

# Supporting information

## Aptamer-Directed Conjugation of DNA to Therapeutic Antibodies

Mikkel B. Skovsgaard,<sup>†</sup> Michael R. Mortensen,<sup>†</sup> Johan Palmfeldt<sup>‡</sup> and Kurt V. Gothelf.<sup>†\*</sup>

<sup>†</sup>iNANO, Department of Chemistry, Gustav Wieds Vej 14, Aarhus University, DK-8000 Aarhus C, Denmark

<sup>‡</sup>Department of Clinical Medicine, Palle Juul-Jensens Boulevard 82, Aarhus University, DK-8200 Aarhus N, Denmark

\*kvg@chem.au.dk

### Table of contents

General information .....	2
Oligonucleotide sequences .....	2
Oligonucleotide modification .....	3
Calculation of the conversion based on gel analysis .....	4
Supporting Figures.....	5
Tandem MS data for the conjugate prepared using the templated conjugation method.....	10
Tandem MS data for the conjugate prepared using the direct conjugation method .....	19
References .....	26

## **General information**

Mouse anti-c-Myc was purchased from Sigma and used without further purification. Herceptin® (trastuzumab), MabThera® (rituximab), Erbitux® (cetuximab) and Vectibix™ (panitumumab) were purchased at the pharmacy of Aarhus University Hospital as the pharmaceutic and buffer exchanged by dissolving them in a phosphate 25 mM pH 6.8 buffer and washed 10 times using Amicon spin filters. The two aptamers were purchased from LGC Biosearch Technologies and used without further purification. All other oligonucleotides were purchased from Integrated DNA Technologies.

A ND-1000 NanoDrop® spectrophotometer was used for the UV-vis measurement of DNA, protein and conjugate absorbance. A Milli-Q® Biocell system was used for purification of the used water. SDS-PAGE analysis was performed using precast NuPAGE®, Novex® 4-12% Bis-Tris gels 1.0 mm. The samples for SDS-analysis was prepared by addition of NuPAGE® LDS 4X sample buffer (final concentration 1X), DTT (final concentration 50 mM) and incubation at 90 °C for 7 min. The SDS-PAGE analysis was performed at 200 V for 55 min. The gel was removed from the cassette, stained using SimplyBlue™ SafeStain and scanned with a Gel Doc™ EZ. The scans were analyzed using Image Lab™ 6.0.

## **Oligonucleotide sequences**

### **Guiding aptamer**

5' GGagg-Fu-gC-Fu-CCGAAAggAA-Fc-Fu-CCGGGCTCATGCGAGGCTTACGAAC 3'

Small letters: RNA

Capital letters: DNA

Fu: 2'-fluoro uridine

Fc: 2'-fluoro cytidine

Black: aptamer region

Green: extended region for DNA hybridization to the reactive strand

Blue: toehold used for strand displacement

### **Aldehyde aptamer (Ald-Apt)**

5' GGagg-Fu-gC-Fu-CCGA\* AAggAA-Fc-Fu-CCGGGCTCATGCGAGGCTTACGAAC 3'

Small letters: RNA

Capital letters: DNA

Fu: 2'-fluoro uridine

Fc: 2'-fluoro cytidine

A\*: amino-modifier C6-dA

Black: aptamer region

Green: extended region for DNA hybridization to the reactive strand

Blue: toehold used for strand displacement

### **Reactive strands:**

OT 5' ACATACAGCCTCGCATGAGCCC-X 3'

2T      5' ACATACAGCCTCGCATGAGCCC-**TT-X** 3'  
 5T      5' ACATACAGCCTCGCATGAGCCC-**TTTTT-X** 3'  
 10T     5' ACATACAGCCTCGCATGAGCCC-**TTTTTTTTT-X** 3'

X: 3' amino modifier

Green: complementary to the extended region of the guiding aptamer.

#### **Scrambled reactive strand:**

TCAGGTGTCATCATGTATGCTT-X

X: 3' amino modifier

#### **Complementary displacement strand**

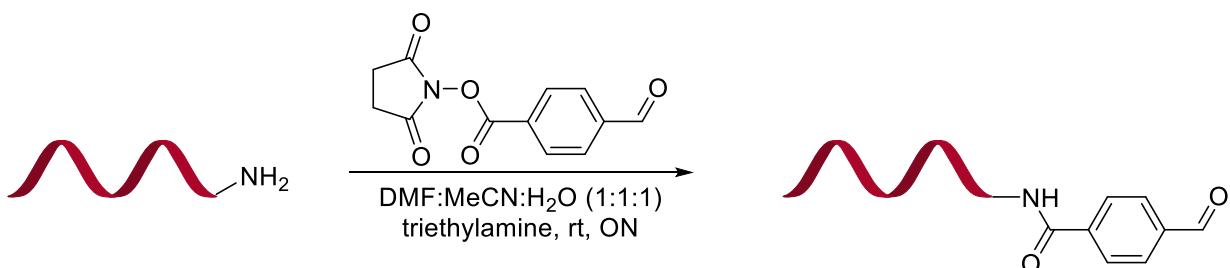
5'- GTTCGTAAGCCTCGCATGAGCCGGAGTTCCCTTCGGAGCACCTCC -3'

#### **Cy5 modified strand**

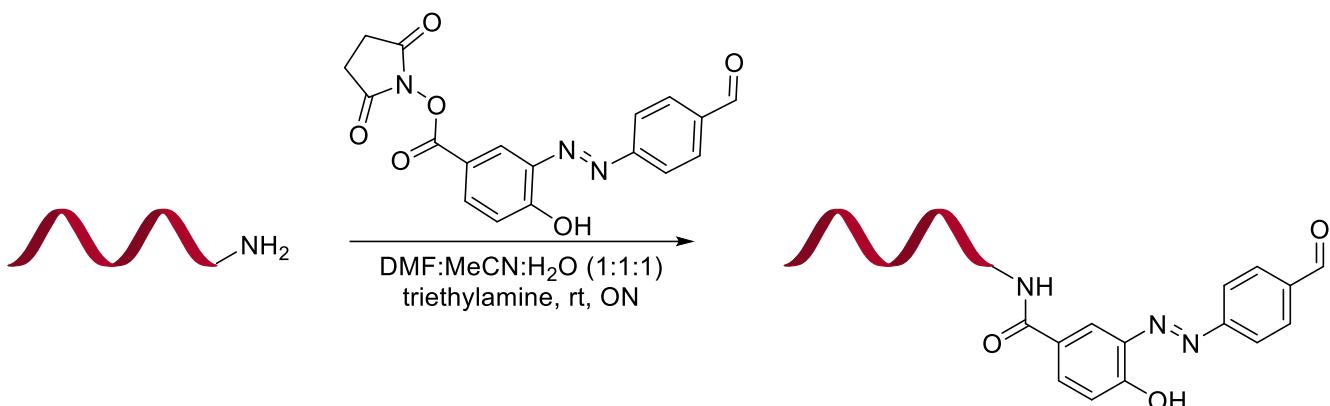
5'-GGGCTCATGCGAGGCTGTATGT-Cy5-3'

#### **Oligonucleotide modification**

Aldehyde modification of amino-DNA either reactive strand or amino-aptamer. The modification reaction was performed as described previously.<sup>1</sup>

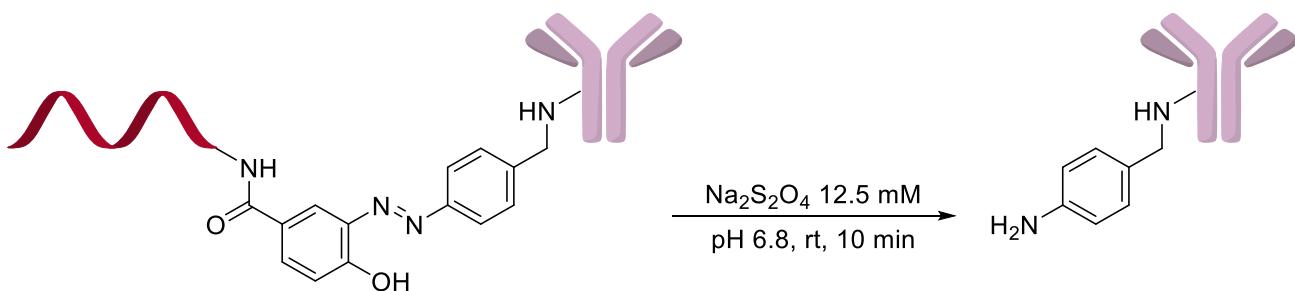


Modification with the cleavable azobenzene aldehyde linker. The NHS ester was synthesized according to literature.<sup>2</sup>



### Oligonucleotide azo benzene cleavage

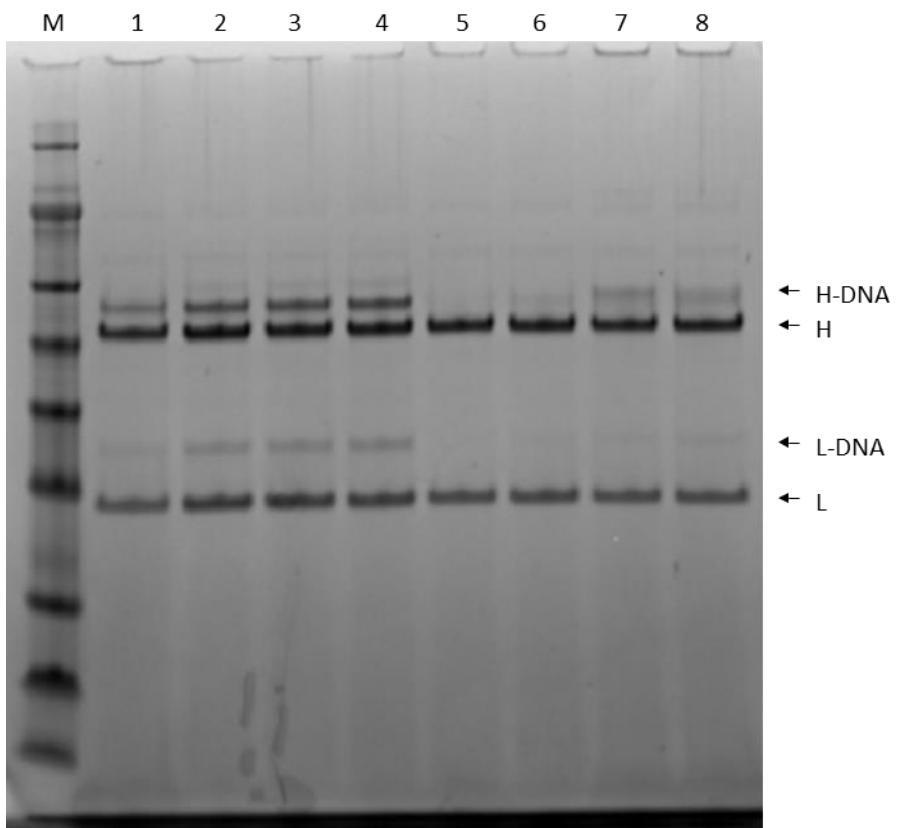
The cleavage of the azo benzene linker was performed as described previously.<sup>2</sup>



### Calculation of the conversion based on gel analysis

Because of the symmetrical nature of an antibody, the conversion calculation cannot solely be based on the intensity of the band of the modified chain. The calculation must also take into account that an antibody modified with one DNA strand on a chain, will give rise to both a modified band and an unmodified band. The intensity of the modified band will be 50% of the total of the two, however, the actual conversion will be 100%. The calculated conversion is based on this premise and is both calculated for the heavy and the light chain. The total estimated conversion is the combined conversion of the two. When some of the antibody is modified with two DNA strands the conversion will exceed 100% conversion.

