential for antitermination, so that it was hypothesized that each monomer of the functional ANTAR dimer binds to one of the two stemloops (P1 and P2). **(b)** Predicted secondary structure of the AmiE leader sequence from *P. aeruginosa*, with the antiterminator consensus motif highlighted in the same color code as in (a). **(c)** This could provide valuable information for the design of new ANTAR detection motifs, e.g. for the establishment of an *E. coli*-compatible reporter assay for the screening of EutV-derived chimeras. Comparing the two structural motifs from (a) and (b), it becomes clear that the P2 terminator in *E. faecalis* is much more pronounced than the P2 hairpin within the AmiE lead sequence, which leads to an effective termination of the transcription within *E. coli*. The shortening of the P2 terminator of the EutP lead sequence through combination of the two motifs presents a conceivable option for creating a reporter assay for screening EutV-derived chimeras within *E. coli*.

6.2.2 Exploration of the biological role of PAL

So far, only little is known about the actinobacterium *Nakamurella multipartita*. The bacterium was first isolated from active sludge in a study by Yoshimi and colleagues [170]. Due to the phylogenetic position and presence of a unique set of 16S rRNA sequence signatures, the new family of Nakamurellaceae (formerly Microsphaeraceae) was established [171], which today includes four published species (*N. multipartita*, *N. panacisegetis, N. flavida* (formerly *Humicoccus flavidu*s) and *N. lactea* (formerly *Saxeibacter lacteus*) that form a robust phylogenetic clan [172]. *N. multipartita* is a coccus-shaped, gram-positive, strictly aerobic, non-motile and nonspore-forming bacterium [170]. The predominant menaquinone MK-8 and meso-diaminopimelic acid within the cell wall were determined as characteristic chemotaxonomic markers. The *N. multipartita* genome consists of a single replicon comprising 6,060,298 bp and a GC content of 70.92 %. Within the genome 5471 genes were predicted, including 5415 protein-coding genes and 56 RNA genes of which 66.5 % were assigned a putative function. The evaluation of the distribution of genes within their 'clusters of orthologous group' (COG) category revealed that the highest number of genes is involved in transcription (400, 9.1 %), followed by genes involved in carbohydrate transport and metabolism (341, 8.3 %), and genes implicated in amino acid transport and metabolism (334, 8.1 %) [172]. The search for additional ANTAR effectors revealed 15 more hits within the genome via BLAST search [173]. The so far characterized members of the ANTAR family are involved in the regulation of bacterial gene expression via transcriptional antitermination. The presence of this large number of ANTAR proteins within *N. multipartita* (in contrast, EutV host organism *E. faecalis* contains 5 ANTAR proteins, while BLAST search in AmiR host *P. aeruginosa* reveals 20 hits of putative ANTAR proteins) indicates that a multitude of processes may be subject to ANTAR-controlled regulatory mechanisms. Furthermore, the BLAST search revealed three other putative PAS/ PAC-sensor proteins, one of which represents another LOV protein that belongs to the group of short-LOV proteins that lack a covalently bound effector module. The search for further blue light receptor domains (BLUF, bacteriophytochrome, bacteriorhodopsins, cryptochromes) via BLAST search revealed no further hits within the genome. Potential blue light-induced effects could thus be attributed to the interaction of one the two LOV proteins in perspective studies.

So far, the *N. multipartita* DSM 44233T strain was the only member of the family Nakamurellaceae for which the complete genome sequence was accessible [174], but through the recent sequencing project of the *N. lactea* DLS-10 type strain genome in 2017 [175], another genome became available. However, the *N. lactea* genome does not seem to involve a PAL-like sequence. We were able to grow the *N. multipartita* cultures in Trypticase Soy Broth medium by DSMZ, as well as in ordinary LB medium in a temperature range of 25 °C - 30 °C. In order to ensure the growth of pure *N. multipartia* cultures, I also tested various antibiotics to determine possible resistance levels, which I detected for

ampicillin (100 µg / mL), gentamycin (10 µg / mL) and nalidixic acid (30 µg / mL), which is in accordance with the findings of Kim et al. [172]. In previous experiments of a several-week incubation of streaked *N. multipartita* cultures on agar plates, I could not detect any signs of differing phenotypic characteristics under light and dark cultivation conditions. To my knowledge, so far no experiments of genetic manipulation have been performed, which complicates the conduction of reverse genetic approaches, such as knock-out or RNA-mediated knock-down studies. Hence, we can only hypothesize about the biological role of PAL in its natural host so far. The efforts of the RIP experiments already initiated are an important step towards the elucidation of the natural function of PAL, as we were able to demonstrate that our generated polyclonal antibody is both specific and suitable for the planned pull-down application, and the cultivation of *N. multipartita* was successfully established in our laboratory. The identification of the natural target sequences would be the first step in the determination of the natural function of PAL, as the analysis of the RIP data will hopefully reveal RNA sequences associated with the genomic location of corresponding operons controlled by PAL. The thereby identified sequences could be further refined through combined bioinformatics approaches. Yet, the detection of candidate genes identified by this technique will require further validation by additional methods, such as quantitative real-time PCR (qPCR), which would permit the quantification of transcript levels of the candidate genes under dark- and blue light conditions. The identification of the natural targets would further permit to assess if the characteristic antitermination mechanism of the so far characterized ANTAR proteins is preserved within *N. multipartita*. As our experiments with the SELEX-derived artificial RNA substrates could already demonstrate that light-activation leads to an increase in affinity of PAL for certain RNA substrates, this suggests an involvement of PAL in posttranscriptional regulatory mechanisms.

6.2.3 Photochemical characterization

UV/Vis absorption spectroscopy was applied to follow the dark state recovery of PAL and the isolated PAL LOV domain after blue light activation. This property plays an important role within *in vivo* applications, since the reversion rate from signaling to dark state importantly defines the effective light-sensitivity of LOV photoreceptors (see Section 3.2.3). For the heterologously purified flPAL, the recovery to the dark-adapted state occurres with a time constant of 1270 ± 100 s at 25° C, the PAL photoreceptor thus falls within the regime of 'intermediate cycling' LOV proteins in terms of the adduct decay [139]. In contrast, the time constant of the isolated PAL LOV domain is significantly accelerated with 470 ± 40 s at 25° C. I further assessed the temperature dependence of the recovery kinetics, which relates to the cleavage of the covalent bond of the photoproduct, for both full-length PAL and PAL-LOV. As most LOV photoreceptors, both constructs show a linear Arrhenius behavior

regarding the temperature dependence of the recovery kinetics, although non-linear Arrhenius behavior for photoreceptors has been reported, e.g. for the photoactive yellow protein or the transcription factor EL222 from *E. litoralis* [176,177]. For EL222, sequence and mutagenesis studies have shown that this effect is due to a glutamine to alanine mutation also found in related LOV proteins from different marine bacteria [177]. The activation energy determined for the full-length PAL receptor is in a comparable range (E_a^{flPAL}= 60.6 ± 3.3 kJ /mol) to that of EL222, whose E_a is 63 ± 2 kJ mol below 45 °C [177].

The dark recovery kinetics were modulated by the introduction of the well-documented mutation of residue T250 (or T32 withinthe isolated LOV domain) of the chromophore binding pocket within the PAL LOV domain. For the mutation of the corresponding residue, a strong decelerating effect on the off-kinetics was found for AsLOV2, YtvA and Vivid LOV domains [56,178,179]. The T32 position corresponds to the residue V416 in AsLOV2 (see Figure 44); V416T substitution was shown to significantly accelerate the AsLOV2 recovery kinetics by a factor of more than 20, whereas V416I and V416L substitution led to a deceleration of dark recovery by a factor of more than ten or 50, respectively (Kawano et al., 2013). Initially, three different variants of the T32 position of the PAL LOV domain were designed and expressed (T32V-/ T32I-/ T32L-PAL-LOV), but only the T32V PAL-LOV mutant yielded sufficient amounts of holoprotein. The introduction of the T32I and T32L mutations resulted in poor protein stability. UV-Vis characterization of the T32V PAL-LOV domain revealed an about 2.5 x decelerated time constant compared to wild-type PAL-LOV ($\tau^{T32V\text{-}PAL\text{-}LOV}$ = 1980± 10 s at 20 °C vs. $\tau^{PAL\text{-}LOL}$ = 765 ± 5 s), and hence a prolonged lifetime of the signaling state of T32V PAL-LOV. The development of further sets of mutations for the generation of optimized PAL variants with different recovery kinetics will be useful for optogenetic applications, since the off-kinetics also affect the effective photosensitivity in the photostationary state under constant illumination (see Section 3.2.3). Figure 44 summarizes some of the previously altered residues for rate-mutating effects within other LOV proteins within an alignment with the PAL LOV domain.

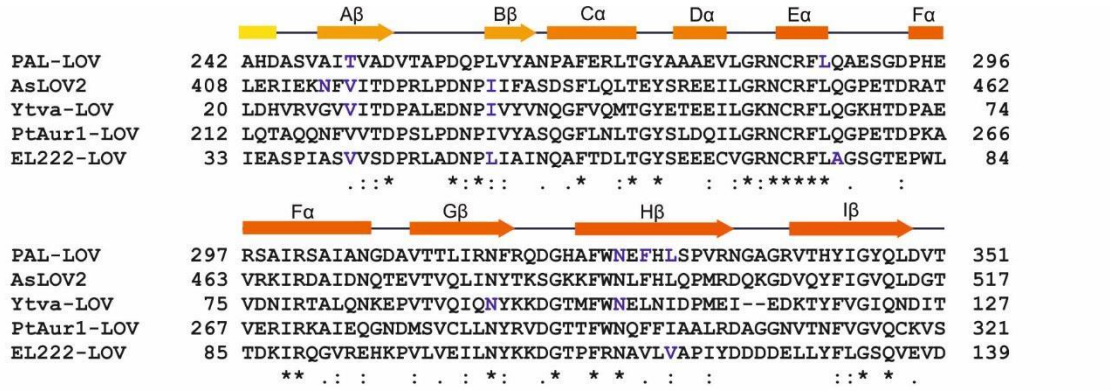


Figure 44: Alignment of different LOV proteins showing previously altered residues resulting in mutating effects of the dark recovery kinetics highlighted in blue [139].

6.2.4 Structural and mechanistic characterization

The unusual domain arrangement of PAL featuring an N –terminal effector attached to the blue light sensing LOV domain is opposite to the sensor-effector topologies found in most LOV proteins. This raises the question how the transmission of the activating light signal and subsequent triggering of the RNA-binding function can be achieved at the mechanistic level. The successful establishment of heterologous protein expression in *E. coli* and the adaptation of buffer conditions that confer an improved protein stability provided the basis for the crystallization attempts and further mechanistic *in vitro* characterization experiments.

