# Conservation of function with diversification of higher-order structure within sensor histidine kinases

Igor Dikiy[1], Danielle Swingle[1,2], Kaitlyn Toy[1,3], Uthama R. Edupuganti[1,2], Giomar Rivera-Cancel[4], Kevin H. Gardner[1,3,5*]

[1]: Structural Biology Initiative, CUNY Advanced Science Research Center, New York, NY 10031

[2]: Biochemistry Ph.D. Program, The Graduate Center – City University of New York, New York, NY 10016

[3]: Department of Chemistry and Biochemistry, City College of New York, New York, NY 10016

[4]: Department of Internal Medicine, UT Southwestern Medical Center, Dallas, TX 75390

[5]: Ph.D. Programs in Biochemistry, Biology, and Chemistry, The Graduate Center – City University of New York, New York, NY 10016

*: direct correspondence to kgardner@gc.cuny.edu

ORCID iDs:

Igor Dikiy: 0000-0002-8155-7788

Danielle Swingle: 0000-0003-3691-4560

Kaitlyn Toy: 0000-0003-2595-8779

Uthama R. Edupuganti: 0000-0003-0571-4101

Giomar Rivera-Cancel: 0000-0002-4909-4103

Kevin H. Gardner: 0000-0002-8671-2556

**Abstract**

Integral to the protein structure/function paradigm, oligomeric state is typically conserved along with function across evolution. However, notable exceptions – such as the hemoglobins – show how evolution can enable new regulatory mechanisms by oligomerization changes. We examine this linkage in the histidine kinases (HKs), widely-distributed prokaryotic environmental sensors. While most HKs are transmembrane homodimers, our finding of a monomeric soluble HK (EL346, a photosensing Light-Oxygen-Voltage (LOV)-HK) raises questions about the relationship between oligomerization and regulation within this family. Here we characterize multiple EL346 homologs and find a diverse range of HK oligomeric states. Three LOV-HK homologs are dimers with differing structural responses to light, while two are Per-ARNT-Sim (PAS)-HKs that interconvert between differentially-active monomers and dimers, suggesting dimerization as a regulatory control. Finally, we examine putative interfaces in a dimeric LOV-HK, finding that multiple domains contribute to dimerization and function. Our findings suggest potential for novel regulation in this important family of environmental sensors.

## Introduction

Sensor histidine kinases (HKs) are signal transduction receptors that are widespread in prokaryotic organisms, enabling them to sense and respond to diverse environmental stimuli (Galperin, 2018). HKs are typically found in two-component systems (TCS), which are most simply composed of a sensor HK and a downstream response regulator (RR) collocated in the same operon. The RR responds to a signal from the HK by acting as a phosphorylation-dependent transcription regulator, usually binding to DNA and controlling gene expression in various bacterial processes, including virulence (Kenney & Anand, 2020; Mechaly et al., 2017; Wang, 2012; West & Stock, 2001). The majority of sensor HKs characterized to date are obligate homodimeric transmembrane receptors, facilitating transmission of signals across the membrane from periplasmic sensor domains to cytoplasmic kinase domains via intermonomer, symmetry-breaking conformational changes (Bhate et al., 2015; Gushchin & Gordeliy, 2018; Szurmant et al., 2007). In contrast, some sensor HKs adopt varying oligomeric states, ranging from large multimeric assemblies observed for bacterial chemosensors (Muok et al., 2020) to the monomeric soluble HK EL346 (Rivera-Cancel et al., 2014). Soluble sensor HKs are relieved of the constraints of transmitting a signal across a lipid bilayer, which typically necessitates signal transmission via motions of one monomer with respect to the other (Bhate et al., 2015; Herrou et al., 2017). This raises the open question: Are such soluble HKs more likely to adopt varied oligomeric states including non-dimeric architectures? We view the relatively understudied HWE/HisKA2 family of HKs (Herrou et al., 2017), which is enriched in soluble cytoplasmic proteins that participate in atypical signaling pathways, to be a tempting testbed for this hypothesis.

Typical sensor HKs contain an N-terminal sensor domain and a C-terminal kinase domain consisting of a two-helix dimerization histidine phosphotransfer (DHp) domain and a catalytic ATP-binding (CA) domain (Bhate et al., 2015). Stimulus detection at the sensor domain, or in some cases at other domains (Kenney & Anand, 2020), modulates the CA domain autophosphorylation of a conserved His residue in the DHp domain, a signal which is subsequently transferred to a conserved Asp in the RR receiver domain (West & Stock, 2001). The homodimeric nature of canonical transmembrane HKs is largely believed to be required for both the structural changes between the “off” and “on” state and determining whether autophosphorylation occurs between monomers (*in trans*) or *in cis* within a monomer (Ashenberg et al., 2013; Bhate et al., 2015). Dimerization is mediated via the DHp domains of two monomers, forming a four-helix bundle and facilitating pivot-, piston-, or scissoring-type

71 movements as part of the activation process (Bhate et al., 2015; Herrou et al., 2017), echoing
72 similar themes in integrins (Luo et al., 2007) and certain other transmembrane receptors.
73 This traditional view of sensor HK structure has been built up from extensive structural
74 and biochemical studies (Bhate et al., 2015), most of which have focused on HKs in only one of
75 several subfamilies of histidine kinases (Pfam (El-Gebali et al., 2019) family HisKA). In contrast,
76 the related HWE-HK and HisKA_2 families (referred to here as the HWE/HisKA2 family) of HKs
77 have not been as extensively studied from a structural perspective (Herrou et al., 2017). This
78 family, which differs from the canonical HisKA family in the DHp and CA domains, appears to be
79 enriched in HKs lacking transmembrane segments and correspondingly likely localized to the
80 cytosol. These HKs also tend to signal in an atypical manner, utilizing phosphorylation to control
81 protein-protein interactions of RRs, rather than directly impacting them as transcription factors
82 *per se* (Herrou et al., 2017). HWE/HisKA2 HKs are also enriched in the *Alphaproteobacteria*
83 (Karniol & Vierstra, 2004; Staroń & Mascher, 2010).
84 Our prior work on the HWE/HisKA2 family has focused on three soluble HKs involved in
85 the general stress response of the marine alphaproteobacterium *Erythrobacter litoralis*
86 HTCC2594 (EL346, EL362, and EL368) (Correa et al., 2013; Rivera-Cancel et al., 2014; Swartz
87 et al., 2007). Each of these proteins senses blue light through a light oxygen voltage (LOV)
88 sensor domain (Crosson et al., 2003; Pudasaini et al., 2015), a sub-class of the versatile Per-
89 ARNT-Sim (PAS) domains that are present in over 30% of HKs (Gao & Stock, 2009). The
90 response to blue light provides an experimentally tractable way to explore the inactive and
91 active states of histidine kinases, typically corresponding to dark and light states of LOV-HKs
92 (Möglich, 2019; Swartz et al., 2007; Tseng et al., 2010). These states can be conveniently
93 distinguished spectrophotometrically: while the dark state spectra of LOV-HKs exhibit a triplet
94 peak at around 450 nm, characteristic of protein-bound flavin chromophores, the light state
95 spectrum has a broad, singlet absorption peak at 390 nm (Zoltowski & Gardner, 2011). This
96 spectroscopic change is due to the formation of a photochemically-generated covalent adduct
97 between the C(4a) of the isoalloxazine ring and a conserved cysteine residue, a bond that is
98 thermally reversible in the dark on a timescale varying from seconds to hours among LOV
99 domains (Pudasaini et al., 2015).
100 While our prior studies indicate that EL368 and EL362 (when reconstituted with flavin)
101 are dimeric, EL346 is a soluble and functional monomer (Correa et al., 2013; Rivera-Cancel et
102 al., 2014). The EL346 crystal structure reveals that it is held in a monomeric state by an
103 intramolecular association between its DHp-like domain and the LOV sensory domain, blocking
104 the face typically used for DHp domains to dimerize (Rivera-Cancel et al., 2014). Despite this

105 substantial deviation from current HK signaling models, the monomeric EL346 fully responds to
106 stimulus with a light-induced increase in autophosphorylation (Rivera-Cancel et al., 2014). More
107 broadly, EL346 suggests that soluble HKs may not require dimerization as a key element in
108 signal transduction, unlike conventional transmembrane HKs. Thus, we set out to characterize
109 the oligomeric and functional states of several novel HWE/HisKA2 HKs.
110 To do so, we used sequence homology searches to identify multiple soluble LOV- and
111 PAS-HKs with similarity to EL346 and used a combination of approaches to test if these
112 proteins were similarly functional monomers. We report here our characterization of three novel
113 LOV-HKs, each of which properly photocycle by UV-visible absorbance spectroscopy and are
114 dimeric with a range of light-triggered conformational changes, as detected by size exclusion
115 chromatography coupled with multi-angle light scattering (SEC-MALS). In contrast, we also
116 detail results from two light-insensitive PAS-HKs which exist in monomeric and dimeric states
117 with differing autophosphorylation activity, suggesting dimerization as a possible regulatory
118 switch in this family of HKs. Finally, we combine structural and sequence information to
119 engineer a "monomerized dimer" by deleting putative dimerization interfaces in a dimeric LOV-
120 HK, finding that a distributed interface of interactions encodes both dimerization and autokinase
121 activity. Taken together, our data indicate that a wider range of oligomerization states are
122 accessible to HKs, raising exciting questions about regulation and evolution potentially
123 analogous to other protein families which exhibit differences in quaternary structure and control,
124 such as the hemoglobins (Pillai et al., 2020).
125
126 **Methods**
127 Bioinformatics
128 EL346 homologs were identified by searching for non-redundant protein sequences in
129 the GenBank database using BLAST with EL346 amino acids 1-346 as query. EL346 and the
130 best 100 sequences were aligned using Clustal Omega (Sievers et al., 2011) and a distance
131 tree was calculated from this alignment using the same software. Graphic display of this tree
132 was accomplished using iTOL (Interactive Tree of Life, http://itol.embl.de/) (Letunic & Bork,
133 2019) and Adobe Illustrator CS6. This tree was used to select HKs for cloning from the strains
134 available at DSMZ (http://www.dsmz.de).
135
136 Cloning, protein expression and purification
137 The HKs selected from each strain are shown in **Table 1**. The proteins were cloned into
138 a pHis-Gβ1 vector (Sheffield et al., 1999). Constructs were verified by DNA sequencing. The

139 proteins were overexpressed in *Escherichia coli* BL21(DE3) (Stratagene) in LB as previously
140 described (Dikiy et al., 2019). Cells were harvested, resuspended in Buffer A (50 mM Tris pH
141 8.0, 100 mM NaCl, and 10 mM $MgCl_2$), and lysed by sonication. The lysate was centrifuged at
142 48,000 x g and 4 °C for 45 min. The supernatant was loaded into a $Ni^{2+}$ Sepharose affinity
143 column (GE Healthcare) and the His6-Gβ1 tagged protein was washed with 4 column volumes
144 of Buffer A supplemented with 15 mM imidazole and eluted with 4 column volumes of Buffer A
145 supplemented with 250 mM imidazole. Eluted proteins were exchanged into Buffer A by dialysis.
146 The fusion tag was cleaved by incubation with His6-TEV protease overnight at 4 °C. The tag-
147 less proteins were separated from the tags and His6-TEV protease by $Ni^{2+}$ affinity
148 chromatography and were further purified by size-exclusion chromatography on a HiLoad
149 16/600 Superdex 200 column (GE Healthcare) equilibrated with HSEC buffer containing 50 mM
150 HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 0.5 mM DTT. For light-sensitive proteins, all
151 purification steps were performed under dim red light. Concentrations were determined from the
152 theoretical absorption coefficient, $\varepsilon_{280}$ for PAS-HKs, calculated from the sequence using the
153 ExPASy ProtParam server (Gasteiger et al., 2005), and $\varepsilon_{446}$ = 11,800 $M^{-1}$ $cm^{-1}$ for flavin-
154 containing proteins.
155 RH376 deletion mutants were generated as follows: RH2 and RH3 were ordered as
156 synthetic constructs from Genewiz (South Plainfield, NJ) and inserted into the same pHis-Gβ1
157 vector as the WT protein. The RH1 deletion was generated by PCR using an N-terminal primer
158 lacking the sequence to be deleted using WT and RH3 plasmids as templates, to generate RH1
159 and RH1+3, respectively.
160
161 Absorbance and dark state reversion kinetics
162 UV-visible absorbance spectroscopy measurements were acquired using a Varian Cary
163 50 spectrophotometer at 24 °C with a 1 cm path length quartz cuvette. First, a UV-visible
164 spectrum in the range 250-550 nm was acquired in light and dark conditions for each LOV-HK.
165 For dark measurements, all steps were performed under dim red light. For light measurements,
166 the samples were illuminated for 1 minute using a blue LED panel (225 LED bulbs in 30.5 ×
167 30.5-cm panel, 13.8 W, 465 nm maximum illumination wavelength; LEDwholesalers [Hayward,
168 CA]) prior to each scan. To measure dark state recovery, each LOV-HK sample was illuminated
169 under a blue LED panel for 1 minute and then transferred to the spectrophotometer. $A_{446}$ of the
170 samples in HSEC buffer with 1 mM ATP was monitored at 5-min intervals for 400-800 min at
171 24 °C. Time constants τ were determined by fitting measurements to a monoexponential decay
172 using the following equation:

