Engineering Small Molecule Switches of Protein Function in Zebrafish Embryos

Wes Brown[a], Joshua Wesalo[a], Michael Tsang[b], and Alexander Deiters*[a]

[a] *Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, United States*

[b] *Department of Developmental Biology, University of Pittsburgh, Pittsburgh, PA 15260, United States*

deiters@pitt.edu

Abstract

Precise temporally regulated protein function directs the highly complex processes that make up embryo development. The zebrafish embryo is an excellent model organism to study development and conditional control over enzymatic activity is desirable to target chemical intervention to specific developmental events and to investigate biological mechanisms. Surprisingly, however, few generally applicable small molecule switches of protein function exist in zebrafish. Genetic code expansion allows for site-specific incorporation of unnatural amino acids into proteins that contain caging groups that are removed through addition of small molecule triggers such as phosphines or tetrazines. This broadly applicable control of protein function was applied to activate several enzymes, including a GTPase and a protease, with temporal precision in zebrafish embryos. Simple addition of the small molecule trigger to the media produces robust and tunable protein activation, which was used to gain insight into the development of a congenital heart defect from a RASopathy mutant of NRAS, and to control DNA and protein cleavage events catalyzed by a viral recombinase and a viral protease, respectively.

Introduction

Zebrafish embryos are an important animal model that have been key to studying vertebrate development and disease.[1-9] They develop into an embryo with heart, eyes, nervous system, and muscle all within 24 hours after fertilization. Crucial developmental steps occur in this short time period, all orchestrated by precise gene and protein function. Conditional control of protein activity is an important tool for studying embryo development.[10-13] Very few general approaches exist that utilize small molecules as triggers for protein function in zebrafish. Rapamycin-induced protein dimerization[14] and tamoxifen-induced nuclear translocation[15, 16] have enabled conditional control of protein activity in embryos (**Figure 1A-B**). However, these methodologies are limited to split proteins and functional control through protein mislocalization, and require protein domain fusions that may not always be functionally equivalent to the native protein. A generally applicable method for directly activating protein function with a small molecule trigger is lacking in zebrafish.

Here, we report the development of a universal technique for controlling protein function in zebrafish embryos, based on small molecule-cleavable caging groups. This builds onto our genetic code expansion approach in zebrafish embryos,[17-19] which relies on the expression of an orthogonal aminoacyl tRNA synthetase and tRNA pair to incorporate unnatural amino acids (UAAs) site-specifically into proteins in response to an amber stop codon.[20, 21] Unlike other approaches for small molecule triggered protein activation in zebrafish, this does not require fusion of protein domains. Caging of enzymatic function is established by simply placing a small molecule-removable group onto an essential active site residue, thereby blocking its function **(Figure 1C**)**.** Protecting groups that can be removed with bioorthogonal small molecule triggers have been used to control a variety of cellular processes.[22, 23] Through unnatural amino acid mutagenesis by way of amber stop codon suppression,[19, 21, 24] protecting groups can be placed into proteins with atomic resolution and complete genetic specificity in pro- and eukaryotic cells. This

includes an azidobenzyloxycarbonyl group that is removed through phosphine-induced Staudinger reduction of the azide to an amine, followed by self-immolation of the caging group to, e.g., release lysine.[25, 26] Similarly, a *trans*-cyclooctenyloxycarbonyl group has been cleaved through reaction with a tetrazine in an inverse electron demand Diels-Alder (ieDDA) cycloaddition, followed by self-immolation.[27, 28]

Here, we report the first site-specific installation of small molecule-activated amino acids into proteins in the zebrafish embryo. This enables fast and tunable activation of the function of several enzymes and we applied the temporal resolution afforded by this technology to gain new insight into the role of a RASopathy mutant in the development of congenital heart defects. These generally applicable small molecule switches are based on the amino acids *para*-azidobenzyloxycarbonyl lysine (PABK), which is decaged with phosphines, and *trans*-cyclooctenyloxycarbonyl lysine (TCOK), which is decaged with tetrazines.

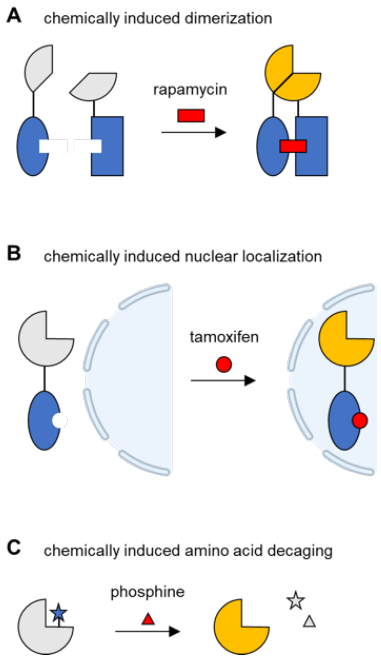

Figure 1. A) Chemically induced dimerization of a split protein using FKBP and FRB fusions that are brought together through addition of rapamycin, thus activating protein function. B) Tamoxifen-induced nuclear localization using a modified estrogen receptor fused to a protein that only functions in the nucleus, such as a transcription factor. C) Site-specific incorporation of a UAA containing a small molecule-cleavable caging group, such as an azidobenzyloxycarbonyl group, into a protein active site, followed by small molecule (phosphine) addition.

Results and Discussion

Conditional activation of a luciferase reporter through genetic incorporation of PABK and TCOK

Injection of mRNA that encodes an evolved pyrrolysyl aminoacyl tRNA synthetase (PylRS), along with tRNA (PylT), mRNA for the protein of interest, and the UAA itself has allowed for site specific incorporation of UAAs into proteins in zebrafish embryos (**Figure 2A**).[17-19] The genes of interest are cloned into the pCS2 vector that is commonly used as a template for *in vitro* transcription of mRNA (**Table S1**). To first demonstrate successful encoding of PABK or TCOK in embryos (**Figure 2A**), we directed incorporation of the caged lysine into the active site of firefly luciferase (fLuc) genetically fused to a *Renilla* luciferase (rLuc) reporter (**Figure 2B**). Both reporters provide high sensitivity and a quantitative readout that directly correlates to expression levels and enzymatic activity. The rLuc is C-terminal to the amber stop codon, hence it acts as a reporter for stop codon suppression as signal is only observed if the full-length dual luciferase protein is produced. Levels of rLuc activity directly correspond to the amount of UAA-protein that is produced, and thereby can be used a measure of UAA incorporation efficiency. We next incorporated PABK in the dual luciferase reporter using an engineered PylRS mutant from *M. barkeri* (L274A, Y349F, C313A).[26] By 24 hours post-fertilization (hpf), incorporation of PABK resulted in 23-fold higher rLuc activity than control embryos injected with all components but lacking UAA (no UAA, **Figure 2C-D**). The low background level of rLuc expression in the absence of the UAA confirms the bioorthogonality of the tRNA/tRNA synthetase pair, suggesting that canonical amino acids are not being incorporated at the amber stop codon. We also tested incorporation of two additional UAAs that are analogues of PABK, however, incorporation efficiency of these were weaker with only about a 5-fold increase in rLuc activity over background (**Figure S1**).

