# Supporting Information for: The molecular features of non-peptidic nucleophilic substrates and acceptor proteins determine the efficiency of sortagging.

Tetiana Bondarchuk, Elena Zhuravel, Oleh Shyshlyk, Mykhaylo Debelyy, Oleksandr Pokholenko, Diana Vaskiv, Alla Pogribna, Mariana Kuznietsova, Yevhenii Hrynyshyn, Oleksandr Nedialko, Volodymyr Brovarets and Sergey Zozulya.

### **Contents:**

Table S1. Percentage yields of the conjugate products and calculated pKa values for nucleophiles

Table S2. Effect of trifluoroacetate on sortagging reaction

Fig. S1. SDS-PAAG of the recombinant proteins

Fig. S2. LC-MS QTOF spectra of the recombinant proteins

**Fig. S3.** Superposition of the docked LPETG motifs in PDB 1T2W and ClusPro protein-protein docking model between Sortase A and the I-TASSER model of human Carbonic Anhydrase II (Sortase-hCAHII-2GS-SRT-6H).

Fig. S4. Sortagging time course experiments for TNF $\alpha$ -GS-SRT - Compound 1 and TNF $\alpha$ -GS-SRT - Compound 7 protein acceptornucleophile pairs.

**Fig. S5.** Reproducibility checks of sortagging experiments analysis for TNFα-GS-SRT - Compound **1** and TNFα-GS-SRT - Compound **13**. **Synthetic procedures and quality control data for biotin derivatives 2, 6, 7, 9, 10, 12** 

						Prot	eins								Software		
Cmpd	EGF-Srt	TNFa-Srt	TNFa-GS- Srt	CAHII-SRT- 6H	CAHII-GS- SRT-6H	CAHII-PAS SRT-6H	EGF-Srt	TNFa-Srt	TNFa-GS- Srt	CAHII-SRT- 6H	CAHII-GS- SRT-6H	CAHII-PAS- SRT-6H	Marvin	ACD/Labs	Schroedin ger	MolGpK	Mean
	0.5/1 mM cmpd concentration, product yield, %			0.5/1mM cmpd concentration, hydrolytic side-product content, %				pKa values									
1	100/100	17/21	95/97	0/0	90/91	90/95	0/0	0/0	1/1	0/0	0/0	0/0	8.58	8.02	7.95	7.90	8.11
2	96/100	14/17	86/93	0/0	89/92	79/89	4/0	0/0	2/2	0/0	0/0	3/0	9.28	9.10	8.33	9.60	9.08
3	96/100	16/19	84/91	0/0	90/91	76/88	4/0	0/0	2/2	0/0	0/0	3/0	8.91	9.00	8.30	9.40	8.90
4	96/100	13/16	83/91	0/0	87/92	79/87	4/0	0/0	2/2	0/0	0/0	0/0	9.29	9.45	8.35	9.60	9.17
5	95/94	10/12	74/56	0/0	90/92	82/92	5/6	0/0	4/2	0/0	3/0	2/0	8.19	9.97	8.18	8.10	8.61
6	95/100	15/17	82/91	0/0	88/91	75/86	5/0	0/0	2/2	0/0	0/0	3/0	8.90	9.07	8.30	9.50	8.94
7	96/94	11/15	76/87	0/0	90/93	78/87	4/6	0/0	4/4	0/0	0/0	6/0	8.37	10.27	8.39	8.70	8.93
8	86/97	11/14	80/90	0/0	93/93	72/84	9/3	0/0	3/2	0/0	0/0	6/3	9.16	9.03	9.51	9.20	9.23
9	95/100	12/15	77/87	0/0	88/95	69/90	5/0	0/0	3/3	0/0	0/0	7/0	9.26	9.17	8.34	9.50	9.07
10	93/100	7/9	65/81	0/0	85/87	58/74	7/0	0/0	4/4	0/0	0/0	7/4	9.29	9.19	8.25	8.70	8.86
11	73/83	9/12	54/72	0/0	86/92	60/53	8/10	0/0	6/4	0/0	4/0	11/9	9.30	8.74	9.39	8.50	8.98
12	84/87	6/7	49/63	0/0	80/82	46/58	16/13	1/0	8/7	0/0	12/7	17/14	8.37	10.27	8.39	8.70	8.93
13	53/60	0/2	10/14	0/0	67/74	17/3	25/27	0/0	8/8	0/0	22/15	21/16	9.30	9.70	9.80	9.80	9.65
14	45/54	2/3	20/28	0/0	36/16	16/27	33/27	1/1	13/11	0/0	46/0	29/25	10.21	10.54	10.10	10.90	10.44
15	0/0	0/0	0/0	0/0	0/0	0/0	58/54	2/2	21/26	0/0	73/76	48/52	9.53	9.90	9.76	9.60	9.70
16	NA/61	NA/0	NA/28	NA/0	NA/20	NA/21	NA/25	NA/0	NA/22	NA/0	NA/23	NA/25	9.16	9.03	9.47	9.50	9.29
17	NA/100	NA/8	NA/60	NA/0	NA/53	NA/44	NA/0	NA/0	NA/7	NA/0	NA/14	NA/7	9.30	8.74	9.39	9.20	9.16