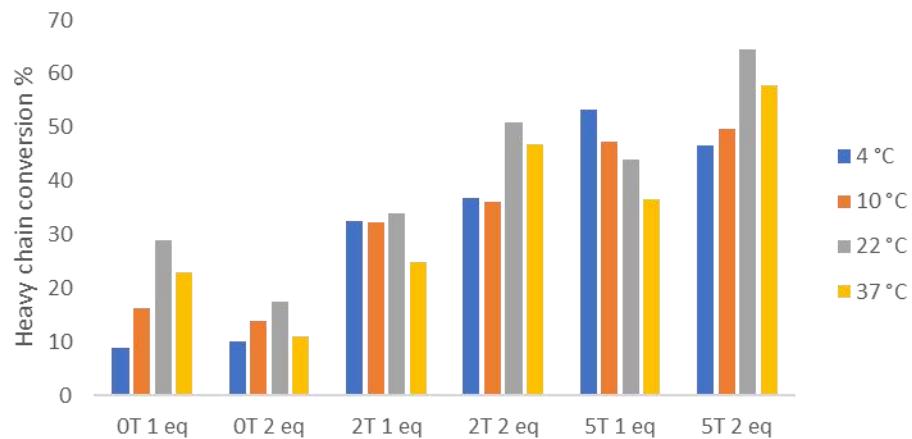
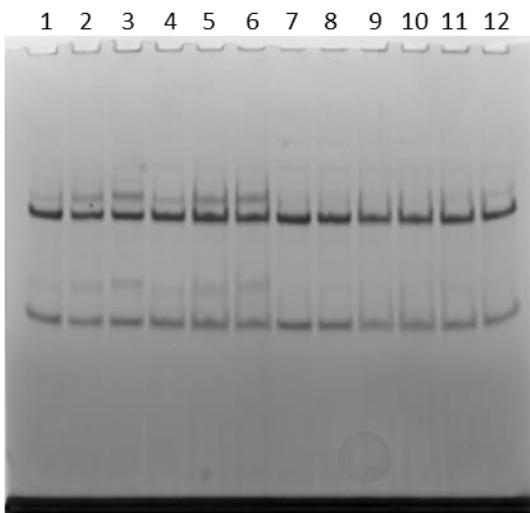
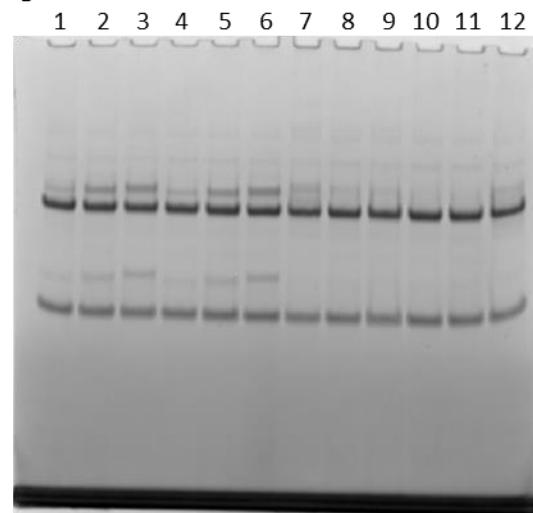
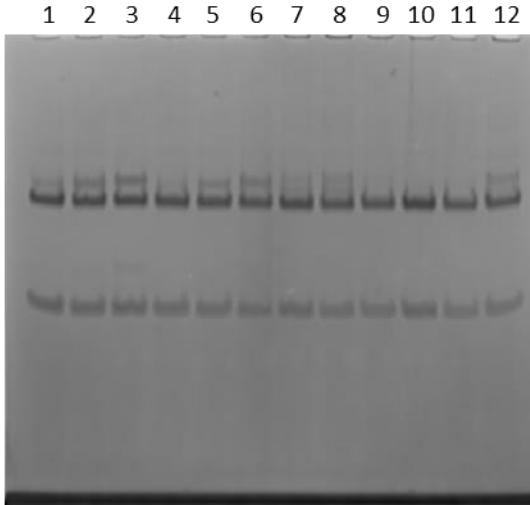
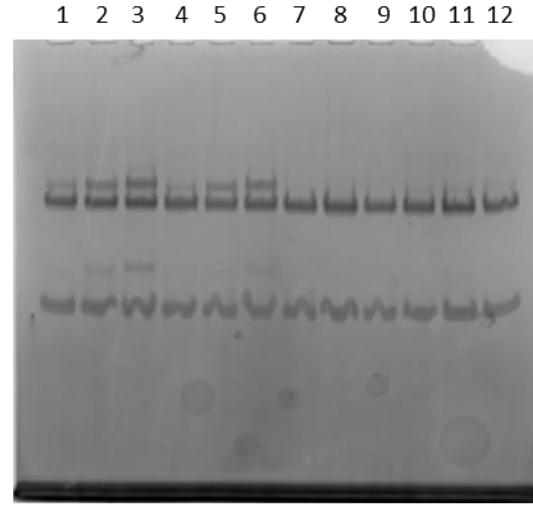
## Supporting Figures



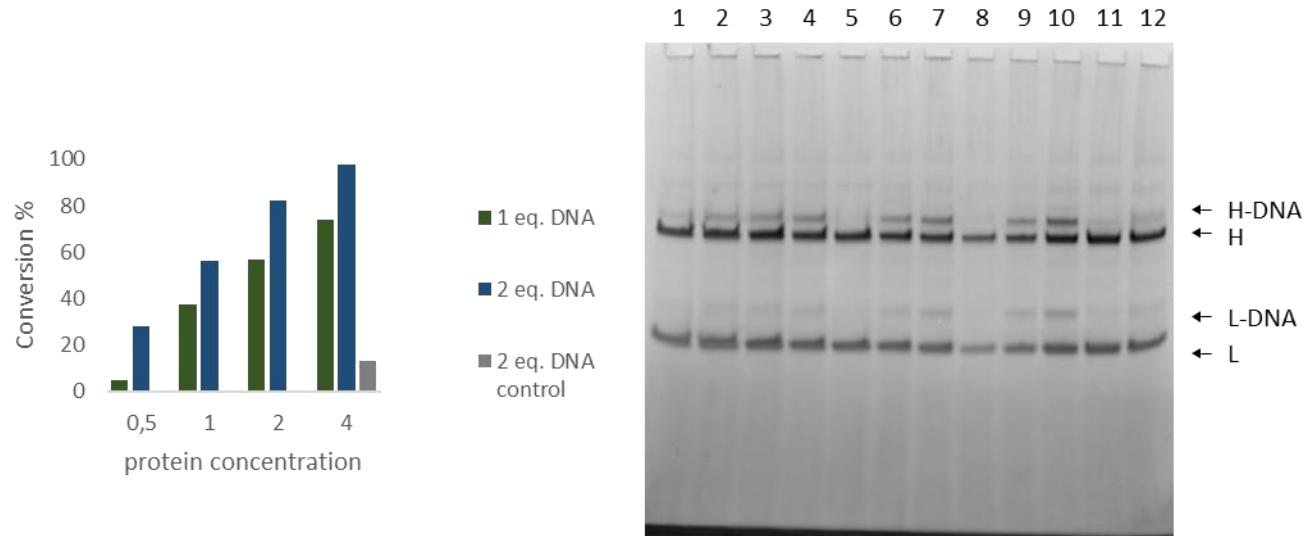
**Figure S1.** Templated conjugation with different amounts of 2T reactive strand analyzed on a 4-12% SDS-PAGE stained for protein. Reactions were performed with Tz (2  $\mu$ M), guiding aptamer (only lane 1-4), 2T reactive strand and NaBH<sub>3</sub>CN. M: ladder SeeBlue Plus2. Lane 1: 1 eq. of DNA. Lane 2: 2 eq. of DNA. Lane 3: 3 eq. of DNA. Lane 4: 5 eq. of DNA. Lane 5: 1 eq. of 2T reactive strand, no aptamer. Lane 6: 2 eq. of 2T reactive strand, no aptamer. Lane 7: 3 eq. of 2T reactive strand, no aptamer. Lane 8: 5 eq. of 2T reactive strand, no aptamer. H: heavy chain. H-DNA: DNA modified heavy chain. L: light chain. L-DNA: DNA modified light chain.

**A**

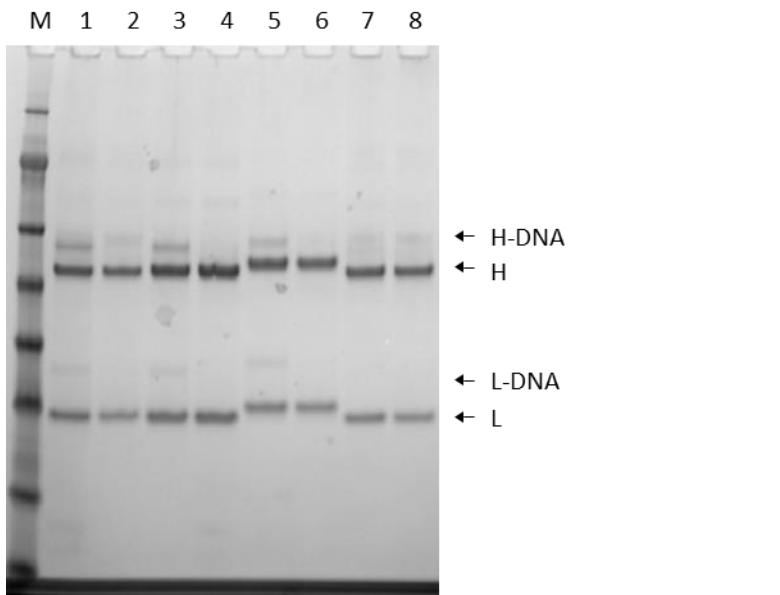
## Temperature study

**B****C****D****E**

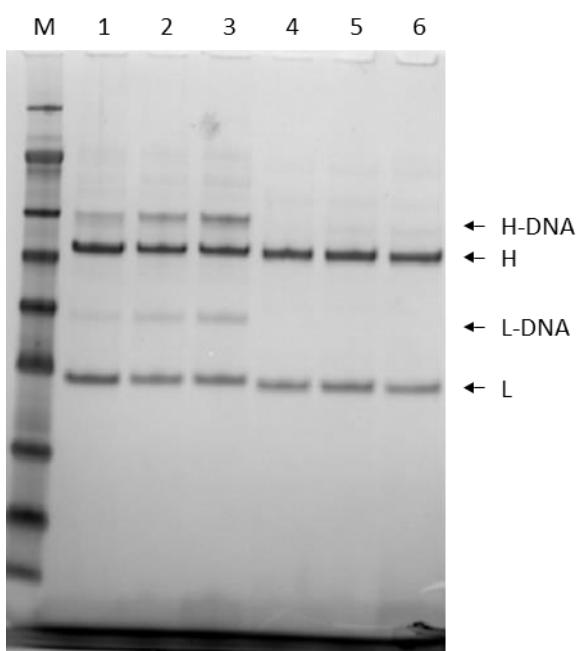
**Figure S2.** Templated conjugation at different temperatures analyzed on a 4-12% SDS-PAGE stained for protein. (A) Bar chart showing the heavy chain conversion for the different temperatures and the different reactive strands used. The conjugation was tested using the OT, 2T and the 5T reactive strands with 1 or 2 eq. of DNA to protein. The reactions were performed at 4, 10, 22 and 37 °C. (B)-(E) Reactions were performed with Tz (2 μM), guiding aptamer (only lane 1-6), reactive strand; OT (lane 1, 4, 7 and 10); 2T (lane 2, 5, 8 and 11); 5T (lane 3, 6, 9 and 12) and NaBH<sub>3</sub>CN. (B) 4 and 10 °C, 1 eq. of DNA. (C) 4 and 10 °C, 2 eq. of DNA. (D) 22 and 37 °C, 1 eq. of DNA. (E) 22 and 37 °C, 2 eq. of DNA.