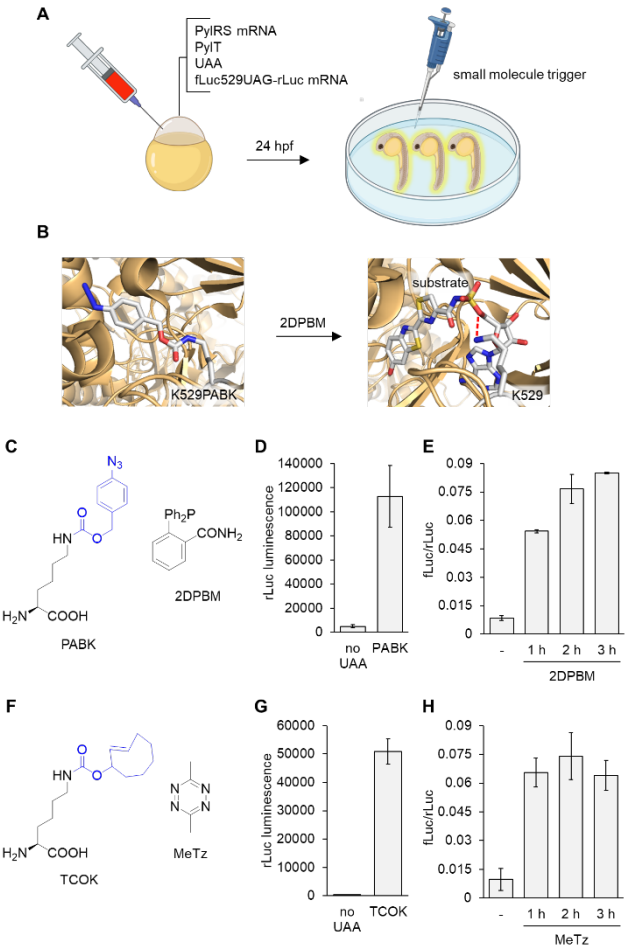

Figure 2. A) Zebrafish embryo injection with the machinery for UAA incorporation and subsequent enzyme activation through addition of the small molecule trigger to the embryo water. B) Active site rendering of PABK-caged fLuc before and after 2DPBM treatment. Red lines indicate hydrogen bonding between K529 and the substrate, a luciferin-adenylate sulfate analogue (PDB: 4G36). C-E) Structures of PABK and 2DPBM and rLuc activity indicate PABK incorporation into the dual-luciferase reporter and fLuc activation through treatment with 2DPBM. F-H) Structures of TCOK and MeTz and rLuc activity indicate TCOK incorporation into the dual-luciferase reporter and fLuc activation through treatment with MeTz. The presence of the small molecule inducer in the fish water before collection of embryos is indicated. Bars represent means and error bars represent standard deviations of measurements from three independent embryo lysates. To normalize the fLuc activity to total protein produced, the fLuc luminescence was divided by the rLuc luminescence (fLuc/rLuc).

For the analysis of small molecule-control of enzyme function using PABK, we measured fLuc activity. The amber stop codon was placed at site K529 of fLuc (fLuc529TAG-rLuc) where a critical lysine residue orients the substrate and stabilizes the adenylated substrate during catalysis through a hydrogen bonding interaction with the phosphate oxygen.[29] Caging this lysine blocks enzymatic activity and small molecule-induced decaging results in full activation of fLuc through formation of the wild-type active site (**Figure 2B**).[30] Previous studies of decaging PABK with a panel of phosphines in mammalian cells found that 2-(diphenylphosphinyl)benzamide (2DPBM) showed the fastest reaction rate and no toxicity.[26] At 24 hpf, we added 2DPBM (100 μM) to the embryo water for 1-3 hours and then collected embryos for assessment of small molecule-triggered protein activation (**Figure 2E**). PABK showed superior small molecule activation, with a 10-fold increase in the fLuc/rLuc ratio, while the other two PABK analogues mentioned previously had no distinguishable activation of fLuc activity with phosphine treatment, most likely due to their much lower incorporation efficiency which made detection of enzyme activation difficult (**Figure 2E, S1**). Overall, PABK demonstrated excellent incorporation into and caging of an enzymatic active site with minimal background activity. The added steric bulk from the azidobenzyloxycarbonyl group was predicted to occlude substrate binding, as well as negate an important hydrogen bond interaction between the substrate and the native lysine. Although the other two azidobenzyloxycarbonyl analogues (**Figure S1**) would have disrupted the active site in a similar way, they were limited by poor incorporation. Of note, enzyme activity plateaued after a 3 h treatment with 2DPBM, consistent with observations in mammalian cells.[25]

The same dual luciferase reporter was utilized for TCOK incorporation and small molecule activation. The PylRS mutant for TCOK incorporation shares one amino acid substitution but is otherwise different than the one used for PABK incorporation (Y271A, Y349F).[27] TCOK incorporation was significant, but, based on rLuc luminescence measurement, was about half of rLuc-PABK expression levels, presumably due to differences in aminoacylation efficiency between

the two UAA/PylRS pairs (**Figure 2F-G**). Background luminescence in the no UAA control was also low, suggesting that the TCOK synthetase also does not utilize any common amino acid as a substrate. For analysis of TCOK decaging, we tested a panel of tetrazines added to the embryo water at 100 μM when injected embryos reached 24 hpf. Luminescence readouts after 1-3 hours of small molecule exposure revealed that dimethyl tetrazine (MeTz) had the best decaging properties (**Figure S2A-B**) and luciferase activation plateaued after just 1 hour of MeTz treatment (**Figure 2H**). Background fLuc activity in the nontreated samples was minimal, not surprising given the larger steric hindrance imposed by TCO. The limiting step in decaging TCOK is in the elimination phase after the tetrazine cycloaddition. The ligated product can exist as two tautomers where only one is capable of immolation. The relative distribution between the two tautomers is dependent on the substituents on the 3- and 6-positions of the tetrazine.[27] Although cycloaddition was complete with all tetrazines that were tested in that study, the amount of decaging varied drastically between tetrazine derivatives, with MeTz having the highest level of decaging (>90%).

The extent of activation of a protein can be tuned based on the dosage of the small molecule trigger.[26] We assessed this by treating embryos expressing the caged dual luciferase reporter with different concentrations of 2DPBM or MeTz (10, 50, or 100 μM). A dose-response relationship was observed for both systems, with 100 μM of small molecule inducer showing the best decaging (**Figure S3**). Decaging for both systems was slightly slower than in cell culture (PABK/2DPBM $t_{1/2}$ = 27 min for a nuclear translocation reporter,[26] TCOK/MeTz $t_{1/2}$ = 10 min for a luciferase reporter[27]), which is likely due to the need for diffusion of the small molecule trigger through the chorion and then throughout the whole embryo, which at 24 hpf consists of over 20,000 cells. Overall, the PABK/2DPBM pair showed the highest protein yields and decaging ability with no toxicity in zebrafish embryos. While the TCOK/MeTz pair showed good incorporation and fast decaging (within 1 hour), MeTz (100 μM) unfortunately led to death of most

embryos by 24 hpf when incubated for one hour at 6 hpf, and the next best tetrazine **2**, while non-toxic (**Figure S2C**), was only about 30% as effective at fully activating the reporter. Therefore, we continued our studies of small molecule activation of protein function in zebrafish embryos, using the PABK/2DPBM-based switch.

