SUPPORTING INFORMATION

Transglutaminase-Catalyzed Bioconjugation Using One-Pot Metal-Free Bioorthogonal Chemistry

Natalie M. Rachel, [†,§,⊥] Jacynthe L. Toulouse, [‡, §,⊥] and Joelle N. Pelletier,*[,†,‡, §,⊥]

[†] *Department of Chemistry, and* [‡]*Department of Biochemistry, Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada*

[§] *PROTEO, the Québec Network for Protein Function, Engineering and Applications, Québec, G1V 0A6, Canada*

[⊥] *CGCC, the Center in Green Chemistry and Catalysis, Montréal, Québec, H3A 0B8, Canada*

Corresponding author:
Joelle N. Pelletier
joelle.pelletier@umontreal.ca

Table of Contents

Materials p. S2
Methods p. S2
Supplemental Tables: p. S7 to S9
Supplemental Figures: p. S10 to S13
High Resolution Mass Spectra: p. S14 to S23
References: p. 24

Materials

The plasmid pDJ1-3 was kindly provided by Professor M. Pietzsch (Martin-Luther-Universität, Halle-Wittenberg, Germany). pDJ1-3 encodes the proenzyme of MTG from *Streptoverticillium mobaraense* inserted between the *Nde*I and *Xho*I restriction sites of the vector pET20b.[1] Deionized water (18Ω) was used for all experiments. HPLC solvents were of analytical grade, and products used for the expression and purification of MTG were of biological grade.

Other chemicals used were purchased from the suppliers listed below. Carboxybenzyl–L-glutaminyl–glycine (Z-Gln-Gly, or ZQG) was from Peptide Institute (Osaka, Japan). Glutathione (reduced) and thiamine were from Bioshop (Burlington, Canada). Dimethyl sulfoxide (99.7%) was purchased from Fisher Scientific (Ontatio, Canada). Dibenzylcyclooctyne-PEG4-Amine, methyltetrazine-PEG4-Amine, *trans*-cyclooctene-Cy5 were purchased from Click Chemistry Tools (Arizona, USA). Formic acid (98 % purity) was from Fluka Analytical (St. Louis, USA). α-lactalbumin from bovine milk (calcium saturated), cadaverine dihydrochloride (98%), and 2-(Diphenylphosphino)terephthalic acid 1-methyl 4-pentafluorophenyl diester were purchased from Sigma Aldrich (St. Louis, MO, USA). Sulfo-cyanine5 azide (Cy5-azide) was purchased from Lumiprobe (Hallandale Beach, FL, USA).

MTG Expression and Purification

MTG was expressed and purified as previously described.[2] Briefly, a 5-mL starter culture of *E. coli* BL21 (DE3) containing the plasmid pET20b-MTG, which expresses a *C*-terminally 6-His-tagged version of MTG, was propagated overnight at 37°C in ZYP-0.8G medium and shaking at 240 rpm. It was used to inoculate 500 mL of auto-inducing ZYP-5052 medium. After 2h of incubation at 37°C and 240 rpm, the temperature was reduced to 22°C overnight. Cells

were collected by centrifugation and resuspended in 50 mM sodium phosphate buffer, 300 mM NaCl, pH 7.5. The cells were lysed using a Constant Systems cell disruptor set at 37 kPSI and cooled to 4°C. After further centrifugation to remove insoluble cellular matter, the inactive form of MTG was incubated with trypsin (1 mg/mL solution, 1:9 ratio of trypsin to MTG, v/v) for the purpose of cleaving its pro-sequence. Activated MTG was purified using a 5-mL His-trap nickel-nitrilotriacetic acid (Ni-NTA) column (GE Healthcare) equilibrated in 50 mM phosphate buffer, pH 7.5, with 300 mM NaCl, and eluted with an imidazole gradient (0 – 250 mM) using an Åtka FPLC (GE Healthcare). After purification, active MTG was dialyzed against 50 mM sodium phosphate buffer, 300 mM NaCl, pH 7.5. The average yield was 25 mg of activated MTG per litre of culture, with ~ 85% purity as estimated by SDS-PAGE and revelation with Coomassie blue stain. Aliquots were snap frozen and stored at -80°C in 15% glycerol.

MTG Activity Assay

The activity of purified MTG was quantified using the hydroxamate assay.[4] Briefly, MTG was incubated with 30 mM Z-Gln-Gly and 100 mM hydroxamate at 37°C for 10 min. A concentrated acidic ferric chloride solution (2.0 M $FeCl_3 \cdot 6\ H_2O$, 0.3 M trichloroacetic acid, 0.8 M HCl) was used to quench the reaction, which was then vortexed and left to stand at room temperature for 10 min. The resulting iron complex was quantified by its absorbance at 525 nm. One unit (U) of MTG produces 1 μmol of L-glutamic acid and γ-monohydroxamate per min at 37°C.

hDHFR Expression and Purification

Recombinant human chromosomal DHFR (hDHFR) was overexpressed in Escherichia coli BL21 (DE3) and purified as previously described[3], with the following minor modifications. The expression was done in Terrific Broth. The purification buffer was 10 mM potassium phosphate pH 8.0 for DEAE, and 50 mM potassium phosphate pH 7.5.

Conjugation Assays

All reactions were performed in triplicate, at least twice. Protein substrate (50 μM; α-Lactalbumin, hDHFR), amine, and complementary azide or *trans*-cyclooctene Cy5 substrates (100 μM) were combined with 5 mM glutathione in 100 mM sodium phosphate buffer, pH 7.5. The conjugation reaction catalyzed by MTG was initiated by the addition of 1 U/mL of MTG, where control reactions had an equivalent volume of buffer added. A second control reaction was performed in addition, in which MTG was added, but the amine substrate was omitted. The final volume of each reaction was 300 μL and all were incubated at 37°C for 24h. 50 μL aliquots were taken after 10 min, 1h, 4h, 8h, to which 3 μL of formic acid was added to quench the reaction. The remaining volume for the 24h aliquot was quenched by adding 6 μL of formic acid. Samples were stored at 4°C for short-term storage, or at -20°C for long-term storage if necessary. In the case of hDHFR, slight precipitation of protein occurred during reactions, independently of the presence of MTG or other reagents. A 5 min microcentrifugation allowed removing the insoluble material.

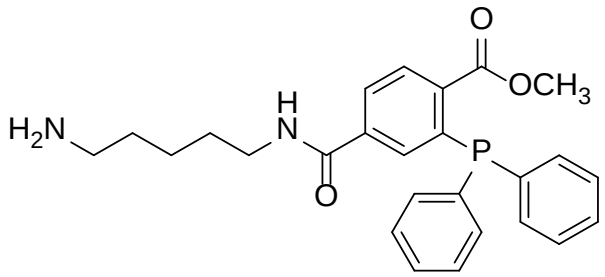
For reactions prepared in a subsequent fashion, the same protocol as above applies, with the exception that the Cy5 substrates were added only after aliquots had been quenched, and then incubated at 4°C overnight prior to analysis.

Aliquots were resolved using tricine SDS-PAGE.[5] The fluorescent bands were visualized and recorded using a Bio Rad ChemiDoc™ MP Imaging System using an excitation filter of 625 nm with a 30 nm bandpass. The gels were then stained with Coomassie brilliant blue to reveal the protein.

HPLC-MS

Samples (10 μL) were injected onto a Aeris peptide 3.6u XB-C18 150 × 2.10 mm LC column (Phenomenex), using an Agilent 1200 series HPLC apparatus, and eluted with a 12-minute 20-80% acetonitrile/H_2O gradient. Masses were detected under positive ionization with a single quadrupole mass detector. Quantification of reagent and product peaks were done by measuring peak heights and corrected for baseline.