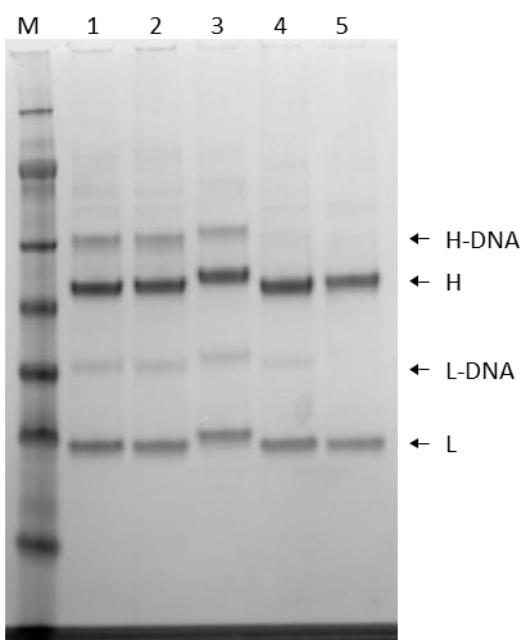
**Figure S3.** Templated conjugation at different concentration of Tz with 1 and 2 eq. of 2T reactive strand analyzed on a 4-12% SDS-PAGE stained for protein. All reactions included NaBH<sub>3</sub>CN. Lane 1: 0.5 μM Tz, 1 eq. of 2T and aptamer. Lane 2: 0.5 μM Tz, 2 eq. of 2T and aptamer. Lane 3: 1 μM Tz, 1 eq. of 2T and aptamer. Lane 4: 1 μM Tz, 2 eq. of 2T and aptamer. Lane 5: 1 μM Tz, 1 eq. of 2T and no aptamer. Lane 6: 2 μM Tz, 1 eq. of 2T and aptamer. Lane 7: 2 μM Tz, 2 eq. of 2T and aptamer. Lane 8: 2 μM Tz, 1 eq. of 2T and no aptamer. Lane 9: 4 μM Tz, 1 eq. of 2T and aptamer. Lane 10: 4 μM Tz, 2 eq. of 2T and aptamer. Lane 11: 4 μM Tz, 1 eq. of 2T and no aptamer. Lane 12: 4 μM Tz, 2 eq. of 2T and no aptamer. H: heavy chain. H-DNA: DNA modified heavy chain. L: light chain. L-DNA: DNA modified light chain.



**Figure S4.** Templated conjugation of 2T reactive strand to different antibodies analyzed on a 4-12% SDS-PAGE stained for protein. All reactions were performed at 2  $\mu$ M protein. M: ladder SeeBlue Plus2. Lane 1, 3, 5, 7: Ab, aptamer and 2T. Lane 2, 4, 6, 8: Ab and 2T. Lane 1-2: Tz. Lane 3-4: Rx. Lane 5-6: Cx. Lane 7-8: Pan. H: heavy chain. H-DNA: DNA modified heavy chain. L: light chain. L-DNA: DNA modified light chain.



**Figure S5.** Direct conjugation of different eq. of Ald-Apt to protein analyzed on a 4-12% SDS-PAGE stained for protein. All reactions were performed at 2  $\mu$ M protein (Lane 1-3: Tz, lane 4-6: mouse anti-c-Myc). M: ladder SeeBlue Plus2. Lane 1+4: 0.5 eq. Ald-Apt. Lane 2+5: 1 eq. Ald-Apt. Lane 3+6: 2 eq. Ald-Apt. H: heavy chain. H-DNA: DNA modified heavy chain. L: light chain. L-DNA: DNA modified light chain.

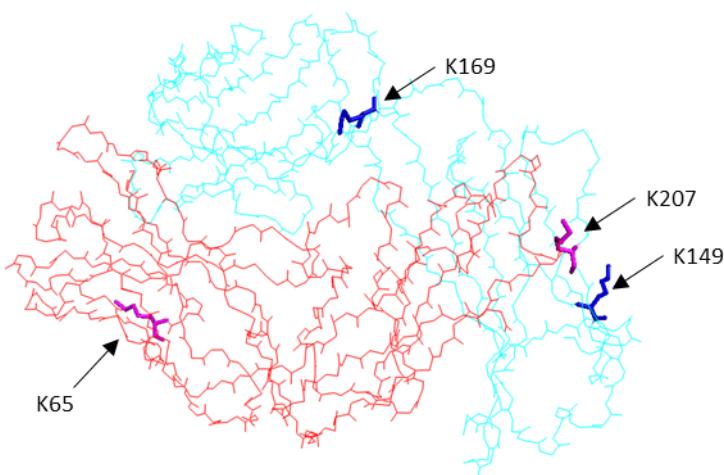


**Figure S6.** Direct conjugation of Ald-Apt to different antibodies analyzed on a 4-12% SDS-PAGE stained for protein. All reactions were performed at 2  $\mu$ M protein with 1 eq. of Ald-Apt. M: ladder SeeBlue Plus2. Lane 1: Tz. Lane 2: Rx. Lane 3: Cx. Lane 4: Pan. Lane 5: mouse anti-c-Myc. H: heavy chain. H-DNA: DNA modified heavy chain. L: light chain. L-DNA: DNA modified light chain.

## Tandem MS data for the conjugate prepared using the templated conjugation method

The first peak eluted during HPLC purification (■)

Residue number	Probability score	Peptide sequence
Light chain		
207	0.013	VYACEVTHQGLSSPVT <u>K</u> SFNR
207	0.0012	VYACEVTHQGLSSPVT <u>K</u>
149/169	$9.8 \times 10^{-9}$	VQW <u>K</u> VVDNALQSGNSQESVTEQDS <u>K</u>
Heavy chain		
65	0.00018	YADSVKGR



The modified lysine residues are indicated on the crystal structure of the Fab domain of Tz. Light chain is colored cyan, while the heavy chain is colored red. Modifications with a good probability score (<0.005) are shown in magenta. Lysine residues colored blue are found on the same peptide, which have a good probability score (<0.005) for being modified. PDB: 1N8Z.

The sequence of Tz is shown below with the modified lysine residues highlighted in red. The peptides detected with a modification are underlined. Two different peptides were found with a modification on lysine 207, they are indicated with over- and underline.

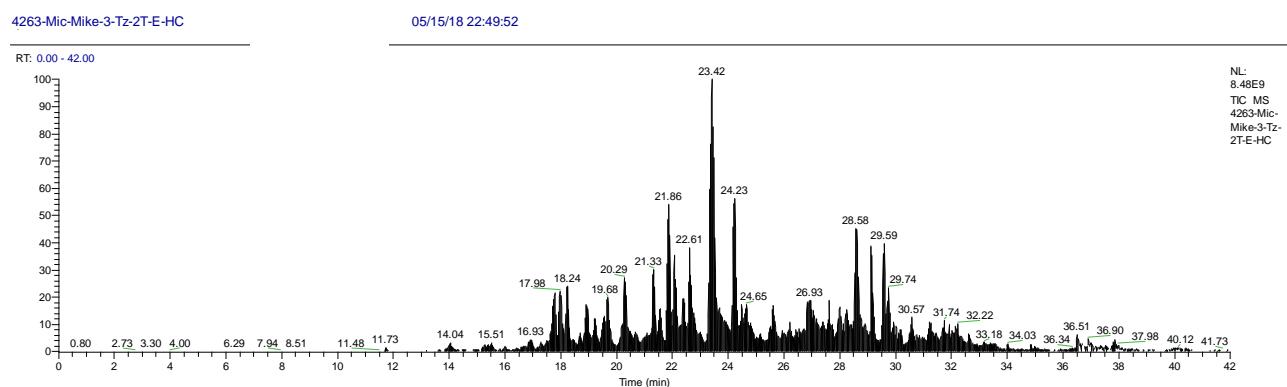
### Heavy chain sequence

10	20	30	40	50	60
EVQLVESGGG LVQPGGSLRL SCAASGFNIK DTYIHWVRQA PGKGLEWVAR IYPTNGYTRY					
70	80	90	100	110	120
ADSV <u>KGRFTI</u> SADTSKNTAY LQMNSLRAED TAVYYCSRWG GDGFYAMDYW GQGTLTVSS					
130	140	150	160	170	180
ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAAVLQSS					
190	200	210	220	230	240
GLYSISSLVVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP PKSCDKTHTC PPCPAPELLG					
250	260	270	280	290	300
GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY					
310	320	330	340	350	360
NSTYRVVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRD					
370	380	390	400	410	420
ELTKNQVSLT CLVKGFYPSD IAVEWESENQ PENNYKTTTPV VLSDSDGSFFL YSKLTVDKSR					
430	440	450			
WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K					

### Light chain sequence

10	20	30	40	50	60
DIQMTQSPSS LSASVGDRVT ITCRASQDVN TAVAWYQQKP GRAPKLLIYS ASFLYSGVPS					
70	80	90	100	110	120
RFSGSRSGTD FTLTISSSLQP EDFATYYCQQ HYTPPPTFGQ GTKVEIKRTV AAPSVFIFPP					
130	140	150	160	170	180
SDEQLKSGTA SVVCLLNNFY PREAKV <u>QWKV</u> DNALQSGNSQ ESVTEQDS <u>KD</u> STYSLSSTLT					
190	200	210			
LSKADYEKHK VYACEVTHQG LSSPVT <u>KSFn</u> RGEC					

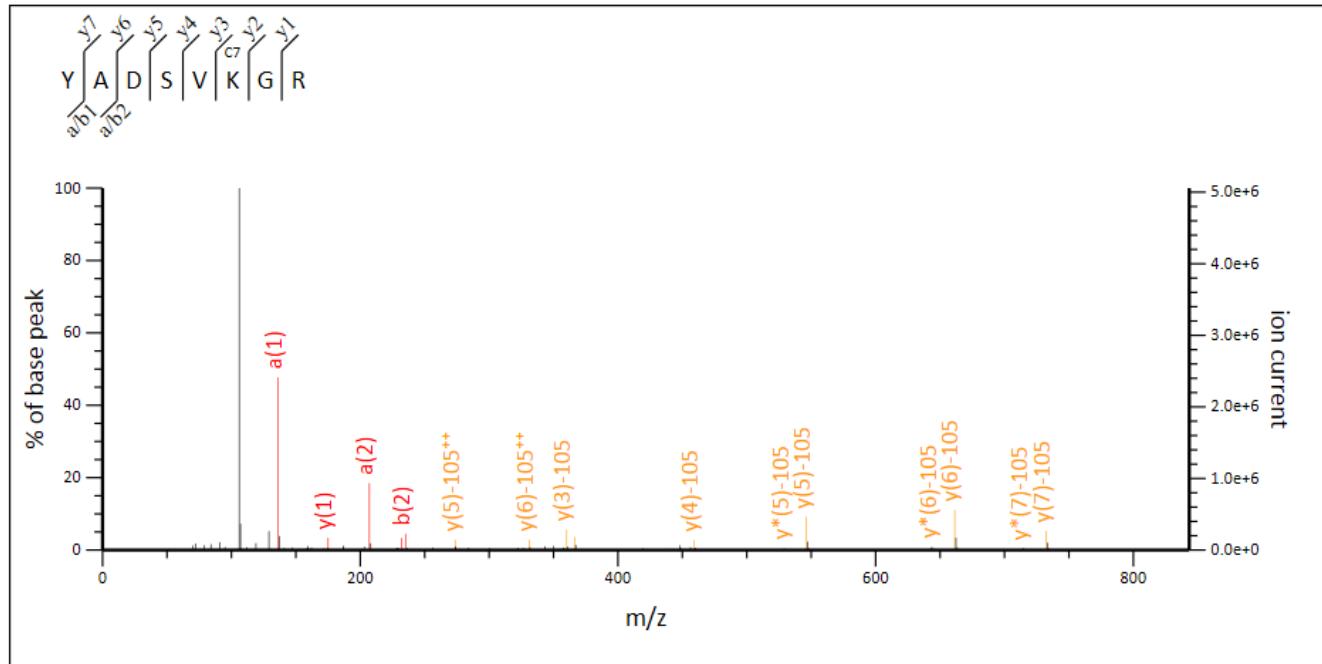
### Full chromatogram for the heavy chain



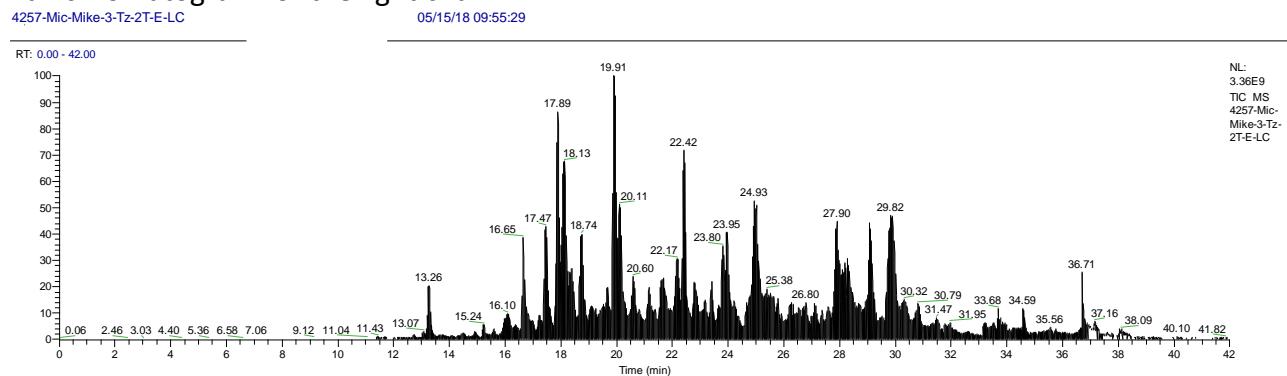
## Fragment spectra of the peptides

The following pages show the fragment spectrum of the heavy chain peptide found with a modification.

