SUPPLEMENTARY MATERIAL

Bile Acid Binding Protein Functionalization Leads to a Fully Synthetic Rhodopsin Mimic

Katiuscia Pagano,[a,†] Marco Paolino,[b,†,*] Stefania Fusi,[b] Vinicio Zanirato,[c] Claudio Trapella,[d]

Germano Giuliani,[b] Andrea Cappelli,[b] Serena Zanzoni,[e] Henriette Molinari,[a] Laura Ragona,[a,*]

Massimo Olivucci[b,f,*]

[a]Istituto per lo Studio delle Macromolecole, CNR, Via A. Corti 12, 20133 Milano, Italy

[b]Dipartimento di Biotecnologie, Chimica e Farmacia (Dipartimento di Eccellenza 2018-2022), Università degli Studi di Siena, Via Aldo Moro 2, 53100 Siena, Italy

[c]Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Ferrara, Via Fossato di Mortara 17, 44121 Ferrara, Italy

[d]Laboratory for Technologies of Advanced Therapies (LTTA), Department of Chemical and Pharmaceutical Sciences, Università di Ferrara, 44121 Ferrara, Italy

[e]Centro Piattaforme Tecnologiche, Università di Verona, Strada Le Grazie, 37134 Verona, Italy

[f]Chemistry Department, Bowling Green State University, Bowling Green, Ohio, USA.

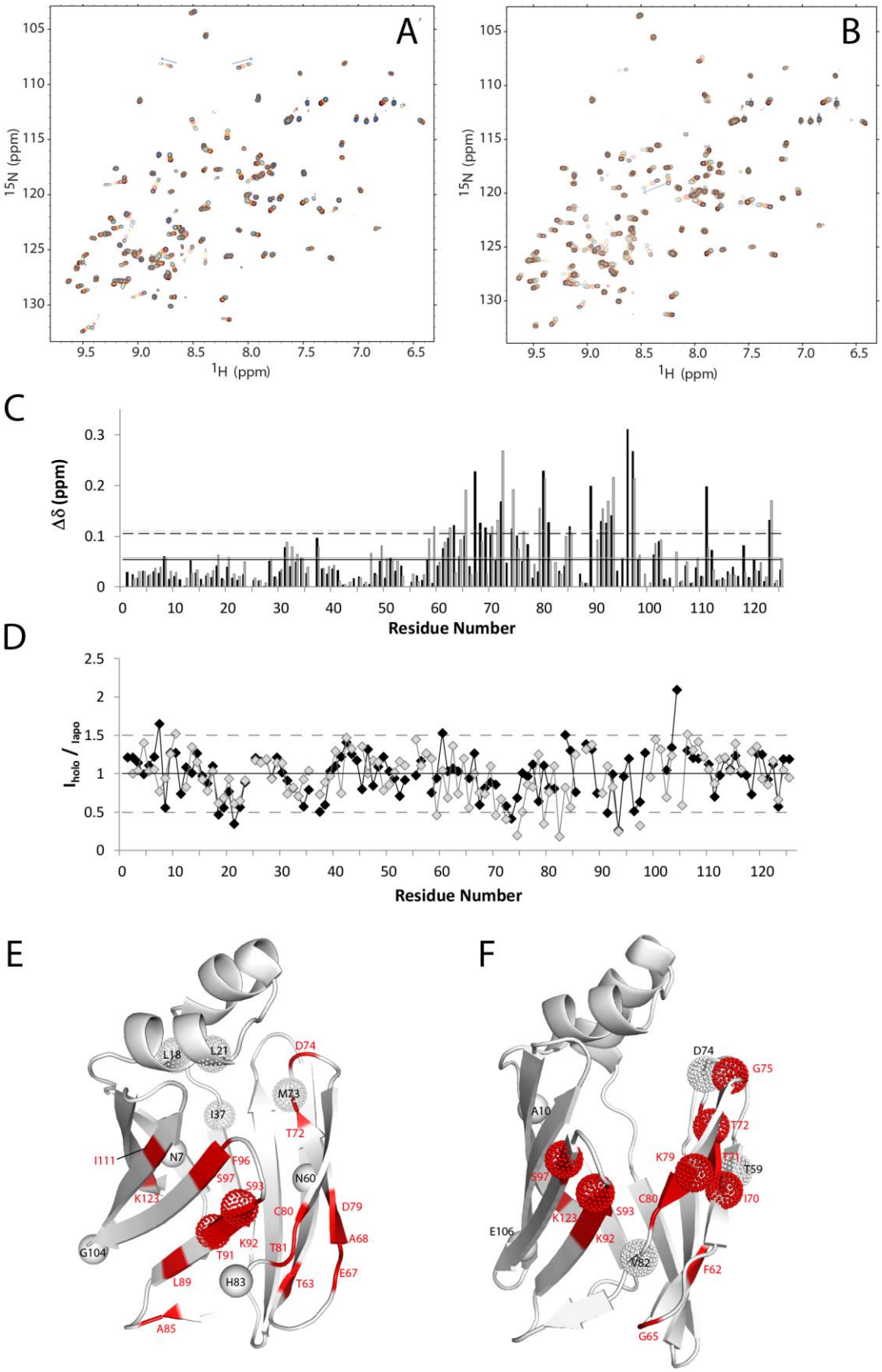


Figure S1. NMR titration data. Superposition of ^{1}H–^{15}N HSQC spectra of BABP (A) and BABP-SS (B) in the absence (black) and in the presence of **MeO-NAIP** at different ratios: BABP:**MeO-NAIP** 1:7 (red), 1:20 (yellow), 1:30 (blue). (C) Graphical representation of the combined H^N and ^{15}N CSP upon **MeO-NAIP** addition at a protein:**MeO-NAIP** 1:30 molar ratio, as observed for

BABP (black) and BABP-SS (grey). Straight and dashed lines indicate < CSP > and < CSP > + 1σ, respectively. (D) $^{1}H-^{15}N$ HSQC cross peaks intensities ratio profiles (defined as normalised holo cross-peak volume/ apo cross-peak volume) upon **MeO-NAIP** addition to BABP (black) and BABP-SS (grey) proteins. The horizontal dashed lines mark intensity changes higher/lower than 50%. (E) BABP and (F) BABP-SS residues showing CSP higher than (⟨CSP⟩ + 1σ) upon **MeO-NAIP** addition are highlighted in red on the structure. Filled and dotted spheres indicate residues showing significant (> 50%) increased or decreased intensity, respectively. 3D structure for graphical representation are (PDB id): 2jn3 and 2lfo for BABP and BABP-SS, respectively. For both proteins, residues significantly affected by the interaction were located in the region 65-96, namely in EF, GH loops and F, G strands, as highlighted on the protein structure (E and F).

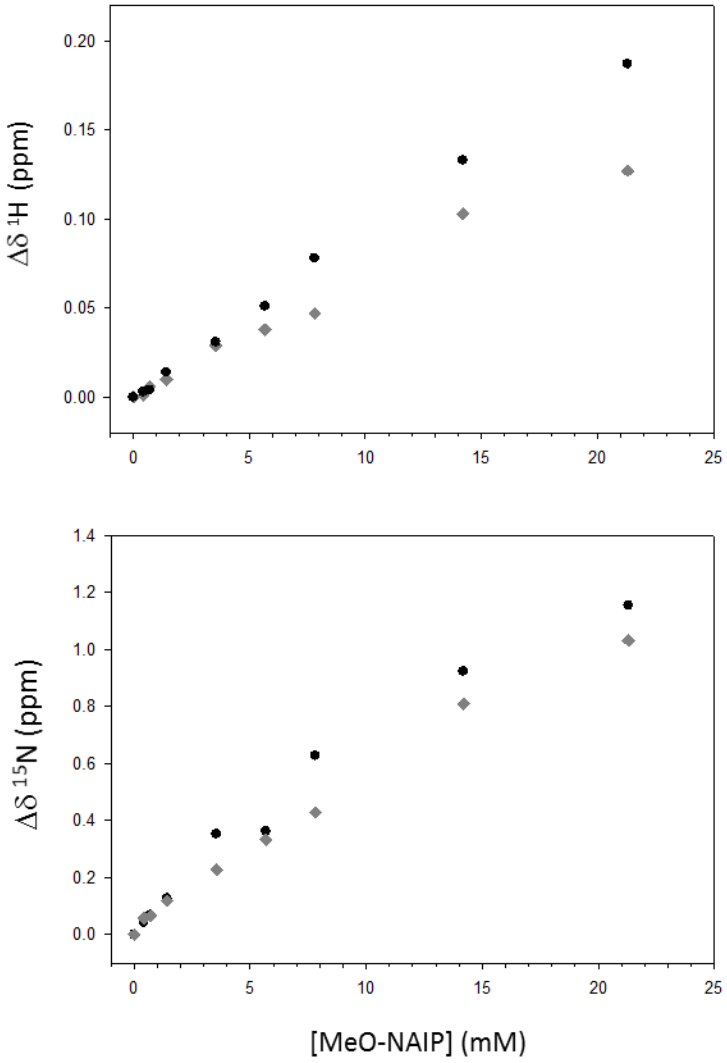


Figure S2. Chemical shift perturbation of ^{1}H (upper panel) and ^{15}N (lower panel) frequencies of residues T72 (circle) and C80 (diamond), as a function of **MeO-NAIP** concentration, as derived from a titration experiment of a 0.7 mM BABAP/SS sample. The titration curves did not reach a plateau even at protein:photoswitch ratios of 1:30, thus indicating a weak affinity (K_d in the mM range).

