Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling

Jihe Liu[†], **James Hemphill[†]**, **Subhas Samanta[†]**, **Michael Tsang[‡]**, and **Alexander Deiters[*,†]**

[†]Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States

[‡]Department of Developmental Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States

Abstract

Site-specific incorporation of unnatural amino acids into proteins provides a powerful tool to study protein function. Here we report genetic code expansion in zebrafish embryos and its application to the optogenetic control of cell signaling. We genetically encoded four unnatural amino acids with a diverse set of functional groups, which included a photocaged lysine that was applied to the light-activation of luciferase and kinase activity. This approach enables versatile manipulation of protein function in live zebrafish embryos, a transparent and commonly used model organism to study embryonic development.

While early examples of the incorporation of unnatural amino acids into proteins and their application in cell-free and cell-based environments involved the synthesis of chemically misacylated tRNAs,[1] recent genetic code expansion approaches through the addition of engineered, orthogonal tRNA/tRNA synthetase pairs to the endogenous protein-biosynthetic machinery provide powerful tools for the study and manipulation of protein function in biological systems.[2] For example, photocaged amino acids enable precise regulation of protein function with light,[3] amino acids bearing bio-orthogonal chemical handles allow for selective protein labeling and imaging in living cells,[4] and amino acids containing biophysical probes facilitate studies of protein structure and function.[2b] In order to expand the genetic code, an orthogonal aminoacyl tRNA synthetase/tRNA$_{CUA}$ pair is added to the biosynthetic machinery of cells to incorporate an unnatural amino acid at a desired position of a protein in response to an amber stop codon.[2] The PylRS system from *Methanosarcina* species (*M. barkeri* or *M. mazei*) is a versatile tool for genetic code expansion.[5] It requires a pyrrolysyl tRNA synthetase (PylRS) and its cognate tRNA$^{Pyl}_{CUA}$ (PylT), and it shows excellent orthogonality to the endogenous tRNA/tRNA synthetase pairs in bacterial, yeast, and mammalian cells. It is a natural amber suppressor, and the synthetase active-site has a

large hydrophobic pocket that can be engineered to accept a wide range of substrates.[6] While site-specific incorporation of unnatural amino acids has been performed in metazoans, [7] previous experiments were limited to reporter genes and no expression of functional proteins that affect animal physiology has been reported. Here, we describe the genetic code expansion of zebrafish embryos with four unnatural amino acids, and its application to the optical control of protein function in live animals.

Zebrafish are a commonly employed model organism for vertebrate development,[8] disease modeling,[9] and drug discovery.[10] The *ex vivo* development and transparency of the embryo make it an excellent system for the application of noninvasive optical tools, including light-activated antisense agents,[11] thereby providing insight into gene regulatory processes and networks with spatial and temporal resolution. Moreover, microinjection of mRNA into the 1-cell stage embryo is a standard and rapid approach for delivery of exogenous genes that can be readily adapted to encode for any gene product and provides homogeneous protein expression in zebrafish.[12] Taken together, these distinct advantages over other model organisms make the zebrafish an ideal system for a wide range of biological studies.[13]

Light regulation of protein activity in zebrafish has been reported using natural photoreceptor domains;[14] however, the genetic encoding of photocaged amino acids will further expand the optogenetic toolbox and will enable a rational design of light-activated proteins based on their function or structure. For example, photocaged lysine analogs have been applied to the optical control of protein localization,[15] kinase function,[16] and CRISPR/Cas9 gene editing[17] in human cells. By genetically encoding a photocaged lysine using the PylRS system in zebrafish embryos, we demonstrate the consequences of optical control of MEK activation at different stages in development. Temporal control of kinase function led to the identification of a critical time window for activity of the MEK/ERK pathway in order to establish dorsal/ventral polarity in the early embryo.

