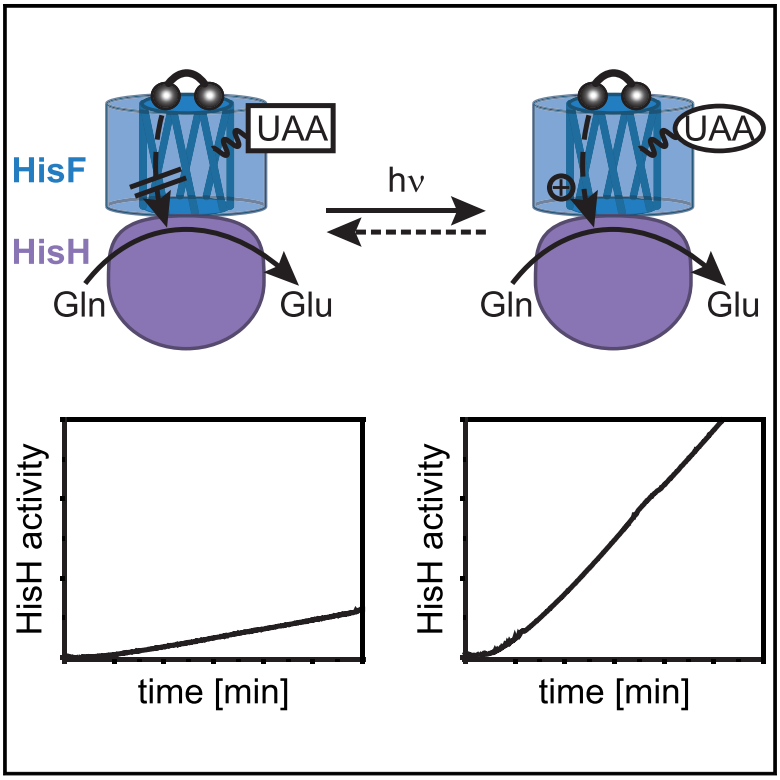
Article

Cell Chemical Biology

Light Regulation of Enzyme Allostery through Photo-responsive Unnatural Amino Acids

Graphical Abstract



Authors

Andrea C. Kneuttinger, Kristina Straub, Philipp Bittner, ..., Burkhard König, Rainer Merkl, Reinhard Sterner

Correspondence

reinhard.sterner@ur.de

In Brief

Within the bienzyme complex imidazole glycerol phosphate synthase, the activity of the subunit HisH was light controlled by photo-responsive unnatural amino acids located in HisF. Our findings demonstrate that light regulation of allostery can be used as a powerful tool for the spatiotemporal control of metabolic enzymes.

Highlights

- Allostery in the HisH-HisF bienzyme complex was regulated by light
- Photo-sensitive unnatural amino acids in HisF were used to control HisH activity
- HisH activity depends on the conformational organization of the catalytic H178

Kneuttinger et al., 2019, Cell Chemical Biology *26*, 1–14
November 21, 2019
https://doi.org/10.1016/j.chembiol.2019.08.006

CellPress

Please cite this article in press as: Kneuttinger et al., Light Regulation of Enzyme Allostery through Photo-responsive Unnatural Amino Acids, Cell Chemical Biology (2019), https://doi.org/10.1016/j.chembiol.2019.08.006

Cell Chemical Biology
Article

Light Regulation of Enzyme Allostery through Photo-responsive Unnatural Amino Acids

Andrea C. Kneuttinger,[1] Kristina Straub,[1] Philipp Bittner,[2] Nadja A. Simeth,[2,3] Astrid Bruckmann,[4] Florian Busch,[5] Chitra Rajendran,[1] Enrico Hupfeld,[1] Vicki H. Wysocki,[5] Dominik Horinek,[6] Burkhard König,[2] Rainer Merkl,[1] and Reinhard Sterner[1,7,*]

[1]Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
[2]Institute of Organic Chemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
[3]Centre for Systems Chemistry, Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands
[4]Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
[5]Department of Chemistry and Biochemistry and Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, OH 43210, USA
[6]Institute of Physical and Theoretical Chemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
[7]Lead Contact
*Correspondence: reinhard.sterner@ur.de
https://doi.org/10.1016/j.chembiol.2019.08.006

SUMMARY

Imidazole glycerol phosphate synthase (ImGPS) is an allosteric bienzyme complex in which substrate binding to the synthase subunit HisF stimulates the glutaminase subunit HisH. To control this stimulation with light, we have incorporated the photo-responsive unnatural amino acids phenylalanine-4′-azobenzene (AzoF), *o*-nitropiperonyl-*O*-tyrosine (NPY), and methyl-*o*-nitropiperonyllysine (mNPK) at strategic positions of HisF. The light-mediated isomerization of AzoF at position 55 (fS55AzoF^{E} ↔ fS55AzoF^{Z}) resulted in a reversible 10-fold regulation of HisH activity. The light-mediated decaging of NPY at position 39 (fY39NPY → fY39) and of mNPK at position 99 (fK99mNPK → fK99) led to a 4- to 6-fold increase of HisH activity. Molecular dynamics simulations explained how the unnatural amino acids interfere with the allosteric machinery of ImGPS and revealed additional aspects of HisH stimulation in wild-type ImGPS. Our findings show that unnatural amino acids can be used as a powerful tool for the spatiotemporal control of a central metabolic enzyme complex by light.

INTRODUCTION

In the last decade, regulation of enzyme activity by light has received increasing attention in the field of synthetic biology. Various approaches have been presented, which range from the binding of light-responsive ligands at the active site to the fusion of an enzyme with a light-oxygen-voltage sensing domain (Szymański et al., 2013; Baker and Deiters, 2014; Hüll et al., 2018; Losi et al., 2018; Lachmann et al., 2019; Schmermund et al., 2019). Less consideration has been dedicated to the photo-control of allosteric interactions, which are crucial regulatory features of enzymes in practically all metabolic pathways (Kastritis and Gavin, 2018). Allostery in enzyme complexes describes the binding of a ligand to one subunit by which the activity of another, associated subunit is influenced (Makhlynets et al., 2015). Thus, light regulation of allostery becomes most interesting for the temporal control of synthetic processes in industrial biocatalysis, such as in enzyme cascades that generally mimic metabolic pathways (Schmidt-Dannert and Lopez-Gallego, 2016).

We have recently described photo-control of allostery within the bienzyme complex imidazole glycerol phosphate synthase (ImGPS) from *Thermotoga maritima* (Figure 1A) (Kneuttinger et al., 2018). ImGPS consists of the glutaminase subunit HisH, which lacks measurable activity as a monomer, and the cyclase subunit HisF, which poorly activates HisH by complexation (List et al., 2012). Although the active sites of the two subunits are 25 Å apart (Douangamath et al., 2002), significant allosteric stimulation of HisH is initiated by binding of the HisF substrate *N′*-[(5′-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (PrFAR) (Lisi et al., 2017) or its analog *N′*-[(5′-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (ProFAR) (List et al., 2012). Several computational studies have described an allosteric network that connects the active sites of HisF and HisH and might transmit the stimulation signal (Rivalta et al., 2012; VanWart et al., 2012; Lisi et al., 2016; Negre et al., 2018). At the stimulation endpoint, chemical activation of glutamine catalysis in HisH has been postulated to occur through a backbone flip in the substrate binding site that leads to the stabilization of the oxyanion reaction intermediate (Chaudhuri et al., 2001, 2003; Lipchock and Loria, 2010; Rivalta et al., 2012). Activated HisH then hydrolyzes glutamine to glutamate and ammonia, which subsequently travels through an intermolecular channel to the active site of HisF (Douangamath et al., 2002). Ammonia there reacts with PrFAR to produce imidazole glycerol phosphate (ImGP) and 5-aminoimidazol-4-carboxamidribotide (AICAR), which are

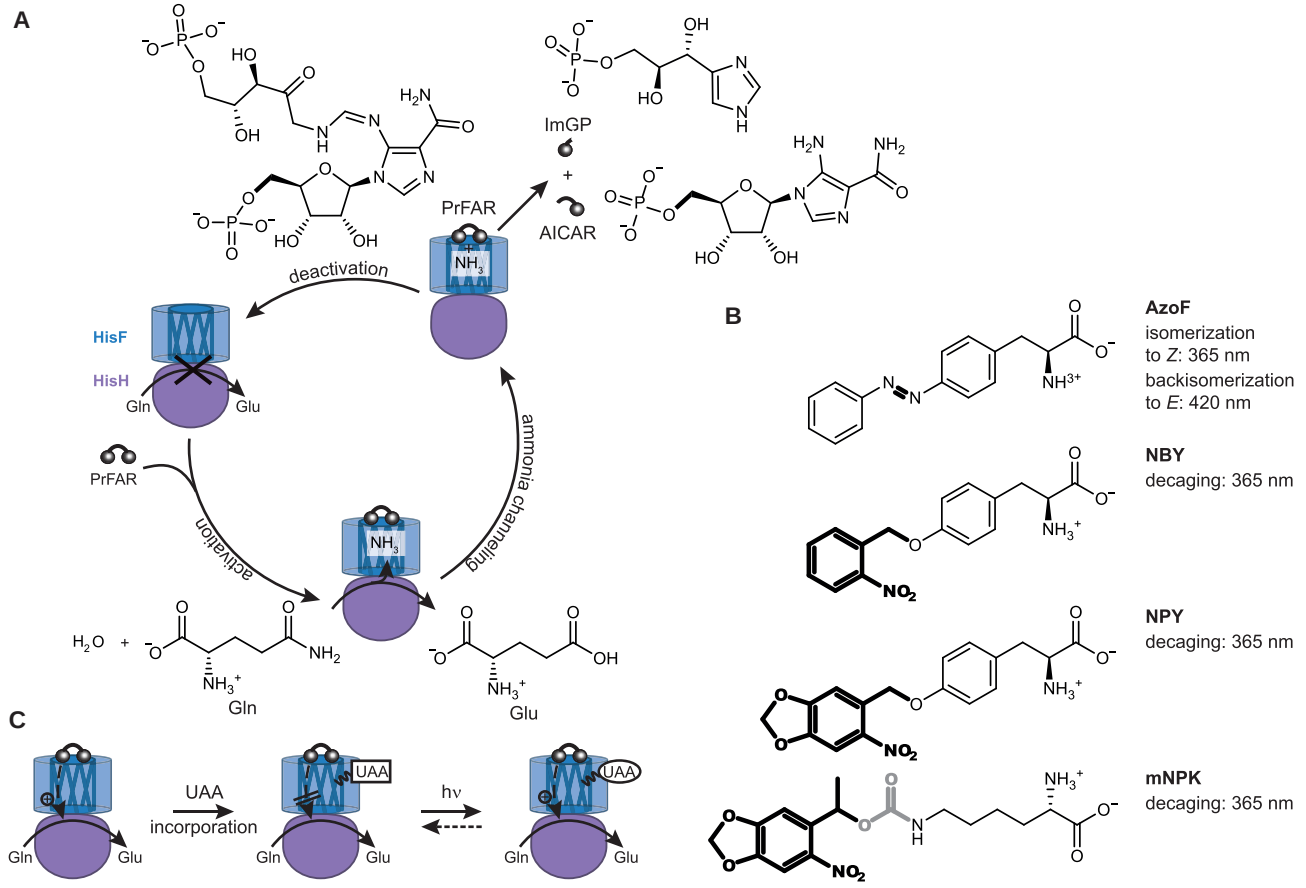


Figure 1. Implementation Strategy of Allosteric Light Regulation of ImGPS by UAAs

(A) Activity cycle of the bienzyme complex ImGPS. PrFAR (barbell) binding to HisF (blue) stimulates HisH (purple) (activation). Glutamine is then turned over to glutamate producing ammonia. After ammonia channeling through the interior of the central β barrel of HisF (dark blue), the active site of HisF cleaves PrFAR into 5-aminoimidazol-4-carboxamidribotide (AICAR) and ImGP (barbell fragments). Dissociation of the HisF products renders HisH inactive (deactivation).

(B) Four light-responsive unnatural amino acids (UAAs) were incorporated into HisF. AzoF isomerizes at the azo-bond (bold) from its *E* state (shown) to its *Z* state. NBY, NPY, and mNPK cleave off their caging group (bold) upon irradiation, which sets free tyrosine or lysine. (In the case of mNPK, carbon dioxide (gray) is also released.)

(C) To regulate the HisH reaction with light, a UAA is incorporated into HisF disrupting the activating allosteric machinery. Upon irradiation, the UAA isomerizes (AzoF) or loses its caging group (NBY, NPY, or mNPK), restoring allosteric activation of HisH activity. This light-induced activation step is irreversible using the caged UAAs, but can be reversed with AzoF (dashed arrow).

used for histidine and *de novo* purine biosynthesis, respectively. By associating a photo-responsive ligand to the active site of HisF, we inhibited proper binding of PrFAR and were, hence, able to photo-control the allosterically stimulated activity of HisH by a factor of 2 (Kneuttinger et al., 2018).

We have now established a more efficient and versatile light regulation of allostery within the ImGPS complex using translational incorporation of unnatural amino acids (UAAs). By means of orthogonal $tRNA_{CUA}$/aminoacyl-tRNA-synthetase pairs, UAAs can be site-specifically incorporated into proteins, providing new opportunities for synthetic biology (Liu and Schultz, 2010). Although light-responsive UAAs have been implemented by targeting positions that are directly associated with the active site of an enzyme either as a substrate (Lemke et al., 2007), a catalytic residue (Luo et al., 2017), or a residue close to substrate binding (Luo et al., 2017; Schlesinger et al., 2018; Wang et al., 2019), light-responsive UAAs have hitherto only rarely been used to control allostery (Luo et al., 2018).

The most commonly used light-responsive UAAs bear a natural amino acid as a scaffold caged with a photolabile protecting group (Figure 1B) (Curley and Lawrence, 1999; Courtney and Deiters, 2018; Bardhan and Deiters, 2019). Irradiation with UV light induces a decaging reaction, setting free the natural amino acid (Klán et al., 2013). One of the first caged UAA that has been incorporated into proteins was *o*-nitrobenzyl-*O*-tyrosine (NBY) (Deiters et al., 2006). Disadvantages of this UAA, however, include slow and incomplete decaging, which could be improved by the addition of a methylenedioxy moiety to form *o*-nitropiperonyl-*O*-tyrosine (NPY) and by an additional methyl substituent (Berroy et al., 2001; Luo et al., 2017). Selective incorporation has also been accomplished for other UAAs using, for example, serine (Lemke et al., 2007), cysteine (Uprety et al., 2014), or lysine (methyl-*o*-nitropiperonyllysine [mNPK]) (Gautier et al., 2010) as a scaffold. A general drawback of caged UAAs from the perspective of light regulation is the irreversibility of the decaging reaction. For reversible photo-control, the azobenzene moiety has

been successfully integrated into many biologically relevant molecules (Morstein et al., 2019; Schmermund et al., 2019) and specifically incorporated into proteins as a derivative of phenylalanine (AzoF) (Bose et al., 2006). Irradiation with UV light converts its *E* isomer into a *Z* isomer, which is thermodynamically less stable and reverts upon irradiation with light above 400 nm (Zimmerman et al., 1958). The switching occurs rapidly, with high quantum yields and little photobleaching (Szymański et al., 2013).

Our strategy to photo-control allosteric stimulation of HisH in ImGPS via the incorporation of UAAs in HisF (AzoF, NBY, NPY, and mNPK) is schematically outlined in Figure 1C. With the intention to disturb activation of HisH by HisF, we chose positions close to three known sites that are part of the previously described allosteric network (Rivalta et al., 2012; VanWart et al., 2012; Lisi et al., 2016; Negre et al., 2018). Furthest away from the HisH active site, the highly flexible loop 1 of HisF is thought to be involved in coupling both catalytic activities of HisF and HisH (Douangamath et al., 2002). Mutational studies (Beismann-Driemeyer and Sterner, 2001) and NMR experiments (Lisi et al., 2016) have supported the significance of loop 1 and suggested that its dynamics (Lisi et al., 2016) and interactions (Rivalta et al., 2012) with hydrophobic residues in the core of HisF change upon PrFAR binding. The other two sites are associated with PrFAR-stimulated motions at the HisH:HisF interface. First, for ammonia to pass into the intermolecular channel, the gate at the entrance, composed of residues fR5, fE46, fK99, and fE167 (f for HisF), needs to open (Chaudhuri et al., 2001; Douangamath et al., 2002). Similarly, PrFAR promotes a so-called "breathing motion" hinged at the cation-π interaction fR249-hW123 (h for HisH) that opens the interface above the glutamine binding site, possibly to allow for glutamine binding (Amaro et al., 2007; Rivalta et al., 2012).

Based on these considerations, we have incorporated the photo-responsive UAAs AzoF, NBY, NPY, and mNPK into positions close to loop 1, the ammonia channel, and the hinge region, and identified three UAA-HisFs that showed the potential to regulate HisH activity by light. We then characterized these proteins with respect to their structure, stability, and function, and optimized the isomerization and decaging processes. In the following, we could directly light regulate HisH activity 10-fold by our AzoF-HisF and 4- to 6-fold by our caged UAA-HisFs. Finally, molecular dynamics (MD) simulations revealed how each UAA affects HisH activity.