Synthesis and Purification of DFFT-CAD



A solution (1 mL) containing DMSO, equimolar (12.5 mM) cadaverine dihydrochloride and 2-(Diphenylphosphino)terephthalic acid 1-methyl 4-pentafluorophenyl diester was prepared. After mixing, the solution was left to sit for 48h in the dark at 4°C. The reaction progress was checked using HPLC-MS. 5 μL was transferred to 495 μL of 18.2 mΩ deionized water containing 0.1% formic acid, and mixed by pipetting. Diluted sample was injected (5 μL) onto a

Synergi 4-μm polar-RP 80 Å, 150 × 4.60 mm LC column (Phenomenex), using an Agilent 1200 series HPLC apparatus and eluted with a 12 minute 5-70% MeOH/H_2O gradient. Masses were detected under positive ionization with a single quadrupole mass detector. The reaction had not gone to completion, and so additional cadaverine (to a total of 25 mM) was added to push the consumption of the phosphine substrate. The reaction was incubated once more for 48h at 4°C, and injected again on the LC-MS. When reaction progress was observed to be sufficient, the reaction mixture was injected onto preparative LC-MS. A 200 μL volume of the DFFT-CAD solution was injected onto a Synergi polar-RP 80 Å, 100 × 21.20 mm AXIA-packed column (Phenomenex), using a Waters 1525 HPLC and Waters 3100 single quadrupole mass detector. Elution was performed with a 12-minute 10-70% MeOH/H_2O gradient. Compounds with the mass corresponding to the expected product were collected and pooled. Methanol was evaporated, the remaining solution was lyophilized to yield the purified product.

Table S1: Fluorescence intensities for reactions with α-LA after 10 min of reaction.[a]

Reaction	Order of addition	Controls		Samples[b]
		No MTG ($\times 10^4$)	No NH_2 ($\times 10^4$)	($\times 10^4$)
Staudinger	Simultaneous	3.4	4.1	34 ± 7
	Subsequent	320	220	200 ± 52
SPAAC	Simultaneous	110	350	1900 ± 150
	Subsequent	70	200	3600 ± 190
Tetrazine	Simultaneous	28	75	5900 ± 480
	Subsequent	160	490	2800 ± 67

[a] Fluorescence intensities were quantified by Image Lab™ on the SDS-PAGE gels of samples taken after 10 min of reaction; [b] Average of triplicate samples.

Table S2: Fluorescence intensities for reactions with α-LA after 4 h of reaction.[a]

Reaction	Order of addition	Controls		Samples[b]
		No MTG ($\times 10^4$)	No NH_2 ($\times 10^4$)	($\times 10^4$)
Staudinger	Simultaneous	2.9	34	210 ± 17
	Subsequent	450	331	1400 ± 100
SPAAC	Simultaneous	850	543	6800 ± 910
	Subsequent	120	177	5700 ± 250
Tetrazine	Simultaneous	45	29	9800 ± 850
	Subsequent	670	407	7100 ± 400

[a] Fluorescence intensities were quantified by Image Lab™ on the SDS-PAGE gels of samples taken after 4 h of reaction; [b] Average of triplicate samples.

Table S2: Fluorescence intensities for reactions with hDHFR after 4 h of reaction.[a]

Reaction	Order of addition	Controls		Samples[b] ($\times 10^4$)
		No MTG ($\times 10^4$)	No NH_2 ($\times 10^4$)	
Staudinger	Simultaneous	310	96	450 ± 84
	Subsequent	3400	59	4200 ± 390
SPAAC	Simultaneous	160	95	210 ± 50
	Subsequent	52	16	130 ± 10
Tetrazine	Simultaneous	30	150	1470 ± 187
	Subsequent	296	380	1640 ± 173

[a] Fluorescence intensities were quantified by Image Lab™ on the SDS-PAGE gels of samples taken after 4 h of reaction; [b] Average of triplicate samples.

Table S4: Fluorescence intensities for reactions with α-LA and hDHFR performed in subsequent format where the click reactions are performed prior to addition of MTG and the protein substrates.[a]

Reaction	Time	Controls		Samples[b]
		No MTG ($\times 10^4$)	No NH_2 ($\times 10^4$)	
α-LA, SPAAC	10 min	300	285	1860 ± 130
α-LA, SPAAC	4 h	521	305	5520 ± 170
hDHFR, SPAAC	4 h	191	50	750 ± 140
α-LA, Tetrazine	10 min	340	52	2320 ± 230
α-LA, Tetrazine	4 h	746	79	5590 ± 280
hDHFR, Tetrazine	4 h	1530	644	3050 ± 210

[a] Fluorescence intensities were quantified by Image Lab™ on the SDS-PAGE gels of samples taken after 4 h of reaction; [b] Average of triplicate samples.

Table S5: Quantification by LC of α-LA conjugation with 3 and 5 by MTG.

	Peak Height (mAU)		
	10 min	**4 h**	**24 h**
Control	750	750	710
Native α-LA[a]	410 ± 10	370 ± 15	320 ± 19
Conjugated α-LA[a]	150 ± 2	180 ± 8	210 ± 12

[a] Average of triplicate samples.

Table S6: Quantification by LC of α-LA conjugation with 3 by MTG.

	Peak Height (mAU)		
	10 min	**4 h**	**24 h**
Control	730	780	700
Native α-LA[a]	480 ± 18	400 ± 31	330 ± 26
Conjugated α-LA[a]	110 ± 5	200 ± 17	230 ± 17

[a] Average of triplicate samples.

Table S7: Relative yields and rates by LC of α-LA conjugation by MTG.

Conjugated α-LA, reagents present	Relative Yield (%)			Rate (nM/s)		
	10 min	**4 h**	**24 h**	**10 min**	**4 h**	**24 h**
3 and **5**	20	24	30	33	1.6	0.4
3	15	26	33	25	1.8	0.4

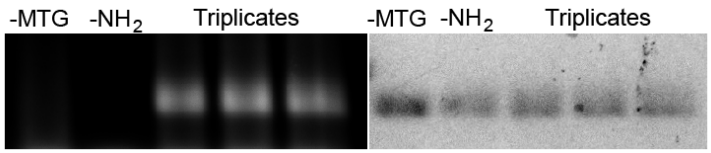


Figure S1: One-pot chemoenzymatic labelling of α-LA using the SPAAC in the absence of glutathione. After 24h, the reactions were quenched as usual with formic acid, and resolved using denaturing tricine SDS-PAGE. The gels were excited with a Cy5 imaging filter, photographed (left panel), and then stained with Coomassie blue to reveal the presence of α-LA (right panel).

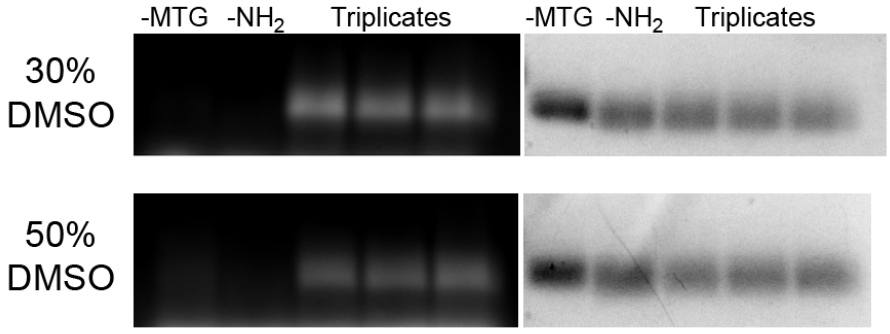


Figure S2: One-pot chemoenzymatic labelling of α-LA using the SPAAC in presence of either 30% or 50% DMSO. After 24h, the reactions were quenched as usual with formic acid, and resolved using denaturing tricine SDS-PAGE. The gels were excited with a Cy5 imaging filter, photographed (left panel), and then stained with Coomassie blue to reveal the presence of α-LA (right panel).

