Preparation of unnatural N-to-N and C-to-C protein fusions

Witte, Martin; Cragnolini, Juan J.; Dougan, Stephanie K.; Yoder, Nicholas C.; Popp, Maximilian W.; Ploegh, Hidde L.; Petsko, Gregory A.

Published in:
Proceedings of the National Academy of Sciences of the United States of America

DOI:
10.1073/pnas.1205427109

Publication date:
2012

Supporting Information

Witte et al. 10.1073/pnas.1205427109

SI Methods

General Experimental. All chemicals were from commercial sources and were used as received. Fmoc-Lys(Mtt)-OH, Fmoc-Gly-OH, Fmoc-Thr-OH, Fmoc-Pro-OH, Fmoc-Glu-OH, Fmoc-Leu-OH, *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium hexafluorophosphate (HBTU), and benzotriazol-1-yl oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) were purchased from EMD Biosciences/Novabiochem. Rink amide resin was purchased from Advanced Chemtech. Cyclooctyne reagents were purchased from Click Chemistry Tools. Water used in biological procedures or as a reaction solvent was purified using a MilliQ purification system (Millipore). DriSolv anhydrous CH2Cl2, DriSolv anhydrous MeOH, and DriSolv anhydrous DMF were purchased from EMD Chemicals. Redistilled, anhydrous *N*,*N*'-diisopropylethylamine (DiPEA), trifluoroacetic acid (TFA), triisopropylsilane (TIS), and *N*-methylpyrrolidone (NMP) were obtained from Sigma-Aldrich.

Mass Spectrometry. Liquid chromatography (LC)-ESI (electrospray ionization)-MS analysis was performed using a Micromass LCT mass spectrometer (Micromass MS Technologies) and a Paradigm MG4 HPLC system equipped with an HTC PAL autosampler (Michrom BioResources) and a Waters Symmetry 5-μm C8 column (2.1 × 50 mm, MeCN:H_2O (0.1% formic acid) gradient mobile phase, 150 μL/min).

HPLC/FPLC. HPLC purifications were achieved using an Agilent 1100 Series HPLC system equipped with a Waters Delta Pak 15-μm, 100-Å C18 column (7.8 × 300 mm, MeCN:H_2O gradient mobile phase, 3 mL/min) as indicated below. Size exclusion and cation exchange chromatography were performed on a Pharmacia AKTA Purifier system equipped with a HiLoad 16/60 Superdex 75 column (Amersham) or a Mono S 5/50 GL column (Amersham), respectively.

UV-Vis Spectroscopy. UV-vis spectroscopy was performed on a Nanodrop ND-1000 spectrophotometer (Thermo Scientific).

In-Gel Fluorescence. Fluorescent gel images were obtained using a Typhoon 9200 Variable Mode Imager (GE Healthcare).

General Procedure for the Solid-Phase Peptide Synthesis of the Probes. Rink-amide resin was solvated in NMP and after removal of the Fmoc group by treating the resin with 20% (vol/vol) piperidine in NMP, the resin was loaded and elongated using consecutive steps: (*i*) The resin was washed with NMP (three times), CH_2Cl_2 (three times), and NMP. (*ii*) Fmoc-protected amino acids were condensed under the agency of HOBt [3 equivalents (equiv.)], PyBOP (3 equiv.), and DiPEA (6 equiv.). (*iii*) The resin was washed again using the same conditions as in step *i*. (*iv*) The coupling was monitored using a Kaiser test and, if complete, (*v*) the Fmoc-protective group was removed using 20% (vol/vol) piperidine in NMP.

In the final step, the peptides were cleaved off resin by agitating the resin in the presence of 95% TFA, 2.5% TIS, 2.5% H_2O (vol/vol/vol) for 3 h. Ice-cold Et_2O was added to the cleavage solution and the formed precipitate was collected by centrifugation of the solution for 30 min at 4 °C. The crude pellet was purified by reverse-phase HPLC purification [buffers used: A, H_2O; B, MeCN; C, 10% (vol/vol) TFA in H_2O].

N-Terminal Probes. ***Azidohexanoic acid-LPETGG-CONH₂ (1).*** Rink amide resin (60 μmol) was loaded with Fmoc-Glyc-OH, elongated with the appropriately protected amino acids, and cleaved off the resin as described in the general method. For the final coupling azidohexanoic acid was used. RP-HPLC [26–35% (vol/vol) B in 12 min (3 column volumes)] gave the title compound (9.5 mg, 13 μmol, 13%) as a white solid.