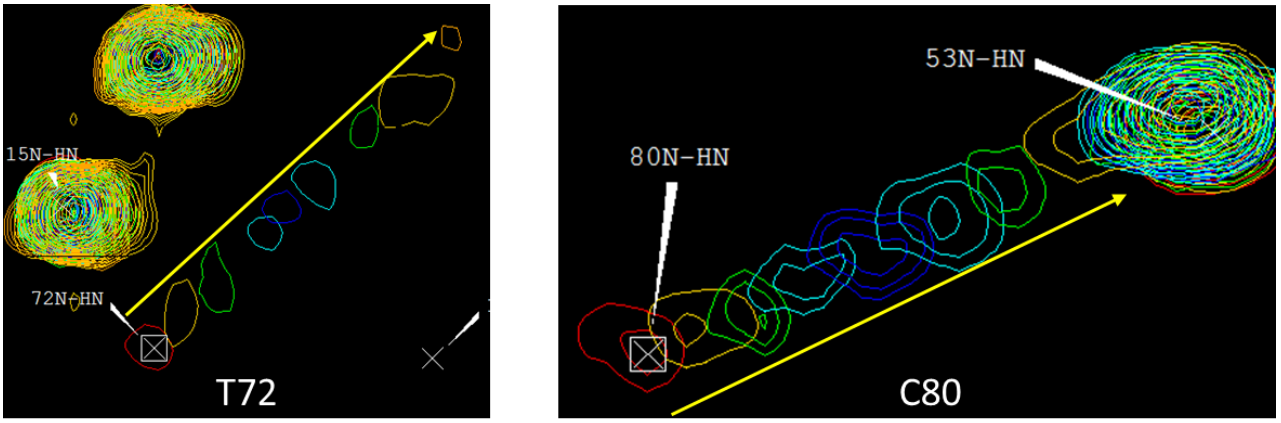


Figure S3. As an example, the superimposition of selected regions of $^{1}H-^{15}N$ HSQC spectra of ^{15}N-BABP-SS, recorded at increasing **MeO-NHIP:**BABP-SS ratios, showing chemical shift perturbation of T72 and C80 resonances. Arrows indicate the observed perturbations upon increasing ligand concentration from 0 to BABP-SS:**MeO-NHIP** 1:7.5 ratio.

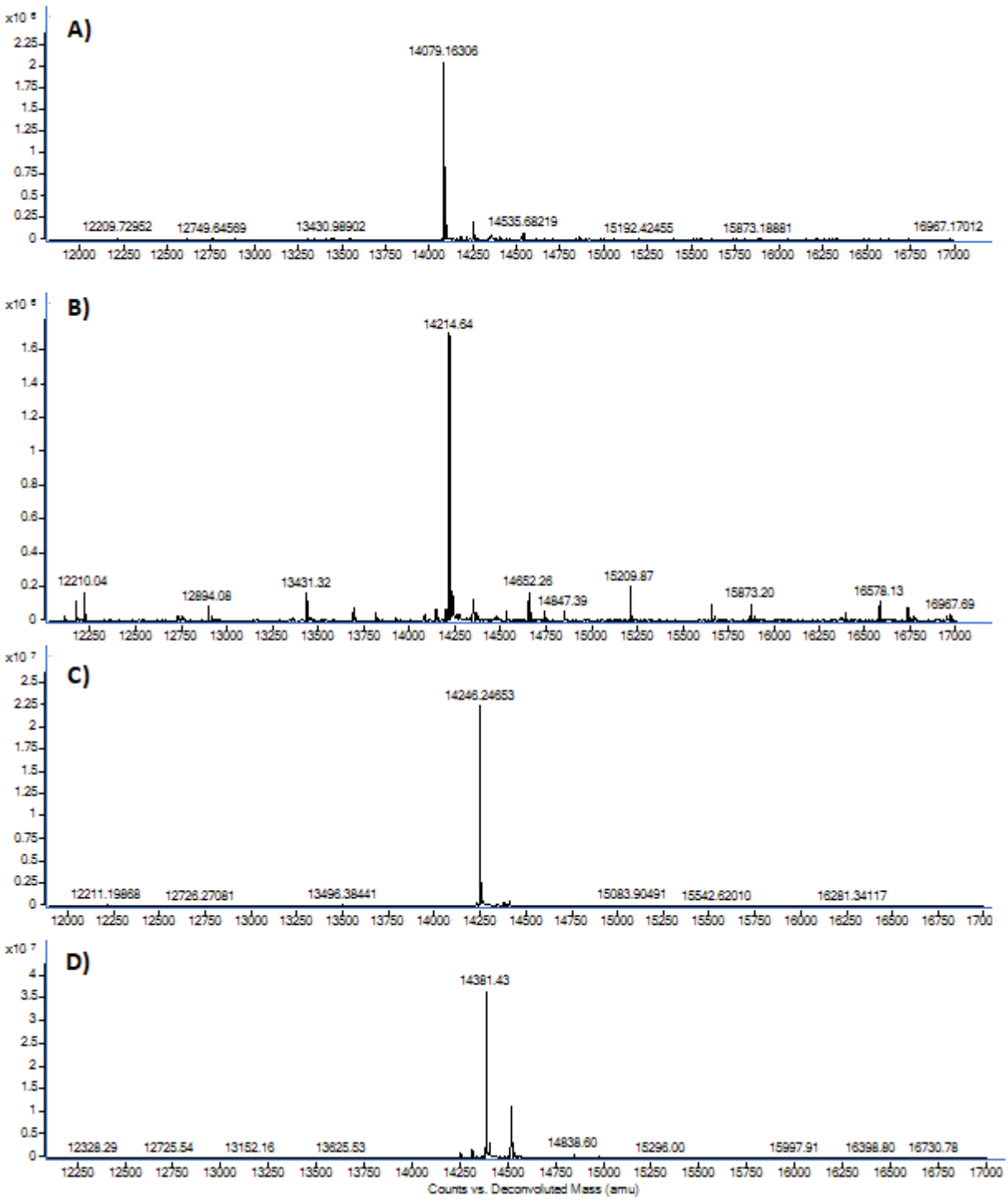


Figure S4. Mass spectra of **U-BABP** (A) and **^{15}N-BABP** (C) and of the new functional protein **U-BABP-Prop** (B) and **^{15}N-BABP-Prop** (D) obtained after reaction with propargyl-maleimide. Peaks at 14079 Da and 14246 Da correspond to the starting unlabeled and ^{15}N labeled BABP samples, respectively (A, C). The peaks at 14214 and 14381 Da are consistent with the addition of 135 Da, corresponding to the added N-propargyl-maleimide (B, D).

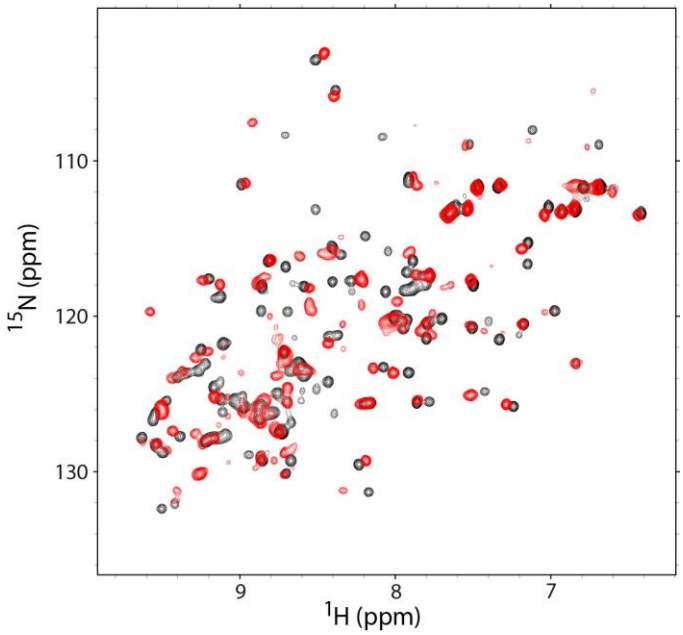


Figure S5. Superposition of the ^{1}H-^{15}N HSQC spectra of **^{15}N-BABP** (black) and **^{15}N-BABP-Prop** (red).