Investigation of light-induced conformational changes

To confirm the dimeric state of the dark-adapted PAL protein in solution, we conducted SEC-MALS experiments. To assess eventual light-induced changes of the oligomeric state for PAL, as well as for the isolated PAL-LOV domain, additional SEC-experiments were performed at 4°C, at which the adduct decay for PAL and PAL-LOV occurs significantly slower than the duration of a complete SEC run (the SEC-MALS device could only be used at room-temperature). Compared to wild-type PAL-LOV, the protein stability of the T32V mutant was compromised, so that we decided to avoid using the T32V PAL-LOV variant for SEC-MALS characterization, and to perform the SEC measurements at lower temperatures instead. The combination of SEC-MALS and SEC experiments provides evidence that PAL occurs as a dimer in solution, regardless of the light conditions. Similarly, the isolated LOV domain does not undergo any detectable changes in its oligomeric state upon blue light illumination. The allosteric regulation mechanisms of PAL are hence presumably not based on association or dissociation, allocating PAL to the group of non-associating photoreceptors (see Section 3.3.1). However, the oligomeric state of proteins is concentration-dependent [180]. Thus, a monomeric conformation cannot be ruled out at lower concentrations. Since the cellular concentration of PAL is

not known under *in vivo* conditions, a regulatory mechanism that relies on light-induced LOV dimerization cannot fully be excluded. Yet, the protein concentrations within our SEC experiments were far below the protein concentrations used by Heintz and Schlichting [68], who previously carried out similar studies on PtAur1a. The SEC-MALS measurements on the full-length protein and the isolated PtAur1a LOV domain were performed at protein concentrations of 100 µM and allowed the observation of a light-induced dimerization effect for the isolated PtAur1a LOV domain. The determination of the concentration range for the monomer-dimer transition using microscale thermophoresis (MST) revealed a K_D of 13.6 ± 1.4 µM for the dark-adapted PtAur1a LOV domain [68]. Using isothermal calibration (ITC) experiments, a K_D of 64 ± 11 µM was determined for the same LOV domain [180]. Such measurements of the concentration dependence of the monomer-dimer equilibrium would also be very interesting for PAL.

To assess the light-induced changes in the content of secondary structure of PAL and the isolated PAL LOV domain, we further conducted CD spectroscopy measurements in the far UV range. Several LOV photosensors exhibit a partial loss of α-helicity in their blue light-induced signaling state relative to their dark-adapted state, e.g., in the most extreme case of AsLOV2 up to 20 % [181]. However, the light-induced changes of the CD spectra are negligible for flPAL. For the isolated PAL-LOV domain, light-treatment induces a slight (≈ 7 %) reduction of the CD signal at 208 nm, indicating partial unfolding of α- helical structural motifs within the LOV domain [182]. The amount of signal reduction is comparable to values that were previously determined for the Aur1a LOV domain from *V. frigida* [183].

Crystal structure of PAL in the dark-adapted state

The solution of the crystal structure of PAL confirmed the dimeric arrangement of the photoreceptor within its dark-adapted state. The dimer interface extends along a vertical axis composed of the coiled-coil α-helical linker between PAS and ANTAR domain and the N-terminal α-helical extensions of the LOV domain (see Figure 35). As outlined in Section 3.2.1, LOV photoreceptors occur in various effector combinations and arrangements in nature. The majority of bacterial LOV proteins feature a C-terminal effector domain. For these more common LOV-proteins [30] the C-terminal Jα element usually acts as important transmitter in signal transduction, as it couples the light-induced changes within the chromophore-binding pocket of the photosensor to the distant effector module (see Section 3.2.4). The opposite orientation of the PAL LOV domain along its horizontal axis presents an intriguing answer to the question how light signaling can be achieved for this rare domain topology by positioning the established Jα signaling module at the interface of the ANTAR effector (see Figure 35). Before interpreting this observation, the information resulting from the crystal structure is

briefly completed here. The rotation of the LOV domain is enabled by the long flexible linker that connects the blue light sensing LOV module to the RNA-binding ANTAR effector. The PAL ANTAR domain comprises an additional helix (12α, see Figure 36 in Section 5.2.6) that maintains multiple polar contacts to Jα and the LOV core region in addition to the characteristic three-helical bundle that defines the characteristic core motif of ANTAR proteins. The pronounced kink within the PAS-ANTAR connecting coiled coil region (built of the 9α helices) results in a relocation of the PAS core domain of chain B, which falls out of the symmetry maintained in the rest of the PAS dimer. The kink in the coiled- is probably due to crystal packing, but a functional relevance cannot be excluded.Signal transduction in LOV photoreceptors with C-terminal sensors

So far, the only described exceptions to the more common LOV architectures with C-terminal effectors domains originate from the aureochrome family only found in stramenopiles, which typically consist of an N-terminal DNA binding bZIP domain and a C-terminal LOV domain [67]. As within PAL, the N-terminal effector is connected to the LOV sensor via a long flexible linker. Many studies have dealt with the question of how signal transmission can be achieved in case of this unusual domain architecture. Upon solving the crystal structures of the PtAur1a LOV domain in the light- and dark-adapted state, Heintz and Schlichting [68] proposed a model for the aureochrome signaling mechanisms. According to this model, the PtAur1a LOV monomers interact directly with the bZIP domain in the dark, thereby inhibiting the DNA-binding function. Illumination with blue light triggers the dissociation of the LOV domain mediated by the terminal A'α and Jα helical extensions from the bZIP domain. Dissociation is followed by dimerization of the two LOV domains, thereby increasing the affinity of PtAu1a for its target DNA sequence. The model is supported through the dark- and light-state crystal structures of the PtAur1a LOV domain [68,180], as well as former FTIR studies on the PtAur1a LOV domain and truncated constructs [78], as well as additional hydrogen/ deuterium-exchange coupled to mass-spectroscopy (HDX-MS) and small-angle x-ray scattering (SAXS) experiments on the full-length protein [68]. The so far collected data on PAL contradicts an association- or dissociation-based activation mechanism (see Figure 40 in Section 5.2.6), since no light-induced effect on the oligomeric state could be detected for flPAL or the isolated PAL-LOV domain. Structural superposition of the PAL-LOV domain with the PtAur1a LOV domain (Figure 45.a) reveals that the dimerization interface of the light-adapted PtAur1a LOV domain (at which the functional dimer is formed) does not match well with the dimerization interface of the PAL LOV domain. Also, the positions of the helical A'α and Jα extensions differ strongly. However, the dimer interface of the PtAur1a LOV domain could be corrupted due to the truncation of the molecule. Parallels between the signal transduction mechanisms of PAL and PtAur1a are thus possible, although there is no clear evidence.

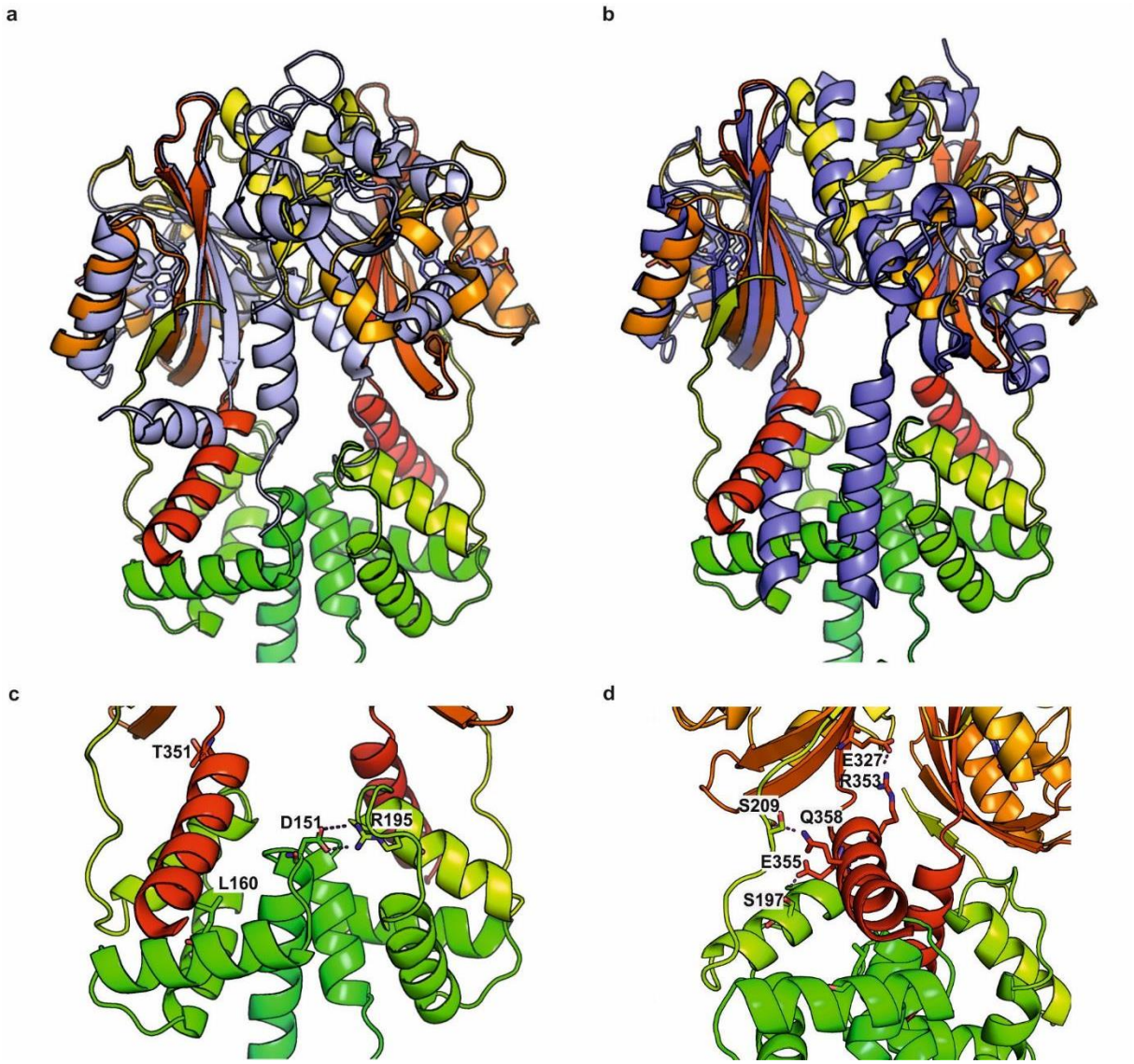


Figure 45: Interpretation of the current data to discuss the mechanisms for signal transmission in PAL. **(a)** Structural overlay of LOV domains from PAL and light-state PtAur1a (PDB code 5DKL; colored in light blue), in which it adapts its functional dimeric state. Neither the helical A'⟨ and J⟨ helical extensions, nor the dimerization interface of the LOV monomers, match well. **(b)** Structural overlay of the PAL LOV domain with the LOV domain of YF1 (PDB code 4GCZ; colored in blue). Here, the dimerization interface and position of the N-terminal A'a extension are in good agreement. For the C-terminal J⟨ helix, only the positions of the 'DIT' motive and the subsequent helical turn match well for the better fitting monomer overlay. **(c)** Depiction of all mutated residues, which resulted in a disruption of the light-sensitivity of the respective mutant variants within the activity test via Ribogreen assay. **(d)** Depiction of the mutant-variant residues, which demonstrated a behavior similar to wild-type PAL in the Ribogreen activity tests, along with the H-bond forming partner residues.