RESULTS

Identification of Three UAA-HisFs Leading to Light-Dependent HisH Activity

We applied three criteria to identify suitable positions in HisF for the incorporation of AzoF, NBY, and mNPK. First, the position in HisF should at least be 10 Å away from the HisH active site avoiding direct interactions of the UAA with catalytic residues. Moreover, UAAs should not hamper substrate binding in either the HisF or HisH active sites, nor impair the overall structure of the ImGPS complex. Based on these criteria, we tested ten positions in *T. maritima* HisF for the incorporation of light-responsive UAAs. The positions include conserved (*) residues and are localized in or close to the allosteric sites of loop 1 (fK13, *fK19, *fF23, fS29, fL35, *fY39, and *fS55), the ammonia gate (*fR5 and *fK99), and the interface hinge (fD74) (Figure 2A). NBY was used to replace tyrosine *fY39, phenylalanine *fF23, and aspartate fD74 (a phenylalanine in yeast ImGPS), whereas mNPK was used to replace lysines fK13, *fK19, and *fK99. For the design of AzoF replacements, we first calculated its rotamer library and incorporated AzoF into each of the ten positions *in silico*. The nine residues where the rotamers minimally overlapped with van der Waals radii of neighboring residues were replaced by AzoF (fR5, fK13, fK19, fF23, fS29, fL35, fY39, fS55, and fK99). Fifteen UAA-HisFs were then expressed in *Escherichia coli* cells using previously designed $tRNA_{CUA}$/aminoacyl-tRNA-synthetase pairs; these pairs have been shown to scarcely incorporate endogenous, natural amino acids, while each UAA was incorporated with high efficiency (Bose et al., 2006; Deiters et al., 2006; Gautier et al., 2010). Purification resulted in yields of 1–87 mg/L expression medium. UAA incorporation was analyzed by tryptic digest coupled to mass spectrometry (MS) analysis and for fK13AzoF with UV-visible spectroscopy (Figure S1). The two variants fK13mNPK and fK19mNPK were discarded after this step, because we could not confirm their identity. For the sake of brevity, we use the term fposUAA to name a HisF protein containing a UAA at position pos, e.g., fS55AzoF designates the HisF protein with AzoF at position S55; likewise ImGPS(fS55AzoF) designates the ImGPS complex containing AzoF at position S55 of HisF.

The remaining 13 UAA-HisFs were screened for their ability to activate HisH in the presence of ProFAR in their "as isolated" states (caged or *E*), after irradiation with UV light at a wavelength of 365 nm (decaged or *Z*) and, for AzoF-HisFs, after additional irradiation with visible light at a wavelength of 420 nm (*E*) (Figure 2B). Suitable candidates for in-depth characterization were selected according to the following criteria: (1) At least 20% wild-type (WT) HisH activity had to be retained in ImGPS complexes containing the irradiated caged UAA-HisFs or the more active isomer of AzoF; (2) HisH activity was altered at least 1.5-fold upon irradiation (light regulation factor [LRF]). HisH activities in ImGPS(fK19AzoF), ImGPS(fF23AzoF), ImGPS(fY39AzoF), ImGPS(fF23NBY), and ImGPS(fY39NBY) did not reach 20% WT activity, though irradiation of fY39NBY should in principle restore the WT situation. Native MS suggests that considerable quantities (20%–60%) of reduced NBY existed in both fF23NBY and fY39NBY, rationalizing missing decaging. Moreover, also the remaining intact NBY-HisF proteins could not be decaged (Figure S2). However, during the course of these studies and as an alternative to NBY, the UAA NPY was introduced and shown to possess higher decaging efficiencies (Luo et al., 2017). Indeed, HisH activities in ImGPS(fF23NPY) and ImGPS(fY39NPY) were >20% WT after irradiation (Figure 2B). In the second selection step, LRFs of HisH in ImGPS(fF23NPY), ImGPS(fY39NPY), and ImGPS containing the remaining eight UAA-HisFs were measured and the results were displayed in a histogram (Figure 2C). Only four UAA-HisFs led to LRFs >1.5, of which fF23NPY was the only caged UAA-HisF whose decaging did not restore WT-HisH activity. This result was rationalized by a control experiment with fF23Y, which showed that a tyrosine residue at position 23 leads to a drastic reduction to ~30% of WT-HisH activity. We therefore excluded fF23NPY from our further studies and finally selected fS55AzoF, fY39NPY, and fK99mNPK (Figure 2D) as most promising UAA-HisFs for light regulation of ImGPS allostery.

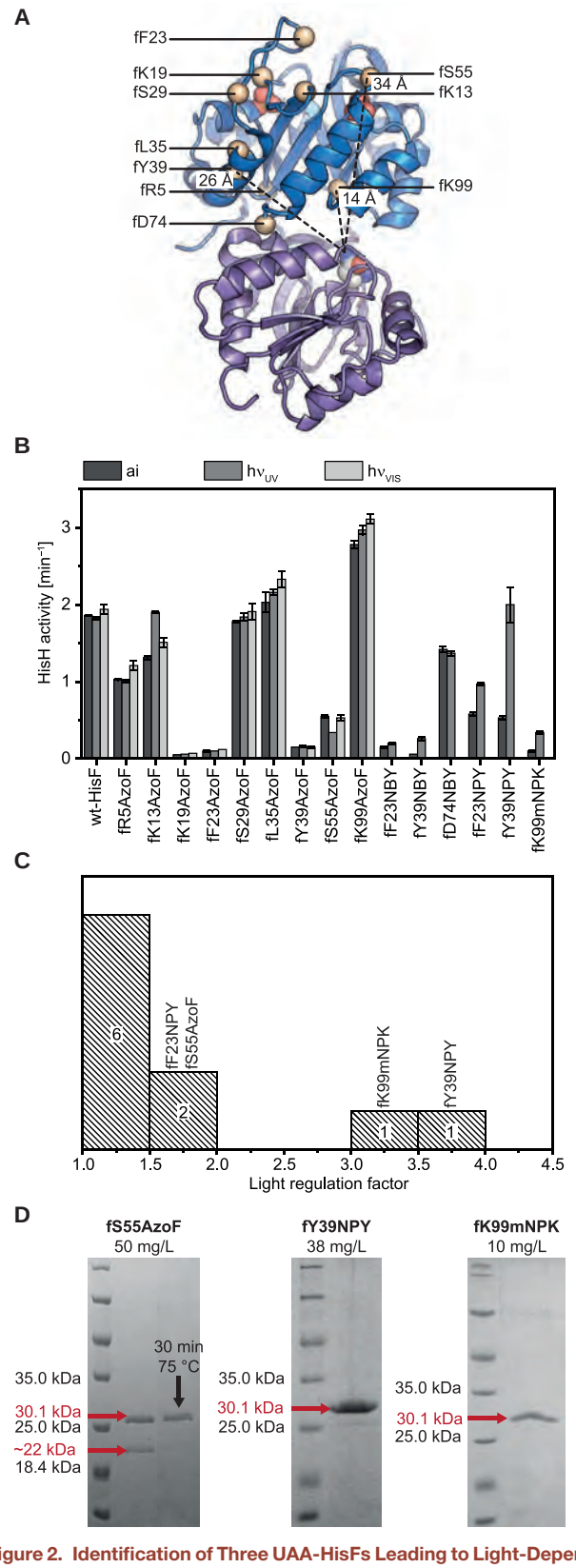


Figure 2. Identification of Three UAA-HisFs Leading to Light-Dependent HisH Activity

(A) Selected positions for incorporation of light-responsive UAAs are shown in beige in the ImGPS overall structure, with HisF in blue and HisH in purple (PDB: 1gpw; Douangamath et al., 2002). The distances of the three finally selected positions to the substrate glutamine, bound at the active site of HisH (super-positioned from PDB: 3zr4; List et al., 2012) are shown.

(B) HisH activities (pH 8.5) in ImGPS complexes containing WT-HisF and UAA-HisFs in the "as isolated" (ai) state, after irradiation with UV light ($h\nu_{UV}$, 365 nm) and, for AzoF-HisFs, after additional irradiation with visible light ($h\nu_{VIS}$, 420 nm). Mean ± SEM for three technical replicates are shown. Incorporation of the UAA at each position was confirmed by tryptic digest coupled to liquid mass spectrometry (MS) analysis (see also Figure S1). The two NBY proteins fF23NBY and fY39NBY showed low decaging efficiencies (see also Figure S2).

(C) By comparison of HisH activities before and after irradiation, light-regulation factors (LRFs) for each of the ten UAA-complexes with >20% WT-HisH activity were calculated and plotted in a histogram. Four proteins stood out with LRFs higher than 1.5.

(D) 12.5% SDS-PAGE of 3 μg protein confirming >90% purity of fS55AzoF, fY39NPY, and fK99mNPK and expression yield per liter TB medium. An additional protein in fS55AzoF, identified as truncated HisF in LC-MS analysis, was removed by a heat step of 75°C for 30 min.

UAAs Weaken HisF Stability and Activity but Do Not Impact Pr(o)FAR or Glutamine Binding

Before analyzing light regulation in detail, we characterized the non-irradiated UAA-HisFs fS55AzoF, fY39NPY, and fK99mNPK with respect to structure, stability, and function. Circular dichroism spectroscopy and analytical size-exclusion chromatography demonstrated that the three UAA-HisFs are properly folded monomers and form stoichiometric complexes with WT-HisH (Figures 3A–3C). Crystal structure analysis of fY39NPY (PDB: 6rtz) and fS55AzoF (PDB: 6ru0) in complex with WT-HisH showed that only the tyrosine and phenylalanine moieties could be resolved (Figures S3A–S3B), confirming the high flexibility of the UAA side chains, as observed in MD simulations (Figures S3C–S3F). Nevertheless, in fY39NPY, an additional electron density cloud indicated the approximate position of the *o*-nitropiperonyl moiety. This moiety shoved aside fH228, causing a destabilization of helix α8′ in HisF (Figure S3G). As a consequence, certain residues of this helix took on unallowed conformations in the Ramachandran plot (Lovell et al., 2003) (Figure S3H). Similarly, the AzoF moiety led to unallowed conformations of residues in loop 2 of fS55AzoF (Figure S3I). These structural effects translated into thermal destabilization: fS55AzoF and fY39NPY, as well as fK99mNPK, which contains a disturbed α4 helix in MD simulations (Figure S3J), showed decreased denaturation midpoints compared with WT-HisF that could be reversed upon irradiation for both caged UAA-HisFs (Figure 3D). Using steady-state enzyme kinetics, we observed that fS55AzoF and fY39NPY, which are located relatively close to the active site of HisF, display reduced (3- to 9-fold for fS55AzoF and 15-fold for fY39NPY) turnover numbers (k_{cat}), while fK99mNPK, which is far remote from the active site has a k_{cat} value similar to that of WT-HisF (~173 min^{-1}) (Table 1; Figure S4). Moreover, we found that the Michaelis constants (K_M^{PrFAR}) of all UAA-HisFs were identical to WT-HisF (~2.5 μM), indicating that substrate binding to the active site was not compromised.

Before we started with the analysis of the allosteric control of HisH activity by the three selected UAA-HisFs, we first measured HisH activities in ImGPS complexes containing these UAA-HisFs. The results showed that, albeit the k_{cat} values were decreased by 7- to 16-fold in ImGPS(fS55AzoF), 12-fold in ImGPS(fY39NPY), and 6-fold in ImGPS(fK99mNPK) compared with WT-HisH (~17 min^{-1}), the K_M^{Gln} values were unaltered (~0.4 mM) indicating that the affinity for the substrate

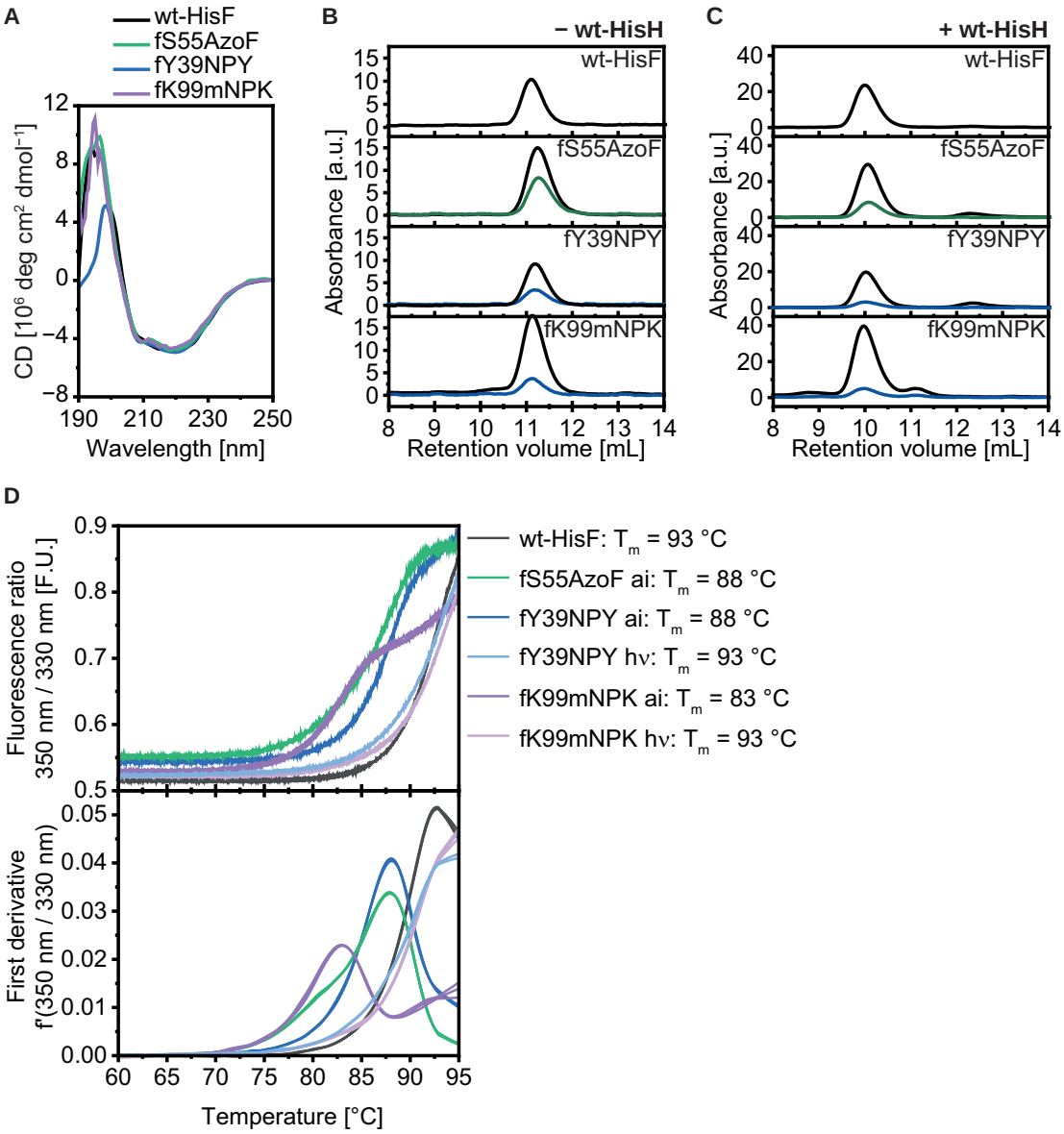


Figure 3. UAAs Weaken HisF Stability but Not the Overall Fold of HisF

(A) Far-UV circular dichroism spectra indicate an intact overall fold of each UAA-HisF.

(B) Analytical gel filtration indicates that each UAA-HisF is monomeric and rules out soluble aggregates that might lower HisF activity.

(C) Analytical gel filtration indicates that each UAA-HisF forms a stoichiometric complex with HisH. Protein absorbance during chromatography was monitored at 280 nm (black). AzoF, NPY, and mNPK signals were detected at their characteristic absorption maxima of 334 nm (Bose et al., 2006) (green) and 360 nm (blue).

(D) Thermal stability measurements with Nano-DSF identified reduced T_m values for all three UAA-HisFs in their "as isolated" (ai) states. Upon irradiation (hν), WT-HisF stability was recovered for fY39NPY and fK99mNPK. Stability of fS55AzoF could only be measured in its *E* state (ai) because the *Z* state is not thermally stable. T_m values were mathematically determined at the transition midpoint (peak maximum of the first derivative). For further structural evaluation see also Figure S3.

glutamine was not compromised (Table 1; Figure S4). Moreover, the ProFAR concentrations that were required for the UAA-HisFs to induce 50% HisH activity (K_{ac}) were also not changed (~29 μM) (Table 1; Figure S4).

