Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Inside a Shell—Organometallic Catalysis Inside Encapsulin Nanoreactors

*Philipp Lohner[+], Mariia Zmyslia[+], Johann Thurn, Jasmin K. Pape, Rūta Gerasimaitė, Jan Keller-Findeisen, Saskia Groeer, Benedikt Deuringer, Regine Süss, Andreas Walther, Stefan W. Hell, Gražvydas Lukinavičius, Thorsten Hugel, and Claudia Jessen-Trefzer**

anie_202110327_sm_miscellaneous_information.pdf

Supporting Information

Table of Contents

I. Abbreviations

6x-His	His-Tag, polyhistidine-tag
Å	Angstrom, 10^{-10}
aa	Amino acid
AA	Acceptor acceptor:: The signal from an acceptor-only specimen using the Acceptor filter set
AF	Alexa Fluor
APBS	All photon burst search
Asn	Asparagine
Asp	Aspartic acid
Au	Gold
BisTris	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
CA	Chloroalkane
Cat.	Catalyst
Ctrl.	Control
Cu	Copper
Da	Dalton
DA	Donor acceptor: The signal from an acceptor-only specimen using the donor filter cube
DCBS	Dual color burst search
DCM	Dichloromethane
DD	Donor donor: The signal from a donor-only specimen using the donor filter cube
DIPEA	*N*,*N*-Diisopropylethylamine
DLS	Dynamic light scattering
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DyP	Dye- decolorizing peroxidase

e.g.	For example (exempli gratia)
EDC*HCl	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
eGFP	Enhanced green fluorescent protein
ELS	Encapsulin localization sequence
Enc	Encapsulin
eq	equivalents
ESI-MS	Electrospray ionization- mass spectrometry
FCS	Fluorescence correlation spectroscopy
Flp	Ferritin- like protein
GSH	Glutathione
HCl	Hydrochloric acid
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HOBt	Benzotriazol-1-ol
HPLC	High performance liquid chromatography
HQ	8- hydroxyquinolinate- motif
ICP- MS	Inductively coupled plasma mass spectrometry
KPi	Potassium phosphate
LB	Lysogeny broth
LC/MS	Liquid chromatography/mass spectrometry
MEA	Cysteamine hydrochloride (2-Mercaptoethylamin hydrochlorid)
MWCO	Molecular weight cut off
NaCl	Sodium chloride
NaH_2PO_4	Monosodium phosphate
NMR	Nuclear magnetic resonance
Ø	Diameter
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pd	Palladium
PDI	Polydispersity index
PEG	Polyethylene glycol

PIE	Pulsed interleaved excitation
PMSF	Phenylmethylsulfonyl fluoride
pos.	Positive
rt	Room temperature
Ru	Ruthenium
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SEM	Standard error of the mean
smFRET	Single- molecule Förster resonance energy transfer
Strep	Strep-tag®
T4 PNK	T4 Polynucleotide Kinase
TEM	Transmission electron microscopy
TLC	Thin layer chromatography
TMR	Tetramethylrhodamine
TOF	Turnover frequency
TON	Turnover number
Tricine	N-(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine
Tris	Tris(hydroxymethyl)aminomethan
Trp	Tryptophan
UV	Ultraviolet

II. Supporting Table

Table S1: Strains, primers and plasmids used in this study.

Strains	
Name	**Source**
E. coli XL1-Blue	Stratagen/Agilent
E. coli BL21 Star (DE3)	Novagen, Merck

Plasmids	
Name	**Source**
pCDFDuet-1	Novagen, Merck
pCDFDuet-1-Strep	Modified from pCDFDuet-1
pUC57-Kan::eGFP-C19	Custom vector purchased from GenScript
pENTR4-HaloTag	Addgene

Primers/Oligos	
Name	**Sequence**
PL1_fw/ PL2_rev (Strep-tag, XhoI/PacI)	**TCGAG**TGGAGCCACCCGCAGTTCGAAAAGTGA**TTAAT**/ **TAA**TCACTTTTCGAACTGCGGGTGGCTCCA**C**
PL3_fw/ PL4_rev (MSMEG_5380, NdeI/KpnI)	TAAGCA**CATATG**ATGAACAACCTCTATCGCGACCTC/ TGCTTA**GGTACC**GGGGGTCAGCGCGACAG
PL5_fw/ PL6_rev (GFP with 19aa, SacI/SalI)	TAAGC**GAGCTC**TGTGAGCAAGGGCGAGGAG/ **GTCGAC**TCAGCGGGTTCCTTTCAG
PL7_fw/ PL8_rev (Halo-tag w/oCterm, EcoRI/HindIII)	TAAGCA**GAATTC**AGCAGAAATCGGTACTGGCTTTC/ TGCTTA**AAGCTT**GCCGGAAATCTCGAGCG
PL9_fw/ PL10_rev (eGFP with 0aa, SacI/HindIII)	TAAGCA**GAGCTC**AGTGAGCAAGGGCGAGG/ TGCTTA**AAGCTT**TCACTTGTACAGCTCGTCCAT
PL11_fw/ PL12_rev (eGFP w/oCterm, SacI/NotI)	TAAGCA**GAGCTC**TGTGAGCAAGGGCGAGGAG/ TGCTTA**GCGGCCGC**CTTGTACAGCTCGTCCATGCC
Oligo_6AA (NotI/AflII)	**GGCCGC**TAGCCTGAAAGGAACCCGCTGA**C**/ **TTAAG**TCAGCGGGTTCCTTTCAGGCTA**GC**
Oligo_9AA (NotI/AflII)	**GGCCGC**TGGAATCGGCAGCCTGAAAGGAACCCGCTGA**C**/ **TTAAG**TCAGCGGGTTCCTTTCAGGCTGCCGATTCCA**GC**
Oligo_12AA (NotI/AflII)	**GGCCGC**TGGCTCACTCGGAATCGGCAGCCTGAAAGGAACCCGCTG A**C/**

	TTAAGTCAGCGGGTTCCTTTCAGGCTGCCGATTCCGAGTGAGCCA**G C**
Oligo_19AA (NotI/AflII)	**GGCCGC**TGCCGACCCGGTCCACACCGACGGCTCACTCGGAATCGG CAGCCTGAAAGGAACCCGCTGA**C**/ **TTAAG**TCAGCGGGTTCCTTTCAGGCTGCCGATTCCGAGTGAGCCGT CGGTGTGGACCGGGTCGGCA**GC**

III. Supporting Figures

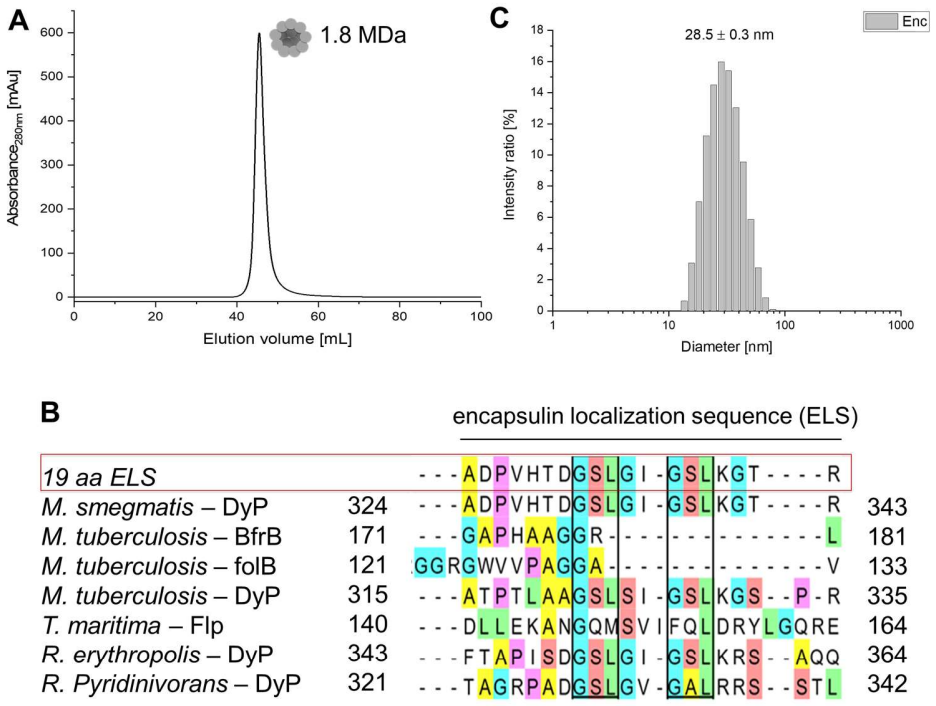


Figure S1: Enc_{SM} from *M. smegmatis*. A) Size exclusion chromatogram of $^{Strep}Enc_{SM}$. B) Alignment of the C-terminal localization sequences of the putative native cargo protein from *Mycobacterium smegmatis* (MSMEG_0529, DyP) and known cargo proteins from various organisms: ferritin BfrB from *Mycobacterium tuberculosis,* dihydroneopterin aldolase FolB from *Mycobacterium tuberculosis,* dyp-type peroxidase DyP from *Mycobacterium tuberculosis,* ferritin from *Thermotoga maritima,* peroxidase from *Rhodococcus erythropolis* and peroxidase from *Rhodococcus pyridinivorans.* The consensus sequence is rich in proline (magenta), alanine (yellow) and glycine (blue) and includes a single or double GSL- motif (glycine= blue, serine =red, leucine =green), which is highlighted in a black box. Alignment of encapsulin localization sequence (ELS) was done using Jalview[1] and T-Coffee[2]. C) DLS measurement of nanocompartment Enc. The hydrodynamic diameter was determined as Z-average of 28.5 ± 0.3 nm. The observed diameter is in good agreement with the results of the TEM measurement.

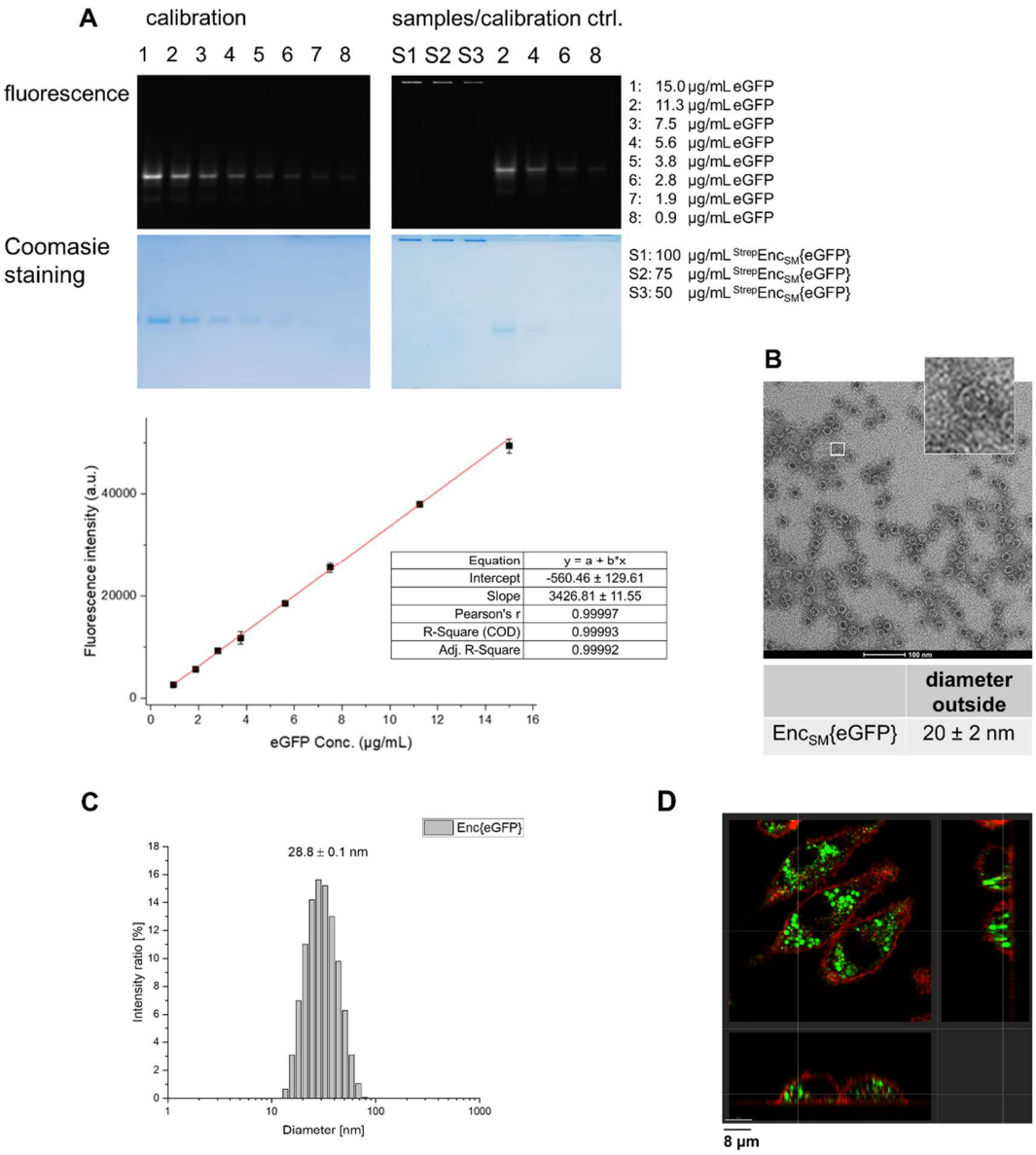


Figure S2: Cargo loading of $^{Strep}Enc_{SM}${eGFP}. $^{Strep}Enc_{SM}$ and eGFP, were co-expressed in *E. coli* and co-purified *via* Streptavidin-tag and size exclusion chromatography. During cloning 19 amino acids ELS were engineered into the C-terminus of eGFP. A) Native PAGE analysis of the purified construct *via* in-gel fluorescence detection and upon coomassie staining. eGFP detection was performed using a blue 535 Y filter, exposure time: 4 s. The experiment was performed 3 times and one representative example is shown. Samples 1-8 were quantified using ImageJ software and fitted using Origin Pro 2019b to generate a calibration curve. Subsequently, samples S1-S3 were quantified using the derived calibration curve. Error bars represent ± SD. B) TEM picture of $^{Strep}Enc_{SM}${eGFP}. Negative staining with uranyl acetate indicates a filled cavity. Number of capsids assessed for diameter calculation (n) = 117. C) DLS analysis of Enc{eGFP}. The hydrodynamic diameter was determined as Z-average [nm] = 28.8 ± 0.1 and the PDI as 0.13 ± 0.01. Analysis was performed in triplicates. D) Z-stack of J774A.1 monocytes treated with 8.5 nM $^{Strep}Enc_{SM}${eGFP} (green color) overnight confirming cellular uptake of the construct. The plasma membrane was stained with MemBright-635 (200 nM, red color). The probes were excited at 488 nm and 633 nm. Scale bar represents 8 µm. PDI: polydispersity index; SD: standard deviation. TEM: transmission electron microscopy. DLS: dynamic light scattering.

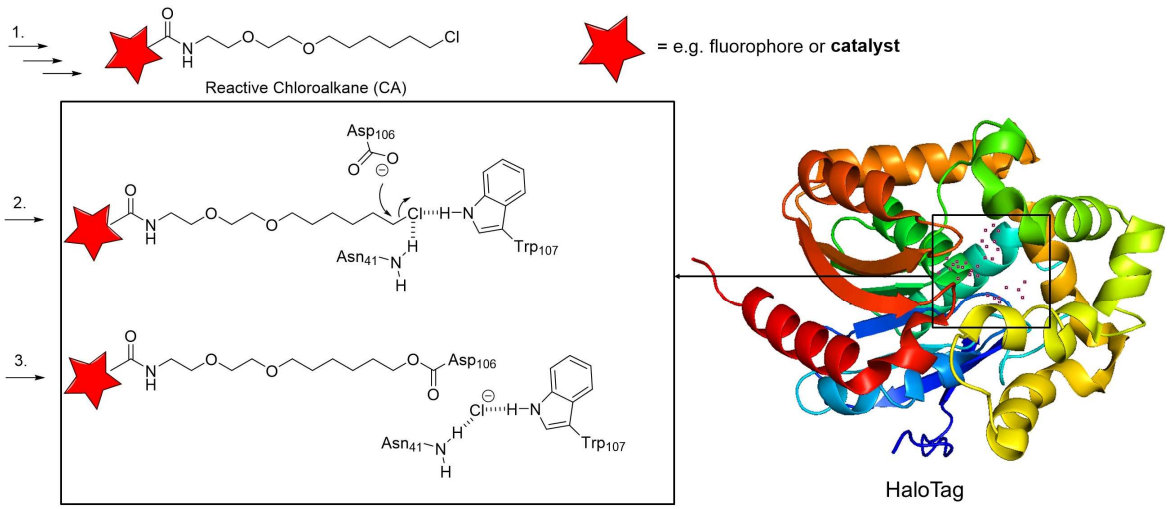


Figure S3: General mechanism of HaloTag covalent labeling with chloroalkane substrates. 1. Step: Synthesis of the reactive chloroalkane (CA), which is connected to a functional group, e.g. fluorescence dye or catalyst. 2. Step: The catalytic triad catalyzes nucleophilic S_N2- substitution of CA terminal chloride with aspartic acid (Asp 106) 3. Step: Reaction leads to irreversible ester formation, resulting in HaloTag that is covalently linked to a synthetic probe of interest (Ref. 19, main text). Cartoon representation of the crystal structure of HaloTag (taken from PDB: 5Y2X). The active site Asn41, Asp106 and Trp107 were highlighted as red dots.