LC/MS: R_t = 6.34 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 711.1 $[M+H]^+$. ^1H NMR (400 MHz, $CDCl_3$) δ ppm 4.65 (dd, J = 10.0, 4.4 Hz, 1H), 4.42 (dd, J = 8.4, 6.0 Hz, 1H), 4.35 (dd, J = 9.2, 5.2 Hz, 1H), 4.30–4.24 (m, 2H), 4.00 (s, 2H), 3.96 (s, 2H), 3.91–3.84 (m, 4H), 3.70–3.64 (m, 1H), 2.48 (t, J = 7.2 Hz, 2H), 2.26 (t, J = 7.6), 2.24–1.96 (m, 6H), 1.78–1.70 (m, 1H), 1.69–1.56 (m, 7H), 1.44–1.38 (m, 3H), 1.21 (d, J 6.4 Hz, 3H), 0.97 (t, 6.4 Hz, 6H).

Aza-dibenzocyclooctyne (DIBAC)-LPETGG-CONH₂ (2). Rink amide resin was loaded with Fmoc-Glyc-OH, elongated with the appropriately protected amino acids, and cleaved off the resin as described in the general method. Precipitation from Et_2O afforded crude H_2N-LPETGG-$CONH_2$ (17.9 mg, 31.3 μmol), which was dissolved in DMF (0.5 mL). DIBAC-OSu (14 mg, 20 μmol) was added and the reaction was stirred overnight. The solution was diluted before being purified by RP-HPLC [25–34% (vol/vol) B in 12 min (3 column volumes)], which gave the title compound (13.1 mg, 12.3 μmol, 39%) as an off-white solid.

LC/MS: R_t = 9.42 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z =1,066.14 $[M+H]^+$. ^1H NMR (400 MHz, $CDCl_3$) δ ppm 7.65 (dd, J =13.2, 7.2 Hz, 1H), 7.46–7.28 (m, 7H), 5.05 (d, J = 14.4 Hz, 1H), 4.72–4.65 (m, 1H), 4.60–4.50 (m, 1H), 4.48–4.38 (m, 2H), 4.36 (d, J = 4 Hz, 1H), 4.26–4.23 (m, 1H), 4.04–3.87 (m, 5H), 3.73–3.62 (m, 2H), 3.52–3.38 (m, 1H), 3.10–2.92 (m, 1H), 2.82–2.67 (m, 1H), 2.56–2.39 (m, 5H), 2.34–2.09 (m, 6H), 2.07–1.98 (m, 4H), 1.94–1.85 (m, 2H), 1.72–1.52 (m, 6H), 1.50–1.40 (m, 1H), 1.20 (d, J = 6.0 Hz, 3H), 0.98–0.90 (m, 6H).

C-Terminal Probes. ***H₂N-GGGK(N₃)K(TAMRA)-CONH₂ (3).*** Rink amide resin (60 μmol) was loaded with Fmoc-Lys(Mtt)-OH and elongated with Fmoc-Azidolysine-OH and Fmoc-GGG-OH as described in the general method. After washing the resin with CH_2Cl_2, the Mtt protective group was removed by treating the resin twice with 1% TFA, 1% TIS in CH_2Cl_2 (vol/vol/vol) for 30 min (or until the yellow color completely disappeared). The resin was washed with CH_2Cl_2 (five times), NMP (five times), and NMP containing DiPEA (43.5 μL, 250 μmol, 5 equiv.). 5(6)-Carboxytetramethylrhodamine (77 mg, 180 μmol, 3 equiv.) was condensed using PyBOP (94 mg, 180 μmol, 3 equiv.) and DiPEA (65 μL, 370 μmol, 6 equiv.). After 16 h shaking, the Kaiser test showed complete conversion. The N-terminal Fmoc group was removed and the peptide was cleaved off resin as described in the general method. Reverse-phase HPLC purification [25–34% (vol/vol) B in 12 min (3 CV)] gave the title compound (41.4 mg, 50.5 μmol, 81%) as a purple solid.

LC/MS: R_t = 5.50 and 6.10 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 883.3 $[M+H]^+$. ^1H NMR (400 MHz, $CDCl_3$) δ ppm 8.78 (d, J = 1.6 Hz, 1H), 8.28 (dd, J = 7.6, 1.6 Hz, 1H), 7.53 (d, J = 8.0 Hz), 7.14 (d, J = 9.6 Hz, 2H), 7.06 (dd, J = 9.6, 2.4 Hz, 2H), 6.98 (d, J = 2.4 Hz, 2H), 4.34 (dd, J = 9.2, 5.2 Hz, 2H), 3.98 (d, 14.8 Hz, 1H), 3.96 (s, 2H), 3.82 (d, 18.4 Hz, 1H), 3.80 (s, 2H), 3.54–3.46 (m, 2H), 3.32–3.28 (m, 14H), 1.94–1.45 (m, 12H).