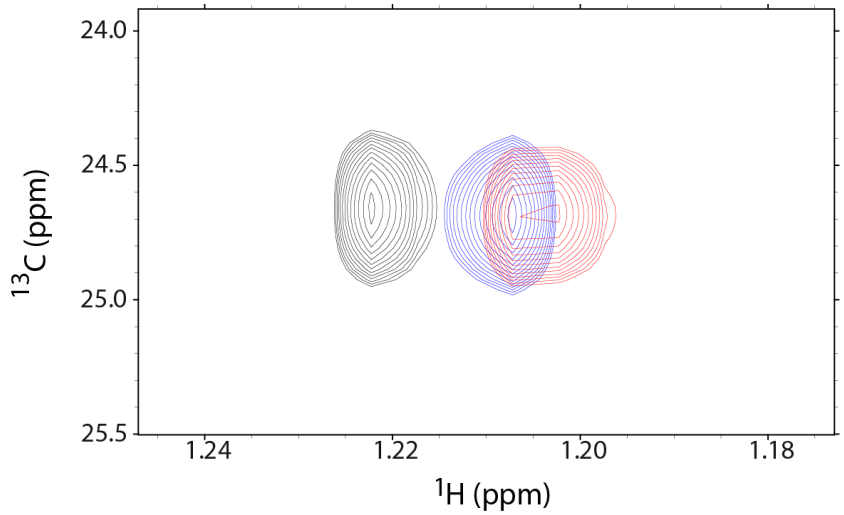


Figure S6. Superposition of ^{1}H-^{13}C HSQC spectra of **^{13C}NHIP-N$_3$** molecule, in the absence (black) and in the presence of the protein at **^{13C}NHIP-N$_3$**:U-BABP 1:1 (blue) and 1:2 (red) ratios. The observation that **^{13C}NHIP-N$_3$** methyl resonances exhibit a proton downfield shift upon U-BABP addition is consistent with a changed structural environment of this chemical group and demonstrates the ability of U-BABP to bind the photoswitch.

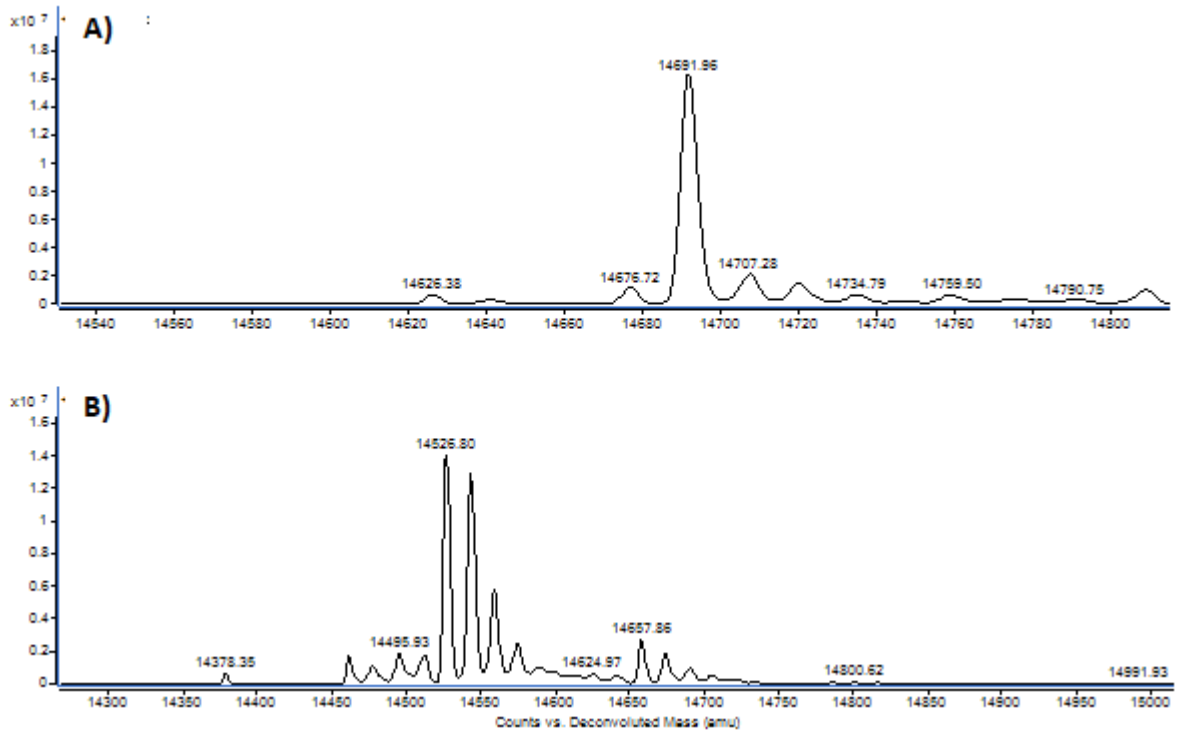


Figure S7. Mass spectra of the CuAAC reaction products. A) Main peak corresponding at **^{15}N-BABP-NHIP** (14691.9 Da) obtained in the reaction of **^{15}N-BABP-Prop** with **NHIP-N$_3$**. B) Main peak corresponding at **U-BABP-^{13C}NHIP** (14526.6 Da) obtained in the reaction of U-BABP-Prop with **^{13C}NHIP-N$_3$**. Minor peaks at (+16 and +32 Da) are due to sample oxidation during the storage period.