We first tested incorporation of unnatural amino acids using a *Renilla* luciferase (Rluc) reporter assay. Wild-type Rluc was active in zebrafish embryos (Figure S1) and was used as a positive control. We predicted that incorporation of an unnatural amino acid at a leucine residue (L95) located at the surface of Rluc would not interfere with Rluc function, thereby generating a highly specific reporter for amber codon suppression (Figure S2). Thus, L95 was mutated to a TAG codon to probe read-through during translation. Wild-type *M. barkeri* PylRS (WTRS) mRNA, Rluc-L95TAG mRNA, and PylT RNA were synthesized through *in vitro* transcription. The purity of PylT was confirmed by agarose gel (Figure S3). WTRS mRNA, Rluc-L95TAG mRNA, and PylT RNA were injected together with the unnatural amino acid (UAA) into zebrafish embryos. After 48 h, zebrafish lysate was collected for luciferase assays and a 217- and 161-fold increase of Rluc activity was observed in the presence of the UAAs **1** and **2**, respectively (Figure 1c). Negligible Rluc activity in the absence of the UAA demonstrated the excellent fidelity of the PylRS system in zebrafish embryos, as none of the common 20 amino acids were recognized as substrates. The effect of PylT on incorporation efficiency was further explored, and we found that chemically synthesized PylT showed similar efficacy compared to *in vitro* transcribed PylT (Figure S4). However, when we tested *in vitro* transcribed PylT without a CCA tail, significantly lower efficacy was noted (Figure S4). The CCA tail is a conserved sequence at the 3′ end of the

tRNA, which is acylated with the amino acid. Although this sequence can be added by CCA-adding enzymes in cells,[18] our results suggested that the direct addition of a CCA tail on the PylT improved incorporation efficiency, and thus was used in all subsequent experiments. We also explored the effect of different PylT concentrations on incorporation efficiency, and 2 ng were found to be critical for high incorporation efficiency (Figure S5). This is in agreement with previous findings of the PylT amount possibly being a limiting factor for genetic encoding of UAAs.[19]

Inspired by the success of using wild-type PylRS for incorporating **1** and **2**, which could be applied in both protein labeling and protein activation experiments,[20] we examined if mutant PylRS enzymes can be employed in embryos to incorporate more structurally complex amino acids. To this end, we synthesized mRNA of HCKRS[21] and OABKRS,[22] which have been shown to incorporate **3** and **4**, respectively. We performed injections as described above, and observed a 70- and 34-fold increase of Rluc activity in the presence of **3** and **4**, respectively (Figure 1c). No toxicity was observed for any of the four UAAs (Figure S6). Taken together, these results demonstrate successful genetic encoding of four different UAAs in zebrafish embryos.

The photocaged lysine **3** has previously been applied to control protein function in mammalian cells using 365, 405, and 760 nm (two-photon) irradiation.[21] With successful genetic encoding of **3** in zebrafish, we tested if protein function could be manipulated with light in developing embryos. As an initial proof-of-concept, we utilized firefly luciferase (Fluc) with a TAG amber codon at position lysine 206, because installation of **3** blocks Fluc activity until light exposure.[21] In order to create an internal control for incorporation efficiency, we fused Rluc to the C-terminus of Fluc-K206TAG (Figure 1d). To this end, Fluc-K206TAG-Rluc mRNA was injected, together with HCKRS mRNA, PylT, and **3**, into zebrafish embryos. After 48 h, injected embryos were either briefly irradiated at 365 nm or kept in the dark. Embryo lysate was subsequently collected, and a luciferase assay was performed for both Fluc and Rluc. Excellent optical OFF to ON switching of Fluc function was observed, with negligible background activity before irradiation (Figure 1e). Normalization of Fluc activity to Rluc activity, as a TAG codon suppression control, revealed a 26-fold increase of Fluc activity upon light triggering. This result shows that light activation of protein function can be achieved in live zebrafish embryos with an expanded genetic code.

We then sought to apply genetic code expansion in zebrafish to an enzyme with endogenous function in order to demonstrate its utility in altering embryonic development. Incorporation of photocaged amino acids into proteins enables precise dissection of signaling pathways with light, and this approach has been applied to study the dynamics of MEK/ERK signaling in mammalian cells.[16] While optical activation on the second to minute time scale in mammalian cells provided further insight into adaptive behavior of the MEK/ERK network in single cells, in the context of zebrafish biology, kinase signaling pathways are important regulators throughout embryogenesis.[23] The MEK/ERK pathway is a well-known downstream target of Fibroblast Growth Factor (FGF) signaling and plays an important role in mesendoderm induction and dorsoventral patterning of the zebrafish embryo. FGF signaling induces expression of chordin and noggin, secreted inhibitors of ventralizing bone