H₂N-GGGC(DIBAC)-CONH₂ (4). Rink amide resin (167 mg, 100 μmol) was loaded with Fmoc-Cys(Trt)-OH, elongated with Fmoc-GGG-OH, and cleaved off the resin as described in the general method, affording crude tetrapeptide, H_2N-GGGC-$CONH_2$, in

quantitative yield. This peptide (38 mg, 83 μmol, 2 equiv.) was dissolved in PBS (0.25 mL) and to this was added DIBAC-maleimide (17 mg, 40 μmol, 1 equiv.) in DMF (0.25 mL). The reaction was stirred overnight, acidified with TFA, and purified by RP-HPLC [20–35% (vol/vol) B in 20 min (5 CV)], giving the title compound (15.3 mg, 22 μmol, 27%) as a white solid.

LC/MS: R_t = 6.90 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 719.3 $[M+H]^+$. ^{1}H NMR (400 MHz, M) δ ppm 7.66 (d, J = 7.2 Hz, 1H), 7.55–7.51 (m, 1H), 7.48–7.45 (m, 3H), 7.38 (dt, J =7.6, 1.4 Hz, 1H), 7.37–7.33 (m, 1H), 7.28 (d, J = 7.2 Hz, 1H), 5.14 (d, J =14 Hz, 1H), 4.69–4.64 (m, 1H), 4.01–3.85 (m, 6H), 3.77 (d, J = 4.8 Hz, 1H) 3.73 (s, 1H), 3.70 (s, 1H), 3.67–3.63 (m, 2H), 3.39 (ddd, J = 14.0, 5.2, 2.8 Hz, 1H), 3.27–3.05 (m, 5H), 2.97 (ddd, J =14, 8.4, 5.2 Hz, 1H), 2.48–2.41 (m, 3H), 2.33–2.87 (m, 2H) 2.08–1.99 (m, 1H).

H_2N-GGGK(Azidohexanoic acid)-$CONH_2$ (5). Rink amide resin (100 mg, 50 μmol) was loaded with Fmoc-Lys(Mtt)-OH and elongated with Fmoc-GGG-OH as described in the general method. After washing the resin with CH_2Cl_2, the Mtt protective group was removed by treating the resin twice with 1% TFA, 1% TIS in CH_2Cl_2 (vol/vol/vol) for 30 min (or until the yellow color completely disappeared). The resin was washed with CH_2Cl_2 (five times), NMP (five times), and NMP containing DiPEA (43.5 μL, 250 μmol, 5 equiv.). Azidohexanoic acid (31 mg, 200 μmol, 4 equiv.) was condensed using PyBOP (104 mg, 200 μmol, 4 equiv.) and DiPEA (70 μL, 400 μmol, 8 equiv.). After 2 h shaking, the Kaiser test showed complete conversion. The N-terminal Fmoc group was removed and the peptide was cleaved off resin as described in the general method. Reverse-phase HPLC purification [15–24% (vol/vol) B in 12 min (3 CV)] gave the title compound (15.4 mg, 33 μmol, 67%) as an off-white solid.

LC/MS: R_t = 2.77 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 456.3 $[M+H]^+$. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 4.35 (dd, J = 9.2, 4.8 Hz, 1H), 3.98 (d, J = 16.8 Hz, 1H), 3.97 (s, 2H), 3.86 (d, J = 16.8 Hz, 1H), 3.78 (s, 2H), 3.29 (t, J = 6.8 Hz, 2H), 3.17 (dt, J =6.8, 2.0 Hz, 2H), 2.20 (t, J = 7.2 Hz, 2H), 1.86–1.81 (m, 1H), 1.73 (ddd, J = 18.4, 9.4, 5.0 Hz, 1H), 1.67–1.57 (m, 4H), 1.55–1.47 (m, 2H), 1.43–1.38 (m, 4H).

Cloning and Expression of Proteins. Ubiquitin N-terminally fused to N-terminal his tag followed by a thrombin cleavage site (MGSSHHHHHHSSGLVPRGGGSH) was cloned into a pET28 vector. The vector was transformed into BL21(DE3)pLysS A starter culture was grown in LB. The expression culture was started at OD_{600} of 0.2. When the culture reached an OD_{600} of 0.6–0.8, the bacteria were induced with 1 mM IPTG and cultured for 6 h at 37 °C. The bacteria were collected by centrifugation at 6,000 × g for 15 min and the pellet was resuspended in lysis buffer [20 mM Tris, pH 8.0, 150 mM NaCl, 10 mM imidazole, 50 μg/mL DNaseI (Roche) and 1 tablet/25 mL complete protease inhibitor (Roche)] and sonificated. The lysate was clarified by centrifugation. Soluble protein was purified by Ni-NTA (Qiagen). The thrombin sequence was removed using a Thrombin CleanCleave kit (Sigma-Aldrich).

