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

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

Authors

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

Correspondence

reinhart.stern@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.
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,6 Dominik Horinek,6 Burkhard König,2 Rainer Merkl,1 and Reinhard Sterner1,7,*

1Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
2Institute of Organic Chemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
3Centre for Systems Chemistry, Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands
4Institute of Biochemistry, Genetics and Microbiology, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
5Department of Chemistry and Biochemistry and Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, OH 43210, USA
6Institute of Physical and Theoretical Chemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany
7Lead 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 → fS55AzoF2) 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 spatio-temporal 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; Hül et al., 2018; Losi et al., 2018; Lachmann 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 allosterity 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 allosterity 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 associated 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'-phosphoribulolysylformimino)-5-aminoimidazol-4-carboxamide ribonucleotide (PrFAR) (Lisi et al., 2017) or its analog N'-(5'-phosphoribosylformimino)-5-aminoimidazol-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; Lipchok 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 then reacts with PrFAR to produce imidazole glycerol phosphate (ImGP) and 5-aminoimidazol-4-carboxamidribotide (AICAR), which are...
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<sub>UAA</sub>/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 allosteric activity (Luo et al., 2019).

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 (Kla´ n 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 (Jpretty 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 (Zymanski 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). Farthest 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 distribution UAA-HisFs for light regulation of ImGPS allostery.

Finally, molecular dynamics (MD) simulations revealed how irradiation affects HisH activity. We then characterized these proteins containing the remaining eight UAA-HisFs were measured step by step, LRFs of HisH in ImGPS(fF23NPY), ImGPS(fY39NPY), and ImGPS(fS55AzoF) designates the ImGPS complex containing AzoF at position S55 of HisH.

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(fF23NPY), and ImGPS(fY39NPY) did not reach 20% WT activity, though irradiation of fY39NPY should in principle restore the WT situation. Native MS suggests that considerable quantities (20%–60%) of reduced NBY existed in both fF23NPY and fY39NPY, 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.
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: 6rz2) 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 a80 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, 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 (kcat), while fK99mNPK, which is far remote from the active site has a kcat value similar to that of WT-HisF (∼173 min⁻¹) (Table 1; Figure S4). Moreover, we found that the Michaelis constants (KMPrFAR) of all UAA-HisFs were identical to WT-HisF (∼2.5 mM), 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 kcat 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 KMPrFAR values were unaltered (∼0.4 mM) indicating that substrate binding to the active site was not compromised.

(PDB: 1gpw; Douangamath et al., 2002). The distances of the three finally selected positions to the substrate glutamine, bound at the active site of HisF (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 (hνUV, 365 nm) and, for AzoF-HisFs, after additional irradiation with visible light (hνVIS, 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 mg 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.
Glutamine was not compromised (Table 1; Figure S4). Moreover, the ProFAR concentrations that were required for the UAA-HisFs to induce 50% HisH activity (Kac) were also not changed (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 to fS55AzoZ (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)**