173 $Abs = A_{inf} - (A_{inf} - A_o)e^{-t/\tau}$ , where $A_{inf} = A_{446}$ at infinity, $A_o = A_{446}$ initial, and t = time
174 Raw data were fit and plotted using RStudio and plots were edited in Adobe Illustrator CS6.
175
176 SEC-MALS
177 All samples and buffers were filtered with a 0.1-um pore size filter before use. ATP was
178 added to each HK sample (~15 μM, 400 μL) to a final concentration of 1 mM. Samples were
179 then injected at a 0.4 mL/min flow rate onto a Superdex 200 GL 10/300 SEC column (GE
180 Healthcare) using HSEC buffer with 1 mM ATP. For lit experiments, samples were illuminated
181 with a blue LED panel for 1 min prior to injection and the blue LED panel was faced toward the
182 column during the run. For dark experiments, all steps were performed under dim red light.
183 Samples were detected post-elution by the inline miniDAWN TREOS light scattering and Optilab
184 rEX refractive index detectors (Wyatt Technology [Santa Barbara, CA]). FPLC runs were
185 performed at 4 °C, while MALS and refractive index measurements were at 25 °C. Data analysis
186 and molecular weight calculations were performed using the ASTRA V software (Wyatt
187 Technology). Raw data were plotted using RStudio and plots were edited in Adobe Illustrator
188 CS6 and overlaid in Adobe Photoshop CS6.
189
190 Autophosphorylation assays
191 Autophosphorylation was measured as described previously (Dikiy et al., 2019; Rivera-
192 Cancel et al., 2014). Two trials were conducted for each protein under each condition. Proteins
193 were added to a reaction buffer containing 50 mM Tris, 100 mM NaCl, 5 mM $MnCl_2$, and 1 mM
194 DTT with pH 8.2. Protein concentrations were confirmed by analysis of dilution series on
195 Coomassie-stained SDS-PAGE. Concentrations typically ranged from 4.7 to 19.8 μM, with
196 EL346 and RH2 able to be more concentrated at 34.1 and 67.1 μM. A mixture of unlabeled ATP
197 and 10 μCi [$\gamma$-$^{32}$P] ATP was added to each mixture to initiate the reaction (final ATP
198 concentration of 3.6 μM). Aliquots were removed at time points of 0.5, 1, 2, 4, 8, 6, and 32 min
199 and placed into a 4x SDS-gel loading buffer to quench the reaction. For dark measurements, all
200 steps were performed under dim red light. For light measurements, the samples were
201 illuminated with a blue LED panel just prior to and throughout the course of the experiment.
202 Samples were subjected to SDS-PAGE analysis; dried gels were exposed overnight and bands
203 were visualized by phosphorimaging. Band intensities were measured using Fiji (Rueden et al.,
204 2017; Schindelin et al., 2012); initial velocity values were calculated in Microsoft Excel from the
205 linear region of the curve and were plotted in RStudio and edited in Adobe Illustrator CS6.
206

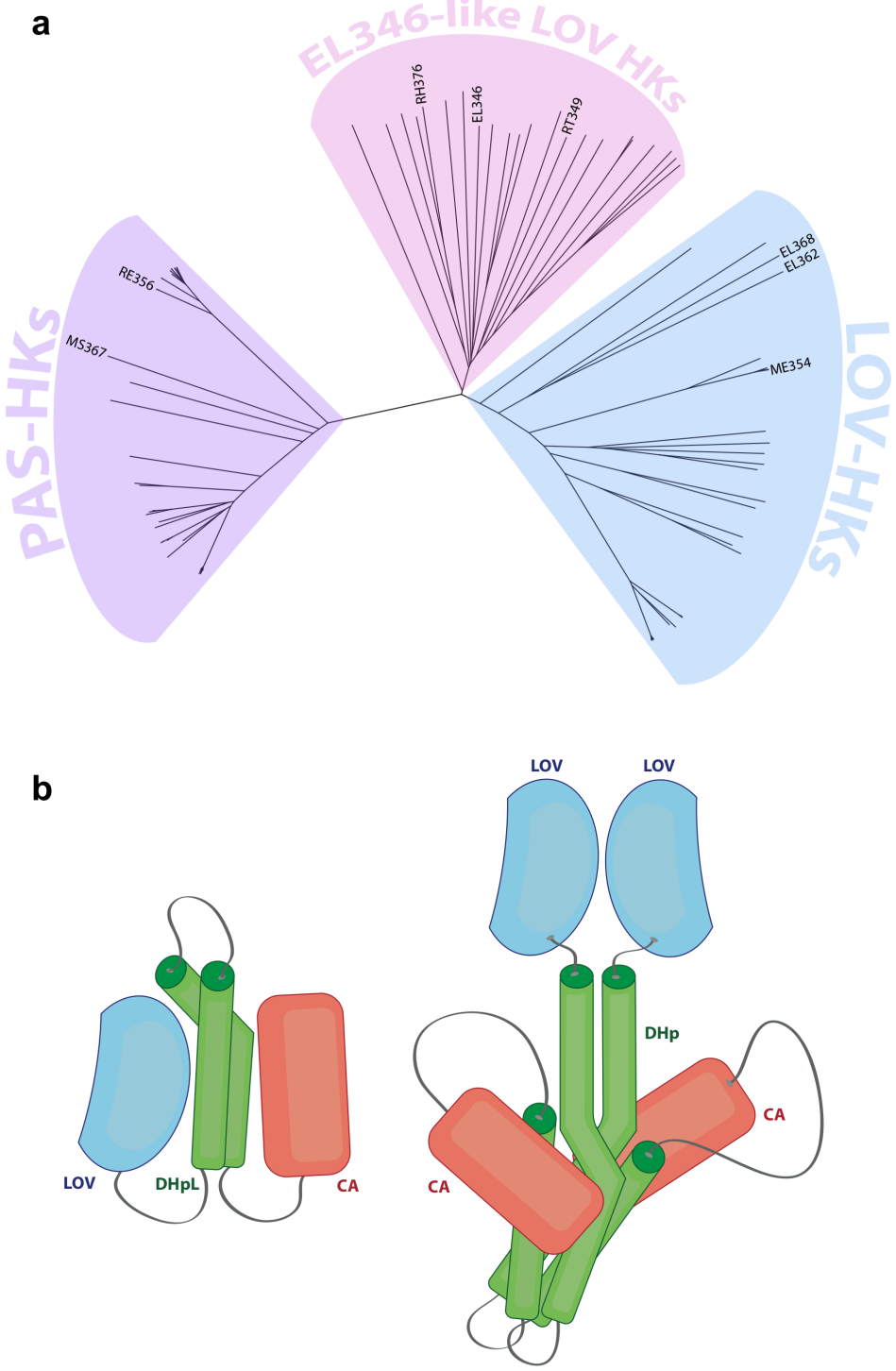


**Figure 1: EL346 homologs form three distinct clades. a)** PAS-HKs (purple), EL346-like LOV HKs (pink), and LOV-HKs (blue). The names of the proteins from each clade that were selected for this study are shown. All other names are omitted for simplicity and can be found in Figure S2. **b)** Schematic structures of a monomeric (left) and dimeric (right) LOV-HK. Monomeric schematic is based on the structure of EL346 (Rivera-Cancel et al., 2014) and dimeric is based on the fusion protein YF1 (Diensthuber et al., 2013).

207

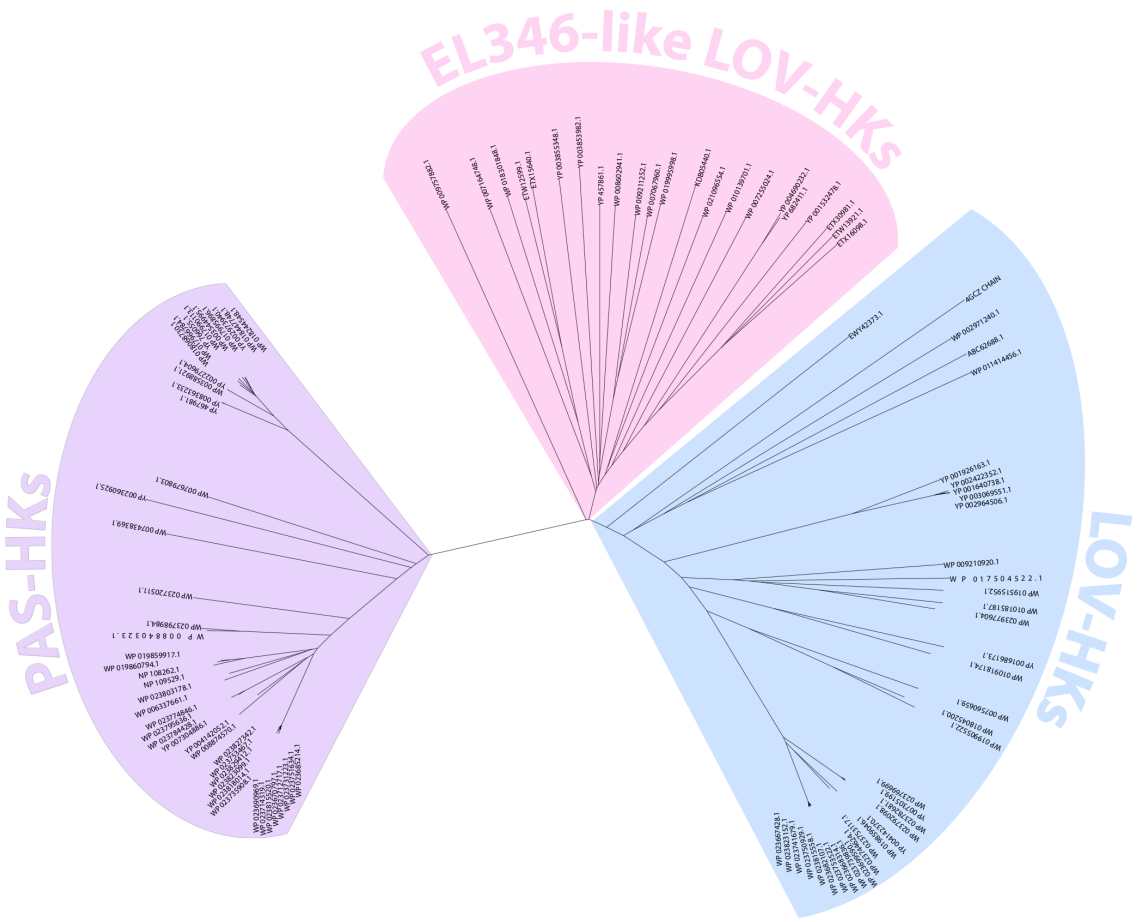


**Figure S2: MSA Tree of EL346 Homologs (Supplemental to Fig. 1).** NCBI reference numbers for each of the 100 EL346 homologues shown in Figure 1 are listed here.