Ubiquitin (1–75) N-terminally fused to a thrombin cleavage site followed by GGG (MGSSHHHHHHSSGLVPRGGG) and C-terminally fused to intein was cloned into pTYB2. The vector was transformed into BL21(DE3)pLysS. The ubiquitin–intein constructed was expressed, purified, and converted into the UbVME adduct as previously described for HA-tagged UbVME. The Thrombin CleanCleave kit was used to expose the N-terminal glycine residues.

A synthetic version of anti-GFP containing a C-terminal LPETGG was subcloned into a pET28A+ vector. The vector was transformed into *Escherichia coli* BL21(DE3)pLysS. A starter culture (250 mL, LB medium) was grown to saturation overnight at 37 °C. An expression culture, started at OD_{600} of 0.2, [2 L Yeast/Tryptone (2YT) medium] was grown at 37 °C until the OD_{600} = 0.6. The bacteria were induced with IPTG (1 mM) and grown for 16 h at 25 °C. Bacteria were collected by centrifugation at 6,000 × g for 15 min and they were lysed by sonification in lysis buffer [20 mM Tris, pH 8.0, 150 mM NaCl, 10 mM imidazole, 50 μg/mL DNaseI (Roche), and 1 tablet/25 mL complete protease inhibitor (Roche)]. The lysate was clarified by centrifugation. Soluble protein was purified by Ni-NTA (Qiagen) followed by size-exclusion chromatography on a Superdex 75.

VHH7 containing a C-terminal LPETGGHHHHHH was cloned into a pHEN vector (kindly provided by the Steyaert Laboratory, Structural Biology Brussels, Vrije Universiteit Brussel, Brussels) N-terminally preceded by the pelB leader sequence. The vector was transformed into *E. coli* WK6. A starter culture (250 mL) was grown in 2YT to saturation overnight at 37 °C. The expression culture was started at OD_{600} of 0.2. When the culture reached an OD_{600} of 0.7, the expression of protein was induced by the addition of 1 mM IPTG. The bacteria were cultured overnight at 37 °C. The periplasmic fraction was isolated by incubating the bacterial pellet in 1 vol of 1× TES buffer (0.2 M Tris, 0.65 mM EDTA, 0.5 M sucrose) for 1 h at 4 °C and subsequently 2 vol of 0.25× TES buffer were added. The resulting suspension was stirred overnight at 4 °C. The solution was clarified by centrifugation and concentrated using amicon ultra 3K spin concentrators and the proteins were subjected to Ni-NTA. The proteins were further purified by size-exclusion chromatography.

Human interleukin-2 lacking the leader sequence and fused at the C terminus to the sequence GGLPETGGHHHHHH was cloned into the pET28a$^+$ vector (Novagen). The vector was transformed into *E. coli* BL21(DE3)pLysS and a starter culture was grown overnight at 37 °C. The starter culture was added to the expression culture (3 L, 2YT) and grown until the OD_{600} reached 0.6. To induce expression, 1 mM IPTG (final concentration) was added and the bacteria were grown at 37 °C for 4 h. The bacteria were collected by centrifugation at 6,000 × g for 15 min at 4 °C. The bacteria were lysed by sonification in lysis buffer [50 mM Tris, pH 7.4, 150 mM NaCl, 50 μg/mL DNaseI (Roche), and 1 tablet/25 mL complete protease inhibitor (Roche)]. The inclusion bodies were collected by centrifugation (12,000 × g for 15 min at 4 °C). Before being dissolved in 50 mM Tris, pH 7.4, 150 mM NaCl, 6 M guanidinium, the inclusions were first washed by resuspending the pellet in lysis buffer (one time), l-butanol (one time), and 50 mM Tris, pH 7.4, 150 mM NaCl, and 1 M guanidinium HCl (two times) and subsequent centrifugation.

The unfolded protein (6 mg/mL, 0.7 mL) was pretreated with tris(2-carboxyethyl)phosphine (1 mM) and subsequently added (0.1 mL/h) to refolding buffer (200 mL, 50 mM Tris, pH 7.4, 150 mM NaCl, 10% (vol/vol) glycerol, 5 mM glutathione, 0.5 mM oxidized glutathione) at 25 °C. The reaction was stirred for 2 d, concentrated on a Ni-NTA column, and subsequently purified by size-exclusion chromatography.

Sortase A of *Staphylococcus aureus* and human IFNα2a were expressed and purified as previously described (1, 2).

Modification of Ubiquitin with N_3-LPETGG (1) and DIBAC-LPETGG (2). Ubiquitin was modified with **1** and **2** as described for UbVME. N_3-Ub: R_t = 7.17 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 9,542 $(M+H)^+$. DIBAC-Ub: R_t = 7.37 min; linear gradient 5→45% (vol/vol) B in 10 min. ESI/MS: m/z = 9,898 $(M+H)^+$.