### Heavy chain peptide: YADSVKGR



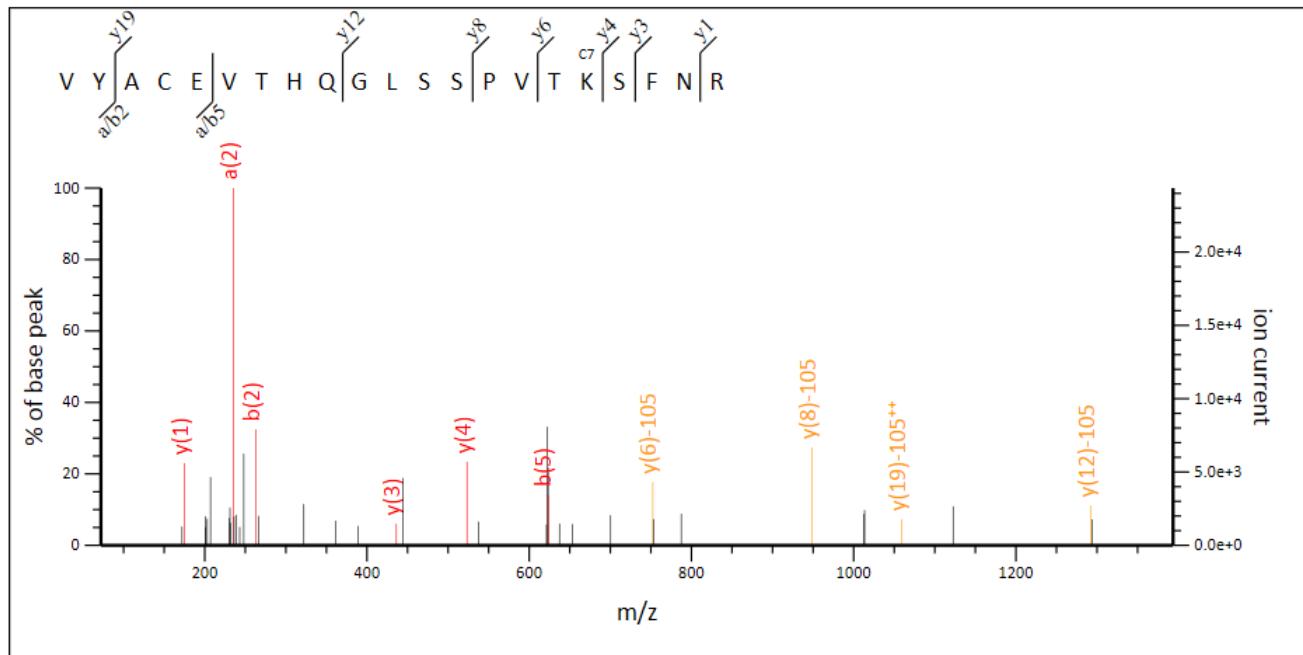
### Full chromatogram for the light chain



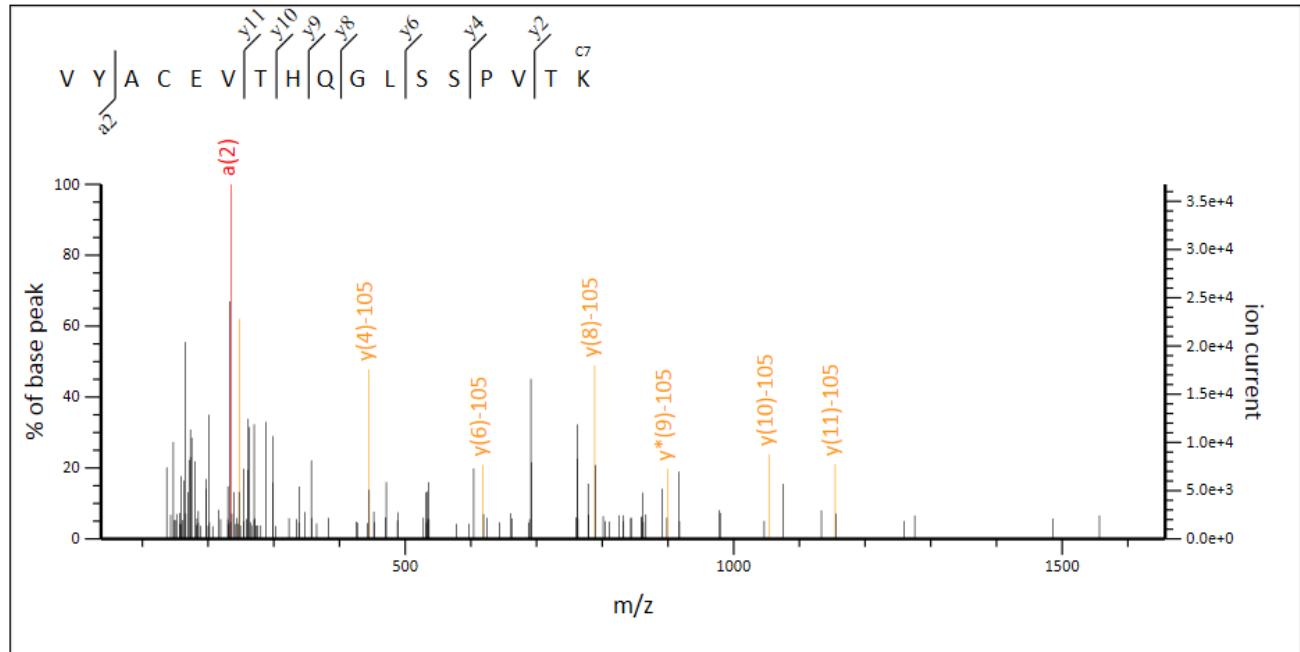
## Fragment spectra of the peptides

The following pages show the fragment spectra of the light chain peptides found with a modification.

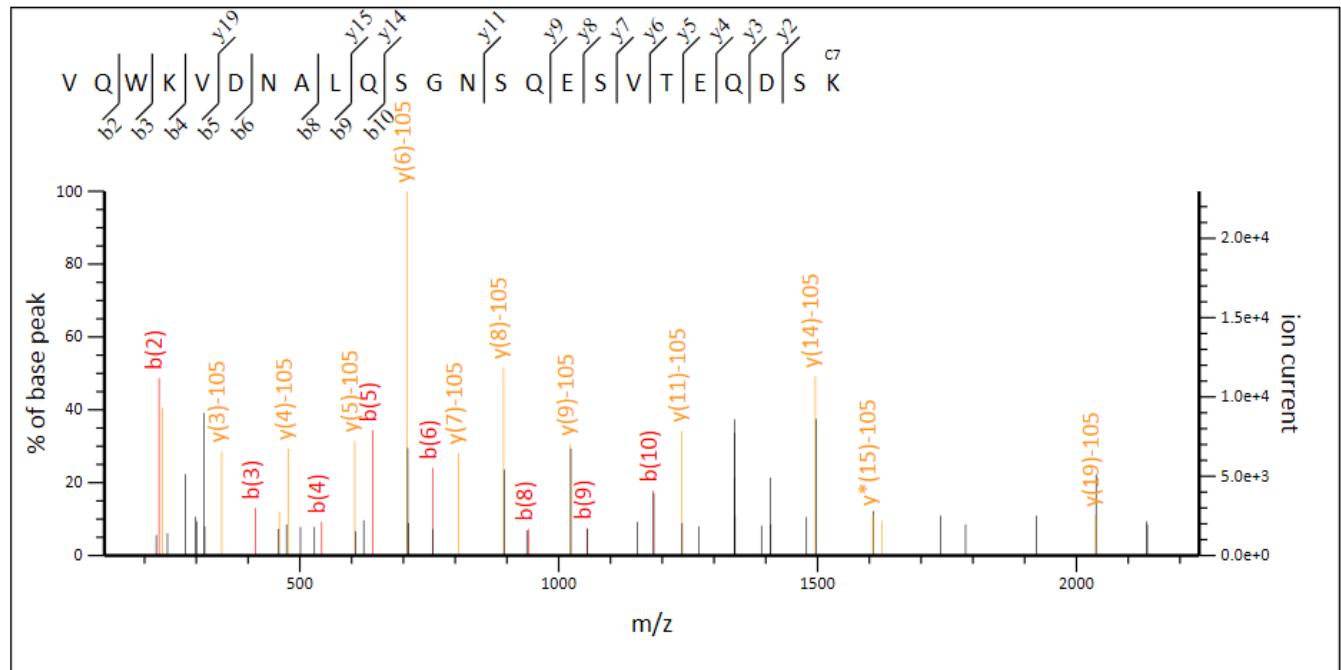
Light chain peptide: VYACEVTHQGLSSPVTKSFNR



Light chain peptide: VYACEVTHQGLSSPVTK

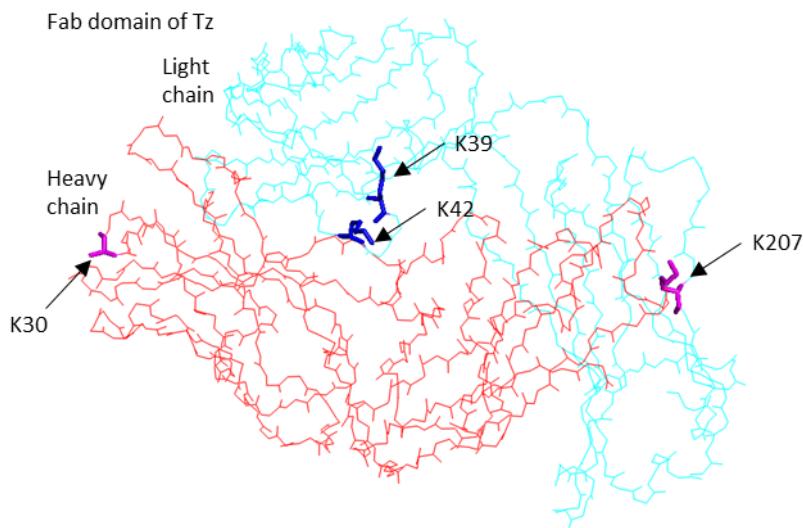


Light chain peptide: VQWKVDNALQSGNSQESVTEQDSK



The second peak eluted on during HPLC purification ( $\blacktriangle$ )

Residue number	Probability score	Peptide sequence
Light chain		
39/42	$2.2 \times 10^{-6}$	ASQDVNTAVAWYQQ <u>KPGK</u>
207	0.00027	VYACEVTHQGLSSPVT <u>K</u>
Heavy chain		
30	$5.3 \times 10^{-6}$	LSCAASGFNI <u>K</u> DTYIHWVR



The modified lysine residues are indicated on the crystal structure of the Fab domain of Tz. Light chain is colored cyan, while the heavy chain is colored red. Modifications with a good probability score ( $<0.005$ ) are shown in magenta. Lysine residues colored blue are found on the same peptide, which have a good probability score ( $<0.005$ ) for being modified. PDB: 1N8Z.