<table>
<thead>
<tr>
<th>Protein State</th>
<th>Protein/ImGPS Activity</th>
<th>kcat (min⁻¹)</th>
<th>Km,ImGPS (μM)</th>
<th>kcat/Km (M⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT-ImGPS</td>
<td>172.6 ± 1.6</td>
<td>2.5 ± 0.1</td>
<td>11.5 ± 0.5 × 10²</td>
</tr>
<tr>
<td>fS55AzoF E</td>
<td>48.6 ± 0.9</td>
<td>4.8 ± 0.1</td>
<td>1.7 ± 0.0 × 10⁵</td>
<td></td>
</tr>
<tr>
<td>fS55AzoF Z</td>
<td>18.6 ± 0.2</td>
<td>2.4 ± 0.2</td>
<td>1.3 ± 0.1 × 10⁵</td>
<td></td>
</tr>
<tr>
<td>fY39NPY</td>
<td>11.3 ± 0.4</td>
<td>3.4 ± 0.5</td>
<td>0.6 ± 0.1 × 10⁵</td>
<td></td>
</tr>
<tr>
<td>fK99mNPK</td>
<td>112.4 ± 4.7</td>
<td>3.2 ± 0.5</td>
<td>5.9 ± 0.9 × 10⁵</td>
<td></td>
</tr>
<tr>
<td>Glutamine-Dependent ImGPS Activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein State</td>
<td>Protein/ImGPS Activity</td>
<td>kcat (min⁻¹)</td>
<td>Km,ImGPS (μM)</td>
<td>kcat/Km (M⁻¹ s⁻¹)</td>
</tr>
<tr>
<td></td>
<td>WT-ImGPS</td>
<td>41.4 ± 1.1</td>
<td>0.78 ± 0.05</td>
<td>8.6 ± 1.1 × 10²</td>
</tr>
<tr>
<td>ImGPS(fS55AzoF) E</td>
<td>6.0 ± 0.2</td>
<td>0.23 ± 0.03</td>
<td>4.3 ± 0.6 × 10²</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fS55AzoF) Z</td>
<td>2.6 ± 0.2</td>
<td>0.33 ± 0.11</td>
<td>1.3 ± 0.4 × 10²</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fY39NPY)</td>
<td>3.5 ± 0.2</td>
<td>0.25 ± 0.05</td>
<td>2.3 ± 0.5 × 10²</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fK99mNPK)</td>
<td>7.3 ± 0.4</td>
<td>0.41 ± 0.09</td>
<td>3.0 ± 0.7 × 10²</td>
<td></td>
</tr>
<tr>
<td>ProFAR-Dependent HisH Activity (pH 7.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein State</td>
<td>Protein/ImGPS Activity</td>
<td>kcat (min⁻¹)</td>
<td>Km,ProFAR (μM)</td>
<td>LRF</td>
</tr>
<tr>
<td></td>
<td>WT-ImGPS</td>
<td>16.8 ± 0.3</td>
<td>32.8 ± 1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>ImGPS(fS55AzoF) E</td>
<td>5.4 ± 0.4</td>
<td>19.5 ± 3.8</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fS55AzoF) Z</td>
<td>2.3 ± 0.2</td>
<td>35.1 ± 5.0</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fY39NPY)</td>
<td>1.2 ± 0.1</td>
<td>32.3 ± 4.2</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>ImGPS(fK99mNPK)</td>
<td>3.5 ± 0.3</td>
<td>19.6 ± 3.7</td>
<td>4.0</td>
<td></td>
</tr>
</tbody>
</table>

**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, kcat, values of HisH reached only ~40% and ~80% of the WT kcat value (~17 min⁻¹) (Table 1). Native MS was used to check whether incomplete decaging was caused by reduction of the nitrophenyl 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) than in WT-ImGPS and why HisH activity is lower in ImGPS(fS55AzoF) than in ImGPS(fS55AzoF) (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 6B); representative MD snapshots are shown in Figure S7. In WT-ImGPS, three ensembles (ensWT,1-ensWT,3) were observed. In ensWT,1 (prevalence 24%), hH178 assumes a pose with relatively short distances to both hE180 (d ~ 3.6 Å) and fD98 (d ~ 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 ensWT,2 (prevalence 44%), the hydrogen-bond distance between hH178 and hE180 remains unaltered, but the mean distance of hH178 to fD98 increases (d ~ 8.6 Å), which makes the formation of a water-mediated hydrogen bond less likely. In ensWT,3 (prevalence 32%) hH178 cannot interact with hE180 (d ~ 5.1 Å) and is still far from fD98 (d ~ 8.6 Å). Obiously, these three specific conformational ensembles of hH178 are optimal to HisH activity. In accordance with this assumption, ImGPS(fS55AzoF) adopts only one conformational ensemble (ens) resembling ensWT,2, whereas ImGPS(fS55AzoF) adopts four ensembles (ensWT,1-ensWT,4) where only ensWT,1 resembles ensWT,2.