UAA-HisFs Allow for Direct Photo-control of HisH Activity

Next, we identified optimum reaction conditions with respect to pH and mode of irradiation. Comparison of the data shown in Table 1 and Figure 2C indicates that the LRFs of fS55AzoF, fY39NPY, and fK99mNPK increased slightly (~1.5-fold) from pH 8.5 to 7.0. Moreover, up to now, we irradiated caged UAA-HisFs for 20 min with a conventional light bulb, to achieve maximum decaging. Using a high-power LED (365 nm) instead, we were able to accelerate this process by ~10-fold. In addition, the new light source improved the switching efficiency of $fS55AzoF^E$ to $fS55AzoF^Z$ (Figure 4A). Whereas we only reached a photostationary state composed of 52% *E* and 48% *Z* with

Table 1. Steady-State Kinetics of WT-ImGPS in Comparison with ImGPS(fS55AzoF), ImGPS(fY39NPY), and ImGPS(fK99mNPK)

PrFAR-Dependent HisF Activity[a]				
Protein	State	k_{cat} (min^{-1})	K_m^{PrFAR} (μM)	k_{cat}/K_m (M^{-1} s^{-1})
WT-HisF		172.6 ± 1.6	2.5 ± 0.1	11.5 ± 0.5 × 10^5
fS55AzoF	*E*	48.6 ± 0.9	4.8 ± 0.1	1.7 ± 0.0 × 10^5
fS55AzoF	*Z*	18.6 ± 0.	2.4 ± 0.2	1.3 ± 0.1 × 10^5
fY39NPY	caged	11.3 ± 0.4	3.4 ± 0.5	0.6 ± 0.1 × 10^5
fK99mNPK	caged	112.4 ± 4.7	3.2 ± 0.5	5.9 ± 0.9 × 10^5
Glutamine-Dependent ImGPS Activity[a]				
Protein	State	k_{cat} (min^{-1})	K_m^{Gln} (mM)	k_{cat}/K_m (M^{-1} s^{-1})
WT-ImGPS		41.4 ± 1.1	0.78 ± 0.05	8.6 ± 1.1 × 10^2
ImGPS(fS55AzoF)	*E*	6.0 ± 0.2	0.23 ± 0.03	4.3 ± 0.6 × 10^2
ImGPS(fS55AzoF)	*Z*	2.6 ± 0.2	0.33 ± 0.11	1.3 ± 0.4 × 10^2
ImGPS(fY39NPY)	caged	3.5 ± 0.2	0.25 ± 0.05	2.3 ± 0.5 × 10^2
ImGPS(fK99mNPK)	caged	7.3 ± 0.4	0.41 ± 0.09	3.0 ± 0.7 × 10^2
ProFAR-Dependent HisH Activity (pH 7.0)[b]				
Protein	State	k_{cat} (min^{-1})	K_{ac}^{ProFAR} (μM)	LRF
WT-ImGPS		16.8 ± 0.3	32.8 ± 1.2	
ImGPS(fS55AzoF)	*E*	5.4 ± 0.4	19.5 ± 3.8	2.3
	Z	2.3 ± 0.2	35.1 ± 5.0	
ImGPS(fY39NPY)	caged	1.2 ± 0.1	32.3 ± 4.2	5.9
	decaged	7.0 ± 0.5	28.7 ± 4.2	
ImGPS(fK99mNPK)	caged	3.5 ± 0.3	19.6 ± 3.7	4.0
	decaged	14.0 ± 0.9	32.1 ± 3.2	

See also Figures S4 and S6.

[a]Values ± SE for k_{cat} and K_m were determined by Michaelis-Menten fitting of the mean ± SEM for at least two technical replicates. Values ± SE for k_{cat}/K_m were calculated according to the Gaussian law of error propagation.

[b]Values ± SE for k_{cat} and K_{ac}^{ProFAR} were determined by fitting the mean ± SEM for at least two technical replicates to a hyperbolic function. LRF = k_{cat}^{E}/k_{cat}^{Z} or $k_{cat}^{decaged}/k_{cat}^{caged}$.

the conventional light bulb, the high-power LED led to the formation of 11% *E* and 89% *Z*.

These results laid the foundation for direct light regulation of HisH activity. To this end, the glutaminase reaction of HisH was stimulated by ProFAR-binding to UAA-HisFs at pH 7.0 and 8.5. The reaction samples were then irradiated with our high-power LED (365 nm) while product formation was still in the linear range (Figure 4B). ImGPS(fS55AzoF) was irradiated a second time with visible light (420 nm), to switch AzoF back into its *E*-enriched state. In parallel, reactions were followed with as isolated complexes and pre-irradiated complexes as negative and positive controls, respectively. ImGPS(fY39NPY) and ImGPS(fK99mNPK) showed only minor pH dependency of their LRFs and allowed us to control HisH activity by ~6- and ~3-fold, respectively, consistent with LRFs of ~6 and ~4 as determined in K_{ac} measurements (Table 1). Activities before and after irradiation matched well with the negative and positive controls. ImGPS(fS55AzoF), on the other hand, showed a stronger pH dependency and allowed us to reversibly control HisH activity by ~10-fold at pH 7.0 and by ~2-fold at pH 8.5. We furthermore tested WT-HisF and as additional control fF23NPY and fF23AzoF; however, neither of the activities were affected >1.5-fold by light (Figure S5). HisH activity was hence unambiguously activated by light with fY39NPY and fK99mNPK, and even more efficiently switched off and on in a pH-dependent manner by fS55AzoF.

Decaging Depends on the Protein Environment

Although HisH can be light regulated with the help of fY39NPY and fK99mNPK, irradiation did not lead to complete decaging of these two UAA-HisFs. Following irradiation of ImGPS complexes containing the two UAA-HisFs, k_{cat} values of HisH reached only ~40% and ~80% of the WT k_{cat} value (~17 min^{-1}) (Table 1). Native MS was used to check whether incomplete decaging was caused by reduction of the nitropiperonyl groups of NPY and mNPK (Figure S6). Compared with NBY proteins (Figure S2), reduction of NPY and mNPK decreased from 20% and 60% to ≤10%, and therefore did not explain incomplete recovery of WT-HisH activity. However, native MS showed that decaging yields were lower for fY39NPY (27%) than for fK99mNPK (70%), consistent with HisH activities following irradiation. The amount of decaging seems to depend on the structural freedom of the UAA side chain: fK99 is part of the gate at the ammonia channel entrance, whereas fY39 is buried by secondary structure elements. The potential for light activation by our caged UAA-HisFs was hence limited by the protein conformation, and neither longer irradiation times nor a stronger light source led to complete decaging (Figure 4A).

The Two fS55AzoF Light Isomers Change Various Structural Characteristics of ImGPS to Different Extents

In the following, we performed MD simulations with PrFAR-bound ImGPS(fS55AzoF), ImGPS(fY39NPY), and ImGPS(fK99mNPK) to

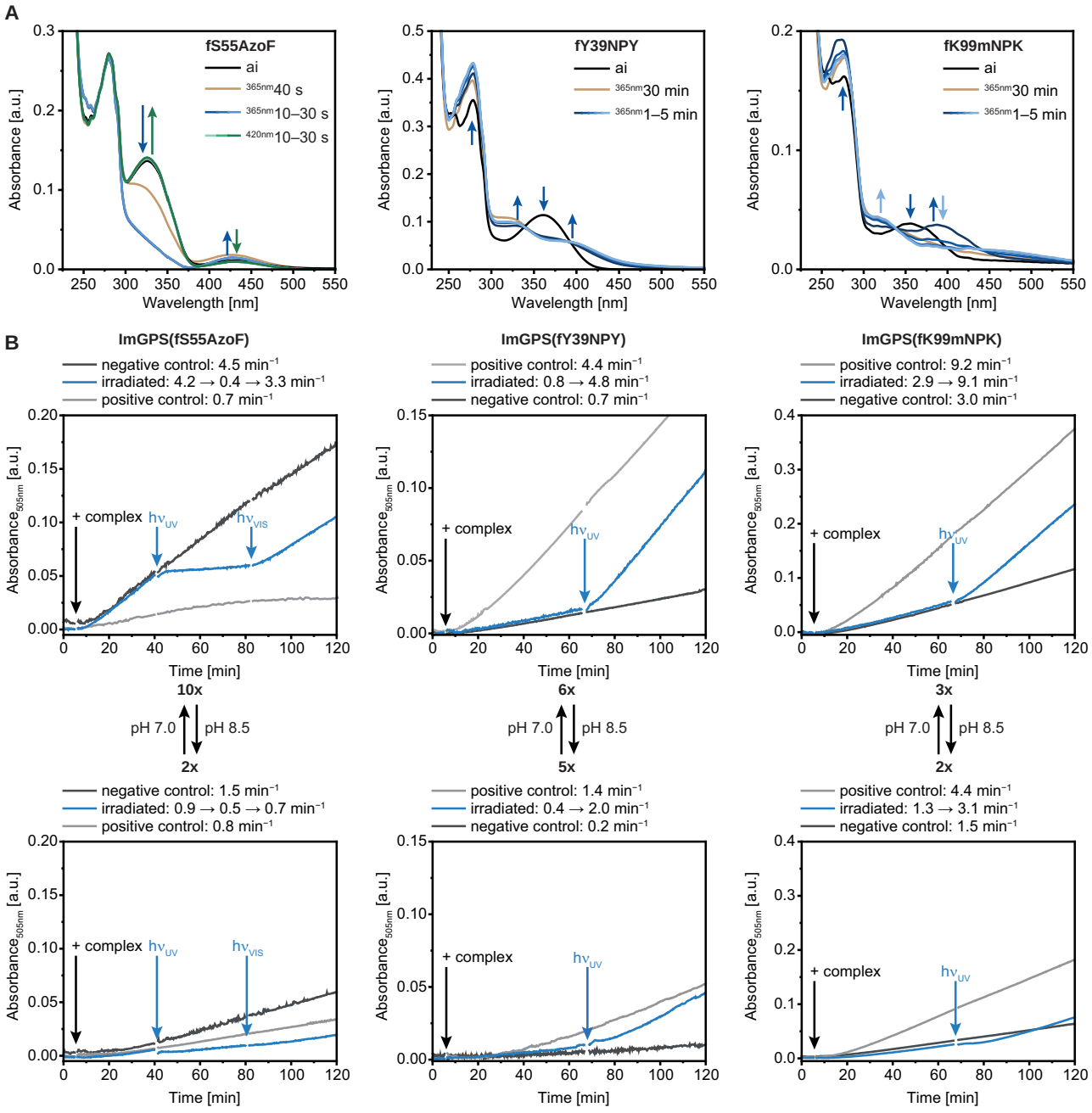


Figure 4. UAA-HisFs Allow for Direct Photo-control of HisH Activity

(A) As described in Bose et al. (2006), fS55AzoF showed two signals in addition to the protein peak at 280 nm in the as isolated (ai) state (black). The signal at ~330 nm corresponds to the population of its *E* state. Upon irradiation with UV light (365 nm), the maximum loses intensity as the *E*/*Z* equilibrium is shifted toward a *Z*-enriched photostationary state (blue arrows). The photostationary composition at this state, achieved with a high-power LED, was approximately 11% *E* to 89% *Z* (blue); with a conventional UV lamp (2 × 8 W), we only achieved 52% *E* and 48% *Z* (brown). Subsequent irradiation with visible light (420 nm) could reverse the equilibrium toward an *E*-enriched state (green arrows). Caged UAA-HisFs fY39NPY and fK99mNPK exhibited a clear peak at ~360 nm next to the protein peak at 280 nm. Upon irradiation with UV light, decaging could be followed due to a loss of intensity at this wavelength (downward blue arrows). In addition, two new peaks were formed (upward arrows), which were caused by the cleaved-off nitrobenzylic species in their monomeric and dimeric forms, respectively (Zuman and Shah, 1994). Formation of these species occurs simultaneously for NPY, but successively for mNPK (light blue arrows). The intensity at 360 nm reached a final minimum after 2 min (blue); with a conventional UV lamp (2 × 8 W), the final minimum was reached after 20–30 min (brown). The final minimum does not represent the fully decaged protein (see also Figure S6).

(B) Light-dependent HisH activities of ImGPS(fS55AzoF), ImGPS(fY39NPY), and ImGPS(fK99mNPK) were followed at pH 7.0 (upper panels) and at pH 8.5 (lower panels). Blue curves show the reaction course of the as isolated complexes that were irradiated with UV light (365 nm) for decaging (fY39NPY, fK99mNPK;

(legend continued on next page)

understand the underpinning mechanism of light regulation of HisH.

We started with UAA-HisF, in which the UAA is positioned furthest away (34 Å) from the HisH active site, fS55AzoF. We searched for reasons why HisH activity is lower in ImGPS(fS55AzoF^E) than in WT-ImGPS and why HisH activity is lower in ImGPS(fS55AzoF^Z) than in ImGPS(fS55AzoF^E) (Figure 5A). For this purpose, we derived normalized frequency distributions for the atom distances (d) between the central catalytic residue hH178 (Massière and Badet-Denisot, 1998; Raushel et al., 1999) and either hE180 or fD98, which are both essential for glutaminase catalysis (Zalkin and Smith, 1998; List et al., 2012), from MD simulations (Figure 6). We concentrated on these distances, because the stabilization of the hydrogen bond between hH178 and hE180, and the shortening of the hH178-fD98 spacing, was previously observed as a consequence of PrFAR binding in WT-HisF and the formation of the glutamyl-thioester reaction intermediate in WT-HisH (Myers et al., 2005). A 2D diagram of these frequency distributions revealed distinct clusters that represent conformational ensembles of hH178 (Figure 5B); representative MD snapshots are shown in Figure S7. In WT-ImGPS, three ensembles (ens$_{WT,1}$-ens$_{WT,3}$) were observed. In ens$_{WT,1}$ (prevalence 24%), hH178 assumes a pose with relatively short distances to both hE180 (d $\sim$ 3.6 Å) and fD98 (d $\sim$ 6.9 Å). This orientation suggests the formation of a direct hydrogen bond between hH178 and hE180 and leaves the possibility of a water-mediated hydrogen bond between hH178 and fD98. In ens$_{WT,2}$ (prevalence 44%), the hydrogen-bond distance between hH178 and hE180 remains unaltered, but the mean distance of hH178 to fD98 increases (d $\sim$ 8.6 Å), which makes the formation of a water-mediated hydrogen bond less likely. In ens$_{WT,3}$ (prevalence 32%) hH178 cannot interact with hE180 (d $\sim$ 5.1 Å) and is still far from fD98 (d $\sim$ 8.6 Å). Obviously, these three specific conformational ensembles of hH178 are optimal for HisH activity. In accordance with this assumption, ImGPS(fS55AzoF^E) adopts only one conformational ensemble (ens$_E$) resembling ens$_{WT,2}$, whereas ImGPS(fS55AzoF^Z) adopts four ensembles (ens$_{Z,1}$-ens$_{Z,4}$) where only ens$_{Z,1}$ resembles ens$_{WT,1}$.

Next, we concentrated on the h49-PGVG-52 motif (Figure 5C). In MD simulations of WT-ImGPS, PrFAR binding has been observed to increase the flexibility of the backbones of hG50 and hV51, which supposedly leads to a backbone flip of the hV51 amine, stabilizing the oxyanion reaction intermediates (Lipchock and Loria, 2010; Rivalta et al., 2012). Our root-mean-square fluctuation (RMSF) analyses of MD simulations albeit showed that these residues are rather rigid in all three ImGPS complexes. However, the adjacent hG52 backbone is more flexible in WT-ImGPS and ImGPS(fS55AzoF^E) than in the least active ImGPS(fS55AzoF^Z), which might allow it to act as oxyanion intermediate stabilizer. In addition, the motif adopts a different configuration in ImGPS(fS55AzoF^Z), with the consequence that hG52 leaves the glutamine binding site.

Finally, we wanted to track the breathing motion at the HisH:HisF interface that was previously shown to occur upon PrFAR binding to WT-HisF and leads to an opening of the interface above the glutamine binding site (Amaro et al., 2007; Rivalta et al., 2012; Negre et al., 2018). We considered the distance between hG52 of the glutamine binding site and the opposite fF120 as an indirect measure of this movement. The mean distances of hG52-fF120 suggest that the interface is more open in WT-ImGPS (d $\sim$ 10.7 Å) and in ImGPS(fS55AzoF^E) (d $\sim$ 11.3 Å) than in the least active ImGPS(fS55AzoF^Z) (d $\sim$ 8.9 Å) (Figure 5D). To elucidate implications of this motion for the structure of HisH, we visually compared RMSF values of secondary structure elements. We orientated the 3D structure such that the hinge region was located in the "left" half of the ImGPS complexes (Figure 5E). Apart from the active site loops of HisF, which are similarly flexible in all three structures, the "right" half of WT-ImGPS and ImGPS(fS55AzoF^E) containing the opening at hG52–fF120 is more flexible with higher RMSF values than the left half containing the hinge region. In contrast, right and left halves remained rigid in the least active ImGPS(fS55AzoF^Z). The mean RMSF values are in full agreement with this postulate: overall flexibility in WT-ImGPS (0.98 Å) and ImGPS(fS55AzoF^E) (1.00 Å) is higher than in ImGPS(fS55AzoF^Z) (0.49 Å). All differences of the mean values we determined here were low, but within the range we expected, due to a recent comprehensive study comparing apo and holo structures (Clark et al., 2019).

In summary, our MD simulations suggest that fS55AzoF^Z compromises the overall structure of HisF most strongly, and affects the orientation of the catalytic residue hH178 and of the oxyanion hole, which are both crucial for HisH function (Figure 5F). In addition, the breathing motion and the concomitant rearrangement of the ImGPS structure seem less pronounced than in WT-ImGPS. Compared with fS55AzoF^Z, the disadvantageous effects of fS55AzoF^E seem to be generally less severe, as indicated by a more structured catalytic site, a typical h49-PGVG-52 orientation, and a HisH:HisF arrangement that resemble the WT-ImGPS.

fY39NPY Causes a More Rigid ImGPS Complex

We continued our MD analyses with the UAA-HisF fY39NPY, in which NPY is located 26 Å away from the HisH active site. We searched for reasons why HisH activities are lower in caged ImGPS(fY39NPY) than in decaged ImGPS(fY39NPY), which is identical to WT-ImGPS (Figure 5A). We observed three structural ensembles of hH178, ens$_{Y,1}$ (prevalence 10%), ens$_{Y,2}$ (prevalence 34%), and ens$_{Y,3}$ (prevalence 49%) that resemble those in WT-ImGPS (Figure 5B). However, the h49-PGVG-52 motif was more rigid than in WT-ImGPS (Figure 5C) and, similarly as deduced for ImGPS(fS55AzoF^Z), the breathing motion at the HisH:HisF interface seems less extensive; the mean hG52-fF120 distance is $\sim$1.5 Å shorter than in WT-ImGPS (Figure 5D). Consequently, both right and left halves of ImGPS(fY39NPY) remain rigid (Figure 5E) and the mean RMSF of the full complex (0.93 Å) is lower than the WT-ImGPS reference value (0.98 Å).