Dimerization of Ubiquitin. Azido-modified ubiquitin (5 μL, 4 μg/μL) and DIBAC-modified ubiquitin (8 μL, 2.5 μg/μL) were mixed (final concentration of the proteins 170 μM) and incubated for 0.5–7 h at 37 °C. The conversion to the dimerized product was analyzed using gel electrophoresis.

1. Popp MW, Dougan SK, Chuang T-Y, Spooner E, Ploegh HL (2011) Sortase-catalyzed transformations that improve the properties of cytokines. *Proc Natl Acad Sci USA* 108: 3169–3174.

2. Popp MW, Antos JM, Ploegh HL (2009) *Current Protocols in Protein Science*, eds Coligan JE, Dunn BM, Speicher DW, Wingfield PT (Wiley, Hoboken, NJ).

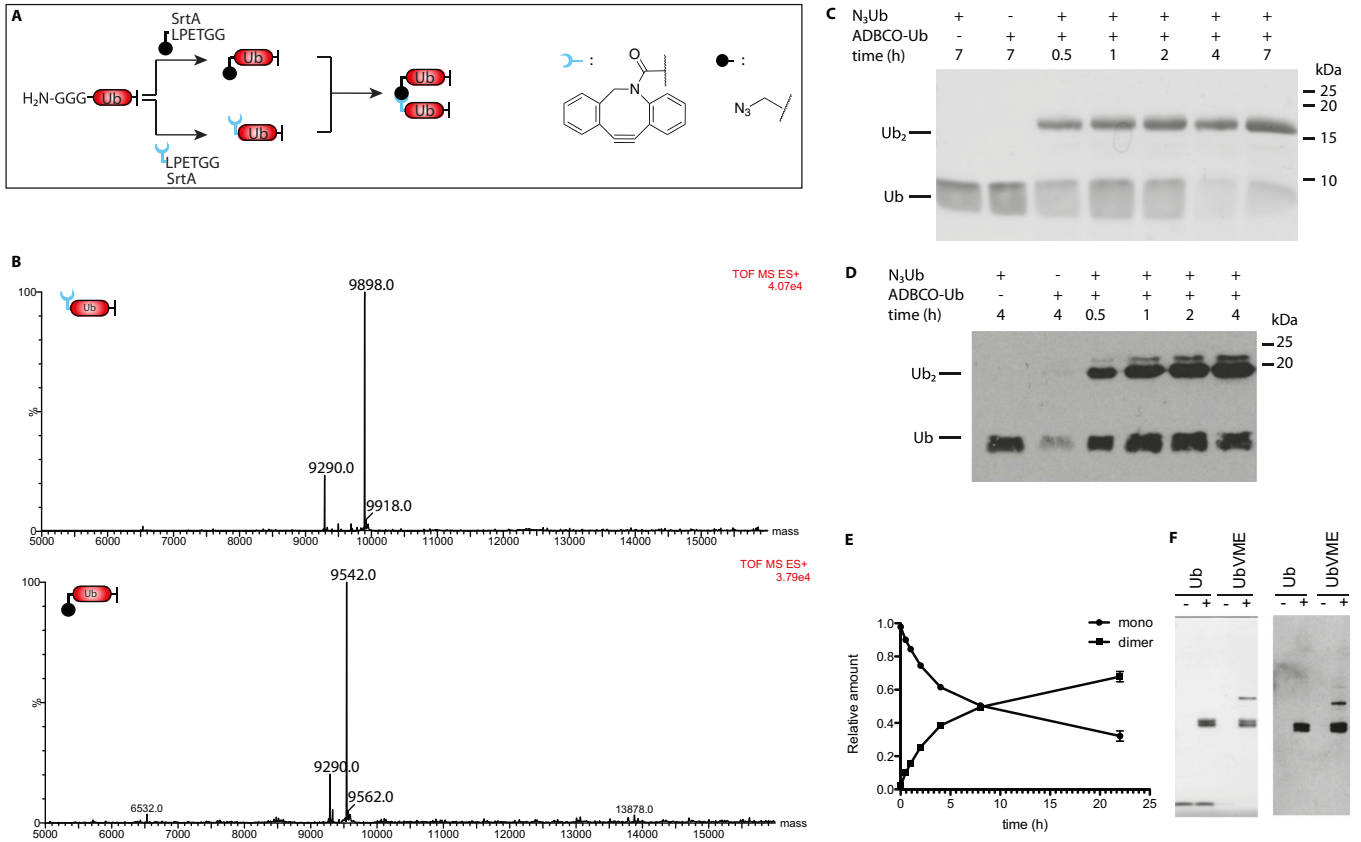

Fig. S1. Requirements for dimerization of ubiquitin. (*A*) Schematic approach. (*B*) Ubiquitin is sortagged with **1** or **2** for 3 h and analyzed with LC/MS. (*C* and *D*) Dimerization of ubiquitin. Azido-modified ubiquitin (2 nmol) is incubated with an equimolar amount of cyclooctyne-equipped ubiquitin in 13 μL H_2O. The dimer was resolved on 15% SDS/PAGE and the proteins were detected by Coomassie staining (*C*) and immunoblotting (*D*) for ubiquitin. (*E*) Azido-ubiquitin (0.1 nmol) incubated with DIBAC-ubiquitin (0.1 nmol) for the indicated time was resolved on a Tris-tricine gel and Coomassie stained, and the resulting protein was quantified by ImageJ. The relative amount of monomer and dimer per lane was determined as follows: relative amount of dimer = intensity of dimer/ total intensity; relative amount of monomer = intensity of monomer/ total intensity. (*F*) Labeling of UCHL3 with either ubiquitin or UbVME. (*Left*) Coomassie-stained gel; (*Right*) immunoblotting for the his_6 tag.