morphogenetic proteins resulting in dorsalization (Figure 2a).[24] An inhibitor-based chemical approach has previously been used for perturbation of the MEK/ERK pathway during zebrafish development;[25] however, pharmacological inhibitors only allow for the deactivation of kinase function, not activation, and their specificity is often limited. We reasoned that optical activation of the MEK/ERK pathway in zebrafish provides an innovative tool to study its role, as site-specific incorporation of the caged amino acid **3** conveys complete kinase specificity. By substituting the critical lysine 97 with **3**, the caging group blocks the ability of the enzyme to correctly position ATP in the MEK1 active site (Figure 2b). We first confirmed the incorporation of **3** into MEK1, and subsequent decaging through UV exposure, by MS/MS analysis of recombinantly expressed MEK1 protein in *E. coli* (Figure S7 and Figure S8). We then generated mRNA of constitutively active MEK1 (caMEK1, containing S218D and S222D mutations) and confirmed that injection of caMEK1 led to dorsalized embryos at 10 hpf (Figure 2d,e), as previously reported.[25] We further generated caMEK1-K97TAG mRNA and injected it into zebrafish embryos, together with HCKRS mRNA, PylT, and **3**. We detected full-length MEK1 by Western blot in the presence of **3** but not in the absence of **3**, suggesting successful incorporation into caMEK1 at position K97 and generation of the photocaged enzyme in live animals (Figure S9). When these embryos were left in the dark, they developed normally, indicating that caged MEK1 was inactive (Figure 2d). To activate caged MEK1 at different developmental stages, we irradiated embryos for 30 s at 2, 5, or 8 h postinjection. Light activation of MEK1 can efficiently increase ERK phosphorylation at all three time points (Figure 2c). Embryos irradiated at 2 and 5 h showed an elongated phenotype at 10 hpf (Figure 2d,e). However, the majority of embryos irradiated at 8 h appeared normal at 10 hpf (Figure 2e), indicating that active MEK was not able to efficiently trigger an elongated phenotype after 8 hpf. As a control, embryos that were injected with caMEK1-K97TAG mRNA, HCKRS mRNA, and PylT, but not **3**, developed normally in both the presence and absence of UV irradiation (Figure S10).

We then tested if optical activation of the MEK/ERK pathway resulted in a change at the gene expression level. We probed expression of the *brachyuary homolog a* (*ta*) gene, a well-known downstream target of the FGF/MEK/ERK pathway.[26] At shield stage (6 hpf), embryos that were exposed to UV light showed broader expression of *ta* in the margin when compared to embryos that were kept in the dark (Figure 3a). In some instances, *ta* expression was detected at the animal pole of the embryos, a pattern that is similar to embryos injected with constitutively active MEK1. At bud stage (~10 hpf), the expression of the *ta* was also wider along the notochord in light-activated embryos compared to embryos that were not irradiated (Figure 3b). We also probed expression of the *chordin* (*chd*) gene, a marker for dorsalized embryos that is known to be induced following activation of the FGF/Ras/MAPK pathway.[27] As expected, embryos that were exposed to UV light showed expanded expression of *chd* at shield stage, compared to embryos that were kept in the dark (Figure S11).

Taken together, the observed *ta* and *chd* expression patterns in response to optical MEK1 activation and the time-resolved phenotypic studies shown in Figure 2d–e demonstrate that the MEK/ERK pathway influences dorsal/ventral patterning in zebrafish development before

8 hpf, thereby providing support for early intervention with pharmacological MEK inhibitors for related congenital defects in humans, such as cardio-faciocutaneous syndrome.[25,28]

In conclusion, we incorporated four unnatural amino acids into proteins in zebrafish embryos through genetic code expansion using injection methods that are applicable to many zebrafish studies. We demonstrated light activation of enzymatic function, specifically luciferase activity, through site-specific incorporation of a photocaged unnatural amino acid in live embryos. We then applied this methodology to the temporal activation of the MEK/ERK pathway in zebrafish and identified a time window for MEK activity that can influence dorsoventral patterning. Besides controlling protein function with light, other potential applications of unnatural amino acids in live zebrafish embryos include small molecule triggered protein activation, site-specific labeling of proteins with fluorescent and biophysical probes, and probing protein interactions through covalent bond formation with electrophilic or photo-cross-linking groups. The zebrafish is a well-established model organism for human development and disease, and we anticipate that the ability to genetically encode a 21st amino acid will become a powerful tool to manipulate and study protein function in animals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