1.5 min) or *E* to *Z* isomerization (fS55AzoF; 30 s). ImGPS(fS55AzoF) was irradiated a second time with visible light (420 nm; 1 min) for back-isomerization to *E*. Dark gray curves (negative control) and light gray curves (positive control) show the reaction courses of complexes that were not irradiated or pre-irradiated, respectively. Reaction rates for the individual curves are given above each chart, LRFs are given below (pH 7.0) or above (pH 8.5) the charts (see also Figure S5).

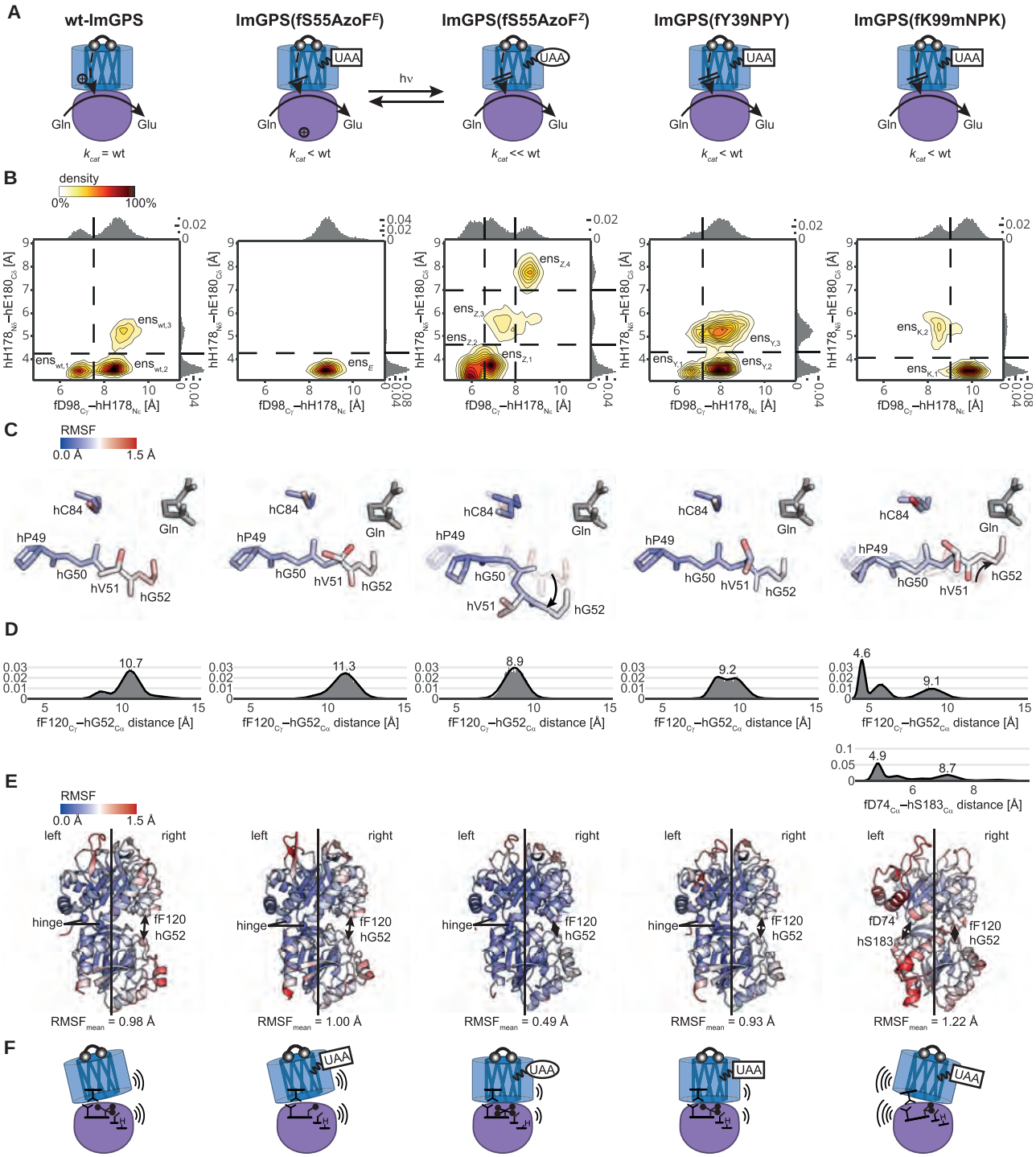


Figure 5. Comparative MD Simulation Analysis of WT-ImGPS, ImGPS(fS55AzoF^E), ImGPS(fS55AzoF^Z), ImGPS(fY39NPY), and ImGPS(fK99mNPK)

(A) Comparison of turnover numbers of HisH (see Table 1) shown to facilitate side-by-side comparison of MD results presented in (B–E).

(B) 1D and 2D frequency distributions for the distances $hH178_{N\delta}$-$hE180_{C\delta}$ and $fD98_{C\gamma}$-$hH178_{N\epsilon}$. The orientation of the hH178 residue in the 2D contour plots suggests for WT-ImGPS three ensembles ($ens_{wt,1}$-$ens_{wt,3}$), for ImGPS(fS55AzoF^E) only one (ens_E) resembling $ens_{wt,2}$, for ImGPS(fS55AzoF^Z) four ($ens_{Z,1}$-$ens_{Z,4}$) with $ens_{Z,1}$ resembling $ens_{wt,1}$, for ImGPS(fY39NPY) three ($ens_{Y,1}$-$ens_{Y,3}$) with all three resembling $ens_{wt,1}$-$ens_{wt,3}$, and for ImGPS(fK99mNPK) two ensembles ($ens_{K,1}$ and $ens_{K,2}$) with $ens_{K,2}$ resembling $ens_{wt,3}$. MD snapshots representative of each of the conformational ensembles are shown in Figure S7.

(legend continued on next page)

In summary, these findings suggest that fY39NPY impairs HisH activity by reducing the h49-PGVG-52 flexibility and the breathing motion (Figure 5F).

fK99mNPK Causes a Rearrangement of the HisH:HisF Interface

In the UAA-HisF fK99mNPK, mNPK is 14 Å away from the HisH active site. Again using MD, we searched for reasons why HisH activities are lower in caged ImGPS(fK99mNPK) than in decaged ImGPS(fK99mNPK), which is identical to WT-HisF (Figure 5A). Here, we observed only two conformational ensembles of hH178, $ens_{K,1}$ (prevalence 56%) and $ens_{K,2}$ (prevalence 29%), with the latter resembling the $ens_{WT,3}$ of WT-ImGPS (Figure 5B). Again, the deviation of these two conformations from the three specific conformations observed in WT-ImGPS seems to impair HisH activity. In addition, a comparison of representative MD snapshots with WT-ImGPS showed a drastic backbone shift for both $ens_{K,1}$ and $ens_{K,2}$ (Figure S7). Moreover, residues of the h49-PGVG-52 motif were also affected by a backbone shift (Figure 5C). Due to this observation, we analyzed the fD74-hS183 distance close to the interface hinge, in addition to the fF120-hG52 distance at the HisH:HisF interface (Figure 5D). While the breathing motion above the glutamine binding site is absent almost completely, the interface opened as far as ~3.7 Å at fD74-hS183. As a consequence, the left half of ImGPS(fK99mNPK) is more flexible than the left half of WT-ImGPS (Figure 5E). The right half is similarly flexible as the right half of WT-ImGPS, likely due to the motion between the different closing states ($d \sim 4.6$ Å $\leftrightarrow$ $d \sim 9.1$ Å) at fF120-hG52. The mean RMSF value of the full complex (1.22 Å) exceeds the WT-ImGPS value (0.98 Å) and is a further indicator for higher flexibility.

In summary, these findings suggest that fK99mNPK causes a drastic rearrangement of the HisH:HisF interface, whereupon the opening of the interface above the glutamine binding site is most prominently affected (Figure 5F).

DISCUSSION

In search for HisF positions that enable photo-control of the activity of its partner protein HisH with UAAs, we identified three promising candidates. Each of the positions is at least 10 Å apart from the HisH active site and, after UAA incorporation, neither Pr(o)FAR nor glutamine binding was affected. ImGPS(fS55AzoF) allowed us to reversibly and pH-dependently control HisH activity by 10-fold. ImGPS(fY39NPY) and ImGPS(fK99mNPK) achieved LRFs of 6 and 4. In direct comparison with photo-control at the active site of an enzyme, our approach of allosteric light regulation reached equivalent LRFs; e.g., TEV protease activity was regulated by 10-fold with NPY (Luo et al., 2017), and horseradish peroxidase activity by 1.5-fold with AzoF (Muranaka et al., 2002). Remarkably, a new computationally aided strategy that cages a residue in close proximity to the active site of an enzyme (CAGE-prox) reduced tumor growth in mice by ~5-fold upon light irradiation (Wang et al., 2019). Recently, AzoF and derivatives were also used to photo-control firefly luciferase from an allosteric position ~16 Å from its active site; however, only a 2-fold LRF was achieved (Luo et al., 2018). We have identified our candidates by choosing known positions close to allosteric sites in ImGPS; however, such knowledge might not be available for other potential target enzymes. Alternatively, computationally guided methods similar to CAGE-prox or the identification of allosteric regulation hotspots from coevolution studies (Reynolds et al., 2011) might be applicable.

Whereas AzoF incorporation and switching occurred straightforwardly, a large fraction of the caged UAA NBY was reduced in *E. coli* and its decaging efficiency was below 10%. Significant improvement was accomplished by the addition of a methylenedioxy moiety, yielding NPY (Luo et al., 2017). Distinct spectral features of NPY allowed easy verification of incorporation, its reduction was minimized in *E. coli*, and decaging increased by up to 80%. On the downside, high levels of decaged species were obtained before irradiation, possibly due to unintentional light exposure during protein synthesis or due to decaging by nitroreductases (Valiauga et al., 2017; Xie et al., 2017) in *E. coli*. Another drawback was the incomplete recovery of WT-HisH activity upon irradiation of ImGPS complexes containing fY39NPY and fK99mNPK. This might be caused by reaction of the cleavage products with the protein to form imines (compare Figure S2A); however, we could not detect any adducts in native MS (Figure S6A). Instead, decaging seems to depend on the protein environment. Specifically, neighboring residues such as fK19, fH228 and fR230 in fY39NPY or fD98, fE46, fR5, and fE167 in fK99mNPK might interfere with the photo-induced, acid-base catalyzed fragmentation process (Figure S6B) (Il'ichev and Wirz, 2000; Klán et al., 2013). Different orientations of the UAA, as observed for both NPY and mNPK in MD simulations, might interrupt fragmentation more or less strongly, so that different ratios of decaged proteins result for each position.

We performed MD simulations to understand how the different UAAs, positioned strategically close to known allosteric elements in HisF, affect HisH activity. In doing so, we also gained further insight into structural events at the stimulation endpoint in HisH. In PrFAR-bound WT-ImGPS, we identified three conformational ensembles of the central catalytic residue hH178. The combination of these three ensembles—with ~70% of hH178 in hydrogen bonding distance to hE180 and ~25% in water-mediated hydrogen bonding distance to fD98—generates the most active state of HisH. This combination is not present in ImGPS(fS55AzoF^{E}), in which only one distinct ensemble is

(C) Simulation endpoint orientation of the h49-PGVG-52 motif and the catalytic HisH residue hC84 color-coded by the residue-specific RMSF value; compare the color bar. The substrate glutamine (Gln) was super-positioned from PDB: 3zr4 (List et al., 2012) and is shown in gray. In ImGPS(fS55AzoF^{Z}) (middle panel) and ImGPS(fK99mNPK) (right panel), the WT localization of h49-PGVG-52 is indicated semi-transparently.

(D) Frequency distributions for the distance $fF120_{C\gamma}$-$hG52_{C\alpha}$, which is an indicator for the extent of the "breathing motion." In addition, frequency distributions for the control distance $fD74_{C\alpha}$-$hS183_{C\alpha}$ of residues close to the WT hinge region are given for ImGPS(fK99mNPK).

(E) Flexibility of secondary structure elements as indicated by their RMSF values. Mean RMSF values deduced from all atoms of the ImGPS complexes are given below their cartoon representation. For easier interpretation, the location of fF120 and hG52, and the hinge region are indicated.

(F) Graphical representation of the effects observed in (B)–(E). Changes of the interface opening, the flexibility of secondary structure elements, and the organization of active site residues in HisH (see Figure 6) are indicated schematically.

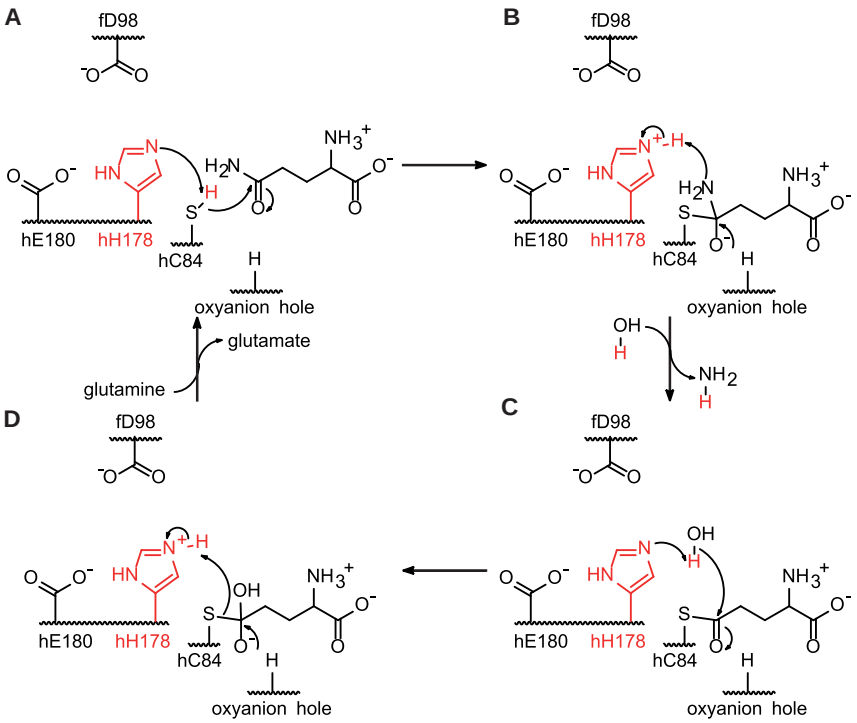


Figure 6. Reaction Mechanism of the Glutaminase HisH

(A) The active site of HisH harbors the catalytic triad consisting of hE180, hH178, and hC84 (Zalkin and Smith, 1998). HisF contributes a fourth residue that was shown to be essential for catalysis, fD98 (List et al., 2012). The central catalytic residue hH178 accepts the hydrogen from hC84, whereupon hC84 starts a nucleophilic attack of the substrate glutamine.

(B) The resulting oxyanion intermediate is stabilized by a residue of the oxyanion hole, while hH178 submits its hydrogen to the amino group of glutamine, setting free ammonia.

(C) hH178 accepts a hydrogen from water, which undergoes a nucleophilic attack with the glutamyl-thioester intermediate.

(D) The resulting oxyanion intermediate is stabilized by a residue of the oxyanion hole, while hH178 submits its hydrogen to hC84, restoring this catalytic residue and setting free glutamate. The mechanism was adapted from Massière and Badet-Denisot (1998).

adopted, or in ImGPS(fS55AzoF^{Z}), in which four distinct ensembles are observed. It is not entirely clear why this different combination of ensembles leads to reduced HisH activity in ImGPS(fS55AzoF^{E}) and in particular in ImGPS(fS55AzoF^{Z}). However, it is plausible to assume that a productive positioning of the catalytically central hH178, which acts as a proton shuffle during the reaction cycle (Figure 6) (Zalkin and Smith, 1998; Raushel et al., 1999), is essential for efficient glutaminase activity. Previous studies proposed a different process of HisH activation, which is caused by a rotation of the hV51 amine into the glutamine binding site that results in the formation of the oxyanion hole (Chaudhuri et al., 2001, 2003; Lipchock and Loria, 2010; Rivalta et al., 2012). With this in mind, we analyzed the conserved h49-PGVG-52 motif and observed significant conformational changes. However, in our MD simulations the backbone of hV51 remained rigid in WT-ImGPS and ImGPS(fS55AzoF^{E}). Only in ImGPS(fS55AzoF^{Z}), the backbone of the h49-PGVG-52 motif rotated, but this orientation was associated with the disruption of the glutamine binding site and an almost complete loss of HisH activity. These findings are in agreement with our previous assumption that, alternatively, hG52 can stabilize the oxyanion intermediate (List et al., 2012). This residue is located close to the bound glutamine, its NH group points into the glutamine binding pocket, and it is highly flexible in PrFAR-bound WT-ImGPS. In contrast, in the least active ImGPS(fS55AzoF^{Z}), hG52 was highly rigid and pointed away from the glutamine binding site. In addition to the different conformational ensembles of hH178 and the different rigidity of the h49-PGVG-52 motif, the lower HisH activity of ImGPS(fS55AzoF^{Z}) compared with ImGPS(fS55AzoF^{E}) correlates with the closure of the subunit interface, consistent with previous studies in which low activity correlated with a closed complex (Rivalta et al., 2016).