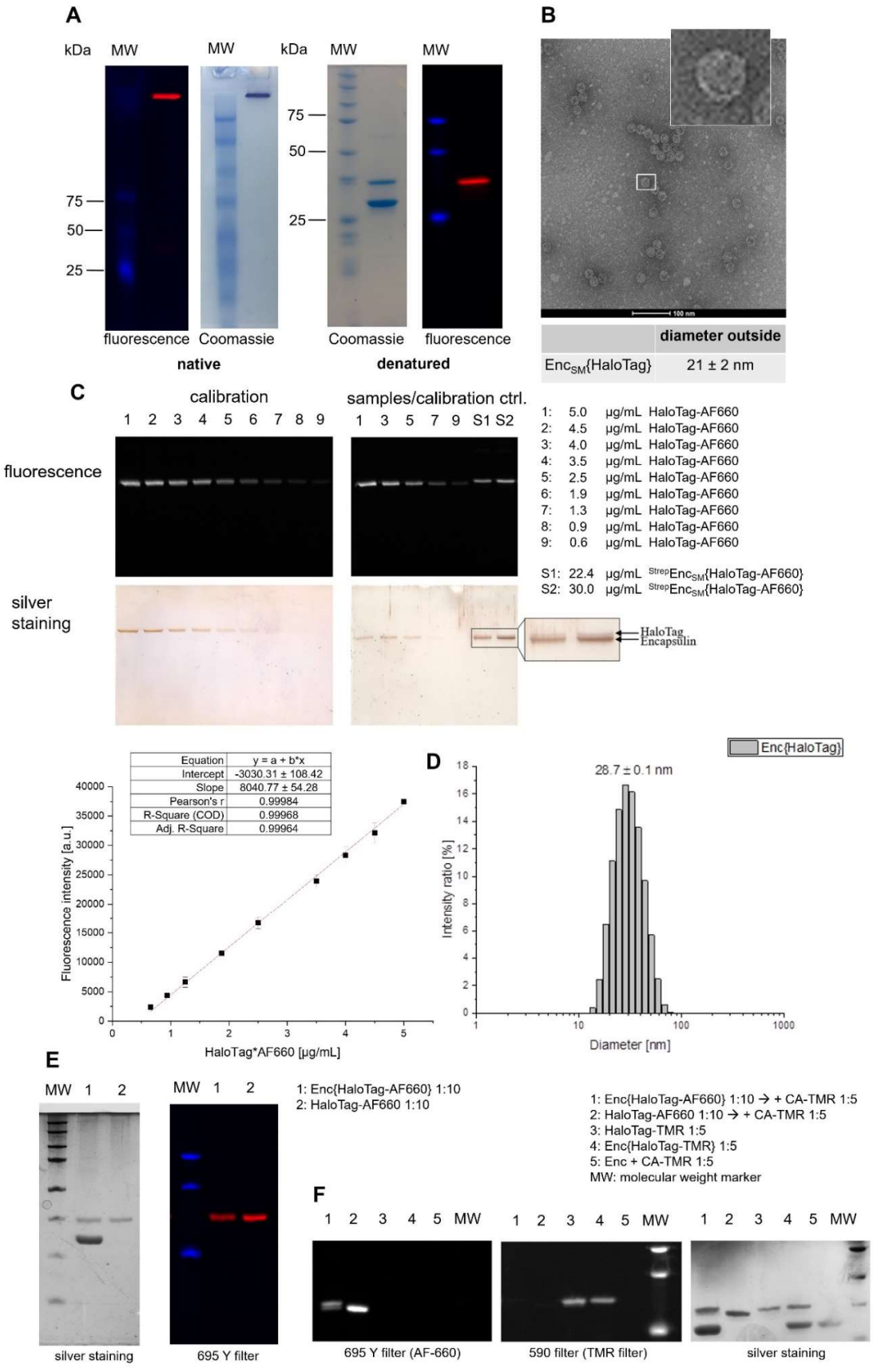

A
kDa
MW
MW
kDa
MW
MW
75
50
25
75
50
25
fluorescence
Coomassie
Coomassie
fluorescence
native
denatured
B
diameter outside
Enc$_{SM}${HaloTag}
21 ± 2 nm
C
calibration
samples/calibration ctrl.
1 2 3 4 5 6 7 8 9
1 3 5 7 9 S1 S2
fluorescence
silver staining
1: 5.0 µg/mL HaloTag-AF660
2: 4.5 µg/mL HaloTag-AF660
3: 4.0 µg/mL HaloTag-AF660
4: 3.5 µg/mL HaloTag-AF660
5: 2.5 µg/mL HaloTag-AF660
6: 1.9 µg/mL HaloTag-AF660
7: 1.3 µg/mL HaloTag-AF660
8: 0.9 µg/mL HaloTag-AF660
9: 0.6 µg/mL HaloTag-AF660
S1: 22.4 µg/mL Strep Enc$_{SM}${HaloTag-AF660}
S2: 30.0 µg/mL Strep Enc$_{SM}${HaloTag-AF660}
HaloTag
Encapsulin
Equation y = a + b*x
Intercept -3030.31 ± 108.42
Slope 8040.77 ± 54.28
Pearson's r 0.99984
R-Square (COD) 0.99968
Adj. R-Square 0.99964
Fluorescence intensity [a.u.]
HaloTag*AF660 [µg/mL]
D
Enc{HaloTag}
28.7 ± 0.1 nm
Intensity ratio [%]
Diameter [nm]
E
MW 1 2
MW 1 2
silver staining
695 Y filter
1: Enc{HaloTag-AF660} 1:10
2: HaloTag-AF660 1:10
1: Enc{HaloTag-AF660} 1:10 → + CA-TMR 1:5
2: HaloTag-AF660 1:10 → + CA-TMR 1:5
3: HaloTag-TMR 1:5
4: Enc{HaloTag-TMR} 1:5
5: Enc + CA-TMR 1:5
MW: molecular weight marker
F
1 2 3 4 5 MW
1 2 3 4 5 MW
1 2 3 4 5 MW
695 Y filter (AF-660)
590 filter (TMR filter)
silver staining

Figure S4: Purification and characterization of Enc_{SM} loaded with HaloTag. $^{Strep}Enc_{SM}$ and HaloTag were co-expressed in *E. coli* and co-purified *via* Streptavidin-tag and subsequent size exclusion chromatography. HaloTag was targeted into the capsid *via* tethering to the encapsulin localization sequence. A) The construct was labeled with the commercially available ligand dye CA-AlexaFluor 660® (Promega). Native and denaturing PAGE analysis of the purified construct, followed by Coomassie staining and in-gel fluorescence. B) TEM picture of $^{Strep}Enc_{SM}${HaloTag}. Negative staining with uranyl acetate indicates a filled cavity. Number of capsids assessed for diameter calculation (n) = 100. C) Loading capacity of successfully labeled $^{Strep}Enc_{SM}${HaloTag-AF660} was estimated by in-gel fluorescence and generation of a standard curve using HaloTag-AF660 only. Denaturing PAGE analysis of the purified construct and analysis by silver staining and in-gel fluorescence. HaloTag-AF660 detection was performed using a red 695 Y filter with an exposure time of 2 s 520 msec. The experiment was repeated 3 times and one representative example is shown. Samples 1-8 were quantified using ImageJ software and fitted using Origin Pro 2019b to generate a calibration curve. Subsequently, samples S1 and S2 were quantified using the derived calibration curve. Error bars represent ± SD. AF660: AlexaFluor 660®. D) DLS analysis of Enc{HaloTag}. The hydrodynamic diameter was determined as Z-average [nm] = 28.7 ± 0.1 and the PDI as 0.11 ± 0.02. The experiment was done in triplicates. E) Analysis of the labeling efficiency of encapsulated HaloTag. Samples were labeled with CA-AlexaFluor 660®, loaded on a Tris-Tricine Gel for separating Enc (29 kDa) and HaloTag (33 kDa) bands and after staining/fluorescence scan analyzed using Image J. HaloTag-AF660 detection was performed using a red 695 Y filter with an exposure time of 0 s 40 msec. In an attempt to quantify the total number of HaloTag proteins found within one capsid, we have additionally quantified single HaloTag / Enc bands on the gels visualize by silver staining using Image J. This analysis allowed us to estimate a ratio of 1:4-1:5, indicating a total number of 12-15 HaloTag proteins within one capsid. F) Samples were additionally treated with excess CA-TMR to confirm saturation of all accessible labelling sites and analysed by denaturing page analysis. Detection was performed using a red 695 Y filter with an exposure time of 0 s 480 msec for AF660, and blue 590 filter with an exposure time of 1 sec for TMR. The analysis was performed in duplicates and one representative gel image is shown.

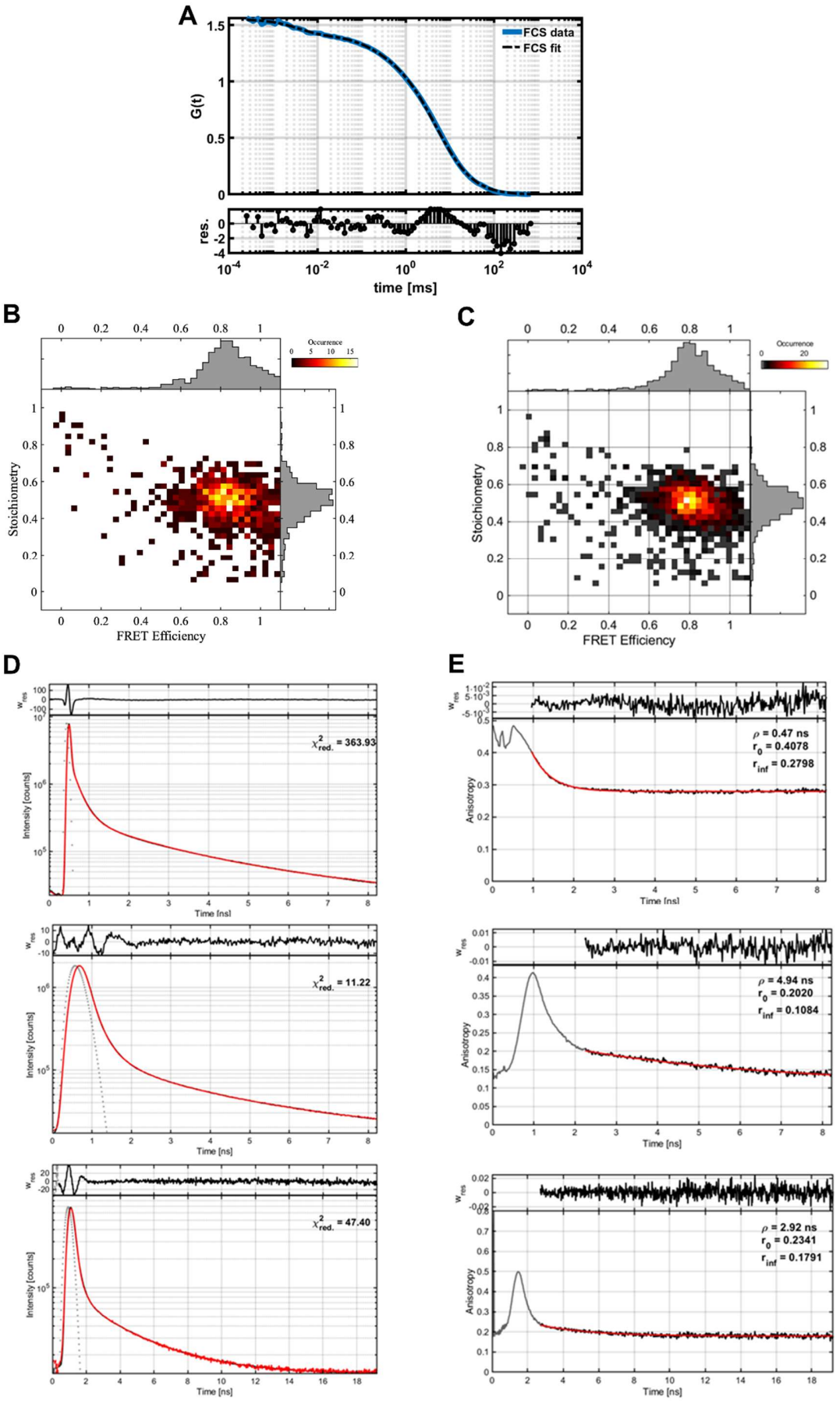

A
FCS data
FCS fit
G(t)
res.
time [ms]
B
Occurrence
Stoichiometry
FRET Efficiency
C
Occurrence
Stoichiometry
FRET Efficiency
D
$\chi^2_{red.}$ = 363.93
$\chi^2_{red.}$ = 11.22
$\chi^2_{red.}$ = 47.40
Intensity [counts]
Time [ns]
E
ρ = 0.47 ns
r0 = 0.4078
rinf = 0.2798
ρ = 4.94 ns
r0 = 0.2020
rinf = 0.1084
ρ = 2.92 ns
r0 = 0.2341
rinf = 0.1791
Anisotropy
Time [ns]

Figure S5: smFRET analysis. The analysis was done with Reconvolution Fits in the PAM software. A) FCS fit of the donor channels (donor cross-correlation between parallel and perpendicular) showing a diffusion time of the vesicles of around 5 ms, this translates to a hydrodynamic diameter of around 11 nm (see methods for details). B+C) smFRET analysis and DNA control. Stoichiometry vs. FRET efficiency plot for a standard DNA sample[4] containing only two fluorophores (one donor and one acceptor). In B) the result of the APBS comparable to the encapsulin (Fig. 3A, main text) is shown. The following parameters were used: DD rate min. 5kHz, DA rate min. 10kHz, AA rate min. 5kHz, FRET-2CDE filter 0-15. The signal is at stoichiometry 0.5, as expected for one donor and one acceptor (i.e. two dyes). In C) the result of the DCBS comparable to the encapsulin (Fig. 3B, main text) is shown. The following parameters were used: time window 500 us, DD rate min. 5kHz, DA rate min. 10kHz, AA rate min. 5kHz, FRET-2CDE filter 0-15. D) Lifetimes of the full unfiltered encapsulin dataset: donor upon donor excitation (top), acceptor upon donor excitation (middle), acceptor upon acceptor excitation (bottom). The instrument response function (IRF) approximated by a gamma distribution is shown in dotted lines. Data is shown in black, the fit in red. The following lifetimes were obtained: tau1 =3.02 ns, tau2 = 1.19 ns (top); tau1 = 2.90 ns, tau2 = 0.49 ns, tau3 = 0.12 ns (middle); tau1 = 3.14 ns, tau2 = 0.54 ns (bottom). The χ^2-values are large due to the high photon count, the residuals show that the fits are very good. E) Anisotropy of the full unfiltered encapsulin. Again top for donor-donor, middle for donor-acceptor and bottom for acceptor-acceptor. The fundamental anisotropy (r_0), the rotational correlation time (ρ) and the residual anisotropy (r_{inf}) determined by tailfits are given in the graphs. We used tailfits, as we are only interested in the residual anisotropy here. Temporal offsets between parallel and perpendicular channels were optimized manually within the PAM software. Please see Experimental Details for further information.