**Table S1.** Percentage yields of the conjugate products for all combinations of the nucleophiles and acceptor proteins for the sortagging reactions done at 0.5 mM and 1 mM nucleophile concentrations (values separated by slash). Percentage content of the hydrolytic side-products for the same reactions. For compounds **16** and **17** (DOTA derivatives) reactions at 0.5 mM nucleophile were not performed. The pKa values for the protonated nucleophilic amino groups of all substrates calculated by four computational tools, as well as their arithmetic means.

	Conjugate Yield, %	at 0.5 mM nucleophile	Conjugate Yield, %	% at 1 mM nucleophile
Acceptor Protein	TNFα-GS-Srt	CAHII-GS-SRT-6His	TNFα-GS-Srt	CAHII-GS-SRT-6His
Biotin Ethylenediamine, HCI, no TFA	70	28	83	26
Biotin Ethylenediamine, HCI, + TFA	72	26	84	32

**Table S2.** Effect of trifluoroacetate on sortagging reactions. Reactions with the 2 acceptor proteins were done in the standard conditions with or without addition of sodium trifluoroacetate in the equimolar concentration to the nucleophilic substrate.



**Fig. S1.** Gradient SDS-PAAG (7-22%) electrophoresis of the recombinant proteins used in this work, Coomassie stain. M – molecular weight markers in kD, 1 – EGF-SRT, 2 – TNF $\alpha$ -SRT, 3 – TNF $\alpha$ -GS-SRT, 4 – CAHII-SRT-6H, 5 – CAHII-GS-SRT-6H, 6 – CAHII-PAS-SRT-6H, 7 - SrtA7M.



Fig. S2. QTOF LC-MS spectra of the recombinant proteins used in the study.



**Fig. S3.** PyMOL visualization of the superposition of the docked LPETG motifs in PDB entry 1T2W (Crystal Structure of Sortase A in Complex with a LPETG peptide, DOI: <u>https://doi.org/10.2210/pdb1T2W/pdb</u>.) with the LPETG-sequence of ClusPro protein-protein docking model between Sortase A and the I-TASSER model of human Carbonic Anhydrase II (Sortase-hCAHII-2GS-SRT-6H). Sulfur atom in the catalytic Cys184 of the active site is highlighted in yellow.



**Fig. S4.** Sortagging time course experiments for TNF $\alpha$ -GS-SRT - Compound **1** and TNF $\alpha$ -GS-SRT - Compound **7** protein acceptornucleophile pairs. Peaks of the initial protein and the conjugate product on mass-spectra are marked by yellow and red stars, correspondingly. Hydrolytic side-product of sortagging was not detectable in these experiments. Calculated product yields in two LC-MS injections for each time point (rounded to the nearest integer) are shown in the tables and their standard deviation values are shown on the graphs. The experiments were done as described in the Materials and Methods section, at 10:1 molar ratio of nucleophile to protein (1 mM:100  $\mu$ M) and consecutive sampling at 4 time points.