The activity of HisH in ImGPS(fY39NPY) was impaired by a similar mechanism but to a lower extent than in ImGPS(fS55AzoF^{Z}). Again, the hG52 backbone was significantly more rigid than in WT-ImGPS, and the ImGPS interface remained closed correlating with a rigid right half of the ImGPS complex. The UAA closest to the HisH active site, mNPK in fK99mNPK, disrupted the structural integrity of the complete interface including HisH catalytic residues, and especially blocked the breathing motion above the glutamine binding site. It is interesting to note that the binding of the substrate glutamine is not disturbed in ImGPS(fK99mNPK), ImGPS(fY39NPY), or ImGPS(fS55AzoF^{Z}), all of which compromise the opening of the ImGPS interface. Hence, the breathing motion appears not to correlate with the binding of glutamine; however, an open interface seems to be associated with high turnover rates of HisH.

As a conclusion, we found evidence that, in contrast to what has been postulated, glutamine turnover by HisH is not induced by the proper formation of the oxyanion hole. Instead, our MD simulations indicate that the conformational organization of the catalytic residue hH178 is essential for HisH activity. This observation will be the basis for further combined biochemical and structural studies to uncover the changes that are induced in the HisH active site upon binding of PrFAR to HisF and transmittance of the signal through an allosteric network.

Most importantly, we have demonstrated that UAAs can be used to regulate allostery within a sophisticated metabolic enzyme complex such as ImGPS. Our results indicate that caged UAAs are not optimal light switches in such complexes, because the effect of irradiation is irreversible and complete decaging and recovery of WT activity is limited by the protein environment. Photo-switchable proteins containing AzoF, on the other hand, generally hold greater promise for allosteric photo-control of enzymes, especially because the effect of irradiation is reversible. However, also with AzoF as the UAA, the identification of

positions leading to high LRFs is not trivial and might remain a cumbersome trial-and-error process in the future. Nevertheless, we envision that the transfer of the approach presented here for ImGPS, to other enzymatic systems, will allow for interesting synthetic or medicinal applications.

SIGNIFICANCE

The artificial spatiotemporal control of biological macromolecules by light is an exciting and rapidly emerging sub-discipline of synthetic biology. Within this framework, one central research goal is the photo-control of monomeric enzymes by reversible obstruction of the active site. The next level of sophistication is the light regulation of allosteric interactions in enzyme complexes, which might pave the way for the regulation of enzyme cascades in industrial biocatalysis. We have made a first step toward this goal and developed a versatile strategy for the light regulation of allostery, namely the manipulation of signal propagation by photo-sensitive unnatural amino acids (UAAs). To implement our new approach, we used the enzyme complex imidazole glycerol phosphate synthase (ImGPS), which consists of the synthase subunit HisF and the glutaminase subunit HisH. Substrate binding to HisF stimulates the glutaminase activity of HisH over a distance of more than 25 Å. To put this long-range allosteric stimulation under the control of light, we have incorporated the light-responsive UAAs phenylalanine-4′-azobenzene (AzoF), *o*-nitropiperonyl-*O*-tyrosine, and (NPY) methyl-*o*-nitropiperonyllysine (mNPK) at ten strategic positions of HisF. The three most promising candidates for the light-dependent regulation of HisH activity were purified and analyzed by various biochemical and biophysical methods. Kinetic measurements showed that HisH activity was light regulated as much as 10-fold by AzoF-HisF and 4- to 6-fold by NPY-HisF and mNPK-HisF. Crystal structure analysis and MD simulations revealed the different mechanisms by which each UAA affects the allosteric machinery in ImGPS and provided additional insight into HisH activation. Taken together, we present an innovative approach for the light regulation of long-range enzyme allostery by the use of photo-sensitive UAAs. Allostery is a crucial regulatory feature of many central metabolic enzymes. Our work demonstrates a general strategy how this feature can be put under the control of light.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chembiol.2019.08.006.

ACKNOWLEDGMENTS

We thank Sabine Laberer and Jeannette Ueckert for excellent technical assistance and would like to further express our gratitude to Peter Schultz for supplying us with the pEVOL-vectors as well as to Gernot Längst for providing the Nano-DSF equipment. This work was supported by a grant of the Deutsche Forschungsgemeinschaft to R.S. (STE 891/12-1) and a grant from the NIH to V.H.W. (P41 GM128577). N.A.S. thanks the Studienstiftung des Deutschen Volkes for a doctoral scholarship.

AUTHOR CONTRIBUTIONS

A.C.K. and R.S. equally conceptualized the idea. A.C.K. was responsible for the acquisition of all data, delegating experiments where necessary, final analysis and interpretation, as well as writing of the original draft. K.S. performed all computational methods and bioinformatical analyses. P.B. (lead) and E.H. (supporting) delivered data for mNPK proteins. P.B. synthesized mNPK and N.S. synthesized NPY and AzoF. A.B. performed tryptic digest and liquid chromatography-MS analysis. F.B. (lead) and V.H.W. (supporting) performed native MS analysis. C.R. supervised X-ray crystallography, as well as collecting and refining structural data. K.S. (lead) and D.H. (supporting) created rotamer libraries for the unnatural amino acids. R.S., R.M., and B.K. supervised the project and acquired funding. A.C.K., K.S., E.H., R.M., and R.S. collaboratively evaluated the data and, together with P.B., N.A.S., F.B., V.H.W., D.H., and B.K. revised and edited the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 18, 2019
Revised: July 31, 2019
Accepted: August 19, 2019
Published: September 5, 2019

REFERENCES

Abascal, J.L.F., and Vega, C. (2005). A general purpose model for the condensed phases of water: TIP4P/2005. J. Chem. Phys. *123*, 234505.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948–1954.

Amaro, R.E., Sethi, A., Myers, R.S., Davisson, V.J., and Luthey-Schulten, Z.A. (2007). A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase. Biochemistry *46*, 2156–2173.

Baker, A.S., and Deiters, A. (2014). Optical control of protein function through unnatural amino acid mutagenesis and other optogenetic approaches. ACS Chem. Biol. *9*, 1398–1407.

Bardhan, A., and Deiters, A. (2019). Development of photolabile protecting groups and their application to the optochemical control of cell signaling. Curr. Opin. Struct. Biol. *57*, 164–175.

Beismann-Driemeyer, S., and Sterner, R. (2001). Imidazole glycerol phosphate synthase from *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex. J. Biol. Chem. *276*, 20387–20396.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Berendsen, H.J.C., van der Spoel, D., and van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. Comput. Phys. Commun. *91*, 43–56.

Berroy, P., Viriot, M.L., and Carré, M.C. (2001). Photolabile group for 5′-OH protection of nucleosides: synthesis and photodeprotection rate. Sens. Actuators B *74*, 186–189.

Bose, M., Groff, D., Xie, J., Brustad, E., and Schultz, P.G. (2006). The incorporation of a photoisomerizable amino acid into proteins in *E. coli*. J. Am. Chem. Soc. *128*, 388–389.

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. J. Chem. Phys. *126*, 14101.

Calbo, J., Weston, C.E., White, A.J.P., Rzepa, H.S., Contreras-García, J., and Fuchter, M.J. (2017). Tuning azoheteroarene photoswitch performance through heteroaryl design. J. Am. Chem. Soc. *139*, 1261–1274.

Chaudhuri, B.N., Lange, S.C., Myers, R.S., Chittur, S.V., Davisson, V.J., and Smith, J.L. (2001). Crystal structure of imidazole glycerol phosphate synthase: a tunnel through a $(\beta/\alpha)_8$ barrel joins two active sites. Structure *9*, 987–997.

Chaudhuri, B.N., Lange, S.C., Myers, R.S., Davisson, V.J., and Smith, J.L. (2003). Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerolphosphate synthase: crystal structures of a ternary complex and the free enzyme. Biochemistry *42*, 7003–7012.

Clark, J.J., Benson, M.L., Smith, R.D., and Carlson, H.A. (2019). Inherent versus induced protein flexibility: comparisons within and between apo and holo structures. PLoS Comput. Biol. *15*, e1006705.

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

Curley, K., and Lawrence, D.S. (1999). Light-activated proteins. Curr. Opin. Chem. Biol. *3*, 84–88.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S., et al. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. *35*, W375–W383.

Davisson, V.J., Deras, I.L., Hamilton, S.E., and Moore, L.L. (1994). A plasmid-based approach for the synthesis of a histidine biosynthetic intermediate. J. Org. Chem. *59*, 137–143.

Deiters, A., Groff, D., Ryu, Y., Xie, J., and Schultz, P.G. (2006). A genetically encoded photocaged tyrosine. Angew. Chem. Int. Ed. *45*, 2728–2731.

Douangamath, A., Walker, M., Beismann-Driemeyer, S., Vega-Fernandez, M.C., Sterner, R., and Wilmanns, M. (2002). Structural evidence for ammonia tunneling across the $(\beta\alpha)_8$ barrel of the imidazole glycerol phosphate synthase bienzyme complex. Structure *10*, 185–193.

Dunbrack, R.L., and Karplus, M. (1993). Backbone-dependent rotamer library for proteins. Application to side-chain prediction. J. Mol. Biol. *230*, 543–574.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. PLoS One *3*, e3647.

Gautier, A., Nguyen, D.P., Lusic, H., An, W., Deiters, A., and Chin, J.W. (2010). Genetically encoded photocontrol of protein localization in mammalian cells. J. Am. Chem. Soc. *132*, 4086–4088.

Gfeller, D., Michielin, O., and Zoete, V. (2013). SwissSidechain: a molecular and structural database of non-natural sidechains. Nucleic Acids Res. *41*, D327–D332.

Green, M.J., and Hill, H.A. (1984). Chemistry of dioxygen. Methods Enzymol. *105*, 3–22.

Hüll, K., Morstein, J., and Trauner, D. (2018). In vivo photopharmacology. Chem. Rev. *118*, 10710–10747.

Il'ichev, Y.V., and Wirz, J. (2000). Rearrangements of 2-nitrobenzyl compounds. 1. Potential energy surface of 2-nitrotoluene and its isomers explored with ab initio and density functional theory methods. J. Phys. Chem. A *104*, 7856–7870.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. *26*, 795–800.

Kastritis, P.L., and Gavin, A.-C. (2018). Enzymatic complexes across scales. Essays Biochem. *62*, 501–514.

Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. Biochim. Biophys. Acta *1751*, 119–139.

Klán, P., Šolomek, T., Bochet, C.G., Blanc, A., Givens, R., Rubina, M., Popik, V., Kostikov, A., and Wirz, J. (2013). Photoremovable protecting groups in chemistry and biology: reaction mechanisms and efficacy. Chem. Rev. *113*, 119–191.

Klem, T.J., and Davisson, V.J. (1993). Imidazole glycerol phosphate synthase: the glutamine amidotransferase in histidine biosynthesis. Biochemistry *32*, 5177–5186.

Kneuttinger, A.C., Winter, M., Simeth, N.A., Heyn, K., Merkl, R., König, B., and Sterner, R. (2018). Artificial light regulation of an allosteric bienzyme complex by a photosensitive ligand. ChemBioChem *19*, 1750–1757.

Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A., and Vriend, G. (2004). Making optimal use of empirical energy functions: force-field parameterization in crystal space. Proteins *57*, 678–683.

Lachmann, D., Lahmy, R., and König, B. (2019). Fulgimides as light-activated tools in biological investigations. Eur. J. Org. Chem. https://doi.org/10.1002/ejoc.201900219.

Lemke, E.A., Summerer, D., Geierstanger, B.H., Brittain, S.M., and Schultz, P.G. (2007). Control of protein phosphorylation with a genetically encoded photocaged amino acid. Nat. Chem. Biol. *3*, 769–772.

Lipchock, J.M., and Loria, J.P. (2010). Nanometer propagation of millisecond motions in V-type allostery. Structure *18*, 1596–1607.

Lisi, G.P., East, K.W., Batista, V.S., and Loria, J.P. (2017). Altering the allosteric pathway in IGPS suppresses millisecond motions and catalytic activity. Proc. Natl. Acad. Sci. U S A *114*, E3414–E3423.

Lisi, G.P., Manley, G.A., Hendrickson, H., Rivalta, I., Batista, V.S., and Loria, J.P. (2016). Dissecting dynamic allosteric pathways using chemically related small-molecule activators. Structure *24*, 1155–1166.

List, F., Vega, M.C., Razeto, A., Häger, M.C., Sterner, R., and Wilmanns, M. (2012). Catalysis uncoupling in a glutamine amidotransferase bienzyme by unblocking the glutaminase active site. Chem. Biol. *19*, 1589–1599.

Liu, C.C., and Schultz, P.G. (2010). Adding new chemistries to the genetic code. Annu. Rev. Biochem. *79*, 413–444.

Losi, A., Gardner, K.H., and Möglich, A. (2018). Blue-light receptors for optogenetics. Chem. Rev. *118*, 10659–10709.

Lovell, S.C., Davis, I.W., Arendall, W.B., Bakker, P.I.W., de Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: Φ, Ψ and Cβ deviation. Proteins *50*, 437–450.

Luo, J., Samanta, S., Convertino, M., Dokholyan, N.V., and Deiters, A. (2018). Reversible and tunable photoswitching of protein function through genetic encoding of azobenzene amino acids in mammalian cells. ChemBioChem *19*, 2178–2185.

Luo, J., Torres-Kolbus, J., Liu, J., and Deiters, A. (2017). Genetic encoding of photocaged tyrosines with improved light-activation properties for the optical control of protease function. ChemBioChem *18*, 1442–1447.

Makhlynets, O.V., Raymond, E.A., and Korendovych, I.V. (2015). Design of allosterically regulated protein catalysts. Biochemistry *54*, 1444–1456.

Marty, M.T., Baldwin, A.J., Marklund, E.G., Hochberg, G.K.A., Benesch, J.L.P., and Robinson, C.V. (2015). Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Anal. Chem. *87*, 4370–4376.

Massière, F., and Badet-Denisot, M.A. (1998). The mechanism of glutamine-dependent amidotransferases. Cell. Mol. Life Sci. *54*, 205–222.

Morstein, J., Hill, R.Z., Novak, A.J.E., Feng, S., Norman, D.D., Donthamsetti, P.C., Frank, J.A., Harayama, T., Williams, B.M., Parrill, A.L., et al. (2019). Optical control of sphingosine-1-phosphate formation and function. Nat. Chem. Biol. *15*, 623–631.

Muranaka, N., Hohsaka, T., and Sisido, M. (2002). Photoswitching of peroxidase activity by position-specific incorporation of a photoisomerizable non-natural amino acid into horseradish peroxidase. FEBS Lett. *510*, 10–12.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Myers, R.S., Amaro, R.E., Luthey-Schulten, Z.A., and Davisson, V.J. (2005). Reaction coupling through interdomain contacts in imidazole glycerol phosphate synthase. Biochemistry *44*, 11974–11985.

Negre, C.F.A., Morzan, U.N., Hendrickson, H.P., Pal, R., Lisi, G.P., Loria, J.P., Rivalta, I., Ho, J., and Batista, V.S. (2018). Eigenvector centrality for characterization of protein allosteric pathways. Proc. Natl. Acad. Sci. U S A *115*, E12201–E12208.

Oliphant, T.E. (2007). Python for scientific computing. Comput. Sci. Eng. *9*, 10–20.

Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G.N., Cohen, S., Perrakis, A., and Noble, M. (2004). Developments in the CCP4 molecular-graphics project. Acta Crystallogr. D Biol. Crystallogr. *60*, 2288–2294.

Raushel, F.M., Thoden, J.B., and Holden, H.M. (1999). The amidotransferase family of enzymes: molecular machines for the production and delivery of ammonia. Biochemistry *38*, 7891–7899.

Reynolds, K.A., McLaughlin, R.N., and Ranganathan, R. (2011). Hot spots for allosteric regulation on protein surfaces. Cell *147*, 1564–1575.

Rivalta, I., Lisi, G.P., Snoeberger, N.-S., Manley, G.A., Loria, J.P., and Batista, V.S. (2016). Allosteric communication disrupted by small molecule binding to the imidazole glycerol phosphate synthase protein-protein interface. Biochemistry *55*, 6484–6494.

Rivalta, I., Sultan, M.M., Lee, N.-S., Manley, G.A., Loria, J.P., and Batista, V.S. (2012). Allosteric pathways in imidazole glycerol phosphate synthase. Proc. Natl. Acad. Sci. U S A *109*, E1428–E1436.

Rohweder, B., Semmelmann, F., Endres, C., and Sterner, R. (2018). Standardized cloning vectors for protein production and generation of large gene libraries in *Escherichia coli*. BioTechniques *64*, 24–26.

Schlesinger, O., Dandela, R., Bhagat, A., Adepu, R., Meijler, M.M., Xia, L., and Alfonta, L. (2018). Photo-switchable microbial fuel-cells. Biotechnol. Bioeng. *115*, 1355–1360.

Schmermund, L., Jurkaš, V., Özgen, F.F., Barone, G.D., Büchsenschütz, H.C., Winkler, C.K., Schmidt, S., Kourist, R., and Kroutil, W. (2019). Photo-biocatalysis: biotransformations in the presence of light. ACS Catal. *9*, 4115–4144.

Schmidt-Dannert, C., and Lopez-Gallego, F. (2016). A roadmap for biocatalysis – functional and spatial orchestration of enzyme cascades. Microb. Biotechnol. *9*, 601–609.

Schrödinger. (2015). The PyMOL Molecular Graphics System (LLC).