In contrast, the structural overlay of PAL-LOV with the LOV domain of YF1 (Figure 45.b), which is known to occur as a dimer in solution [65,185], results in relatively good agreement regarding the dimerization interface and position of the N-terminal A'a extension. The position of C-terminal Jα helices, on the other hand, differ rather strongly, although the positions of the conserved 'DIT'-motif and the subsequent helical turn match well for the better fitting monomer overlay. In contrast to the YF1 LOV domain, within which the C-terminal Jα extensions merge into the coiled-coil linker that runs parallel to the backbone of the molecule, the Jα helical extensions of the PAL LOV domain deviate from the symmetry axis of the molecule and instead interact with the PAL-specific 12α helix (see Figure 36).

The 12α helix maintains multiple polar contacts to Jα and the LOV core, and is hence presumably important for the transmission of the light stimulus and the establishment of the RNA-binding function of the ANTAR effector. This assumption is supported by the fact that the mutation of L160, that makes hydrophobic contact with Ja helix, to a charged arginine residue results in a constitutively inactive PAL variant in tests with the Ribogreen assay. Also, the mutation of the 195 position (R195L), which maintains a polar contact to D151 of helix 9α of the ANTAR domain (that merges into the coiled-coil linker) results in a constitutively inactive PAL variant, as does the reciprocal D151L mutation (see Figure 45.c). The two LOV domain-lacking [PA]-variants turned out to be constitutively inactive. However, this observed effect might be due to a reduced protein stability. Surprisingly, the mutation of resi R353, E355 and Q358 within the Jα helix does not seem to have any impact on the functioning of PAL, even though all of them are involved in polar contacts with either the LOV domain core (R353), helix 12α of the ANTAR domain (E355) or the long flexible linker (Q358) (see Figure 45.d). However, their mutation does not seem to have any negative effect on the propagation of the light stimulus as the resulting PAL variants behaved similar to wild-type PAL in the Ribogreen assay. Based on the current data, one can only speculate on the concrete mechanisms of light-induced conformational changes. Yet, the polar contact between D151 and R195 seems to be essential; the light-induced change of the H-bond network must therefore extend to this point via a path still to be clarified. The central position of the Jα helix suggests an important role as transmitting element. A likely candidate for an involved transmitter residue is R356, which is situated proximal to D151. Since no suitable high-throughput screening assay was available until recently (but has now been successfully established by Jennifer Kaiser), the testing of individual mutants was very time-consuming, so that only few selected mutant constructs could be tested.

Outlook

The clarification of the exact mechanisms of light-induced PAL activation will require the continuation of mechanistic and structural investigations. Among the possible follow-up experiments, the co-crystallization of PAL in its light-adopted state together with its RNA target substrate would be a major challenge. The already made progress in determining the minimal required motif of the target substrate (19 nt vs. 80 nt at the beginning of the experiments) will certainly be of advantage for that purpose. Moreover, the already established production of milligram quantities of fairly pure PAL protein permits the screening of suitable crystallization conditions for the light-activated state in sparse-matrix screens. The photoinduced activation of photoreceptors often requires tertiary and quaternary structural rearrangements. Due to crystal lattice limitations, these can only be triggered to a limited extent by the illumination of protein crystals at the dark-adapted state. To detect the full extent of light-induced structural changes, light crystals should be grown under continuous blue light illumination. However, the increase in structural dynamics and aggregation tendency under light-conditions as well as potential photobleeching complicate the crystallization of photoreceptors in their signaling state. For that reason, a successful crystallization of light-adapted photoreceptor domains has so far only been achieved only for the isolated LOV domains of Vivid [186]and PpsB1 [187], which belong to the group of so-called 'short LOV domains' that lack a covalently associated effector module, as well as for the isolated PtAur1a LOV domain [68,180]. In addition, Takala et al. were able to succeed in the crystallization of a truncated phytochrome construct that comprises the complete photosensor-core module, consisting of three individual domains, in its inactive and its signaling state [188] . All of these photoreceptors either feature a slow return to the dark state or, as in the case of Vivid, have been modified accordingly (the light-state dimer was obtained by employing a M135I:M165I variant [186]). Therefore, the use of the T250V PAL variant, or another decelerated photocycle mutant (see Figure 44), may be of advantage for the pursuit of light-adapted PAL crystals. The crystallization attempts of the light-adapted signaling state could be further combined with Small-angle x-ray scattering (SAXS) measurements that provide a medium-resolution view of protein structures in solution. This would permit the pursuit of light and target-association induced global conformational changes of PAL. In case of a successful crystallization of the isolated PAL-LOV domain in its light-adapted state, the x-ray data could be combined with SAXS data of the full-length PAL receptor under illumination conditions to develop a more advanced signaling model [68,189].

Further on, it would be interesting to replace the N-terminal PAS domains of PAL with different LOV domains for the dual aim of understanding its role in signaling and of obtaining optogenetic actuators with divergent light-regulated properties. Since LOV domains represent a subfamily of the PAS familiy, this exchange benefits from the conservation of the PAS core structure, as well as the

presence of the previously mentioned highly conserved DIT consensus sequence motif at the C-terminus of most PAS modules [35] that precisely delineates the domain boundary. The YtvA LOV domain of *B. subtilis* would be a particularly interesting candidate for the replacement of the N-terminal PAS domain, due to numerous mutational studies that have led to a relatively deep understanding of the structure-function relationships [21,57,155]. For example, the previously identified single mutantions D21V and H22P led to inversion of the signal polarity in YF1, converting the histidine kinase function from light-suppressed to light-activated. If these results could be transferred to the PAL-derived fusion constructs, this would greatly extend the optogenetic repertoire of RNA-binding, light-regulated actuators. To assess whether the additional N-terminal LOV domains actually contribute to light regulation, the original C-terminal PAL-LOV domain could be shut down by mutagenesis, for example by the already employed C284A point mutation [158], although caution is advised with increased light intensities.

6.2.5 Optogenetic potential

So far there is no example of an optogenetic tool that mediates control over RNA-associated processes. In addition to the crucial role of RNA in the translation of genetic information to the protein level, many other cell functions have been demonstrated for the category of ncRNAs in recent years that go far beyond the long-known functions of messenger RNA, ribosomal RNA and transfer RNA. For many of these RNA molecules, the natural function is not yet well understood, so that an RNA-binding optogenetic actuator could be of great benefit for the ongoing research in the future. The results so far indicate a great potential of PAL for use in optogenetics, as they afford the hitherto unavailable possibility of generating light-dependent RNA-protein interactions with high affinity. With the structural and mechanistic studies of PAL, we were able to gain first important insights into the signaling mechanisms of this unusual sensor-effector architecture, which will surely prove to be useful for the development of further, orthogonal optogenetic tools. In RNA-aptamer applications, RNA aptamers are used to control diverse biological functions, as well as for the construction of cellular logic gates [190–193]. Since the unique architecture of PAL enables light-control over the RNA binding function, it should be possible to generate light-dependent modules for use in cell culture (efforts in this regard have succeeded by our collaboration partners from the Mayer group at the University of Bonn) and *in vivo*. If such light-regulated aptames would be combined with ncRNA regions of unknown function, their role could be scanned by adept optogenetic experiments. By combining the PAL-binding aptamers with sequences from untranslated mRNA regions, it should further be possible to achieve light-control over the specific process of translation in unprecedented manner.

7. Bibliography

0. http://www.bbc.com/news/science-environment-35344664

1. Hegemann P. Algal sensory photoreceptors. Annu Rev Plant Biol. 2008;59: 167–189. doi:10.1146/annurev.arplant.59.032607.092847

2. Deisseroth K, Feng G, Majewska AK, Miesenböck G, Ting A, Schnitzer MJ. Next-generation optical technologies for illuminating genetically targeted brain circuits. J Neurosci Off J Soc Neurosci. 2006;26: 10380–10386. doi:10.1523/JNEUROSCI.3863-06.2006

3. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2016;44: D7–D19. doi:10.1093/nar/gkv1290

4. Möglich A, Moffat K. Engineered photoreceptors as novel optogenetic tools. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2010;9: 1286–1300. doi:10.1039/c0pp00167h

5. Schmidt D, Cho YK. Natural photoreceptors and their application to synthetic biology. Trends Biotechnol. 2015;33: 80–91. doi:10.1016/j.tibtech.2014.10.007

6. Ziegler T, Möglich A. Photoreceptor engineering. Biophysics. 2015;2: 30. doi:10.3389/fmolb.2015.00030

7. Möglich A, Yang X, Ayers RA, Moffat K. Structure and function of plant photoreceptors. Annu Rev Plant Biol. 2010;61: 21–47. doi:10.1146/annurev-arplant-042809-112259

8. Brown BA, Cloix C, Jiang GH, Kaiserli E, Herzyk P, Kliebenstein DJ, et al. A UV-B-specific signaling component orchestrates plant UV protection. Proc Natl Acad Sci U S A. 2005;102: 18225–18230. doi:10.1073/pnas.0507187102

9. Conrad KS, Manahan CC, Crane BR. Photochemistry of flavoprotein light sensors. Nat Chem Biol. 2014;10: 801–809. doi:10.1038/nchembio.1633

10. Losi A, Mandalari C, Gärtner W. The Evolution and Functional Role of Flavin-based Prokaryotic Photoreceptors. Photochem Photobiol. 2015; doi:10.1111/php.12489

11. Pudasaini A, El-Arab KK, Zoltowski BD. LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. Biophysics. 2015;2: 18. doi:10.3389/fmolb.2015.00018

12. Ziegler T, Schumacher CH, Möglich A. Guidelines for Photoreceptor Engineering. Methods Mol Biol Clifton NJ. 2016;1408: 389–403. doi:10.1007/978-1-4939-3512-3_27

13. Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H. Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. Chem Rev. 2014;114: 126–163. doi:10.1021/cr4003769

14. Rockwell NC, Su Y-S, Lagarias JC. Phytochrome structure and signaling mechanisms. Annu Rev Plant Biol. 2006;57: 837–858. doi:10.1146/annurev.arplant.56.032604.144208

15. Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, Bhattacharya D, et al. Eukaryotic algal phytochromes span the visible spectrum. Proc Natl Acad Sci U S A. 2014;111: 3871–3876. doi:10.1073/pnas.1401871111

16. Ikeuchi M, Ishizuka T. Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2008;7: 1159–1167. doi:10.1039/b802660m

17. Rockwell NC, Lagarias JC. A brief history of phytochromes. Chemphyschem Eur J Chem Phys Phys Chem. 2010;11: 1172–1180. doi:10.1002/cphc.200900894

18. Tsukamoto H, Terakita A. Diversity and functional properties of bistable pigments. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2010;9: 1435–1443. doi:10.1039/c0pp00168f

19. Zhou XX, Chung HK, Lam AJ, Lin MZ. Optical Control of Protein Activity by Fluorescent Protein Domains. Science. 2012;338: 810–814. doi:10.1126/science.1226854

20. Way JC, Collins JJ, Keasling JD, Silver PA. Integrating biological redesign: where synthetic biology came from and where it needs to go. Cell. 2014;157: 151–161. doi:10.1016/j.cell.2014.02.039

21. Diensthuber RP, Bommer M, Gleichmann T, Möglich A. Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators. Structure. 2013;21: 1127–1136. doi:10.1016/j.str.2013.04.024

22. Gasser C, Taiber S, Yeh C-M, Wittig CH, Hegemann P, Ryu S, et al. Engineering of a red-light-activated human cAMP/cGMP-specific phosphodiesterase. Proc Natl Acad Sci U S A. 2014;111: 8803–8808. doi:10.1073/pnas.1321600111

23. Whippo CW, Hangarter RP. Phototropism: bending towards enlightenment. Plant Cell. 2006;18: 1110–1119. doi:10.1105/tpc.105.039669

24. Huala E, Oeller PW, Liscum E, Han IS, Larsen E, Briggs WR. Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain. Science. 1997;278: 2120–2123.