(1). (a) Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. Science. 1989; 244:182–188. [PubMed: 2649980] (b) Nowak MW, Kearney PC, Sampson JR, Saks ME, Labarca CG, Silverman SK, Zhong W, Thorson J, Abelson JN, Davidson N. Science. 1995; 268:439–42. [PubMed: 7716551]

(2). (a) Chin JW. Annu. Rev. Biochem. 2014; 83:379–408. [PubMed: 24555827] (b) Liu CC, Schultz PG. Annu. Rev. Biochem. 2010; 79:413–44. [PubMed: 20307192]

(3). Baker AS, Deiters A. ACS Chem. Biol. 2014; 9:1398–407. [PubMed: 24819585]

(4). Lang K, Chin JW. Chem. Rev. 2014; 114:4764–806. [PubMed: 24655057]

(5). Wan W, Tharp JM, Liu WR. Biochim. Biophys. Acta, Proteins Proteomics. 2014; 1844:1059–70.

(6). (a) Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Green-Church KB, Chan MK, Krzycki JA. Nature. 2004; 431:333–5. [PubMed: 15329732] (b) Neumann H, Peak-Chew SY, Chin JW. Nat. Chem. Biol. 2008; 4:232–4. [PubMed: 18278036] (c) Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. Chem. Biol. 2008; 15:1187–97. [PubMed: 19022179]

(7). (a) Bianco A, Townsley FM, Greiss S, Lang K, Chin JW. Nat. Chem. Biol. 2012; 8:748–50. [PubMed: 22864544] (b) Chen Y, Ma J, Lu W, Tian M, Thauvin M, Yuan C, Volovitch M, Wang Q, Holst J, Liu M, Vriz S, Ye S, Wang L, Li D. Cell Res. 2017; 27:294–297. [PubMed: 27934867] (c) Ernst RJ, Krogager TP, Maywood ES, Zanchi R, Beranek V, Elliott TS, Barry NP, Hastings MH, Chin JW. Nat. Chem. Biol. 2016; 12:776–8. [PubMed: 27571478] (d) Greiss S, Chin JW. J. Am. Chem. Soc. 2011; 133:14196–9. [PubMed: 21819153] (e) Parrish AR, She X, Xiang Z, Coin I, Shen Z, Briggs SP, Dillin A, Wang L. ACS Chem. Biol. 2012; 7:1292–302. [PubMed: 22554080]

(8). Grunwald DJ, Eisen JS. Nat. Rev. Genet. 2002; 3:717–24. [PubMed: 12209146]

(9). Lieschke GJ, Currie PD. Nat. Rev. Genet. 2007; 8:353–67. [PubMed: 17440532]

(10). Zon LI, Peterson RT. Nat. Rev. Drug Discovery. 2005; 4:35–44. [PubMed: 15688071]

(11). (a) Hemphill J, Liu Q, Uprety R, Samanta S, Tsang M, Juliano RL, Deiters A. J. Am. Chem. Soc. 2015; 137:3656–62. [PubMed: 25734836] (b) Shestopalov IA, Sinha S, Chen JK. Nat. Chem. Biol. 2007; 3:650–1. [PubMed: 17717538] (c) Tang X, Maegawa S, Weinberg ES, Dmochowski IJ. J. Am. Chem. Soc. 2007; 129:11000–1. [PubMed: 17711280] (d) Wang Y, Wu L, Wang P, Lv C, Yang ZJ, Tang XJ. Nucleic Acids Res. 2012; 40:11155–11162. [PubMed: 23002141] (e) Yamazoe S, Liu Q, McQuade LE, Deiters A, Chen JK. Angew. Chem., Int. Ed. 2014; 53:10114–8.

