Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

In vitro Characterization of 3-Chloro-4-hydroxybenzoic Acid Building Block Formation in Ambigol Biosynthesis

I Dewa Made Kresna,[a,b] Luis Linares-Otoya,[a,b] Tobias Milzarek,[c] Elke R. Duell,[c,d] Mahsa Mir Moseni,[e,f] Ute Mettal,[a,b] Gabriele M. König,[e] Tobias A. M. Gulder[c,g] and Till F. Schäberle[a,b,*]

Supporting Information

Table S1. Conserved motifs of DAHP synthase Ab7.

Organism	Motif			Accession Number
	DxxHxN	KPRT	xGxR	
S. cerevisiae	I**D**YS**H**G**N**S 285	**KPRT** 117	I**G**A**R** 180	GI: 48425087
C. albicans	V**D**CS**H**G**N**S 283	**KPRT** 108	I**G**A**R** 170	GI: 3647668
E. coli	I**D**FS**H**A**N**S 271	**KPRT** 102	I**G**A**R** 165	GI: 12932715
H. influenzae	V**D**FS**H**A**N**S 279	**KPRT** 109	I**G**A**R** 172	GI: 950411
S. typhimurium	V**D**CS**H**G**N**S 272	**KPRT** 103	I**G**A**R** 166	GI: 1252280
F. ambigua	I**D**CS**H**G**N**S 273	**KPRT** 103	I**G**A**R** 166	This study

Table S2. Multiple sequence alignment of the metal binding amino acid (AA) residues of DAHP synthase Ab7.

Organism	AA residue binding to the metal ion				Accession Number
	C	H	E	D	
S. cerevisiae	GP**C**SI 78	S**H**GN 284	I**E**SN 318	T**D**AC 344	GI: 48425087
C. albicans	GP**C**SI 72	S**H**GN 282	I**E**SN 316	T**D**AC 345	GI: 3647668
E. coli	GP**C**SI 63	S**H**AN 270	V**E**SH 304	T**D**AC 328	GI: 12932715
H. influenzae	GP**C**SI 70	S**H**AN 278	V**E**SH 312	T**D**AC 336	GI: 950411
S. typhimurium	GP**C**SI 64	S**H**GN 271	I**E**SN 305	T**D**AC 330	GI: 1252280
F. ambigua	GP**C**SI 64	S**H**GN 272	L**E**SN 306	T**D**KC 331	This study

Table S3. Influence of metal salt addition on the enzymatic activity of Ab7. One unit of enzyme activity is defined as the consumption of 1 µmol PEP per minute.

Metal Salt	Activity Unit
$MnCl_2$	7.5±0.5
$MgCl_2$	7.0±1.0
$ZnCl_2$	7.5±0.5
$CuSO_4$	2.0±1.0
$CdSO_4$	5.0±1.0
EDTA	0.0±0.0

Table S4. The results of the Fre assay for measurement of UV absorbance at 340 nm. Consumption of NADH by Fre was calculated by the difference of $A_{340\ nm}{}^{1}$ minus $A_{340\ nm}{}^{2}$.

Reaction in HEPES Buffer (a)

Incubation time (min)	$A_{340\ nm}{}^{1}$ (Control)	$A_{340\ nm}{}^{2}$ (with Fre)	NADH consumption by Fre $A_{340\ nm}{}^{1} - A_{340\ nm}{}^{2}$
5	0,582	0,322	0,26
30	0,505	0,038	0,467
60	0,431	0,035	0,396
120	0,317	0,035	0,282
180	0,254	0,034	0,22

Reaction in Phosphate Buffer (b)

Incubation time (min)	$A_{340\ nm}{}^{1}$ (Control)	$A_{340\ nm}{}^{2}$ (with Fre)	NADH consumption by Fre $A_{340\ nm}{}^{1} - A_{340\ nm}{}^{2}$
5	1,332	1,105	0,227
30	1,269	0,602	0,667
60	1,203	0,455	0,748
120	1,102	0,422	0,68
180	1,021	0,411	0,61

Reaction in Tris-Cl Buffer (c)

Incubation time (min)	$A_{340\ nm}{}^{1}$ (Control)	$A_{340\ nm}{}^{2}$ (with Fre)	NADH consumption by Fre $A_{340\ nm}{}^{1} - A_{340\ nm}{}^{2}$
5	2,169	2,029	0,14
30	1,995	1,58	0,415
60	1,912	1,519	0,393
120	1,775	1,394	0,381
180	1,688	1,308	0,38