Triggering of DNA recombination with phosphine treatment

To demonstrate broader utility of PABK for small molecule control of enzyme function, we next turned to Cre recombinase. This protein is a crucial tool in the field of zebrafish biology where it is used to conditionally activate or deactivate gene expression.[31] Generating a small molecule-triggered Cre recombinase based on decaging will allow for precise temporal control of activity with little to no background activity based on our luciferase results. Thus, PABK was incorporated in place of the critical lysine K201, which assists the 5'-O of the DNA backbone as a leaving group to facilitate the cleavage reaction (**Figure 3A**).[32] We expressed PABK-caged Cre recombinase (Cre-PABK) in a transgenic fish line that harbors a fluorescence reporter for DNA recombination.[18, 33] In this homozygous line, a *loxp*-EGFP-*loxp*-mCherry construct is expressed under a ubiquitous promotor. The EGFP-flanking Cre recombinase recognition motif *loxp*, a 34 bp DNA sequence, directs excision of EGFP and its stop codon, thereby activating mCherry expression. In summary, active Cre recombinase causes embryos to switch from EGFP to mCherry fluorescence (**Figure 3B**), providing a facile readout in live animals. Embryos producing Cre-PABK were treated with 2DPBM at 4, 5, or 6 hpf and then imaged at 48 hpf (**Figure 3C, Figure S4**). After 2DPBM treatment, the amount of mCherry is drastically increased to levels matching the wild-type Cre recombinase control. The earlier activation timepoint, as expected due to lower cell numbers, showed the most mCherry, and hence, the highest recombination frequency. Importantly, no background Cre recombinase activity was seen in the absence of

2DPBM. Some EGFP is still present in the 2DPBM treated Cre-PABK expressing embryos because gene recombination was activated after EGFP protein had already been expressed. These results further demonstrate the utility of PABK for caging enzymatic activity, here for controlling gene expression at defined timepoints in embryo development through DNA recombination.

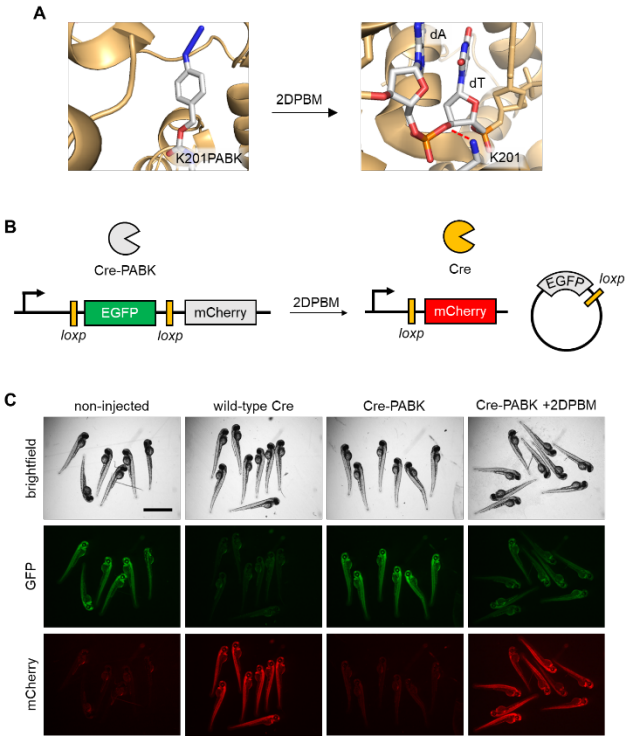

Figure 3. A) Active site rendering of PABK-caged Cre recombinase before and after 2DPBM treatment. Red lines denote hydrogen bonding (PDB: 1CRX). B) Recombination-based reporter gene activation and deactivation in response to 2DPBM treatment of PABK-caged Cre recombinase (Cre-PABK). C) Micrographs of 48 hpf transgenic embryos with the treatment conditions indicated, showing recombinase-based switching of fluorescence from green to red. Treatment with 2DPBM was initiated at 4 hpf. Scale bar = 2 mm.

Conditional control of a novel protease through active site incorporation of PABK

Next, we tested the ability of our protein activity on-switch to control protease function. Conditional control of proteases can allow for temporal control of protein knockout or protein localization.[34] To our knowledge, the use of proteases as chemical biological tools in zebrafish embryos has not been utilized, so we hoped to establish in this model organism a small molecule-activated exogenous protease. The Blotched Snakehead Virus VP4 protease relies on a serine/lysine catalytic dyad, in which the lysine acts as a general base to activate the serine for nucleophilic attack of the substrate protein backbone (**Figure 4A-B**).[35] This makes it an excellent target for installation of a PAB group, we expected that a K173PABK mutation would remove its function as a general base as well as sterically block the peptide substrate from entering the binding pocket. In order to construct this small molecule-activated protease, K173 in VP4 was mutated to an amber stop codon. PylT, PylRS mRNA, VP4 K173UAG mRNA, and PABK were injected into embryos at the one-cell stage and expression of caged VP4 (VP4-PABK) was confirmed by western blot at 24 hpf (**Figure 4C**). The virus uses VP4 for segmentation of its polyprotein into its individual proteins for replication and assembly of new viral particles through cleavage at the recognition motif PXAA.[35] We cloned this motif in between a nuclear exclusion sequence (NES) and mCherry to generate a fluorescent nuclear translocation reporter that can be imaged in live embryos. Activation of VP4-PABK with 2DPBM leads to removal of the NES from mCherry, resulting in diffusion from the cytoplasm into the nucleus and eventually even distribution throughout the cell (N/C = 1) (**Figure 4D**). Embryos co-expressing the caged VP4 and the reporter were treated with 2DPBM for 3 hours and then imaged (**Figure 4E**). The reporter showed nuclear exclusion when injected alone or in the presence of VP4-PABK. When wild-type VP4 was expressed, the N/C ratio increased from 0.64 to 0.84, indicating mCherry translocation from the nucleus to the cytoplasm (**Figure 4F**). Importantly, activation of VP4-PABK with 2DPBM resulted in a similar N/C ratio of 0.82, which is indicative of successful small

molecule-activation of the protease, NES cleavage, and subsequent translocation of the reporter into the nucleus. No background cleavage of the reporter was observed in the absence of 2DPBM. Thus, we were able to demonstrate conditional control of a new viral protease in zebrafish embryos, achieving a response on par with the wild-type enzyme. To our knowledge, this is the first demonstration of an exogenous viral protease being used for selective protein cleavage in the zebrafish embryo. This offers a unique tool with many capabilities such as releasing a membrane bound transcription factor with a cleavable linker to act as a sensor for protein-protein interactions or to activate signaling pathways,[36, 37] inactivating an engineered protein or activating a protein by removing an appended degron,[38, 39] and controlling cellular localization with cleavage of signaling peptides as we demonstrated in this work.