## Results

Bioinformatics analysis reveals numerous alphaproteobacterial EL346 homologs

To identify potential additional monomeric histidine kinases, we searched for homologs to the full-length EL346 sequence as a well-validated, functional monomer. Taking the top 100 BLASTp hits (in addition to EL346), all belonging to the *Alphaproteobacteria*, we generated a multiple sequence alignment of proteins which all contained the same domain architecture of a single N-terminal PAS or LOV domain followed by a C-terminal HWE/HisKA2 histidine kinase, and without any predicted transmembrane segments. A distance tree analysis revealed three clades (**Fig. 1a**) which split on sensor type: PAS-HKs (non-LOV), LOV-HKs, and EL346-like LOV-HKs. Interestingly, a subgroup of the EL346-like LOV-HKs is missing the conserved phosphoacceptor histidine (**Fig. S1**). We also included the related EL368 (dimer) and EL362 (blind monomer) histidine kinases from *E. litoralis* HTCC2594 that we had previously identified

221 (Correa et al., 2013), as well as the *Caulobacter crescentus* LovK (dimer) (Purcell et al., 2010)
222 in the alignment; all three of these proteins clustered in the LOV-HKs clade. Additionally, our
223 search identified a third PAS-HK clade with proteins containing non-LOV PAS domains with 17-
224 24% sequence identity to the EL346 LOV domain. Notably, many of these domains contained a
225 conserved tryptophan in place of the cysteine utilized in LOV-type photochemistry (Crosson &
226 Moffat, 2001) (**Fig. S1**). These data show that the PAS/LOV-HK architecture is widespread
227 throughout *Alphaproteobacteria* and contains members which could well respond to stimuli
228 other than blue light.
229

| | Organism | Protein Name | Uniprot ID | NCBI ref. | Identity to EL346 | Oligomeric State |
|---|---|---|---|---|---|---|
| ***EL346-like LOV-HKs*** | *Erythrobacter litoralis HTCC2594* | EL346 | Q2NB77 | YP_457861.1 | N/A | Monomer |
| | *Rubellimicrobium thermophilum DSM 16684* | RT349 | S9SB69 | WP_021096554.1 | 42.3% | Dimer |
| | *Roseivivax halodurans JCM 10272* | RH376 | X7EIB8 | ETX15640.1 | 35.8% | Dimer |
| ***LOV-HKs*** | *Methylorubrum extorquens DM4* | ME354 | C7CCX6 | YP_003069551.1 | 29.6% | Dimer |
| | *Erythrobacter litoralis HTCC2594* | EL368 | Q4TL45 | ABC62688.1 | 25.5% | Dimer |
| | *Erythrobacter litoralis HTCC2594* | EL362 | Q2N9L9 | WP_011414456.1 | 25.8% | Monomer |
| ***PAS-HKs*** | *Rhizobium etli CFM 42* | RE356 | Q2KD32 | YP_4679981.1 | 30.1% | Monomer/ Dimer |
| | *Methylocella silvestris BL2* | MS367 | B8ELH4 | YP_002360925.1 | 28.7% | Monomer/ Dimer |
| ***Other*** | *Bacillus subtilis/ Bradyrhizobium japonicum* | YF1 (engineered YtvA/FixL fusion) | N/A | N/A | 25.5% | Dimer |

**Table 1: Oligomeric states of known LOV-HKs and PAS-HKs include mix of monomers and dimers.** Soluble HKs shown here are separated by clade with the same coloring as Fig. 1 (EL346-like LOV HKs – pink, LOV-HKs – blue, PAS-HKs – purple, Other – grey) and include those investigated in this study (EL346, RT349, RH376, ME354, RE356, and MS367) as well as those identified in literature (EL368, EL362, and YF1) (Correa et al., 2013; Diensthuber et al., 2013). Oligomeric states were determined in this study by SEC-MALS.

230

231 LOV-HK homologs are light-sensing and appear mostly dimeric

232 We posited that the EL346-like LOV-HK clade might contain other monomeric HKs
233 similar to EL346, in addition to other members which might adopt the more commonly-observed
234 dimeric assembly seen in the engineered LOV-HK YF1 (Diensthuber et al., 2013), as illustrated
235 in **Fig. 1b**. To test this hypothesis, we selected several HKs from each clade and characterized
236 their oligomeric state and function with a variety of biochemical assays. We started by
237 determining the ability of the LOV-containing proteins to bind flavin chromophores and undergo
238 typical LOV photochemistry via UV-visible absorbance spectroscopy. The dark and light spectra
239 of four representative proteins, EL346, ME354, RT349, and RH376 (see **Table 1** for protein
240 details) all show the characteristic flavin-LOV triplet (~450 nm) in the dark state which
241 disappears upon illumination (**Fig. 2a**). Upon incubation in the dark, the 450 nm absorbance
242 returns with exponential kinetics and a time constant that varies by protein: the other LOV-HKs
243 measured here, such as ME354 (236 min), RT349 (240 min), and RH376 (62 min), had longer
244 reversion constants $\tau$ than EL346 (33 min) in the presence of 1 mM ATP (**Fig. 2b**). The notably
245 long reversion time constants of ME354 and RT349 can be explained by the presence of
246 several "slowing" mutations in their LOV domains relative to that of EL346, specifically at
247 positions 19, 21, 32, and 101 (EL346 numbering), as collated in a review by Pudasaini and
248 coworkers (Pudasaini et al., 2015).

249 We determined the oligomeric state of these LOV-HKs in the presence of ATP and either
250 dark or light state, using size exclusion chromatography coupled to multi-angle light scattering
251 (SEC-MALS), starting by confirming that EL346 is monomeric under both conditions with
252 minimal change in shape upon illumination as revealed by SEC elution profiles (**Fig. 2c**). In
253 contrast, the three other LOV-HK proteins are stable dimers (via MALS data) but undergo a
254 range of light-dependent shape changes which impact SEC elution times. These range from
255 minimal, with ME354 eluting as a single peak that does not noticeably change between dark
256 and light, to substantial shifts to later elution times for both RT349 and RH376 consistent with a
257 light-induced compaction. Overall, our results suggest that other LOV-HKs are generally dimeric
258 and exhibit varying global conformational changes in response to light.

259

260 Non-LOV PAS-HKs show mix of monomers and dimers, both of which are active

261 The two PAS-HKs we characterized, from *Rhizobium etli* CFN 42 (RE356) and from
262 *Methylocella silvestris* BL2 (MS367), appear as a mixture of dimer and monomer, as well as an
263 oligomeric or aggregated fraction, as confirmed by SEC-MALS (**Fig. 3a**). These two proteins
264 exhibit 45% sequence identity to one another (32% over the PAS domain) and behave very

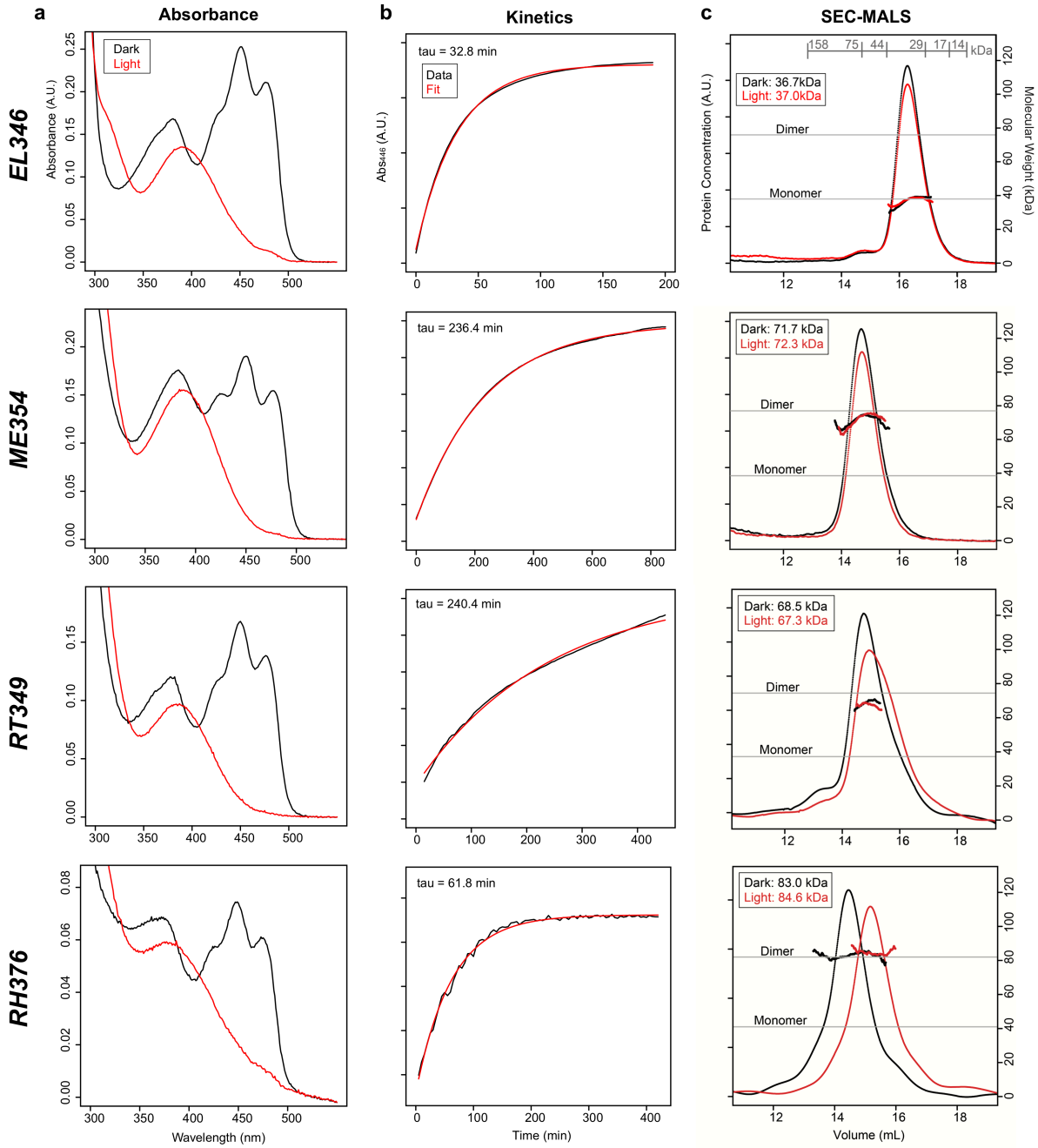


**Figure 2: UV-visible absorption spectra and SEC-MALS reveal diversity of LOV-HK dark state reversion kinetics and large scale responses to illumination. a)** UV-visible absorption spectra of EL346, ME354, RT349, and RH376, superimposing dark state scans (black) with scans post-illumination (red). All spectra display the disappearance of the 450 nm triplet upon exposure to light, consistent with LOV photochemistry. **b)** Dark state reversion kinetics, measured by the return of absorbance at 446 nm post illumination (black), fit to single exponentials (red) reveal an 8-fold range of time constants. **c)** SEC-MALS elution peaks are represented by dRI in arbitrary units and MALS-derived masses are solid lines under the peaks. MWs and elution volumes of six standard proteins are shown in grey in the top panel. MALS traces show that EL346 is monomeric in solution, while ME354, RT349 and RH376 are all predominantly dimeric. A light-dependent shift in elution peak shape and position is seen in both RT349 and RH376.