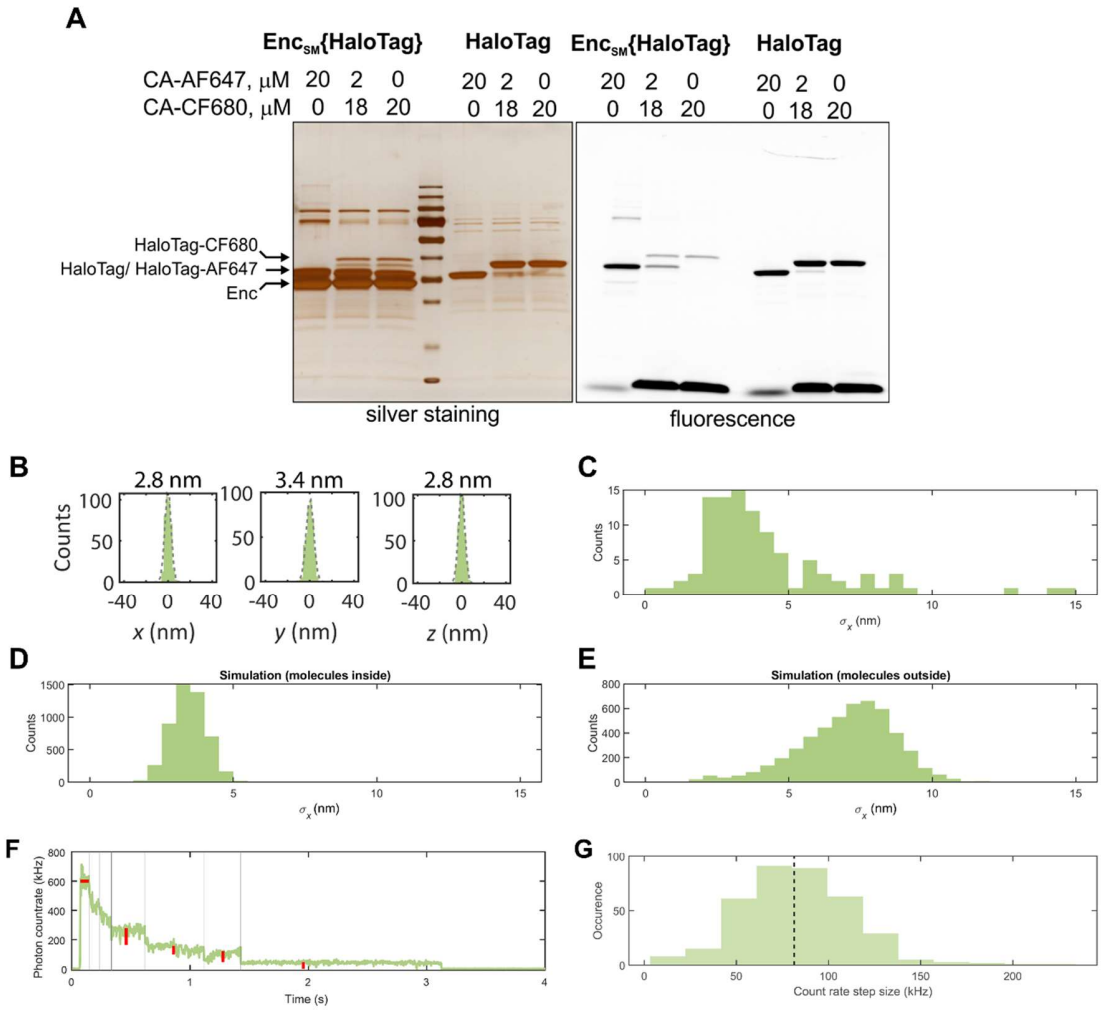
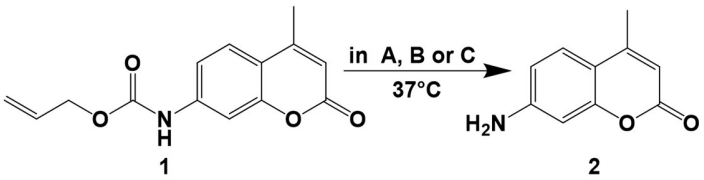


Figure S6: MINFLUX analysis. A) Labeling of free and encapsulated HaloTag with CA-AF647 and CA-CF680 substrates. The samples containing 0.8 mg/ml Enc$_{SM}${HaloTag} or comparable amount (1 µM) free HaloTag were incubated in 50 mM HEPES pH 7.4, 0.1% Tween 20 with indicated amount of substrates for 1h at room temperature and fractionated by SDS-PAGE. Left, staining with Pierce Silver Stain Kit; right, fluorescence in Cy5 channel. CF680-labeled HaloTag can be detected by electrophoretic mobility shift. Given e^{AF647} = 270000 $cm^{-1}{\cdot}M^{-1}$ and e^{CF680} = 140000 $cm^{-1}{\cdot}M^{-1}$, a rough estimation indicates that ~30% of molecules are labeled with CF680 in the sample labeled with both substrates. B) Distribution of localizations for the data shown in Figure 4C. A Gaussian fit delivers standard deviation values of 2 – 4 nm. C) 1D standard deviation values along *x* for all MINFLUX measurements (*n*=125). D) 1D standard deviation of simulated localizations from seven randomly distributed molecules within a shell of 14 nm diameter. E) Same as E, but with molecules randomly distributed within a spherical shell of 22 – 26 nm diameter. F) Off-switching count trace with initial intensity (horizontal red bar) and intensity step heights of manually detected steps (vertical red bars) for a single ^{Strep}Enc$_{SM}${HaloTag-AF647} construct. G) Histogram of intensity step heights observed during off-switching for n=125 off-switching traces and median of the distribution (dashed line) for a single ^{Strep}Enc$_{SM}${HaloTag-AF647} construct.

A



A: H_2O/DMSO (200:1)

Cat. [eq]	5 min	1 h	4 h	24 h
0.1	4 ± 1	34 ± 1	46 ± 1	46 ± 1
0.05	2 ± 0.1	26 ± 2	42 ± 1	41 ± 1

B: Kpi, pH 7.4

Cat. [eq]	5 min	1 h	4 h	24 h
0.1	9 ± 0	45 ± 0.2	49 ± 1	44 ± 1
0.05	7 ± 1	44 ± 1	48 ± 1	41 ± 1

C: 50 mM HEPES pH 7.4 + 0.05% Tween 20

Cat. [eq]	5 min	1 h	4 h	24 h
0.1	2 ± 0	22 ± 0.4	30 ± 1	29 ± 1
0.05	1 ± 0.2	14 ± 1	15 ± 1	15 ± 1

D: 50 mM HEPES pH 7.4 + 0.05% Tween 20 (anaerob)

Cat. [eq]	5 min	1 h	4 h	24 h
0.1	1 ± 0	83 ± 3	88 ± 2	> 90
0.05	1 ± 0.1	70 ± 4	84 ± 1	> 90

B

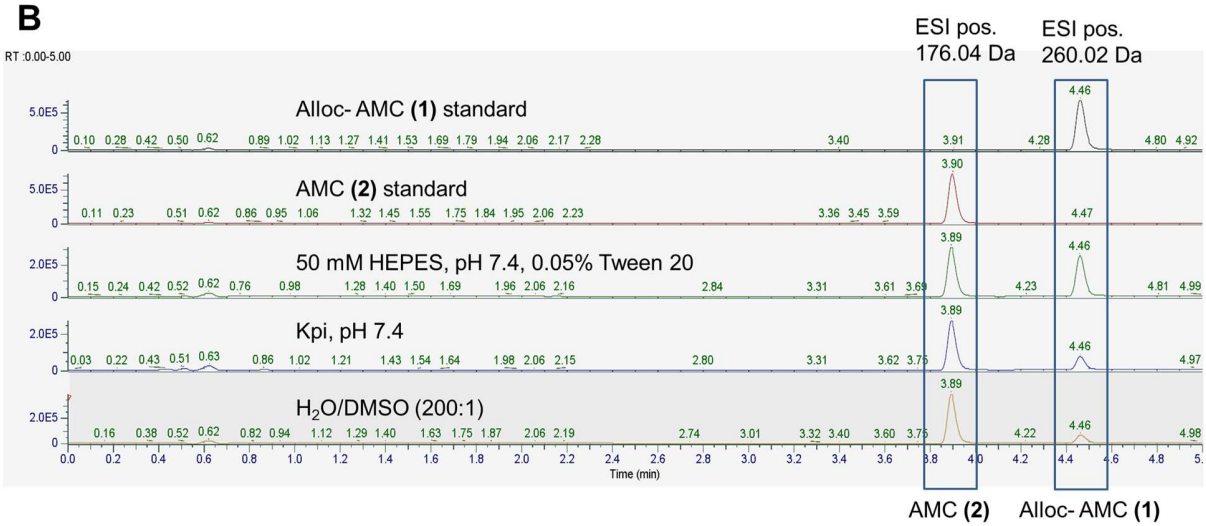


Figure S7: Catalytic deprotection of alloc-AMC (**1**) with CA-PEG3-Ru (**5**) in different buffer systems. A) Analysis of the de-*N*-allylation reaction of **1** to **2** under biologically relevant conditions (37° C, water based solvents A, B or C, in the presence of 5 mM glutathione, 125 µM in the presence of air (A-C) or under anaerobic conditions (D). Yields as determined by means of fluorescence detection (product formation in %). Cat.: CA-PEG_3-Ru (**5**). B) Product formation was monitored by means of LC-MS analysis. Reaction conditions: 125 µM 1, 5 mM GSH, 12.5 µM 5, 50 mM HEPES, pH 7.4 with 0.05% Tween 20, H_2O/DMSO (200:1) or Kpi, pH 7.4. The reactions were performed in LC-MS vials and incubated for 15 hours at 37° C. Subsequently, aliquots of each reaction were analyzed by LC-MS (330 nm). The experiment was performed in triplicates and one representative dataset is shown.

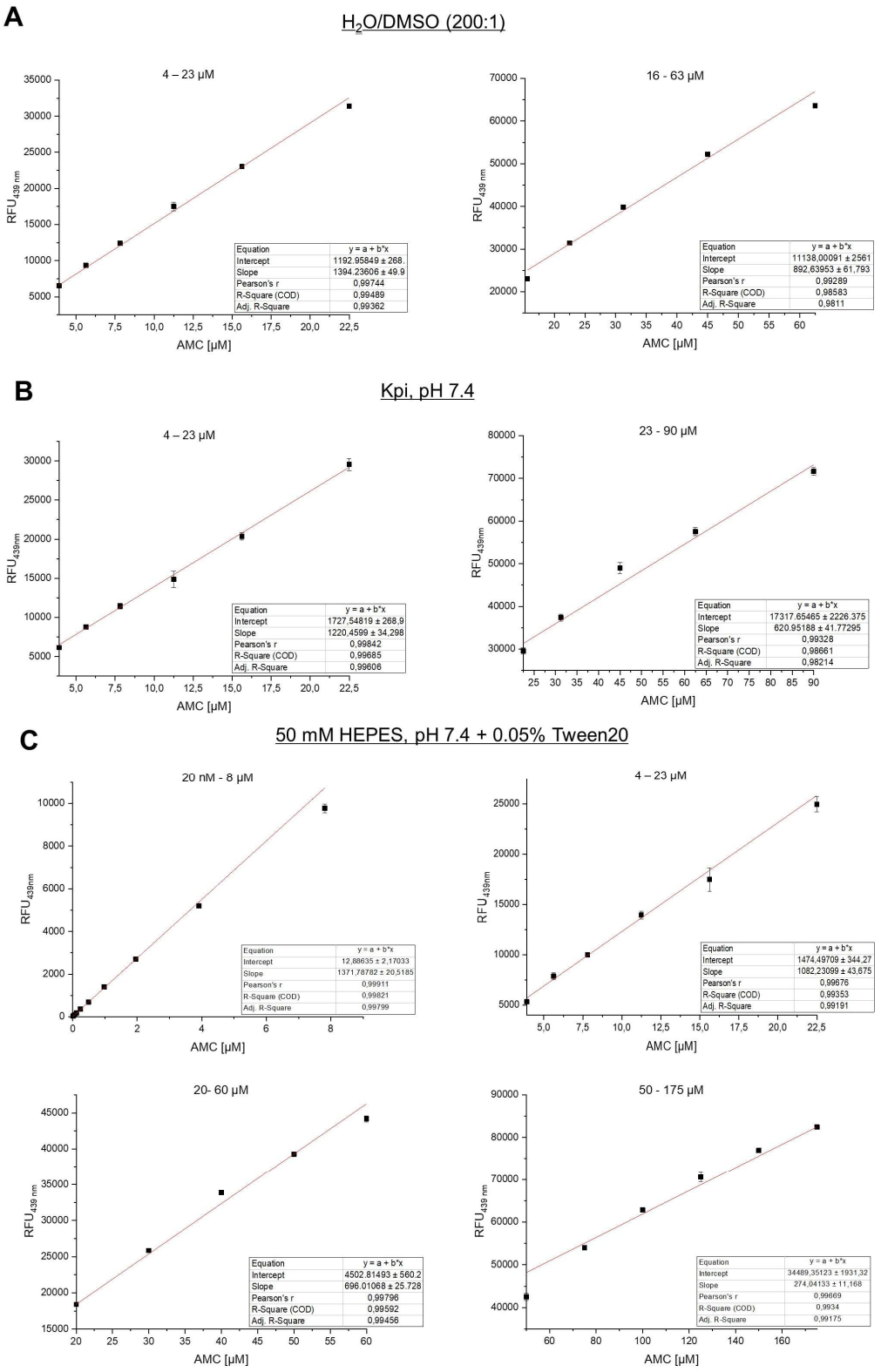
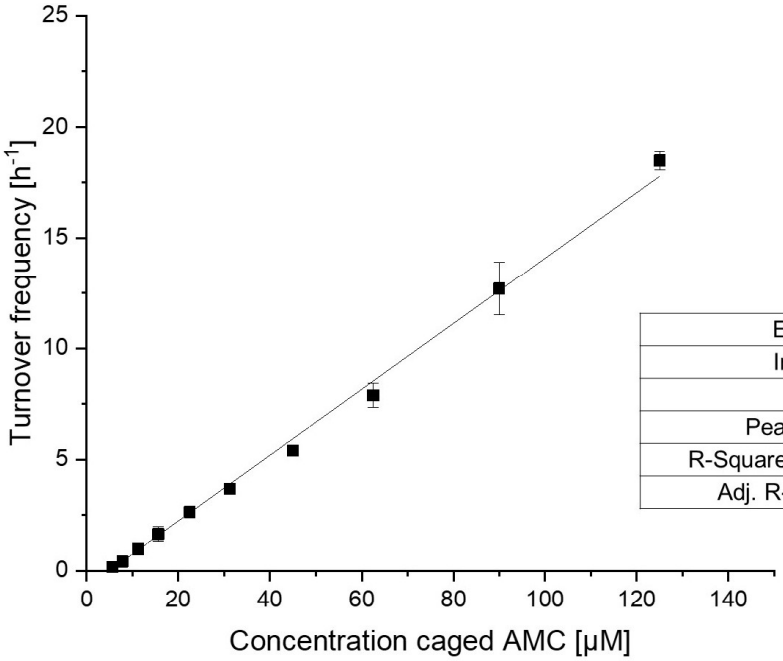


Figure S8: Fluorescence calibration curves of 7-Amino-4-methycoumarin (**2,** AMC) in various buffer systems (A: H_2O/DMSO (200:1); B: Kpi, pH 7.4 and C: 50 mM HEPES, pH 7.4 containing 0.05% Tween20) at different AMC concentrations. Fluorescence intensity measurement with two independent experiments (λ_{ex}=339 nm, λ_{em}= 439 nm) are shown. Fluorescence was detected using a 96-well plate reader (Tecan Sparks).



Equation	y = a + b*x
Intercept	-0.721 ± 0.05394
Slope	0.14807 ± 0.00307
Pearson's r	0.99829
R-Square (COD)	0.99658
Adj. R-Square	0.99615

Figure S9: Turnover frequency and reaction order of the uncaging process of **1**. Reaction conditions: 125-5.63 µM of **1**, 12.5 µM of **5** and 5 mM GSH in Water/DMSO (200:1) at 37° C. The turnover frequency was determined after 60 min. Yields were determined by fluorescence intensity measurement with three independent experiments (λ_{ex}=339 nm, λ_{em}= 439 nm), linear regression curve and calculated curve for linear reaction are shown.

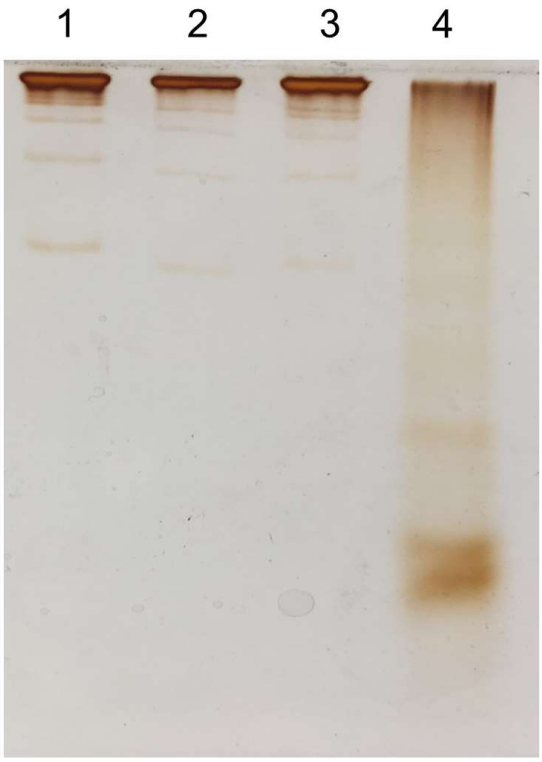


Figure S10: Native PAGE analysis followed by silver staining after completed labeling, washing (anaerobic) and fluorescence measurements. 1) $^{Strep}Enc_{SM}$ 2) $^{Strep}Enc_{SM}${GFP} 3) $^{Strep}Enc_{SM}${HaloTag-PEG_3-Ru} and 4) HaloTag-PEG_3-Ru.