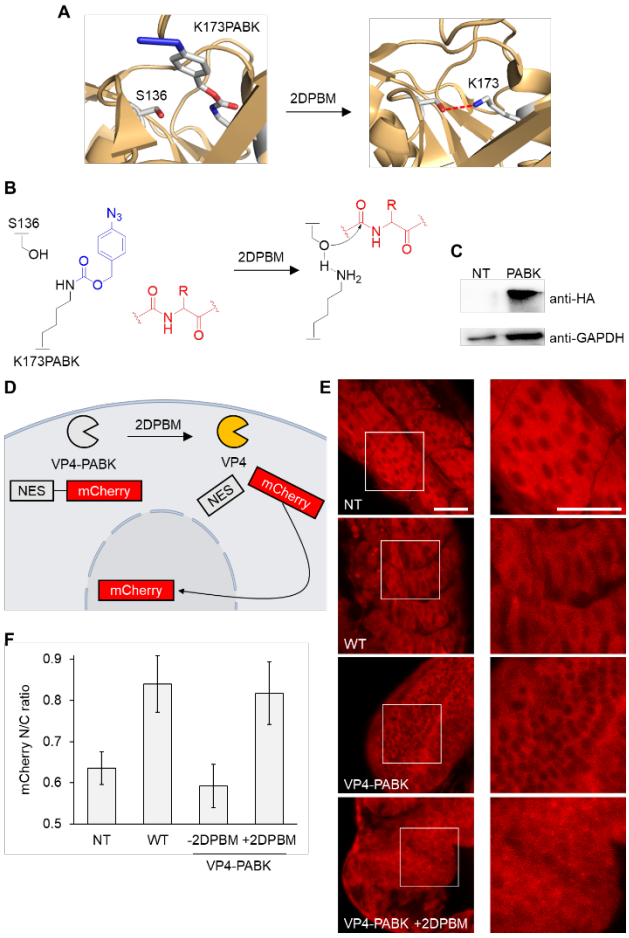

Figure 4. A) Active site rendering of PABK-caged VP4 protease before and after 2DPBM treatment. Red line denotes hydrogen bond (PDB: 2GEF). B) K173 to PABK mutation inactivates the function of the lysine as a base and blocks the ability for the substrate to bind the active site of VP4. C) Western blot showing UAA-dependent expression of K173PABK VP4-HA in zebrafish embryos. D) Fluorescent nuclear translocation reporter for VP4 activity: activation of caged VP4 by 2DPBM leads to removal of the nuclear export signal (NES) from mCherry, allowing for diffusion into the nucleus. E) Confocal micrographs of embryos at 24 hpf expressing only the reporter, wild-type VP4, or caged VP4 with or without 2DPBM treatment. Scale bar = 40 μm F) Average nucleus/cytoplasm fluorescence ratio of 20 cells for each condition. Error bars represent standard deviations.

Temporal activation of NRAS reveals insight into the mechanism of RASopathy-induced heart defects

After successfully controlling DNA recombination (gene expression) and protease function (protein translocation) using our small molecule switch, we next attempted to conditionally regulate cell signaling, another essential and highly dynamic biological process during embryo development. Here, we turned to NRAS, a GTPase involved in the RAS/MAPK signaling pathway.[18, 33] RAS/MAPK signaling plays several important roles during early embryo development, like coordinating gastrulation and organogenesis by regulating cell proliferation and migration.[40] Indeed, when these developmental programs are inappropriately reactivated, they can lead to cancer in which hyperactive cell proliferation and migration are characteristic to the disease.[41] When RAS GTPases are bound to GTP, they bind downstream effectors, activating the signaling pathway. Although RAS has intrinsic GTPase activity, it is slow and is usually assisted by a GTPase activating protein (GAP). Hydrolysis of GTP to GDP, causes a conformational change and inactivates its binding to effectors. A guanine nucleotide exchange factor (GEF) is then able to remove the GDP and GTP can freely diffuse back into the active site. K16 of NRAS undergoes electrostatic interactions with both the β and γ-phosphate of GTP, holding it in the active site and orienting it for hydrolysis to GDP.[42] We reasoned that replacing this lysine with PABK would block those interactions and also sterically occlude the GTP binding pocket, rendering NRAS inactive (**Figure 5A**). While kinases have been successfully caged with this method before, this is the first example of this approach being applied to cage a GTPase.[43, 44] The nucleotide binding is very similar, but the functions of the protein classes are very different. RAS GTPases catalyze hydrolysis to induce conformational changes to alter target binding, whereas kinases transfer the phosphate group to protein targets. We opted to study a constitutively active RASopathy mutant of NRAS (NRAS G60E) , which has reduced GAP binding, thus slower GTPase

function and a longer residence in its active GTP form to decouple it from dependence on upstream receptor tyrosine kinase activation.[45, 46]

RASopathies are a class of diseases caused by activating mutations of RAS/MAPK signaling proteins that share several clinical phenotypes such as craniofacial and heart defects.[47] It is one of the largest families of congenital diseases, affecting nearly 1 out of every 1000 live births.[48] While the many different known mutations that have been identified all lead to the same upregulation of the RAS/MAPK pathway, it is a surprisingly heterogenous group of diseases that can be sub-classified based on the prevalent phenotypes. The diversity of phenotypes includes craniofacial deformities, neurocognitive defects, skin abnormalities, and heart defects.[47, 48] The kinase mutations tend to be weakly activating, since strongly activating mutations, as seen in cancer, tend to be developmentally lethal.[48] There is a lot we do not know about RASopathies, particularly when it comes to understanding the mechanism linking the different mutations and the different malformations. We sought to learn more about the molecular underpinnings behind RASopathy-induced disruption of development through conditional control of this NRAS mutant. Specifically, we focused on early embryonic events that set the stage for proper zebrafish morphogenesis, gastrulation, and heart development. The exact mechanism for the induction of heart defects by RASopathy mutations is not clearly understood, and we aimed to gain insight into which cardiogenic step was being disrupted through the temporal control of the NRAS mutant. PABK caged NRAS G60E (NRAS-PABK) was expressed in zebrafish embryos and western blot of embryo lysate confirmed expression of the caged construct (**Figure 5B**). The RASopathy mutant is known to induce gastrulation defects in zebrafish embryos,[45] so we treated embryos starting at 3 hpf with 2DPBM to ensure complete decaging by the beginning of gastrulation at 6 hpf.[49] Embryos were scored at 24 hpf for gastrulation defects, manifesting here as dorsalization, where dorsoanterior structures like the head and brain are formed while ventroposterior structures like the tail are malformed (**Figure 5C, S5**). In this case, shortening of the

tail was the most prominent phenotype. Of note, non-injected embryos treated with 2DPBM at the same timepoint showed no defects (**Figure S5**), again supporting the compatibility of the small molecule with embryo development. When NRAS K16PABK was expressed, a large increase in dorsalized embryos was observed after 2DPBM treatment, compared to nontreated NRAS-PABK animal. The NRAS G60E mRNA injection (referred to as NRAS to help distinguish it from NRAS-PABK) resulted in most embryos dying before 24 hpf due to severe dorsalization defects incompatible with life, as seen in a prior study.[45] This level of dorsalization was not observed in the case of small molecule-activated NRAS-PABK, highlighting the ability to control the extent of enzyme activation through 2DPBM addition. Importantly, the number of dorsalized embryos in the nontreated NRAS-PABK was also low, indicating minimal background activity in the caged GTPase.