(12). Cheung CY, Webb SE, Meng A, Miller AL. Int. J. Dev. Biol. 2006; 50:561–9. [PubMed: 16741871]

(13). (a) Amatruda JF, Shepard JL, Stern HM, Zon LI. Cancer Cell. 2002; 1:229–31. [PubMed: 12086858] (b) Hill AJ, Teraoka H, Heideman W, Peterson RE. Toxicol. Sci. 2005; 86:6–19. [PubMed: 15703261] (c) Norton W, Bally-Cuif L. BMC Neurosci. 2010; 11:90. [PubMed: 20678210]

(14). (a) Portugues R, Severi KE, Wyart C, Ahrens MB. Curr. Opin. Neurobiol. 2013; 23:119–26. [PubMed: 23246238] (b) Reade A, Motta-Mena LB, Gardner KH, Stainier DY, Weiner OD, Woo S. Development. 2017; 144:345–55. [PubMed: 27993986] (c) Yoo SK, Deng Q, Cavnar PJ, Wu YI, Hahn KM, Huttenlocher A. Dev. Cell. 2010; 18:226–236. [PubMed: 20159593] (d) Arrenberg AB, Stainier DY, Baier H, Huisken J. Science. 2010; 330:971–974. [PubMed: 21071670]

(15). Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. J. Am. Chem. Soc. 2010; 132:4086–8. [PubMed: 20218600]

(16). Gautier A, Deiters A, Chin JW. J. Am. Chem. Soc. 2011; 133:2124–7. [PubMed: 21271704]

(17). Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A. J. Am. Chem. Soc. 2015; 137:5642–5. [PubMed: 25905628]

(18). Xiong Y, Steitz TA. Curr. Opin. Struct. Biol. 2006; 16:12–7. [PubMed: 16364630]

(19). Schmied WH, Elsässer SJ, Uttamapinant C, Chin JW. J. Am. Chem. Soc. 2014; 136:15577. [PubMed: 25350841]

(20). (a) Li J, Yu J, Zhao J, Wang J, Zheng S, Lin S, Chen L, Yang M, Jia S, Zhang X, Chen PR. Nat. Chem. 2014; 6:352–61. [PubMed: 24651204] (b) Nguyen DP, Lusic H, Neumann H, Kapadnis PB, Deiters A, Chin JW. J. Am. Chem. Soc. 2009; 131:8720–1. [PubMed: 19514718]

(21). Luo J, Uprety R, Naro Y, Chou C, Nguyen DP, Chin JW, Deiters A. J. Am. Chem. Soc. 2014; 136:15551–8. [PubMed: 25341086]

(22). Luo J, Liu Q, Morihiro K, Deiters A. Nat. Chem. 2016; 8:1027–1034. [PubMed: 27768095]

(23). Krens SF, He S, Spaink HP, Snaar-Jagalska BE. Gene Expression Patterns. 2006; 6:1019–26. [PubMed: 16774848]

(24). Furthauer M, Van Celst J, Thisse C, Thisse B. Development. 2004; 131:2853–64. [PubMed: 15151985]

(25). Anastasaki C, Estep AL, Marais R, Rauen KA, Patton EE. Hum. Mol. Genet. 2009; 18:2543–54. [PubMed: 19376813]

(26). Mizoguchi T, Izawa T, Kuroiwa A, Kikuchi Y. Dev. Biol. 2006; 300:612–22. [PubMed: 17026981]

(27). Xiong C, Liu X, Meng A. J. Biol. Chem. 2015; 290:25512–21. [PubMed: 26306042]