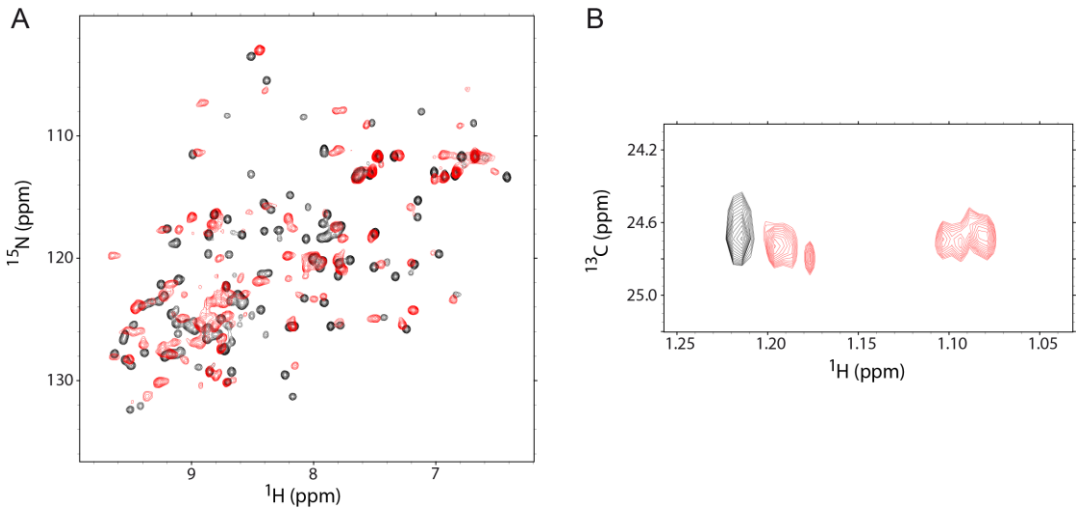
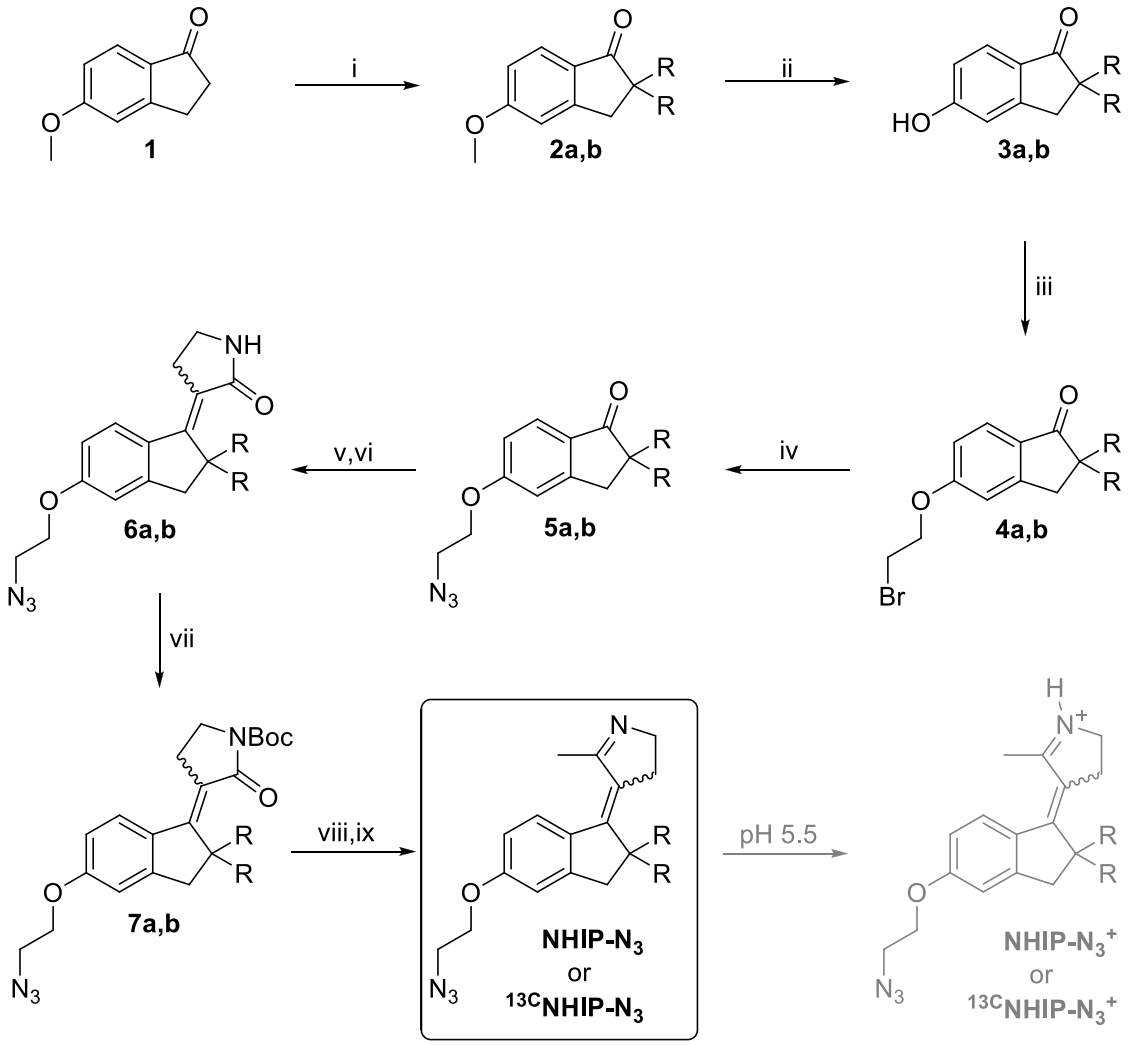


Figure S8. Superposition of ^{1}H-^{15}N HSQC spectra of ^{15}N-BABP (black) and **^{15}N-BABP-^{13C}NHIP** (red) (A). The protein appeared to be folded and stable for up to 15 days, indicating a tolerance to the mild reaction conditions. (B) Superposition of ^{1}H-^{13}C HSQC spectra of free **^{13C}NHIP-N$_3$** (black) and conjugated to BABP protein in **^{15}N-BABP-^{13C}NHIP** (red). Notice that the linkage of the sulfhydryl group to *N*-propargyl-maleimide gives rise to two diastereoisomers, which, in turn, may exhibit multiple conformers, consistently with the heterogeneity observed only in the bound **^{13C}NHIP** chromofore.

Synthesis of the NHIP-N_3 and ^{13C}NHIP-N_3 photoswitches.

The compounds **NHIP-N_3** and **^{13C}NHIP-N_3** were prepared as reported in Scheme S1. The commercially available starting material 5-methoxyindanone (**1**) was bismethylated at position C2 using methyl iodide or ^{13}C-enriched methyl iodide to yield the indanone derivatives **2a,b**. The insertion of ^{13}C enriched methyl groups in compound **2b** (and consequently in the final compound **^{13C}NHIP-N_3**) provides a suitable probe for NMR investigations. The methoxy group of **2a,b** was then deprotected with BBr_3 to get the phenolic compounds **3a,b**, which were reacted with 1,2-dibromoetane to obtain the intermediates **4a,b**. The treatment of **4a,b** with sodium azide afforded the indanone derivatives **5a,b,** which were reacted with *N*-Boc-pyrrolidinone to obtain E/Z isomeric mixtures of **6a,b** through dehydration of the aldol condensation intermediates. The amidic nitrogen of **6a,b** was then protected using di-*t*-butyl dicarbonate to obtain the intermediates **7a,b**. Finally, the target compounds **NHIP-N_3** and **^{13C}NHIP-N_3** were obtained using methylmagnesium bromide and subsequently TFA. The E/Z configuration was assigned on the basis of NOESY contacts, while the isomeric ratios were determined using ^{1}H NMR spectroscopy, on the basis of the integral ratio of the aromatic resonances. The thermostationary states of the amidic compounds **6a,b** and **7a,b** were found to consist mainly of the stereoisomer with the carboxylic group pointing in the direction of the two methyl groups on the C2 carbon. The amide group of the minor isomer points instead in the direction of the aromatic residue. Therefore, the main isomer of compounds **6a** and **7a** must have an E configuration, while the main isomer of compounds **6b** and **7b** assumes a Z configuration, due to the priority of the ^{13}C enriched methyl groups. In the final **NHIP-N_3** and **^{13C}NHIP-N_3** the main stereoisomers had the imine group pointing in the direction of the aromatic residue. Thus the Z configuration is the most stable for **NHIP-N_3** while the E configuration is the most stable for the ^{13}C enriched compound **^{13C}NHIP-N_3**. The target compounds possess a basic imine nitrogen to favor an easier entry into the hydrophobic protein cavity. However, they can be easily protonated at pH 5.5 providing *in vitro* the desired biomimetic photoswitch.



Scheme S1. Regents: i) CH_3I (for **2a**) or $^{13}CH_3I$ (for **2b**), *t*-BuOK, *t*-BuOH, Et_2O; ii) BBr_3, DCM; iii) 1,2-dibromoethane, K_2CO_3, CH_3CN/DMF (5:1); iv) NaN_3, CH_3CN/DMF (5:1); v) N-Boc-2-pyrrolidinone, LiHDMS, $BF_3(Et)_2O$, THF; vi) TFA, DCM; vii) di-*t*-butyl dicarbonate, DMPA, TEA, DCM; viii) CH_3MgBr, THF; ix) TFA, DCM. **Substituents:** R = CH_3 for **2-7a, NHIP-N_3**; R = $^{13}CH_3$ for **2-7b**, and **^{13C}NHIP-N_3**.

MATERIAL and METHODS

Materials. All chemicals used were of reagent grade. Yields refer to purified products and are not optimized. Merck silica gel 60 (230-400 mesh) were used for column chromatography. Merck TLC plates, silica gel 60 F_{254} were used for TLC. NMR spectra were recorded using a Bruker DRX-400 AVANCE spectrometer in the indicated solvents (TMS as internal standard): the values of the chemical shifts are expressed in ppm and the coupling constants (*J*) in Hz. An Agilent 1100 LC/MSD operating with a electrospray source was used in mass spectrometry experiments. The absorption spectra were recorded with a PerkinElmer Lambda 40 in the indicated solvent.

Photoswitch synthesis. Compound **MeO-NAIP** and **MeO-NHIP** (Scheme 1) were prepared as previously reported in reference 1 and 2, obtaining an isomeric excess of the Z isomer (85%).

2,3-Dihydro-5-methoxy-2,2-dimethyl-1*H*-inden-1-one (2a).[3]

To a mixture of compound **1** (1.50 g, 9.25 mmol) and *t*-BuOK (3.40 g, 30.3 mmol) in THF (20 mL) cooled (0 °C) was added dropwise methyl iodide (2.9 mL, 46.2 mmol). The resulting mixture was stirred at 0 °C in nitrogen atmosphere for 4 h. Then, a saturated NH_4Cl solution was added; the reaction mixture was extracted with dichloromethane, and the organic layer was dried over sodium sulfate and evaporated under reduced pressure. The residue was purified by flash chromatography (diethylether/petroleum ether 3:7) to give **2a** (1.50 g, 85%) as a colourless oil. ^{1}H NMR (400 MHz, $CDCl_3$): 1.16 (s, 6H), 2.89 (s, 2H), 3.81 (s, 3H), 6.82–6.90 (m, 2H), 7.67 (d, J = 8.4, 1H). MS (EI): *m/z* 191.1 ($M + H^+$).