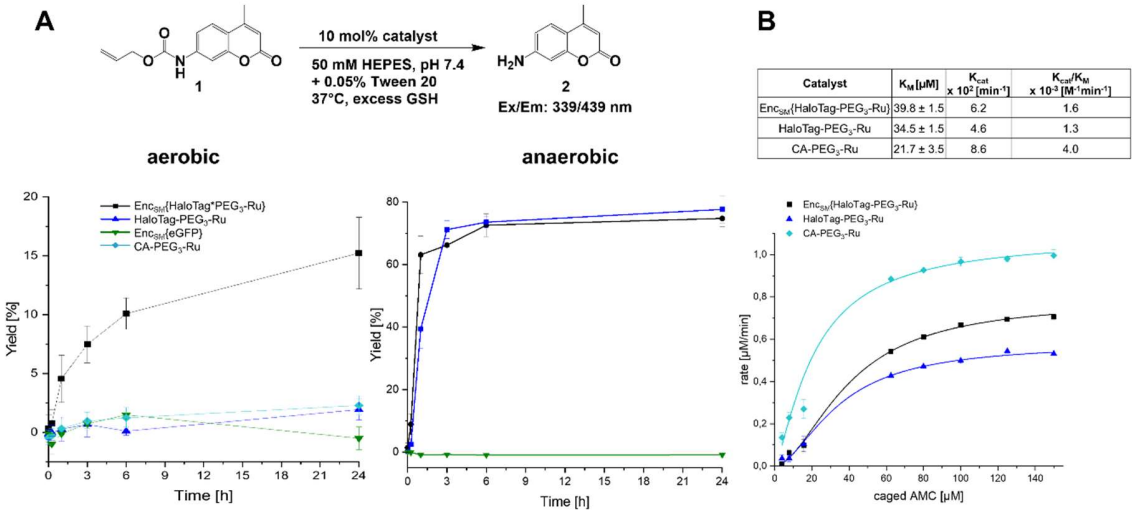


Figure S11: Ruthenium catalyzed uncaging of **1**. A) Uncaging reaction of **1** (125 µM, 1 eq.) to **2**. Aerobic labeling and washing steps and aerobic reaction conditions or anaerobic labeling/washing steps and anaerobic reaction conditions. Yields were determined by fluorescence intensity measurement in three independent experiments (λ_{ex}=339 nm, λ_{em}= 439 nm). Error bars represent ± SEM. Black square: Enc_{SM}(HaloTag-PEG_3-Ru}; blue triangle up: HaloTag-PEG_3-Ru, green triangle down: Enc_{SM}(GFP}, light blue diamond: CA-PEG_3-Ru. See also Figure S7A. B) Kinetic study of Ru-labelled proteins and Michaelis-Menten plots for the deallyation of **1** to **2** under standard conditions (anaerobic labeling/washing and aerobic reaction). Yields were determined by fluorescence intensity measurement in two independent experiments (λ_{ex}=339 nm, λ_{em}= 439 nm). Error bars represent ± SEM. Black square: Enc_{SM}(HaloTag-PEG_3-Ru}; blue triangle up: HaloTag-PEG_3-Ru, light blue diamond: CA-PEG_3-Ru. Please see Experimental Details for more information.

IV. Experimental Details

a. Molecular biology

General note. Microbial work was performed under sterile conditions in a laminar flow hood. PCR amplifications were performed on a Gene Amp PCR System 9700 (Thermo Scientific, Waltham, USA) or Mastercycler epi*gradient* (Eppendorf AG, Hamburg, Germany).

Strains, plasmids and growth conditions. Strains, plasmids and primers used in this study are listed in table S1. For cloning procedures *E. coli* XL1-Blue was cultivated in LB medium or LB-agar. For selection purposes streptomycin or spectinomycin were used at final concentrations of 50 µg/mL.

Predictions and BLAST alignment. The closest homologues of Enc_{SM} (MSMEG_5380, CFP29) were identified using blastp suite (National Library of Medicine). Alignment of encapsulin localization sequence (ELS) was done using Jalview[1] and T-Coffee[2].

Cloning of pCDFDuet-1::$^{Strep}Enc_{SM}$. Cloning was achieved by using standard cloning techniques with restriction enzymes. Firstly, commercially available pCDFDuet-1 was modified, by replacing the containing S-tag by Strep-tag using the restriction enzymes XhoI and PacI. Strep-tag was embedded by annealed oligo cloning using the primer pair PL1_fw/ PL2_rev as indicated in table S1. The gel-purified PCR product (Wizard® SV Gel and PCR Clean-Up System, Promega) was ligated into pCDFDuet-1 using T4 ligase. Correct sequence incorporation was validated by Sanger sequencing. Subsequently, the open reading frame encoding MSMEG_5380 was amplified from genomic DNA of *M. smegmatis* mc^2155 (isolated with the Wizard® Genomic DNA Purification Kit, Promega) using the primer pair PL3_fw/ PL4_rev as indicated in table S1. The obtained gel-purified PCR product (Wizard® SV Gel and PCR Clean-Up System, Promega) and pCDFDuet1-Strep (Novagen, Merck) were digested using the restriction enzymes NdeI and KpnI, again gel-purified and subsequently ligated using T4 ligase. After transformation into *E. coli* XL1-blue, colonies were selected on LB-agar plates with streptomycin or spectinomycin and sequenced to validate correct sequence incorporation.

Cloning of pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP, pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP-ELS0aa, pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP-ELS19aa and pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$HaloTag. The genomic sequence of eGFP followed by encapsulin localization sequence encoding for 19 amino acid long peptide (designed from MSMEG_5829) were purchased from GenScript (pUC57-Kan::eGFP-C19aa) and amplified using the primer pair PL5_fw/ PL6_rev (pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{N\text{-}6xHis}$eGFP-ELS19aa) or using the primer pair PL9_fw or PL10_rev for pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP-ELS0aa and PL11_fw and PL12_rev for pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP as indicated in table S1. HaloTag was amplified from pENTR4-HaloTag (Addgene) using the primer pair PL7_fw/ PL8_rev (pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$HaloTag) as indicated in table S1. All constructs were amplified using primers as indicated in table S1, gel-purified and ligated into pCDFDuet-1::$^{Strep}Enc_{SM}$ using the restriction sites as indicated in table S1.

Cloning of pCDFDuet-1:: $^{6xN\text{-}His}$eGFP-ELS19aa, pCDFDuet-1:: $^{6xN\text{-}His}$HaloTag. The constructs were cloned as described above, however ligated into empty pCDFDuet-1 (Strep-tag, no Enc_{SM}).

Cloning of pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP-ELS6-12aa, pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$HaloTag-ELS19aa and pCDFDuet-1:: $^{6xN\text{-}His}$HaloTag-ELS19aa Encapsulin localization sequence of varying size (6, 9, 12 amino acids for GFP and 19 amino acids for HaloTag) were ligated into pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$eGFP or pCDFDuet-1::$^{Strep}Enc_{SM}$/$^{6xN\text{-}His}$HaloTag using self-annealing oligos as described in table S1 and the indicated restriction enzymes. Oligos were purchased from Eurofins, Germany. Briefly, two complementary oligos were annealed and phosphorylated using following conditions: 1 µL each oligo (100 µM stock), 1µL T4 ligation buffer, 6.5 µL H_2O and 0.5 µL T4 PNK (NEB Biolabs). The reaction was heated in a thermo cycler: 37°C for 30 min, 95°C for 5 min and then ramped down to 25°C at 5°C/min. The phosphorylated and annealed oligo duplex was (after dilution 1:200 in 1x ligation buffer) directly ligated into the respective target vector (pre-treated with respective restriction enzymes and gel-purified).

b. Protein production and purification

Production and purification of $^{Strep}Enc_{SM}$, $^{Strep}Enc_{SM}${$^{6xN\text{-}His}$eGFP-ELS0-19aa} or **$^{Strep}Enc_{SM}${$^{6xN\text{-}His}$HaloTag-ELS19aa}**. Corresponding plasmids were transformed into chemically competent *E. coli* BL21 Star (DE3) cells and transformants were selected on LB-agar plates containing streptomycin or spectinomycin. Single colonies were isolated and validated by PCR analysis using the primer pairs as indicated in table S1. One colony was transferred into LB-medium (containing streptomycin or spectinomycin) and grown overnight at 37°C, 180 rpm. 5 mL pre-culture was transferred into 1 L auto induction medium (containing streptomycin or spectinomycin) and grown at 18° C for 60 hours. Cells were harvested by centrifugation (5000 x g, 4°C, 30 min) and lysed in cell lysis buffer (5 mL per gram wet weight, 50 mM NaH_2PO_4, 300 mM NaCl, 0.1 % Tween 20, 1 mM DTT and 1 mM PMSF, pH 8.0) by lysozyme (1 mg/mL, incubation on ice for 30-45 min) followed by sonication (6 x 10 s bursts at 40 % power with a 10 s cooling period between each burst, Sonoplus from Bandelin). The obtained crude cell lysate was centrifuged (10000 x g, 4°C, 30 min) and the supernatant was incubated with affinity beads (10 mL, 50% slurry/mL lysate). Subsequently, beads were washed with lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.1% Tween 20, pH 8.0) to remove any non-specifically bound proteins. Proteins were eluted from the beads using elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 2.5 mM desthiobiotin, 0.1% Tween 20, pH 8.0). The obtained fractions were transferred to SDS sample buffer (100 mM Tris-HCl (pH 6.8), 24% glycerol, 8% SDS, 5% mercaptoethanol, 0.02% bromophenol blue) and boiled at 95°C for 5 minutes. Proteins were separated *via* SDS-PAGE using 16.5% Tris-Tricine gels. Samples containing the desired protein were pooled, dialyzed overnight (50 mM HEPES, 0.05 % Tween 20, pH 7.4) and concentrated using a Vivaspin concentrator with a molecular weight cutoff of 10 kDa (Sartorius). Further purification was achieved by size exclusion chromatography on a HiLoad Superdex 200 PG column (Äkta Protein Purification System, GE Healthcare Germany). Protein was eluted from the column with elution buffer (50 mM HEPES, 0.05 % Tween 20, pH 7.4). Fractions were analyzed by SDS-PAGE, pooled and concentrated (Vivaspin 10 kDa cut off, Sartorius) and stored at -80°C until further use. Purification yielded 31 mg $^{Strep}Enc_{SM}$, 33 mg $^{Strep}Enc_{SM}${$^{6xN\text{-}His}$HaloTag-ELS19aa} and 29 mg $^{Strep}Enc_{SM}${$^{6xN\text{-}His}$eGFP-ELS19aa} per liter of culture.

Production and purification of $^{6xN\text{-}His}$eGFP-ELS19aa or **$^{6xN\text{-}His}$HaloTag-ELS19aa**. Corresponding plasmids were transformed into chemically competent *E. coli* BL21 Star (DE3) cells and transformants were selected on LB-agar plates containing streptomycin or spectinomycin. Single colonies were isolated and validated by PCR analysis using the primer pairs as indicated in table S1. One colony was transferred into LB-medium (containing streptomycin or spectinomycin) and cultured overnight at 37°C, 200 rpm. 5 mL pre-culture was transferred into 1 L auto induction medium (containing streptomycin or spectinomycin) and grown at 18° for 60 hours. Cells were harvested by centrifugation (5000 x g, 4°C, 10 min) and lysed in cell lysis buffer (5 mL per gram wet weight, 50 mM NaH_2PO_4, 300 mM NaCl, 0.1% Tween 20, 1 mM DTT and 1 mM PMSF, 10 mM imidazole, pH 8.0) by lysozyme (1 mg/mL, incubation for 30-45 min on ice) followed by sonication (6 x 10 s bursts at 40 % power with a 10 s cooling period between each burst, Sonoplus from Bandelin). The obtained crude cell lysate was centrifuged (10000 x g, 4°C, 30 min) and the supernatant was incubated with affinity beads (2.5 mL, 50% slurry/mL lysate) for 30 minutes at 4°C. Subsequently, beads were washed with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.1% Tween 20, 20 mM imidazole, pH 8.0) to remove any non-specifically bound proteins. Proteins were eluted from the beads using elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). The obtained fractions were mixed (1:1) with SDS sample buffer (100 mM Tris-HCl (pH 6.8), 24% glycerol, 8% SDS, 5% mercaptoethanol, 0.02% bromophenol blue) and boiled at 95°C for 5 minutes. Proteins were separated *via* SDS-PAGE using 16.5% Tris-Tricine gels. Samples containing the desired protein were pooled and concentrated using a Vivaspin concentrator with a molecular mass cutoff of 10 kDa (Sartorius). Further purification was achieved by size exclusion chromatography on a HiLoad Superdex 200 PG column (Äkta Protein Purification System, GE Healthcare Germany). Protein was eluted from the column with elution buffer (50 mM HEPES, 0.05% Tween 20, pH 7.4). Fractions were analyzed by SDS-PAGE, pooled, concentrated (Vivaspin 10 kDa cut off, Sartorius) and stored at -80°C until further use. Purification yielded 28 mg $^{6xN\text{-}His}$eGFP-ELS19aa and 26.6 mg $^{6xN\text{-}His}$HaloTag-ELS19aa per liter of culture.

c. Characterization of $^{Strep}Enc_{SM}$ as nanocompartment

SDS and native PAGE analysis. Purified protein complexes were mixed with an equal volume of loading buffer (native: 62.5 mM Tris-HCl (pH 6.8), 25% glycerol, 1% bromophenol blue; denatured: 100 mM Tris-HCl (pH 6.8), 24% glycerol, 8% SDS, 5% mercaptoethanol, 0.02% bromophenol blue) and incubated at room temperature for 10 minutes (native) or at 95°C for 5 min (denaturing). Samples were loaded on 8% native or 16.5 % Tris-Tricine gels (denatured) and run at 100 V, 60 mA for 90 min (native) (tank buffer: 25 mM Tris and 192 mM glycine) or 30 V (60 min), stepwise increasing every 10 min by 15 V until 100 V is reached, then 100 V (120 min), 60 mA for denatured conditions (tank buffer: 0.1 M Tris, 0.1 M tricine, 0.1% SDS). Gels were analyzed by in-gel fluorescence (gel imager: Vilber Lourmat, Fusion SL) and stained using Coomassie Brilliant Blue or silver staining.

eGFP- influence of the encapsulin localization sequence. Purified $^{Strep}Enc_{SM}$, $^{Strep}Enc_{SM}${eGFP-ELS0aa}, $^{Strep}Enc_{SM}${eGFP-ELS6aa}, $^{Strep}Enc_{SM}${eGFP-ELS9aa}, $^{Strep}Enc_{SM}${eGFP-ELS12aa} and $^{Strep}Enc_{SM}${eGFP-ELS19aa} were diluted to 0.5 mg/mL in 50 mM HEPES (pH 7.4) containing 0.05% Tween 20. eGFP detection was performed using a blue 535 Y filter, exposure time was 1 s (Vilber Lourmat, Fusion SL). Fluorescence intensity was quantified using ImageJ software and data were fitted in Origin Pro 2019b.

The protein concentrations were determined by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

Quantification of eGFP cargo protein *via* direct in-gel fluorescence detection. Firstly, purified eGFP-ELS19aa was used to prepare a dilution series in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20 for a calibration curve (Figure S2). The samples were loaded on a native gel. Fluorescence intensity of protein bands was quantified *via* ImageJ and fitted using Origin Pro 2019b. A linear correlation between protein concentration and fluorescence intensity was established. Dilutions of $^{Strep}Enc_{SM}${eGFP-ELS19aa} with concentrations 100 µg/mL, 75 µg/mL and 50 µg/mL were prepared in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20 and loaded on the native gel. The fluorescence intensity was determined and quantified using the derived calibration curve. The concentration was converted into molar concentrations (eGFP: 30.6 kDa, Enc_{SM}: 1800 kDa) and the molar ratio was calculated. eGFP detection was performed using a blue 535 Y filter, exposure time was 4 s (Vilber Lourmat, Fusion SL). The protein concentrations were determined by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

Z-stack analysis of Enc{eGFP}. For live cell imaging, 400 µL of $2.5x10^5$ cell suspension were seeded in 35 mm µ-dishes with a polymer coverslip bottom (Ibidi) and allowed to attach for 5 h. 1.6 mL of growth media was added and monocytes were allowed to proliferate for 24 h. Media was removed and 1.8 mL fresh media containing 8.5 nM $^{Strep}Enc_{SM}${eGFP} was added and incubated overnight. After incubation the medium was removed and the attached cells were washed twice with serum-free DME. The plasma membrane was stained with MemBright-635 (200 nM) prior to imaging (red color). Fluorescence images were obtained with inverted Zeiss LSM 880 laser scanning microscope using LD LCI PlanApochromat 40x/1.2 autocorr objective and applying Fast Airyscan technology for super-resolution microscopy. Fluorescence was recorded at 488 (green, $^{Strep}Enc_{SM}${eGFP}) and 633 nm (red, MemBright-635). Z-stacks were acquired with an interval of 0.227 µm. The Zeiss ZEN Black 2.3 software was used for Airyscan Processing. Figures were assembled with the software Imaris (9.7) from Bitplane.