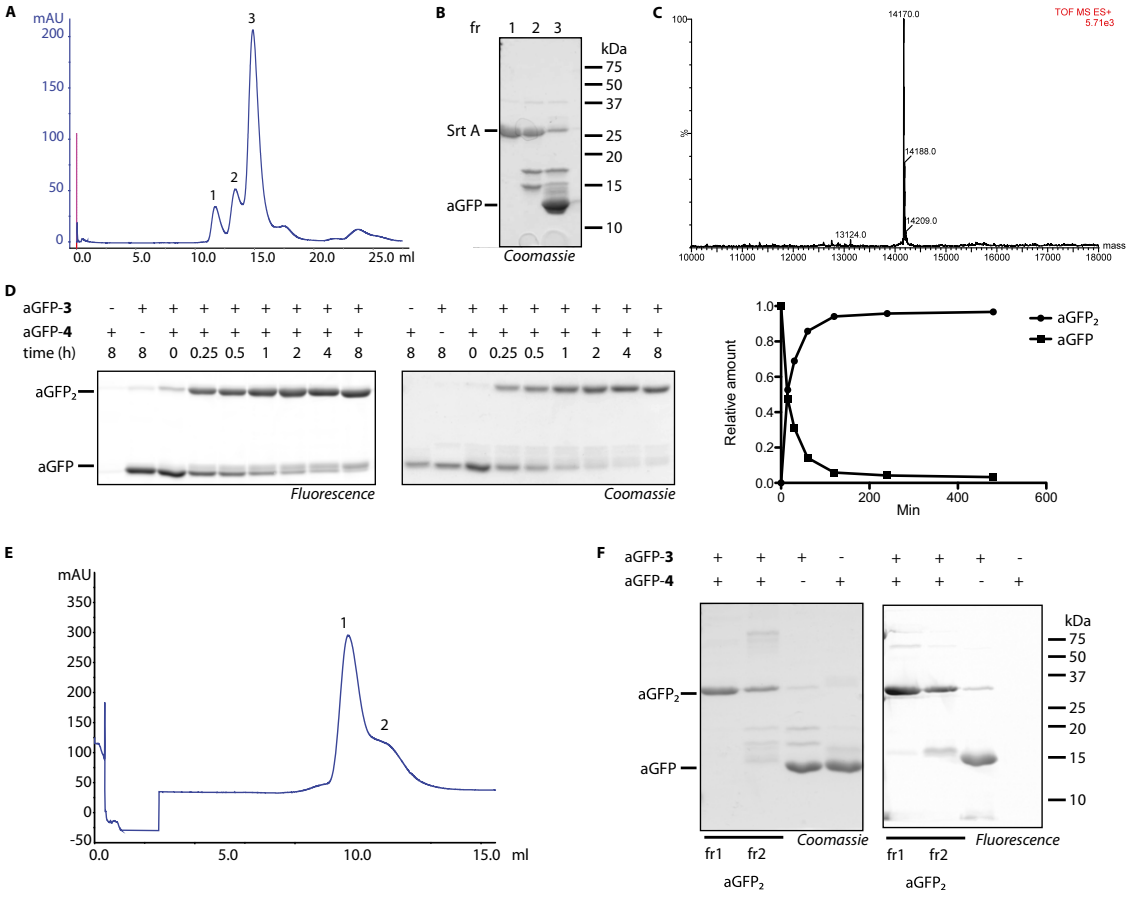

Fig. S2. (*A*) purification of anti-GFP sortagged with probe **4**. (*B* and *C*) Coomassie brilliant-blue–stained gel (*B*) and mass spectrum (*C*) of purified anti-GFP labeled with **4**. (*D*) Dimerization of aGFP-**3** and aGFP-**4**. aGFP-**3** (2.5 μg, 0.17 nmol) in Tris (50 mM, pH 7.4, 150 mM NaCl) was incubated with an equimolar amount of aGFP-**4** for the indicated time at room temperature. The dimerized product was resolved from the monomer on a Tris-tricine SDS/PAGE. Proteins