25. Briggs WR, Christie JM, Salomon M. Phototropins: a new family of flavin-binding blue light receptors in plants. Antioxid Redox Signal. 2001;3: 775–788. doi:10.1089/15230860152664975

26. Gallagher S, Short TW, Ray PM, Pratt LH, Briggs WR. Light-mediated changes in two proteins found associated with plasma membrane fractions from pea stem sections. Proc Natl Acad Sci U S A. 1988;85: 8003–8007.

27. Hager A. Properties of a blue-light-absorbing photoreceptor kinase localized in the plasma membrane of the coleoptile tip region. Planta. 1996;198: 294–299.

28. Palmer JM, Short TW, Gallagher S, Briggs WR. Blue Light-Induced Phosphorylation of a Plasma Membrane-Associated Protein in Zea mays L. Plant Physiol. 1993;102: 1211–1218.

29. Short TW, Briggs WR. Characterization of a Rapid, Blue Light-Mediated Change in Detectable Phosphorylation of a Plasma Membrane Protein from Etiolated Pea (Pisum sativum L.) Seedlings. Plant Physiol. 1990;92: 179–185.

30. Glantz ST, Carpenter EJ, Melkonian M, Gardner KH, Boyden ES, Wong GK-S, et al. Functional and topological diversity of LOV domain photoreceptors. Proc Natl Acad Sci U S A. 2016;113: E1442-1451. doi:10.1073/pnas.1509428113

31. Anantharaman V, Balaji S, Aravind L. The signaling helix: a common functional theme in diverse signaling proteins. Biol Direct. 2006;1: 25. doi:10.1186/1745-6150-1-25

32. Möglich A, Ayers RA, Moffat K. Design and signaling mechanism of light-regulated histidine kinases. J Mol Biol. 2009;385: 1433–1444. doi:10.1016/j.jmb.2008.12.017

33. Möglich A, Ayers RA, Moffat K. Addition at the molecular level: signal integration in designed Per-ARNT-Sim receptor proteins. J Mol Biol. 2010;400: 477–486. doi:10.1016/j.jmb.2010.05.019

34. Rockwell NC, Ohlendorf R, Möglich A. Cyanobacteriochromes in full color and three dimensions. Proc Natl Acad Sci U S A. 2013;110: 806–807. doi:10.1073/pnas.1220690110

35. Möglich A, Ayers RA, Moffat K. Structure and signaling mechanism of Per-ARNT-Sim domains. Struct Lond Engl 1993. 2009;17: 1282–1294. doi:10.1016/j.str.2009.08.011

36. Taylor BL, Zhulin IB. PAS domains: internal sensors of oxygen, redox potential, and light. Microbiol Mol Biol Rev. 1999;63: 479–506.

37. Zoltowski BD, Gardner KH. Tripping the Light Fantastic: Blue-Light Photoreceptors as Examples of Environmentally Modulated Protein−Protein Interactions. Biochemistry. 2011;50: 4–16. doi:10.1021/bi101665s

38. Halavaty AS, Moffat K. N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa. Biochemistry. 2007;46: 14001–14009. doi:10.1021/bi701543e

39. Nash AI, McNulty R, Shillito ME, Swartz TE, Bogomolni RA, Luecke H, et al. Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. Proc Natl Acad Sci U S A. 2011;108: 9449–9454. doi:10.1073/pnas.1100262108

40. Zoltowski BD, Crane BR. Light activation of the LOV protein vivid generates a rapidly exchanging dimer. Biochemistry. 2008;47: 7012–7019. doi:10.1021/bi8007017

41. Zoltowski BD, Schwerdtfeger C, Widom J, Loros JJ, Bilwes AM, Dunlap JC, et al. Conformational Switching in the Fungal Light Sensor Vivid. Science. 2007;316: 1054–1057. doi:10.1126/science.1137128

42. Lokhandwala J, Hopkins HC, Rodriguez-Iglesias A, Dattenböck C, Schmoll M, Zoltowski BD. Structural biochemistry of a fungal LOV domain photoreceptor reveals an evolutionarily conserved pathway integrating light and oxidative stress. Struct Lond Engl 1993. 2015;23: 116–125. doi:10.1016/j.str.2014.10.020

43. Salomon M, Christie JM, Knieb E, Lempert U, Briggs WR. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. Biochemistry. 2000;39: 9401–10.

44. Swartz TE, Corchnoy SB, Christie JM, Lewis JW, Szundi I, Briggs WR, et al. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. J Biol Chem. 2001;276: 36493–500.

45. Alexandre MT, Domratcheva T, Bonetti C, van Wilderen LJ, van Grondelle R, Groot ML, et al. Primary reactions of the LOV2 domain of phototropin studied with ultrafast mid-infrared spectroscopy and quantum chemistry. Biophys J. 2009;97: 227–37.

46. Bittl R, Kay CWM, Weber S, Hegemann P. Characterization of a flavin radical product in a C57M mutant of a LOV1 domain by electron paramagnetic resonance. Biochemistry. 2003;42: 8506–8512. doi:10.1021/bi034123i

47. Corchnoy SB, Swartz TE, Lewis JW, Szundi I, Briggs WR, Bogomolni RA. Intramolecular proton transfers and structural changes during the photocycle of the LOV2 domain of phototropin 1. J Biol Chem. 2003;278: 724–31.

48. Dittrich M, Freddolino PL, Schulten K. When light falls in LOV: a quantum mechanical/molecular mechanical study of photoexcitation in Phot-LOV1 of Chlamydomonas reinhardtii. J Phys Chem B. 2005;109: 13006–13013. doi:10.1021/jp050943o

49. Holzer W, Penzkofer A, Fuhrmann M, Hegemann P. Spectroscopic characterization of flavin mononucleotide bound to the LOV1 domain of Phot1 from Chlamydomonas reinhardtii. Photochem Photobiol. 2002;75: 479–487.

50. Iwata T, Yamamoto A, Tokutomi S, Kandori H. Hydration and temperature similarly affect light-induced protein structural changes in the chromophoric domain of phototropin. Biochemistry. 2007;46: 7016–21.

51. Kennis JT, Crosson S, Gauden M, van Stokkum IH, Moffat K, van Grondelle R. Primary reactions of the LOV2 domain of phototropin, a plant blue-light photoreceptor. Biochemistry. 2003;42: 3385–92.

52. Kennis JT, van Stokkum IH, Crosson S, Gauden M, Moffat K, van Grondelle R. The LOV2 domain of phototropin: a reversible photochromic switch. J Am Chem Soc. 2004;126: 4512–3.

53. Kottke T, Heberle J, Hehn D, Dick B, Hegemann P. Phot-LOV1: photocycle of a blue-light receptor domain from the green alga Chlamydomonas reinhardtii. Biophys J. 2003;84: 1192–1201.

54. Nihongaki Y, Suzuki H, Kawano F, Sato M. Genetically engineered photoinducible homodimerization system with improved dimer-forming efficiency. ACS Chem Biol. 2014;9: 617–621. doi:10.1021/cb400836k

55. Schleicher E, Kowalczyk RM, Kay CW, Hegemann P, Bacher A, Fischer M, et al. On the reaction mechanism of adduct formation in LOV domains of the plant blue-light receptor phototropin. J Am Chem Soc. 2004;126: 11067–76.

56. Zoltowski BD, Vaccaro B, Crane BR. Mechanism-based tuning of a LOV domain photoreceptor. Nat Chem Biol. 2009;5: 827–834. doi:10.1038/nchembio.210

57. Diensthuber RP, Engelhard C, Lemke N, Gleichmann T, Ohlendorf R, Bittl R, et al. Biophysical, mutational, and functional investigation of the chromophore-binding pocket of light-oxygen-voltage photoreceptors. ACS Synth Biol. 2014;3: 811–819. doi:10.1021/sb400205x

58. Okajima K, Kashojiya S, Tokutomi S. Photosensitivity of kinase activation by blue light involves the lifetime of a cysteinyl-flavin adduct intermediate, S390, in the photoreaction cycle of the LOV2 domain in phototropin, a plant blue light receptor. J Biol Chem. 2012;287: 40972–40981. doi:10.1074/jbc.M112.406512

59. Domratcheva T, Fedorov R, Schlichting I. Analysis of the Primary Photocycle Reactions Occurring in the Light, Oxygen, and Voltage Blue-Light Receptor by Multiconfigurational Quantum-Chemical Methods. J Chem Theory Comput. 2006;2: 1565–1574. doi:10.1021/ct0600114

60. Yamamoto A, Iwata T, Tokutomi S, Kandori H. Role of Phe1010 in light-induced structural changes of the neo1-LOV2 domain of Adiantum. Biochemistry. 2008;47: 922–928. doi:10.1021/bi701851v

61. Freddolino PL, Gardner KH, Schulten K. Signaling mechanisms of LOV domains: new insights from molecular dynamics studies. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2013;12: 1158–1170. doi:10.1039/c3pp25400c

62. Harper SM, Neil LC, Gardner KH. Structural basis of a phototropin light switch. Science. 2003;301: 1541–1544. doi:10.1126/science.1086810

63. Harper SM, Christie JM, Gardner KH. Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. Biochemistry. 2004;43: 16184–16192.

64. Nash AI, Ko W-H, Harper SM, Gardner KH. A conserved glutamine plays a central role in LOV domain signal transmission and its duration. Biochemistry. 2008;47: 13842–13849. doi:10.1021/bi801430e

65. Engelhard C, Raffelberg S, Tang Y, Diensthuber RP, Möglich A, Losi A, et al. A structural model for the full-length blue light-sensing protein YtvA from Bacillus subtilis, based on EPR spectroscopy. Photochem Photobiol Sci. 2013;12: 1855–1863. doi:10.1039/c3pp50128k

66. Tang Y, Cao Z, Livoti E, Krauss U, Jaeger K-E, Gärtner W, et al. Interdomain signalling in the blue-light sensing and GTP-binding protein YtvA: a mutagenesis study uncovering the importance of specific protein sites. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2010;9: 47–56. doi:10.1039/b9pp00075e

67. Takahashi F, Yamagata D, Ishikawa M, Fukamatsu Y, Ogura Y, Kasahara M, et al. AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. Proc Natl Acad Sci U A. 2007;104: 19625–19630.