Quantification of HaloTag cargo *via* direct in-gel fluorescence detection. Purified HaloTag-ELS19aa and $^{Strep}Enc_{SM}${HaloTag-ELS19aa} were labeled with CA-AlexaFluor 660®. Following assumption was made to calculate the amount of CA-AlexaFluor 660® for labeling reaction: six cargo proteins (HaloTag: 37.6 kDa, Enc_{SM}{HaloTag}: 2036 kDa) are loaded per one $^{Strep}Enc_{SM}${HaloTag} nanocompartment. The labeling reaction was performed in high excess of CA-AlexaFluor 660® at molar ratio of 1:10 (HaloTag:dye) for HaloTag and 1:15 for $^{Strep}Enc_{SM}${HaloTag-ELS19aa}. 150 µL HaloTag (20 µg/mL, 0.53 µM) and 100 µL $^{Strep}Enc_{SM}${HaloTag-ELS19aa} (total protein concentration 22.4 µg/mL, 0.011 µM; corresponds to 0.066 µM HaloTag) were mixed with 0.23 µL CA-AlexaFluor 660® (3.5 mM) and with 0.29 µL CA-AlexaFluor 660® (0.35 mM), respectively, and incubated under gentle shaking at 4°C overnight. Labeled HaloTag was used to prepare a dilution series in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20 for a calibration curve (Figure S4). The samples were loaded on SDS gel (denatured). Fluorescence intensity of protein bands was quantified *via* ImageJ and fitted using Origin Pro 2019b. A linear correlation between protein concentration and fluorescence intensity was established. Dilutions of labeled $^{Strep}Enc_{SM}${HaloTag-ELS19aa} with 22.4 µg/mL and 30.0 µg/mL were prepared in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20 and loaded on the denatured gel. The fluorescence intensity was determined and quantified using the derived calibration curve. The concentration was converted into molar

concentrations and the molar ratio was calculated. The fluorescence detection of the labelled protein was performed using a red 695 Y filter with an exposure time of 2 s 520 msec (Vilber Lourmat, Fusion SL). The protein concentrations were determined by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

Labelling efficiency of $^{Strep}Enc_{SM}${HaloTag-ELS19aa} *vs* HaloTag alone and saturation of labeling within Enc. Following the described procedure as stated above, we furthermore determined the labelling efficiency of HaloTag within the capsid (Figure S4) by quantification of SDS-page gel bands/fluorescence intensity using Image J. Secondly, we confirmed saturation of all accessible HaloTag sites within Enc. The labelling reaction was performed at molar ratios of 1:10 (HaloTag:CA-AlexaFluor 660®; 15.2 µM; for HaloTag and for $^{Strep}Enc_{SM}${HaloTag-ELS19aa}, and incubated under gentle shaking at 4°C overnight. The samples were split in two, while one half was loaded and analyzed by SDS-page. The second half was washed 5 times using Amicon Ultra 0.5 mL Centrifugal Filters (Merck Millipore, 10k MWCO cut-off), treated with a 5-fold excess of CA-TMR (Promega, 7.6 µM) and incubated for 3 h at room temperature. To confirm saturation of all binding sites samples were loaded and analyzed by SDS-page. The fluorescence detection of the labelled protein was performed using a red 695 Y filter with an exposure time of 40 msec or 480 msec for AlexaFluor 660® and using a blue 590 Y filter with an exposure time of 1 sec for TMR (Vilber Lourmat, Fusion SL).

HaloTag- cargo protein quantification *via* Single- molecule Förster resonance energy transfer (smFRET). Purified $^{Strep}Enc_{SM}${HaloTag-ELS19aa} was labelled with CA-JaneliaFluor 549® and CA-JaneliaFluor 646® (1:1 dye-ratio). The following assumption was made: $^{Strep}Enc_{SM}${HaloTag-ELS19aa} contains seven HaloTag cargo proteins (Enc_{SM}{HaloTag-ELS19aa}: 2031 kDa). The labelling reaction of HaloTag enclosed within Enc was performed in 1:10 molar ratio (HaloTag: dye). 100 µL of $^{Strep}Enc_{SM}${HaloTag-ELS19aa} (total protein concentration 180 µg/mL, 0.09 µM; corresponds to 0.63 µM HaloTag) was mixed with 3 µL CA-JaneliaFluor 549® (200 µM stock in DMSO) and with 3 µL CA-JaneliaFluor 646® (200 µM stock in DMSO) and incubated under gently shaking at 4°C overnight. Excess of unreacted dyes was removed by extensive wash with 50 mM HEPES, pH 7.4 containing 0.05% Tween 20 using Amicon Ultra 0.5 mL Centrifugal Filters (Merck Millipore, 50k MWCO). Single molecule measurements were carried out on a home-build confocal microscope.[3] Pulsed green and red laser light (532nm, LDH-P-FA-530 and 640nm, LDH-D-C-640, respectively, PicoQuant) was polarized, overlayed and focused on the sample by an 60x water immersion objective (CFI Plan Apo VC 60XC/1.2 WI, Nikon). Excitation light was separated from the emitted light by a dichroic mirror (F53-534 Dual Line beam splitter z 532/633, AHF). The emitted light was then guided through a further dichroic mirror (F33-647 beam splitter 640 DCXR, AHF) to separate donor and acceptor fluorescence. After spectral separation, pinholes with a diameter of 150 mm defined the detection volume. The effective volume is about 16 fl with an eccentricity of 4 (this corresponds to about 1.8 µm in diameter in x and y and 7.2 µm in z). We determined this by FCS calibration measurements with Rhodamine 6G and Atto 655 maleimide and using bead scans (in analogy to the Picoquant FCS AppNote). We chose the size of the volume larger than a diffraction limited volume on purpose to obtain longer bursts with better photon statistics. We adjusted the concentration such that the time between two bursts was much larger than the time of a burst. Finally, the two photon streams were separated by polarizing beam splitters into their parallel and perpendicular parts and recorded by single-photon detectors (two SPCM-AQR-14, PerkinElmer and two PDM series APDs, Micro Photon Devices). Time-correlated single photon counting with picosecond resolution and data collection was performed by a HydraHarp400 (PicoQuant)

and the Symphotime 32 software (PicoQuant). To reach the single-molecule level we adjusted the protein concentration to about 50-200 pM. Measurements were recorded for 30 min. All experiments were carried out at 22°C in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20. We used fluorescence correlation spectroscopy (FCS) to confirm that we have freely diffusing vesicles in solution (Figure S5A). Fig. S5A shows the donor cross-correlation and a fit resulting in a diffusion constant of about 40 µm²/s. This translates to a hydrodynamic radius of about 5.6 nm using the Stokes-Einstein relation, i.e. a hydrodynamic diameter of about 11 nm. This fits reasonably well with the diameter of the vesicles determined by TEM and DLS considering that bleaching and blinking increases the apparent diffusion constant in the FCS experiments and therefore results in an apparent smaller size.
Fig. S5D shows the lifetimes of the dyes. The longer lifetimes are in the range that are expected for free Rhodamine based dyes. The shorter lifetimes indicate quenching and FRET.
Fig. S5E show the anisotropy decays of the dyes. The residual anisotropy of the acceptor (0.18) hints towards a reasonably free acceptor. The donor seems to be (partially) stuck to the encapsulin as it shows an anisotropy of 0.28. For the FRET species the residual anisotropy is even lower.[3]
Altogether, the dyes are clearly affected by the constraint environment, but not more than what has been seen for dyes attached to proteins.

smFRET data evaluation: All photon burst search (APBS). The data was evaluated and filtered by a FRET-2CDE filters using the PAM software.[4] Further filters on the minimum count rates for DD, DA and AA fluorescence were used as specified in the figure captions. The histograms were correct such that the largest FRET peak is at stoichiometry 0.5, because we mainly observed one large FRET population with a relatively narrow peak. If many more dyes were present, FRET efficiencies and stoichiometries would be much more homogeneously spread across a wide range of FRET efficiencies and stoichiometries. In a FRET-Efficiency (E) *vs* stoichiometry (S) plot the position of species with one or two dyes are well known: Donor only is around S=1 and E=0, Acceptor only is around S=0, one Donor + one Acceptor is at S=0.5.[5] Although we do not fully understand the asymmetry in the E-S-plot for the APBS, we are convinced that more than two dyes are necessary to obtain the observed population between S=0.6 and S=0.8. In theory, two donor plus one acceptor would result in S=0.66 and three donor plus one acceptor in S=0.75.[6]

Dual color burst search (DCBS): A dual colour burst search with a time window of 500 µs results in a histogram showing a large FRET population that can be fitted with two Gaussians in the filtered data, indicating at least two different possible distances within the vesicle. This also indicates that there are at least three labelling sites (two labelling sites would only give one distance). Interestingly, these sites seem to be quite reproducible, i.e. each vesicle has the labelling sites at the same position. Otherwise the two populations would not be that well separated.

Correction factors for APBS and DCBS (same for both burst searches): (The data was corrected under the assumption that the main population is at S=0.5, as we had a 1:1 labeling).

gamma 0.5803

beta	0.3200
crosstalk	0.0907
direct exc	0.2199
G(D)	1.2
G(A)	1.16
l1	0
l2	0
BG DD par	2.4412
BG DD per	0.8406
BG DA par	1.5085
BG DA per	0.7615
BG AA par	0.7080
BG AA per	0.3959

HaloTag- cargo protein quantification *via* MINFLUX nanoscopy

Sample preparation. We incubated 0.8 mg/ml $^{Strep}Enc_{SM}${HaloTag} with 20 µM CA-AlexaFluor 647 or CA-AlexaFluor 647 and CA-CF680 mixture (molar ratio 1:9) in 50 µl of 50 mM HEPES pH 7.4 with 0.1% Tween 20 at room temperature for 1 h protected from light. Then, we dialyzed the reaction mixture against 13 ml of the same buffer for 3 h at room temperature, followed by overnight at 8°C against a fresh buffer, using Slide-A-Lyzer™ 10K MWCO (ThermoFisher Scientific). Immediately before measuring, the sample was diluted 20000-fold, to ~40 ng/ml.

For active sample stabilization during MINLFUX acquisition, we added fiducial markers (gold nanorods, A12-25-980-CTAB-DIH-1-25, Nanopartz Inc., Loveland, CO, USA) to all samples. The supplied solution of gold nanorods was diluted 1:3 in MilliQ water and sonicated for 5–10 min. The coverslips were incubated with about 50 µl of nanorod-solution for few seconds only. We rinsed the coverslips with MilliQ water and dried them using pressurized air. Then we spotted 50 µl of 0.01% poly-L-lysine, incubated the coverslips for 10 min at room temperature, rinsed with MilliQ water and dried in the air stream. Next, we applied 50 µl of the diluted sample and incubated a coverslip at room temperature for 10 min, briefly rinsed and stored it in the buffer protected from light before mounting. For sample mounting and imaging buffer we followed previously described protocols.[7] We used a concentration of 10mM cyteamine hydrochloride (MEA) in all measurements.

MINFLUX imaging. For MINFLUX single-molecule experiments, we used a previously presented custom-built microscope system.[7] During measurements, we first selected a fluorescent spot on a widefield camera image and subsequently switched all molecules to a long-lived dark state by applying ~100 µW of focused excitation light (entering the objective lens). For off-switching we targeted the regularly focused beam in a cross-like pattern with a distance of 300 nm between opposing coordinates.[7] Illuminating the molecules with low intensities enabled the observation of single-molecule off-switching steps in the photon count trace.

Once all molecules were switched off, we applied UV light until a single molecule was photo-activated and started to fluoresce. We used an iterative MINFLUX scheme[7] for localization, choosing between regularly focused and doughnut-shaped excitation beams, different beam separations L and collecting N photons before continuing in the next iteration. We used the following sequence: (1) Regularly focused beam, L_{xy}= 300 nm,

N= 100; (2) Regularly focused beam, L_{xy}= 300 nm, N= 150; (3) 3D doughnut beam, L_{xy}= 0, L_z= 300 nm, N=100; (4) 3D doughnut beam, L_{xy} = 0, L_z= 200 nm, N=100; (5) 3D doughnut beam, L_{xy}= 100 nm, L_z= 100 nm, N= 150; (6) 3D doughnut beam, L_{xy}= 70 nm, L_z= 70 nm, N= 250; (7) 3D doughnut beam, L_{xy}= 70 nm, L_z= 70 nm, N=10000. We repeated the activation-localization loop, slowly increasing the power of the activation laser (0.5 – 10 µW in the backfocal plane of the objective lens) until no more emission bursts occurred.

Data analysis and simulations. For counting molecules from the off-switching count traces, we applied a smoothing filter (Matlab function sgolayfilt) using a frame length of about 10 ms. We then manually identified a time period of constant intensity at the beginning of the off-switching process and used the average signal within this time period as an estimate of the initial brightness. We also measured the intensity of all manually identifiable switching steps in each trace and computed the molecule number from each off-switching trace (n=125) as the ratio of the initial brightness and the average switching step size.

We obtained single molecule localizations from the photon emission bursts after off-switching using a Matlab-based framework as previously described.[7] We filtered the localization data using $p_0 < 0.1$, $r_{\text{est, relative}} > 25$ nm and excluded localizations for which the maximum likelihood position estimator had not converged.[8]

We used the 1D standard deviation of the localizations as a figure of merit to compare different molecular distribution models with the experimental data (Figure S6). Due to clearly visible one-directional drift in the data, we excluded standard deviations along y from the analysis. For simulations, we randomly generated a Poissonian mean of 50 localizations within a Gaussian distribution of σ=2 nm around molecule positions that we arranged according to different model scenarios. In scenario 1, we randomly distributed all seven molecules within a sphere of 14 nm diameter, mimicking Alexa Fluor 647 molecules located inside the $^{\text{Strep}}\text{Enc}_{\text{SM}}$ shell. In scenario 2, we assumed molecules to be located in a spherical shell ranging from 22 – 26 nm, corresponding to a situation, where Alexa Fluor 647 is located outside the $^{\text{Strep}}\text{Enc}_{\text{SM}}$ shell. We repeated the generation of molecule positions and localizations (n=5000) and compared the resulting distribution of 1D standard deviations to the experimental data (Figure S6).

Control measurements. To draw the conclusion that the CA-AlexaFluor 647 molecules reside inside the $^{\text{Strep}}\text{Enc}_{\text{SM}}${HaloTag-AF647} shell, it is strictly required to observe more than one molecule per measurement. By step-based analysis of off-switching traces we confirmed the presence of around seven molecules in the observed region. There is a certain probability that molecules remain in a permanent dark state after off-switching, however, rather than being photo-activated by the UV light. To ensure that we observed multiple molecules during MINFLUX localization also, we performed two control experiments:

First, we labeled $^{\text{Strep}}\text{Enc}_{\text{SM}}${HaloTag} with AlexaFluor 647 and CF680 (see Figure S6) and performed a MINFLUX localization experiment as before, but additionally classified the emitter species based on the photon counts on the two spectral detection channels. About 17% of the observed $^{\text{Strep}}\text{Enc}_{\text{SM}}$ (n=60) carried both dye species, suggesting that at least in a subset of measurements we observe two or more molecules. However, this experiment did not show both dye species in >95% of the cases, which would be expected

for an equal incorporation and activation efficiency of CF680 and AF647 in the presence of seven molecules. There are at least two reasons for that: (1) While free HaloTag reacts with both substrates similarly, CA-CF680 labels encapsulated HaloTag much less efficiently than CA-AF647. Most likely, the bulky structure of the dye (MW~3000 Da) impedes diffusion of the substrate into the encapsulin shell, and as a result, only ≤30% of molecules are labeled with CA-CF680 even though this substrate is supplied in 9-fold excess. This explains a low number of particles having both labels. (2) Moreover, we observed a tendency for a lower activation efficiency for CF680 than for AF647, further reducing the ratio of observed CF680 molecules.

Secondly, we immobilized the Halo-ligand coupled to AlexaFluor 647 without $^{Strep}Enc_{SM}$ on the coverslip, and followed the same acquisition and data analysis procedure. On average we obtained roughly a fifth of the overall photon counts compared to measurements with the full $^{Strep}Enc_{SM}$\{HaloTag-AF647\} construct.

Negative stain-transmission electron microscopy (TEM). Nanocompartment samples were diluted to an initial concentration of 0.1-0.2 mg/mL in 50 mM HEPES buffer containing 0.05% Tween 20, 7.4 pH, and adsorbed for 1 min on carbon-coated grids (Electron Microscopy Sciences, CF300-CU) that had been prior exposed to a 25 s glow discharge. The grid was washed once with water, then stained with 1% uranyl acetate for 20 s. The excess stain was removed using filter paper. TEM images were recorded with a Thermo Scientific Talos L120C transmission electron microscope, operated at 120 kV. The images were analyzed using ImageJ software.

Fluorescence detection. Reactions were followed by LC-MS (see "Chemistry" section) and by following fluorescence detection of the reaction product at λ_{ex}=339 nm and λ_{em}= 439 nm in 96-well plates using a microplate reader (Spark Control, Tecan). The conversion of **1** to **2** was performed under following standard conditions: 125 µM **1**, 12.5 µM Ru-catalyst and 5 mM GSH in buffer, 37°C. Reactions in 96-well plates were carried out in a final volume of 100 µL buffer. The measurement was carried out at 339 nm, 20 nm bandpass excitation filter, the emission was monitored at 439 nm, 20 nm bandpass emission filter and following set-up: gain 46; number of flashes 30; integration time 40 µs and z-position 17300 µm. Stock solutions of alloc-AMC (**1**), AMC (**2**), and CA-PEG_3-Ru (**5**) were prepared in 10 mM DMSO, aliquoted and stored at -80°C. GSH stock was prepared freshly before use.

Turnover frequency (TOF). The turnover frequency of CA-PEG_3-Ru **(5)** under water/DMSO (200:1) conditions at different concentrations of **1** was determined. Following reaction mixture was used: 50 µL of a solution with water/DMSO (200:1) containing 5 mM GSH and 2.5 µM **5** were transferred into a 96-well plate. 50 µL of a solution of water/DMSO (200:1) containing 5 mM GSH, starting with 250 µM and 180 µM **1** were added to the first and second well, respectively. After a log2 dilution series to 5.6 µM **1**, the solution was mixed and incubated for 37°C. The kinetic was followed after 1h by fluorescence detection (λ_{ex}=339 nm, λ_{em}= 439 nm). The intensities were plotted against the substrate concentration using the software OriginPro2019b to determine the conversion by calibration curve.