The sequence of Tz is shown below with the modified lysine residues highlighted in red. The peptides detected with a modification are underlined.

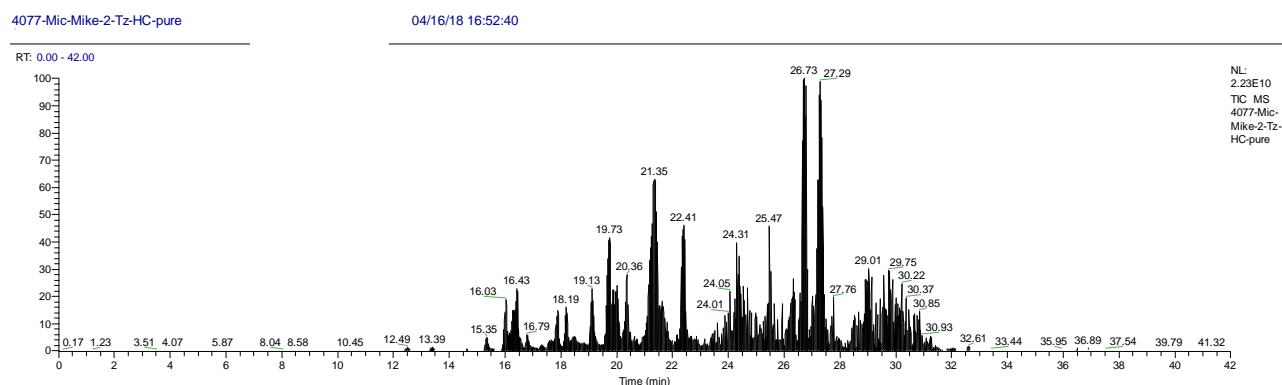
### Heavy chain sequence

10	20	30	40	50	60
EVQLVESGGG LVQPGGSLRL <u>SCAASGFNIK</u> DTYIHWVRQA PGKGLEWVAR IYPTNGYTRY					
70	80	90	100	110	120
ADSVVKGRFTI SADTSKNTAY LQMNSLRAED TAVYYCSRWG GDGFYAMDYW GQGTLTVSS					
130	140	150	160	170	180
ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAAVLQSS					
190	200	210	220	230	240
GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP PKSCDKTHTC PPCPAPELLG					
250	260	270	280	290	300
GPSVFLFPPK PKDTLMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY					
310	320	330	340	350	360
NSTYRVSVSL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRD					
370	380	390	400	410	420
ELTKNQVSILT CLVKGFYPSDL IAVEWESENQ PENNYKTTTPV VLSDSDGSFFL YSKLTVDKSR					
430	440	450			
WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K					

### Light chain sequence

10	20	30	40	50	60
DIQMTQSPSS LSASVGDRVT <u>ITCRASQDVN</u> TAVAWYQQKP <u>GK</u> APKLLIYS ASFLYSGVPS					
70	80	90	100	110	120
RFSGSRSGTD FTLTISSSLQP EDFATYYCQQ HYTTPPTFGQ GTKVEIKRTV AAPSVFIFPP					
130	140	150	160	170	180
SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYSLSSLT					
190	200	210			
LSKADYEKHK <u>VYACEVTHQG</u> <u>LSSPVT</u> <u>K</u> SFN RGEc					

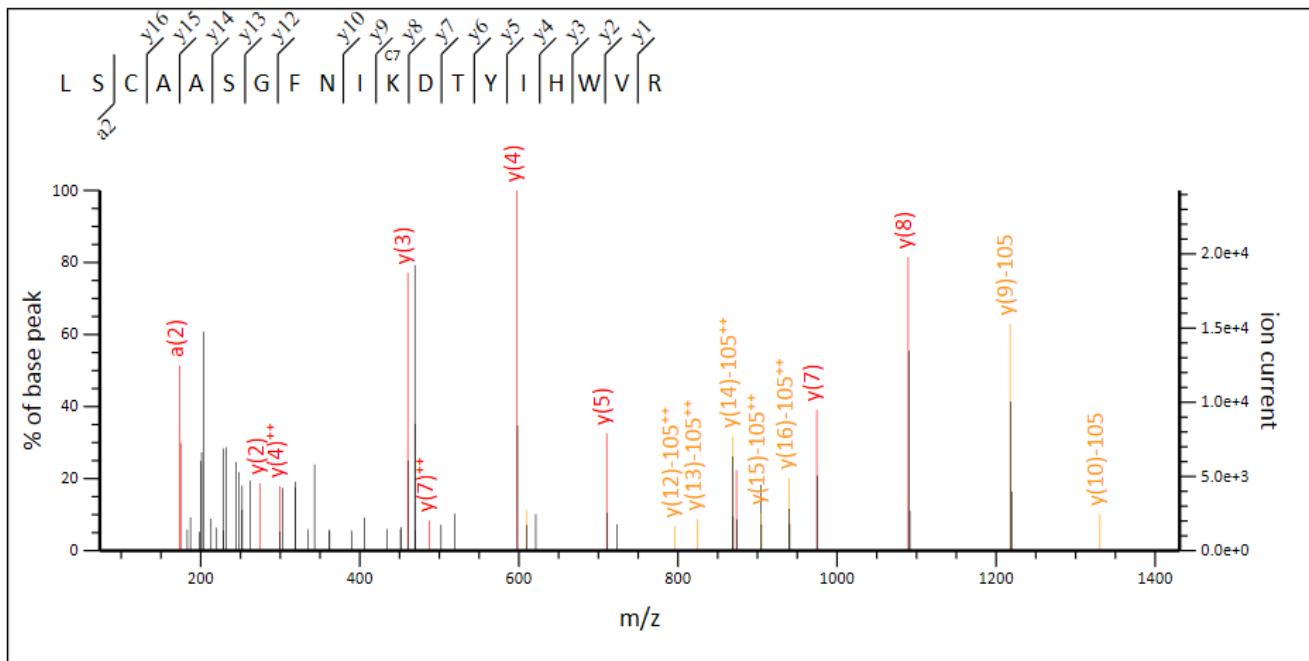
### Full chromatogram for the heavy chain



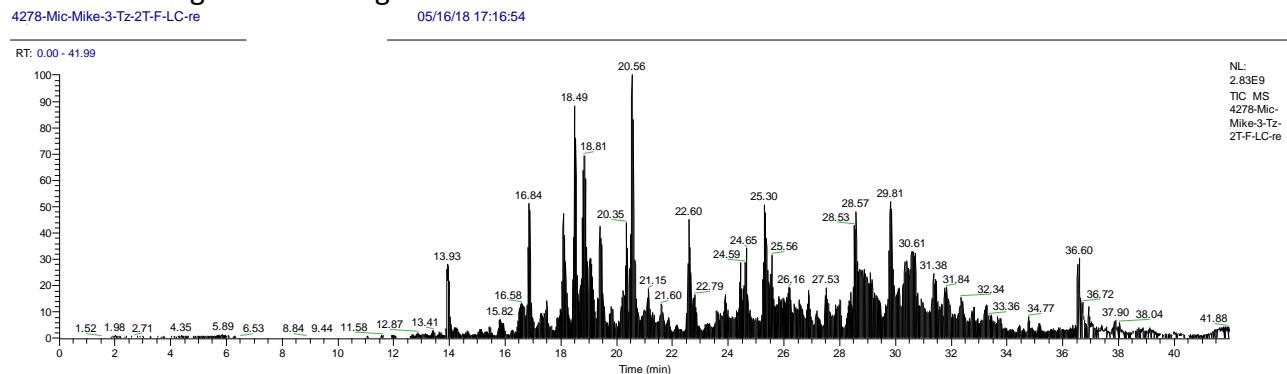
## Fragment spectra of the peptides

The following pages show the fragment spectrum of the heavy chain peptide found with a modification.

### Heavy chain peptide: LSCAASGFNIKDTYIHWVR



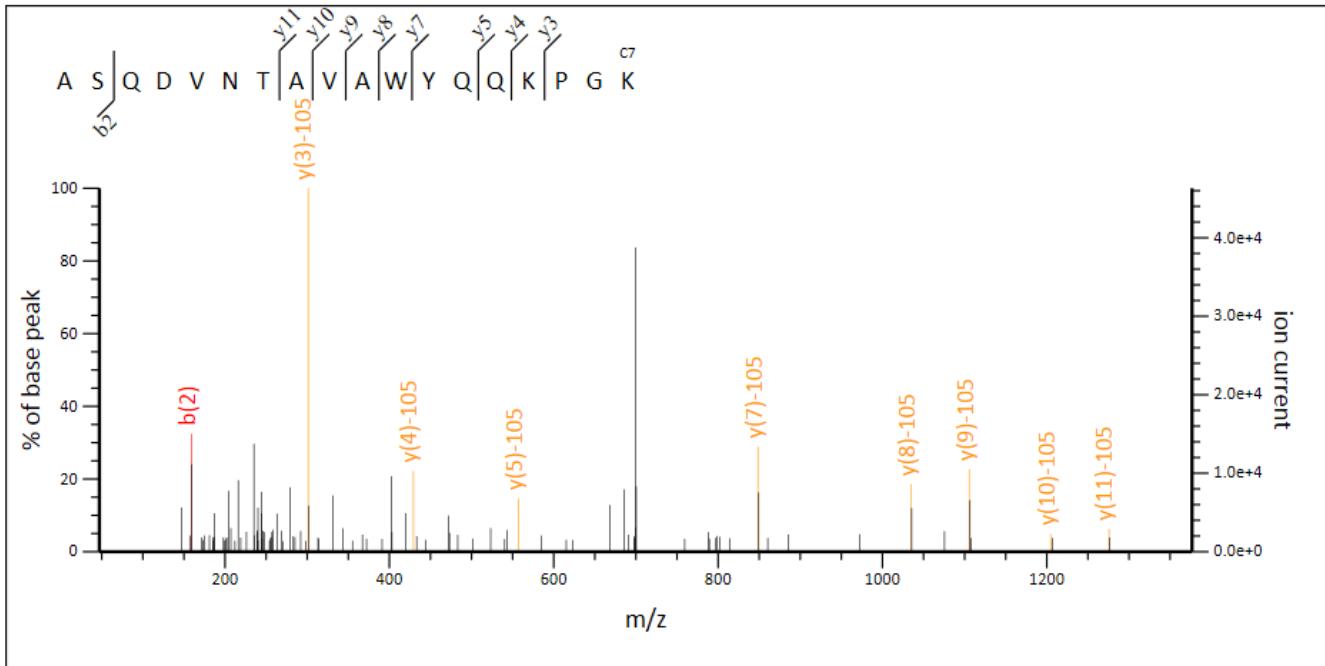
### Full chromatogram for the light chain