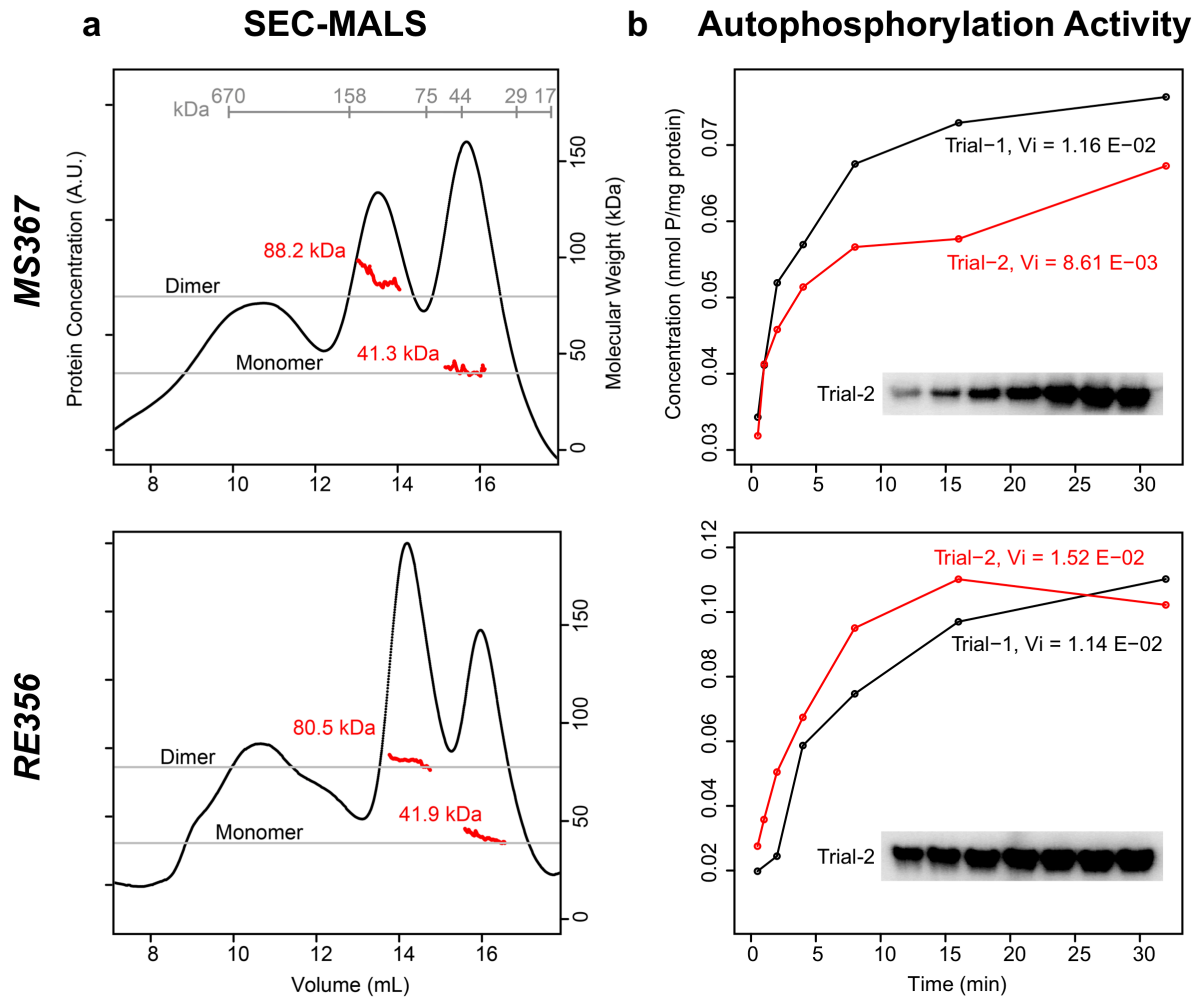


**Figure 3: The MS367 and RE356 PAS-HKs are monomer/dimer mixtures with autophosphorylation activities. a)** SEC-MALS measurements of PAS-HKs MS367 and RE356 reveal the presence of both monomeric and dimeric states (mass, red traces). The elution peaks (black trace) of each protein are consistent with the expected volumes, suggesting both adopt globular states. **b)** Autophosphorylation activity curves suggest that both PAS-HKs are active. Two trials were conducted for each protein (mixture of monomer/dimer) and both measurements are shown here. One representative gel is shown for each protein. Initial velocities for each trial were calculated from the linear region of these curves.

similarly in our biophysical tests and activity assays. While these proteins lack a known trigger analogous to light for the LOV-HKs, we still wanted to determine whether both proteins exhibited any autophosphorylation in their default state as purified from *E. coli*. Both MS367 and RE356 displayed autophosphorylation activity approximately one order of magnitude higher than that of EL346 as determined previously (Dikiy et al., 2019; Rivera-Cancel et al., 2014) and confirmed in this study (see below), with the characteristic plateau indicating an equilibrium between kinase and phosphatase activity (**Fig. 3b**). The activity of RE356 was slightly higher than that of MS367, with approximately 1.5-fold greater initial velocity. It is unclear whether this level of

activity indicates a high basal activity or whether binding to the putative trigger turns off the kinase.

We wondered whether oligomeric state (monomer *vs* dimer) affects autophosphorylation activity. We found that the monomeric and dimeric states of MS367 and RE356 are in slow equilibrium, and could be separately injected onto SEC-MALS, as shown for MS367 in **Fig. 4a**. We also determined autophosphorylation activity separately for the monomer and dimer fractions of MS367 and found that the dimer is approximately 4-fold more active than the monomer fraction (**Fig. 4b**). However, the dimeric fraction also exhibited a large proportion of oligomer/aggregate, suggesting that the dimeric state is less stable or prone to further oligomerization (**Fig. 4a**). Using size-exclusion chromatography, we determined that the monomeric state of RE356 is stable on the order of days at 4 °C, while the dimer slowly oligomerizes or aggregates on this timescale (**Fig. 4c**). These data illustrate that active monomeric HKs are to be found in this family, but dimerization may serve as a regulatory mechanism.

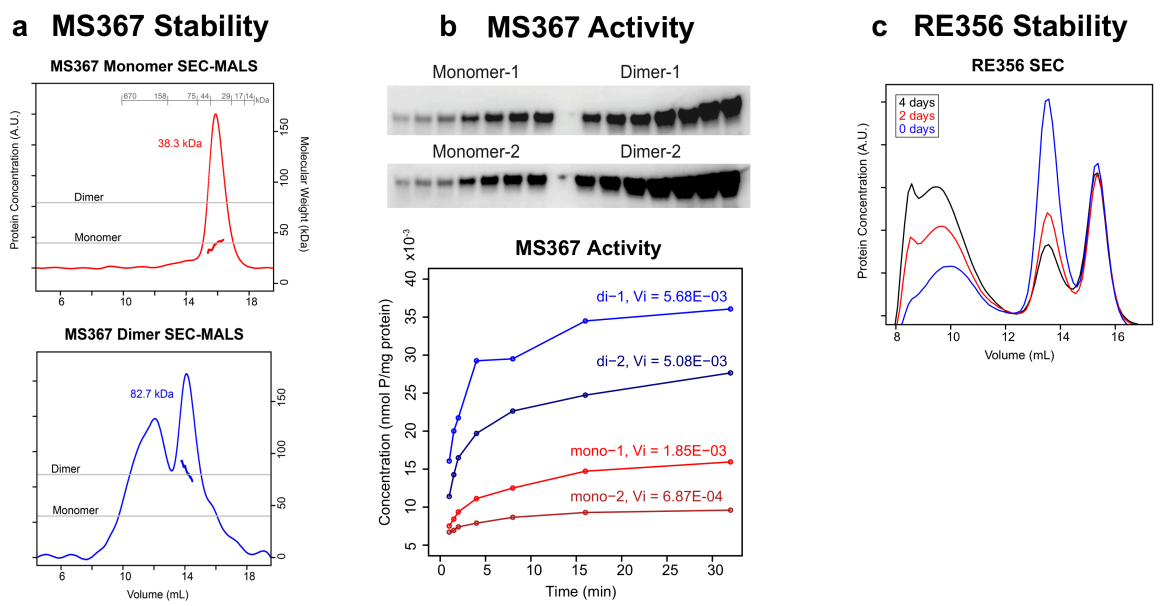


**Figure 4: PAS-HK monomer/dimer states can be separated, dimer is more active than monomer, and monomer is more stable over time. a)** MS367 monomers and dimers were separated and injected onto a Superdex 200 GL 10/300 SEC column. The resulting MALS-derived masses (solid lines under elution peaks) show that both remain stable upon separation. **b)** MS367 autophosphorylation assays (phosphor images, top and plots, bottom) of the monomer (red traces) and dimer (blue traces) of MS367 show that the dimer has more autophosphorylation activity. **c)** SEC traces of RE356 dimer and monomer after 0, 2, and 4 days at 4 °C are shown, revealing that the monomer is stable over time, while the dimer tends to aggregate.

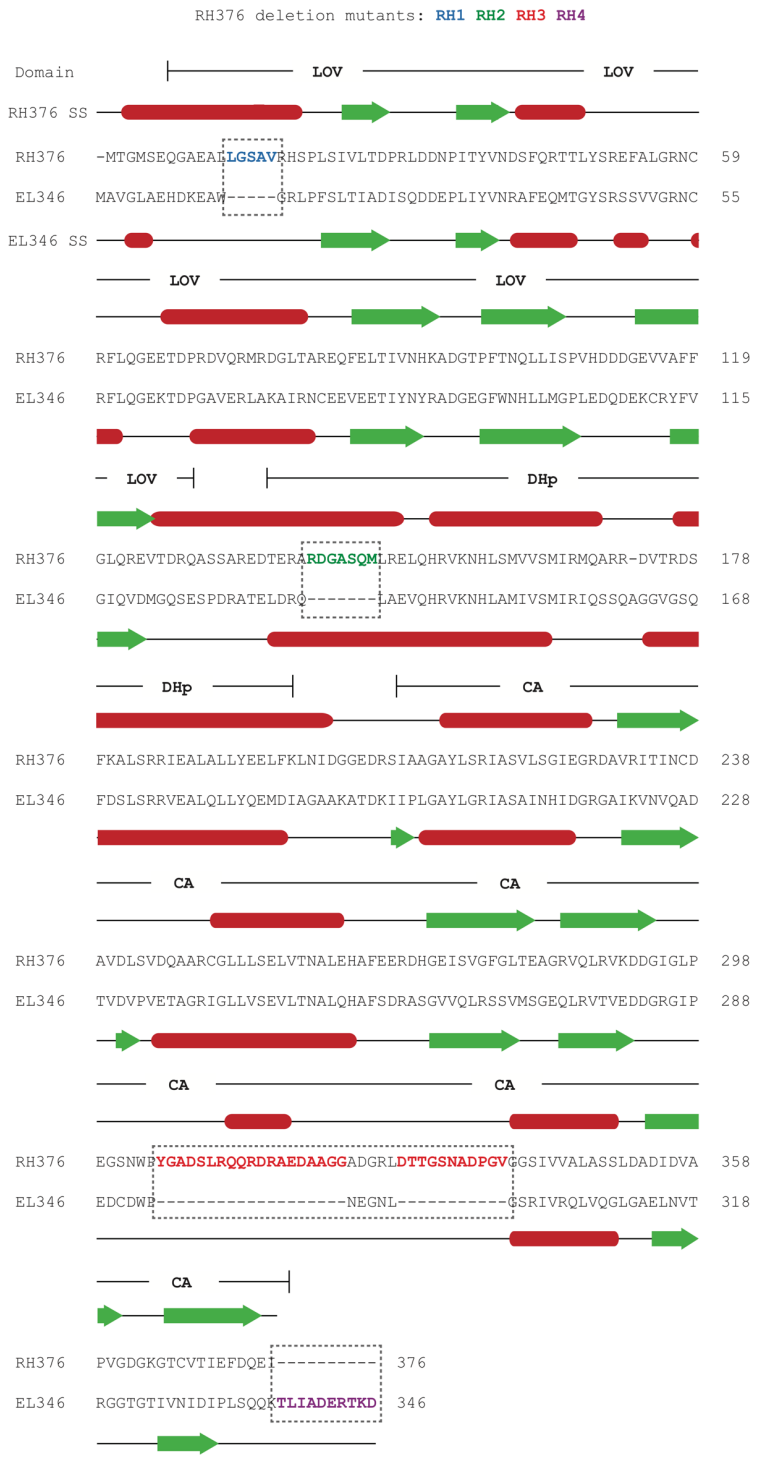


**Figure S3: Pairwise alignment of EL346 and RH376 reveals "gaps" in EL346 sequence (Supplemental to Fig. 5).** The pairwise alignment of EL346 and RH376 sequences was performed by MUSCLE (Edgar, 2004). "Insertion" regions of RH376 termed RH1-4 are highlighted with dashed boxes. Secondary structure prediction by JPred (Drozdetskiy et al., 2015) (red = helices, green = sheets) for RH376 and from the 4R3A PDB entry for EL346 (Rivera-Cancel et al., 2014) as well as domain locations (LOV, DHp, and CA) are shown next to the amino acid sequences.