2,3-Dihydro-5-methoxy-2,2-(^{13}C)dimethyl-1*H*-inden-1-one (2b).

To a mixture of compound **1** (0.50 g, 3.08 mmol) and *t*-BuOK (1.13 g, 10.1 mmol) in THF (20 mL) cooled (0 °C) was added dropwise ^{13}C-methyl iodide (0.97 mL, 15.4 mmol). The resulting mixture was stirred at 0 °C in nitrogen atmosphere for 4 h. Then, a saturated NH_4Cl solution was added; the reaction mixture was extracted with dichloromethane, and the organic layer was dried over sodium

sulfate and evaporated under reduced pressure. The residue was purified by flash chromatography (diethylether/petroleum ether 3:7) to give **2b** (0.50 g, 85%) as a colourless oil. ^{1}H NMR (400 MHz, $CDCl_3$): 1.21 (dd, *J* =127.0, 4.9, 6H), 2.93 (t, *J* = 3.4, 2H), 3.86 (s, 3H), 6.84 (s, 1H), 6.89 (dd, *J* = 8.5, 2.0, 1H), 7.67 (d, *J* = 8.5, 1H). MS (EI): *m/z* 193.1 ($M + H^+$).

2,3-Dihydro-5-hydroxy-2,2-dimethyl-1*H*-inden-1-one (3a).

To a solution of compound **2a** (1.0 g, 5.26 mmol) in dichloromethane (20 mL) cooled to 0 °C was added dropwise a solution (1 M in CH_2Cl_2) of BBr_3 (26.3 mL, 26.3 mmol). The resulting mixture was stirred at room temperature for 3 h. Then, a saturated $NaHCO_3$ solution was added until the gas evolution ceased. The reaction mixture was extracted with dichloromethane, and the organic layer was dried over sodium sulfate and evaporated under reduced pressure. Purification of residue by flash chromatography with petroleum ether-ethyl acetate (7:3) as the eluent gave pure compound **3a** (0.83 g, yield 90%) as a white solid melting at 131-132 °C. ^{1}H NMR (400 MHz, $CDCl_3$): 1.22 (s, 6H), 2.92 (s, 2H), 6.86-6.94 (m, 2H), 7.22 (s, 1H), 7.67 (d, *J* = 8.3, 1H). MS (EI): *m/z* 177.1 ($M + H^+$).

2,3-Dihydro-5-hydroxy-2,2-(^{13}C)dimethyl-1*H*-inden-1-one (3b).

To a solution of compound **2b** (0.50 g, 2.60 mmol) in dichloromethane (20 mL) cooled to 0 °C was added dropwise a solution (1 M in CH_2Cl_2) of BBr_3 (13.0 mL, 13.0 mmol). The resulting mixture was stirred at room temperature for 3 h. Then, a saturated $NaHCO_3$ solution was added until the gas evolution ceased. The reaction mixture was extracted with dichloromethane, and the organic layer was dried over sodium sulfate and evaporated under reduced pressure. Purification of residue by flash chromatography with petroleum ether-ethyl acetate (7:3) as the eluent gave pure compound **3b** (0.44 g, yield 95%) as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.22 (dd, *J* =127.2, 4.8, 6H), 2.92 (t, *J* = 3.4, 2H), 6.68 (br s, 1H), 6.84-6.88 (m, 2H), 7.67 (d, *J* = 8.1, 1H). MS (EI): *m/z* 179.1 ($M + H^+$).

5-(2-Bromoethoxy)-2,3-dihydro-2,2-dimethyl-1*H*-inden-1-one (4a).

Compound **4a** was prepared by optimizing the previously reported alkylation procedure.[4] In detail, to a mixture of compound **3a** (0.75 g, 4.26 mmol) and K_2CO_3 (5.88 g, 42.6 mmol) in CH_3CN/DMF (5:1, 30 mL) was added dropwise 1,2-dibromoethane (0.74 mL, 8.52 mmol). The reaction mixture was stirred overnight at room temperature under a nitrogen atmosphere. Then, the reaction mixture was diluted with dichloromethane (30 mL) and washed with a saturated NH_4Cl solution. The organic layer was dried over sodium sulfate and evaporated under reduced pressure. The residue was purified by flash chromatography with petroleum ether-ethyl acetate (8:2) as the eluent to obtain compound **4a** (0.66 g, yield 55%) as a off-white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.16 (s, 6H), 2.89 (s, 2H), 3.61 (t, *J* = 6.2, 2H), 4.31 (t, *J* = 6.2, 2H), 6.82 (s, 1H), 6.86 (dd, *J* = 8.5, 2.0, 1H), 7.63 (d, *J* = 8.5, 1H). MS (EI): *m/z* 304.8, 306.8 ($M + Na^+$).

5-(2-Bromoethoxy)-2,3-dihydro-2,2-(^{13}C)dimethyl-1*H*-inden-1-one (4b).

Compound **4b** was prepared by optimizing the previously reported alkylation procedure.[4] In detail, to a mixture of compound **3b** (0.40 g, 2.24 mmol) and K_2CO_3 (3.1 g, 22.4 mmol) in CH_3CN/DMF (5:1, 20 mL) was added dropwise 1,2-dibromoethane (0.39 mL, 4.48 mmol). The reaction mixture was stirred overnight at room temperature under nitrogen atmosphere, then diluted with dichloromethane (30 mL), and washed with a saturated NH_4Cl solution. The organic layer was dried over sodium sulfate and concentrated under reduced pressure. The residue was purified by flash chromatography with petroleum ether-ethyl acetate (8:2) as the eluent to obtain compound **4b** (0.36 g, yield 56%) as an off-white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.21 (dd, *J* =127.2, 4.8, 6H), 2.93 (t, *J* = 3.4, 2H), 3.65 (t, *J* = 6.4, 2H), 4.35 (t, *J* = 6.4, 2H), 6.85 (d, *J* = 1.6, 1H), 6.90 (dd, *J* = 8.4, 1.6, 1H), 7.69 (d, *J* = 8.4, 1H). MS (EI): *m/z* 306.9, 308.9 ($M + Na^+$).

5-(2-Azidoethoxy)-2,3-dihydro-2,2-dimethyl-1H-inden-1-one (5a).

Compound **5a** was prepared by optimizing the previously reported procedure.[5] In detail, a mixture

of compound **4a** (0.60 g, 2.12 mmol) in CH_3CN/DMF (5:1, 30 mL) containing NaN_3 (1.38 g, 21.2 mmol) was heated under reflux for 48 h and then concentrated under reduced pressure. The residue was partitioned between CH_2Cl_2 and a saturated NH_4Cl solution. The organic layer was dried over sodium sulfate and evaporated under reduced pressure. Purification of the residue by flash chromatography with petroleum ether-ethyl acetate (8:2) as the eluent gave compound **5a** (0.41 g, yield 78%) as an off-white solid melting at 65-66 °C. ^{1}H NMR (400 MHz, $CDCl_3$): 1.21 (s, 6H), 2.93 (s, 2H), 3.62 (t, *J* = 5.0, 2H), 4.20 (t, *J* = 5.0, 2H), 6.80–6.85 (m, 2H), 7.70 (d, *J* = 8.4, 1H). MS (EI): *m/z* 267.9 ($M + Na^+$).

5-(2-Azidoethoxy)-2,3-dihydro-2,2-(^{13}C)dimethyl-1H-inden-1-one (5b).

Compound **5b** was prepared by optimizing the previously reported procedure.[5] In detail, a A mixture of compound **4b** (0.35 g, 1.23 mmol) in CH_3CN/DMF (5:1, 30 mL) containing NaN_3 (0.80 g, 12.3 mmol) was heated under reflux for 48 h and then concentrated under reduced pressure. The residue was partitioned between CH_2Cl_2 and a saturated NH_4Cl solution. The organic layer was dried over sodium sulfate and evaporated under reduced pressure. Purification of the residue by flash chromatography with petroleum ether-ethyl acetate (8:2) as the eluent gave compound **5b** (0.25 g, yield 82%) as a an off-white solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.19 (dd, *J* =127.6, 4.8, 6H), 2.93 (t, *J* = 2.8, 2H), 3.62 (t, *J* = 5.0, 2H), 4.20 (t, *J* = 5.0, 2H), 6.86 (s, 1H), 6.91 (dd, *J* = 8.4, 1.6, 1H), 7.68 (d, *J* = 8.4, 1H). MS (EI): *m/z* 270.3 ($M + Na^+$).