were visualized by fluorescent imaging (λ_{ex} = 532, λ_{em} = 580, *Left*) and Coomassie brilliant blue (*Center*) and quantified (*Right*). The relative amount of monomer vs. dimer was determined as described for ubiquitin. (*E*) Purification of anti-GFP dimer on a Superdex 75 10/30. (*F*) Analysis of the concentrated purified protein on 15% SDS/PAGE.

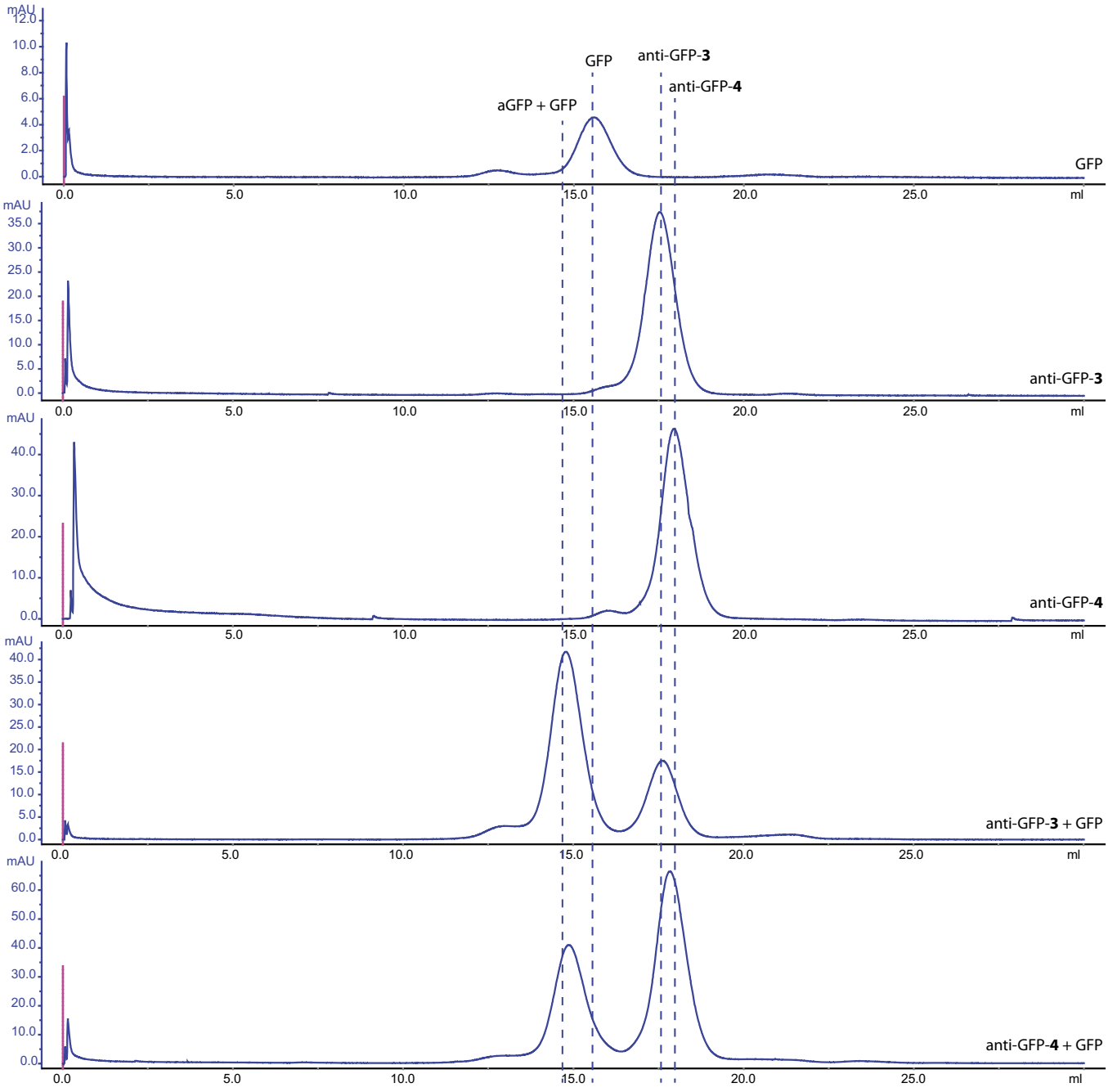

Fig. S3. Superdex 200 10/30 elution profile of monomer anti–GFP-**3** and anti–GFP-**4** incubated in the presence and absence of GFP.