Reproducibility of LC-MS screening,  $\text{TNF}\alpha\text{-}\text{GS-SRT}$  with Compound 13 Compound 13/ Well 1 Compound 13/ Well 2 Injection 1 Injection 1  $\star$ x10 x10 <sup>5</sup> 18536.25 18536.25 1 \* ★ ★ 17720.27 13 18850.82 18850.82 17720.31 18194.31 17389.00 17389.00 18194.64 0 17500 18000 18500 Counts vs. Deconvoluted Mass (amu) 19000 19000 17500 18000 18500 Counts vs. Deconvoluted Mass (amu) Injection 2 Injection 2 x10 5 ★ x10 18536.26 18536.23 Product Yield, % 17.13 17.80 3 17.75 17.54 \* \* 2 ┪ ★ Mean 17.56, StDev 0.31 2 17720.26 18850.82 18850.68

17395.00

0

19000

0

17500 18000 18500 Counts vs. Deconvoluted Mass (amu) 18194

19000

17500 18000 18500 Counts vs. Deconvoluted Mass (amu)

**Fig. S5.** Reproducibility checks of sortagging experiments analysis for TNF $\alpha$ -GS-SRT - Compound **1** and TNF $\alpha$ -GS-SRT - Compound **13** acceptor-nucleophile pairs. Peaks of the initial protein, the conjugate product and hydrolytic side-product on mass-spectra are marked by yellow, red and green stars, correspondingly. Peaks of SrtA7M are marked by blue triangles. Calculated product yields in two separate reactions with two consecutive LC-MS system injections for each reaction are shown in the tables. The experiments were done as described in the Materials and Methods section, at 10:1 molar ratio of nucleophile to protein (1 mM:100  $\mu$ M) and 24 h incubation.



### List of compounds

No.	No. Structure		
compound 2		95	HCI salt
compound 6		95	HCI salt
compound 7		90	HCI salt
compound 9	H NH2 NH2	91	
compound 10		91	HCI salt
compound 12		90	HCI salt



**Synthesis** Synthesis was carried out following the scheme given below:

#### Scheme 1





1







Scheme 3



Scheme 4



2



Scheme 5



Scheme 1 (compound 2)

#### Step A

Aldehyde 1(0.4 g, 1.7 mmol, 1 eq.) was dissolved in MeNH<sub>2</sub> (20% in MeOH, 20 ml) and the resulting mixture was stirred at room temperature overnight. Then NaBH<sub>4</sub> (0.07 g, 1.7 mmol, 1 eq.) was added portion wise and the resulting mixture was stirred at room temperature for 3 h. The reaction mixture was quenched with K<sub>2</sub>CO<sub>3</sub> (50 mL), extracted with EtOAc (3\*40 mL), washed with brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude amine 2(0.29 g, 93% purity, 1.16 mmol, 63% yield) obtained as yellow oil was used without further purification.

#### Step B

EDC (0.22 g, 1.74 mmol, 1.5 eq.) was added to a mixture of biotin (0.28 g, 1.16 mmol, 1 eq.), DIPEA (0.42 g, 3.48 mmol, 0.56 ml, 3 eq.) and HOBt (0.39 g, 2.9 mmol, 2.5 eq.) in DMF and the resulting mixture was stirred at room temperature for 30 min. Amine 2(0.29 g, 1.16 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude Boc-amine 3 (0.43 g, 63% purity, 0.9 mmol, 49% yield) obtained as brown oil was used without further purification.

#### Step C

Boc-amine **3** (0.43 g, 0.9 mmol) was treated with 3M HCI (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (0-0-25-100% 0-2-7-7.1min; 30ml/min water+HCI-



MeCN (loading pump 4ml/min water); target mass 377 column PFP 5uM 19\*100mm(R)) to afford compound 2 (0.058 g, 95% purity, 0.14 mmol, 15% yield) as light brown oil.

#### Scheme 2 (compound 6)

#### Step A

Aldehyde 1(0.4 g, 1.7 mmol, 1 eq.) was dissolved in MeNH<sub>2</sub> (20% in MeOH, 20 ml) and the resulting mixture was stirred at room temperature overnight. Then NaBH<sub>4</sub> (0.07 g, 1.7 mmol. 1 eq.) was added portion wise and the resulting mixture was stirred at room temperature for 3 h. The reaction mixture was quenched with K<sub>2</sub>CO<sub>3</sub> (50 mL) and extracted with EtOAc (3\*40 mL), washed with brine (20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude amine 2(0.38 g, 82% purity, 1.52 mmol, 73% yield) obtained as yellow oil was used without further purification.