Ru-labeled proteins under anaerobic and aerobic conditions. The protein stock solutions were diluted in 50 mM HEPES buffer, pH 7.4 containing 0.05% Tween 20. Following assumption was made: $^{Strep}Enc_{SM}$\{HaloTag-ELS19aa\} contains two cargo proteins ($^{Strep}Enc_{SM}$\{HaloTag-ELS19aa\}: 1866 kDa). The protein samples

$^{Strep}Enc_{SM}$\{HaloTag-ELS19aa\} and $^{6xN\text{-}His}$Halo-ELS19aa were diluted to 30 µM (related to HaloTag) and $^{Strep}Enc_{SM}$ and $^{Strep}Enc_{SM}$\{GFP-ELS19aa\} to 15 µM, respectively. 5 eq. of the CA-PEG_3-Ru **(5)** were added and the protein samples were incubated at room temperature for 6 hours either under aerobic or anaerobic conditions in an anaerobic glove box (Coy Laboratory Products, Grass Lake, MI, USA) with gentle shaking. At first, the protein samples were dialyzed with a Slide-A-Lyzer™ (Thermo Fisher Scientific, 10k MWCO) at room temperature overnight in 50 mM HEPES, 0.05% Tween 20, pH 7.4 with 2% DMSO. Then the samples were dialyzed in 50 mM HEPES, 0.05% Tween 20, pH 7.4 for 1 to 2 hours at room temperature to remove DMSO. Secondly, the samples were washed 15 times with 50 mM HEPES, 0.05% Tween 20, pH 7.4 using Amicon Ultra 0.5 mL Centrifugal Filters (Merck Millipore, 10k MWCO). The dialysis and wash steps were performed under aerobic or anaerobic conditions. The labeling, dialysis and wash steps were performed under exclusion of light. For monitoring of the wash process 40 µL of the flow-through was incubated with 10 µL alloc-AMC (**1**) (250 µM, 50 mM HEPES, 0.05% Tween 20, pH 7.4) and 1 µL freshly prepared GSH (250 mM, 50 mM HEPES, 0.05% Tween 20, pH 7.4) for 1 hour at 37°C and checked by TLC (cyclohexane:ethyl acetate 1:1). The protein samples all underwent the same labeling and wash procedure, as well as CA-PEG_3-Ru (**5**). The final solution consisted of 125 µM alloc-AMC (**1**), 12.5 µM Ru-labeled protein or 6.25 µM nanocompartment, 5 mM GSH in 50 mM HEPES, 0.05% Tween-20, pH 7.4. Reactions were incubated in a 96-well plate at 37°C protected by aluminum foil and kinetics was followed for 24 hours (reading time: 0 h, 0.25 h, 1 h, 3 h, 6 h and 24 h). The protein concentration was determined by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). The intensities were plotted against the substrate concentration using the software OriginPro2019b to determine the yield by calibration curve.

Ruthenium quantification *via* inductively coupled plasma mass spectrometry. Samples were processed as describe above ("Ru-labeled proteins under anaerobic and aerobic conditions"). Samples were diluted to a final concentration of 20 µM related to HaloTag or 10 µM related to nanocompartment in a final volume of 550 µL 50 mM HEPES (containing 0.05 % Tween20, pH 7.4) storing buffer. Ruthenium content was quantified in µg/mL at the Spurenanalytisches Laboratorium Dr. Baumann (Maxhütte-Haidhof, Germany). Samples were analyzed in biological duplicates and technical duplicates. Mean error indicates +/- SEM.

Kinetic studies of Ru-labeled proteins. Samples were processed as describe above ("Ru-labeled proteins under anaerobic and aerobic conditions"). The conversion of **1** to **2** was performed under following conditions: 150 µM, 125 µM, 100 µM, 80 µM, 62.5 µM, 15.6 µM, 7.8 µM and 3.9 µM of **1**, 12.5 µM Ru-catalyst and 5 mM GSH in 50 mM HEPES containing 0.05% Tween 20, pH 7.4, 37°C. Reactions in 384-well plates were carried out in a final volume of 20 µL. The measurement was carried out at 339 nm, 20 nm bandpass excitation filter, the emission was monitored at 439 nm, 20 nm bandpass emission filter and following set-up: gain 60; number of flashes 30; integration time 40 µs and z-position 17300 µm (Spark Control, Tecan). The kinetic was assessed after 30 min by fluorescence detection (λ_{ex}=339 nm, λ_{em}= 439 nm). The intensities were plotted against the substrate concentration using the software OriginPro2019b to determine the conversion by calibration curve.[9] Afterwards the rate was determined and the curve was fitted with the Michaelis-Menten Function (Hill Fit, OrginPro2019b).

Fluorescence microscopy and bright field images of J774A.1. Monocytes were treated with Enc_{SM}\{HaloTag-PEG_3-Ru\} + **1**, **5** + **1** or **1** only (negative ctrl.). Cells were

seeded in 96-well plates (2000 cells/well) and incubated overnight in DMEM (10 % FCS, Penicillin/Streptomycin, 37 °C, 5 % CO_2, humidified incubator) to allow cell attachment. Subsequently, cells were washed and treated with 200 µL of 1 µM solution either Enc_{SM}{HaloTag-PEG_3-Ru} or **5** in DMEM overnight (or DMEM only as ctrl.). Subsequently, **1** (20 µM) was applied to the cell supernatant and after incubation (4 hours) and washing, pictures were taken using a digital inverted microscope (EVOS FL, ThermoFisher Scientific, bright field, EVOS LED cube, v2, DAPI: excitation: 357/44 nm/emission: 447/60 nm, LPlanFL PH2 20x/0.40, LED light intensity 50%. Spectral Properties of **2** (7-Amino-4-methylcoumarin, AMC, CAS: 26093-31-2): Ex/Em: 344/437 nm in 50 mM HEPES, pH 7.4 containing 0.05% Tween 20, molar extinction coefficient at the maximum absorption wavelength: 5200 $cm^{-1}M^{-1}$. Image overlays were generated using ImageJ.

Dynamic light scattering (DLS). DLS data were collected on a Zetasizer Nano ZS (Malvern Instruments; Malvern, UK). Measurements were performed at 25°C using ultra low volume quartz cuvette (ZEN2112) containing 1 mg/mL of nanocompartment (Enc, Enc{eGFP} or Enc{HaloTag}) in 50 mM HEPES containing 0.05% Tween 20, pH 7.4. Three measurements were performed with the following set up: attenuator: 8; mean count rate (kcps): 260-420. The data were analyzed and presented with Zetasizer Nanoseries software (Malvern Instruments, Malvern, UK) using the general purpose analysis model. The intensity-weighted mean hydrodynamic diameter for each measurement was calculated and recorded.

d. Chemistry – including synthesis scheme SS1 and SS2

Chemicals, analytics and general remarks

Reactions were carried out in an open flask equipped with a magnetic stirrer at room temperature, unless otherwise noted.
Reagents were purchased from commercial suppliers (Sigma Aldrich, Acros) and used as received, unless noted otherwise.
Solvents were obtained in analytical grade and used as received.
TLC was performed using Macherey-Nagel 0.20 mm silica gel 60 with fluorescent indicator UV_{254nm} and visualized by UV fluorescence or $KMnO_4$ staining
HPLC solvents were used as obtained for analytical HPLC
Deuterated solvents for NMR were obtained from Euriso-Top, Germany, in the indicated purity grade and used as received for NMR spectroscopy.
HPLC analysis were performed using 1) an Agilent Technologies 1100 analytical HPLC equipped with a C_{18} column (3.5 µm, 4.6 × 100 mm, Waters XBridge) coupled to a UV and LC/MSD detector (Agilent Technologies 1100 Series) or 2) a Thermo Scientific (Dionex, Ultimate 3000) analytic HPLC equipped with a C_{18} column (Rapid Resolution HD 2.1x50mm 1.8-Micron, Zorbax Exlipse Plus C18) coupled to a UV and ESI-MS- detector (Thermo Fisher, TSQ Quantum Access Max). The methods were as follows:

Time (min)	Flow (mL/min)	% (H_2O:HAc = 99.5:0.5 (v/v))	% CH_3CN:HAc = 99.5:0.5 (v/v))
0.00	0.5	80	20
6.00	0.5	80	20
7.00	0.5	70	30

25.00	0.5	5	95
28.00	0.5	5	95
30.00	0.5	20	80
35.00	Stop run	Stop run	Stop run

1): UV detection at 230 nm, 254 nm and 330 nm

Time (min)	Flow (mL/min)	% (H_2O:HAc = 99.5:0.5 (v/v))	% CH_3CN:HAc = 99.5:0.5 (v/v))
-5.00	0.3	95	5
0.00	0.3	95	5
0.50	0.3	95	5
3.00	0.3	5	95
4.70	0.3	5	95
5.00	0.3	95	5

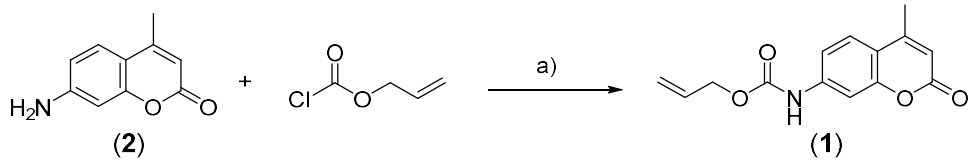
2) UV detection at 330 nm

^{1}H-NMR spectra were recorded on a Bruker 400 MHz spectrometers in the indicated deuterated solvent. Data are reported as follows: chemical shift (δ, ppm), multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; p quintet; m, multiplet; br, broad signal), coupling constant(s) (*J*, Hz), integration. All signals were referenced to the internal solvent signal as standard ($CDCl_3$, δ 7.26)

^{13}C- NMR spectra were recorded with ^{1}H-decoupling on Bruker 101 MHz (with cryoprobe) spectrometers at 298K in the indicated deuterated solvent. All signals were referenced to the internal solvent signal as standard ($CDCl_3$, δ 77.0).

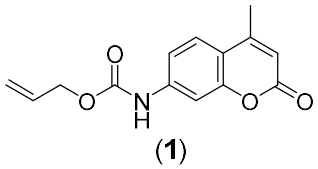
Mass spectra were recorded at the mass spectrometry service at the University of Freiburg on Finnigan TSQ 700 MS and Thermo Scientific EXACTIVE spectrometers with Orbitrap analyzer.

Synthetic procedures:

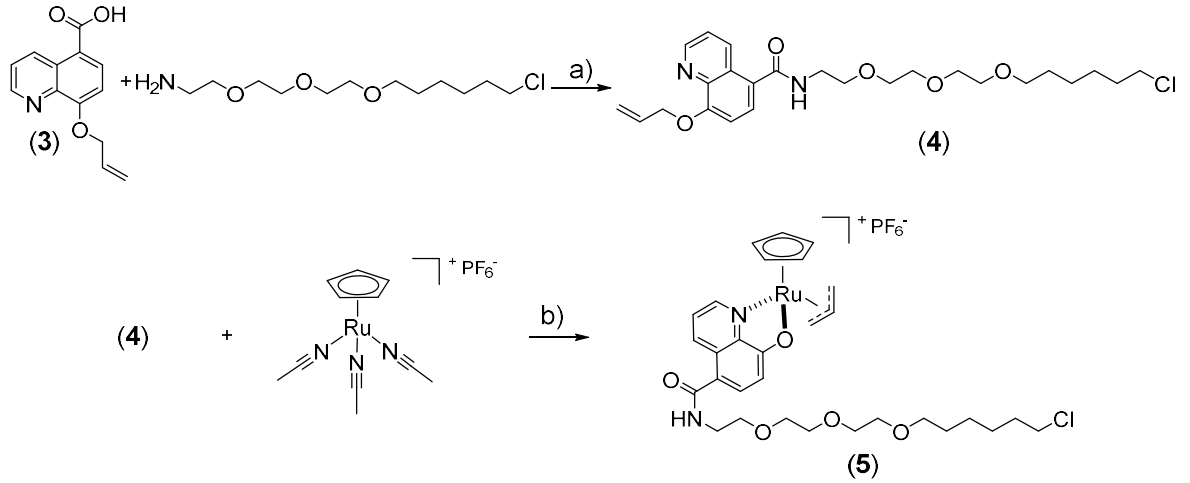


Scheme SS1: Synthesis of alloc- AMC (**1**). Reagent and conditions: a) Indium (10 mol%), solvent DMF_{dry}, rt, 24 hours.

Ally-(4-methyl-2-oxo-*2H*-chromen-7-yl) carbamate (1)

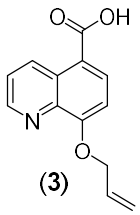


7-Amino-4-methylcoumarin (**2**) (150.00 mg, 0.86 mmol, 1 eq.) and Indium (9.9 mg, 0.09 mmol, 0.1 eq.) were mixed in 5 mL DMF_{dry} under a N_2- atmosphere. Allyl chloroformate (153.7 mg, 136 µL, 1.28 mmol, 1.5 eq.) was added and the mixture was stirred for 24 hours. The solvent was removed *in vacuo* and the residue purified by column chromatography on silica gel (hexane: ethyl acetate 1:1). Isolated yield: 49.6 mg, 0.19 mmol, 22% as white powder.
^{1}H NMR (400 MHz, chloroform-*d*): δ= 7.53 (d, *J* = 8.6 Hz, 1H), 7.45 (d, *J* = 2.2 Hz, 1H), 7.39 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.96 (s, 1H), 6.19 (d, *J* = 1.3 Hz, 1H), 6.03 – 5.92 (m, 1H), 5.39 (dq, *J* = 17.2, 1.6 Hz, 1H), 5.29 (dq, *J* = 10.4, 1.3 Hz, 1H), 4.70 (dt, *J* = 5.7, 1.4 Hz, 2H), 2.41 (d, *J* = 1.2 Hz, 3H) ppm.
^{13}C NMR (101 MHz, chloroform-*d*): δ= 161.15, 154.56, 152.75, 152.29, 141.43, 132.07, 125.47, 118.85, 115.67, 114.45, 113.31, 106.04, 66.41, 18.67 ppm
HRMS (ESI) *m/z* calculated for $C_{14}H_{13}NO_4$, $[M+H]^+$: 260.0845, observed: 260.0918



Scheme SS2: Synthesis of CA-PEG3-Ru (**5**). Reagent and conditions: a) HOBt, DIPEA, EDCI, DMF, 0°C → rt, overnight; b) DCM, 30 min, rt.

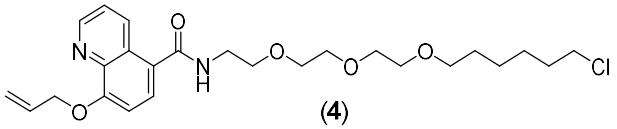
8-(allyloxy)quinolone-5-carboxylic acid (3):



(3)

Compound **3** was synthesized as described in the literature.[10] The analytical data were identical with the reported literature values.

8-(allyloxy)-*N*-(2-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethoxy)ethyl-quionoline-5-carboxamide (4)

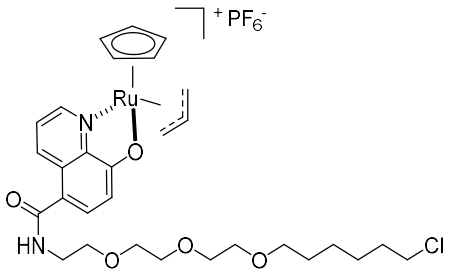


(4)

To a solution of compound **3** (30.0 mg, 0.13 mmol, 1 eg.) and 2-[2-[2-(6-chlorohexoxy)ethoxy]ethoxy]ethanamine (39.6 mg, 0.13 mmol, 1 eq.; AA Blocks, San Diego, USA, CAS number: 1261238-19-0) as hydrochloride in 1.5 mL DMF_{dry} at rt under inert gas, HOBt (24.5 mg, 0.16 mmol, 1.2 eq.) and *N, N*-diisoprpyolethylamine (67.2 mg, 91 µL, 0.52 mmol, 4 eq.) were added. The reaction mixture was cooled to 0°C and EDC*HCl (30.7 mg, 0.16 mmol, 1.2 eq.) was added to the solution. The solution was stirred for 30 min at 0°C and after removal of the ice bath, the solution was stirred for 22 hours at rt. The mixture was quenched at 0°C with water. The obtained solution was extracted three times with ethyl acetate and the combined organic extracts were washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (ethyl acetate: methanol: ammonia solution 14:1:1%). Isolated yield as brown oil: 59.8 mg, 0.12 mmol, 95%.