A consequence of most RASopathy mutations is the development of heart defects in the patient.[50, 51] The impacts of NRAS RASopathy mutants on heart development have never been studied during development to our knowledge.[52, 53] We used a transgenic zebrafish line, Tg(*myl7:egfp*) with fluorescent cardiomyocytes, in order to image heart morphology at 48 hpf, and expressed NRAS-PABK in the corresponding embryos. NRAS was activated through 2DPBM treatment at 8 or 24 hpf. We observed that heart looping defects were present in 70% of embryos in which NRAS-PABK was activated at 8 hpf, while only about 35% had looping defects with activation at 24 hpf (**Figure 5D-E**). There were minimal heart defects in the no 2DPBM condition, similar to the non-injected condition. Other GTPases in the RAS/MAPK pathway with RASopathy mutants like RIT1 and KRAS cause heart looping defects.[54, 55] This was found to be due to disruption of Kupffer's Vesicle (KV) function, which is crucial for establishing the first signs of asymmetry in embryos, such as the rightward looping of the heart tube.[56] The 8 hpf timepoint is before the KV forms (around 10 hpf) and is right after gastrulation so temporal activation is less likely to cause gastrulation defects that would confound gross heart anatomy

scoring (**Figure S6**). The 24 hpf timepoint is after KV has already formed, performed its function in development, and dissolved.[57] Therefore, the higher prevalence of looping defects with 8 hpf activation suggests the looping defects we saw were likely related to disruption of KV function. However, the 24 hpf timepoint still had slightly higher rates of looping defects than embryos expressing non-activated NRAS, so other involvement in later steps of looping is possible. This has important implications for the pathogenesis of RASopathy-induced heart defects, suggesting that the various types of heart defects seen in patients such as septal defects, hypertrophic cardiomyopathy, and pulmonic stenosis may be linked to initial disruption of heart looping. Before, it was unknown whether the heart defects seen in zebrafish embryos with RASopathy mutant expression were primarily due to disruption of KV function, or if there was disruption at later steps as well. The results from our temporally controlled activation of NRAS G60E supports the hypothesis that the majority of heart defects due to this RASopathy mutant are linked to KV disruption, supporting findings for other RASopathy GTPases.[54, 55] Because heart looping is an early process that sets the stage for proper morphogenesis, disruption can cause many problems such as septal defects and outflow tract malformations.[58, 59]

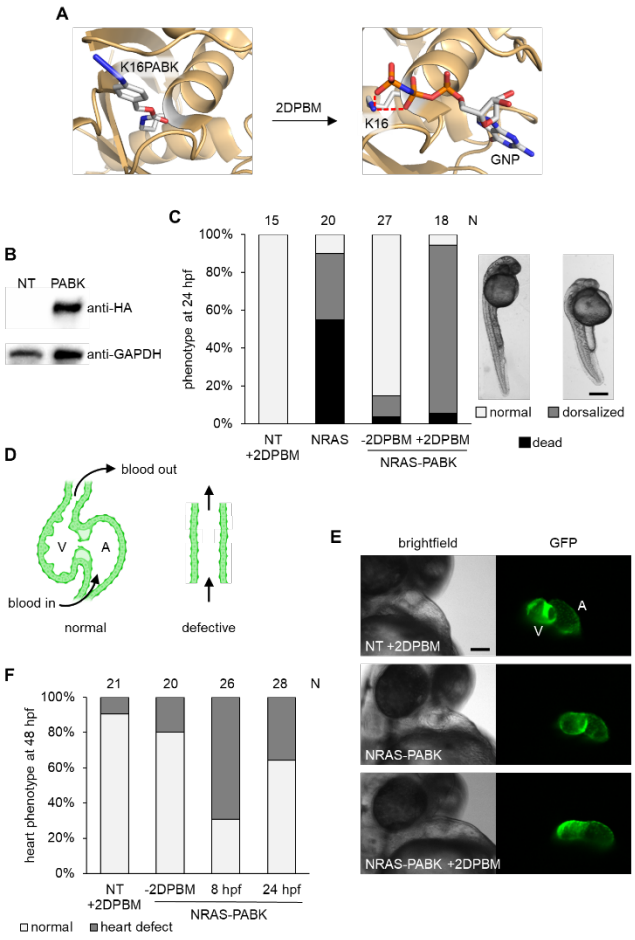

Figure 5. A) Active site rendering of the caged NRAS before and after 2DPBM treatment. Red lines denote hydrogen bonding (PDB: 5UHV). B) Western blot demonstrating incorporation of PABK into HA-tagged NRAS in zebrafish embryos. C) Expression of the RASopathy mutant NRAS G60E causes embryo dorsalization defects. Embryos expressing caged NRAS were treated with 2DPBM at 3 hpf and scored at 24 hpf. Representative examples of normal or dorsalized embryos are shown to the right. Scale bar = 0.4 mm. D) A diagram of a normally developed (left) and defective (right) heart at 48 hpf. A = atrium and V = ventricle. E-F) The RASopathy mutant NRAS G60E induces heart defects in zebrafish. Embryos expressing caged NRAS were treated with 2DPBM at 8 hpf or 24 hpf in a fluorescent heart transgenic line Tg(*myl7*:*egfp*). Scale bar = 100 μm. Heart defects were scored at 48 hpf. N = number of embryos per condition.

Summary

In summary, we have developed two different small molecule-triggered lysine decaging systems in zebrafish embryos, one based on the Staudinger reduction of an aryl azide and the other based on a cycloaddition between a tetrazine and a *trans*-cyclooctene. The corresponding lysine analogs PABK and TCOK were genetically encoded for the first time in zebrafish embryos and good incorporation into several proteins was observed. Small molecule-triggered protein activation was efficient for both PABK and TCOK, however, the best tetrazine small molecule trigger for TCOK showed toxicity in embryos, prompting us to explore the general applicability of the phosphine (2DPBM) triggered PABK switch. PABK showed very high incorporation efficiency into proteins in zebrafish embryos and was versatile for caging a variety of enzymes including luciferase, Cre recombinase, VP4 protease, and the GTPase NRAS. Fast and tunable activation of enzyme function was observed through simple addition of the small molecule trigger to the fish water, suggesting good diffusibility into the embryo without observable toxicity. In fact, addition of small molecules to the embryo water is a common technique used in high-throughput drug screens with zebrafish.[60] We also pioneered the first use of a conditionally controlled protease in zebrafish embryos, a useful approach that has enabled valuable biological studies in mammalian cells. Temporal control of the RASopathy mutant NRAS offered insight into the mechanism of RASopathy-induced heart looping defects, suggesting disordered looping may be a common initiator for congenital heart defects for RASopathies. Detailing the mechanism behind RASopathy-associated heart defects may aid in discovering diagnostics and treatments for this pathogenic process. Thus, we showcase mutagenesis with PABK as a powerful and generalizable approach for the rational design of conditionally controlled enzymes in zebrafish embryos. The placement of the caging group is readily based on structural or mechanistic information about the protein of interest and the machinery for unnatural amino acid mutagenesis is easily produced and is highly efficient in zebrafish embryos. While conditional control

and perturbation of biological processes is important for understanding them, generally applicable approaches for small molecule control of protein function are rare in aquatic embryos. The approach developed and validated new opportunities for biological studies in this important model organism for human health and development.