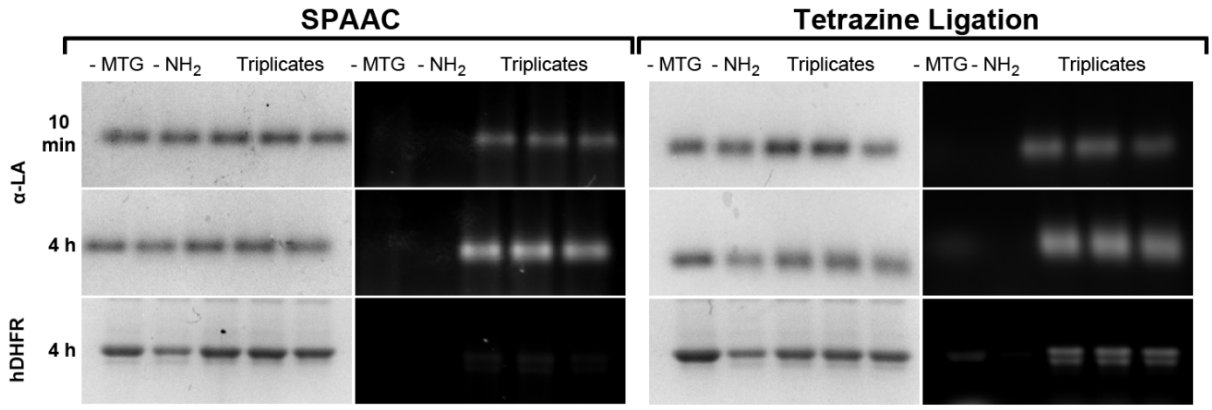


Figure S3. One-pot chemoenzymatic protein labeling of α-LA and hDHFR catalyzed by MTG, performed in subsequent format where the click reactions are performed prior to addition of MTG and the protein substrates. Reagents **2** and **4**, or **3** and **5**, were incubated in buffer at 37°C for 4 h to allow the formation of amines **7** and **8**, respectively. After this initial incubation, protein, glutathione, and MTG were added to the solution. Once all reagents were present, aliquots were taken after 10 min and 4 hours, quenched by the addition of formic acid and resolved by tricine SDS-PAGE. The negative controls lacked either MTG (- MTG) or the amine substrate (- NH_2) and were run in parallel. Left-hand panels: Coomassie brilliant blue staining, performed after the gel was excited with a Cy5 filter to detect fluorescence (right-hand panels).

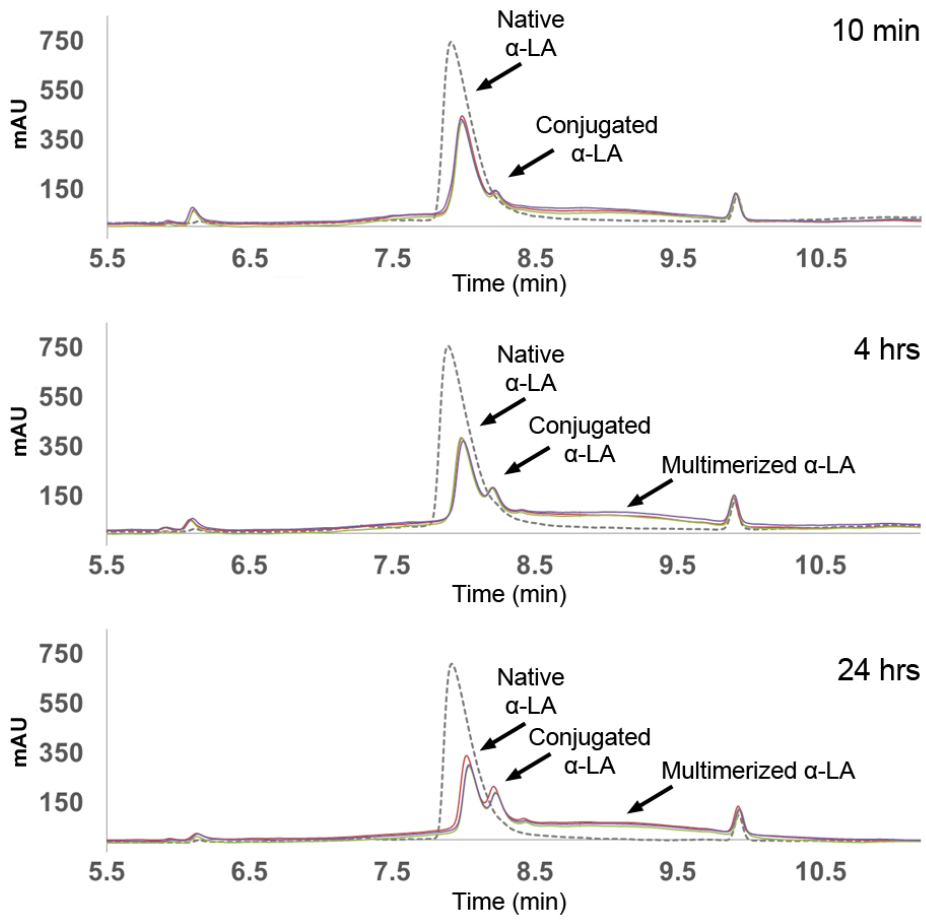


Figure S4. LC chromatograms of α-LA upon conjugation with **3** and **5** by MTG. Detection was performed at 214 nm. Grey dashed traces and colored traces correspond to a negative control lacking MTG and reaction triplicates, respectively.