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 (Lipchick and Loria, 2010; Rivarla 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) than in the least active ImGPS(fS55AzoF), which might allow it to act as oxyanion intermediate stabilizer. In addition, the motif adopts a different configuration in ImGPS(fS55AzoF), 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; Rivarla et al., 2012; Negre et al., 2018). We considered the distance between hG52 of the glutamine binding site and the opposite fI120 as an indirect measure of this movement. The mean distances of hG52-fI120 suggest that the interface is more open in WT-ImGPS (d ~ 10.7 Å) and in ImGPS(fS55AzoF) (d ~ 11.3 Å) than in the least active ImGPS(fS55AzoF) (d ~ 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) containing the opening at hG52-fI120 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). The mean RMSF values are in full agreement with this postulate: overall flexibility in WT-ImGPS (0.98 Å) and ImGPS(fS55AzoF) (1.00 Å) is higher than in ImGPS(fS55AzoF) (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 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, the disadvantageous effects of fS55AzoF seem to be generally less severe, as indicated by a more structured catalytic site, a typical h49-PGVG-partial orientation, and a HisH:HisF arrangement that resemble the WT-ImGPS.

**F39NPY Causes a More Rigid ImGPS Complex**

We continued our MD analyses with the UAA-HisF F39NPY, 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(F39NPY) than in decaged ImGPS(F39NPY), which is identical to WT-ImGPS (Figure 5A). We observed three structural ensembles of hH178, ensY,1 (prevalence 10%), ensY,2 (prevalence 34%), and ensY,3 (prevalence 49%) that resemble those in WT-ImGPS. However, the h49-PGVG-52 motif was more rigid than in WT-ImGPS (Figure 5C) and, similarly as deduced for ImGPS(fS55AzoF), the breathing motion at the HisH:HisF interface seems less extensive: the mean hG52-fI120 distance is ~1.5 Å shorter than in WT-ImGPS (Figure 5D).

Consequently, both right and left halves of ImGPS(F39NPY) 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).
Figure 5. Comparative MD Simulation Analysis of WT-ImGPS, ImGPS(fS55AzoF), ImGPS(fS55AzoF), 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 hH178N-d-hE180C and fD98C-g-hH178N. The orientation of the hH178 residue in the 2D contour plots suggests for WT-ImGPS three ensembles (ensWt,1-ensWt,3), for ImGPS(fS55AzoF) only one (ensE) resembling enswt,2, for ImGPS(fS55AzoF) four (ensZ,1-ensZ,4) with ensZ,1 resembling enswt,1, for ImGPS(fY39NPY) three (ensY,1-ensY,3) with all three resembling enswt,1-enswt,3, and for ImGPS(fK99mNPK) two ensembles (ensK,1 and ensK,2) with ensK,2 resembling enswt,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, ensK,1 (prevalence 56%) and ensK,2 (prevalence 29%), with the latter resembling the ensVT,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 ensK,1 and ensK,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 fF120-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 ~ 4.6 Å ↔ d ~ 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 UAs, 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-con

**DISCUSSION**

In search for HisF positions that enable photo-control of the activity of its partner protein HisH with UAs, 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-con
adopted, or in ImGPS(fS55AzoF), 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) and in particular in ImGPS(fS55AzoF). 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). However, in our MD simulations the backbone of the h49-PGVG-52 motif and observed significant conformational changes. 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). However, in our MD simulations 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 glutamate, its NH group points into the glutamate binding pocket, and it is highly flexible in PrFAR-bound WT-ImGPS. In contrast, in the least active ImGPS complex, 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) compared with ImGPS(S55AzoF) 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).
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-o-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:

- **KEY RESOURCES TABLE**
- **LEAD CONTACT AND MATERIALS AVAILABILITY**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - **Chemistry, Materials and Instrumentation**
  - **Synthesis of AzoF**
  - **Synthesis of NPY**
  - **Synthesis of mNPK**
- **Biochemical Synthesis of ProFAR**
- **Auxiliary Enzymes**
- **Subcloning of the hisF and hisH Genes**
- **Subcloning of the aaRS/trNA_{CUA} Pairs**
- **Site-Directed Mutagenesis of hisF**
- **Building a Rotamer Library**
- **Expression and Purification of ImGPS**
- **Tryptic Digest and MS Analysis**
- **Screening of HisF Positions**
- **Native MS Analysis**
- **CD Analysis**
- **UV/Vis Analysis**
- **Analytical Size-Exclusion Chromatography**
- **Steady-State Kinetics**
- **Direct Photo-control of HisH Activity**
- **Crystallization**
- **Data Collection, Structure Solution and Refinement**
- **Molecular Dynamics (MD) Simulations**
- **Nano Differential Scanning Fluorimetry**
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND CODE AVAILABILITY**