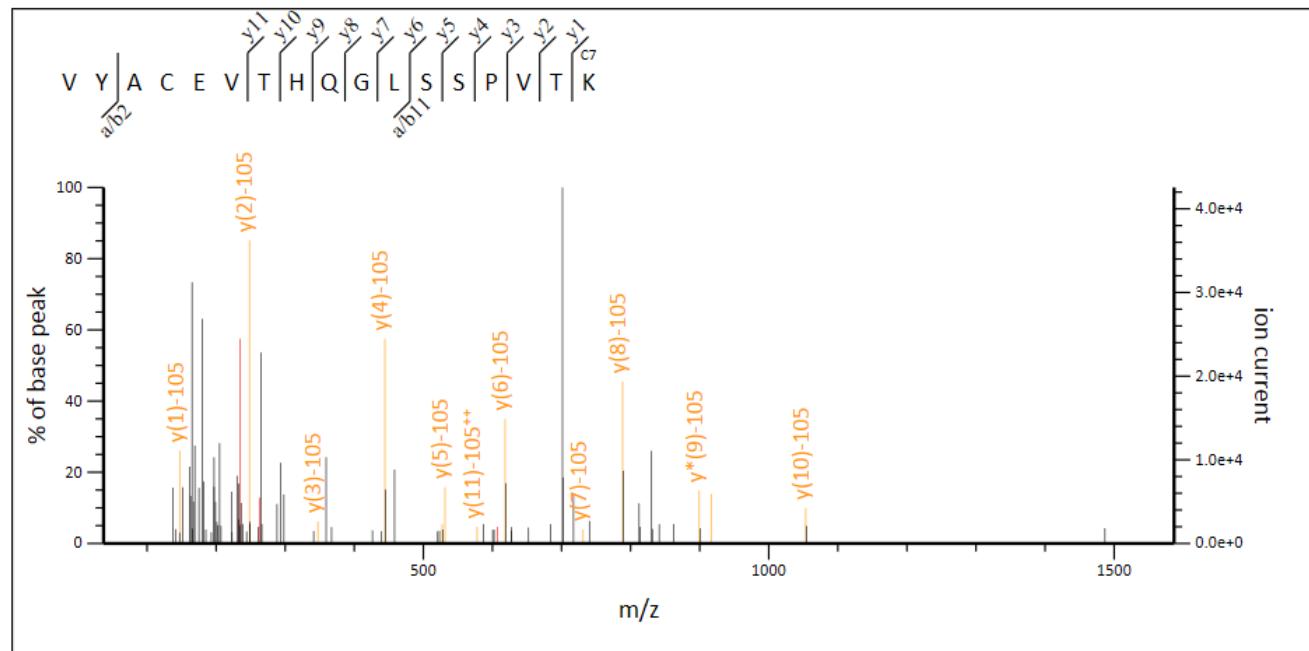
## Fragment spectra of the peptides

The following pages show the fragment spectra of the light chain peptides found with a modification.

### Light chain peptide: ASQDVNTAVAWYQQKPGK

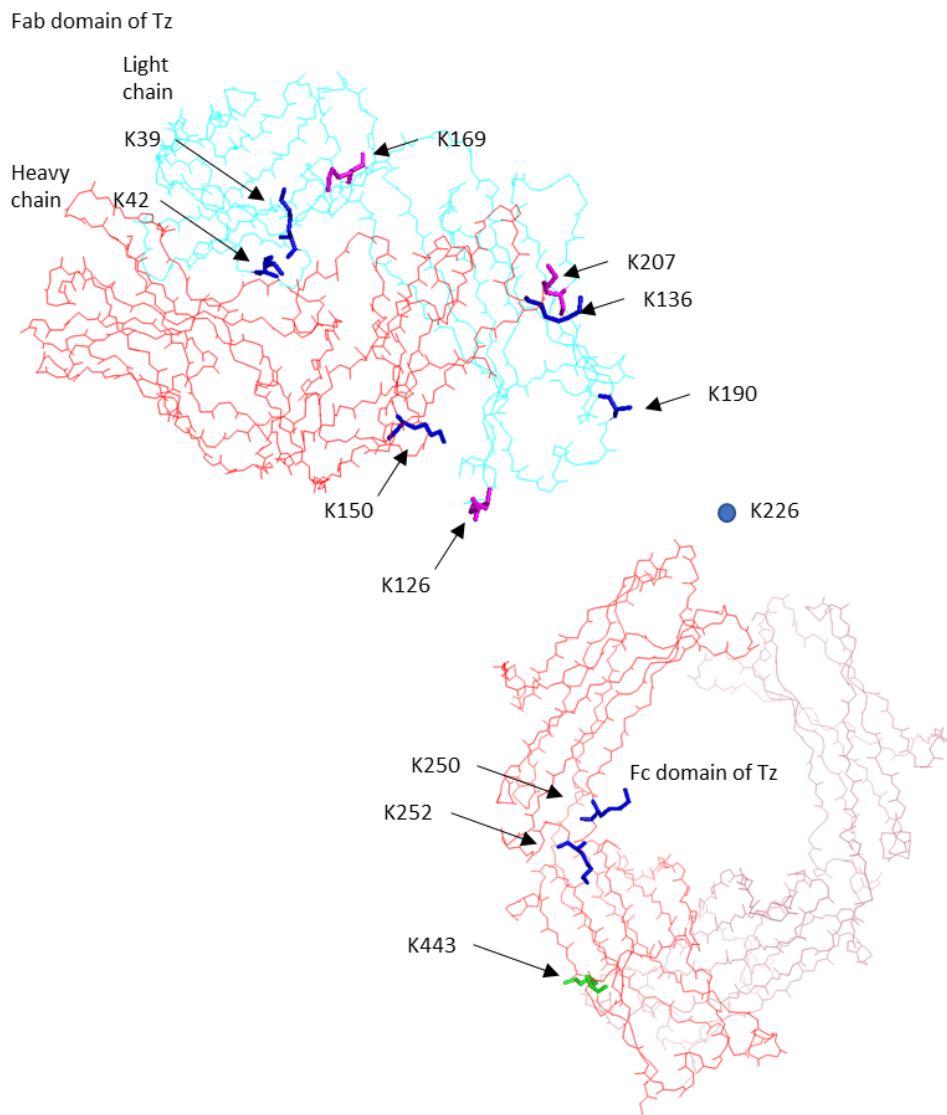


### Light chain peptide: VYACEVTHQGLSSPVTK



### Tandem MS data for the conjugate prepared using the direct conjugation method

Residue number	Probability score	Peptide sequence
Light chain		
39/42	0.0018	ASQDVNTAVAWYQQ <b>KPGK</b>
169	$2.3 \times 10^{-7}$	VDNALQSGNSQESVTEQDS <b>K</b>
190/207	$1.5 \times 10^{-8}$	<b>H</b> KVYACEVTHQGLSSPVT <b>K</b>
207	$1.2 \times 10^{-8}$	VYACEVTHQGLSSPVT <b>KSFNR</b>
126	0.00053	RTVAAPSVFIFPPSDEQL <b>K</b>
Heavy chain		
136/150	$8.2 \times 10^{-7}$	GPSVFPLAPSS <b>K</b> STSGGTAALGCLV <b>K</b>
226/250/252	$9.6 \times 10^{-7}$	SCD <b>K</b> THTCPPCPAPELLGGPSVFLFPP <b>KPK</b>
443	0.019	WQQGNVFSCSVMHEALHNHYT <b>QK</b>



The modified lysine residues are indicated on the crystal structure of the Fab and the Fc domain of Tz. Light chain is colored cyan, while the heavy chain is colored red. The other heavy chain in the Fc domain is shown in a darker shade of red without modifications. Modifications with a good probability score ( $<0.005$ ) are indicated with magenta. Lysine residues colored blue are found in the same peptide, which have a good probability score ( $<0.005$ ) for being modified. Green is used for the lysine having a moderately good probability score ( $0.05 > X > 0.005$ ). LC: Light chain. HC: Heavy chain. PDB, Fab: 1N8Z, Fc: 3D6G.

The sequence of Tz is shown below with the modified lysine residues highlighted in red. The peptides detected with a modification are underlined.

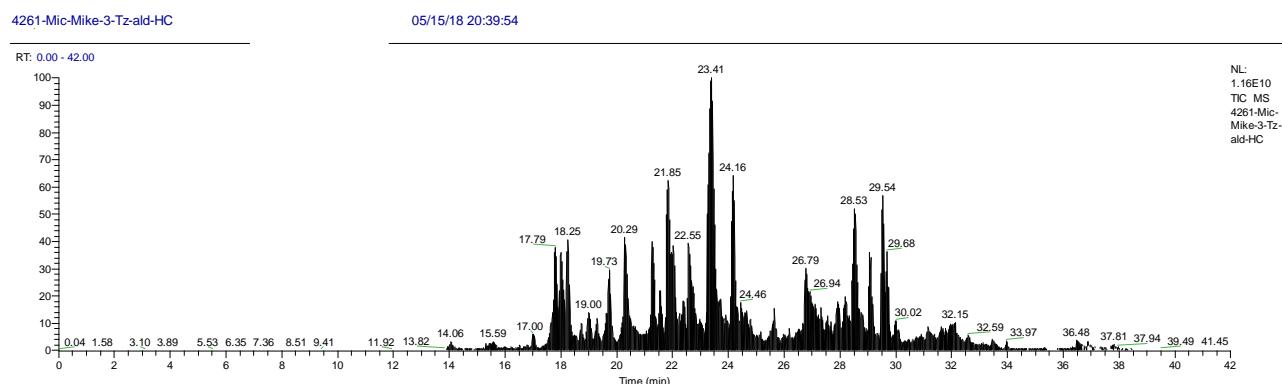
## Heavy chain sequence

10            20            30            40            50            60  
 EVQLVESGGG LVQPGGSLRL SCAASGFNIK DTYIHWVRQA PGKGLEWVAR IYPTNGYTRY  
 70            80            90            100          110          120  
 ADSVKGRFTI SADTSKNTAY LQMNSLRAED TAVYYCSRWG GDGFYAMDYW GQGTLTVSS  
 130          140          150          160          170          180  
ASTKGPSVFP LAPSSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV HTFPAVLQSS  
 190          200          210          220          230          240  
 GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP PKSCDKTHTC PPCPAPELLIG  
 250          260          270          280          290          300  
GPSVFLFPPPK PKDTLTMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN AKTKPREEQY  
 310          320          330          340          350          360  
 NSTYRVSVSL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKAKGQPREP QVYTLPPSRD  
 370          380          390          400          410          420  
 ELTKNQVSLT CLVKGFYPSDL IAVEWESNGQ PENNYKTTPP VLDSDGSFFL YSKLTVDKSR  
 430          440          450  
WQQGNVFSCS VMHEALHNHY TQKSLSLSPG K

## Light chain sequence

10            20            30            40            50            60  
 DIQMTQSPSS LSASVGDRVT ITCRASQDVN TAVAWYQQKP GKAPKLLIYS ASFLYSGVPS  
 70            80            90            100          110          120  
 RFSGSRSGTD FTLTISSLQP EDFATYYCQQ HYTPPPTFGQ GTKVEIKRTV AAPSVFIFPP  
 130          140          150          160          170          180  
SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYSLSSTLT  
 190          200          210  
 LSKADYEKHK VYACEVTHQG LSSPVTKSFN RGEc