Methods

Zebrafish care and microinjection. The zebrafish experiments were performed according to a protocol approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Pittsburgh (protocol no. 19075360). Embryos were collected after natural mating. Injection solutions were prepared on ice and a volume of 2 nL was injected into the yolk of 1-2 cell stage embryos using a World Precision Instruments Pneumatic PicoPump injector. Embryos were incubated in E3 (embryo) water at 28.5 °C in the dark. For all UAA incorporation experiments, a total of 400 pg of PylRS mRNA, 16 ng of PylT, and 20 pmol of UAA were injected along with the mRNA for the protein of interest: fLuc-529UAG-rLuc (400 pg) Cre recombinase-201UAG (400 pg) wild-type Cre recombinase (25 pg), VP4-729UAG (400 pg), wild-type VP4 (400 pg), NES-mCherry reporter (200 pg), NRAS-G60E-16UAG (100 pg), NRAS-G60E (50 pg; 25 pg for heart defect experiments due to the toxicity of the protein).

Zebrafish small molecule treatment and imaging. For embryo small molecule treatment, a solution of the trigger (100 μM) in E3 water was prepared. Embryos were suspended in the solution and incubated at 28.5 °C. 2DPBM oxidizes in aqueous solutions, and at 28.5 °C the half-life of 2DPBM in E3 water was 90 minutes.[61] Therefore, for PABK decaging, the 2DPBM solution was replaced at 90 minutes and incubated an additional 90 minutes for a total of 180 minutes for full decaging. For TCOK decaging, embryos were incubated with tetrazine for 1 hour for full decaging. Stereoscope imaging of embryos was performed with a Leica M205 FA

microscope with a DsRed filter (ex: 510-560, em: 590-650), an EGFP filter (ex: 450-490 nm, em: 500-550 nm), and the bright field channel. For VP4 protease studies, a Zeiss LSM 700 laser scanning confocal microscope was used for imaging. Embryos at 24 hpf were treated with 2DPBM for 3 hours and then embedded in 500 µl of 1.5 % low-melting-point agarose in a glass bottom 35 mm dish and then submerged in E3 water. The 555 nm laser was used for imaging mCherry fluorescence at between 40-80% laser power with a 20x water immersion objective. A 1024 x 1024 pixel imaging window was used with a 15 µs pixel dwell time. Fluorescent nucleus/cytoplasm (N/C) intensity ratios were calculated using ImageJ software. Background-subtracted measurements of fluorescent intensity were taken from selected regions in the nucleus and cytoplasm for 20 cells per condition. This was performed by measuring the mean grey value of the nuclear or cytoplasmic regions, subtracting mean grey value of the background, and then calculating the ratio for each cell independently.

Luciferase assays. At 24 hpf, 4 embryos were collected in a 1.5 ml microcentrifuge tube. The water was removed and 50 µl of passive lysis buffer (Promega) was added. Embryos were manually homogenized with a p200 pipette tip. Samples were centrifuged at 16,200 rcf for 8 minutes at 4 °C. Lysate aliquots (30 µl) were added to a white-bottom 96-well plate and loaded into a plate reader with autoinjection function (Tecan Infinite M1000 pro). The Dual-Luciferase Reporter 1000 Assay System (Promega) was used 20 µl of fLuc assay reagent was injected (injection speed = 200 µl/sec), followed by a pause for 2 seconds, and a luminescence reading with auto-attenuation mode. Then, 20 µl of the stop and glo rLuc assay reagent was injected, followed by a pause for 2 seconds, and another luminescence reading. Relative fLuc values were calculated by dividing the fLuc value by the rLuc value (fLuc/rLuc).

Acknowledgments

We acknowledge financial support from the National Institutes of Health (R01GM132565, R01HL142788). W.B. was supported by a University of Pittsburgh Mellon Fellowship. We thank Dr. Rohan Kumbhare and Dr. Yaniv Tivon for supplying the tetrazines used in this study. We thank Dr. Joseph M. Fox for suppling the *trans*-cyclooctene alcohol for TCOK synthesis. Parts of figures were generated using BioRender.com.

Supporting Information

Supporting methods, DNA sequences, and figures.

References

1. Teame, T.; Zhang, Z.; Ran, C.; Zhang, H.; Yang, Y.; Ding, Q.; Xie, M.; Gao, C.; Ye, Y.; Duan, M.; Zhou, Z., The use of zebrafish (Danio rerio) as biomedical models. *Animal Frontiers* **2019,** *9* (3), 68-77.

2. Adams, M. M.; Kafaligonul, H., Zebrafish—A Model Organism for Studying the Neurobiological Mechanisms Underlying Cognitive Brain Aging and Use of Potential Interventions. *Frontiers in Cell and Developmental Biology* **2018,** *6* (135).

3. Bradford, Y. M.; Toro, S.; Ramachandran, S.; Ruzicka, L.; Howe, D. G.; Eagle, A.; Kalita, P.; Martin, R.; Taylor Moxon, S. A.; Schaper, K.; Westerfield, M., Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *ILAR J* **2017,** *58* (1), 4-16.

4. Liu, J.; Stainier, D. Y. R., Zebrafish in the study of early cardiac development. *Circ Res* **2012,** *110* (6), 870-874.

5. Bakkers, J., Zebrafish as a model to study cardiac development and human cardiac disease. *Cardiovasc Res* **2011,** *91* (2), 279-288.

6. Delvecchio, C.; Tiefenbach, J.; Krause, H. M., The zebrafish: a powerful platform for in vivo, HTS drug discovery. *Assay Drug Dev Technol* **2011,** *9* (4), 354-61.

7. Xu, C.; Zon, L. I., 9 - The zebrafish as a model for human disease. In *Fish Physiology*, Perry, S. F.; Ekker, M.; Farrell, A. P.; Brauner, C. J., Eds. Academic Press: 2010; Vol. 29, pp 345-365.

8. Veldman, M. B.; Lin, S., Zebrafish as a Developmental Model Organism for Pediatric Research. *Pediatric Research* **2008,** *64* (5), 470-476.

9. Lieschke, G. J.; Currie, P. D., Animal models of human disease: zebrafish swim into view. *Nature Reviews: Genetics* **2007,** *8* (5), 353-367.

10. Kalvaitytė, M.; Balciunas, D., Conditional mutagenesis strategies in zebrafish. *Trends in Genetics*.

11. Varady, A.; Distel, M., Non-neuromodulatory Optogenetic Tools in Zebrafish. *Frontiers in Cell and Developmental Biology* **2020,** *8*.