3-(5-(2-Azidoethoxy)-2,3-dihydro-2,2-dimethyl-1H-inden-1-ylidene)pyrrolidin-2-one (6a).

To a solution of N-Boc-2-pyrrolidinone (0.27 g, 1.47 mmol) in anhydrous THF (10 mL), a 1 M solution of lithium hexamethyldisilazide (LiHMDS) in anhydrous THF (1.63 mL, 1.63 mmol) was added at −78 °C under a nitrogen atmosphere. After 1 h, a solution of compound **5a** (0.40 g, 1.63 mmol) and $BF_3{\cdot}Et_2O$ (0.21 mL, 1.63 mmol) in anhydrous THF (8.0 mL) was added dropwise. The reaction mixture was stirred at −78 °C for 3 h. Then, a saturated NH_4Cl solution was added, and the crude was extracted with CH_2Cl_2. The combined organic layers were dried over sodium sulfate and

concentrated under reduced pressure. The oily residue was dissolved in CH_2Cl_2 (10 mL), and trifluoroacetic acid (1.0 mL) was added. The resulting reaction mixture was stirred at room temperature for 30 min. Then, a saturated $NaHCO_3$ solution was added and the crude was extracted with CH_2Cl_2. The organic layer was dried over sodium sulfate and evaporated under reduced pressure. The residue was purified by flash chromatography with petroleum ether-ethyl acetate (1:1) as the eluent to give an E/Z mixture (isomeric ratio ca. 8:2) of compound **6a** (0.34 g, yield 74%) as a pale yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.25 (s, 6H Z), 1.54 (s, 6H E), 2.77 (s, 2H Z), 2.91 (s, 2H E), 3.03 (t, *J* = 6.6, 2H Z), 3.13 (t, *J* = 6.6, 2H E), 3.44 (m, 2H Z, 2H E), 3.56 (m, 2H Z, 2H E), 4.12 (m, 2H Z, 2H E), 5.19 (br s, 1H E), 5.28 (br s, 1H Z), 6.69 (s, 1H Z), 6.72 (d, *J* = 8.9, 1H Z), 6.79 (m, 2H E), 7.41 (d, *J* = 8.5, 1H E), 8.58 (d, *J* = 8.8, 1H Z). MS (EI): *m/z* 334.9 (M + Na^+).

3-(5-(2-Azidoethoxy)-2,3-dihydro-2,2-(^{13}C)dimethyl-1H-inden-1-ylidene)pyrrolidin-2-one (6b).

To a solution of N-Boc-2-pyrrolidinone (0.13 g, 0.70 mmol) dissolved in anhydrous THF (10 mL), a 1 M solution of lithium hexamethyldisilazide (LiHMDS) in anhydrous THF (0.80 mL, 0.80 mmol) was added at −78 °C under a nitrogen atmosphere. After 1 h, a solution of compound **5b** (0.25 g, 1.01 mmol) and $BF_3{\cdot}Et_2O$ (0.10 mL, 0.80 mmol) in anhydrous THF (5.0 mL) was added dropwise. The reaction mixture was stirred at −78 °C for 3 h. A saturated NH_4Cl solution was then added, and the crude was extracted with CH_2Cl_2. The combined organic layers were dried over sodium sulfate and concentrated under reduced pressure. The oily residue was dissolved in CH_2Cl_2 (10 mL), and trifluoroacetic acid (1.0 mL) was added. The resulting reaction mixture was stirred at room temperature for 30 min. A saturated $NaHCO_3$ solution was added and the crude was extracted with CH_2Cl_2. The organic layer was dried over sodium sulfate and evaporated under reduced pressure. The residue was purified by flash chromatography with petroleum ether-ethyl acetate (1:1) as the eluent to give an E/Z mixture (isomeric ratio ca. 2:8) of compound **6b** (0.17 g, yield 77%) as a pale yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.28 (dd, *J* =128.0, 4.3, 6H E), 1.55 (dd, *J* =128.0,

4.3, 6H Z), 2.79 (s, 2H E), 2.90 (s, 2H Z), 3.03 (t, *J* = 6.4, 2H E), 3.11 (t, *J* = 6.4, 2H Z), 3.35-3.45 (m, 2H E, 2H Z), 3.50-3.60 (m, 2H E, 2H Z), 4.08-4.16 (m, 2H E, 2H Z), 6.69-6.81 (m, 2H E, 2H Z), 7.01 (br s, 1H Z), 7.41 (d, *J* = 8.8, 2H Z), 7.52 (br s, 1H E), 8.58 (d, *J* = 8.8, 2H E). MS (EI): *m/z* 337.4 ($M + Na^+$).

_t_-Butyl 3-(5-(2-azidoethoxy)-2,3-dihydro-2,2-dimethyl-1*H*-inden-1-ylidene)-2-oxopyrrolidine-1-carboxylate (7a).

A mixture of compound **6a** (0.30 g, 0.96 mmol), di-*t*-butyl dicarbonate (0.42 g, 2.0 mmol), *N*,*N*-dimethylpyridin-4-amine (DMPA, 0.023 g, 0.19 mmol) in dichloromethane (30 mL) was stirred at room temperature under an argon atmosphere for 16 h. The reaction mixture was then washed with water and the organic layer was dried over sodium sulfate and concentrated under reduced pressure. The residue was purified by flash chromatography using petroleum ether-ethyl acetate (8:2) as the eluent to obtain the E/Z mixture (isomeric ratio ca. 8:2) of compound **7a** (0.38 g, 96%) as a pale yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.24 (s, 6H Z), 1.47-1.56 (m, 15H E, 9H Z), 2.77 (s, 2H Z), 2.86 (t, *J* = 7.1, 2H Z), 2.90 (s, 2H E), 2.96 (t, *J* = 7.2, 2H E), 3.55 (m, 2H Z, 2H E), 3.71 (m, 2H Z, 2H E), 4.12 (m, 2H Z, 2H E), 6.69 (m, 2H Z), 6.79 (m, 2H E), 7.39 (d, *J* = 8.6, 1H E), 8.55 (d, *J* = 8.8, 1H Z). MS (EI): *m/z* 434.9 ($M + Na^+$).

_t_-Butyl 3-(5-(2-azidoethoxy)-2,3-dihydro-2,2-(^{13}C)dimethyl-1*H*-inden-1-ylidene)-2-oxopyrrolidine-1-carboxylate (7b).

A mixture of compound **6b** (0.15 g, 0.48 mmol), di-*t*-butyl dicarbonate (0.21, 0.96 mmol), *N*,*N*-dimethylpyridin-4-amine (0.012 g, 0.096 mmol) in dichloromethane (30 mL) was stirred at room temperature under an argon atmosphere for 16 h. The reaction mixture was then washed with water and the organic layer was dried over sodium sulfate and concentrated under reduced pressure. The residue was purified by flash chromatography using petroleum ether-ethyl acetate (8:2) as eluent to obtain the E/Z mixture (isomeric ratio ca. 2:8) of compound **7b** (0.19, 96%) as pale yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): 1.28 (dd, *J* = 127.6, 4.3, 6H E), 1.52 (dd, *J* = 127.6, 4.3, 6H Z), 1.53

(m, 9H E, 9H Z), 2.79 (s, 2H E), 2.85-3.05 (m, 4H Z, 2H E), 3.53 (m, 2H E, 2H Z), 3.68-3.81 (m, 2H E, 2H Z), 4.12 (m, 2H E, 2H Z), 6.62-6.73 (m, 2H E), 6.75-6.83 (m, 2H Z), 7.41 (d, *J* = 8.8, 2H Z), 8.53 (d, *J* = 8.8, 2H E). MS (EI): *m/z* 437.0 ($M + Na^+$).

4-(5-(2-Azidoethoxy)-2,3-dihydro-2,2-dimethyl-1*H*-inden-1-ylidene)-3,4-dihydro-5-methyl-2H-pyrrole (NHIP-N_3).