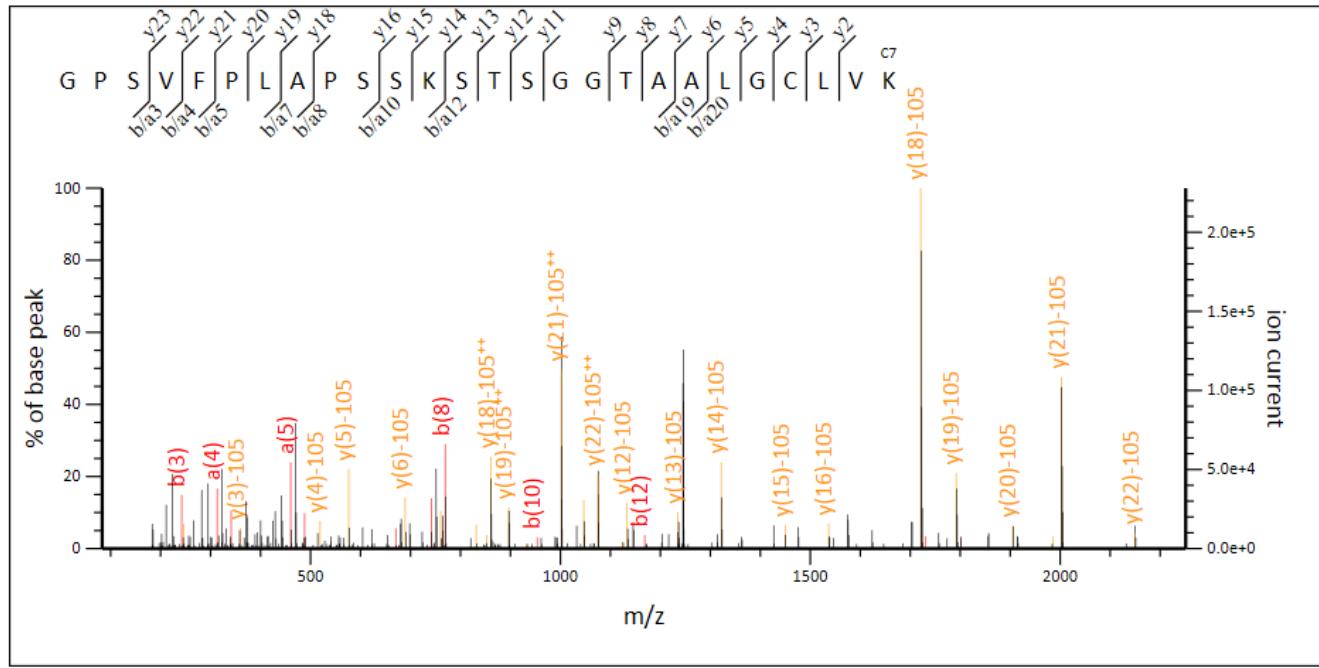
## Full chromatogram for the heavy chain



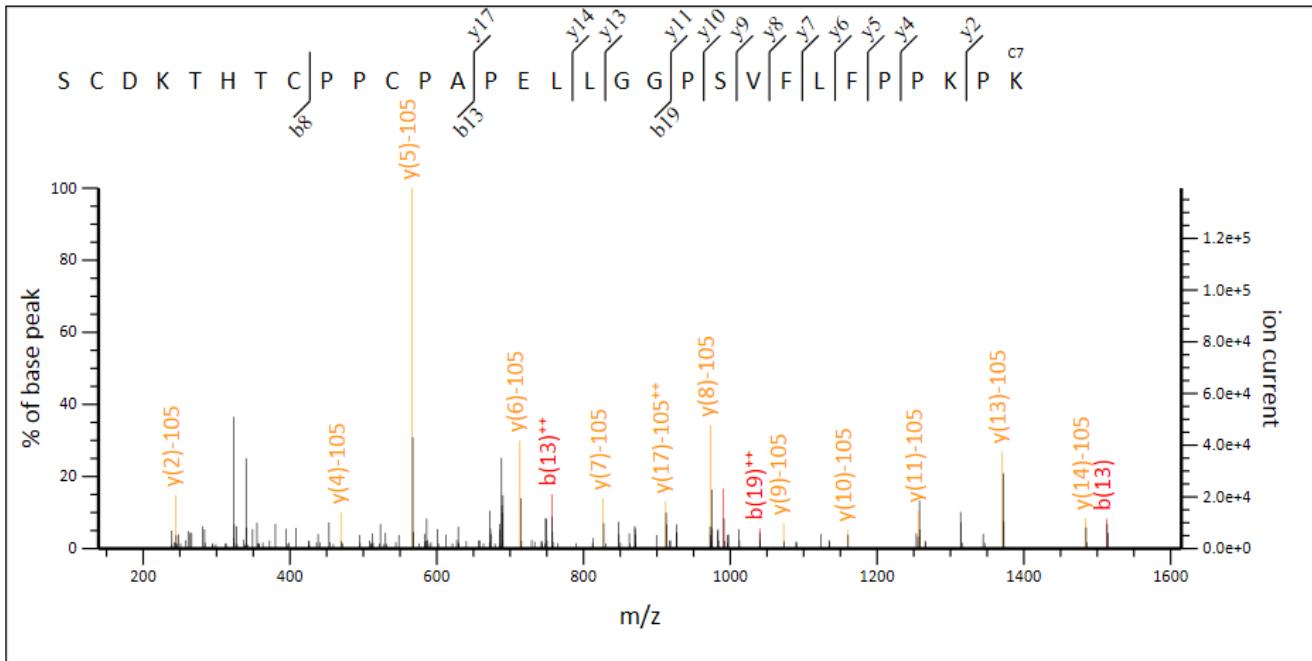
## Fragment spectra of the peptides

The following pages show the fragment spectra of the heavy chain peptides found with a modification.

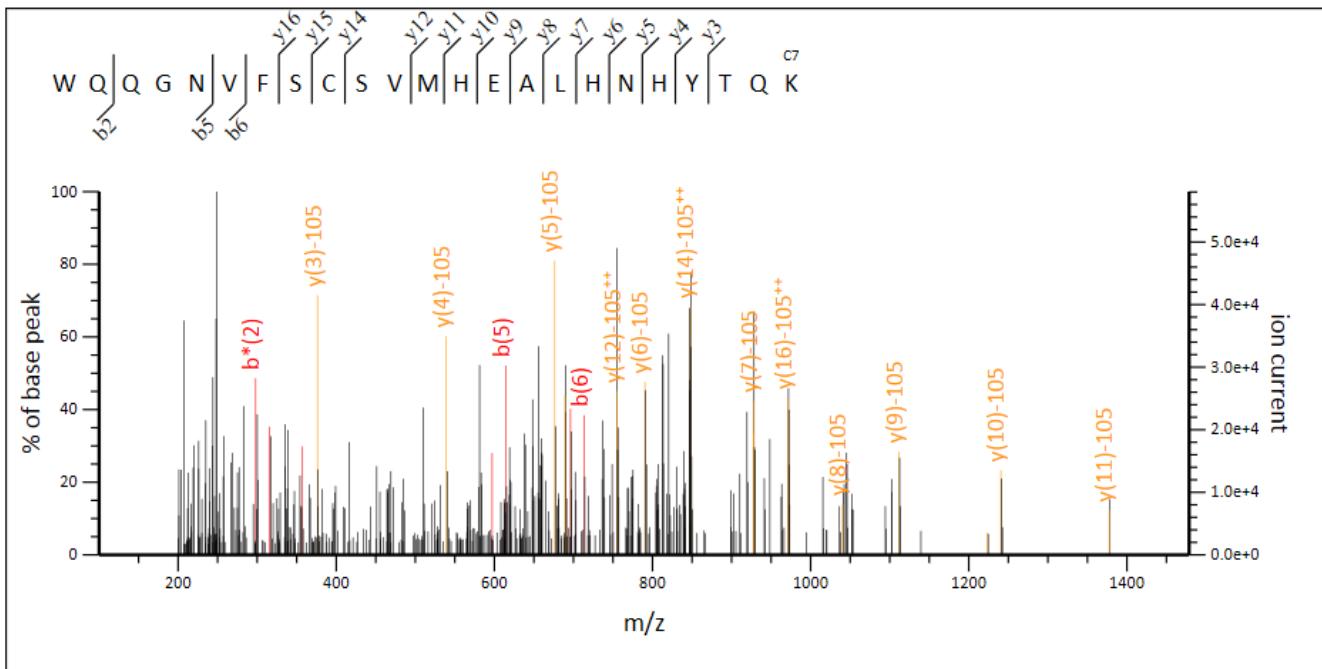
### Heavy chain peptide: GPSVFPLAPSSKSTSGGTAAALGCLVK



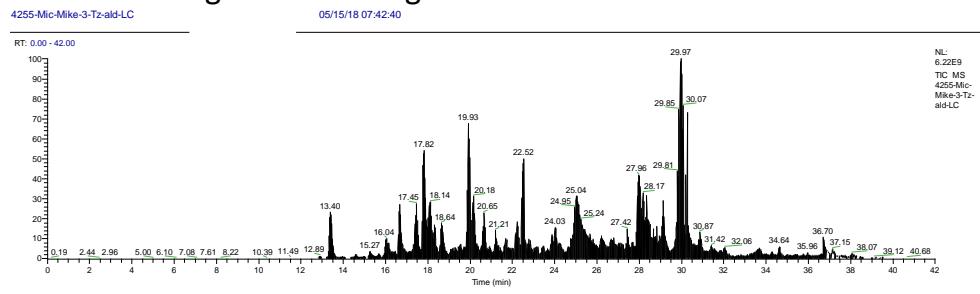
### Heavy chain peptide: SCDKTHTCPCPAPELLGGPSVFLFPPKPK



Heavy chain peptide: WQQGNVFSCSVMHEALHNHYTQK



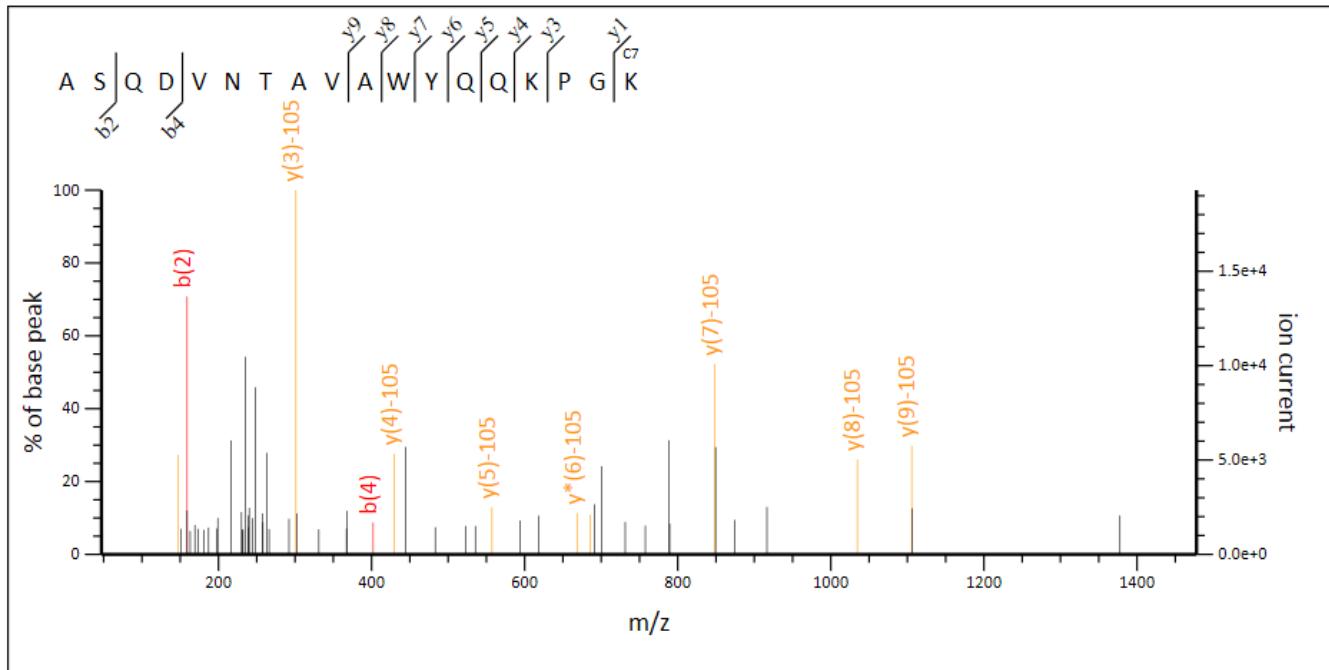
**Full chromatogram for the light chain**