68. Heintz U, Schlichting I. Blue light-induced LOV domain dimerization enhances the affinity of Aureochrome 1a for its target DNA sequence. eLife. 2016;5: e11860. doi:10.7554/eLife.11860

69. Herman E, Kottke T. Allosterically regulated unfolding of the A'α helix exposes the dimerization site of the blue-light-sensing aureochrome-LOV domain. Biochemistry. 2015;54: 1484–1492. doi:10.1021/bi501509z

70. Lamb JS, Zoltowski BD, Pabit SA, Crane BR, Pollack L. Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering. J Am Chem Soc. 2008;130: 12226–12227. doi:10.1021/ja804236f

71. Taslimi A, Zoltowski B, Miranda JG, Pathak G, Hughes RM, Tucker CL. Optimized second generation CRY2/CIB dimerizers and photoactivatable Cre recombinase. Nat Chem Biol. 2016;12: 425–430. doi:10.1038/nchembio.2063

72. Levskaya A, Weiner OD, Lim WA, Voigt CA. Spatiotemporal control of cell signalling using a light-switchable protein interaction. Nature. 2009;461: 997–1001.

73. Shimizu-Sato S, Huq E, Tepperman JM, Quail PH. A light-switchable gene promoter system. Nat Biotechnol. 2002;20: 1041–1044. doi:10.1038/nbt734

74. Rivera-Cancel G, Ko W, Tomchick DR, Correa F, Gardner KH. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. Proc Natl Acad Sci U S A. 2014;111: 17839–17844. doi:10.1073/pnas.1413983111

75. Harper SM, Neil LC, Gardner KH. Structural basis of a phototropin light switch. Science. 2003;301: 1541–1544. doi:10.1126/science.1086810

76. Christie JM, Blackwood L, Petersen J, Sullivan S. Plant flavoprotein photoreceptors. Plant Cell Physiol. 2015;56: 401–413. doi:10.1093/pcp/pcu196

77. Wu YI, Frey D, Lungu OI, Jaehrig A, Schlichting I, Kuhlman B, et al. A genetically encoded photoactivatable Rac controls the motility of living cells. Nature. 2009;461: 104–108. doi:10.1038/nature08241

78. Herman E, Kottke T. Allosterically regulated unfolding of the A'α helix exposes the dimerization site of the blue-light-sensing aureochrome-LOV domain. Biochemistry. 2015;54: 1484–1492. doi:10.1021/bi501509z

79. Conrad KS, Bilwes AM, Crane BR. Light-induced subunit dissociation by a light-oxygen-voltage domain photoreceptor from Rhodobacter sphaeroides. Biochemistry. 2013;52: 378–391. doi:10.1021/bi3015373

80. Yazawa M, Sadaghiani AM, Hsueh B, Dolmetsch RE. Induction of protein-protein interactions in live cells using light. Nat Biotechnol. 2009;27: 941–945. doi:10.1038/nbt.1569

81. Wang X, Chen X, Yang Y. Spatiotemporal control of gene expression by a light-switchable transgene system. Nat Methods. 2012;9: 266–269. doi:10.1038/nmeth.1892

82. Nihongaki Y, Suzuki H, Kawano F, Sato M. Genetically engineered photoinducible homodimerization system with improved dimer-forming efficiency. ACS Chem Biol. 2014;9: 617–621. doi:10.1021/cb400836k

83. Grusch M, Schelch K, Riedler R, Reichhart E, Differ C, Berger W, et al. Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. EMBO J. 2014;33: 1713–1726. doi:10.15252/embj.201387695

84. Chang K-Y, Woo D, Jung H, Lee S, Kim S, Won J, et al. Light-inducible receptor tyrosine kinases that regulate neurotrophin signalling. Nat Commun. 2014;5: 4057. doi:10.1038/ncomms5057

85. Grusch M, Schelch K, Riedler R, Reichhart E, Differ C, Berger W, et al. Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. EMBO J. 2014;33: 1713–1726. doi:10.15252/embj.201387695

86. Kim N, Kim JM, Lee M, Kim CY, Chang K-Y, Heo WD. Spatiotemporal control of fibroblast growth factor receptor signals by blue light. Chem Biol. 2014;21: 903–912. doi:10.1016/j.chembiol.2014.05.013

87. Möglich A, Ayers RA, Moffat K. Design and signaling mechanism of light-regulated histidine kinases. J Mol Biol. 2009;385: 1433–1444. doi:10.1016/j.jmb.2008.12.017

88. Ohlendorf R, Schumacher CH, Richter F, Möglich A. Library-Aided Probing of Linker Determinants in Hybrid Photoreceptors. ACS Synth Biol. 2016;5: 1117–1126. doi:10.1021/acssynbio.6b00028

89. Ryu M-H, Kang I-H, Nelson MD, Jensen TM, Lyuksyutova AI, Siltberg-Liberles J, et al. Engineering adenylate cyclases regulated by near-infrared window light. Proc Natl Acad Sci U S A. 2014;111: 10167–10172. doi:10.1073/pnas.1324301111

90. Stierl M, Penzkofer A, Kennis JTM, Hegemann P, Mathes T. Key residues for the light regulation of the blue light-activated adenylyl cyclase from Beggiatoa sp. Biochemistry. 2014;53: 5121–5130. doi:10.1021/bi500479v

91. Jansen V, Alvarez L, Balbach M, Strünker T, Hegemann P, Kaupp UB, et al. Controlling fertilization and cAMP signaling in sperm by optogenetics. eLife. 2015;4. doi:10.7554/eLife.05161

92. Schumacher CH, Körschen HG, Nicol C, Gasser C, Seifert R, Schwärzel M, et al. A Fluorometric Activity Assay for Light-Regulated Cyclic-Nucleotide-Monophosphate Actuators. Methods Mol Biol Clifton NJ. 2016;1408: 93–105. doi:10.1007/978-1-4939-3512-3_7

93. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet. 2013;9: e1003569. doi:10.1371/journal.pgen.1003569

94. Mattick JS. The central role of RNA in human development and cognition. FEBS Lett. 2011;585: 1600–1616. doi:10.1016/j.febslet.2011.05.001

95. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009;136: 615–628. doi:10.1016/j.cell.2009.01.043

96. Chai W, Stewart V. NasR, a novel RNA-binding protein, mediates nitrate-responsive transcription antitermination of the Klebsiella oxytoca M5al nasF operon leader in vitro. J Mol Biol. 1998;283: 339–351. doi:10.1006/jmbi.1998.2105

97. Drew R, Lowe N. Positive control of Pseudomonas aeruginosa amidase synthesis is mediated by a transcription anti-termination mechanism. J Gen Microbiol. 1989;135: 817–823. doi:10.1099/00221287-135-4-817

98. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44: D279-285. doi:10.1093/nar/gkv1344

99. Mascher T, Helmann JD, Unden G. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol Rev. 2006;70: 910–938.

100. Ramesh A, DebRoy S, Goodson JR, Fox KA, Faz H, Garsin DA, et al. The Mechanism for RNA Recognition by ANTAR Regulators of Gene Expression. PLOS Genet. 2012;8: e1002666. doi:10.1371/journal.pgen.1002666

101. Morth JP, Feng V, Perry LJ, Svergun DI, Tucker PA. The Crystal and Solution Structure of a Putative Transcriptional Antiterminator from Mycobacterium tuberculosis. Structure. 2004;12: 1595–1605. doi:10.1016/j.str.2004.06.018

102. O'Hara BP, Norman RA, Wan PTC, Roe SM, Barrett TE, Drew RE, et al. Crystal structure and induction mechanism of AmiC–AmiR: a ligand-regulated transcription antitermination complex. EMBO J. 1999;18: 5175–5186. doi:10.1093/emboj/18.19.5175

103. Norman RA, Poh CL, Pearl LH, O'Hara BP, Drew RE. Steric hindrance regulation of the Pseudomonas aeruginosa amidase operon. J Biol Chem. 2000;275: 30660–30667. doi:10.1074/jbc.M000813200

104. O'Hara BP, Norman RA, Wan PT, Roe SM, Barrett TE, Drew RE, et al. Crystal structure and induction mechanism of AmiC-AmiR: a ligand-regulated transcription antitermination complex. EMBO J. 1999;18: 5175–5186. doi:10.1093/emboj/18.19.5175

105. Boudes M, Lazar N, Graille M, Durand D, Gaidenko TA, Stewart V, et al. The structure of the NasR transcription antiterminator reveals a one-component system with a NIT nitrate receptor coupled to an ANTAR RNA-binding effector. Mol Microbiol. 2012;85: 431–444. doi:10.1111/j.1365-2958.2012.08111.x

106. Shu CJ, Zhulin IB. ANTAR: an RNA-binding domain in transcription antitermination regulatory proteins. Trends Biochem Sci. 2002;27: 3–5. doi:10.1016/S0968-0004(01)02036-9

107. Shu CJ, Ulrich LE, Zhulin IB. The NIT domain: a predicted nitrate-responsive module in bacterial sensory receptors. Trends Biochem Sci. 2003;28: 121–124. doi:10.1016/S0968-0004(03)00032-X

108. Peters JM, Vangeloff AD, Landick R. Bacterial Transcription Terminators: The RNA 3'-End Chronicles. J Mol Biol. 2011;412: 793–813. doi:10.1016/j.jmb.2011.03.036

109. Stewart V, van Tilbeurgh H. Found: the elusive ANTAR transcription antiterminator. PLoS Genet. 2012;8: e1002773. doi:10.1371/journal.pgen.1002773

110. Wilson SA, Wachira SJ, Norman RA, Pearl LH, Drew RE. Transcription antitermination regulation of the Pseudomonas aeruginosa amidase operon. EMBO J. 1996;15: 5907–5916.

111. DebRoy S, Gebbie M, Ramesh A, Goodson JR, Cruz MR, Hoof A van, et al. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. Science. 2014;345: 937. doi:10.1126/science.1255091

112. Mellin JR, Koutero M, Dar D, Nahori M-A, Sorek R, Cossart P. Riboswitches. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. Science. 2014;345: 940–943. doi:10.1126/science.1255083

113. Cech TR, Steitz JA. The noncoding RNA revolution-trashing old rules to forge new ones. Cell. 2014;157: 77–94. doi:10.1016/j.cell.2014.03.008

114. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6: 343–345. doi:10.1038/nmeth.1318

115. Beta-Galactosidase Assay (A better Miller) - OpenWetWare [Internet]. [cited 28 Mar 2017]. Available: http://www.openwetware.org/index.php?title=Beta-Galactosidase_Assay_%28A_better_Miller%29&oldid=620416

116. Zhang X, Bremer H. Control of the Escherichia coli rrnB P1 promoter strength by ppGpp. J Biol Chem. 1995;270: 11181–9.