Skjærven, L., Yao, X.-Q., Scarabelli, G., and Grant, B.J. (2014). Integrating protein structural dynamics and evolutionary analysis with Bio3D. BMC Bioinformatics *15*, 399.

Smith, D.W., and Ames, B.N. (1964). Intermediates in the early steps of histidine biosynthesis. J. Biol. Chem. *239*, 1848–1855.

Sousa da Silva, A.W., and Vranken, W.F. (2012). ACPYPE – AnteChamber PYthon Parser interfacE. BMC Res. Notes *5*, 367.

Szymański, W., Beierle, J.M., Kistemaker, H.A.V., Velema, W.A., and Feringa, B.L. (2013). Reversible photocontrol of biological systems by the incorporation of molecular photoswitches. Chem. Rev. *113*, 6114–6178.

Tietze, L.F., Müller, M., Duefert, S.-C., Schmuck, K., and Schuberth, I. (2013). Photoactivatable prodrugs of highly potent duocarmycin analogues for a selective cancer therapy. Chem. Eur. J. *19*, 1726–1731.

Uprety, R., Luo, J., Liu, J., Naro, Y., Samanta, S., and Deiters, A. (2014). Genetic encoding of caged cysteine and caged homocysteine in bacterial and mammalian cells. ChemBioChem *15*, 1793–1799.

Valiauga, B., Williams, E.M., Ackerley, D.F., and Čėnas, N. (2017). Reduction of quinones and nitroaromatic compounds by *Escherichia coli* nitroreductase A (NfsA): characterization of kinetics and substrate specificity. Arch. Biochem. Biophys. *614*, 14–22.

VanAernum, Z.L., Gilbert, J.D., Belov, M.E., Makarov, A.A., Horning, S.R., and Wysocki, V.H. (2019). Surface-induced dissociation of noncovalent protein complexes in an extended mass range orbitrap mass spectrometer. Anal. Chem. *91*, 3611–3618.

VanWart, A.T., Eargle, J., Luthey-Schulten, Z., and Amaro, R.E. (2012). Exploring residue component contributions to dynamical network models of allostery. J. Chem. Theor. Comput. *8*, 2949–2961.

Waitt, G.M., Xu, R., Wisely, G.B., and Williams, J.D. (2008). Automated in-line gel filtration for native state mass spectrometry. J. Am. Soc. Mass Spectrom. *19*, 239–245.

Wang, J., Liu, Y., Liu, Y., Zheng, S., Wang, X., Zhao, J., Yang, F., Zhang, G., Wang, C., and Chen, P.R. (2019). Time-resolved protein activation by proximal decaging in living systems. Nature *569*, 509–513.

Wang, J., Wang, W., Kollman, P.A., and Case, D.A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. *25*, 247–260.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. J. Comput. Chem. *25*, 1157–1174.

Xie, X., Li, X.-M., Qin, F., Lin, J., Zhang, G., Zhao, J., Bao, X., Zhu, R., Song, H., Li, X.D., et al. (2017). Genetically encoded photoaffinity histone marks. J. Am. Chem. Soc. *139*, 6522–6525.

Young, T.S., Ahmad, I., Yin, J.A., and Schultz, P.G. (2010). An enhanced system for unnatural amino acid mutagenesis in *E. coli*. J. Mol. Biol. *395*, 361–374.

Zalkin, H., and Smith, J.L. (1998). Enzymes utilizing glutamine as an amide donor. Adv. Enzymol. Relat. Areas Mol. Biol. *72*, 87–144.

Zimmerman, G., Chow, L.-Y., and Paik, U.-J. (1958). The photochemical isomerization of azobenzene. J. Am. Chem. Soc. *80*, 3528–3531.

Zuman, P., and Shah, B. (1994). Addition, reduction, and oxidation reactions of nitrosobenzene. Chem. Rev. *94*, 1621–1641.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
AzoF	This paper	N/A
NBY	Accela ChemBio	Cat#SY008208
NPY	This paper	N/A
mNPK	This paper	N/A
ProFAR	This paper	N/A
HisA	This paper	N/A
HisG/IE	This paper	N/A
wt/UAA-HisF	This paper	N/A
HisH	This paper	N/A
Glutamate dehydrogenase (GDH)	Sigma-Aldrich	CAT#10197734001
NAD^+, free acid	Sigma-Aldrich	CAT#124542
Glutamine	Sigma-Aldrich	CAT#G8540
Glutamate-oxidase (GOX)	Sigma-Aldrich	CAT#G1924
Hydrogen peroxidase (HRP)	Sigma-Aldrich	CAT#P8125
Phenol	Merck	CAT#100206
2-aminoantipyrine	Sigma-Aldrich	CAT#06800
Deposited Data		
ImGPS structure from *Thermotoga maritima*	Douangamath et al., 2002	PDB-ID 1gpw
ImGPS structure from *Saccharomyces cerevisiae*	Chaudhuri et al., 2003	PDB-ID 1ox5
ImGPS(fS55AzoF) structure	This paper	PDB-ID 6ru0
ImGPS(fY39NPY) structure	This paper	PDB-ID 6rtz
Oligonucleotides		
Primers: see Table S1	This paper	N/A
Recombinant DNA		
pET28aTEV	This paper	N/A
pET28a_HisFwt	This paper	N/A
pET28aTEV_BsaI	This paper	N/A
pET28a_HisHwt	This paper	N/A
pEVOL_mNPK	This paper	N/A
pET28a vectors containing stop codon mutations: see Table S1	This paper	N/A
pEVOL_AzoF	Peter Schultz (Scripps Research Institute, La Jolla, USA) (Young et al., 2010; Bose et al., 2006)	N/A
pEVOL_NBY	Peter Schultz (Scripps Research Institute, La Jolla, USA) (Young et al., 2010; Deiters et al., 2006)	N/A
Software and Algorithms		
Gromacs version 5.1.2	Berendsen et al., 1995	https://www.gromacs.org
Acpype	Sousa da Silva and Vranken, 2012	https://www.pypi.org/project/acpype/
AmberTools14	Wang et al., 2006	https://www.ambermd.org/AmberTools.php
Origin 2018	OriginLab	https://www.originlab.com
Data Analysis 4.2	Bruker Daltonics	https://www.bruker.com

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Yasara version 17.4.17	Krieger et al., 2004	https://www.yasara.org
Pymol 2.0	Schrödinger, 2015	https://www.pymol.org/2/
plotly	Plotly	https://plot.ly
Python 2.7 with packages numpy, scipy, yasara, plotly	Python	https://www.python.org
R version 3.6.0	The R Foundation	https://www.r-project.org
UniDec version 2.6.5.	Marty et al., 2015	http://unidec.chem.ox.ac.uk/
MolProbity	Davis et al., 2007	http://molprobity.biochem.duke.edu/
Other		
"Conventional UV lamp": two F8 T5 BLB black light bulbs (8 W)	Sylvania	EAN#5410288000244
High Power LED, UV Z5 series, 420 nm	Seoul Viosys	CAT#CUN26A1B
High Power LED, LZ4 series, 365 nm	LED Engin	CAT#LZ4-04UV00-0000

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Reinhard Sterner (reinhard.sterner@ur.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

E. coli BL21 Gold (DE3) cells were originally purchased from Agilent Technologies and *E. coli* NEB Turbo from New England Biolabs. The strains were further maintained following the manufacturer's guidelines for the preparation of chemically competent cells. Cells were grown in lysogenic broth (LB) medium at 37°C, and stored in 80% glycerol at −80°C.

METHOD DETAILS

Chemistry, Materials and Instrumentation

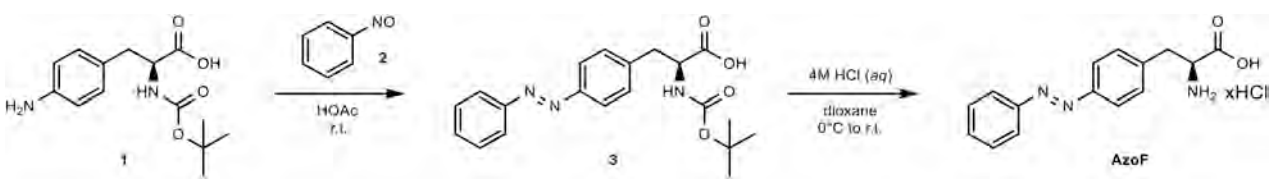
All reagents and solvents other than **AzoF**, **NPY**, **mNPK** and ProFAR were purchased in analytical grade or higher from commercial sources and were used without further purification, if not otherwise stated. **AzoF**, **NPY**, **mNPK** and ProFAR were synthesized and purity confirmed as described below. **NBY** was purchased from Accela ChemBio (96% pure).

All reactions were performed under magnetic stirring at room temperature (r.t.) unless otherwise specified. Light-sensitive compounds were kept in the dark using aluminum foil to cover reaction vessels or were worked-up under red light (LEDs > 650 nm). For TLC, silica coated aluminum plates (Macherey-Nagel GmbH & Co. KG 60 Å silica gel, 0.25 mm) were utilized. Visualization was carried out with a UV lamp at 254 or 366 nm or through suitable staining. For MPLC, a Biotage Isolera One Flash Purification System with manually packed columns using Macherey-Nagel GmbH & Co. KG 60M (0.04–0.063 mm, 230–400 grain diameter).

Nuclear magnetic resonance spectroscopy (NMR) was carried out using a Bruker Avance 300 MHz (^{1}H: 300 MHz, ^{13}C: 75 MHz, T = 300 K) or Bruker Avance 400 MHz spectrometer (^{1}H: 400 MHz, ^{13}C: 101 MHz, T = 300 K). Chemical shifts are reported in δ [ppm] relative to an internal standard (solvent residual peak). The solvents used are indicated for each spectrum. Coupling constants are reported in Hertz [Hz]. Characterization of the signals: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, bs = broad singlet, dd = doublet of doublet, dt = doublet of triplet. Integration is directly proportional to the number of the protons.

Synthesis of AzoF

AzoF was synthesized following a protocol developed by (Bose et al., 2006). The identity and purity of products was determined through ^{1}H-NMR.



(S,E)-2-((tert-butoxycarbonyl)amino)-3-(4-(phenyldiazenyl)phenyl)propanoic acid (3)

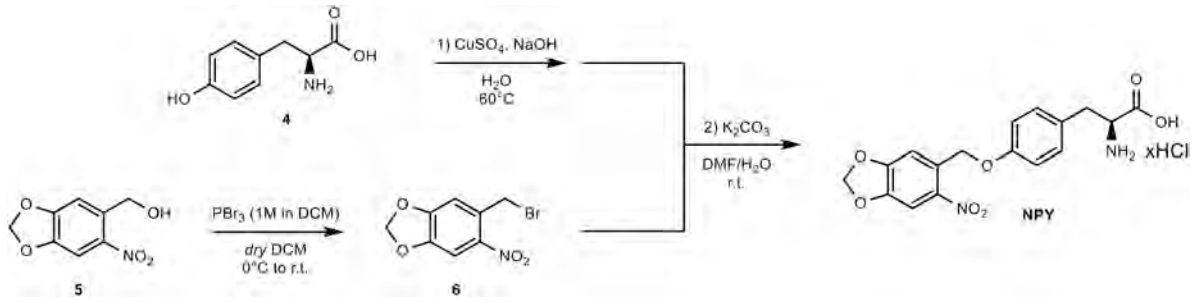
Compound **3** was synthesized adapting a previously reported procedure (Bose et al., 2006). Nitrosobenzene (1.91 g, 17.84 mmol) was added to butoxycarbonyl-*p*-aminophenylalanine (5.00 g, 17.84 mmol) dissolved in glacial acetic acid (1.0 L). The flask was covered with aluminum foil and the reaction mixture was stirred at ambient temperature for 18 h. The reaction mixture was poured onto crushed ice/water (~ 1.5 L) and extracted with EtOAc (2 x 300 mL). The combined organic phases were washed with water (2 x 200 mL) and brine (1 x 200 mL) and dried over Na_2SO_4. The volatiles were removed *in vacuo* and the crude product was purified by MPLC (10% MeOH in DCM). Compound **3** was isolated as orange solid (4.09 g, 11.06 mmol, 62% yield). **^{1}H-NMR** (300 MHz, Chloroform-*d*): δ = 7.95 – 7.83 (m, 4H), 7.59 – 7.45 (m, 3H), 7.34 (d, J = 8.1 Hz, 2H), 4.75 – 4.57 (m, 1H), 3.35 – 3.08 (m, 2H), 1.42 (s, 9H).

(S,E)-2-Amino-3-(4-(phenyldiazenyl)phenyl)propanoic acid (AzoF)

AzoF was synthesized following the previously reported protocol (Bose et al., 2006) employing compound **3** (4.09 g, 11.06 mmol). **AzoF**×HCl was dried via lyophilization to yield an orange solid (3.38 g, 11.06 mmol, quantitative; purity > 98% by NMR). **^{1}H-NMR** (300 MHz, DMSO-d6): δ = 7.97 – 7.83 (m, 4H), 7.72 – 7.55 (m, 3H), 7.54 – 7.43 (m, 2H), 4.17 (t, J = 6.4 Hz, 1H), 3.21 (t, J = 6.8 Hz, 2H).

Synthesis of NPY

NPY was synthesized following a protocol developed by (Luo et al., 2017). The identity and purity of products was determined through ^{1}H-NMR.



5-(Bromomethyl)-6-nitrobenzo[d][1,3]dioxole (6)

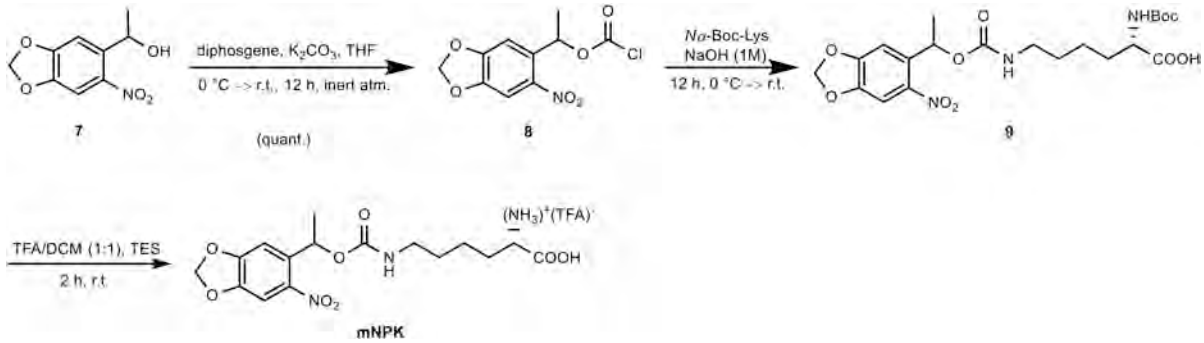
Compound **6** was synthesized following the previously reported protocol (Luo et al., 2017) employing **5** (1.00 g, 5.07 mmol). After purification by MPLC (20% EtOAc in petroleum ether), the product was obtained as yellow powder (685 mg, 2.64 mmol, 52% yield). **^{1}H NMR** (400 MHz, Chloroform-d): δ = 7.57 (s, 1H), 6.95 (s, 1H), 6.14 (s, 2H), 4.79 (s, 2H) (Tietze et al., 2013).

(2S)-2-Amino-3-(4-(1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)phenyl)propanoic acid x HCl (NPY)

NPY was synthesized following the previously reported protocol (Luo et al., 2017) employing L-tyrosine (490 mg, 2.70 mmol) and compound **6** (537 mg, 2.08 mmol). **NPY**×HCl was isolated as yellow powder (361 mg, 0.87 mmol, 42% yield; purity ~ 85% by NMR). **^{1}H-NMR** (400 MHz, DMSO-d6): δ = 7.74 (s, 1H), 7.25 (s, 1H), 7.19 (d, J = 8.3 Hz, 2H), 6.96 (d, J = 8.3 Hz, 2H), 6.24 (s, 2H), 3.92–3.87 (m, 1H), 3.10–2.92 (m, 2H).

Synthesis of mNPK

mNPK synthesis was adopted from a protocol developed by (Gautier et al., 2010). The identity and purity of products was determined through ^{1}H-NMR.



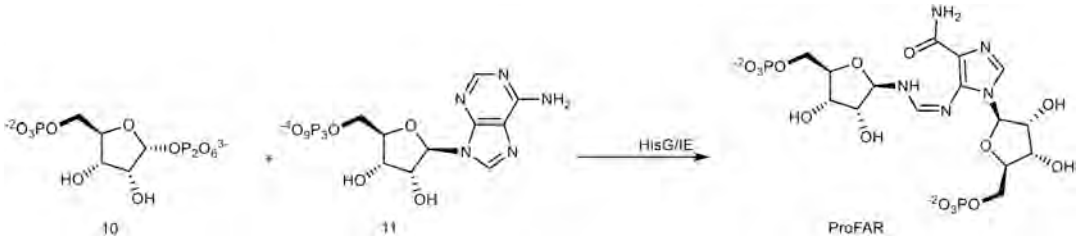
N^2-(tert-butoxycarbonyl)-N^6-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)-L-lysine (9)

1-(6-Nitro-1,3-benzodioxol-5-yl)ethanol **7** (1.00 g, 4.7 mmol) was dissolved in dry THF (10 mL) and supplemented with K_2CO_3 (0.66 g, 4.7 mmol) at 0°C under nitrogen atmosphere. Diphosgene (0.6 mL, 4.7 mmol) was added dropwise and stirring continued for 12 h allowing the reaction mixture to warm to r.t. After filtration the quantitatively converted chloroformate **8** (1.30 g, 4.7 mmol) was dissolved in THF and added dropwise to a solution of N_α-Boc-lysine (1.28 g, 5.2 mmol) in NaOH (13 mL, 1 M, aq) under stirring at 0°C. Stirring was continued for 12 h allowing the mixture to warm to r.t. The deep red aqueous phase was washed with Et_2O (10 ml), cooled to 0°C and acidified with ice-cold HCl (1 M, aq) to pH 1 producing a bright yellow suspension. The crude product was extracted with EtOAc (3 x 30 mL). The organic phases were combined, dried over Mg_2SO_4 and filtered. The volatiles were evaporated to afford **9** (1.94 g, 4.1 mmol, 85% yield) as a yellow foam. **^{1}H-NMR** (400 MHz, $CDCl_3$): δ, [ppm] = 7.47 (s, 1H), 7.00 (s, 1H), 6.28-6.21 (s, 1H.), 6.20-6.10 (m, 2H), 5.04-5.24 (m, 1H), 4.28 (bs, 1H), 3.12 (bs, 2H), 1.28-1.80 (m, 18H).