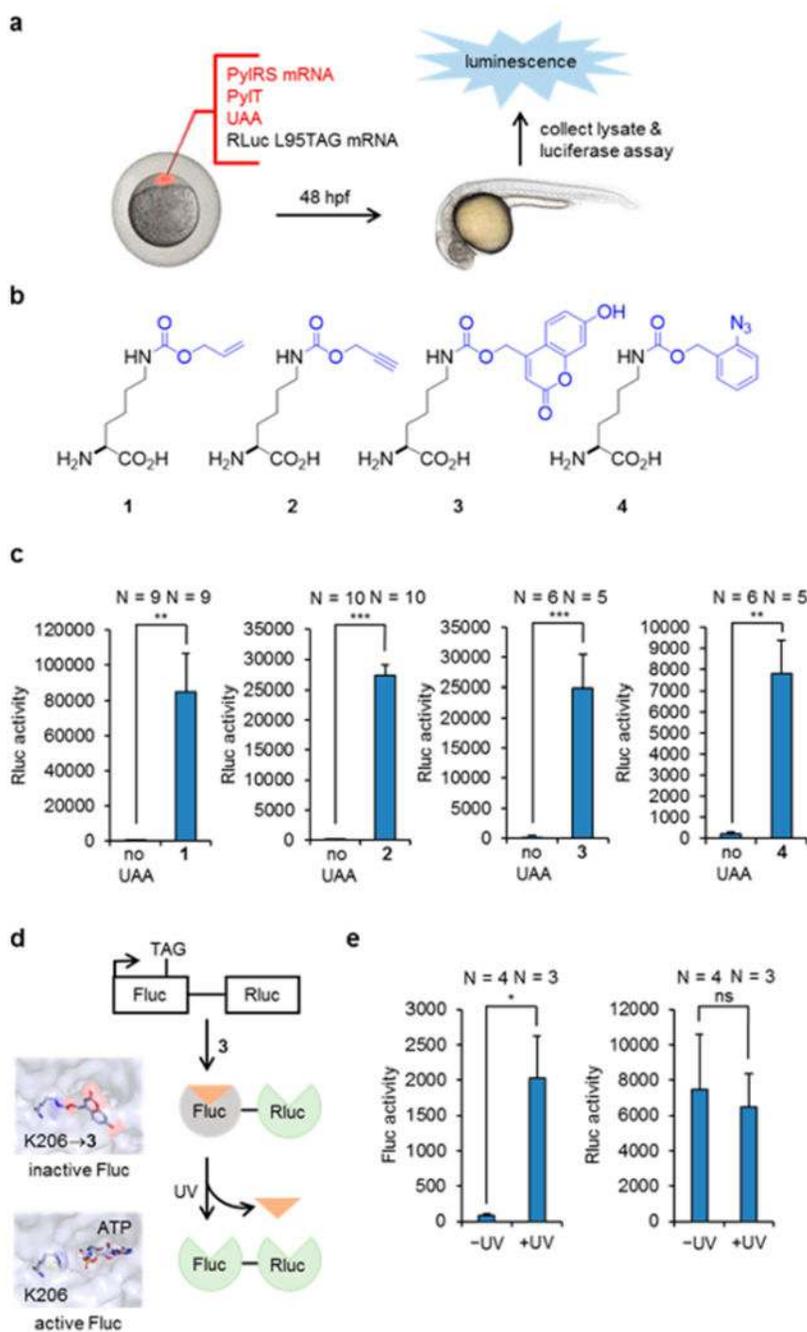(28). Anastasaki C, Rauen KA, Patton EE. Dis. Models & Mech. 2012; 5:546–52.

Figure 1.
Genetic encoding of the unnatural amino acids **1–4** and optical control of enzyme function. (a) Injection, gene expression, and Rluc reporter assay. (b) Structures of the unnatural lysine derivatives modified with an alkene, an alkyne, a coumarin caging group, and an azido benzyl group. (c) Incorporation of **1–4** into Rluc-L95TAG quantified by Rluc activity. (d) Construct used for generating a photocaged firefly luciferase followed by light activation. *Renilla* luciferase was used as an internal control for incorporation efficiency. Fluc containing **3** at position K206 is inactive, as ATP is blocked from the active site (PDB