292

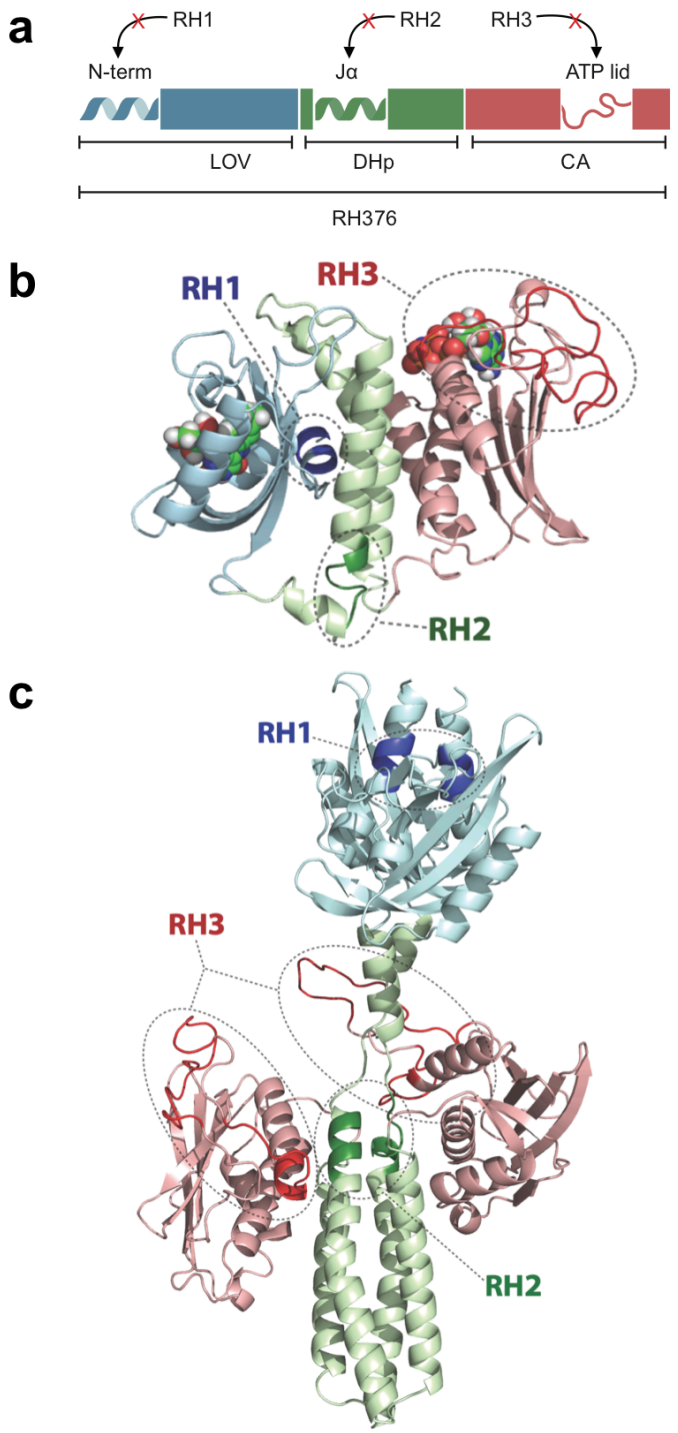


**Figure 5: RH376 mutant deletion regions are designed in predicted dimer interfaces.** RH376 deletion mutants RH1, RH2, and RH3 are highlighted in 3 schematics: **a)** Domain architecture with the predicted secondary structure at the location of each deletion. RH1 is located in the A'α helix of the LOV domain, RH2 in the linker between the LOV domain and α1 of the DHp domain, and RH3 is in the ATP lid. **b)** Monomeric homology model of RH376 based on the structure of EL346 (Rivera-Cancel et al., 2014). **c)** Dimeric homology model of RH376 based on the engineered YF1 sensor kinase (Diensthuber et al., 2013). All LOV domains are colored blue, DHp domains green, and CA domains red. Mutations are highlighted and circled.

293

294 Engineering a “monomerized” LOV-HK

295 With the possibility of monomer-dimer transition as a regulatory mechanism, we sought
296 to isolate the determinants of dimerization in this family of sensor HKs. During evolution, many
297 factors (e.g. changes to protein sequence) can influence the multimeric arrangement of proteins
298 belonging to the same family (Perica et al., 2012; Pillai et al., 2020). To investigate such
299 determinants of the oligomeric state, we compared monomeric EL346 with selected dimeric HKs
300 from all three clades using sequence alignment. These pairwise alignments highlighted four
301 regions of difference in the primary sequence between EL346 and the other HKs; these regions
302 were distributed differently throughout the three clades. However, the dimeric RH376 from the
303 EL346-like clade contained all four regions of insertions of amino acids between the two
304 proteins (**Fig. S3, 5a**). We named the regions RH1-4 and set out to test the hypothesis that one
305 or more of these might be determinants of oligomeric state. Region RH1 is an insertion of 5
306 amino acids in RH376 compared to EL346, likely in the A’$\alpha$ helix of the LOV domain. Region
307 RH2 is an insertion of 7 amino acids in RH376 compared to EL346, located in the linker
308 between the LOV domain and $\alpha$1 of the DHp domain. These residues may be in the
309 unstructured linker or form part of an elongated $\alpha$1 helix; secondary structure prediction by
310 JPred (Drozdetskiy et al., 2015) suggests that these residues are helical (**Fig. S3**). Region RH3
311 consists of two insertions, the largest by far, of 19 and 11 amino acids in RH376 compared to
312 EL346, in the ATP lid region. This insert region is predicted to have some helical structure (**Fig.
313 S3**), unlike the ATP lid of EL346. Region RH4, located at the very C-terminus, is the only unique
314 insertion in EL346, of 10 residues. Coincidentally, RH4 corresponds almost entirely to the highly
315 dynamic C-terminal residues we deleted in an NMR construct of EL346, removal of which had
316 no effect on the monomeric state of that protein (Dikiy et al., 2019). Thus, we did not pursue
317 RH4 any further as a driver of monomer/dimer state.

318 Observing that the insert regions are either in previously reported dimerization interfaces
319 for LOV domains (RH1, A’$\alpha$ helix) (Banerjee et al., 2016; Diensthuber et al., 2013) and HKs
320 (RH2, DHp domain) (Qin et al., 2000; Tomomori et al., 1999), or large insertions with predicted
321 secondary structure elements (RH3), we hypothesized that they could contribute to the
322 dimerization of RH376 and perhaps be important for dimerization in other HKs. A homology
323 model of the RH376 sequence based on the engineered dimeric LOV-HK YF1 (Diensthuber et
324 al., 2013) illustrates the potential involvement of these residues in dimerization interfaces (**Fig.
325 5c**). We thus deleted each of the RH1, RH2, and RH3 regions from the RH376 sequence,
326 expressed the proteins, and compared them to RH376 WT in terms of LOV photocycle,
327 oligomeric state, and autophosphorylation activity.

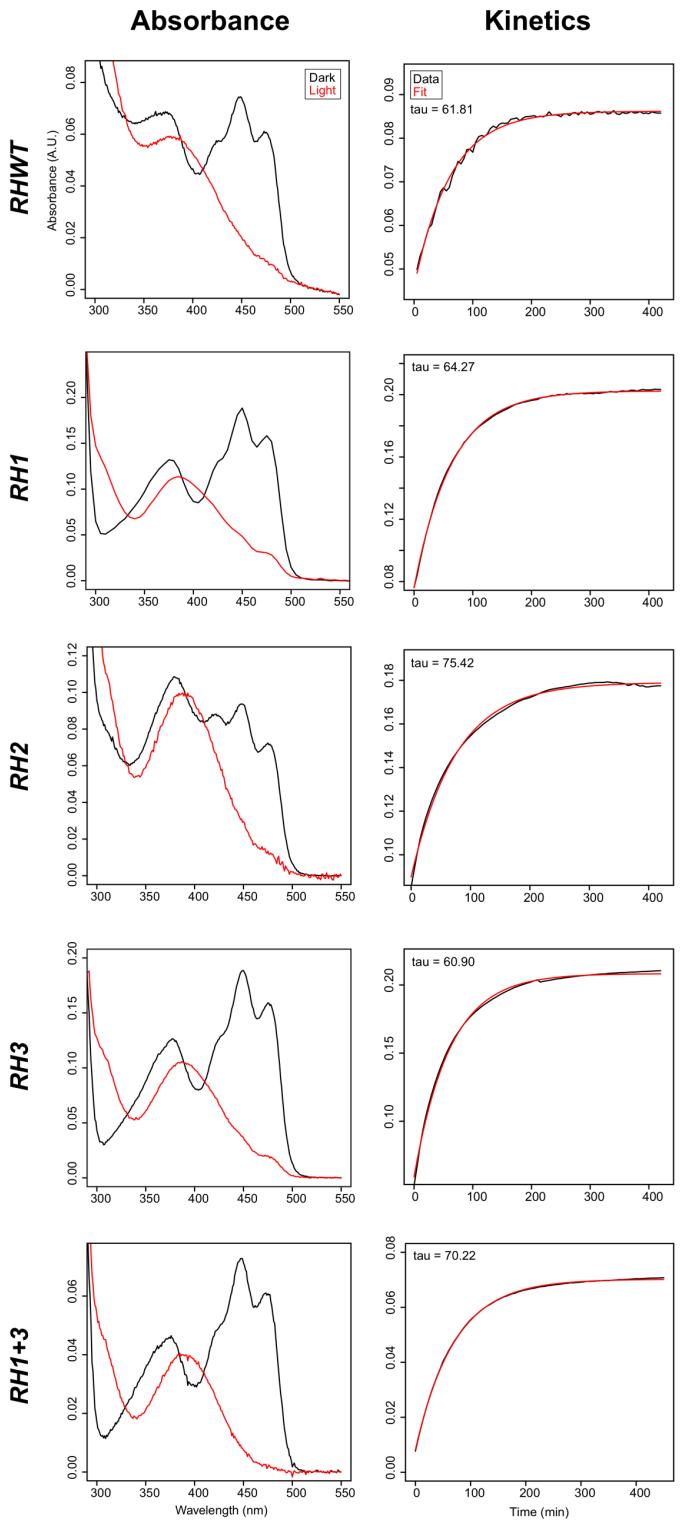


**Figure S4: RH376 mutant UV/visible absorbance and dark state reversion kinetics show that mutants are light sensitive with recovery times comparable to the WT (Supplemental to Fig. 6). a)** Absorbance of dark and light states along with **b)** dark state reversion kinetics show that RH376 WT and mutants are light sensitive. Based on tau values extracted from a single exponential decay fit, the photocycle of these mutants is comparable to the WT.