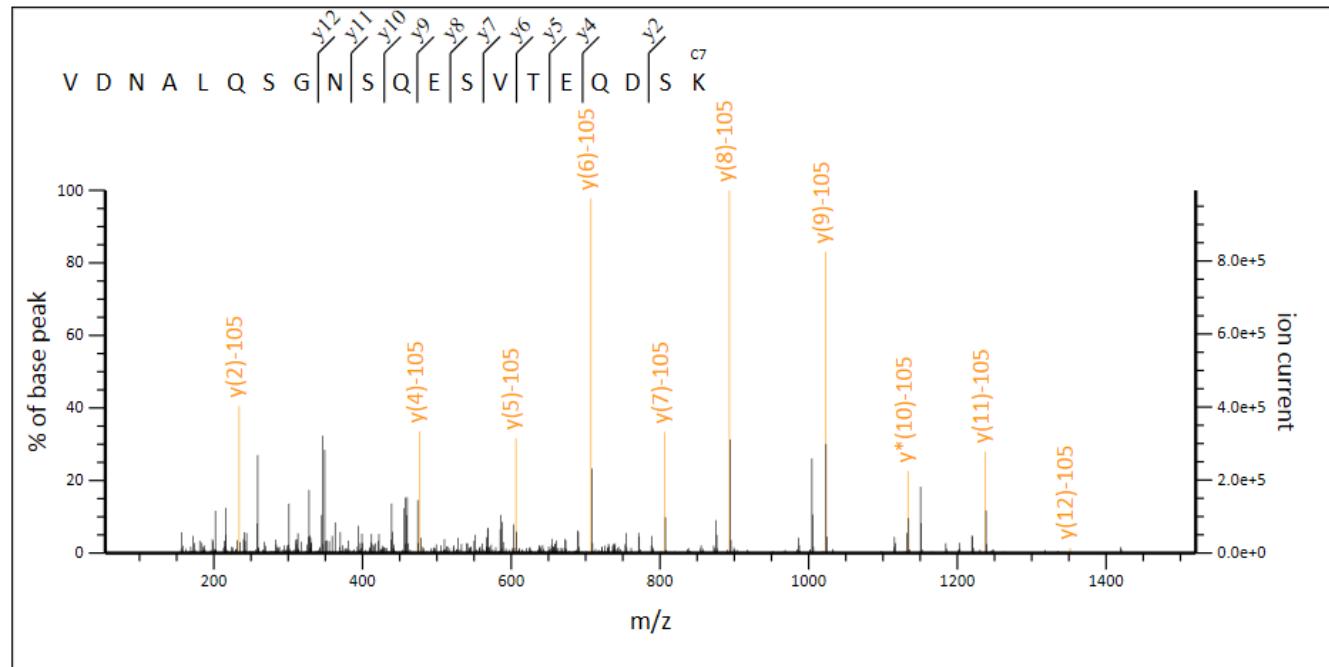
## Fragment spectra of the peptides

The following pages show the fragment spectra of the light chain peptides found with a modification.

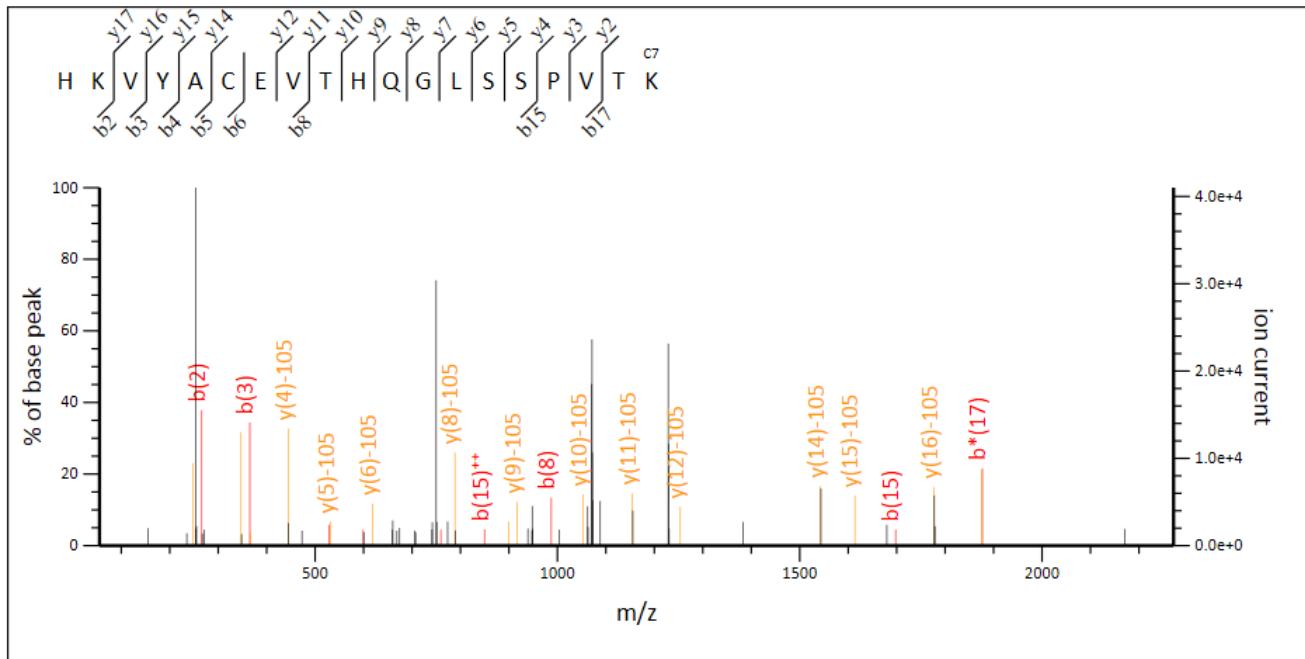
Light chain peptide: ASQDVNTAVAWYQQKPGK



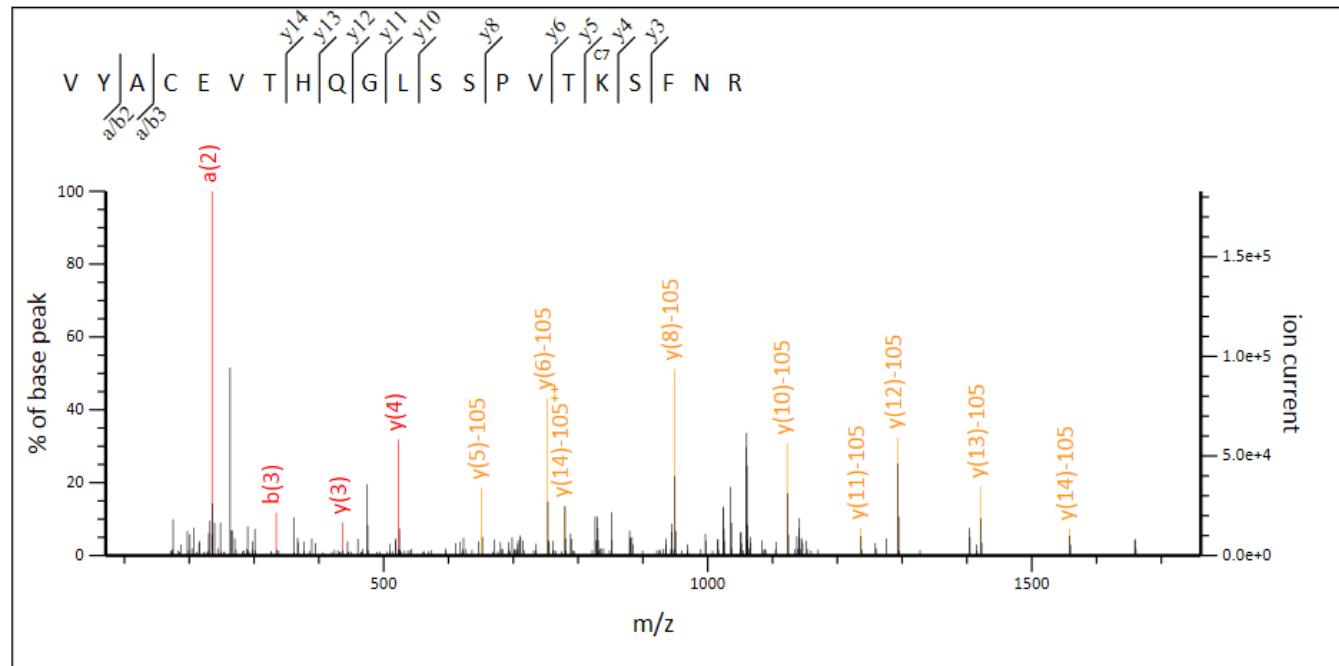
Light chain peptide: VDNALQSGNSQESVTEQDSK



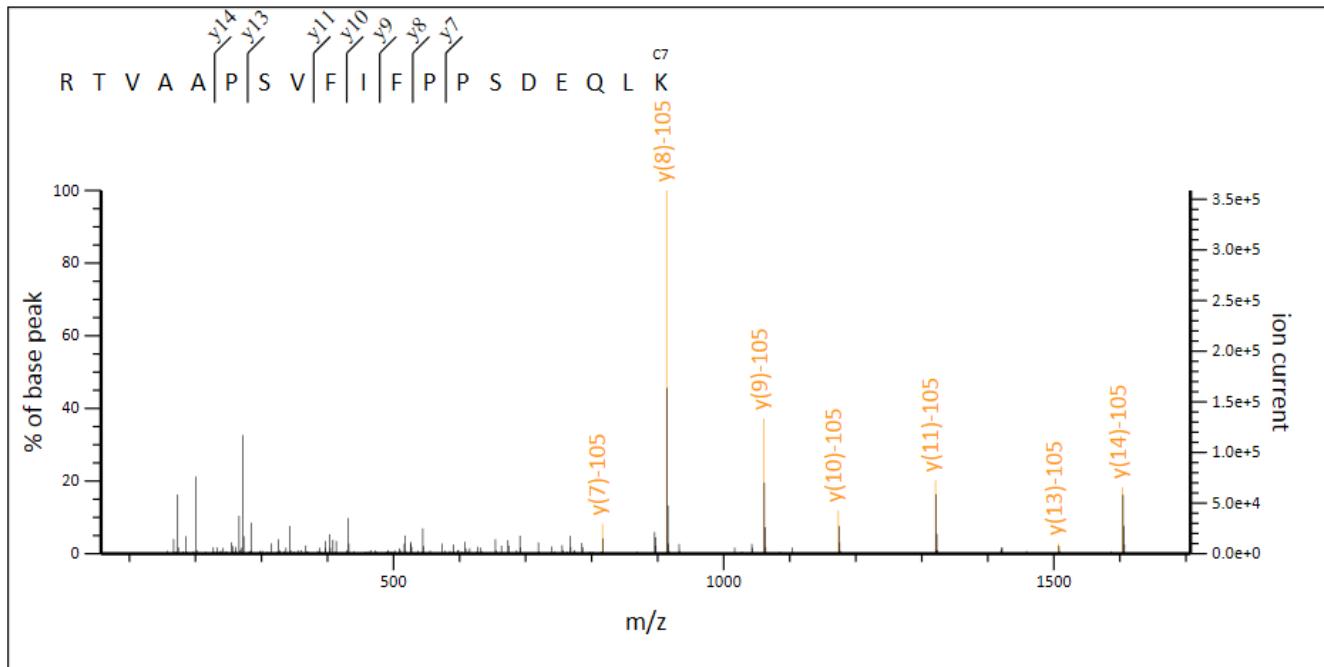
Light chain peptide: HKVYACEVTHQGLSSPVTK



Light chain peptide: VYACEVTHQGLSSPVTKSFNR



Light chain peptide: RTVAAPSVFIFPPSDEQLK



**References**

1. Kodal, A. L. B., Rosen, C. B., Mortensen, M. R., Tørring, T. & Gothelf, K. V. (2016) DNA-Templated Introduction of an Aldehyde Handle in Proteins. *ChemBioChem* 17, 1338–1342.
2. Mortensen, M. R., Skovsgaard, M. B., Okholm, A. H., Scavenius, C., Dupont, D. M., Rosen, C. B., Enghild, J. J., Kjems, J. & Gothelf, K. V. (2018) Small-Molecule Probes for Affinity-Guided Introduction of Biocompatible Handles on Metal-Binding Proteins. *Bioconjug. Chem.* 29, 3016–3025.