#### Step B

EDC (0.29 g, 2.28 mmol, 1.5 eq.) was added to a mixture of biotin (0.56 g, 1.52 mmol, 1 eq.), DIPEA (0.56 g, 4.56 mmol, 0.73 ml, 3 eq.) and HOBt (0.51 g, 3.8 mmol, 2.5 eq.) in DMF and the resulting mixture was stirred at room temperature for 30 min. Amine 2(0.38 g, 1.52 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The crude Boc-amine **3** (0.61 g, 37% purity, 1.28 mmol, 31% yield) obtained as yellow oil was used without further purification.

#### Step C

Boc-amine **3** (0.61 g, 1.28 mmol) was treated with 3M HCl (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (0-0-25-100% 0-2-7-7.1min; 30ml/min water+HCl-MeCN (loading pump 4ml/min water); target mass 377 column Chromatorex C18 5uM 19\*100mm(R)) to afford compound 6 (0.087 g, 95% purity, 0.21 mmol, 16% yield) as brown solid.

#### Scheme 3 (compound 7)

#### Step A

EDC (0.22 g,1.71 mmol, 1.5 eq.) was added to a mixture of biotin (0.27 g, 1.14 mmol, 1 eq.), DIPEA (0.42 g, 3.42 mmol, 0.55 ml, 3 eq.) and HOBt (0.38 g, 2.85 mmol, 2.5 eq.) in DMF and the resulting mixture was stirred at room temperature for 30 min. Amine 1(0.25 g, 1.14 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude Boc-amine 2 (0.5 g, 15% purity, 1.12 mmol, 14% yield) obtained as yellow oil was used without further purification.

#### Step B

Boc-amine **2** (0.5 g, 1.12 mmol) was treated with 3M HCI (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (0-0-25-100% 0-2-7-7.1min; 30ml/min water+HCI-MeCN (loading pump 4ml/min water); target mass 345 column Chromatorex C18 5uM 19\*100mm(R)) to afford compound 7 (0.051 g, 90% purity, 0.13 mmol, 10% yield) as pale-yellow oil.

#### Scheme 4 (compound 9)

#### Step A

HATU (1 g, 2.7 mmol, 1.5 eq.) was added portion wise to a mixture of biotin (0.43 g, 1.8 mmol, 1 eq.) and DIPEA (0.66 g, 5.4 mmol, 0.87 ml, 3 eq.) in DMF and the resulting mixture was stirred at



room temperature for 30 min. Amine 1(0.4 g, 1.8 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude Boc-amine **2** (0.53 g, 28% purity, 1.18 mmol, 18% yield) obtained as yellow oil was used without further purification.

#### Step B

Boc-amine **2** (0.53 g, 1.18 mmol) was treated with 3M HCl (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (3-10-40-100% 0-2-7-7.1 30ml/min water-MeOH (loading pump 4ml/min water); target mass 349; column SunFireC18 19\*100mm (L)) to afford compound 9 (0.096 g, 91% purity, 0.27 mmol, 21% yield) as brown solid.

#### Scheme 5 (compound 10)

#### Step A

Piperazine (1.02 g, 11.91 mmol, 7 eq.) was added to mixture of aldehyde 1(0.4 g, 1.7 mmol, 1 eq.) in methanol (15 ml) and the resulting mixture was stirred at room temperature overnight. Then NaBH<sub>4</sub> (0.07 g, 1.7 mmol, 1 eq.) was added portion wise and the resulting mixture was stirred at room temperature for 3 h. The reaction mixture was quenched with K<sub>2</sub>CO<sub>3</sub> (50 mL) and extracted with EtOAc (3\*40 mL), washed with brine (20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude amine 2(0.19 g, 94% purity, 0.62 mmol, 34% yield) obtained as colorless oil was used without further purification.

#### Step B

EDC (0.12 g, 0.93 mmol, 1.5 eq.) was added to a mixture of biotin (0.15 g, 0.62 mmol, 1 eq.), DIPEA (0.22 g, 1.86 mmol, 0.3 ml, 3 eq.) and HOBt (0.2 g, 1.55 mmol, 2.5 eq.) in DMF and the resulting mixture was stirred at room temperature for 30 min. Amine 2(0.19 g, 0.62 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude Boc-amine 3 (0.22 g, 50% purity, 0.44 mmol, 33% yield) obtained as brown oil was used without further purification.