328

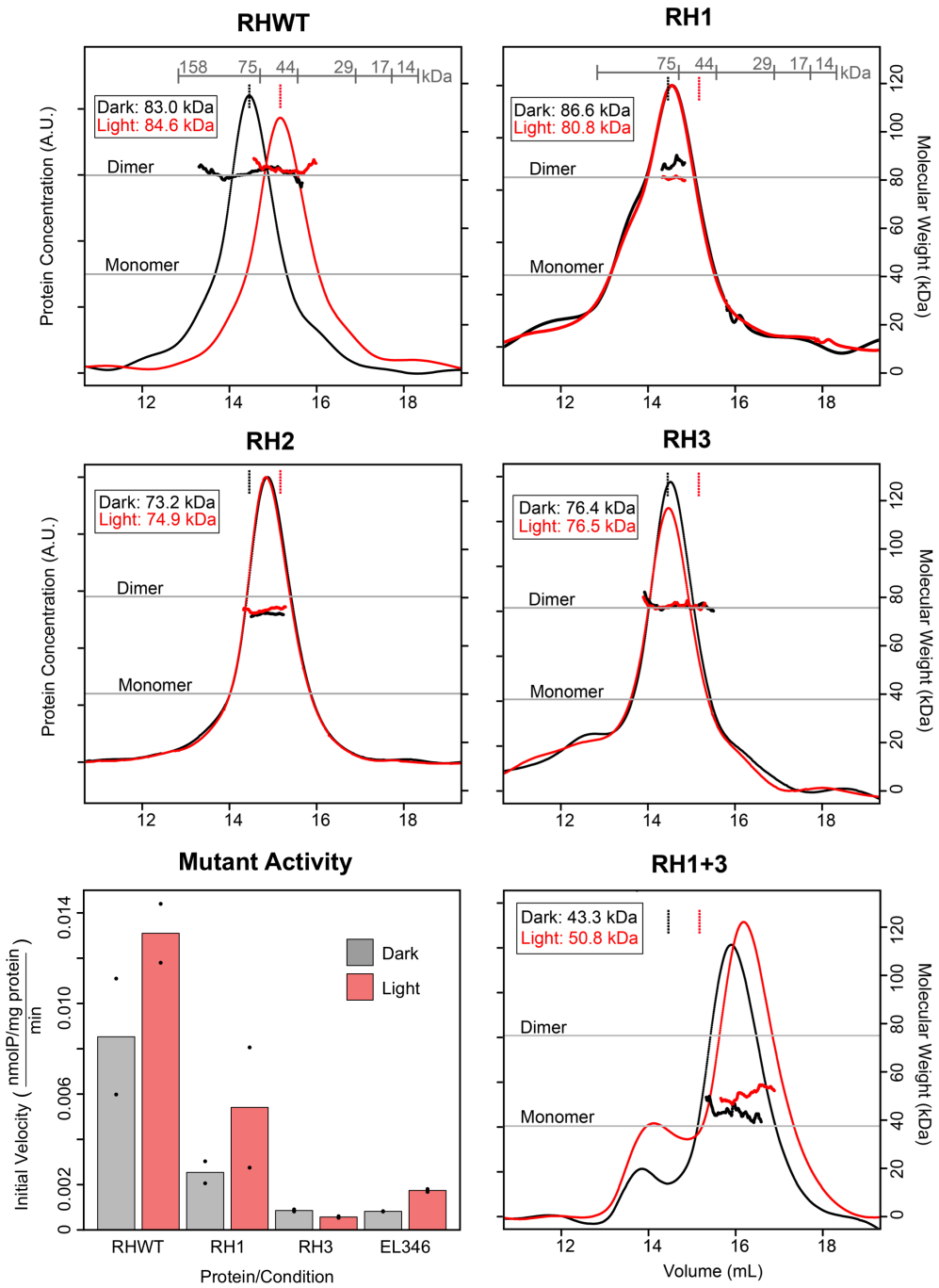


**Figure 6: RH376 mutants RH1 and RH3 are active dimers, while double mutant RH1+3 is an inactive monomer.** SEC-MALS results show that RH376 WT, RH1 and RH3 are clear dimers, while RH1+3 is primarily monomeric. The mass calculated for RH2 falls between a monomer and dimer, suggesting that the protein exists as a monomer/dimer mixture. The monomeric RH1+3 exhibits a light-dependent shift in elution peak similar to RHWT that is not present in RH1, RH2 and RH3. Expected MWs were calculated based on protein sequence. Black and red hash marks represent the position of RH376 WT D and L peaks. Mutant activity, represented by initial velocity calculated from the linear part of the curve in two trials (trials, dots; average, bars), is shown in the bottom left panel, with EL346 values for reference. No activity was detected for RH2 or RH1+3. Light activation can be seen in RH376 WT and RH1, but not in RH3.

329

330 RH1, RH2, and RH3 proteins display the characteristic LOV absorbance triplet centered
331 at 450 nm and thermal reversion to the dark state with time constants comparable to that of the
332 WT protein (**Fig. S4**), suggesting that the deletions do not perturb the folding and function of the
333 LOV domain. In RH2, we detected a small change in the relative heights of the triplet peaks,
334 indicating that flavin binding is likely adversely affected (**Fig. S4**) despite the deletion and flavin
335 binding occurring in separate domains (DHp, LOV). We determined the oligomeric state of each
336 deletion mutant by SEC-MALS and found that none of the single deletion constructs elutes as a
337 monomer regardless of dark or light conditions (**Fig. 6**). However, in each case, the apparent
338 light-induced conformational shift seen in the WT protein is lost. In addition, the measured mass
339 for RH2 is lower than that expected for a full dimer, suggesting, along with the changed elution
340 volume, that the protein samples both dimeric and monomeric conformations.

| Protein | Deletion | Oligomeric State | Tau (min) | Initial Velocity $\frac{nmol\ PO_4}{mg\ protein} min^{-1}$ | | |
|---|---|---|---|---|---|---|
| | | | | Dark | Light | Light:Dark |
| **RHWT** | N/A | Dimer | 61.8 | 8.53E-03 ± 59.9% | 1.31E-02 ± 19.8% | 1.54 |
| **RH1** | A'α | Dimer | 64.3 | 2.54E-03 ± 37.8% | 5.41E-03 ± 98.4% | 2.13 |
| **RH2** | LOV/DHp linker | Monomer/ Dimer | 75.4 | n.d. | n.d. | N/A |
| **RH3** | ATP lid | Dimer | 60.9 | 8.57E-04 ± 12.1% | 5.71E-04 ± 11.6% | 0.67 |
| **RH1+3** | A'α and ATP lid | Monomer | 70.2 | n.d. | n.d. | N/A |

**Table S1: RH376 Deletion Mutant Measurements (Supplemental to Fig. 6).** Oligomeric states were determined by SEC-MALS. Tau represents the time constant of the thermal reversion of the light state as fit to a first-order exponential decay (Fig. S4). Initial velocities were calculated from the linear region of autophosphorylation activity curves. Error is represented by the percent bias between trials of the same conditions. The ratio of light to dark initial velocities represents the degree of light activation that each protein undergoes.

341
342 We suspected that these deletions would not have a major impact on
343 autophosphorylation activity of RH376, given their modest effects on photocycle and
344 oligomerization state. As a baseline, we found that the RH376 WT protein displays robust
345 autophosphorylation activity in the dark, comparable to that of RE356, as well as a ~1.5x
346 increase in initial velocity upon illumination (**Fig. 6**). The RH1 mutant also shows
347 autophosphorylation activity, although lower than that of the WT protein, with an initial velocity in
348 the dark of about 0.3x of the WT. However, there is still an increase in the initial velocity of

349 autophosphorylation in the light. By contrast, RH2 shows no detectable autophosphorylation in
350 the dark or light (**Table S1, Fig. S5**), while RH3 shows detectable but severely attenuated
351 activity in the dark (~0.08x of the WT initial velocity) with no light-dependent increase in activity
352 (**Fig. 6**). By comparison, EL346 shows autophosphorylation activity dramatically lower than that
353 of RH376 WT, on the same order as that of RH3, but with a detectable light-mediated increase
354 of ~2x (**Fig. 6**). In the case of RH376, both the RH2 and RH3 variants dramatically decrease
355 autophosphorylation activity despite having different effects on dimerization.

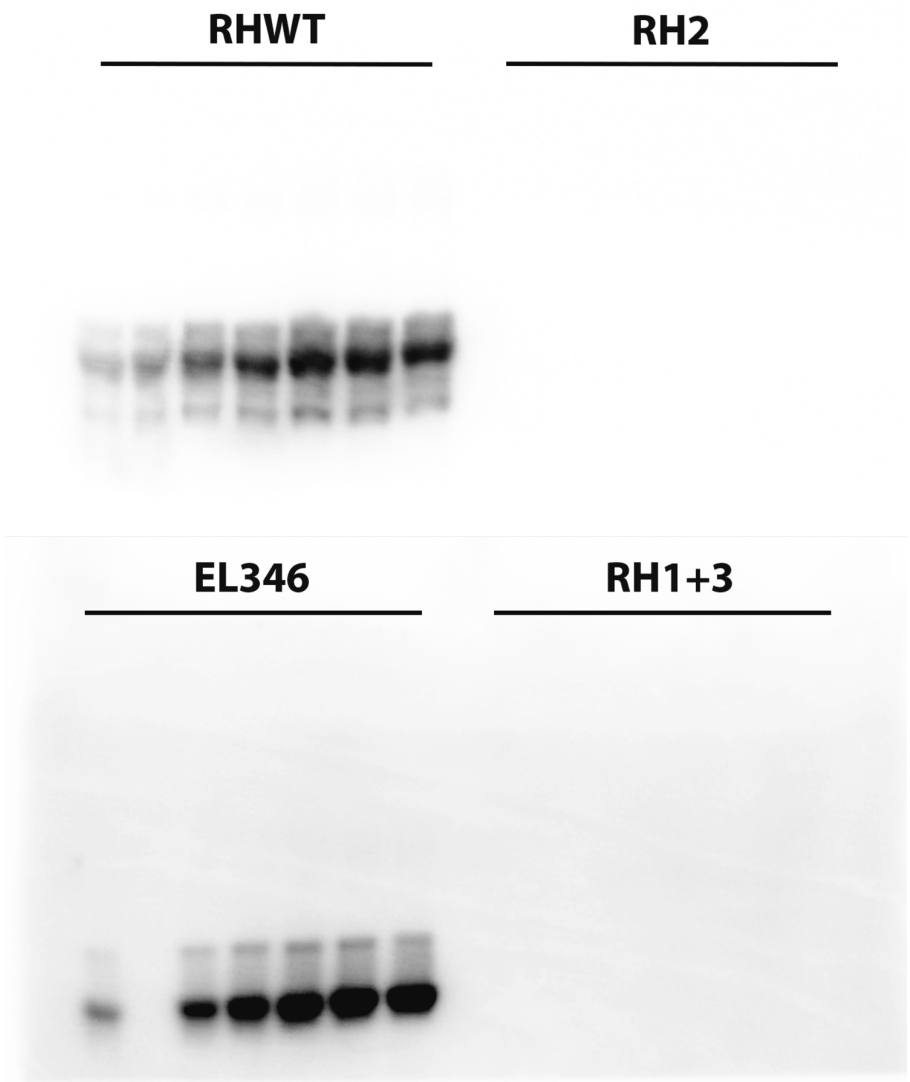


**Figure S5: RH2 and RH1+3 display no detectible autophosphorylation activity (Supplemental to Fig. 6).** The phosphor image of RH2 and RH1+3 autophosphorylation assay gels, showing no detectable signal for each of the proteins. Proteins EL346 and RH376 WT were run on the same gel and are shown here for reference.

356

357 Given that neither RH1 nor RH3 perturbed dimeric state yet had widely divergent effects
358 on activity and light-activation, we sought to combine these deletions to see whether there

359 would be a synergistic effect. We characterized the LOV absorbance spectrum and photocycle
360 of the so-called RH1+3 (**Fig. S4**) and saw no large perturbations. By contrast, SEC-MALS
361 showed a dramatic change in oligomeric state, with RH1+3 existing mainly as a monomer (**Fig.**
362 **6**), though neither deletion alone showed monomeric character. Interestingly, this monomeric
363 protein also retains the apparent compaction of the protein upon illumination, which was not
364 seen for either RH1 or RH3 separately. Additionally, we detected no autophosphorylation
365 activity in RH1+3 at conditions similar to the other proteins tested (**Table S1, Fig. S5**). Thus,
366 loss of activity seems to be correlated with RH2 and RH3 deletions, with loss of RH3 region
367 driving a loss of activity in combined mutants as well.
368 Our data suggest that neither of the segments targeted by RH1 and RH3 are sufficient
369 for dimerization, as detected by SEC-MALS. However, they are apparently both necessary, as
370 loss of both regions results in a primarily monomeric protein. For autokinase activity, all regions
371 contribute to activity, with the order of RH1<RH3<RH2 in terms of increasing necessity for
372 activity. Thus, activity and dimeric state are not entirely directly correlated, with the dimeric RH3
373 deletion lacking activity and light-activation as well as the monomeric RH1+3.
374
375 **Discussion**
376 Prior to this study, the vast majority of characterized sensor HKs were known to be
377 dimeric; we are aware of only the monomeric LOV-containing EL346 and EL362 from *E. litoralis*
378 HTCC2594 (Correa et al., 2013; Rivera-Cancel et al., 2014) as exceptions. Our work further
379 expands the behavior of these enzymes to include two proteins in a slow monomer:dimer
380 equilibrium along with three obligate dimers which exhibit a range of large-scale conformational
381 changes upon activation. Our approach to identify these novel sensor HKs relied on sequence
382 homology searches to find homologs of EL346, revealing a large family of related LOV- and
383 PAS-HKs in the *Alphaproteobacteria*. Most of these proteins were Pfam (El-Gebali et al., 2019)
384 HisKA_2 or HWE-HK-type histidine kinases, which have not been thoroughly characterized
385 structurally (Herrou et al., 2017; Rinaldi et al., 2016; Rivera-Cancel et al., 2014), and we
386 hypothesized that a subset of these sequences with the highest homology to monomeric EL346
387 might contain other monomers.
388 Contrary to this simple expectation, we found a diverse range of assemblies and
389 conformational changes among five new sensor HKs characterized here. For the LOV-HK
390 proteins, we found several dimers, including two members of the EL346-like clade of LOV-HKs
391 (RT349 and RH376) that were dimeric in SEC-MALS analyses in both light and dark conditions.
392 A newly-characterized member of the other LOV-HK clade (ME354) is also clearly dimeric, as

393 are the established EL368 and *C. crescentus* LovK (Correa et al., 2013; Purcell et al., 2010),
394 which also cluster within that second clade.