^{1}H NMR (400 MHz, chloroform-*d*) δ= 9.01 (dd, *J* = 4.2, 1.7 Hz, 1H), 8.95 (d, *J* = 8.6 Hz, 1H), 7.71 (d, *J* = 8.1 Hz, 1H), 7.54 (dd, *J* = 8.6, 4.1 Hz, 1H), 7.03 (d, *J* = 8.1 Hz, 1H), 6.73 (s, 1H), 6.21 (ddt, *J* = 17.3, 10.7, 5.4 Hz, 1H), 5.48 (dq, *J* = 17.3, 1.6 Hz, 1H), 5.37 (dq, *J* = 10.5, 1.4 Hz, 1H), 4.91 (dt, *J* = 5.5, 1.5 Hz, 2H), 3.77 – 3.62 (m, 8H), 3.61 – 3.55 (m, 2H), 3.54 – 3.43 (m, 4H), 3.32 (t, *J* = 6.7 Hz, 2H), 1.72 (m, *J* = Hz, 2H), 1.54 – 1.44 (m, 2H), 1.43 – 1.34 (m, 2H), 1.33 – 1.22 (m, 2H) ppm.
^{13}C NMR (101 MHz, chloroform-*d*): δ= 168.41, 156.14, 149.70, 140.13, 134.81, 132.70, 127.49, 126.74, 126.11, 122.57, 118.89, 107.73, 71.33, 70.69, 70.68, 70.42, 70.13, 70.10, 69.95, 45.17, 39.91, 32.61, 29.46, 26.76, 25.47 ppm.
HRMS (ESI): *m/z* calculated for $C_{25}H_{35}ClN_2O_5$, $[M+H]^+$: 479.2234, observed: 479.2307

CA-PEG$_3$- Ru (5)



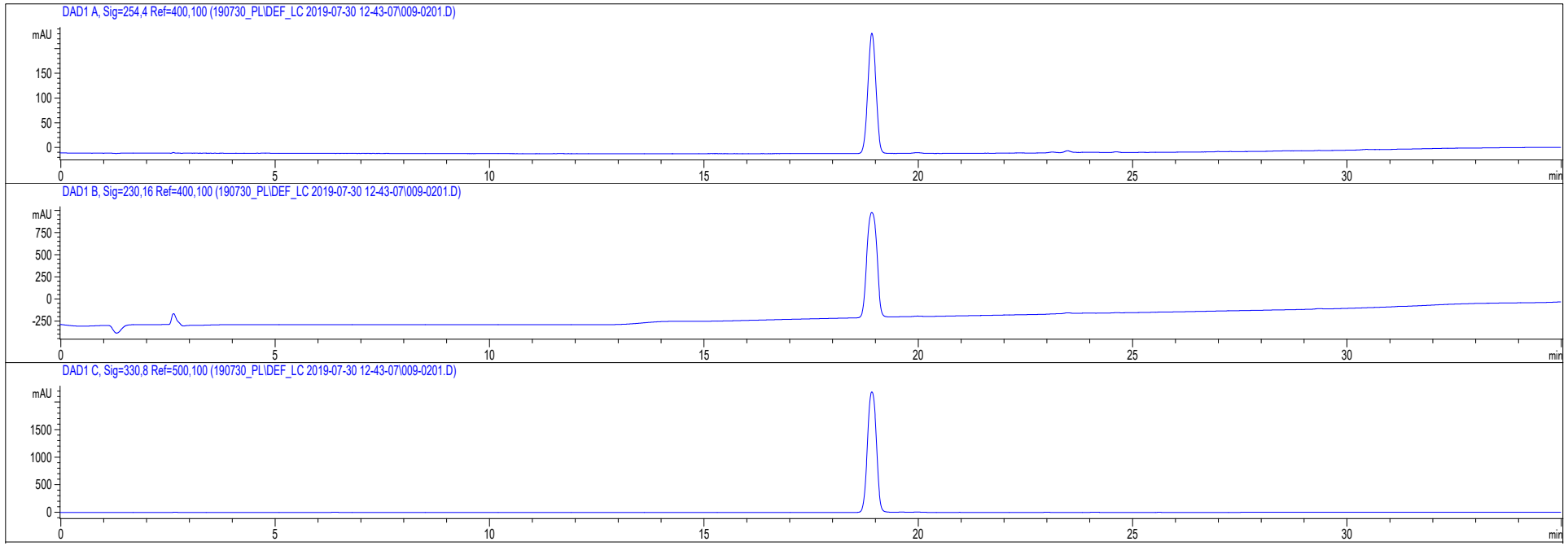
(**5**)

Under a N_2- atmosphere Tris(acetonitrile)cyclopentadienylruthenium(II) hexafluorophosphate (8.6 mg, 0,20 mmol, 1.05 eq., Sigma Aldrich) was dissolved in degassed DCM_{dry}. At room temperature compound **4** (9.0 mg, 0.19 mmol, 1 eq.) was added and stirred for 30 min. The solvent was removed under a nitrogen stream and the residue was dried under reduced pressure. The raw product was washed three times with ice cold chloroform. Isolated yield: 14.8 mg, 0.17 mmol, 88% as brown-orange solid.
^{1}H NMR (400 MHz, Chloroform-*d*) δ 9.24 (d, *J* = 8.6 Hz, 1H), 8.83 (d, *J* = 4.9 Hz, 1H), 7.67 (d, *J* = 8.2 Hz, 1H), 7.61 (dd, *J* = 8.7, 5.0 Hz, 1H), 6.89 (d, *J* = 8.2 Hz, 1H), 6.73 – 6.67 (m, 1H), 6.00 (s, 5H), 4.56 (d, *J* = 9.0 Hz, 1H), 4.44 – 4.31 (m, 2H), 4.28 – 4.16 (m, 2H), 3.74 – 3.59 (m, 10H), 3.57 – 3.48 (m, 4H), 3.40 (t, *J* = 6.6 Hz, 2H), 1.79 – 1.70 (m, 2H), 1.55 (q, *J* = 7.3 Hz, 2H), 1.48 – 1.37 (m, 2H), 1.37 – 1.21 (m, 2H) ppm.
^{13}C NMR (101 MHz, chloroform-*d*): δ= 171.23, 167.32, 155.93, 145.51, 139.58, 130.68, 129.51, 124.94, 120.84, 118.56, 114.85, 99.07, 95.74, 71.36, 70.67, 70.40, 70.17, 69.89, 69.47, 64.32, 45.22, 39.92, 32.62, 29.52, 26.78, 25.51 ppm.
HRMS (ESI): *m/z* calculated for for $C_{30}H_{40}N_2O_5ClRu$, $[M\text{-}PF_6]^+$: 645.1669, observed: 645.1664 and for $[PF_6]^-$: 144.9642, observed: 144.964