2D1S). Decaging restores Fluc activity. (e) Fluc activity was observed after UV exposure, while Rluc activity was not affected and was used as an internal control. N indicates the number of pooled samples (4 embryos each). Statistical significance is indicated by ns (not significant), $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$ (unpaired t tests).
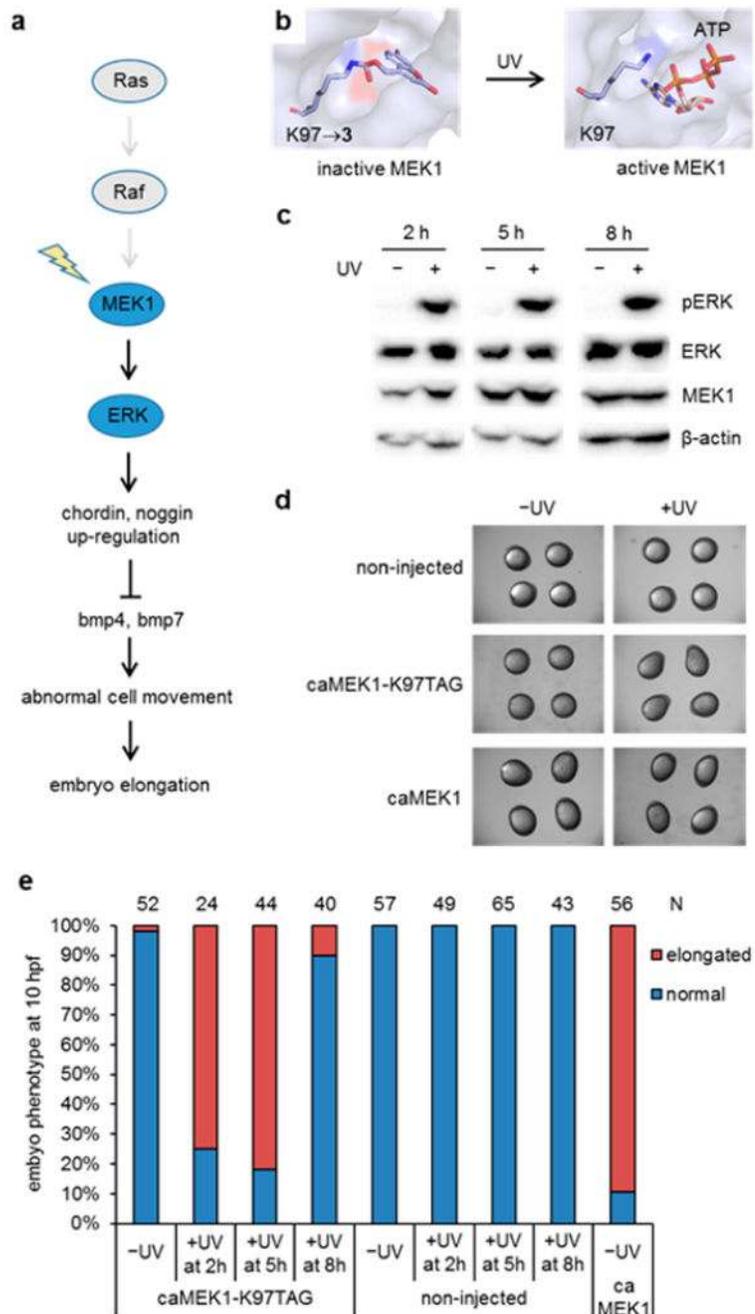
Figure 2.
Light-activation of MEK1 leads to elongated zebrafish embryos. (a) Activation of caged MEK1 induces an elongated phenotype through the secreted bmp inhibitors chordin and noggin. (b) MEK1 containing **3** at position K97 is inactive, as ATP is blocked from the active site (PDB 1S9J). Removal of the caging group through light exposure restores MEK1 activity. (c) Time-course analysis of ERK phosphorylation by activated MEK1. (d) Micrographs of embryos imaged at 10 hpf; irradiation was performed at 5 h postinjection. Embryos expressing caged MEK1 only displayed an elongation phenotype when activated

through light exposure. (e) Temporal activation of MEK1 reveals a critical time window for activity of the MEK/ERK pathway in the early embryo. N indicates the number of phenotypically scored embryos.

embryos fixed at shield stage (6 hpf)

embryos fixed at bud stage (10 hpf)