N^6-((1-(6-nitrobenzo[d][1,3]dioxol-5-yl)ethoxy)carbonyl)-L-lysine TFA salt (mNPK)

The Boc-protected lysine **9** (500 mg, 1.0 mmol) was dissolved in DCM (6 mL), supplemented with triethylsilane (0.3 mL, 2.1 mmol) and the mixture was added dropwise to a 0°C solution of TFA (6 mL) under nitrogen atmosphere. After the reaction mixture was stirred for 2 h at r.t., the volatiles were evaporated under a constant N_2-stream without heating. The residue was dissolved in MeOH (2 mL) and added dropwise to ice-cold Et_2O (200 mL) under vigorous stirring. The supernatant was decanted, the precipitate was washed with ice-cold Et_2O (100 mL) and decanted again. The volatiles were evaporated to afford the caged lysine TFA salt **mNPK** (360 mg, 0.8 mmol, 73 % yield; purity ~ 85% by NMR) as a light-yellow, hygroscopic solid. **^{1}H-NMR** (400 MHz, D_2O): δ [ppm] = 7.38 (s, 1H), 6.98 (s, 1H), 6.00 (s, 2H), 5.97-5.92 (m, 1H), 3.55-3.52 (m, 1H), 2.91 (m, 2H), 1.85-1.73 (m, 2H), 1.42-1.24 (m, 7H).

Biochemical Synthesis of ProFAR



ProFAR was synthesized following the standard protocol developed by (Davisson et al., 1994) from 5-phospho-D-ribosyl α-1-pyrophosphate **10** and adenosine triphosphate **11** in 50 mM NH_4COOCH_3 pH 7.8. Reaction progress was tracked spectrophotometrically as described (Davisson et al., 1994). Products were purified with ion-exchange chromatography using a POROS column (HQ 20, 10 mL, Applied Biosystems) and a linear gradient of NH_4COOCH_3 (50 mM → 1 M). ProFAR concentration was determined at 300 nm (ϵ_{300} = 6069 $M^{-1}cm^{-1}$) (Klem and Davisson, 1993) and ProFAR purity through the absorbance ratio A_{290}/A_{260} for each fraction. Highly concentrated and > 90% pure (A_{290}/A_{260} = 1.1–1.2 accounts for > 95% purity) (Smith and Ames, 1964) fractions were lyophilized and stored at −80°C. ProFAR identity was confirmed by total turnover measurements to ImGP and AICAR, with HisA and HisF, using the HisF assay described in the methods section (in ammonia saturation). Concentrations determined by total turnover were compared to spectrophotometrically measured concentrations. For steady-state kinetics, ProFAR-concentrations were always determined by total turnover.

Auxiliary Enzymes

HisA from *T. maritima* was produced by heterologous gene expression in *E. coli* BL21 Gold (DE3) (Agilent Technologies) from pET21_HisA (List et al., 2012). The recombinant protein was purified from the soluble fraction of the crude extract by heat precipitation (15 min at 73°C) of the host proteins, followed by nickel-affinity chromatography. To this end, the extract was loaded onto a column (HisTrap FF crude 5 ml; GE Healthcare) that had been equilibrated with 50 mM potassium phosphate, 300 mM NaCl, 1 mM imidazole (pH 7.5). The column was washed with the equilibration buffer, and the bound protein was eluted by applying a linear gradient of 1–300 mM imidazole. Fractions with pure protein were pooled and dialyzed extensively against 50 mM potassium phosphate (pH 7.5). Based on SDS-PAGE analysis, the purity of all samples was at least 90%. The proteins were dripped into liquid nitrogen and stored at −80°C

HisG/IE was produced by heterologous gene expression in *E. coli* BL21 Gold (DE3) (Agilent Technologies) from p_hisGIE_tac (Davisson et al., 1994). The strains containing the respective expression vectors were grown in lysogeny broth medium (2 L) at 37°C to an OD_{600} of 0.7. Protein expression was induced by isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM) and incubated overnight. Bacterial pellets were harvested and resuspended in 10 mM KP pH 7.5, 2.5 mM EDTA, and 1 mM DTT. Sonication and repeated centrifugation yielded the protein-containing supernatant, which was subjected to ion-exchange chromatography with a MonoQ column (HR 16/10, 20 mL, Pharmacia). Proteins were eluted with a linear gradient of KP (10 → 500 mM), and fractions containing either

HisG or HisIE or both were pooled and dialyzed against 50 mM KP pH 7.5, 2.5 mM EDTA, and 1 mM DTT. Proteins were identified based on their molecular weight using SDS-PAGE analysis and simultaneously checked for > 90% purity. Concentrated proteins were dripped into liquid nitrogen and stored at −80°C.

Subcloning of the *his*F and *his*H Genes

The *hisF* and *hisH* genes were cloned into vectors encoding an N-terminal His_6-tag. For this purpose, the thrombin cleavage site of pET28a(+) (Novagen) was substituted for a TEV-cleavage site. For *hisF* incorporation, the vector was amplified using primers flanking the region encoding the thrombin cleavage site (P1 and P2, Table S1) using PWO Polymerase (Sigma Aldrich). Both primers contained a part of the TEV cleavage site (bold), creating an 18 bp overlap facilitating circularization. pET28a(+) was amplified with each primer separately and, after digestion with DpnI (NEB), a mixture of both reactions was transformed into *E. coli* NEB Turbo (NEB) to yield the target plasmid pET28aTEV independent of ligation. The correct exchange was confirmed by Sanger Sequencing (Microsynth Seqlab) starting from the T7 terminator.

The *hisF* gene was amplified from pET11c_HisF (Beismann-Driemeyer and Sterner, 2001) with primers P3 and P4 (Table S1) using PWO polymerase (Sigma Aldrich). The primers carried an overhang containing an *Nde*I and a *Hind*III restriction site, respectively. After digestion of both pET28aTEV and the PCR product with the appropriate enzymes (HF-enzymes, NEB), both fragments were purified via agarose gel electrophoresis by means of a Cleanup Kit (Thermo Scientific). Afterwards, the fragments were mixed and ligated with T4 Ligase (NEB). The correct sequence of the *hisF* gene was confirmed by Sanger Sequencing (Microsynth Seqlab) starting from the T7 terminator.

For *hisH* incorporation, pET28a_BsaI (Rohweder et al., 2018) designed for golden-gate cloning, was modified to encode for a TEV-cleavage site C-terminal of the N-terminal His_6-tag. For this purpose, the plasmid was amplified in a standard PCR reaction using primers P5 and P6 (Table S1). P5 contained the TEV cleavage site (bold).The purified (agarose gel Cleanup Kit, Thermo Scientific), linear PCR fragment was re-cyclized in a ligation reaction with addition of T4 PNK (NEB), yielding pET28aTEV_BsaI.

The gene encoding *hisH* from *T. maritima* was optimized for *E. coli* codon usage and purchased as a linear gene string (GeneArt, Thermo Scientific) carrying two terminal *Bsa*I cleavage sites for golden-gate cloning. *hisH* was cloned into pET28aTEV_BsaI as described previously (Rohweder et al., 2018). Because the N-terminus of HisH is shielded and therefore inaccessible for proteolytic cleavage, an additional two amino acid linker (MG) was inserted between the TEV cleavage site and the *hisH* gene. To this end, the plasmid carrying the gene was amplified in a standard PCR reaction using primers P7 and P8 (Table S1). After purification, the PCR product was cyclized in a ligation reaction with addition of T4 PNK (NEB). The correct sequence of both pET28aTEV_BsaI and pET28a_HisHwt was confirmed by Sanger Sequencing (Microsynth Seqlab) starting from the T7 terminator.

Subcloning of the aaRS/$tRNA_{CUA}$ Pairs

Plasmids for incorporation of **AzoF** and **NBY/NPY** were provided by Prof. Peter Schultz (Scripps Research Institute, La Jolla, USA). The pEVOL-vector system was especially designed for the incorporation of UAAs with two copies of the respective aaRS (Young et al., 2010). pEVOL_NBY carries the *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase / $tRNA_{CUA}$ (*Mj*TyrRS-TyrT) pair evolved for **NBY** incorporation with the mutations Y32G, L65G, F108E, D158S, and L162E (Deiters et al., 2006). The same construct was used for **NPY** incorporation as previously confirmed (Luo et al., 2017). pEVOL_AzoF carries the *Mj*TyrRS-TyrT pair evolved for **AzoF** incorporation with the mutations Y32G, L65E, F108A, Q109E, D158G, and L162H (Bose et al., 2006). Both vectors were checked by Sanger Sequencing (Microsynth Seqlab) starting from the *araBAD* promoter.

pEVOL_AzoF was further used as a template for the design of pEVOL_mNPK. The genes encoding the tyrosyl-tRNA synthetase and its respective $tRNA_{CUA}$ were substituted for the pyrollysyl-tRNA synthetase and its respective $tRNA_{CUA}$ from *Methanosarcina barkeri* (*Mb*PylRS-PylT), which was evolved for caged-lysine derivatives with the mutations M241F, A267S, Y271C, and L274M (Gautier et al., 2010). Initially, *tyrT*-tRNA was substituted by PCR-amplification of pEVOL_AzoF with primers P9 and P10 (Table S1) containing *pylT*-tRNA (bold) and overlaps to the sequences flanking the original *tyrT*-tRNA. The PCR product was cyclized by blunt end ligation using T4 polynucleotide kinase (NEB) and T4 Ligase (NEB) according to the manufacturer's instructions to obtain pEVOL_AzoF'_pylT.

Exchange of the first copy of *MjtyrRS* was prepared by amplifying pEVOL_AzoF'_pylT with primers P11 and P12 (Table S1) containing overhangs with a *BsaI* restriction site and flanking the *MjtyrRS* gene. *MbpylRS*, synthesized (GeneArt) with *BsaI* restriction sites complementary to those in the PCR product, was then joined with the linearized vector in a golden gate cloning procedure (Engler et al., 2008) to build pEVOL_mNPK'. The deletion 3' of the open reading frame (ORF), resulting from deleting the *BsaI* restriction site, does not include the transcription terminator and thus leaves the regulation of transcription intact. Analogous to the first copy, the second *MjtyrRS* gene was substituted with *MbpylRS* by the same strategy using primers P13 and P14 (Table S1) leading to no further deletions and resulting in the final pEVOL_mNPK vector. After each cloning step, the correct modification of the target site was confirmed by Sanger Sequencing (Microsynth Seqlab) with the sequencing primers (Table S1) P15 for pEVOL_AzoF'_pylT, P16 for pEVOL_mNPK', and P17 for pEVOL_mNPK.

Site-Directed Mutagenesis of *his*F

Stop codon point mutations (TAG, underlined) were introduced into pET28a_HisFwt according to the protocol of the Phusion™ site-directed mutagenesis kit from Finnzymes with 5' phosphorylated (PHO) and HPLC-purified primers (Metabion) P18–P37 (Table S1). Correct mutagenesis was checked by Sanger Sequencing (Microsynth Seqlab) starting from the T7 terminator.

Building a Rotamer Library

In order to deduce a backbone-dependent rotamer library (Dunbrack and Karplus, 1993), MD simulations of a tri-peptide (Ala-UAA-Ala) were analyzed to sample the conformational space of the respective side chains. MD simulations were performed using Gromacs (version 5.1.2) (Berendsen et al., 1995) with GAFF (Wang et al., 2004) topologies generated by means of Acpype (Sousa da Silva and Vranken, 2012), which is part of the AmberTools14 (Wang et al., 2006). The tripeptide has been solvated using tip4p/2005 water (Abascal and Vega, 2005) in a 3.5 Å box, which was sufficiently large to accommodate the tri-peptide. After equilibration, the temperature was kept at 300 K using a stochastic velocity-rescaling thermostat (Bussi et al., 2007) with a relaxation time of 1 ps, and the pressure was set to 1 bar using a Berendsen barostat (Berendsen et al., 1984) with a relaxation time of 1 ps. The electrostatic interactions were calculated by smooth particle-mesh Ewald summation, and the Lennard-Jones interactions were smoothly truncated between 0 and 10 Å. The simulation time was 50 ns; during this period, all chi, phi, and psi angles were determined resulting in the joint probability distributions of each angle combination. Binning the probabilities and calculating the median angle of each bin led to several distinct angle combinations, their probability of occurrence, and their relative free energy. As described for the rotamer library of SwissSidechain (Gfeller et al., 2013), the angle combinations were processed to enable their import into PyMOL. This import provided the prerequisites for the interactive assessment of mutation experiments introducing UAAs. For MD simulations, conformations of minimal clashes with the adjacent residues were chosen for each UAA.

Expression and Purification of ImGPS

HisF and HisH proteins were produced by heterologous gene expression in *E. coli* BL21 Gold (DE3) (Agilent Technologies). Strains containing the respective vector were grown in lysogeny broth (LB) medium (2–4 L) at 37°C to an OD_{600} of 0.6. Protein expression was induced after 30 min at 30°C with 0.5 mM IPTG. Incubation overnight at 30°C was followed by harvesting of bacterial pellets and suspension in either 50 mM Tris•HCl pH 7.5, 100 mM NaCl, and 10 mM imidazole (HisF) or 50 mM KP pH 7.5, 100 mM NaCl, and 10 mM imidazole (HisH). Proteins were obtained from the supernatant after sonication and repeated centrifugation steps. The majority of *E. coli* proteins were precipitated in a heat step (15 min, 75°C HisF and 15 min, 70°C HisH) and subsequent centrifugation. Proteins were subjected to nickel-affinity chromatography (HisTrap™ FF Crude column, 5 mL, GE Healthcare) and eluted with a linear gradient of imidazole (10 → 750 mM). Fractions containing the proteins were identified by SDS-PAGE analysis, pooled and further purified with a size-exclusion chromatography column (Superdex 75 HiLoad 26/600, GE Healthcare) by using 50 mM HEPES pH 7.5 and 100 mM NaCl as running buffer. Fractions were checked on SDS-PAGE analysis for > 90% purity, pooled, concentrated, and dripped into liquid nitrogen for storage at −80°C.

For incorporation of UAA into HisF, a slightly adjusted expression protocol was used. After cotransformation of the pET28a vector, carrying the desired TAG codon, and pEVOL_NBY, pEVOL_AzoF or pEVOL_mNPK, respectively, into *E. coli* BL21 Gold (DE3), the strains were grown in LB medium (6 L) at 37°C to an OD_{600} of 0.6. Then, bacterial pellets were harvested by centrifugation at room temperature and suspended in terrific broth (TB) medium (600 mL). Bacterial growth was resumed to an OD_{600} of approximately 10 at 37°C and incorporation was induced by addition of 1 mM UAA and 0.02% L-arabinose. Omission of IPTG turned out to be favorable for expression of UAA-HisF proteins, as higher yields could be obtained in its absence. Cultures were incubated overnight at 30°C and proteins were purified as described above.

Tryptic Digest and MS Analysis

Recombinant *T. maritima* HisF proteins were run on a 15% SDS-PAGE and stained with Coomassie G250 (SimplyBlue SafeStain, Lifetech). Protein bands were cut out from the gel, washed with 50 mM NH_4HCO_3, 50 mM NH_4HCO_3/acetonitrile (3/1), 50 mM NH_4HCO_3/acetonitrile (1/1) and lyophilized. After a reduction/alkylation treatment and additional washing steps, proteins were *in gel* digested with trypsin (Trypsin Gold, mass spectrometry grade, Promega) overnight at 37°C. The resulting peptides were sequentially extracted with 50 mM NH_4HCO_3 and 50 mM NH_4HCO_3 in 50% acetonitrile. After lyophilization, peptides were reconstituted in 20 μL 1% TFA and separated by reversed-phase chromatography. An UltiMate 3000 RSLCnano System (Thermo Fisher Scientific, Dreieich) equipped with a C18 Acclaim Pepmap100 preconcentration column (100 μm i.D.x20 mm, Thermo Fisher Scientific) and an Acclaim Pepmap100 C18 nano column (75 μm i.d.x250 mm, Thermo Fisher Scientific) was operated at flow rate of 300 nL/min and a 60 min linear gradient of 4% to 40% acetonitrile in 0.1% formic acid. The LC was online-coupled to a maXis plus UHR-QTOF System (Bruker Daltonics) via a CaptiveSpray nanoflow electrospray source. Acquisition of MS/MS spectra after CID fragmentation was performed in data-dependent mode at a resolution of 60,000. The precursor scan rate was 2 Hz processing a mass range between m/z=175 and m/z=2,000. A dynamic method with a fixed cycle time of 3 s was applied via the Compass 1.7 acquisition and processing software (Bruker Daltonics). Prior to database searching with Protein Scape 3.1.3 (Bruker Daltonics) connected to Mascot 2.5.1 (Matrix Science), raw data were processed in Data Analysis 4.2 (Bruker Daltonics). A customized database comprising the *T. maritima* entries from UniProt as well as manually added sequences of the mutated HisF proteins and common contaminants was used for database search with the following parameters: enzyme specificity trypsin with two missed cleavages allowed, precursor tolerance 10 ppm, MS/MS tolerance 0.04 Da. As general variable modifications, deamidation of asparagine and glutamine, oxidation of methionine, carbamidomethylation or propionamide modification of cysteine were set. Specific variable modifications for identification of unnatural amino acids were *o*-nitrobenzyltyrosine (**NBY**), *o*-nitropiperonyltyrosine (**NPY**), methyl-*o*-nitropiperonyllysine (**mNPK**), or phenylazophenylanine (**AzoF**). **AzoF** was detected as modification of phenylalanine, why each position of **AzoF**

incorporation was changed to a phenylalanine in the query, respectively. **NBY** and **NPY** were detected as modification of tyrosine, why each position of **NBY** and **NPY** incorporation was changed to a tyrosine in the query, respectively. Spectra of peptides containing unnatural amino acids were inspected manually.