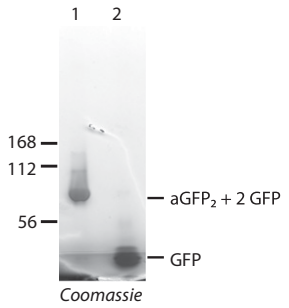

Fig. S4. The peaks eluting at 12.5 mL (1) and 15.5 mL (2) of anti-GFP dimer incubated with 30 μL GFP were concentrated and loaded on a native page.

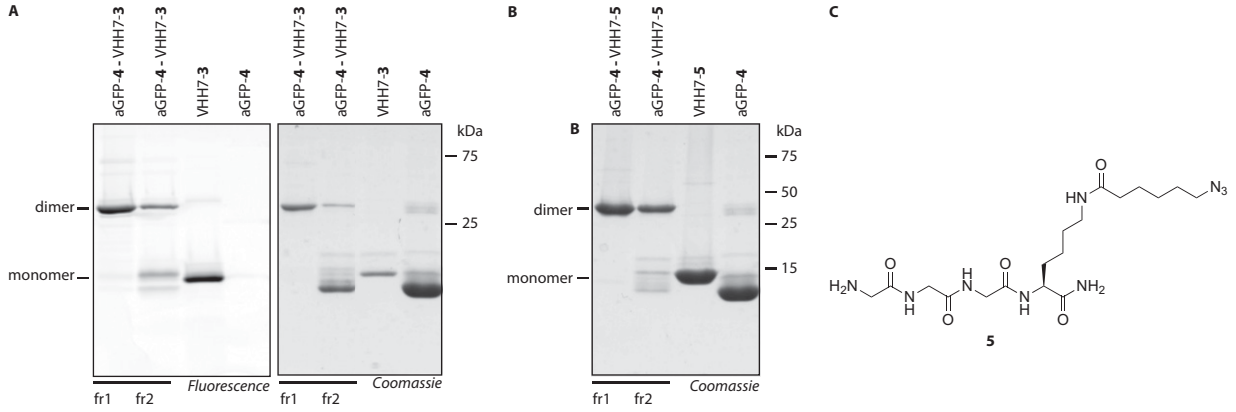

Fig. S5. (*A* and *B*) Dimerization and purification of fluorescent anti-GFP-**4**–VHH7-**3** (*A*) and nonfluorescent anti-GFP-**4**–VHH7-**5** (*B*). (*C*) Structure of 5.

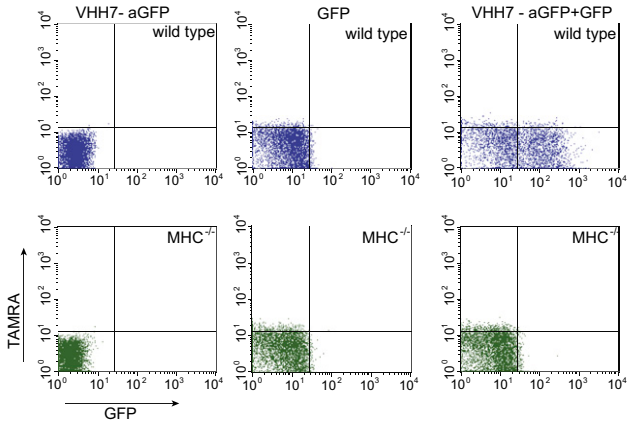

Fig. S6. FACs staining of mouse lymph node cells with anti-MHC II–anti-GFP antibodies. (*Upper*) Staining observed in wild-type cells. (*Lower*) Staining of MHC class II-deficient cells.

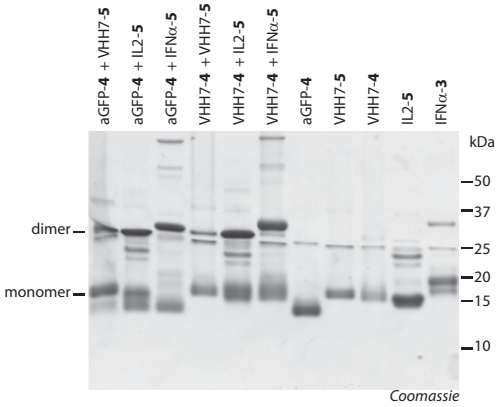

Fig. S7. Production of heterodimers of aGFP with VHH7, IL2, and IFNα.

Other Supporting Information Files

SI Appendix (PDF)