12. Antinucci, P.; Dumitrescu, A.; Deleuze, C.; Morley, H. J.; Leung, K.; Hagley, T.; Kubo, F.; Baier, H.; Bianco, I. H.; Wyart, C., A calibrated optogenetic toolbox of stable zebrafish opsin lines. *eLife* **2020,** *9*, e54937.

13. Deiters, A.; Yoder, J. A., Conditional Transgene and Gene Targeting Methodologies in Zebrafish. *Zebrafish* **2006,** *3* (4), 415-429.

14. Courtney, T. M.; Darrah, K. E.; Horst, T. J.; Tsang, M.; Deiters, A., Blue Light Activated Rapamycin for Optical Control of Protein Dimerization in Cells and Zebrafish Embryos. *ACS Chemical Biology* **2021,** *16* (11), 2434-2443.

15. Hans, S.; Kaslin, J.; Freudenreich, D.; Brand, M., Temporally-controlled site-specific recombination in zebrafish. *PLoS One* **2009,** *4* (2), e4640.

16. Akerberg, A. A.; Stewart, S.; Stankunas, K., Spatial and Temporal Control of Transgene Expression in Zebrafish. *PLOS ONE* **2014,** *9* (3), e92217.

17. Liu, J.; Hemphill, J.; Samanta, S.; Tsang, M.; Deiters, A., Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. *Journal of the American Chemical Society* **2017,** *139* (27), 9100-9103.

18. Brown, W.; Liu, J.; Tsang, M.; Deiters, A., Cell-Lineage Tracing in Zebrafish Embryos with an Expanded Genetic Code. *Chembiochem* **2018,** *19* (12), 1244-1249.

19. Brown, W.; Liu, J.; Deiters, A., Genetic Code Expansion in Animals. *ACS Chemical Biology* **2018,** *13* (9), 2375-2386.

20. Chin, J. W., Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annual Review of Biochemistry* **2014,** *83* (1), 379-408.

21. Shandell, M. A.; Tan, Z.; Cornish, V. W., Genetic Code Expansion: A Brief History and Perspective. *Biochemistry* **2021,** *60* (46), 3455-3469.

22. Li, J.; Chen, P. R., Development and application of bond cleavage reactions in bioorthogonal chemistry. *Nature Chemical Biology* **2016,** *12* (3), 129-137.

23. Shieh, P.; Bertozzi, C. R., Design strategies for bioorthogonal smart probes. *Organic & Biomolecular Chemistry* **2014,** *12* (46), 9307-9320.

24. Chung, C. Z.; Amikura, K.; Söll, D., Using Genetic Code Expansion for Protein Biochemical Studies. *Frontiers in Bioengineering and Biotechnology* **2020,** *8*.

25. Luo, J.; Liu, Q.; Morihiro, K.; Deiters, A., Small-molecule control of protein function through Staudinger reduction. *Nat Chem* **2016,** *8* (11), 1027-1034.

26. Wesalo, J. S.; Luo, J.; Morihiro, K.; Liu, J.; Deiters, A., Phosphine-Activated Lysine Analogues for Fast Chemical Control of Protein Subcellular Localization and Protein SUMOylation. *Chembiochem* **2020,** *21* (1-2), 141-148.

27. Li, J.; Jia, S.; Chen, P. R., Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells. *Nature Chemical Biology* **2014,** *10* (12), 1003-1005.

28. Zhang, G.; Li, J.; Xie, R.; Fan, X.; Liu, Y.; Zheng, S.; Ge, Y.; Chen, P. R., Bioorthogonal Chemical Activation of Kinases in Living Systems. *ACS Cent Sci* **2016,** *2* (5), 325-331.

29. Jazayeri, F. S.; Amininasab, M.; Hosseinkhani, S., Structural and dynamical insight into thermally induced functional inactivation of firefly luciferase. *PloS one* **2017,** *12* (7), e0180667-e0180667.

30. Liu, L.; Liu, Y.; Zhang, G.; Ge, Y.; Fan, X.; Lin, F.; Wang, J.; Zheng, H.; Xie, X.; Zeng, X.; Chen, P. R., Genetically Encoded Chemical Decaging in Living Bacteria. *Biochemistry* **2018,** *57* (4), 446-450.

31. Carney, T. J.; Mosimann, C., Switch and Trace: Recombinase Genetics in Zebrafish. *Trends in Genetics* **2018,** *34* (5), 362-378.

32. Gibb, B.; Gupta, K.; Ghosh, K.; Sharp, R.; Chen, J.; Van Duyne, G. D., Requirements for catalysis in the Cre recombinase active site. *Nucleic Acids Res* **2010,** *38* (17), 5817-32.

33. Brown, W.; Deiters, A., Chapter Thirteen - Light-activation of Cre recombinase in zebrafish embryos through genetic code expansion. In *Methods in Enzymology*, Deiters, A., Ed. Academic Press: 2019; Vol. 624, pp 265-281.

34. Chung, H. K.; Lin, M. Z., On the cutting edge: protease-based methods for sensing and controlling cell biology. *Nature Methods* **2020,** *17* (9), 885-896.

35. Da Costa, B.; Soignier, S.; Chevalier, C.; Henry, C.; Thory, C.; Huet, J.-C.; Delmas, B., Blotched snakehead virus is a new aquatic birnavirus that is slightly more related to avibirnavirus than to aquabirnavirus. *J Virol* **2003,** *77* (1), 719-725.

36. Barnea, G.; Strapps, W.; Herrada, G.; Berman, Y.; Ong, J.; Kloss, B.; Axel, R.; Lee, K. J., The genetic design of signaling cascades to record receptor activation. *Proc Natl Acad Sci U S A* **2008,** *105* (1), 64-9.

37. Morsut, L.; Roybal, K. T.; Xiong, X.; Gordley, R. M.; Coyle, S. M.; Thomson, M.; Lim, W. A., Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors. *Cell* **2016,** *164* (4), 780-91.

38. Pratt, M. R.; Schwartz, E. C.; Muir, T. W., Small-molecule-mediated rescue of protein function by an inducible proteolytic shunt. *Proc Natl Acad Sci U S A* **2007,** *104* (27), 11209-14.

39. Uhlmann, F.; Wernic, D.; Poupart, M. A.; Koonin, E. V.; Nasmyth, K., Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. *Cell* **2000,** *103* (3), 375-86.

40. Krens, S. F. G.; Spaink, H. P.; Snaar-Jagalska, B. E., Functions of the MAPK family in vertebrate-development. *FEBS Letters* **2006,** *580* (21), 4984-4990.

41. Dhillon, A. S.; Hagan, S.; Rath, O.; Kolch, W., MAP kinase signalling pathways in cancer. *Oncogene* **2007,** *26* (22), 3279-3290.

42. Chen, J.; Zeng, Q.; Wang, W.; Hu, Q.; Bao, H., Q61 mutant-mediated dynamics changes of the GTP-KRAS complex probed by Gaussian accelerated molecular dynamics and free energy landscapes. *RSC Advances* **2022,** *12* (3), 1742-1757.