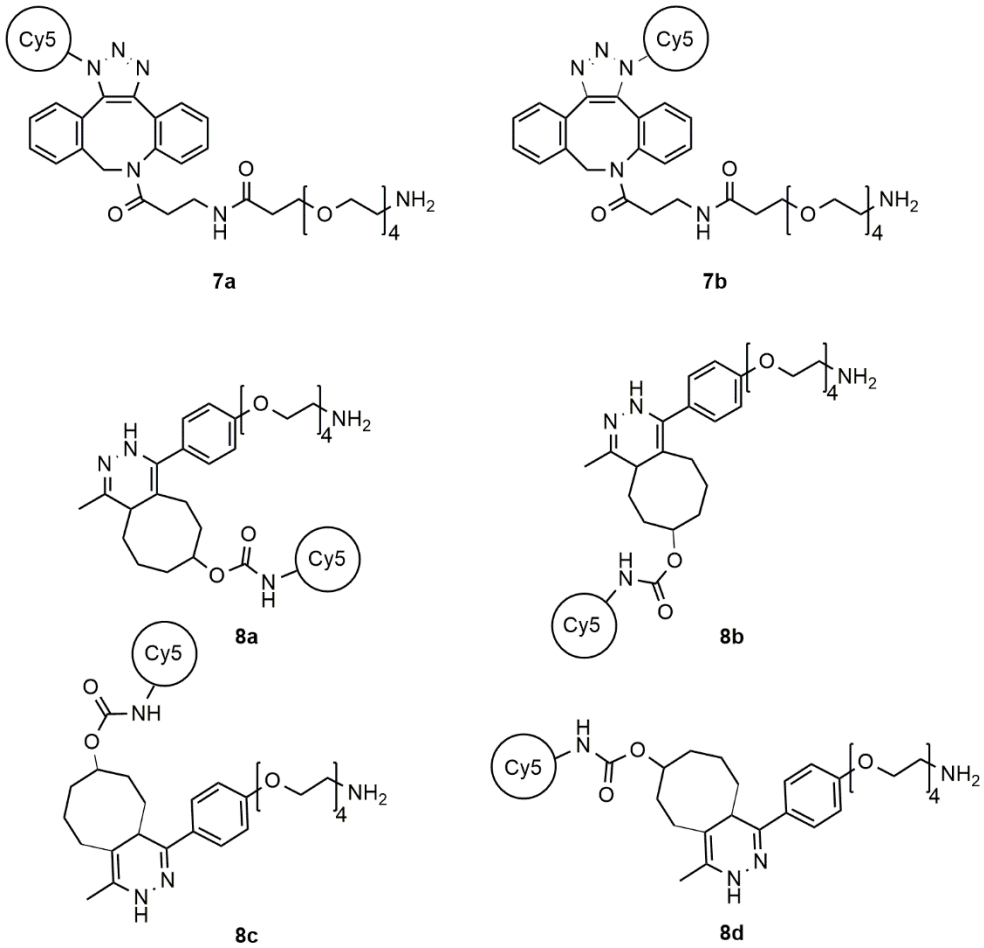


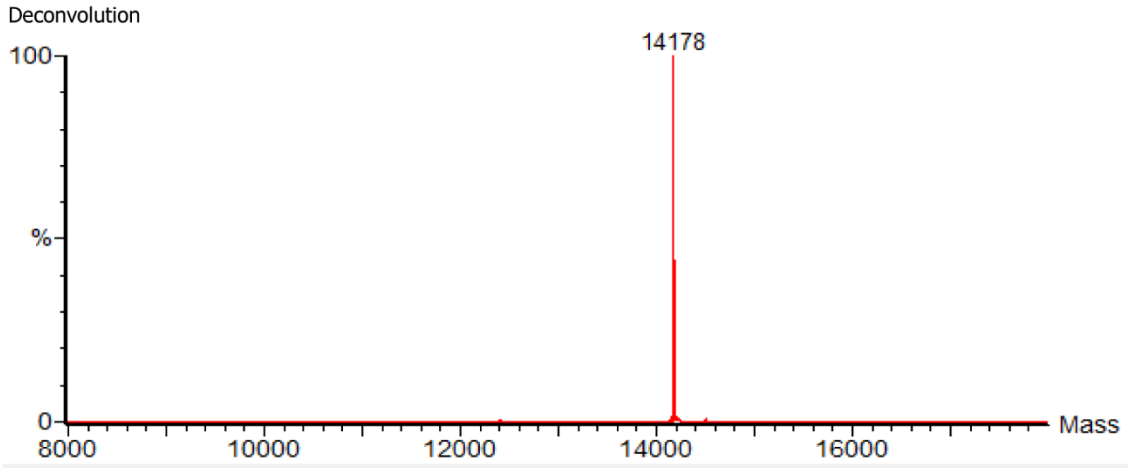
Figure S5: Alternate regiochemical outcomes for reaction of **2** and **4** (**7a** and **7b**) and of **3** and **5** (**8a-d**).

High Resolution MS Spectra of Conjugated Protein Products

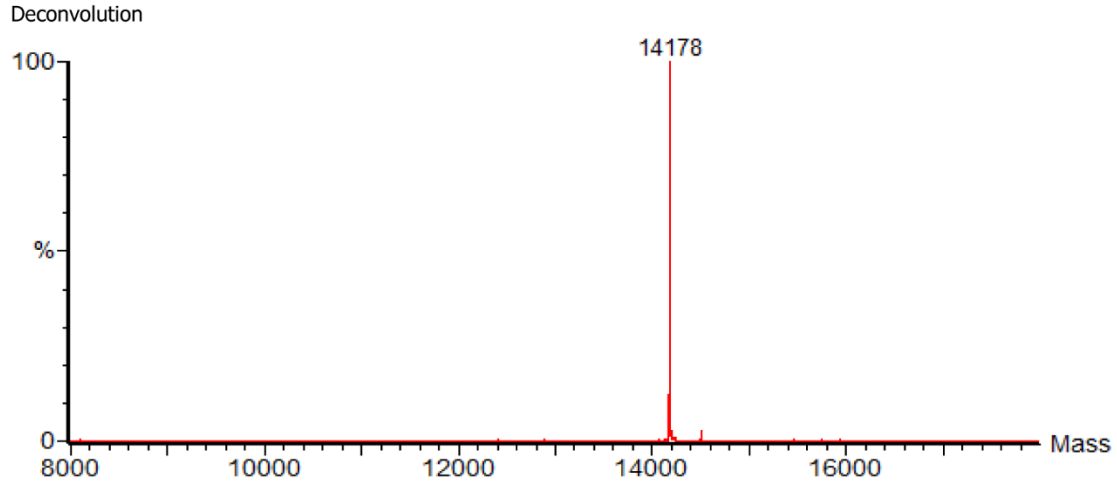
Each page displays a set of reactions, for which there are two spectra: the first is the control in which MTG is absent, preventing conjugation. The second is the reaction. For the latter, two potential products can be observed: the protein conjugated with its amine (compounds **1-3** in Figure 1 in the main text), and/or the protein conjugated with its amine after it has been clicked with its corresponding probe (compounds **4** or **5** in Figure 1). The masses for these products are calculated in the captions.

In some spectra, many weaker peaks are observed. They are observed in our control reaction as well as our reaction sample, and for this reason, were not listed; we listed only the most prominent peaks.

➔ ***SPAAC on α-Lactalbumin***

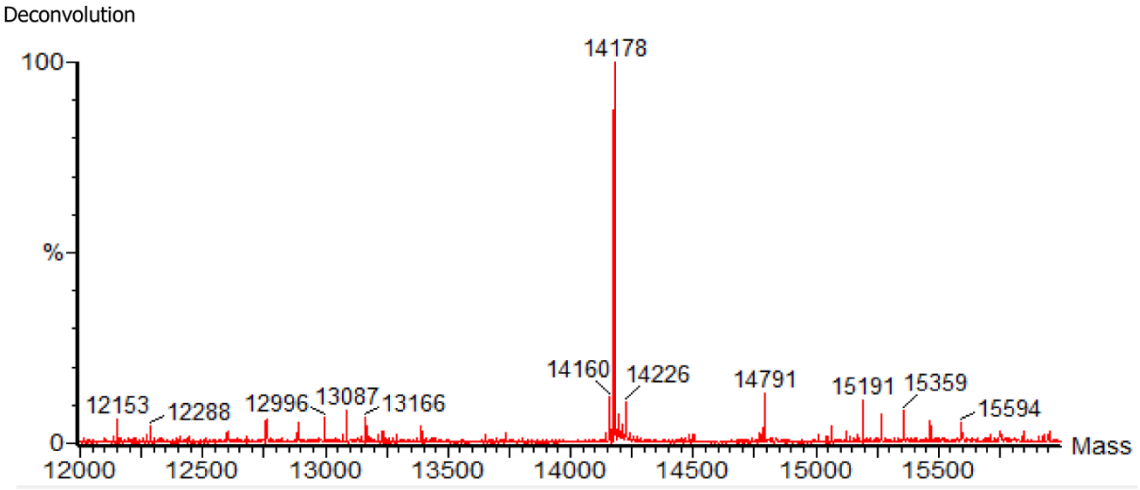


HRMS (ESI) m/z calculated for control containing α-Lactalbumin + 2, no MTG: 14178; found 14178.

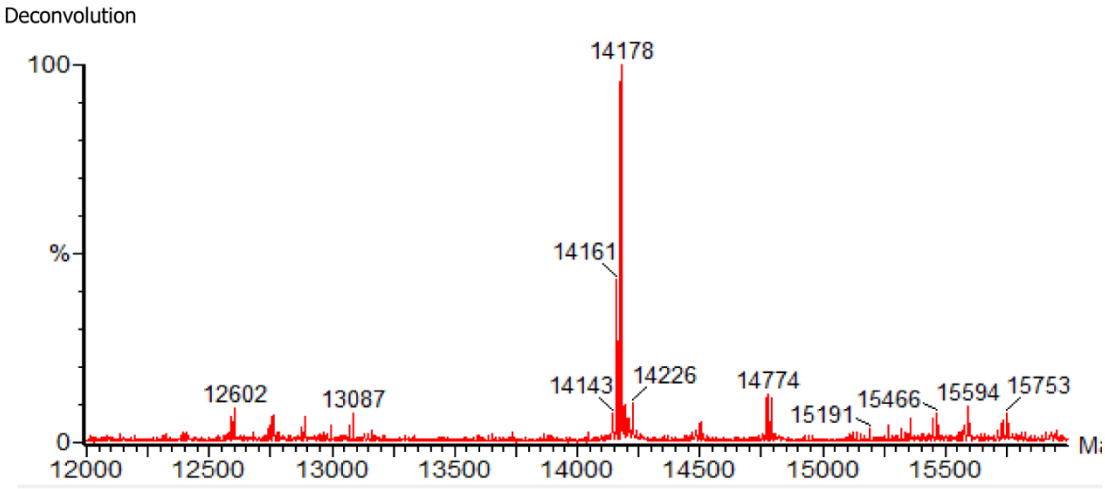


HRMS (ESI) m/z calculated for reaction containing α-Lactalbumin + 2: 14685; m/z calculated for reaction containing α-Lactalbumin + 2 + 4: 15433. Found: 14178.