117. Meyer TS, Lamberts BL. Use of coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide-gel strips. Biochim Biophys Acta. 1965;107: 144–145.

118. Garner MM, Revzin A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. 1981;9: 3047–3060.

119. Stoltenburg R, Reinemann C, Strehlitz B. SELEX—A (r)evolutionary method to generate high-affinity nucleic acid ligands. Biomol Eng. 2007;24: 381–403. doi:10.1016/j.bioeng.2007.06.001

120. Folta-Stogniew E. Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors. Methods Mol Biol Clifton NJ. 2006;328: 97–112. doi:10.1385/1-59745-026-X:97

121. Su X-D, Zhang H, Terwilliger TC, Liljas A, Xiao J, Dong Y. Protein Crystallography from the Perspective of Technology Developments. Crystallogr Rev. 2015;21: 122–153. doi:10.1080/0889311X.2014.973868

122. Krug M, Weiss MS, Heinemann U, Mueller U. XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. J Appl Crystallogr. 2012;45: 568–572. doi:10.1107/S0021889812011715

123. Sheldrick GM. A short history of SHELX. Acta Crystallogr A. 2008;64: 112–122. doi:10.1107/S0108767307043930

124. Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr. 2008;64: 61–69. doi:10.1107/S090744490705024X

125. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006;62: 1002–1011. doi:10.1107/S0907444906022116

126. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012;68: 352–367. doi:10.1107/S0907444912001308

127. The PyMOL Molecular Graphics System [Internet]. San Carlos, CA, USA; 2002. Available: http://www.pymol.org

128. Henry J, Crosson S. Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context. Annu Rev Microbiol. 2011;65: 261–286. doi:10.1146/annurev-micro-121809-151631

129. Surette MG, Stock JB. Role of alpha-helical coiled-coil interactions in receptor dimerization, signaling, and adaptation during bacterial chemotaxis. J Biol Chem. 1996;271: 17966–17973.

130. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet TIG. 2000;16: 276–277.

131. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinforma Oxf Engl. 2007;23: 2947–2948. doi:10.1093/bioinformatics/btm404

132. Möglich A, Moffat K. Engineered photoreceptors as novel optogenetic tools. Photochem Photobiol Sci Off J Eur Photochem Assoc Eur Soc Photobiol. 2010;9: 1286–1300. doi:10.1039/c0pp00167h

133. Deuschle U, Kammerer W, Gentz R, Bujard H. Promoters of Escherichia coli: a hierarchy of in vivo strength indicates alternate structures. EMBO J. 1986;5: 2987–2994.

134. Schleif R. AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. FEMS Microbiol Rev. 2010;34: 779–796. doi:10.1111/j.1574-6976.2010.00226.x

135. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42: D222. doi:10.1093/nar/gkt1223

136. Consortium TU. The universal protein resource (UniProt). Nucleic Acids Res. 2008;36: D190-5.

137. Möglich A, Moffat K. Structural Basis for Light-dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA. J Mol Biol. 2007;373: 112–126.

138. Lamb JS, Zoltowski BD, Pabit SA, Crane BR, Pollack L. Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering. J Am Chem Soc. 2008;130: 12226–12227. doi:10.1021/ja804236f

139. Pudasaini A, El-Arab KK, Zoltowski BD. LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. Biophysics. 2015;2: 18. doi:10.3389/fmolb.2015.00018

140. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. 2009;37: W202. doi:10.1093/nar/gkp335

141. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol AMB. 2011;6: 26. doi:10.1186/1748-7188-6-26

142. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31: 3406–3415.

143. Hellman LM, Fried MG. Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions. Nat Protoc. 2007;2: 1849–1861. doi:10.1038/nprot.2007.249

144. Alexandre MT, Arents JC, van Grondelle R, Hellingwerf KJ, Kennis JT. A base-catalyzed mechanism for dark state recovery in the Avena sativa phototropin-1 LOV2 domain. Biochemistry. 2007;46: 3129–3137.

145. Walden H. Selenium incorporation using recombinant techniques. Acta Crystallogr D Biol Crystallogr. 2010;66: 352–357. doi:10.1107/S0907444909038207

146. Doublié S. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. 1997;276: 523–530.

147. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60: 2126–2132. doi:10.1107/S0907444904019158

148. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66: 486–501. doi:10.1107/S0907444910007493

149. Crosson S, Moffat K. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. Proc Natl Acad Sci U A. 2001;98: 2995–3000.

150. Hulme EC, Trevethick MA. Ligand binding assays at equilibrium: validation and interpretation. Br J Pharmacol. 2010;161: 1219–1237. doi:10.1111/j.1476-5381.2009.00604.x

151. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 2012;40: W597-603. doi:10.1093/nar/gks400

152. Smale ST. Beta-galactosidase assay. Cold Spring Harb Protoc. 2010;2010: pdb.prot5423. doi:10.1101/pdb.prot5423

153. Ohlendorf R, Vidavski RR, Eldar A, Moffat K, Möglich A. From dusk till dawn: one-plasmid systems for light-regulated gene expression. J Mol Biol. 2012;416: 534–542. doi:10.1016/j.jmb.2012.01.001

154. Garcia HG, Lee HJ, Boedicker JQ, Phillips R. Comparison and Calibration of Different Reporters for Quantitative Analysis of Gene Expression. Biophys J. 2011;101: 535–544. doi:10.1016/j.bpj.2011.06.026

155. Gleichmann T, Diensthuber RP, Möglich A. Charting the Signal Trajectory in a Light-Oxygen-Voltage Photoreceptor by Random Mutagenesis and Covariance Analysis. J Biol Chem. 2013;288: 29345–29355. doi:10.1074/jbc.M113.506139

156. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The Shape and Structure of Proteins. 2002; Available: https://www.ncbi.nlm.nih.gov/books/NBK26830/

157. Chambers P, Pringle CR, Easton AJ. Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. J Gen Virol. 1990;71 (Pt 12): 3075–3080. doi:10.1099/0022-1317-71-12-3075

158. Yee EF, Diensthuber RP, Vaidya AT, Borbat PP, Engelhard C, Freed JH, et al. Signal transduction in light–oxygen–voltage receptors lacking the adduct-forming cysteine residue. Nat Commun. 2015;6. doi:10.1038/ncomms10079

159. Miyazaki K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. Methods Enzymol. 2011;498: 399–406. doi:10.1016/B978-0-12-385120-8.00017-6

160. Losi A, Gardner KH, Möglich A. Blue-Light Receptors for Optogenetics. Chem Rev. 2018; doi:10.1021/acs.chemrev.8b00163

161. Iseki M, Matsunaga S, Murakami A, Ohno K, Shiga K, Yoshida K, et al. A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis. Nature. 2002;415: 1047–1051.

162. Ryu M-H, Moskvin OV, Siltberg-Liberles J, Gomelsky M. Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. J Biol Chem. 2010;285: 41501–41508. doi:10.1074/jbc.M110.177600

163. Schröder-Lang S, Schwärzel M, Seifert R, Strünker T, Kateriya S, Looser J, et al. Fast manipulation of cellular cAMP level by light in vivo. Nat Methods. 2007;4: 39–42. doi:10.1038/nmeth975

164. Motta-Mena LB, Reade A, Mallory MJ, Glantz S, Weiner OD, Lynch KW, et al. An optogenetic gene expression system with rapid activation and deactivation kinetics. Nat Chem Biol. 2014;10: 196–202. doi:10.1038/nchembio.1430

165. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science. 1990;249: 505–510. doi:10.1126/science.2200121

166. Woodbury CP, von Hippel PH. On the determination of deoxyribonucleic acid-protein interaction parameters using the nitrocellulose filter-binding assay. Biochemistry. 1983;22: 4730–4737.

167. Forwood JK, Jans DA. Quantitative analysis of DNA-protein interactions using double-labeled native gel electrophoresis and fluorescence-based imaging. Electrophoresis. 2006;27: 3166–3170. doi:10.1002/elps.200500872

168. Li Y, Jiang Z, Chen H, Ma W-J. A modified quantitative EMSA and its application in the study of RNA--protein interactions. J Biochem Biophys Methods. 2004;60: 85–96. doi:10.1016/j.jbbm.2004.03.008

169. Baker KA, Perego M. Transcription antitermination by a phosphorylated response regulator and cobalamin-dependent termination at a B_{12} riboswitch contribute to ethanolamine utilization in Enterococcus faecalis. J Bacteriol. 2011;193: 2575–2586. doi:10.1128/JB.00217-11

170. YOSHIMI Y, HIRAISHI A, NAKAMURA K. Isolation and Characterization of Microsphaera multipartita gen. nov., sp. nov., a Polysaccharide-Accumulating Gram-Positive Bacterium from Activated Sludge. Int J Syst Evol Microbiol. 1996;46: 519–525. doi:10.1099/00207713-46-2-519

171. Tao T-S, Yue Y-Y, Chen W-X, Chen W-F. Proposal of Nakamurella gen. nov. as a substitute for the bacterial genus Microsphaera Yoshimi et al. 1996 and Nakamurellaceae fam. nov. as a substitute for the illegitimate bacterial family Microsphaeraceae Rainey et al. 1997. Int J Syst Evol Microbiol. 2004;54: 999–1000. doi:10.1099/ijs.0.02933-0

172. Kim KK, Lee J-S. The Family <Emphasis Type="Italic">Nakamurellaceae</Emphasis>. The Prokaryotes. Springer, Berlin, Heidelberg; 2014. pp. 577–585. doi:10.1007/978-3-642-30138-4_278

173. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25: 3389–402.

174. Tice H, Mayilraj S, Sims D, Lapidus A, Nolan M, Lucas S, et al. Complete genome sequence of Nakamurella multipartita type strain (Y-104). Stand Genomic Sci. 2010;2: 168–175. doi:10.4056/sigs.721316

175. Nouioui I, Göker M, Carro L, Montero-Calasanz M del C, Rohde M, Woyke T, et al. High quality draft genome of Nakamurella lactea type strain, a rock actinobacterium, and emended description of Nakamurella lactea. Stand Genomic Sci. 2017;12. doi:10.1186/s40793-016-0216-0

176. Van Brederode ME, Hoff WD, Van Stokkum IH, Groot ML, Hellingwerf KJ. Protein folding thermodynamics applied to the photocycle of the photoactive yellow protein. Biophys J. 1996;71: 365–380. doi:10.1016/S0006-3495(96)79234-2

177. Zoltowski BD, Nash AI, Gardner KH. Variations in protein-flavin hydrogen bonding in a light, oxygen, voltage domain produce non-Arrhenius kinetics of adduct decay. Biochemistry. 2011;50: 8771–8779. doi:10.1021/bi200976a

178. Kawano F, Aono Y, Suzuki H, Sato M. Fluorescence Imaging-Based High-Throughput Screening of Fast- and Slow-Cycling LOV Proteins. PLoS ONE. 2013;8: e82693. doi:10.1371/journal.pone.0082693

179. Raffelberg S, Mansurova M, Gärtner W, Losi A. Modulation of the Photocycle of a LOV Domain Photoreceptor by the Hydrogen-Bonding Network. J Am Chem Soc. 2011;133: 5346–5356. doi:10.1021/ja1097379

180. Banerjee A, Herman E, Kottke T, Essen L-O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from Phaeodactylum tricornutum. Struct Lond Engl 1993. 2016;24: 171–178. doi:10.1016/j.str.2015.10.022

181. Zayner JP, Antoniou C, Sosnick TR. The amino-terminal helix modulates light-activated conformational changes in AsLOV2. J Mol Biol. 2012;419: 61–74. doi:10.1016/j.jmb.2012.02.037

182. Fasman GD, editor. Circular Dichroism and the Conformational Analysis of Biomolecules [Internet]. Springer US; 1996. Available: //www.springer.com/de/book/9780306451423

183. Hisatomi O, Takeuchi K, Zikihara K, Ookubo Y, Nakatani Y, Takahashi F, et al. Blue light-induced conformational changes in a light-regulated transcription factor, aureochrome-1. Plant Cell Physiol. 2013;54: 93–106. doi:10.1093/pcp/pcs160

184. Ayers RA, Moffat K. Changes in quaternary structure in the signaling mechanisms of PAS domains. Biochemistry. 2008;47: 12078–12086.