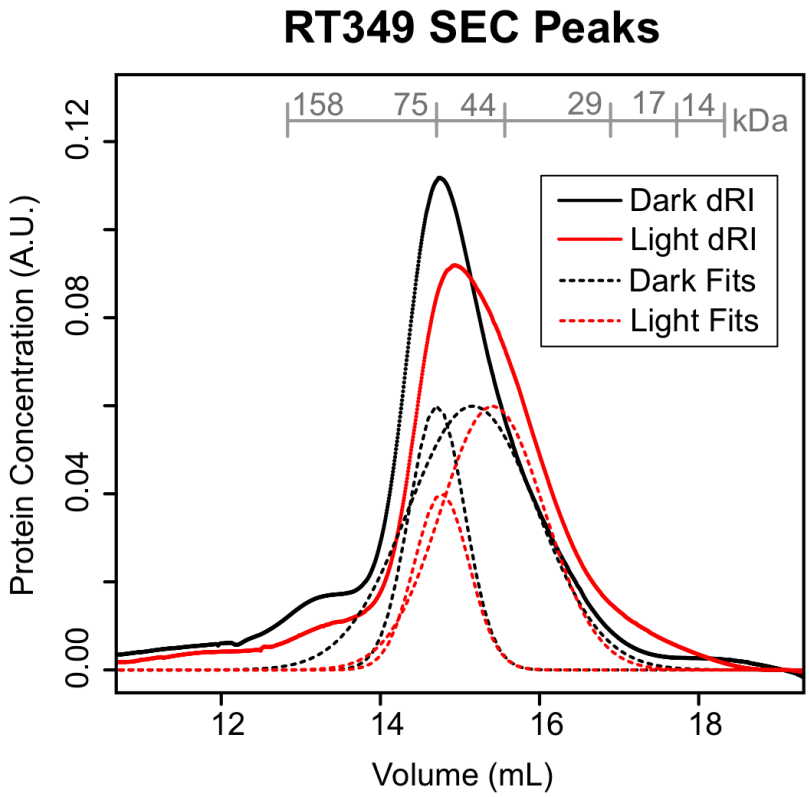


**Figure S6: Deconvolution of RT349 Superdex S200 SEC elution profile reveals two peaks (Supplemental to Fig. 6).** Two Gaussian peaks were fit to each of the asymmetrical RT349 SEC peaks. From left to right, the centers of the dark peak fits are at 14.71 and 15.16 mL, and light peak fits are at 14.77 and 15.41 mL. These two states likely correspond to globular dimer (14.71, 14.77 mL) and an expanded monomer (15.16, 15.41 mL).

395

396 While these new LOV-HKs are all dimeric, we observed an unexpected diversity of
397 large-scale structural changes among them in response to blue light activation as assessed by
398 SEC-MALS. While ME354 shows minimal change upon illumination, both RT349 and RH376
399 exhibit light-dependent shifts towards later elution volumes, suggesting that the proteins adopt
400 more compact structures in the light while not changing the apparent MW of either protein.
401 Notably, RT349 in the light appears to consist of two overlapped peaks, indicating multiple
402 conformations, one of which corresponds to a globular dimer and the other to an expanded
403 monomer state (**Fig. S6**); in contrast, RH376 appears to shift to a single, compacted structure.
404 These shifts appear to be more substantial than the light-dependent motions reported in the
405 activation of either the engineered dimeric YF1 or natural monomeric EL346 LOV-HKs
406 (Berntsson et al., 2017; Diensthuber et al., 2013; Dikiy et al., 2019; Rivera-Cancel et al., 2014).
407 Further studies are needed to more fully characterize these motions, particularly to determine

408 how they relate to the large conformational changes postulated to accompany activation from
409 crystal structures of inactive HKs (Casino et al., 2014; Mechaly et al., 2014).
410 Turning to the two PAS-HKs, both RE356 and MS367 interconvert slowly enough
411 between dimeric and monomeric states that we could separate them by size exclusion
412 chromatography. The stable PAS-HK MS367 monomer is also active in autophosphorylation,
413 albeit at a lower rate than the dimer, suggesting that shifts in the monomer/dimer equilibrium
414 could regulate sensor HK enzymatic activity, as has been demonstrated or postulated for a
415 range of signaling proteins including certain sensor HKs (Lee et al., 2008), receptor tyrosine
416 kinases (Yarden & Schlessinger, 1987), and photoreceptors (Liao et al., 2020; Sang et al.,
417 2005; Zoltowski & Crane, 2008; Zoltowski et al., 2013). Dimerization may be modulated by
418 potential ligand binding in the yet-uncharacterized PAS domains of these HKs, as in other PAS-
419 domain containing signaling proteins (Scheuermann et al., 2009), or via other mechanisms. We
420 also suggest that the relative instability of the dimer state, which forms higher-molecular-weight
421 aggregates over time, may serve as an intrinsic timer on the more active signaling species to
422 temporally limit a biological response post-activation.
423 Given the integral role of dimerization in this possible regulation, along with its obvious
424 necessity in *trans*-phosphorylation (Ashenberg et al., 2013) within HKs, we sought to
425 understand the sequence determinants of dimerization with an engineering approach. Using
426 RH376 as a model dimer, we systematically removed combinations of the three insertions it
427 contains compared to the monomeric EL346 (**Figs 5 and S3**), hypothesizing that one or more of
428 these changes could create a monomeric LOV-HK. While each deletion mutant was well-folded
429 as evidenced by binding flavin chromophores and undergoing canonical LOV photochemistry,
430 we found that all three single deletion mutants (RH1, RH2, RH3) remained largely dimeric
431 despite the removal of 5-30 amino acid residues apiece. Each, however, lost the light-induced
432 compaction seen in SEC-MALS analysis of the wildtype protein. Remarkably, the combination of
433 two of these deletions to generate the RH1+3 construct both monomerized the protein and
434 restored the light-induced conformational shift seen in the WT parent sequence. From these
435 studies, we arrive at two general conclusions: 1) multiple regions determine oligomeric states in
436 this family of sensor HKs, and 2) light-induced conformational changes have a complex
437 relationship with autophosphorylation activity and do not appear to be restricted to dimeric or
438 monomeric states.
439 The first of these conclusions, the use of distributed interfaces for dimerization, is
440 consistent with structural studies on the dimeric YF1 engineered sensor kinase (Diensthuber et
441 al., 2013) showing intermonomer contacts via residues in both the LOV domain A'$\alpha$ helix and

442 DHp domain, which are typical dimerization interfaces for LOV photosensors (Diensthuber et al.,
443 2013; Key et al., 2007) and canonical HisKA histidine kinases (Ashenberg et al., 2013; Dutta et
444 al., 1999) respectively. These correspond to our RH1 and RH2 deletions; neither of these
445 changes, nor the RH3 deletion in a large loop near the kinase ATP binding site, were sufficient
446 on its own to generate monomers, but each had markedly different effects on
447 autophosphorylation. These data suggest different types of dimers may be generated by these
448 deletions, some of which retain function better than others. Of surprise to us, the RH3 deletion –
449 which removes a large loop near the ATP binding site – affected dimerization when paired with
450 RH1. While this site has not been previously described as being involved in dimerization to the
451 best of our knowledge, we note that ATP itself bridges interactions between multiple domains in
452 both known HWE/HisKA2 HK crystal structures (Rinaldi et al., 2016; Rivera-Cancel et al., 2014),
453 laying some precedent for interactions at or near the nucleotide to be involved in dimerization.
454 Clearly, further structural work is needed to more fully assess how dimerization is encoded
455 among multiple HKs.
456 In addition, our work underscores the diversity of structural changes which occur as part
457 of the functional activation process among HKs, another long-discussed topic within the
458 community. Considering the low-resolution conformational changes we observe by SEC-MALS
459 upon photoactivation of the LOV-HKs tested here, we see a wide span of responses. It is clear
460 from the detailed structures available among HKs, chiefly on inactive forms, that substantial
461 rearrangements must be required to let catalytic domains transit long distances to
462 phosphorylate substrate histidines within DHp domains (Diensthuber et al., 2013; Marina et al.,
463 2005; Rivera-Cancel et al., 2014). Solution integrated structural biology approaches are
464 beginning to reveal such changes with fewer experimental caveats than crystallographic
465 approaches (Berntsson et al., 2017; Dikiy et al., 2019; Ghosh et al., 2019) but this essential
466 process still remains undercharacterized at present.
467 The diversity of oligomeric states in a family of sensor HKs that we uncover here opens
468 the door to novel regulatory modes within this family, such as control of dimerization, whether by
469 ligand or inherent thermodynamic stability, or cooperativity. The closely related bacterial
470 chemoreceptor proteins, for instance, form large arrays in order to increase signaling output
471 through cooperativity (Muok et al., 2020). Similar use of monomer/dimer transitions and
472 evolution of variable quaternary structure is at the heart of acquiring cooperativity and novel
473 regulatory modes in the classic hemoglobin family of oxygen-binding proteins, which started as
474 ancestral monomers prior to dimerizing by gene duplication events and the acquisition of a
475 relatively small number of mutations (Pillai et al., 2020). Our findings suggest that different

oligomeric states within the sensor HK family could allow for different potential for cooperativity, perhaps depending on the input signal or output pathway, or integrating multiple points of control or regulation to be added to the canonical signal transduction pathway of two-component systems.

**Acknowledgements**

The authors would like to acknowledge Bright Shi for generating the RH1+3 construct, as well as Matthew Cleere, Zaynab Jaber, Jaynee Hart, and other Gardner lab members for fruitful discussions. This work was supported by NIH grants R01 GM106239 (to K.H.G.) and F32 GM1119311 (to I.D.).

**Competing interests**

The authors declare no competing interests.