To a solution of **7a** (0.35 g, 0.85 mmol) in dry THF (10 mL) cooled at -20 °C under a nitrogen atmosphere, a solution of methylmagnesium bromide (1 M in diethylether 8.5 mL, 8.5 mmol) was added, and resulting mixture was stirred at the same temperature for 3 h. Then, the excess of the Grignard reagent was destroyed by the dropwise addition of HCl until the gas evolution ceased. The mixture was dried over sodium sulfate and concentrated under reduced pressure. The oily residue was dissolved in dichloromethane (10 mL) and trifluoroacetic acid (1.0 mL) was added at 0 °C under a nitrogen atmosphere. The reaction mixture was stirred at 0 °C for 1 h, and the excess of acid was quenched by addition of solid sodium bicarbonate until the gas evolution ceased. The organic layer was washed with water, dried over sodium sulfate and concentrated under reduced pressure. The oily residue was purified by flash chromatography using ethyl acetate-methanol (9:1) as the eluent to obtain the E/Z mixture (isomeric ratio ca. 2:8) of **NHIP-N_3** (0.10 g, 38%) as a brown oil. ^{1}H NMR (400 MHz, $CDCl_3$): 1.32 (s, 6H Z), 1.48 (s, 6H E), 2.58 (s, 3H Z), 2.82 (s, 2H E), 2.92 (s, 2H Z), 3.02 (s, 2H E), 3.16 (t, *J* = 7.2, 2H Z), 3.35 (t, *J* = 7.2, 2H E), 3.62 (m, 2H Z, 2H E), 3.98 (t, *J* = 7.2, 2H E), 4.03 (t, *J* = 6.8 2H Z), 4.19 (m, 2H Z, 2H E), 6.80 (dd, *J* = 7.2, 2.4, 1H Z), 6.83-6.88, (m, 2H E, 1H Z), 7.27 (d, *J* = 8.6, 1H Z), 7.50 (d, *J* = 8.8, 1H E). MS (EI): *m/z* 311.1 ($M + H^+$).

4-(5-(2-Azidoethoxy)-2,3-dihydro-2,2-(^{13}C)dimethyl-1*H*-inden-1-ylidene)-5-methyl-3,4-dihydro-2H-pyrrole (^{13C}NHIP-N_3).

To a solution of **7b** (0.18 g, 0.43 mmol) in dry THF (10 mL) cooled at -20 °C under a nitrogen atmosphere a solution of methylmagnesium bromide (1 M in diethylether 4.3 mL, 4.3 mmol) was added, and resulting mixture was stirred for 3 h. Then, the excess of the Grignard reagent was

destroyed by the dropwise addition of HCl until the gas evolution ceased. The mixture was dried over sodium sulfate and concentrated under reduced pressure. The oily residue was dissolved in dichloromethane (10 mL) and trifluoroacetic acid (0.50 mL) was added at 0 °C under nitrogen atmosphere. The reaction mixture was stirred at 0 °C for 1 h. Then, the excess of acid was quenched by addition of solid sodium bicarbonate until the gas evolution ceased. The organic layer was washed with water, dried over sodium sulfate and concentrated under reduced pressure. The oily residue was purified by flash chromatography using ethyl acetate-methanol (9:1) as eluent to obtain the E/Z mixture (isomeric ratio ca. 8:2) of **^{13C}NHIP-N_3** (0.050 g, 37%) as brown oil. ^{1}H NMR (400 MHz, CD_3OD): 1.31 (dd, *J* = 126.8, 4.4, 6H E), 1.52 (dd, *J* = 127.6, 4.6, 6H Z), 2.40 (s, 3H E), 2.66 (s, 3H Z), 2.91 (t, *J* = 3.4, 2H E), 2.99 (t, *J* = 3.5, 2H Z), 3.11 (t, *J* = 6.5, 2H E), 3.34 (t, *J* = 6.6, 2H Z), 3.51-3.65 (m, 2H Z, 2H E), 3.83 (t, *J* = 5.7, 2H Z), 3.91 (t, *J* = 6.1, 2H E), 4.10 (m, 2H Z), 4.22 (t, *J* = 4.8, 2H E), 6.74-6.98, (m, 2H E, 2H Z), 7.40 (d, *J* = 8.6, 1H E), 7.57 (d, *J* = 8.7, 1H Z). MS (EI): *m/z* 313.0 ($M + H^+$).

***Protein expression and purification*.** Unlabeled and ^{15}N-labeled chicken liver bile acid binding proteins BABP and BABP-SS were expressed and purified as previously described.[6,7] The expression plasmids for BABP-SS and for BABP R120Q were obtained from that of wild-type BABP using the Quickchange (Stratagene, La Jolla, CA, USA) mutagenesis kit. The presence of the desired mutation was confirmed by plasmid sequencing. The presence of the disulfide bridge was confirmed by MS. Protein concentrations for sample preparations were determined spectrophotometrically.

Maleimide and N-Propargyl-maleimide labeling of ^{15}N-BABP.

The solvent accessibility of cysteine residue in ^{15}N-BABP under native conditions was assessed using maleimide labeling. Maleimide (0.75 M stock prepared in dimethyl formamide) was added in

about ten-fold molar excess to protein in 30 mM phosphate buffer at pH 6.5, and the reaction was carried out at 25 °C. The substitution was completed in 10 minutes. The reaction was quenched with dithiothreitol (DTT) in two-fold excess with respect to maleimide, and dialyzed against 30 mM phosphate buffer at pH 7.2. NMR and mass spectrometry were employed to check the reaction products. To increase cysteine accessibility the reaction was performed in the presence of 2 M urea. It is worth noting that, although BABP spends a significant amount of time in a barrel-open conformation to allow ligand binding and experiences greater flexibility at the ligand-entry portal with respect to CRABP,[8,9] a complete functionalization could be obtained only carrying out the reaction in the presence of 2M urea, at room temperature at protein:maleimide ratio of 1:10. We have previously reported[10] that BABP maintains its fold upon addition of up to 2 M urea. Propargyl maleimide labelling was performed in identical conditions.

CuAAC Cycloaddition.

The CuAAC reaction was performed using $CuSO_4$ and sodium ascorbate as a catalytic couple[11] under argon atmosphere employing the following molar ratio of the reagents with respect to BABP-Prop (**NHIP-N_3** or **^{13C}NHIP-N_3** 15:1, $CuSO_4$ 1.5:1, sodium ascorbate 100:1, aminoguanidinium chloride 100:1 and THPTA 15:1).[12]

The protein concentration was 30 μM and the final reaction volume was 3 mL. The reaction was quenched by EDTA addition (in three-fold excess with respect to $CuSO_4$) followed by dialysis against 30 mM phosphate buffer at pH 7.2. Stock solutions of $CuSO_4$, THPTA, sodium ascorbate, and aminoguanidinium chloride were prepared in water, while **NHIP-N_3** and **^{13C}NHIP-N_3** were dissolved in DMF. All solutions were kept under argon. The reagents were combined in a Schlenk tube in the following order: 1) biomolecule-alkyne in phosphate buffer + **NHIP-N_3** or **^{13C}NHIP-N_3**; 2) + a premixed solution of $CuSO_4$ and THPTA; 3) + aminoguanidinium chloride; 4) + sodium ascorbate. Specifically, aminoguanidine was employed to avoid side reactions between

dehydroascorbate and protein arginines, while tris(3-hydroxypropyltriazolyl-methyl)amine (THPTA) has served the dual purpose of protecting the biomolecule from hydrolysis by Cu(II) by-products, and intercepting the radicals and/or peroxides, derived from O_2/Cu/ascorbate, that oxidize histidine and other residues.

The reaction was conducted for one hour at 40 °C under stirring. Copper ions were then removed at room temperature by dialysis against buffer solutions containing the chelating agent ethylenediamine tetraacetic acid (EDTA) in three-fold excess with respect to $CuSO_4$. The final pH of the samples was corrected to 5.5 for maximizing the protonation of the imine. Both NMR and mass spectrometry have been employed to analyze each preparation and evaluate the presence of by-products.

Mass spectrometry

The LC/HRMS analysis were performed with an Agilent 6520 ESI-Q-TOF Nano-HPLC-CHIP cube® instrument (Agilent Technologies USA) using a linear gradient (0.4 mL/min) from 0% solvent A (97% water/3% acetonitrile/ +0.1% formic acid) to 90% solvent B (97% acetonitrile /3% water/ 0.1% formic acid in 15 min using a Zorbax C18 column (43 mm X 75 mm, 5 mm) equipped with a enrichment column (4 mm, 40 nL). The multicharge profile of the proteins were deconvoluted using the Masshunter Workstation software Qualitative analysis version B.06.00.

Photoisomerization.

The sample was irradiated using a Xenon lamp (450 W) equipped with a Gemini 180 monochromator for the selection of the excitation wavelength (400 nm). The acquisition of the first HSQC spectrum typically started 9 minutes after turning the lamp off and lasted 19 minutes.

NMR titration experiments.