43. Courtney, T.; Deiters, A., Recent advances in the optical control of protein function through genetic code expansion. *Curr Opin Chem Biol* **2018,** *46*, 99-107.

44. Rahman, S. M. T.; Zhou, W.; Deiters, A.; Haugh, J. M., Optical control of MAP kinase kinase 6 (MKK6) reveals that it has divergent roles in pro-apoptotic and anti-proliferative signaling. *Journal of Biological Chemistry* **2020,** *295* (25), 8494-8504.

45. Runtuwene, V.; van Eekelen, M.; Overvoorde, J.; Rehmann, H.; Yntema, H. G.; Nillesen, W. M.; van Haeringen, A.; van der Burgt, I.; Burgering, B.; den Hertog, J., Noonan syndrome gain-of-function mutations in NRAS cause zebrafish gastrulation defects. *Dis Model Mech* **2011,** *4* (3), 393-399.

46. Cirstea, I. C.; Kutsche, K.; Dvorsky, R.; Gremer, L.; Carta, C.; Horn, D.; Roberts, A. E.; Lepri, F.; Merbitz-Zahradnik, T.; König, R.; Kratz, C. P.; Pantaleoni, F.; Dentici, M. L.; Joshi, V. A.; Kucherlapati, R. S.; Mazzanti, L.; Mundlos, S.; Patton, M. A.; Silengo, M. C.; Rossi, C.; Zampino, G.; Digilio, C.; Stuppia, L.; Seemanova, E.; Pennacchio, L. A.; Gelb, B. D.; Dallapiccola, B.; Wittinghofer, A.; Ahmadian, M. R.; Tartaglia, M.; Zenker, M., A restricted spectrum of NRAS mutations causes Noonan syndrome. *Nat Genet* **2010,** *42* (1), 27-9.

47. Tajan, M.; Paccoud, R.; Branka, S.; Edouard, T.; Yart, A., The RASopathy Family: Consequences of Germline Activation of the RAS/MAPK Pathway. *Endocrine Reviews* **2018,** *39* (5), 676-700.

48. Rauen, K. A., The RASopathies. *Annu Rev Genomics Hum Genet* **2013,** *14*, 355-369.

49. Kimmel, C. B.; Ballard, W. W.; Kimmel, S. R.; Ullmann, B.; Schilling, T. F., Stages of embryonic development of the zebrafish. *Dev Dyn* **1995,** *203* (3), 253-310.

50. Jhang, W. K.; Choi, J. H.; Lee, B. H.; Kim, G. H.; Yoo, H. W., Cardiac Manifestations and Associations with Gene Mutations in Patients Diagnosed with RASopathies. *Pediatr Cardiol* **2016,** *37* (8), 1539-1547.

51. Calcagni, G.; Limongelli, G.; D'Ambrosio, A.; Gesualdo, F.; Digilio, M. C.; Baban, A.; Albanese, S. B.; Versacci, P.; De Luca, E.; Ferrero, G. B.; Baldassarre, G.; Agnoletti, G.; Banaudi, E.; Marek, J.; Kaski, J. P.; Tuo, G.; Russo, M. G.; Pacileo, G.; Milanesi, O.; Messina, D.; Marasini, M.; Cairello, F.; Formigari, R.; Brighenti, M.; Dallapiccola, B.; Tartaglia, M.; Marino, B., Cardiac defects, morbidity and mortality in patients affected by RASopathies. CARNET study results. *Int J Cardiol* **2017,** *245*, 92-98.

52. Altmüller, F.; Lissewski, C.; Bertola, D.; Flex, E.; Stark, Z.; Spranger, S.; Baynam, G.; Buscarilli, M.; Dyack, S.; Gillis, J.; Yntema, H. G.; Pantaleoni, F.; van Loon, R. L. E.; MacKay, S.; Mina, K.; Schanze, I.; Tan, T. Y.; Walsh, M.; White, S. M.; Niewisch, M. R.; García-Miñaúr, S.; Plaza, D.; Ahmadian, M. R.; Cavé, H.; Tartaglia, M.; Zenker, M., Genotype and phenotype spectrum of NRAS germline variants. *European Journal of Human Genetics* **2017,** *25* (7), 823-831.

53. Riller, Q.; Rieux-Laucat, F., RASopathies: From germline mutations to somatic and multigenic diseases. *Biomedical Journal* **2021,** *44* (4), 422-432.

54. Aoki, Y.; Niihori, T.; Banjo, T.; Okamoto, N.; Mizuno, S.; Kurosawa, K.; Ogata, T.; Takada, F.; Yano, M.; Ando, T.; Hoshika, T.; Barnett, C.; Ohashi, H.; Kawame, H.; Hasegawa, T.; Okutani, T.; Nagashima, T.; Hasegawa, S.; Funayama, R.; Nagashima, T.; Nakayama, K.; Inoue, S.-i.; Watanabe, Y.; Ogura, T.; Matsubara, Y., Gain-of-Function Mutations in RIT1 Cause Noonan Syndrome, a RAS/MAPK Pathway Syndrome. *The American Journal of Human Genetics* **2013,** *93* (1), 173-180.

55. Razzaque, M. A.; Komoike, Y.; Nishizawa, T.; Inai, K.; Furutani, M.; Higashinakagawa, T.; Matsuoka, R., Characterization of a novel KRAS mutation identified in Noonan syndrome. *American Journal of Medical Genetics Part A* **2012,** *158A* (3), 524-532.

56. Bonetti, M.; Paardekooper Overman, J.; Tessadori, F.; Noël, E.; Bakkers, J.; den Hertog, J., Noonan and LEOPARD syndrome Shp2 variants induce heart displacement defects in zebrafish. *Development* **2014,** *141* (9), 1961-1970.

57. Essner, J. J.; Amack, J. D.; Nyholm, M. K.; Harris, E. B.; Yost, H. J., Kupffer's vesicle is a ciliated organ of asymmetry in the zebrafish embryo that initiates left-right development of the brain, heart and gut. *Development* **2005,** *132* (6), 1247-1260.

58. Ramsdell, A. F., Left–right asymmetry and congenital cardiac defects: Getting to the heart of the matter in vertebrate left–right axis determination. *Developmental Biology* **2005,** *288* (1), 1-20.

59. Sherrid, M. V.; Männer, J.; Swistel, D. G.; Olivotto, I.; Halpern, D. G., On the Cardiac Loop and Its Failing: Left Ventricular Outflow Tract Obstruction. *J Am Heart Assoc* **2020,** *9* (3), e014857.

60. Zon, L. I.; Peterson, R. T., In vivo drug discovery in the zebrafish. *Nature Reviews Drug Discovery* **2005,** *4* (1), 35-44.

61. Darrah, K.; Wesalo, J.; Lukasak, B.; Tsang, M.; Chen, J. K.; Deiters, A., Small Molecule Control of Morpholino Antisense Oligonucleotide Function through Staudinger Reduction. *Journal of the American Chemical Society* **2021,** *143* (44), 18665-18671.