➔ ***SPAAC on α-Lactalbumin in 30% DMSO***

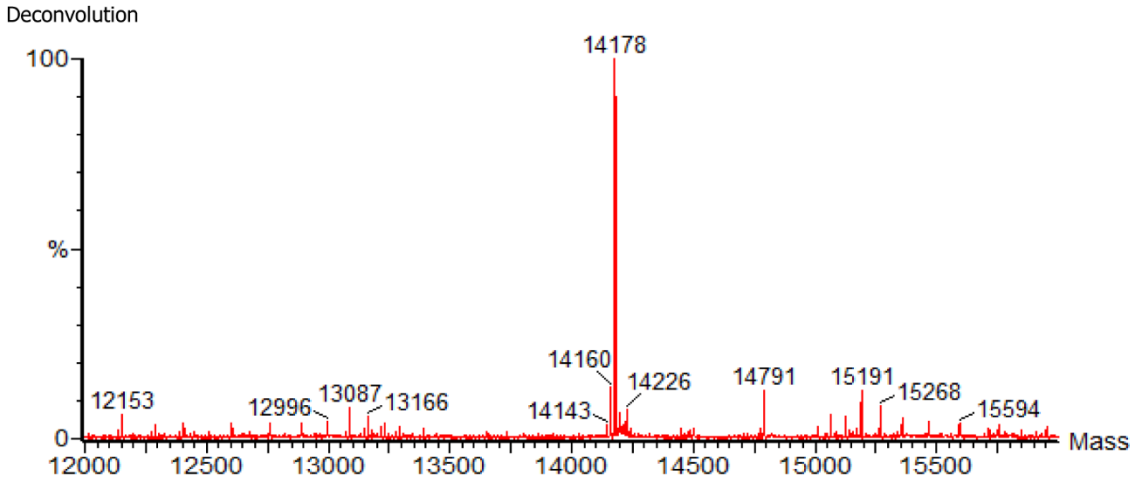


HRMS (ESI) m/z calculated for control containing 30% DMSO, α-Lactalbumin + 2, no MTG: 14178; found 14178.

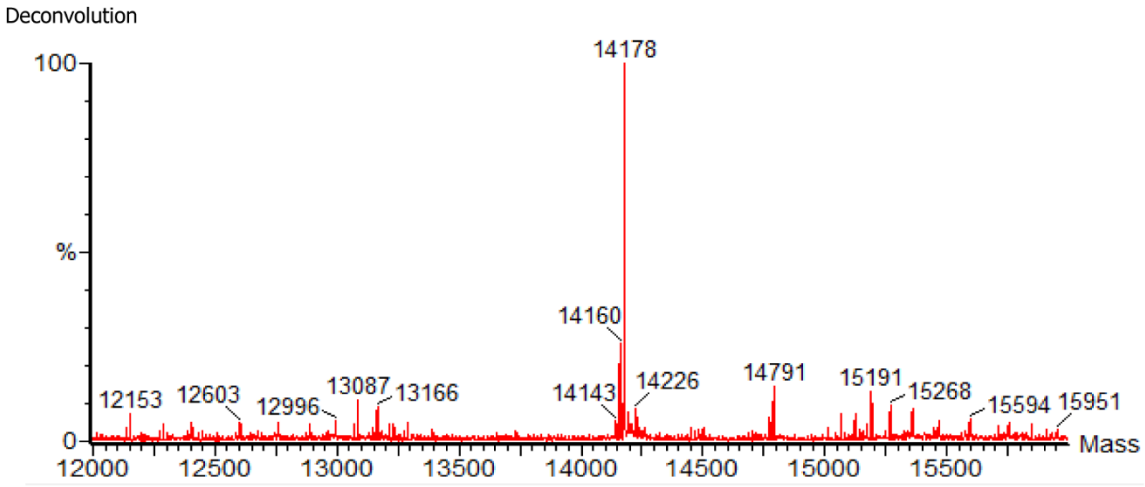


HRMS (ESI) m/z calculated for reaction containing 30% DMSO, α -Lactalbumin + 2: 14684; m/z calculated for reaction containing α-Lactalbumin + 2 + 4: 15433. Found 14178.

➔ ***SPAAC on α-Lactalbumin in 50% DMSO***

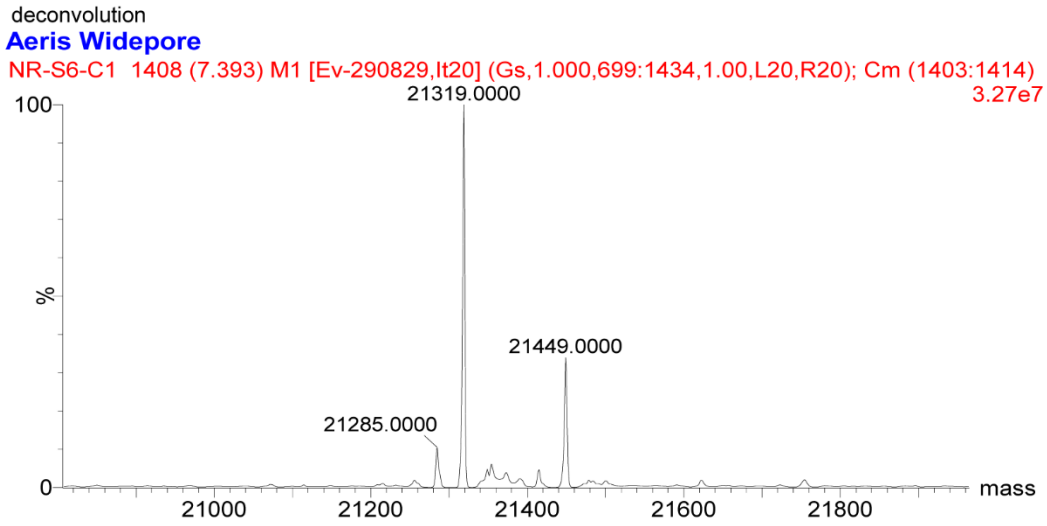


HRMS (ESI) m/z calculated for control containing 50% DMSO, α-Lactalbumin + 2, no MTG: <u>14178</u>; found <u>14178</u>.

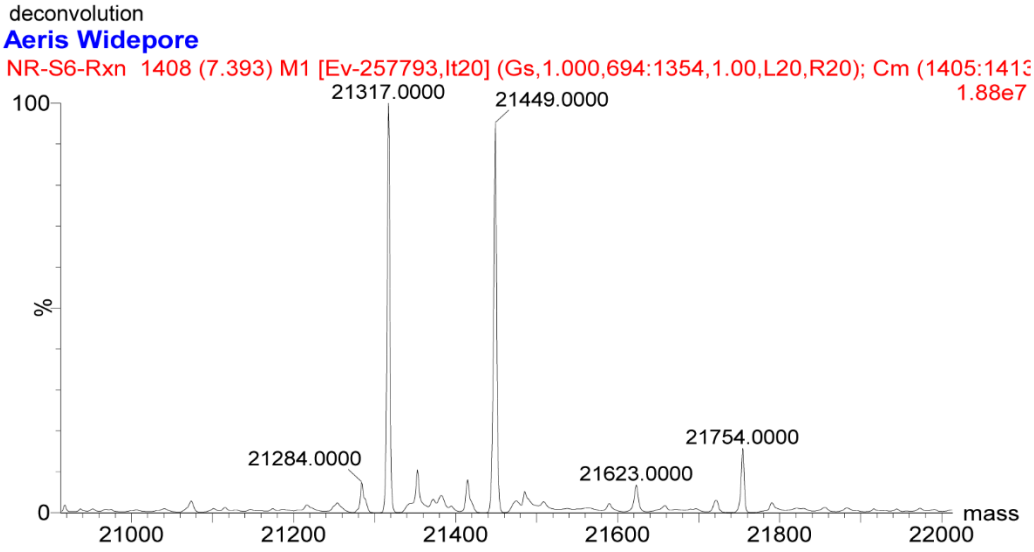


HRMS (ESI) m/z calculated for reaction containing 50% DMSO, α -Lactalbumin + 2: 14684; m/z calculated for reaction containing α-Lactalbumin + 2 + 4: <u>15433</u>. found <u>14178</u>.