BABP (0.69 mM) and BABP-SS (0.71 mM) were dissolved in a 30 mM sodium phosphate buffer, pH 7.2, containing 0.03% NaN_3 and 95%/5% H_2O/ D_2O. Unlabeled **MeO-NAIP** photoswitch, from 10 mM stock solution in phosphate buffer, was incrementally added from 0 to 2.1 mM and 1.8 mM to BABP-SS and BABP to the protein solution, respectively, in a 10-step titration.

BABP-SS (0.65 mM) was titrated with **U-MeO-NHIP (**MeO-NHIP 100 mM stock solution in tetrahydrofuran (THF), which was incrementally added from 0 to 2.6 mM, in a 12-step titration. For each titration step the desired amount of stock solution was deposited at the bottom of the vial, THF was evaporated by the insufflation of nitrogen gas and then the protein solution was added.

^{13C}NHIP-N_3 (0.28 mM) was titrated with 0.28, and 0.56 mM U-BABP-SS. Due to the limited protein solubility higher ligand:protein ratios could not be prepared.

U-BABP-^{13C}NHIP sample was treated with ^{15}N-GCDA, following the same procedure described for NHIP, starting from a 95 mM bile salt stock solution in THF.

NMR data were collected on a Bruker DMX NMR spectrometer operating at 600 MHz. Typically sweep widths of 14 and 40 ppm, 2048 × 128 data points, were used in proton and nitrogen dimensions, respectively, for the 2D ^{1}H−^{15}N HSQC spectra. Sweep widths of 10 and 20 ppm, with 4096 and 80 data points, in proton and carbon dimensions, respectively, were used for the 2D ^{1}H−^{13}C HSQC spectra. Experiments were performed at 298 K. Data were acquired and processed using Topspin 3.9 (Bruker Biospin).

Chemical shift perturbations were calculated as a weighted sum of ^{1}H and ^{15}N chemical shift changes.[11] The peaks in the HSQC spectra of the protein, recorded as a function of photoswitch additions, were integrated by Sparky software (T. D. Goddard and D. G. Kneller, SPARKY3, University of California, San Francisco). Normalized intensities were obtained dividing the intensities of each peak by the sum of all peak intensities. NMR titration experiments allowed for an estimate of a dissociation constant (K_d), as obtained from CSP data fitted to the following equation:

$$\Delta\delta_{obs} = (\frac{\delta_b - \delta_f}{2C_p})\delta\left[(C_P + C_L + K_d) - \sqrt{(C_P + C_L + K_d)^2 - 4C_P C_L}\right] \quad [1]$$

where δ_{obs}, δ_b, and δ_f are the chemical shifts of the protein signals in the observed, bound, and free states, and C_P and C_L are the total concentrations of the protein and the ligand, respectively.

Docking Simulations.

Docking of **MeO-NHIP** to BABP-SS was performed using the data driven HADDOCK2.2 docking software.[14,15] The ensemble of 10 holo structures, obtained by removing the two bile salts from the ternary complex (PDB ID: 2LFO), were used as starting structures. **MeO-NHIP** geometric coordinates and parameters were calculated using the PROGDG server.[16] Ambiguous interaction restraints (AIR) were introduced on the basis of NMR chemical shift perturbation (CSP) data from spectra of **BABP-SS**:**MeO-NHIP** at 1:7.5 molar ratio. Residues showing significant CSP upon binding were defined as active (F62, G65, T71, T72, C80, S93, S97), and their neighbours as "passive" (21, 49, 58, 60, 64, 70, 73, 74, 75, 76, 78, 82, 91, 96, 98, 111, 113), according to HADDOCK definition. The rigid body docking step was performed with 4000 structures, the best 200 of which were refined in the semi-flexible stage and subsequently in explicit water. Docking solutions were clustered with 0.8 Å cut off and a minimum of 10 structures per cluster, resulting in 3 clusters ranked according to the average HADDOCK score of their top 5 members.

References

(1) Rossi Paccani, R.; Donati, D.; Fusi, S.; Latterini, L.; Farina, G.; Zanirato, V.; Olivucci, M. Toward a stable alpha-cycloalkyl amino acid with a photoswitchable cationic side chain. *J. Org. Chem*. **2012**, *77* (4), 1738-1748.

(2) Zanirato, V.; Pollini, G. P.; De Risi, C.; Valente, F.; Melloni, A.; Fusi, S.; Barbetti, J.; Olivucci, M. Synthesis of biomimetic light-driven molecular switches via a cyclopropyl ring-opening/nitrilium ion ring-closing tandem reaction. *Tetrahedron* **2007**, *63* (23), 4975-4982.

(3) Paolino, M.; Gueye, M.; Pieri, E.; Manathunga, M.; Fusi, S.; Cappelli, A.; Latterini, L.; Pannacci, D.; Filatov, M.; Leonard, J.; Olivucci, M. Design, synthesis, and dynamics of a green fluorescent protein fluorophore mimic with an ultrafast Switching function. *J. Am. Chem. Soc.* **2016**, *138* (31), 9807-9825.

(4) Cappelli, A.; Paolino, M.; Anzini, P.; Giuliani, G.; Valenti, S.; Aggravi, M.; Donati, A.; Mendichi, R.; Zetta, L.; Boccia, A. C.; Bertini, F.; Samperi, F.; Battiato, S.; Paccagnini, E.; Vomero, S. Structure-property relationships in densely grafted π-stacked polymers *J. Polym. Sci., Part A: Polym. Chem.* **2010**, *48*, 2446–2461.

(5) Cappelli, A.; Paolino, M.; Grisci, G.; Giuliani, G.; Donati, A.; Mendichi, R.; Boccia, A. C.; Samperi, F.; Battiato, S.; Paccagnini, E.; Giacomello, E.; Sorrentino, V.; Licciardi, M.; Giammona, G.; Vomero, S. A click chemistry-based ‘‘grafting through’’ approach to the synthesis of a biorelevant polymer brush *Polym. Chem.* **2011**, *2*, 2518-2527.

(6) Ragona, L.; Catalano, M.; Luppi, M.; Cicero, D.; Eliseo, T.; Foote, J.; Fogolari, F.; Zetta, L.; Molinari, H. NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J. Biol. Chem*. **2006**, *281* (14), 9697-9709.

(7) Cogliati, C.; Tomaselli, S.; Assfalg, M.; Pedo, M.; Ferranti, P.; Zetta, L.; Molinari, H.; Ragona, L. Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins. *FEBS J.* **2009**, *276* (20), 6011-6023.

(8) Ferrolino, M. C.; Zhuravleva, A.; Budyak, I. L.; Krishnan, B.; Gierasch, L. M. Delicate balance between functionally required flexibility and aggregation risk in a beta-rich protein. *Biochemistry* **2013**, *52* (49), 8843-8854.

(9) Ragona, L.; Pagano, K.; Tomaselli, S.; Favretto, F.; Ceccon, A.; Zanzoni, S.; D'Onofrio, M.; Assfalg, M.; Molinari, H. The role of dynamics in modulating ligand exchange in intracellular lipid binding proteins. *Biochim. Biophys. Acta* **2014**, *1844* (7), 1268-1278.

(10) D'Onofrio, M.; Ragona, L.; Fessas, D.; Signorelli, M.; Ugolini, R.; Pedo, M.; Assfalg, M.; Molinari, H. NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviors. *Arch. Biochem. Biophys.* **2009**, *481* (1), 21-29.

(11) Cappelli, A.; Paolino, M.; Reale, A.; Razzano, V.; Grisci, G.; Giuliani, G.; Donati, A.; Bonechi, C.; Lamponi, S.; Mendichi, R.; Battiato, S.; Samperi, F.; Makovec, F.; Licciardi, M.; Depau, L.; Botta, C. Hyaluronan-based graft copolymers bearing aggregation-induced emission fluorogens. *RSC Advances* **2018**, *8* (11), 5864-5881.

(12) Presolski, S. I.; Hong, V. P.; Finn, M. G., Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Curr. Protoc. Chem. Biol.* **2011**, *3* (4), 153-162.

(13) Mulder, F. A.; Schipper, D.; Bott, R.; Boelens, R., Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. *J. Mol. Biol.* **1999**, *292* (1), 111-123.

(14) de Vries, S. J.; van Dijk, A. D.; Krzeminski, M.; van Dijk, M.; Thureau, A.; Hsu, V.; Wassenaar, T.; Bonvin, A. M., HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* **2007**, *69* (4), 726-733.

(15) Dominguez, C.; Boelens, R.; Bonvin, A. M., HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731-1737.

(16) Schuttelkopf, A. W.; van Aalten, D. M., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1355-1363.