**References**

Ashenberg, O., Keating, A. E., & Laub, M. T. (2013). Helix bundle loops determine whether histidine kinases autophosphorylate in cis or in trans. *J Mol Biol, 425*(7), 1198-1209. doi:10.1016/j.jmb.2013.01.011

Banerjee, A., Herman, E., Kottke, T., & Essen, L. O. (2016). Structure of a Native-like Aureochrome 1a LOV Domain Dimer from Phaeodactylum tricornutum. *Structure, 24*(1), 171-178. doi:10.1016/j.str.2015.10.022

Berntsson, O., Diensthuber, R. P., Panman, M. R., Bjorling, A., Gustavsson, E., Hoernke, M., . . . Westenhoff, S. (2017). Sequential conformational transitions and alpha-helical supercoiling regulate a sensor histidine kinase. *Nat Commun, 8*(1), 284. doi:10.1038/s41467-017-00300-5

Bhate, M. P., Molnar, K. S., Goulian, M., & DeGrado, W. F. (2015). Signal transduction in histidine kinases: insights from new structures. *Structure, 23*(6), 981-994. doi:10.1016/j.str.2015.04.002

Casino, P., Miguel-Romero, L., & Marina, A. (2014). Visualizing autophosphorylation in histidine kinases. *Nat Commun, 5*, 3258. doi:10.1038/ncomms4258

Correa, F., Ko, W. H., Ocasio, V., Bogomolni, R. A., & Gardner, K. H. (2013). Blue light regulated two-component systems: enzymatic and functional analyses of light-oxygen-voltage (LOV)-histidine kinases and downstream response regulators. *Biochemistry, 52*(27), 4656-4666. doi:10.1021/bi400617y

Crosson, S., & Moffat, K. (2001). Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. *Proc Natl Acad Sci U S A, 98*(6), 2995-3000. doi:10.1073/pnas.051520298

Crosson, S., Rajagopal, S., & Moffat, K. (2003). The LOV domain family: photoresponsive signaling modules coupled to diverse output domains. *Biochemistry, 42*(1), 2-10. doi:10.1021/bi026978l

Diensthuber, R. P., Bommer, M., Gleichmann, T., & Moglich, A. (2013). Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators. *Structure, 21*(7), 1127-1136. doi:10.1016/j.str.2013.04.024

Dikiy, I., Edupuganti, U. R., Abzalimov, R. R., Borbat, P. P., Srivastava, M., Freed, J. H., & Gardner, K. H. (2019). Insights into histidine kinase activation mechanisms from the monomeric blue light sensor EL346. *Proc Natl Acad Sci U S A, 116*(11), 4963-4972. doi:10.1073/pnas.1813586116

Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res, 43*(W1), W389-394. doi:10.1093/nar/gkv332

Dutta, R., Qin, L., & Inouye, M. (1999). Histidine kinases: diversity of domain organization. *Mol Microbiol, 34*(4), 633-640. doi:10.1046/j.1365-2958.1999.01646.x

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res, 32*(5), 1792-1797. doi:10.1093/nar/gkh340

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., . . . Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res, 47*(D1), D427-D432. doi:10.1093/nar/gky995

Galperin, M. Y. (2018). What bacteria want. *Environ Microbiol, 20*(12), 4221-4229. doi:10.1111/1462-2920.14398

Gao, R., & Stock, A. M. (2009). Biological insights from structures of two-component proteins. *Annu Rev Microbiol, 63*, 133-154. doi:10.1146/annurev.micro.091208.073214

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In J. M. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). Totowa, NJ: Humana Press.

Ghosh, M., Wang, L. C., Huber, R. G., Gao, Y., Morgan, L. K., Tulsian, N. K., . . . Anand, G. S. (2019). Engineering an Osmosensor by Pivotal Histidine Positioning within Disordered Helices. *Structure, 27*(2), 302-314 e304. doi:10.1016/j.str.2018.10.012

544 Gushchin, I., & Gordeliy, V. (2018). Transmembrane Signal Transduction in Two-Component
545 Systems: Piston, Scissoring, or Helical Rotation? *BioEssays : news and reviews in*
546 *molecular, cellular and developmental biology, 40*(2). doi:10.1002/bies.201700197
547 Herrou, J., Crosson, S., & Fiebig, A. (2017). Structure and function of HWE/HisKA2-family
548 sensor histidine kinases. *Curr Opin Microbiol, 36*, 47-54. doi:10.1016/j.mib.2017.01.008
549 Karniol, B., & Vierstra, R. D. (2004). The HWE histidine kinases, a new family of bacterial two-
550 component sensor kinases with potentially diverse roles in environmental signaling. *J*
551 *Bacteriol, 186*(2), 445-453. doi:10.1128/jb.186.2.445-453.2004
552 Kenney, L. J., & Anand, G. S. (2020). EnvZ/OmpR Two-Component Signaling: An Archetype
553 System That Can Function Noncanonically. *EcoSal Plus, 9*(1).
554 doi:10.1128/ecosalplus.ESP-0001-2019
555 Key, J., Hefti, M., Purcell, E. B., & Moffat, K. (2007). Structure of the redox sensor domain of
556 Azotobacter vinelandii NifL at atomic resolution: signaling, dimerization, and mechanism.
557 *Biochemistry, 46*(12), 3614-3623. doi:10.1021/bi0620407
558 Lee, J., Tomchick, D. R., Brautigam, C. A., Machius, M., Kort, R., Hellingwerf, K. J., & Gardner,
559 K. H. (2008). Changes at the KinA PAS-A dimerization interface influence histidine
560 kinase function. *Biochemistry, 47*(13), 4051-4064. doi:10.1021/bi7021156
561 Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new
562 developments. *Nucleic Acids Res, 47*(W1), W256-W259. doi:10.1093/nar/gkz239
563 Liao, X., Liu, W., Yang, H. Q., & Jenkins, G. I. (2020). A dynamic model of UVR8 photoreceptor
564 signalling in UV-B-acclimated Arabidopsis. *New Phytol, 227*(3), 857-866.
565 doi:10.1111/nph.16581
566 Luo, B. H., Carman, C. V., & Springer, T. A. (2007). Structural basis of integrin regulation and
567 signaling. *Annu Rev Immunol, 25*, 619-647.
568 doi:10.1146/annurev.immunol.25.022106.141618
569 Marina, A., Waldburger, C. D., & Hendrickson, W. A. (2005). Structure of the entire cytoplasmic
570 portion of a sensor histidine-kinase protein. *Embo j, 24*(24), 4247-4259.
571 doi:10.1038/sj.emboj.7600886
572 Mechaly, A. E., Sassoon, N., Betton, J. M., & Alzari, P. M. (2014). Segmental helical motions
573 and dynamical asymmetry modulate histidine kinase autophosphorylation. *PLoS Biol,*
574 *12*(1), e1001776. doi:10.1371/journal.pbio.1001776
575 Mechaly, A. E., Soto Diaz, S., Sassoon, N., Buschiazzo, A., Betton, J. M., & Alzari, P. M. (2017).
576 Structural Coupling between Autokinase and Phosphotransferase Reactions in a
577 Bacterial Histidine Kinase. *Structure, 25*(6), 939-944 e933. doi:10.1016/j.str.2017.04.011

Möglich, A. (2019). Signal transduction in photoreceptor histidine kinases. *Protein Sci, 28*(11), 1923-1946. doi:10.1002/pro.3705

Muok, A. R., Briegel, A., & Crane, B. R. (2020). Regulation of the chemotaxis histidine kinase CheA: A structural perspective. *Biochim Biophys Acta Biomembr, 1862*(1), 183030. doi:10.1016/j.bbamem.2019.183030

Perica, T., Chothia, C., & Teichmann, S. A. (2012). Evolution of oligomeric state through geometric coupling of protein interfaces. *Proc Natl Acad Sci U S A, 109*(21), 8127-8132. doi:10.1073/pnas.1120028109

Pillai, A. S., Chandler, S. A., Liu, Y., Signore, A. V., Cortez-Romero, C. R., Benesch, J. L. P., . . . Thornton, J. W. (2020). Origin of complexity in haemoglobin evolution. *Nature, 581*(7809), 480-485. doi:10.1038/s41586-020-2292-y

Pudasaini, A., El-Arab, K. K., & Zoltowski, B. D. (2015). LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front Mol Biosci, 2*, 18. doi:10.3389/fmolb.2015.00018

Purcell, E. B., McDonald, C. A., Palfey, B. A., & Crosson, S. (2010). An analysis of the solution structure and signaling mechanism of LovK, a sensor histidine kinase integrating light and redox signals. *Biochemistry, 49*(31), 6761-6770. doi:10.1021/bi1006404

Qin, L., Dutta, R., Kurokawa, H., Ikura, M., & Inouye, M. (2000). A monomeric histidine kinase derived from EnvZ, an Escherichia coli osmosensor. *Mol Microbiol, 36*(1), 24-32. doi:10.1046/j.1365-2958.2000.01837.x

Rinaldi, J., Arrar, M., Sycz, G., Cerutti, M. L., Berguer, P. M., Paris, G., . . . Goldbaum, F. A. (2016). Structural Insights into the HWE Histidine Kinase Family: The Brucella Blue Light-Activated Histidine Kinase Domain. *J Mol Biol, 428*(6), 1165-1179. doi:10.1016/j.jmb.2016.01.026

Rivera-Cancel, G., Ko, W. H., Tomchick, D. R., Correa, F., & Gardner, K. H. (2014). Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. *Proc Natl Acad Sci U S A, 111*(50), 17839-17844. doi:10.1073/pnas.1413983111

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics, 18*(1), 529. doi:10.1186/s12859-017-1934-z

Sang, Y., Li, Q. H., Rubio, V., Zhang, Y. C., Mao, J., Deng, X. W., & Yang, H. Q. (2005). N-terminal domain-mediated homodimerization is required for photoreceptor activity of Arabidopsis CRYPTOCHROME 1. *Plant Cell, 17*(5), 1569-1584. doi:10.1105/tpc.104.029645

612 Scheuermann, T. H., Tomchick, D. R., Machius, M., Guo, Y., Bruick, R. K., & Gardner, K. H.
613 (2009). Artificial ligand binding within the HIF2alpha PAS-B domain of the HIF2
614 transcription factor. *Proc Natl Acad Sci U S A, 106*(2), 450-455.
615 doi:10.1073/pnas.0808092106
616 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . .
617 Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nat*
618 *Methods, 9*(7), 676-682. doi:10.1038/nmeth.2019
619 Sheffield, P., Garrard, S., & Derewenda, Z. (1999). Overcoming expression and purification
620 problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr Purif,*
621 *15*(1), 34-39. doi:10.1006/prep.1998.1003
622 Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011).
623 Fast, scalable generation of high-quality protein multiple sequence alignments using
624 Clustal Omega. *Molecular systems biology, 7*, 539. doi:10.1038/msb.2011.75
625 Staroń, A., & Mascher, T. (2010). General stress response in alpha-proteobacteria: PhyR and
626 beyond. *Mol Microbiol, 78*(2), 271-277. doi:10.1111/j.1365-2958.2010.07336.x
627 Swartz, T. E., Tseng, T. S., Frederickson, M. A., Paris, G., Comerci, D. J., Rajashekara, G., . . .
628 Bogomolni, R. A. (2007). Blue-light-activated histidine kinases: two-component sensors
629 in bacteria. *Science, 317*(5841), 1090-1093. doi:10.1126/science.1144306
630 Szurmant, H., White, R. A., & Hoch, J. A. (2007). Sensor complexes regulating two-component
631 signal transduction. *Curr Opin Struct Biol, 17*(6), 706-715. doi:10.1016/j.sbi.2007.08.019
632 Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S. K., Zhu, Y., . . . Ikura, M. (1999).
633 Solution structure of the homodimeric core domain of Escherichia coli histidine kinase
634 EnvZ. *Nat Struct Biol, 6*(8), 729-734. doi:10.1038/11495
635 Tseng, T. S., Frederickson, M. A., Briggs, W. R., & Bogomolni, R. A. (2010). Light-activated
636 bacterial LOV-domain histidine kinases. *Methods Enzymol, 471*, 125-134.
637 doi:10.1016/S0076-6879(10)71008-9
638 Wang, S. (2012). Bacterial Two-Component Systems: Structures and Signaling Mechanisms. In
639 C. Huang (Ed.), *Protein Phosphorylation in Human Health*: InTech.
640 West, A. H., & Stock, A. M. (2001). Histidine kinases and response regulator proteins in two-
641 component signaling systems. *Trends Biochem Sci, 26*(6), 369-376. doi:10.1016/s0968-
642 0004(01)01852-7
643 Yarden, Y., & Schlessinger, J. (1987). Epidermal growth factor induces rapid, reversible
644 aggregation of the purified epidermal growth factor receptor. *Biochemistry, 26*(5), 1443-
645 1451. doi:10.1021/bi00379a035

646 Zoltowski, B. D., & Crane, B. R. (2008). Light activation of the LOV protein vivid generates a
647 rapidly exchanging dimer. *Biochemistry, 47*(27), 7012-7019. doi:10.1021/bi8007017
648 Zoltowski, B. D., & Gardner, K. H. (2011). Tripping the light fantastic: blue-light photoreceptors
649 as examples of environmentally modulated protein-protein interactions. *Biochemistry,*
650 *50*(1), 4-16. doi:10.1021/bi101665s
651 Zoltowski, B. D., Motta-Mena, L. B., & Gardner, K. H. (2013). Blue light-induced dimerization of
652 a bacterial LOV-HTH DNA-binding protein. *Biochemistry, 52*(38), 6653-6661.
653 doi:10.1021/bi401040m
654