#### Step C

Boc-amine **3** (0.22 g, 0.44 mmol) was treated with 3M HCl (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (0-0-40-100% 0-2-7-7.1 min; 30ml/min water-MeOH+HCl (loading pump 4 ml/min water); target mass 432 column Chromatorex C18 5uM 19\*100 uM) to afford compound 10 (0.036 g, 91% purity, 0.07 mmol, 16% yield) as brown oil.

#### Scheme 6 (compound 12)

#### Step A

EDC (0.22 g,1.71 mmol, 1.5 eq.) was added to a mixture of biotin (0.27 g, 1.14 mmol, 1 eq.), DIPEA (0.42 g, 3.42 mmol, 0.55 ml, 3 eq.) and HOBt (0.38 g, 2.85 mmol, 2.5 eq.) in DMF and the resulting mixture was stirred at room temperature for 30 min. Amine 1(0.25 g, 1.14 mmol, 1 eq.) was then added slowly and reaction mixture was stirred at room temperature further 12 h. The reaction mixture was diluted with EtOAc (120 mL), washed with NaHSO<sub>4</sub> aq. (25 mL), NaHCO<sub>3</sub> aq. (25 mL), brine (3\*20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. The combined organic extract was concentrated under reduced pressure to dryness. The crude Boc-amine 2 (0.32 g, 8% purity, 0.72 mmol, 5% yield) obtained as brown oil was used without further purification.



#### Step B

Boc-amine **2** (0.32 g, 0.72 mmol) was treated with 3M HCl (dioxane solution, 10 ml) and resulting mixture was stirred at room temperature overnight. After consumption of the starting material (LCMS control) the resulting mixture was concentrated under reduced pressure to dryness. The residue obtained was subjected for prep HPLC purification (0-0-25-100% 0-2-7-7.1min; 30ml/min water+HCl-MeCN (loading pump 4ml/min water); target mass 345 column Chromatorex C18 5uM 19\*100mm(R)) to afford compound 12 (0.061 g, 90% purity, 0.16 mmol, 20% yield) as pale-yellow solid.



Analytical Tests	Results
LCMS	[M-H] <sup>-</sup> : 375.0
LCMS (>95%)	Rt= 0.618 min, 100% (215nm)
1H-NMR	Consistent with structure





![](_page_14_Figure_1.jpeg)

Analytical Tests	Results
LCMS	[M-H] <sup>-</sup> : 375.0
LCMS (>95%)	Rt= 0.624 min, 100% (215nm)
1H-NMR	Consistent with structure

![](_page_15_Figure_0.jpeg)

![](_page_16_Figure_0.jpeg)

![](_page_17_Figure_1.jpeg)

Analytical Tests	Results
LCMS	[M+H]+: 345.0
LCMS (>95%)	Rt= 0.685 min, 96.37% (215nm)
1H-NMR	Consistent with structure

![](_page_18_Figure_0.jpeg)

![](_page_19_Figure_0.jpeg)

Mol Structure	H N N H <sub>2</sub> N H
Mol Formula	$C_{17}H_{24}N_4O_2S$
Compound Name	Z2451166862
Mol. Wt.	348.16
Date Synthesized	October 20, 2023

Analytical Tests	Results
LCMS	[M-H] <sup>-</sup> : 347.2
LCMS (>95%)	Rt= 0.687 min, 97.98% (215nm)
1H-NMR	Consistent with structure

![](_page_21_Figure_0.jpeg)

![](_page_22_Figure_0.jpeg)

![](_page_23_Figure_0.jpeg)

![](_page_24_Figure_1.jpeg)

Analytical Tests	Results
LCMS	[M-H] <sup>-</sup> : 429.2
LCMS (>95%)	Rt= 0.330 min, 100% (215nm)
1H-NMR	Consistent with structure

![](_page_25_Figure_0.jpeg)

![](_page_26_Figure_0.jpeg)

![](_page_27_Figure_1.jpeg)

Analytical Tests	Results
LCMS	[M+H]+: 345.0
LCMS (>95%)	Rt= 0.682 min, 97.52% (215nm)
1H-NMR	Consistent with structure

![](_page_28_Figure_0.jpeg)

![](_page_29_Figure_0.jpeg)