185. Engelhard C, Diensthuber RP, Möglich A, Bittl R. Blue-light reception through quaternary transitions. Sci Rep. 2017;7: 1385. doi:10.1038/s41598-017-01497-7

186. Vaidya AT, Chen C-H, Dunlap JC, Loros JJ, Crane BR. Structure of a light-activated LOV protein dimer that regulates transcription. Sci Signal. 2011;4: ra50. doi:10.1126/scisignal.2001945

187. Circolone F, Granzin J, Jentzsch K, Drepper T, Jaeger K-E, Willbold D, et al. Structural Basis for the Slow Dark Recovery of a Full-Length LOV Protein from Pseudomonas putida. J Mol Biol. 2012;417: 362–374. doi:10.1016/j.jmb.2012.01.056

188. Takala H, Björling A, Berntsson O, Lehtivuori H, Niebling S, Hoernke M, et al. Signal amplification and transduction in phytochrome photosensors. Nature. 2014;509: 245–248. doi:10.1038/nature13310

189. Berntsson O, Diensthuber RP, Panman MR, Björling A, Hughes AJ, Henry L, et al. Time-Resolved X-Ray Solution Scattering Reveals the Structural Photoactivation of a Light-Oxygen-Voltage Photoreceptor. Struct Lond Engl 1993. 2017;25: 933-938.e3. doi:10.1016/j.str.2017.04.006

190. Ausländer D, Wieland M, Ausländer S, Tigges M, Fussenegger M. Rational design of a small molecule-responsive intramer controlling transgene expression in mammalian cells. Nucleic Acids Res. 2011;39: e155. doi:10.1093/nar/gkr829

191. Ausländer S, Ausländer D, Müller M, Wieland M, Fussenegger M. Programmable single-cell mammalian biocomputers. Nature. 2012;487: 123–127. doi:10.1038/nature11149

192. Beilstein K, Wittmann A, Grez M, Suess B. Conditional control of mammalian gene expression by tetracycline-dependent hammerhead ribozymes. ACS Synth Biol. 2015;4: 526–534. doi:10.1021/sb500270h

193. Culler SJ, Hoff KG, Smolke CD. Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. Science. 2010;330: 1251–1255. doi:10.1126/science.1192128

8. Appendix

8.1. Abbreviations

Abbreviation	Description
Å	Ångström
aa	amino acid
ami-lead	leader sequence of the AmiR operon
Amp	ampicilin
Ara	arabinose
ANTAR	AmiR and NasR transcription antitermination regulator domain
AsLOV2	phototropin 1 LOV2 domain from *Avena sativa*
ATP	adenosine trisphosphate
Aur1a	blue-light-sensitive transcription factor from *Vaucheria frigida*
A_{xxx}	absorbance at xxx nm
β–gal	β–galactosidase
BLUF	sensors of blue light using flavin adenine dinucleotide
bp	basepair
bZIP	basic leucine zipper domain
cAMP	cyclic adenosine monophosphate
Chl	chloramphenicol
CIB1	cryptochrome 2-interacting basic helix-loop-helix 1
C-terminal/ C-terminus	carboxy terminus
D	Dark-adapted state
DEPC	Diethyl pyrocarbonate
DHp	dimerization and histidine phosphotransfer
DIT	aspartate-isoleucine-threonine
DNA	deoxyribonucleic acid
dNTP	deooxynucleotide
EMSA	electrophoretic mobility shift assay
EL222	light-regulated transcription factor from
FACS	fluorescence-activated cell sorting
FAD	flavin-adenine dinucleotide
FMN	flavin mononucleotide
GAF	cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA
GTP	guanosine trisphosphate
H-bond	hydrogen bond
HCl	hydrogen chloride
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HPLC	high-performance liquid chromatography
ICB	intracellular buffer
IP	immunoprecipitation
IPTG	isopropyl-β-d-thiogalactopyranoside
Kan	kanamycin
kb	kilo bases
kDa	kilo Dalton

LB	lysogeny broth
LED	light-emitting diode
LOV	light oxygen voltage
M	molar (mol L-1)
$MgCl_2$	magnesium chloride
min	minute
MU	Miller unit
NaCl	sodium chloride
nt	nucleotide
N-terminal / N-terminus	amino terminal / terminus
OD	optical density
OD_{xxx}	optical density at xxx nm
ONPG	ortho-nitrophenyl-β-galactosid
PAGE	poly-acrylamide gel electrophoresis
PAS	Per-ARNT-Sim
PCR	polymerase chain reaction
PDE	phosphodiesterase
PEG	polyethylene glycol
PtAur1a	Aur1a homologue from *Phaeodactylum tricornutum*
PIF	Phytochrome interacting partner
PYP	photoactive yellow protein
RBS	ribosome binding site
RIP	RNA immunoprecipitation
resi	residue
RR	response regulator
RT	Reverse transcription
S	signaling state
SD	standard deviation
SDS	sodium dodecyl sulfate
sec	second
SELEX	Systematic evolution of ligands by exponential enrichment
SEM	standard deviation
STAS	sulfate transporter/anti-sigma-factor antagonist
TBE	Tris-borate-EDTA
Tris	Tris(hydroxymethyl)aminomethane
UV	ultra violet
UVR8 U	UV-B resistance 8
w/	with
ε	extinction coefficient
> / <	smaller / larger than

Table A1: Abbreviations from Figure 3 (extract from Glantz et al. PNAS 2016;113:E1442-E1451).

Abbreviated Name	Full Name(s)	Conserved Domain ID(s)
ANTAR	AmiR and NasR transcription antitermination regulator domain	PF03861
bHTH	Helix-turn-helix_10 domain	PF04967
bZIP	bZIP transcription factor and Basic region leucine zipper domains	PF00170, PF07716
CheY	CheY-like superfamily domain	IPR011006
Cyc	Adenylate and Guanylate cyclase catalytic domain	PF00211
Endoribonuclease	Endoribonuclease L-PSP	PF01042
F-box	F-box, F-box-like domains	PF00646, PF12937
GAF	GAF domain	PF01590, PF13185, PF13492
GATase	Glutamine amidotransferase clase II domain	PF13537
GCII	GTP cyclohydrolase II domain	PF00925
GerE	LuxR-type DNA-binding HTH domain	PF00196
GGDEF	GGDEF domain	PF00990
HAMP	HAMP linker domain	PF00672
HATP	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase domains	PF13581, PF02518
HisKA	Histidine kinase domain	PF00512, PF07568, PF07730
HLH	Helix-loop-helix DNA-binding domain	PF00010
HTH	Helix-turn-helix_18 domain	PF12833
HWE	HWE histidine kinase domain	PF07536
Lipase	GDSL-like Lipase/Acylhydrolase	PF00657
Mase1	MASE1 domain	PF05231
NAD	Ferric reductase NAD binding domain	PF08030
PAS	PAS fold domain, PAS-associated, C-terminal motif	PF00989, PF08446, PF08447, PF08448, PF12860, PF13188, PF13426, PF13596, PF14598, PS50113
PHY	Phytochrome region domain	PF00360
Pkinase	Protein kinase domain	PF00069
RGS	Regulator of G protein signaling, Regulator of G protein signaling-like domains	PF00615, PF09128
RR	Response regulator receiver domain	PF00072
SAM	Sterile alpha motif domain	PF07647
SpoIIE	Stage II sporulation protein E	PF07228
STAS	STAS domain	PF01740
UNK	Domain of Unknown Function 3700	PF12481

8.2 Supplementary material

Table A2: Plasmid variants employed for the establishment of the reporter screening system. **(a)** Overview of reporter constructs varying in their promoter, RBS or reporter attributes. **(b)** Plasmid backbones for provision of the LOVA library constructs.

	Name	Vector	Ori / copy number	Description
a				
	pE_DsRed	pET28	pMB1 / low – med.	Original T7 promoter and RBS replaced by ami-lead sequence followed by DsRed2 gene
	newRBS-pE_DsRed	pET28	pMB1 / low – med.	RBS from pet28c restored
	T7_DsRed	pET28	pMB1 / low – med.	pE replaced by IPTG-dependent T7 promoter within ami-lead sequence
	pAraIO2I1I2_DsRed	pET28	pMB1 / low – med.	pE replaced by arabinose-dependent pBAD within ami-lead sequence
	pAraI1I2_DsRed	pET28	pMB1 / low – med.	pE replaced by arabinose-dependent pBAD missing the O2 regulatory-binding site
	pRha_DsRed	pET28	pMB1 / low – med.	pE replaced by rhamnose-dependent pRha within ami-lead sequence
	T5-DsRed	pET28	pMB1 / low – med.	pE replaced by T5 promoter within ami-lead
	pJ5-DsRed	pET28	pMB1 / low – med.	pE replaced by pJ5 within ami-lead sequence
	pBla-DsRed	pET28		pE replaced by pBla within ami-lead sequence
	pACYC-T7_DsRed	pACYC184	P15A / low	pE replaced by T7 promoter followed by terminator loops
	pE_ β-gal	pET28	pMB1 / low – med.	DsRed reporter replaced by β-gal
b				
	Library_pKT	pKT270 =pACYC184	P15A / low	pBAD and AraC sequences were transferred from pBAD30-M >> Arabinose-dependent expression of library constructs
	Library_pBAD	pBAD30-M	pUC / high	Arabinose-dependent expression of library constructs