➔ ***SPAAC on hDHFR***

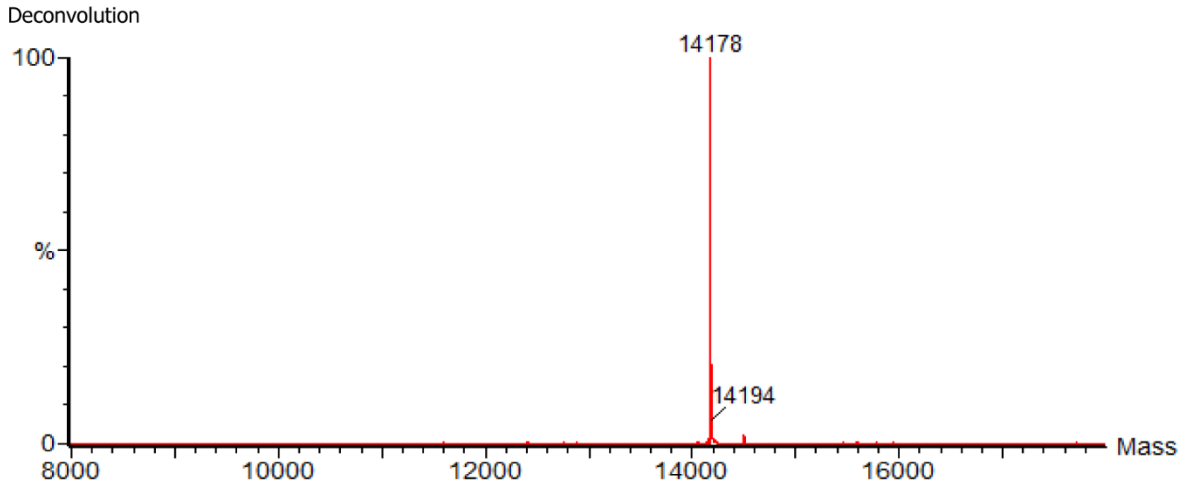


HRMS (ESI) m/z calculated for control containing hDHFR + 2, no MTG: 21324; found 21319.0, 21449.0.

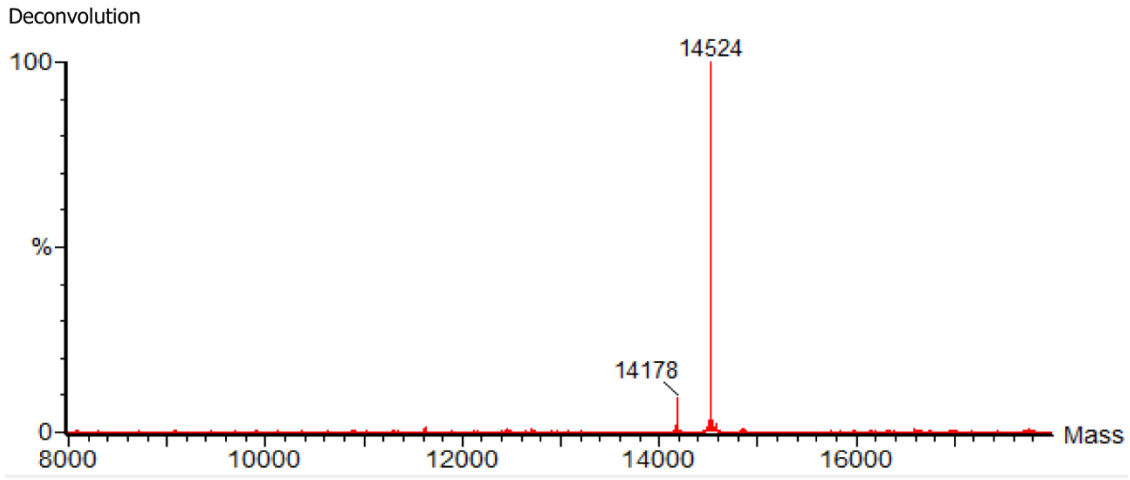


HRMS (ESI) m/z calculated for reaction containing hDHFR + 2: 21829; m/z calculated for reaction containing hDHFR + 2 + 4; 22574. Found 21317.0, 21449.0.

➔ ***TL on α-Lactalbumin***

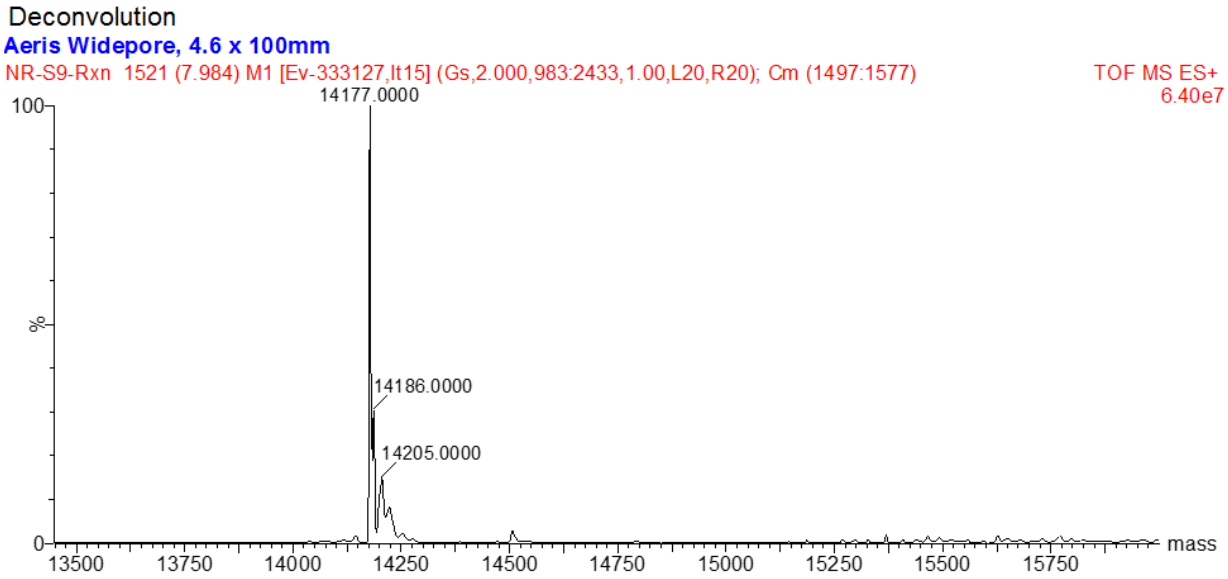


HRMS (ESI) m/z calculated for control containing α-Lactalbumin + 3, no MTG: <u>14178</u>; found <u>14178</u>.