**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.


### STAR METHODS

#### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AzoF</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>NPY</td>
<td>Accela ChemBio</td>
<td>Cat#SY008208</td>
</tr>
<tr>
<td>mNPK</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>ProFAR</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>HisA</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>HisG/IE</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>wt/UAAP-HisF</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>HisH</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (GDH)</td>
<td>Sigma-Aldrich</td>
<td>CAT#10197734001</td>
</tr>
<tr>
<td>NAD+, free acid</td>
<td>Sigma-Aldrich</td>
<td>CAT#124542</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Sigma-Aldrich</td>
<td>CAT#G8540</td>
</tr>
<tr>
<td>Glutamate-oxidase (GOX)</td>
<td>Sigma-Aldrich</td>
<td>CAT#G1924</td>
</tr>
<tr>
<td>Hydrogen peroxidase (HRP)</td>
<td>Sigma-Aldrich</td>
<td>CAT#P8125</td>
</tr>
<tr>
<td>Phenol</td>
<td>Merck</td>
<td>CAT#100206</td>
</tr>
<tr>
<td>2-aminoantipyrine</td>
<td>Sigma-Aldrich</td>
<td>CAT#06800</td>
</tr>
<tr>
<td><strong>Deposited Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ImGPS structure from <em>Thermotoga maritima</em></td>
<td>Douangamath et al., 2002</td>
<td>PDB-ID 1gpw</td>
</tr>
<tr>
<td>ImGPS structure from <em>Saccharomyces cerevisiae</em></td>
<td>Chaudhuri et al., 2003</td>
<td>PDB-ID 1ox5</td>
</tr>
<tr>
<td>ImGPS(fS55AzoF) structure</td>
<td>This paper</td>
<td>PDB-ID 6ru0</td>
</tr>
<tr>
<td>ImGPS(fY39NPY) structure</td>
<td>This paper</td>
<td>PDB-ID 6rtz</td>
</tr>
<tr>
<td><strong>Oligonucleotides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primers: see Table S1</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Recombinant DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pET28aTEV</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pET28a_HisFwt</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pET28aTEV_Bsal</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pET28a_HisHwt</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEVOL_mNPK</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pET28a vectors containing stop codon mutations: see Table S1</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEVOL_AzoF</td>
<td>Peter Schultz (Scripps Research Institute, La Jolla, USA) (Young et al., 2010; Bose et al., 2006)</td>
<td>N/A</td>
</tr>
<tr>
<td>pEVOL_NPY</td>
<td>Peter Schultz (Scripps Research Institute, La Jolla, USA) (Young et al., 2010; Deiters et al., 2006)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Software and Algorithms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gromacs version 5.1.2</td>
<td>Berendsen et al., 1995</td>
<td><a href="https://www.gromacs.org">https://www.gromacs.org</a></td>
</tr>
<tr>
<td>AmberTools14</td>
<td>Wang et al., 2006</td>
<td><a href="https://www.ambermd.org/Ambertools.php">https://www.ambermd.org/Ambertools.php</a></td>
</tr>
<tr>
<td>Origin 2018</td>
<td>OriginLab</td>
<td><a href="https://www.originlab.com">https://www.originlab.com</a></td>
</tr>
<tr>
<td>Data Analysis 4.2</td>
<td>Bruker Daltonics</td>
<td><a href="https://www.bruker.com">https://www.bruker.com</a></td>
</tr>
</tbody>
</table>

(Continued on next page)
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 (1H: 300 MHz, 13C: 75 MHz, T = 300 K) or Bruker Avance 400 MHz spectrometer (1H: 400 MHz, 13C: 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 1H-NMR.