Screening of HisF Positions

Caged HisF proteins were either used "as isolated", or irradiated for 20 min at 365 nm (conventional UV lamp, two fluorescent black light bulbs with 8 W, Sylvania, settings: 250 mA and 220 V). The proteins were each diluted to 30 μM and cooled in a metal rack during irradiation. HisF proteins containing **AzoF** were also used either "as isolated", irradiated for 40 s at 365 nm (conversion to the *Z* state) or further irradiated for 40 s at 420 nm (back conversion to the *E* state) (High Power LED, Seoul Viosys, settings: 600 mA and 4.5 V). Irradiation occurred in ~ 1 cm distance to the light source. UAA-HisFs were then mixed at equimolar concentrations with wt-HisH in 50 mM HEPES pH 7.5, and 100 mM NaCl to form the ImGPS complex and kept in the dark. HisH activity was determined in a coupled enzymatic assay (Beismann-Driemeyer and Sterner, 2001) with glutamate-dehydrogenase (GDH, 99% pure, Roche) as auxiliary enzyme and NAD^+ as co-substrate. Reaction conditions included 50 mM Tris-acetate pH 8.5, glutamine in saturation (5 mM), 1 g/L GDH, 70 μM ProFAR, and 10 mM NAD^+. All reactions were set up in a 96-well plate, started with 2 μM ImGPS and followed continuously at λ = 340 nm [$\Delta\epsilon_{340}(NADH-NAD^+)$ = 6,300 M^{-1} cm^{-1}] using a plate reader (Tecan Infinite M200 Pro).

Native MS Analysis

Identity of fF23NBY, fY39NBY, fY39NPY, and fK99mNPK and decaging efficiencies were analyzed by online buffer exchange MS using an UltiMate™ 3000 RSLC (Thermo Fisher Scientific) coupled to an Exactive Plus EMR Orbitrap instrument (Thermo Fisher Scientific) modified to incorporate a quadrupole mass filter and allow for surface-induced dissociation (VanAernum et al., 2019). Proteins were either analyzed in their "as isolated" state or after exposure to UV light (UVP BL-15; Analytik Jena US; CA 91786) for 20 min. 100 pmol protein were injected and online buffer exchanged to 200 mM ammonium acetate, pH 6.8 (AmAc) by a self-packed buffer exchange column (Waitt et al., 2008) (P6 polyacrylamide gel, BioRad) at a flow-rate of 100 μL per min. Mass spectra were recorded for 1,000–8,000 *m/z* at 35,000 resolution as defined at 200 *m/z*. The injection time was set to 200 ms. Voltages applied to the ion optics were optimized to allow for efficient ion transmission while minimizing unintentional ion activation. fY39NPY and fK99mNPK showed a minor tendency to form higher oligomers under the given conditions. Only *m/z* corresponding to the monomer were considered for deconvolution and subsequent relative quantitation. Mass spectra were deconvoluted with UniDec version 2.6.5 (Marty et al., 2015) using the following processing parameters: sample mass every 0.3 Da; peak FWHM 0.3 Thompson, Gaussian peak shape function.

CD Analysis

Circular dichroism (CD) spectra in the far-UV range of 190–250 nm were recorded in a Jasco J-815 spectrophotometer with five accumulations. The spectra were measured with 5–12 μM protein in 50 mM potassium phosphate, pH 7.5 in a 0.1 cm cuvette at 25°C. The curves were smoothed in Origin 2018 (OriginLab). Data were normalized to obtain the mean residue ellipticity as described in (Kelly et al., 2005).

UV/Vis Analysis

UV/Vis spectra in the range of 230–550 nm were recorded in a Jasco V650 spectrophotometer. The spectra were measured with 15–30 μM protein in 50 mM HEPES, pH 7.5 and 100 mM NaCl in a 1 cm cuvette at 25°C. Irradiation of fS55AzoF took place by removing the cuvette from the spectrophotometer and exposing it to 365 nm UV light (conventional UV lamp, two fluorescent black light bulbs with 8 W, Sylvania, settings: 250 mA and 220 V) or subsequently to 420 nm visible light (High Power LED, Seoul Viosys, settings: 600 mA and 4.5 V). Caged proteins were irradiated by removing the cuvette and exposing it to 365 nm UV light with the conventional UV lamp. High-power irradiation was performed with a 365 nm LED (High Power LED, LED Engin, settings: 700 mA and 16 V) and a 420 nm LED (High Power LED, Seoul Viosys, settings: 300 mA and 4 V) installed perpendicular to the measurement beam. After each irradiation step, spectra were recorded and baseline corrected. Estimation of photostationary state composition from UV/Vis spectra was performed as described in (Calbo et al., 2017).

Analytical Size-Exclusion Chromatography

30 μM HisF monomer or 30 μM HisF mixed with 30 μM wt-HisH were subjected to a S75 10/300 GL (GE Healthcare) column pre-equilibrated in 50 mM HEPES, pH 7.5 and 100 mM NaCl. Samples were eluted in the same buffer, and protein as well as UAA peaks were detected at 280 nm, 334 nm (**AzoF**), and 360 nm (**NPY**, **mNPK**).

Steady-State Kinetics

UAA-HisFs were either used "as isolated", or irradiated with 365 nm light as described above (screening). Proteins were then used as single subunit for the determination of steady-state constants of HisF activity or mixed at equimolar concentrations with HisH in 50 mM HEPES, pH 7.5, and 100 mM NaCl. The ammonia-dependent HisF activity and the ImGPS activity were measured continuously following PrFAR turnover at 300 nm [$\Delta\epsilon_{300}$(PrFAR−AICAR) = 5,637 M^{-1} cm^{-1}] (Beismann-Driemeyer and Sterner, 2001). Reaction conditions of HisF activity measurements were 50 mM Tris-acetate pH 8.5, 100 mM ammonium acetate (saturated), 0.6 μM HisA (to convert ProFAR to PrFAR), varying concentrations of ProFAR (1–40 μM), and 0.1–0.3 μM HisF at 25°C. Reaction conditions of HisH activity measurements were 50 mM Tris-acetate pH 8.5, 0.6 μM HisA (to convert ProFAR to PrFAR), 40 μM ProFAR (saturated),

varying concentrations of glutamine (0.1–10 mM), and 0.1–0.2 μM ImGPS complex at 25°C. In order to circumvent the stimulating effect of NAD^+ on HisH activity as in the GDH-coupled assay used in the screening (Beismann-Driemeyer and Sterner, 2001), a different continuous assay was performed for determination of ProFAR activation constants (K_{ac}). For this, the HisH reaction was coupled to glutamate-oxidase (GOX, 99% pure, Sigma-Aldrich) producing α-ketoglutarate and hydrogen peroxide out of glutamate. In a second step, hydrogen peroxide was turned over by horse-radish peroxidase (HRP, 99% pure, Sigma-Aldrich) together with colorless 2-aminoantipyrine and phenol to form red-colored quinoneimine. The reaction was followed at 505 nm with ϵ_{505}(quinoneimine) = 6,400 M^{-1} cm^{-1} (Green and Hill, 1984). Initial optimization of this coupled reaction identified pH 7.0 as an optimum, hence, reaction conditions included 20 mM Tris•HCl pH 7.0, 5 mM glutamine (saturated), 20 mU/mL GOX, 0.15 g/L HRP, 1 mM 4-aminoantipyrin, 1 mM phenol, varying concentrations of ProFAR (5–70 μM), and 0.05–1 μM ImGPS complex at 25°C. All reactions were started by the addition of the ImGPS complex. Measurements were performed either with a Jasco V650-UV/Vis spectrophotometer or for the GOX assay with a plate reader (Tecan Infinite M200 Pro). Activities were deduced from the initial slopes of the transition curves and analyzed using the Michaelis-Menten equation.

Direct Photo-control of HisH Activity

ProFAR-stimulated HisH-activity was followed in the GOX-assay with increased concentrations of auxiliary enzymes to prevent irradiation effects on the light-sensitive cofactors of GOX and HRP. Reaction conditions included 20 mM Tris•HCl pH 7.0 or pH 8.5, 5 mM glutamine, 100 mU/mL GOX, 70 μM ProFAR, 0.75 g/L HRP, 3 mM 2-aminoantipyrin, 3 mM phenol, and 0.2 μM ImGPS complex. All reactions were started by the addition of the ImGPS complex and followed at λ = 505 nm with ϵ_{505}(quinoneimine) = 6,400 M^{-1} cm^{-1} (Green and Hill, 1984). Measurements were performed in a plate reader (Tecan Infinite M200 Pro) and activities were deduced from the slopes of the transition curves. Each sample was measured simultaneously with three controls in a 96-well plate. Negative and positive controls, started with non-irradiated and pre-irradiated ImGPS complex, respectively, were kept in the dark. Pre-irradiation was performed with 365 nm for 1.5 min (High Power LED, LED Engin, settings: 700 mA and 16 V). In a fourth control, buffer was added instead of the ImGPS complex, in order to monitor the remaining effect on the auxiliary enzymes. Irradiation of each sample and its buffer control occurred by pausing the measurement in the plate reader, taking out the plate, and irradiating the two wells while keeping the other two wells (positioned furthest away), containing both negative and positive control, covered. wt and caged proteins were irradiated once after 60 min for 1.5 min with 365 nm (High Power LED, LED Engin, settings: 700 mA and 16 V). **AzoF**-proteins were irradiated after 40 min for 30 s with the 365 nm High Power LED and after another 40 min for 1 min with a 420 nm High Power LED (High Power LED, Seoul Viosys, settings: 300 mA and 4 V). The turnover curves of irradiated samples were baseline corrected by subtraction of the buffer control. Negative and positive controls were baseline corrected by subtraction of a linear fit of the first 40–60 min of the buffer control.

Crystallization

ImGPS(fY39NPY) and ImGPS(fS55AzoF) were crystallized following a protocol described for PDB-ID 1gpw (Douangamath et al., 2002). Prior to crystallization, His_6-tags of HisF and HisH proteins were removed by cleavage with TEV protease (20 μg per 1 mg protein) at 20°C overnight. Pure complex was obtained by size-exclusion chromatography (Superdex 75 HiLoad 26/600, GE Healthcare) of equimolar amounts of HisH and HisF in 10 mM Tris•HCl pH 8.0. Fractions containing the complex were identified by SDS-PAGE analysis, pooled and concentrated.

1 μL of ~ 20 mg/mL complex was then mixed with 1 μL of reservoir solution, containing 13–17% PEG 8000, 0.7–0.9 M ammonium nitrate, 0.1 M HEPES•NaCl pH 8.5, 10 mM DTT, and 5% (v/v) MPD. Crystals grew within one week in a rod-like morphology similar to wt-ImGPS using the hanging drop vapor diffusion method. Crystals were mounted onto a nylon loop and shock-frozen in liquid nitrogen without addition of further cryoprotectants.

Data Collection, Structure Solution and Refinement

Data sets were collected to ~ 2.8 Å resolution each, using synchrotron radiation from the Swiss Light Source (SLS), Switzerland at beamline PXIII and PXI. Data collection for both ImGPS(fY39NPY) or ImGPS(fS55AzoF) was done at cryogenic temperature. Data were processed using XDS (Kabsch, 1993), and the data quality was assessed using the program PHENIX (Adams et al., 2002). Structures were solved by molecular replacement with MOLREP within the CCP4i (Potterton et al., 2004) suite using PDB-ID 1gpw (Douangamath et al., 2002) as search model. Initial refinement was performed using REFMAC (Murshudov et al., 1997). The model was further improved in several refinement rounds using automated restrained refinement with PHENIX (Adams et al., 2002) and interactive modelling with *Coot* (Emsley and Cowtan, 2004), in which the mutation D11N present in PDB-ID 1gpw was restored to wt. The final model was analyzed using the program MolProbity (Davis et al., 2007). Structural and refinement statistics are listed in Tables S2 and S3.

Molecular Dynamics (MD) Simulations

MD simulations were conducted with Yasara (Krieger et al., 2004), version 17.4.17, force field AMBER03 on the basis of the ImGPS complex (PDB-ID 1gpw, chains C, D). Residue 11 in HisF was mutated back to the wt aspartate, missing side chains in flexible loops of HisF were replaced, and PrFAR was positioned as in the *Saccharomyces cerevisiae* ImGPS structure (PDB-ID 1ox5). UAAs were incorporated into HisF and sterically oriented by using the previously generated rotamer library. A hexagonal simulation cell was created, which was 5 Å larger than the protein along each axis, filled with water to a density of 0.997 g/mL and with counterions

to a final concentration of 0.9% NaCl. Three independent simulation runs of 100 ns length were performed with slightly different temperatures (25 ± 0.01°C) adjusted by the Berendsen thermostat to alter the starting conditions. Each minimization consisted of an initial equilibration step of 14 ps length. During the next 100 ns of simulation, a snapshot was recorded every 20 ps. The root-mean-square fluctuation (RMSF) value representing the deviation from the mean position for each atom of the ImGPS complexes was deduced for each series of snapshots. For this calculation, the function R:rmsf (version 3.6.0) was used (Skjærven et al., 2014); corresponding values resulting from the three temperature-specific runs were averaged. By means of PyMol (Schrödinger, 2015), RMSF values were mapped onto the 3D structure and visualized. RMSF values were read as B factors, and a color gradient ranging from blue over white to red was used to indicate RMSF values in the range of [0 Å, 1.5 Å]; values $\geq$ 1.5 Å are shown in red. Distances between specific atoms were calculated using the Yasara function Distance; corresponding values resulting from the three MD runs were pooled in histograms (bin size 0.05 Å) and plotted in the form of probability distributions. Minima and maxima were determined by utilizing a sliding window of size five. By means of the plotly (https://plot.ly) function Histogram2DContour, a 2D contour density plot was generated for the pair of distances $hH178_{N\delta}$–$hE180_{C\delta}$ and $fD98_{C\gamma}$–$hH178_{N\epsilon}$. Scripts for automated processing were written in Python 2.7 (Oliphant, 2007).

Nano Differential Scanning Fluorimetry

The Nano-DSF technique was used to analyze the thermal stability of proteins. All samples were prepared in 30 μM concentration, filled into capillaries and scanned at 330 nm and 350 nm during a heating gradient of 35–95°C (1°C min^{-1}) in a NanoDSF Prometheus NT.48 (Nanotemper technologies). For analysis of melting temperatures (T_m), the ratio 350 nm/330 nm was plotted against the temperature. T_m was determined as the transition midpoint with Origin 2018 (OriginLab), also defined as the peak maximum in the first derivative (*f'*) of the denaturing curve. Melting temperatures were measured for wt-HisF, fY39NPY, fS55AzoF, and fK99mNPK in their "as isolated" forms. fY39NPY and fK99mNPK were also analyzed after 30 min irradiation (365 nm, conventional UV lamp, two fluorescent black light bulbs with 8 W, Sylvania, settings: 250 mA and 220 V). fS55AzoF, however, could only be measured in its *E*-enriched state, since its *Z*-enriched state is thermally unstable.

QUANTIFICATION AND STATISTICAL ANALYSIS

All activity data points shown in the figures are the mean ± SEM (standard error of mean) of at least two technical replicates. Fitting of data points using Origin 2018 (OriginLab) resulted in the fitting values ± SE (standard error; scaled with square root of reduced Chi-square). Detailed description of the used fitting functions can be found in the figure legends, table footer, and the Method Details section. For data calculated from two values ± SE(M.), the SE was calculated according to the Gaussian law of error propagation. All analyses and data representations were performed and assembled in Origin 2018 (OriginLab) except for MS analysis (Data Analysis 4.2, Bruker Daltonics), native MS analysis (UniDec version 2.6.5 (Marty et al., 2015)), crystal structure analysis (MolProbity (Davis et al., 2007)), and computational analysis (Yasara (Krieger et al., 2004), plotly (https://plot.ly), Gromacs (version 5.1.2) (Berendsen et al., 1995)). Details can be found in the method details section and data collection and refinement statistics of crystal structure analysis can be found in the Supplemental Information.

DATA AND CODE AVAILABILITY

The crystal structures generated during this study are available at the Protein Data Bank (http://www.rcsb.org) with the following accession numbers: ImGPS(fY39NPY) PDB: 6rtz and ImGPS(fS55AzoF) PDB: 6ru0.