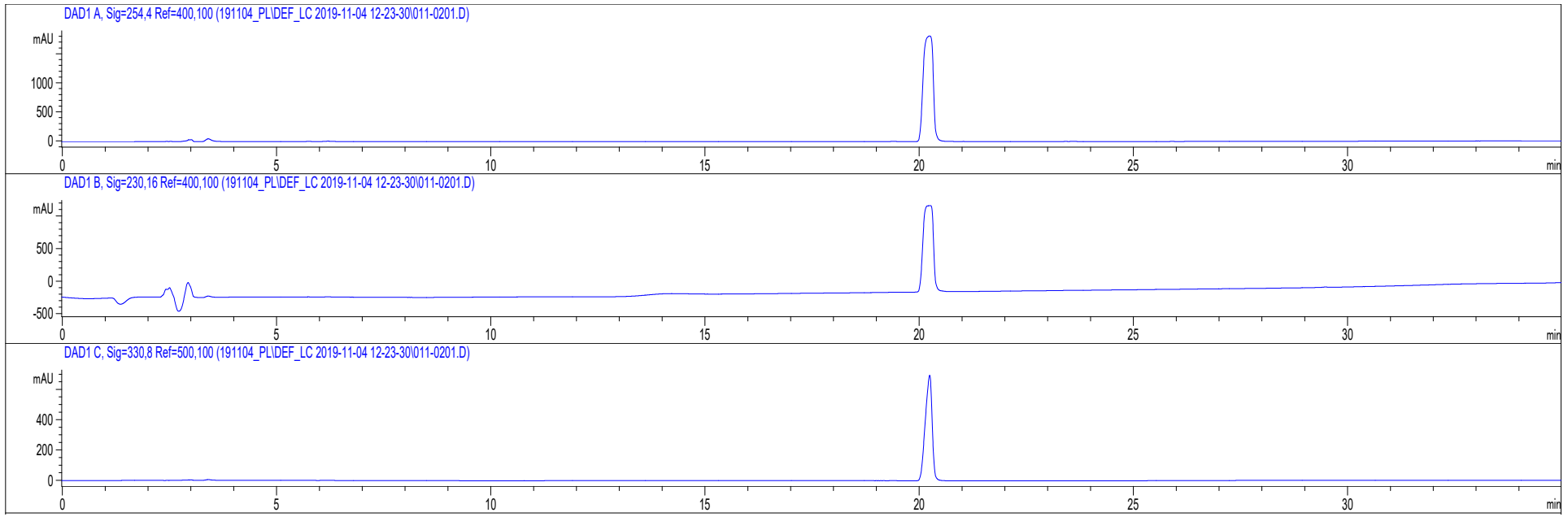
V. Attachment

a. LC chromatograms

Compound (1): Alloc- AMC



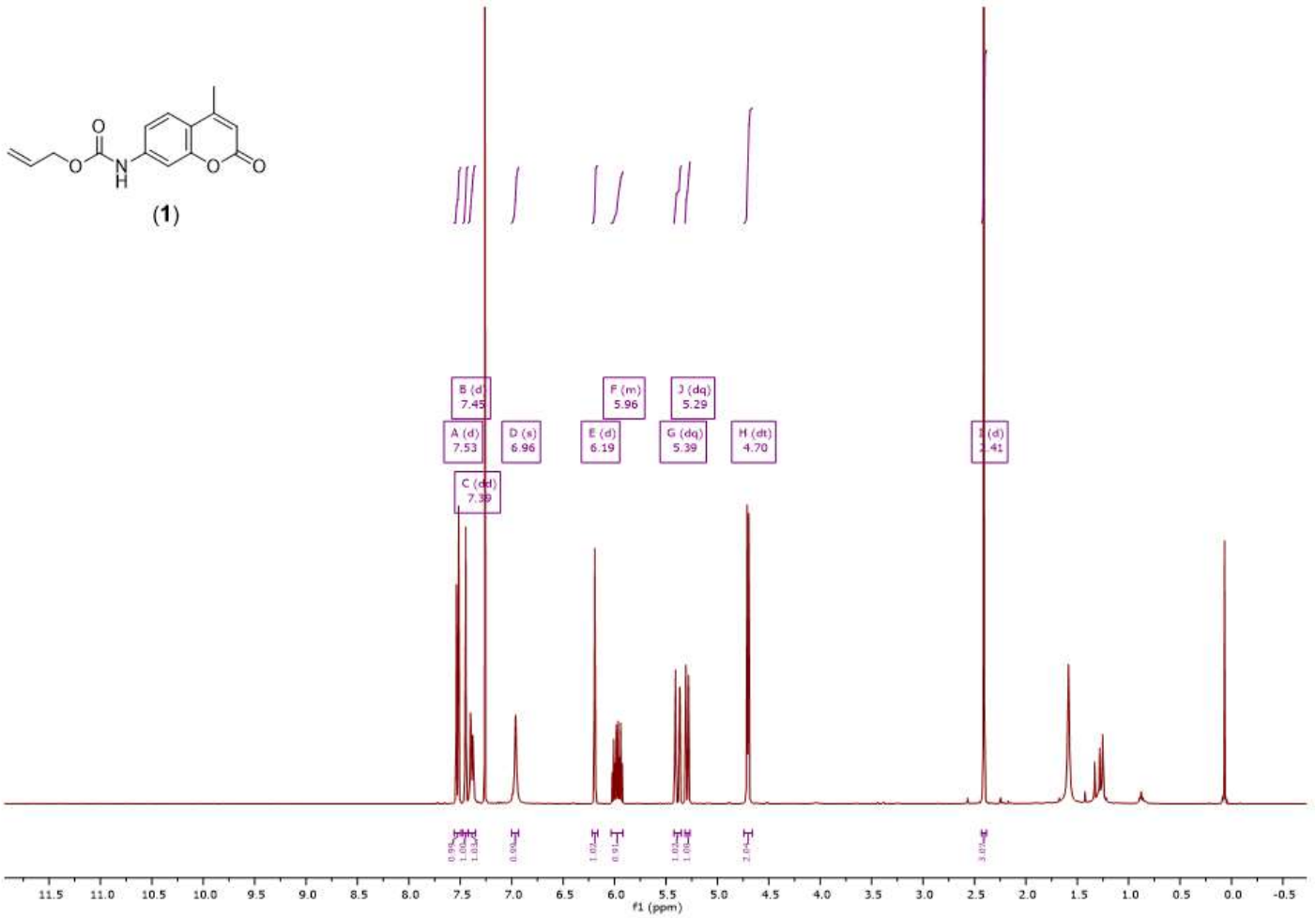
Compound (4): CA-PEG$_3$-Linker



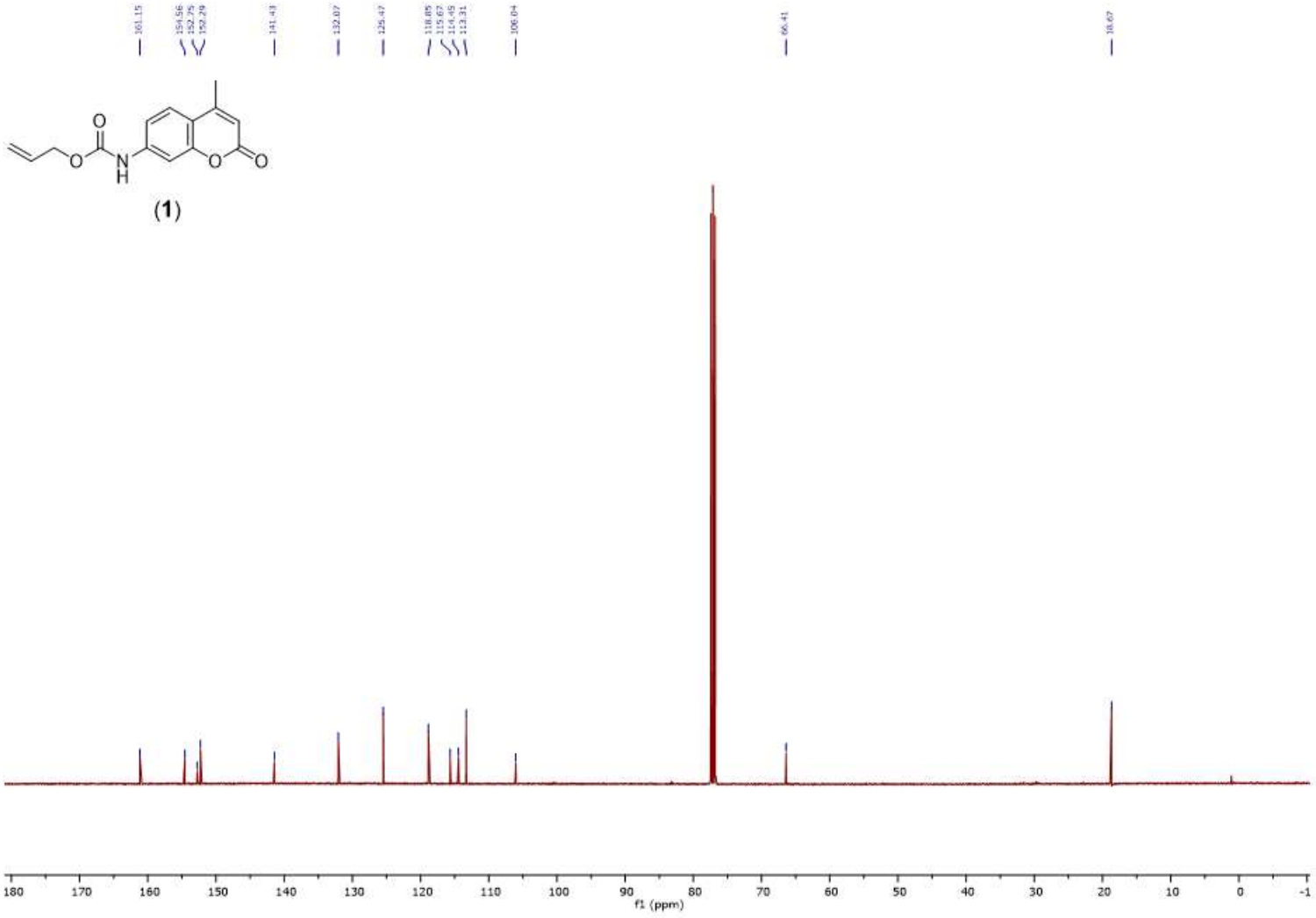
b. NMR files

Compound (1): Alloc- AMC

^{1}H-NMR

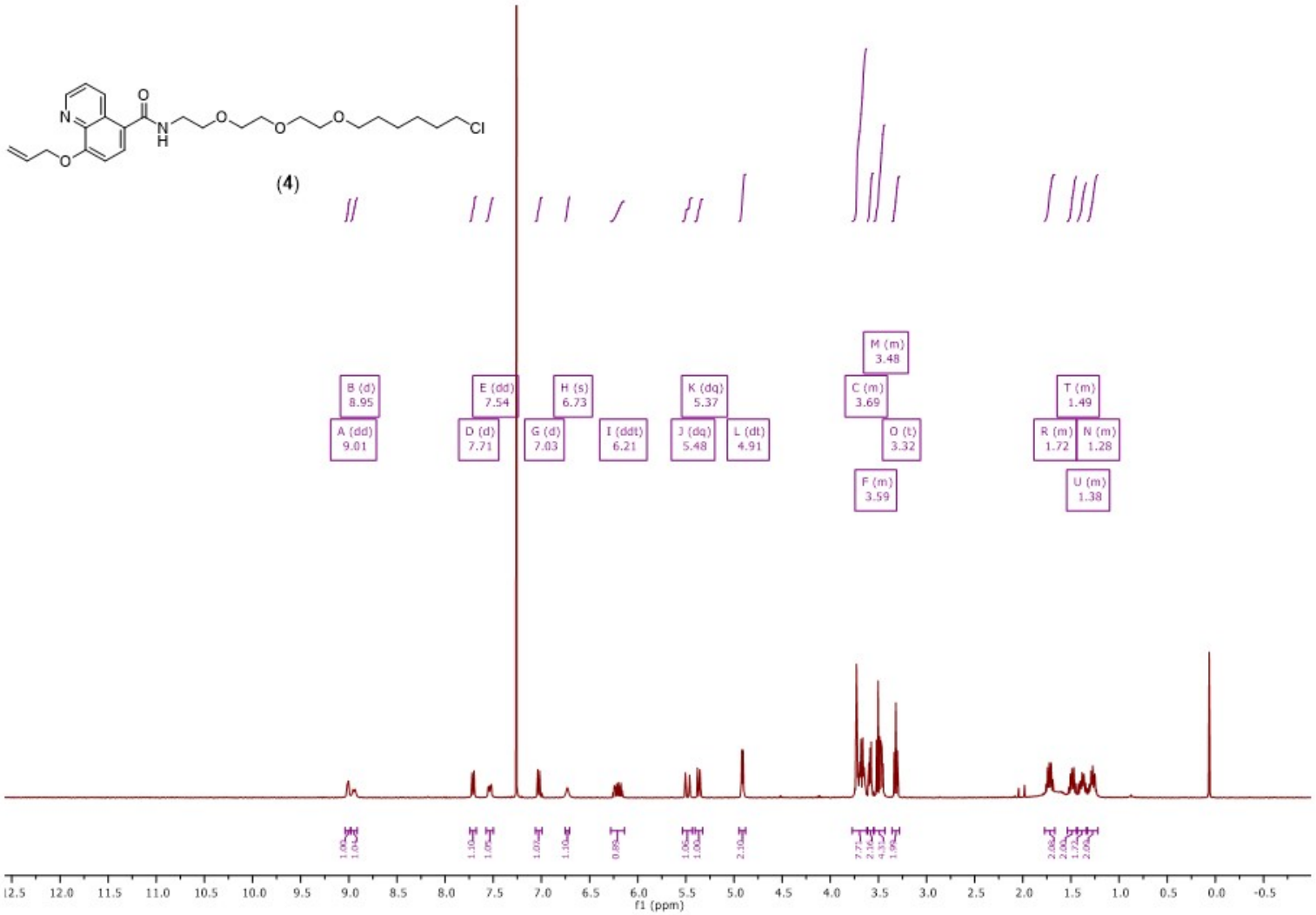


^{13}C-NMR

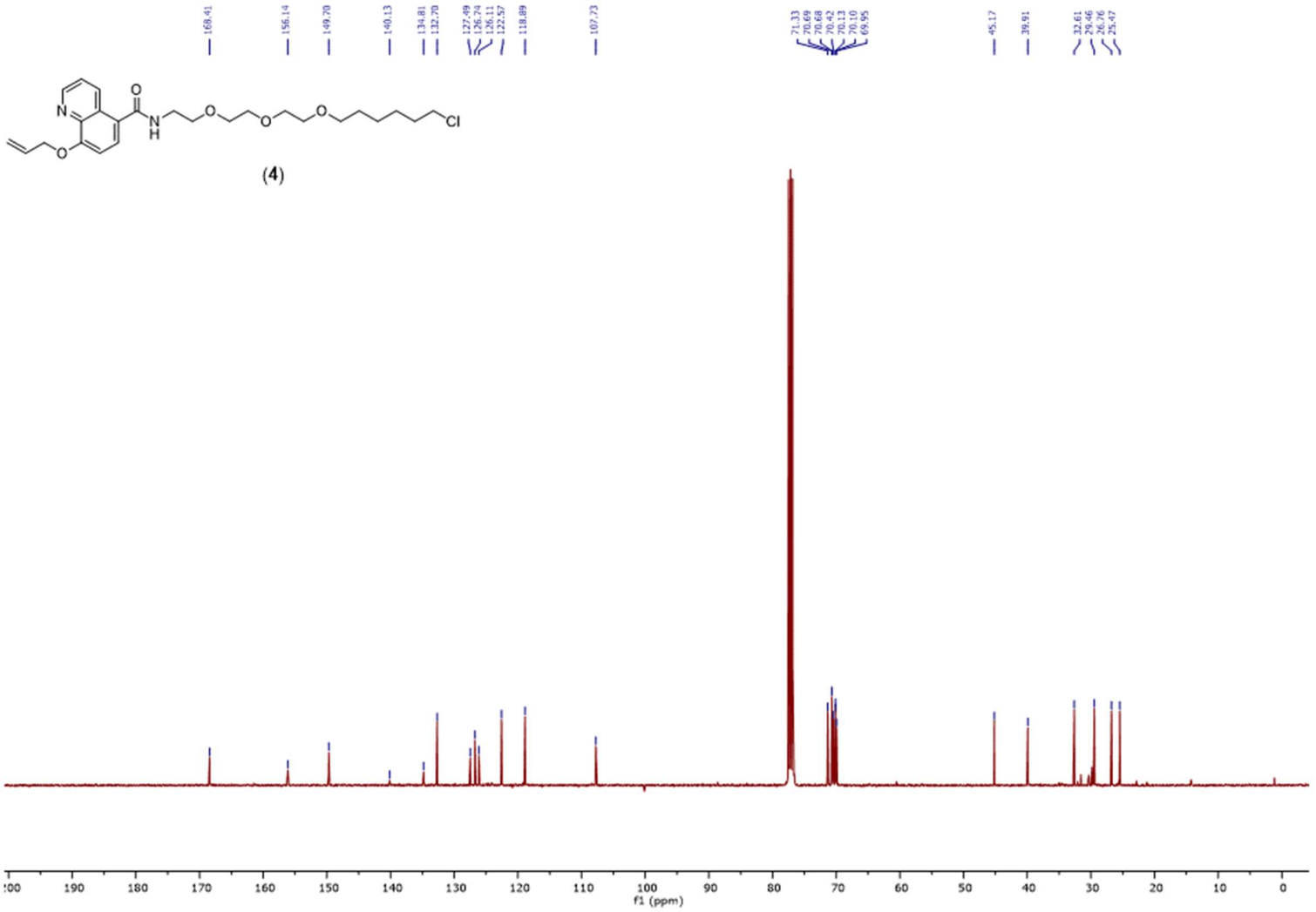

161.15
154.56
152.75
152.29
141.43
132.07
125.47
118.85
115.67
114.45
113.31
106.04
66.41
18.67
(1)
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)

Compound (4): CA-PEG$_3$-Linker

^{1}H-NMR:

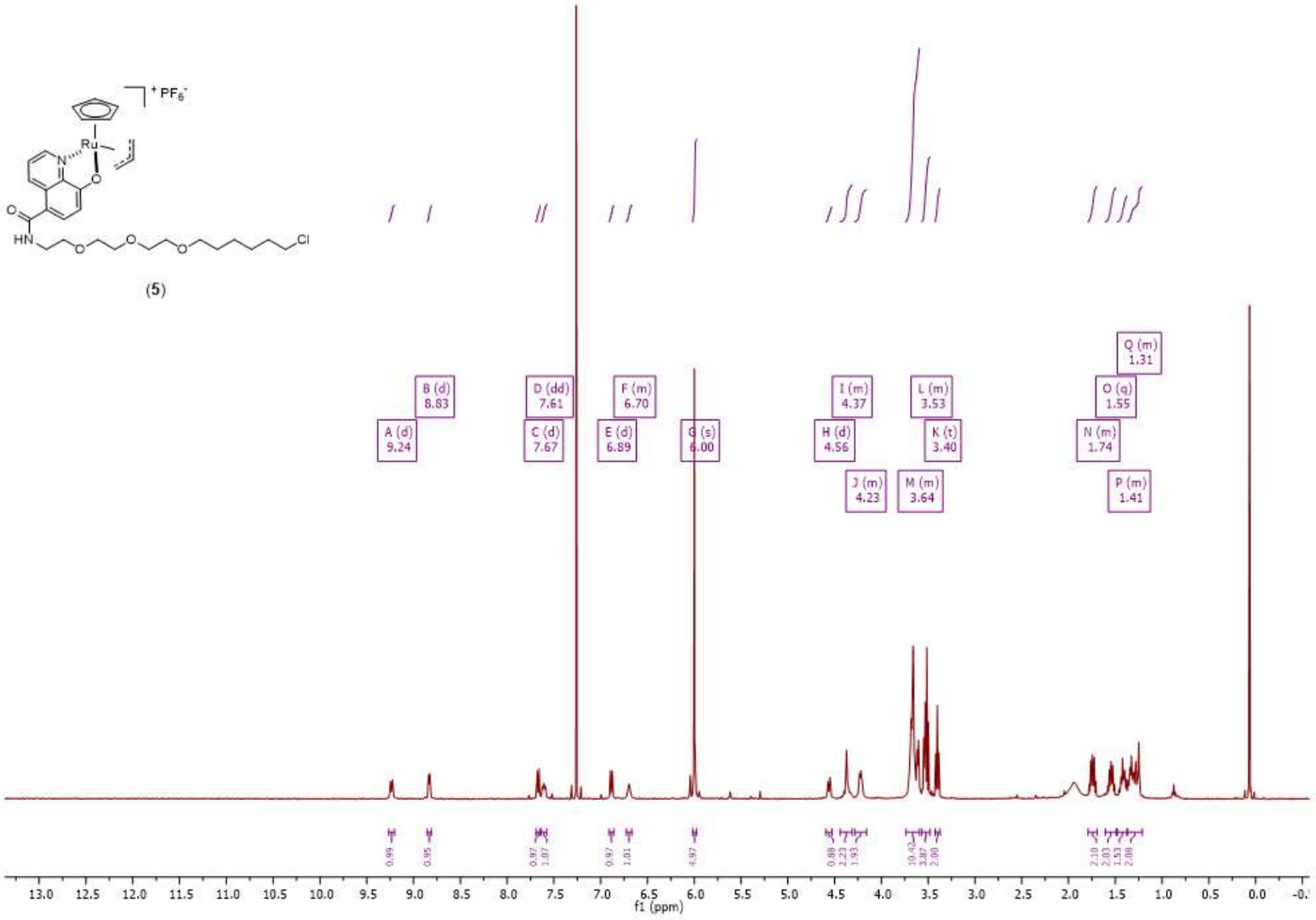


^{13}C-NMR:

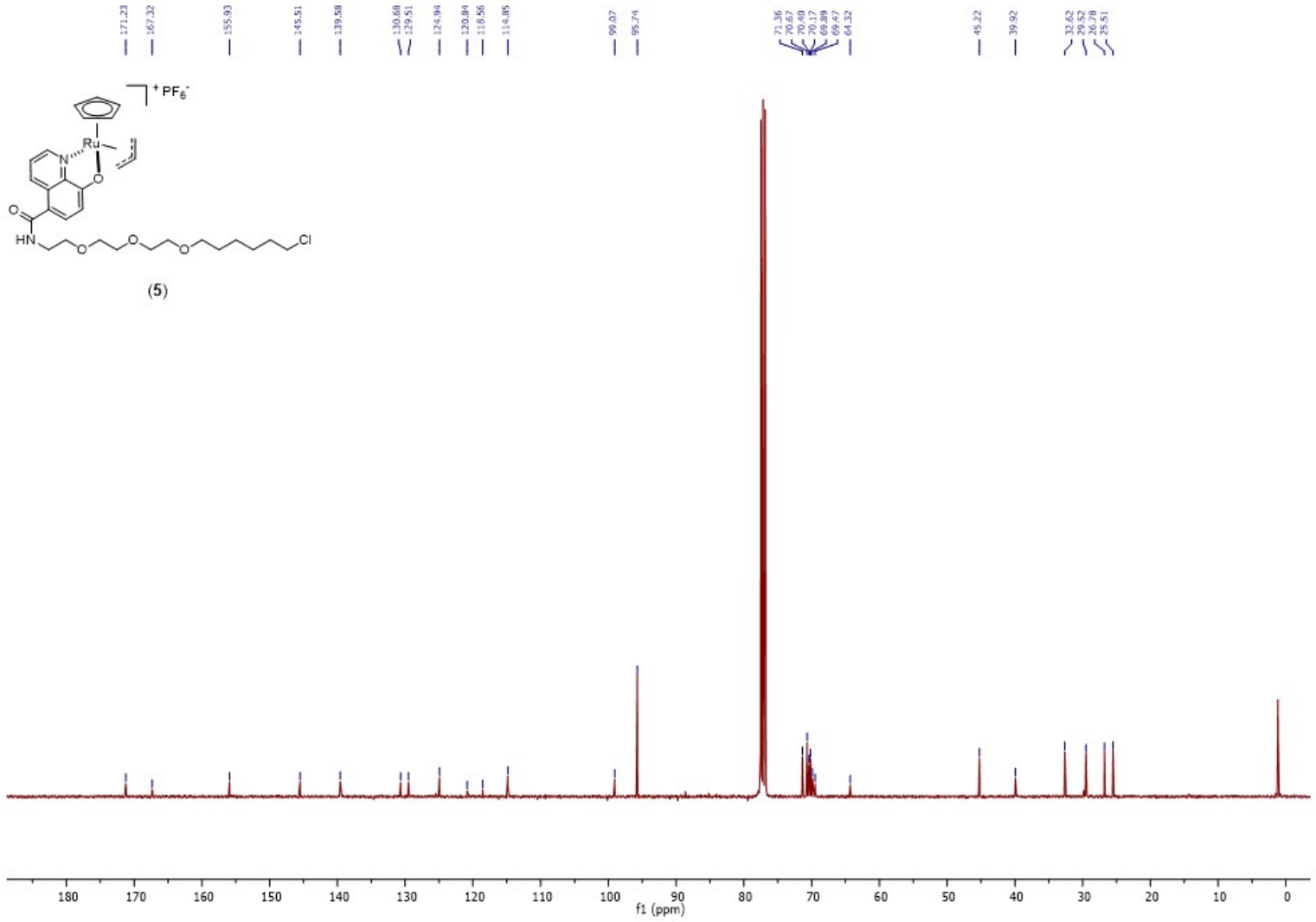

168.41
156.14
149.70
140.13
134.81
132.70
127.49
126.74
126.11
122.57
118.89
107.73
71.33
70.66
70.68
70.42
70.13
70.10
69.95
45.17
39.91
32.61
29.46
26.76
25.47
(4)
f1 (ppm)

Compound (5): CA-PEG$_3$-Ru

^{1}H-NMR:



^{13}C-NMR:


171.23
167.32
155.93
145.51
139.58
130.68
129.51
124.94
120.84
118.56
114.85
99.07
95.74
71.36
70.62
70.40
70.17
69.89
69.47
64.32
45.22
39.92
32.62
29.52
26.78
25.51
+ PF6-
Ru
N
O
HN
O
O
O
Cl
(5)
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

VI. References

[1] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, *Bioinformatics* **2009**, *25*, 1189.
[2] C. Notredame, D. G. Higgins, J. Heringa, *J. Mol. Bio.* **2000**, *302*, 205.
[3] Björn Hellenkamp, Sonja Schmid, Olga Doroshenko, Oleg Opanasyuk, Ralf Kühnemuth, Soheila Rezaei Adariani, Benjamin Ambrose, Mikayel Aznauryan, Anders Barth, Victoria Birkedal et al., *Nat. Methods* **2018**, *15*, 669.
[4] W. Schrimpf, A. Barth, J. Hendrix, D. C. Lamb, *Biophys. J.* **2018**, *114*, 1518.
[5] A. N. Kapanidis, N. K. Lee, T. A. Laurence, S. Doose, E. Margeat, S. Weiss, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 8936.
[6] A. Coullomb, C. M. Bidan, C. Qian, F. Wehnekamp, C. Oddou, C. Albigès-Rizo, D. C. Lamb, A. Dupont, *Sci. Rep.* **2020**, *10*, 6504.
[7] Klaus C. Gwosch, Jasmin K. Pape, Francisco Balzarotti, Philipp Hoess, Jan Ellenberg, Jonas Ries, Stefan W. Hell, *Nat. Methods* **2020**, *17*, 217.
[8] J. K. Pape, T. Stephan, F. Balzarotti, R. Büchner, F. Lange, D. Riedel, S. Jakobs, S. W. Hell, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 20607.
[9] R. Cao-Milán, L. D. He, S. Shorkey, G. Y. Tonga, L.-S. Wang, X. Zhang, I. Uddin, R. Das, M. Sulak, V. M. Rotello, *Mol. Syst. Des. Eng.* **2017**, *2*, 624.
[10] T. Völker, E. Meggers, *Chembiochem* **2017**, *18*, 1083.