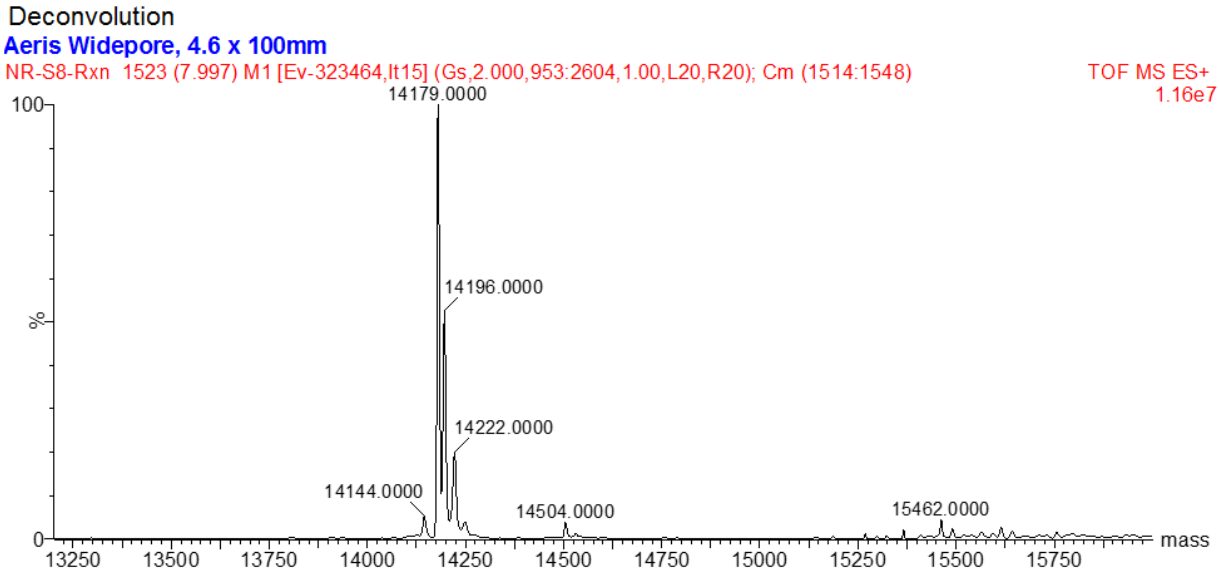


HRMS (ESI) m/z calculated for reaction containing α-Lactalbumin + 3: <u>14524</u>; found <u>14524</u>, <u>14178</u>.

➔ ***TL on α-Lactalbumin***

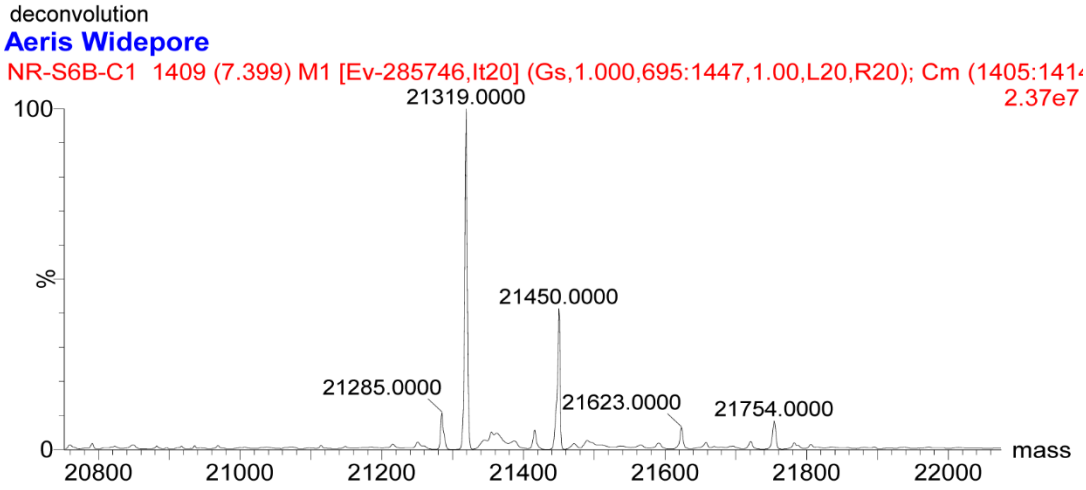


HRMS (ESI) m/z calculated for control containing hDHFR + 3 + 5, no MTG: 14178; found 14177.

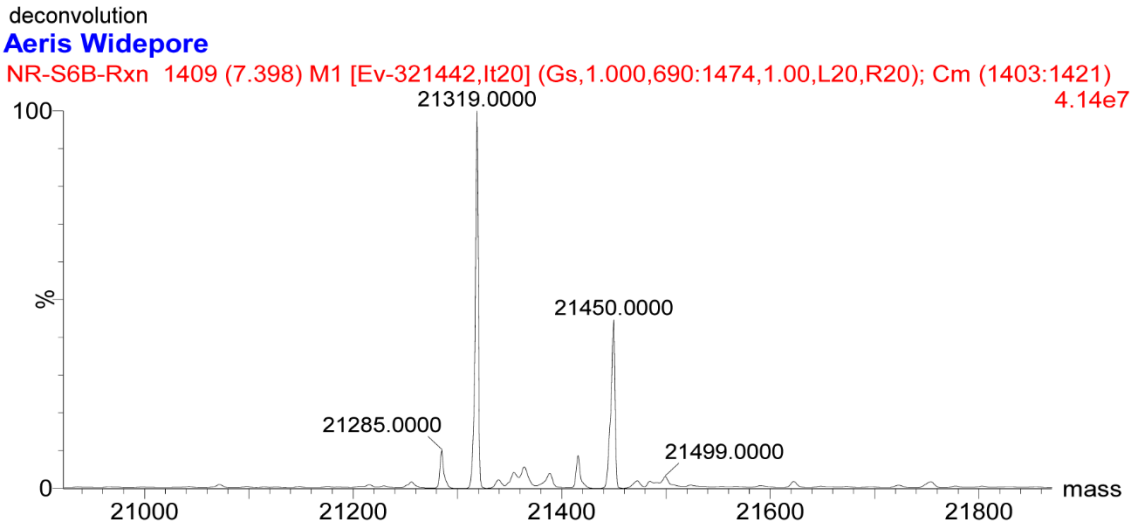


HRMS (ESI) m/z calculated for control containing hDHFR + 3 + 5, no MTG: 15460; found 15462.

➔ ***TL on hDHFR***

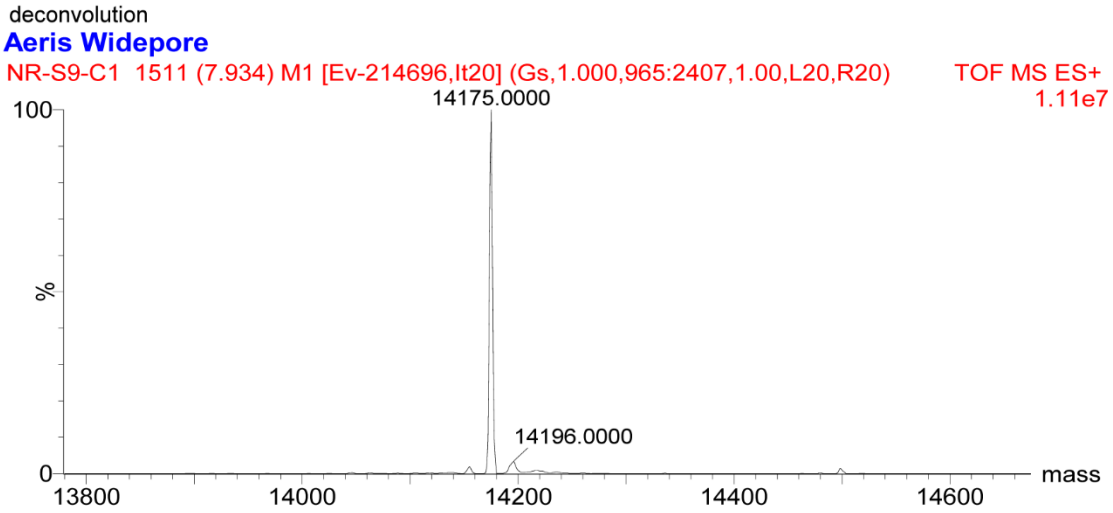


HRMS (ESI) m/z calculated for control containing hDHFR + 3, no MTG: 21324; found 21319.0, 21450.0.