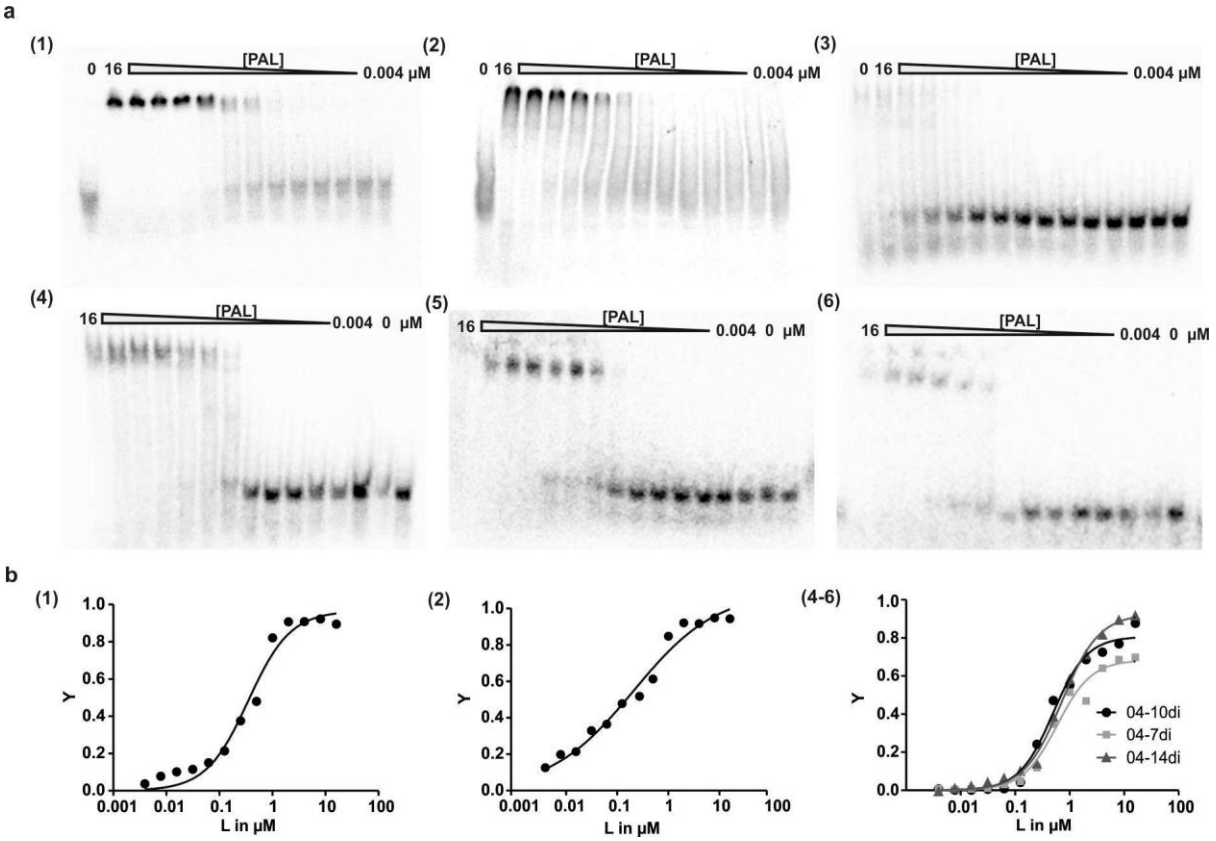


Figure A1: Comparison of binding behavior of SELEX clone 04 and a shortened version featuring only the 37 last nucleotides of the 3' end via EMSA. **(a)** EMSAs of PAL in the presence of SELEX clone 04.37 (1), 04-ori (2) under constant illumination, as well as the dimer variants 04-di10 under dark (3) and light conditions (4), and 04-di7 (5) and 04di14 (6) under light conditions. **(b)** RNA-dissociation curves of PAL of the EMSAs depicted in (a).

Table A3: Computed fitting values of the RNA affinity curves from (1), (2) and (4) – (6) shown in Figure A1.

Best-fit values	(1) 04.37	(2) 04-ori	(4) 04-di10	(5) 04-di7	(6) 04-di14
Y_{max}	0.97 ± 0.05	1.0 ± 0.1	0.80 ± 0.03	0.68 ± 0.03	0.93 ± 0.02
h	1.1 ± 0.2	0.53 ± 0.08	1.4 ± 0.2	1.4 ± 0.2	1.27 ± 0.09
K_d [nM]	350± 70	200 ± 90	490 ± 60	590 ± 90	710 ± 50

8.3 List of publications

1. Weber AM*, Kaiser J*, **Ziegler T***, Pilsl S, Renzl C, Sixt L, Pietruschka G, Moniot S, Kakoti A, Juraschitz M, Schrottke S, Lledo Bryant L, Steegborn C, Bittl R, Mayer G, Möglich A. Nat Chem Biol. 2019; doi: 10.1038/s41589-019-0346-y.

2. **Ziegler T**, Möglich A. Photoreceptor engineering. Biophysics. 2015;2: 30. doi:10.3389/fmolb.2015.00030

3. **Ziegler T***, Schumacher CH*, Möglich A. Guidelines for Photoreceptor Engineering. Methods MolBiol Clifton NJ. 2016;1408: 389–403. doi:10.1007/978-1-4939-3512-3_27

* equal contribution

8.4 Acknowledgments

An erster Stelle danke ich Prof. Dr. Andreas Möglich für die lehrreiche Zeit in seiner Arbeitsgruppe sowie die kontinuierliche wissenschaftliche Unterstützung bei der Weiterentwicklung des Projekts.

Die Arbeit an dieser Dissertation hat es mir nicht bloß ermöglicht mich wissenschaftlich weiterzuentwickeln, sondern mich ab einem bestimmten Punkt auch gezwungen zu überdenken aus welcher Quelle ich meine Motivationen und Antriebe schöpfe. Obwohl ich erkennen musste, dass für den Antrieb zur langjährigen Forschung an dieser Arbeit nicht nur intrinsische Faktoren eine Rolle gespielt haben, ist dieser Anteil doch groß genug, um mich selbst nach dem Entzug aller extrinsischer Ansporne mit Stolz und Freude zu erfüllen. Das persönliche Ziel, das ich mir für meine Doktorarbeit gesetzt habe, ging über die Beantwortung einiger isolierter Fragestellungen hinaus. Mein Wunsch war es stattdessen, diese Fragen zu einem Bild mit einer eigenen kleinen Geschichte zusammenzufügen, einem weiteren Puzzleteil innerhalb des großen Spiels der Wissenschaft (anstelle einer weiteren, abgeschnittenen LOV-Kristallstruktur). Dass meine Forschungsarbeit am Ende eine so spannende kleine Geschichte ergeben hat, wird mich für immer mit tiefster Freude erfüllen.

Nicht zuletzt ist dies der Auflösung der einzigartigen Volllängenkristallstruktur des PAL-Rezeptors zu verdanken, bei der ich durchgehend große Unterstützung erhalten habe. Mein größter Dank geht daher an Dr. Sébastien Moniot, da ich bei der Auswertung der Kristall-Daten andernfalls mit Sicherheit verzweifelt wäre. Dr. Martin Humenik danke ich für seine Hilfe bei den SEC-MALS-Messungen, und Prof. Dr. Olaf Stemmann für die Möglichkeit der Nutzung seines Radioaktiv-Labors.

Weiterhin danke ich unseren Kooperationspartner Prof. Dr. Günter Mayer und Anna-Maria Weber, durch deren hohe Expertise und Unterstützung auf der ‚anderen' (RNA) Seite das PAL Projekt sich so toll entwickeln konnte. Auch Dr. Hannes Ruwe und Prof. Schmitz-Linneweber danke ich für ihre Hilfe bei den Anfängen des Projekts.

Robert S. danke ich für die schöne gemeinsame Zeit in der Bayreuther Phase, und auch den Kollegen der benachbarten AG Steegborg für die gute Aufnahme. Auch der Zuwachs der AG Höcker war aus meiner Sicht eine große Bereicherung für das Institutsleben in Bayreuth, für die ich sehr dankbar bin.

Besonderer Dank geht an Nicteh, Przemek, und der wohl wärmsten Wohngemeinschaft, die man sich nur wünschen kann – dank euch ist die Zeit in Bayreuth (Forkendorf) zu einer Erfahrung geworden, die ich in meinem Leben nicht missen möchte.

Bei Carlos, Roman, Ralph, Florian und Charlotte möchte ich mich für die tolle Atmosphäre innerhalb der Berliner AG bedanken, nicht zu vergessen sind auch die Mitglieder der benachbarten AG Hegemann. Auch Prof. Hegemann selbst danke ich seine große Vorbildsfunktion – auch wenn es manchmal schwer ist, sich mit den Erinnerungen an die schöne „Blase“, die er um sich herum am Institut für Biophysik geschaffen hat, anschließend wieder mit der etwas raueren Realität zu arrangieren.

Natürlich möchte ich auch Robert, Johannes, Katrina und Uli danken, deren wissenschaftlicher Rat oft sehr hilfreich war, und die zudem enge Freunde sind. Robert danke ich außerdem für die schöne gemeinsame Laborzeit sowie für die Korrekturen an dieser Arbeit. Großer Dank geht auch an Laura für ihre rasche Promotion und die damit verbundene rechtzeitige Wiederkehr.

Von ganzem Herzen danke ich auch meinen Eltern, Harald und meiner Oma Heidi, deren unentwegter Rückhalt es mir ermöglicht hat mich zu dem glücklichen Menschen zu entwickeln, der ich bin. Insbesondere meine Mutter hat es geschafft mir in meiner Kindheit erfolgreich zu vermitteln, die Maßstäbe für meine Ziele, Erfolge und Misserfolge ausschließlich bei mir selbst zu suchen. Ich glaube mittlerweile, dass dies eines der größten Geschenke ist, die man einem Kind machen kann.

Nicht zuletzt danke ich Theresa, Max, Michael und Marc für ihr Dasein. Freunde zu haben, die trotz ihrer verschiedenen Hintergründe ein so stimmiges Verständnis von Freundschaft und Leben haben, ist eine immense Glücksquelle. Auch Daniel danke ich für seine nie erlöschende positive Energie, die die mir die letzte mühsame Etappe des Schreibens sehr erleichtert hat und mir selbst in dieser schweren Phase dabei geholfen hat, mich auf die vielen positiven Inhalte meines Lebens zu konzentrieren.

8.5 Eidesstattliche Versicherungen und Erklärungen

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit versichere ich eidesstattlich, dass ich die Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe (vgl. Art. 64 Abs. 1 Satz 6 BayHSchG).

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit erkläre ich, dass ich die Dissertation nicht bereits zur Erlangung eines akademischen Grades eingereicht habe und dass ich nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden habe.

(§ 8 Satz 2 Nr. 4 PromO Fakultät)

Hiermit erkläre ich, dass ich Hilfe von gewerblichen Promotionsberatern bzw. –vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe noch künftig in Anspruch nehmen werde.

(§ 8 Satz 2 Nr. 7 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass die elektronische Fassung der Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung unterzogen werden kann.

(§ 8 Satz 2 Nr. 8 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass bei Verdacht wissenschaftlichen Fehlverhaltens Ermittlungen durch universitätsinterne Organe der wissenschaftlichen Selbstkontrolle stattfinden können.

……

Ort, Datum, Unterschrift