Continued
(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 (2 x 300 mL) 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 Na2SO4. 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). 1H-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). 1H-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 1H-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). 1H 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 > 98% by NMR). 1H-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 1H-NMR.
**Biochemical Synthesis of ProFAR**

ProFAR was synthesized following the standard protocol developed by (Davisson et al., 1994) from 5-phospho-α-riboyl α-1-pyrophosphate 10 and adenosine triphosphate 11 in 50 mM NH₄COOCH₃ 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 (HA16/10, 20 mL, Pharmacia). ProFAR concentration was determined at 300 nm (ε₂₉₀ = 6069 M⁻¹ cm⁻¹) (Klem and Davisson, 1993) and ProFAR purity through the absorbance ratio A₂₉₀/A₂₆₀ for each fraction. Highly concentrated and > 90% pure (A₂₉₀/A₂₆₀ = 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 HisFc, 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**

HistA 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 (HiTrap 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, 300 mM NaCl.

HistG/IE was produced by heterologous gene expression in *E. coli* BL21 Gold (DE3) (Agilent Technologies) from pET21_HisGIE/IE (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 (HiTrap 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.

HistG/IE was produced by heterologous gene expression in *E. coli* BL21 Gold (DE3) (Agilent Technologies) from p_p_hisGIE/IE_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 OD₆₀₀ 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 hisF and hisH Genes**

The hisF and hisH genes were cloned into vectors encoding an N-terminal His6-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 NdeI and a HindIII 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_Bsal (Rohweder et al., 2018) designed for golden-gate cloning, was modified to encode for a TEV-cleavage site C-terminal of the N-terminal His6-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 from Finnzymes with 5’ phosphorylated (PHO) and HPLC-purified primers (Metabion) P18–P37 (Table S1). 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 (MjTyrRS-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 MjTyrRS-TyrT pair evolved for AzoF incorporation with the mutations Y32G, L65G, F108E, D158S, 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 pyrrolysyl-tRNA synthetase and its respective tRNA_CUA from Methanosarcina barkeri (MbpYRS-PyIT), 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 pyIT-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_pyIT.

Exchange of the first copy of MjtyrRS was prepared by amplifying pEVOL_AzoF_pyIT with primers P11 and P12 (Table S1) containing overhangs with a Bsal restriction site and flanking the MjtyrRS gene. MbpYRS, synthesized (GeneArt) with Bsal 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 Bsal 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 by MbpYRS 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_pyIT, P16 for pEVOL_mNPK+, and P17 for pEVOL_mNPK.

**Site-Directed Mutagenesis of hisF**

Stop codon 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 Acype (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-scaling 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 UAs. 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 OD600 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 OD600 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 OD600 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 NH4HCO3, 50 mM NH4HCO3/acetonitrile (3/1), 50 mM NH4HCO3/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 NH4HCO3 and 50 mM NH4HCO3 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 Scrape 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-nitropiperonylllysine (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 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 ImGMP and followed continuously at λ = 340 nm [ΔmM ([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 ProFAR turnover at 300 nm [ΔmM (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 HisF 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 (Kact). 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) coupled to glutamate-oxidase (GOX, 99% pure, Sigma-Aldrich) together with colorless 2-aminantipyrine and phenol to form red-colored quinoneimine. The reaction was followed at 505 nm with ε505(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-Cl pH 7.0, 5 mM glutamine (saturated), 20 μM/mL GOX, 0.15 g/L HRP, 1 mM 4-aminantipyrin, 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-assy 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-Cl pH 7.0 or pH 8.5, 5 mM glutamine, 100 μM/mL GOX, 70 μM ProFAR, 0.75 g/L HRP, 3 mM 2-aminantipyrin, 3 mM phenol, and 0.2 μM ImGPS complex. All reactions were started by the addition of the ImGPS complex and followed at = 505 nm with ε505(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, His6-tags of HisF and HisH proteins were removed by cleavage with TEV protease (20 μg/mL). Pools of HisF and HisH proteins were replaced, and PrFAR was positioned as in 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). UAs 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.

**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). UAs 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.

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
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. 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 hH178N–hE180C and fD98C–hH178N. 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 (Tm), the ratio 350 nm/330 nm was plotted against the temperature. Tm 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.