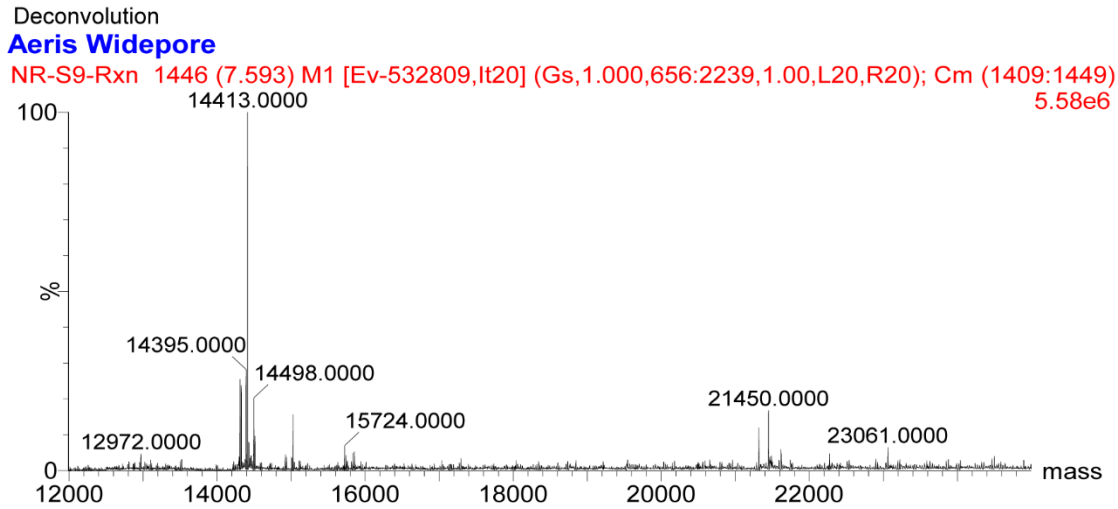


HRMS (ESI) m/z calculated for reaction containing hDHFR + 3: 21668; found 21319.0, 21450.0.

➔ ***Staudinger ligation on α-Lactalbumin***

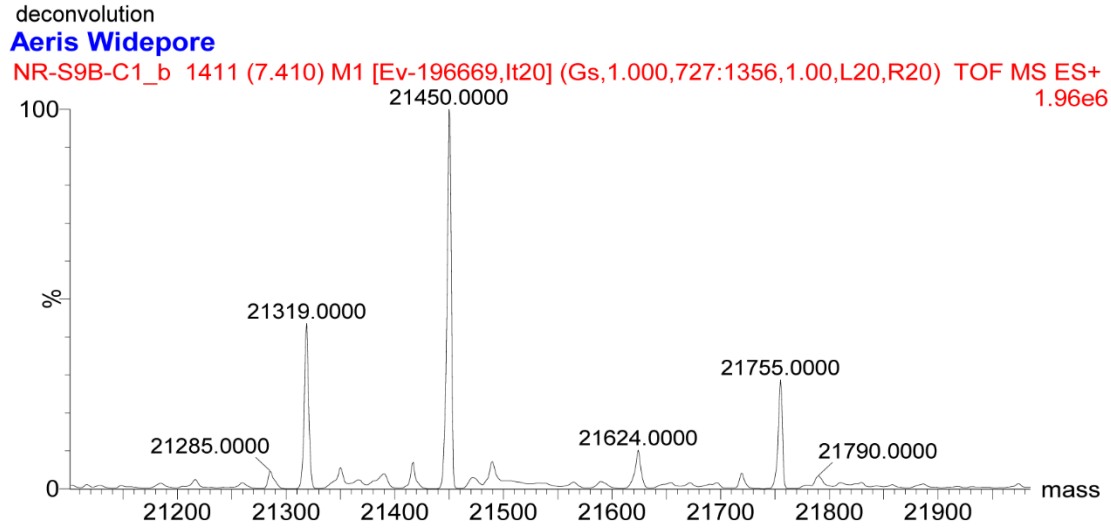


HRMS (ESI) m/z calculated for control containing α-Lactalbumin + 1, no MTG: 14178; found 14175.0.

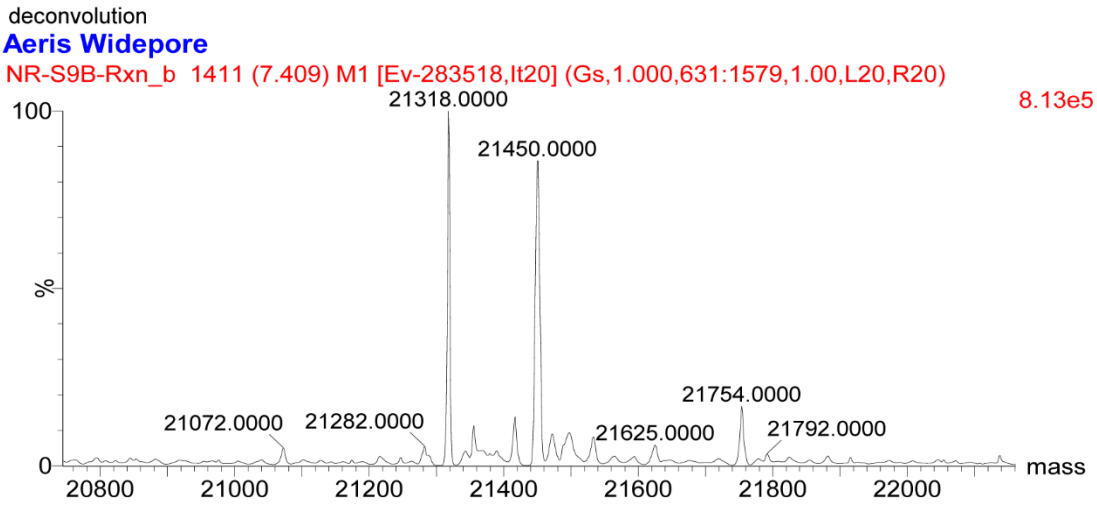


HRMS (ESI) m/z calculated for reaction containing α-Lactalbumin + 1: 14610; m/z calculated for reaction containing α-Lactalbumin + 1 + 4: 15291. Found 14413.0, 14498, 15724.

➔ ***Staudinger ligation on hDHFR***



HRMS (ESI) m/z calculated for control containing hDHFR + 1, no MTG: 21324; found 21319.0, 21450.



HRMS (ESI) m/z calculated for reaction containing hDHFR + 1: 22437; m/z calculated for reaction containing hDHFR + 1 + 4: 23184. Found 21318.0, 21450.

References

(1) Marx C.K., Hertel T.C., Pietzsch M. (2007) Soluble expression of a pro-transglutaminase from Streptomyces mobaranesis in Escherichia coli. *Enzyme Microb. Technol. 40*, 1543-1550.

(2) Gundersen, M.T., Keillor, J.W., Pelletier, J.N. (2014) Microbial transglutaminase displays broad acyl-acceptor substrate specificity. *Appl. Microbiol. Biot*. *98*, 219-230.

(3) Volpato J.P., Yachnin B.J., Blanchet J., Guerrero V., Poulin L., Fossati E., Berghuis, A.M., Pelletier, J.N. (2009) Multiple conformers in active site of human dihydrofolate reductase F31R/Q35E double mutant suggest structural basis for methotrexate resistance. *J. Biol. Chem. 284*, 20079–20089.

(4) Folk J.E., Cole P.W. (1966) Transglutaminase: mechanistic features of the active site as determined by kinetic and inhibitor studies. *Biochim. Biophys. Acta Enzymol. Biol. Oxid. 122*, 244– 264.

(5) Schägger,H. and Jagow,G.V. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem*., *166*, 368–379.