V ⎯ D   −UV

V ▬ D   +UV

ant / pos   −UV

ant / pos   +UV

a caMEK1-K97TAG, −UV   caMEK1-K97TAG, +UV   caMEK1

shield   5/6   6/6   7/7
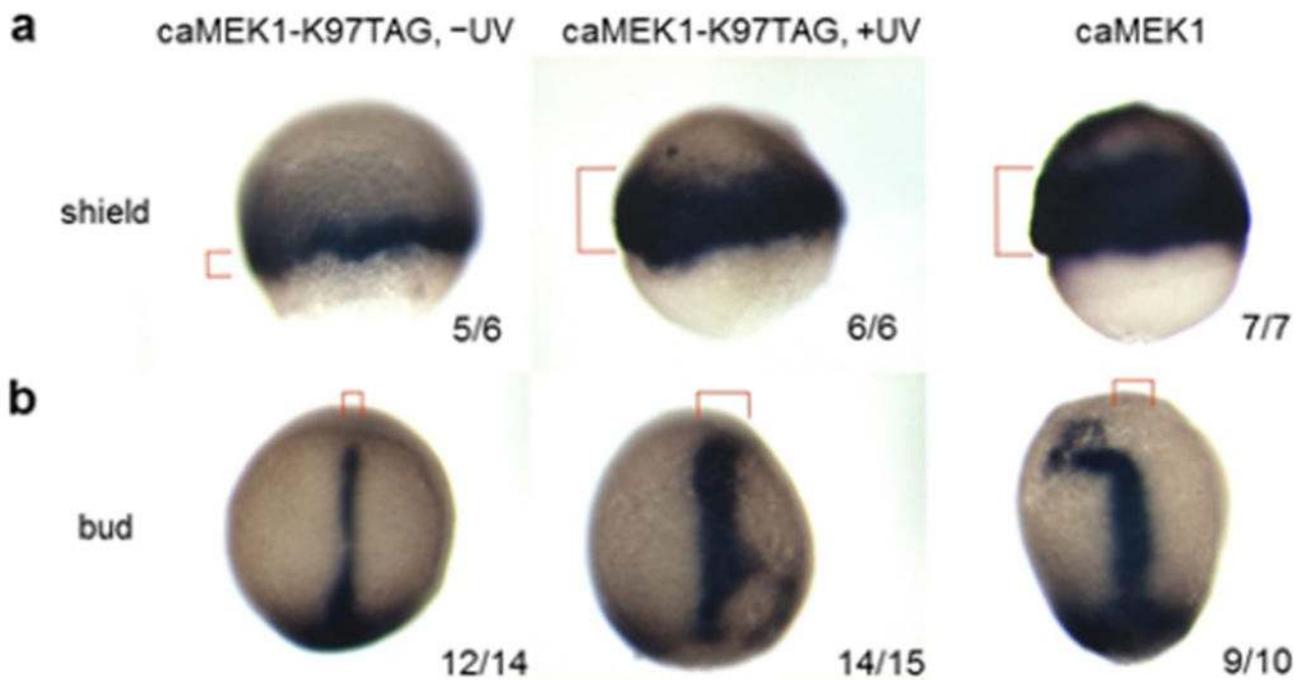
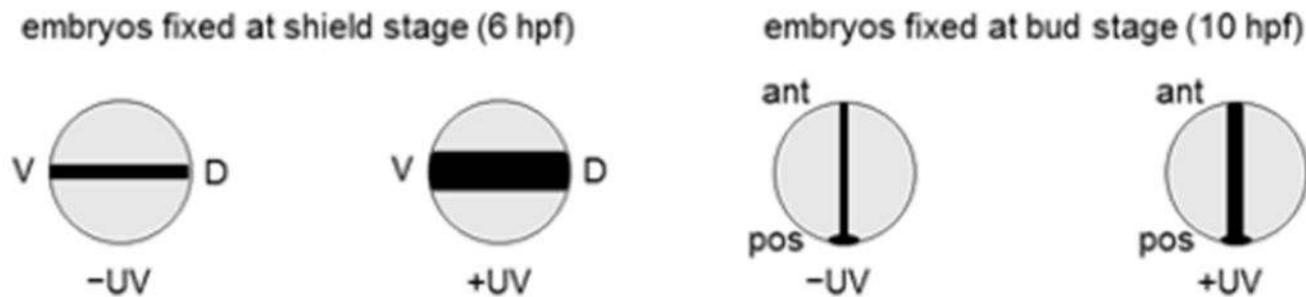**b** bud   12/14   14/15   9/10

Figure 3.
Upregulation of downstream targets in response to light-activated MEK1. (a) Expression of the *brachyuary homolog a* (*ta*) gene at shield stage. Embryos that expressed caged MEK1 showed broader expression after exposure to UV light, compared to embryos that were kept in the dark (see red brackets). (b) The same experiments conducted at bud stage showed wider expression along the notochord after light activation. The number of embryos with the displayed expression and the total number of embryos is indicated. Red brackets mark the *ta* expression area. Lateral views (a) and dorsal views with the anterior at the top (b).