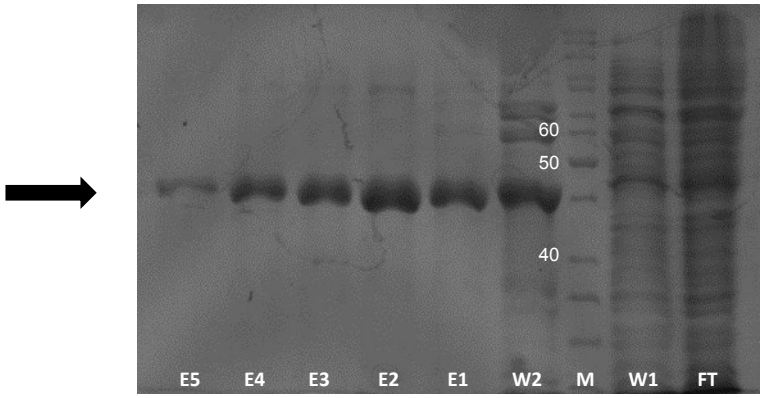


Figure S1. SDS page analysis of Ab7 after purification. FT, flow through; W1 and W2, washing steps (with 30 and 50 mM imidazole, respectively); E1 – E5, elution fractions (100- 350 mM imidazole); M, marker (kDa). The arrow indicates the band corresponding to Ab7. Elution fractions 1 – 5 were pooled, concentrated and redissolved in Tris HCl buffer (pH 7). The protein concentration was determined spectrophotometrically to be around 1 mg/mL, using Thermo Scientific NanoDrop.

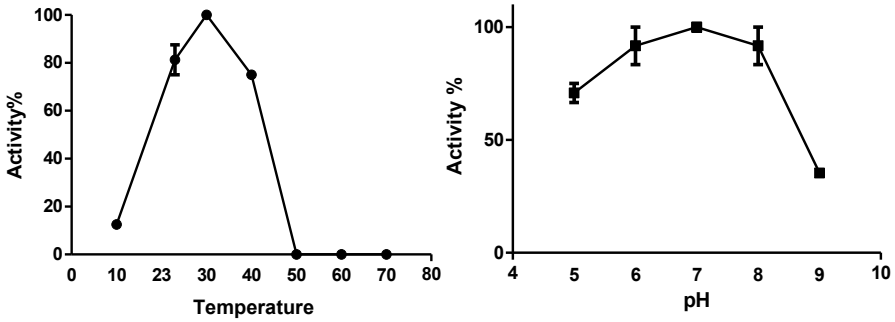
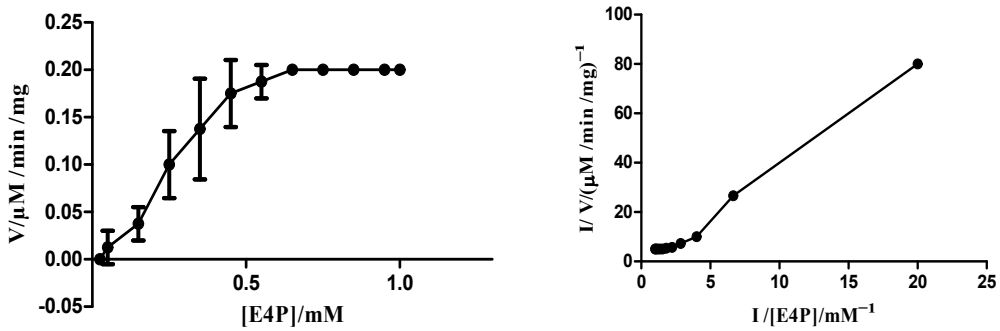


Figure S2. Effect of (a) incubation temperature and (b) pH on Ab7 activity.



(a) (b)

Figure S3. Lineweaver-Burk plot of Ab7 with various E4P concentrations (0.025 - 1 mM) and a constant concentration of PEP. Using Graf Pad Prism 5, kinetic constants of Ab7 were determined according to double-reciprocal curves. A double-reciprocal plot of Ab7 was calculated against the substrate concentration. The concentration of PEP was always preserved at 80 µM while the concentration of E4P was varied between 0.025 and 1 mM. The maximum velocity achieved by the system, at saturating substrate concentrations (v_{max}), was determined to be 0.3725 ± 0.083 (U/mg^{-1}) and the obtained K_m and k_{cat} values were shown to be 0.6993 ± 0.3746 mM and 0.0281 ± 0.063 s^{-1}, respectively.

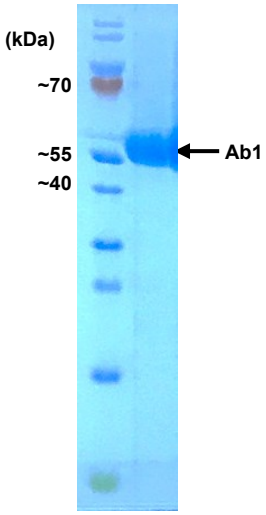


Figure S4. Purification of Ab1 as analyzed by SDS-Page.

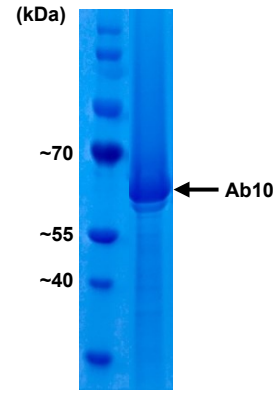


Figure S5. Purification of Ab10 as analyzed by SDS-Page.

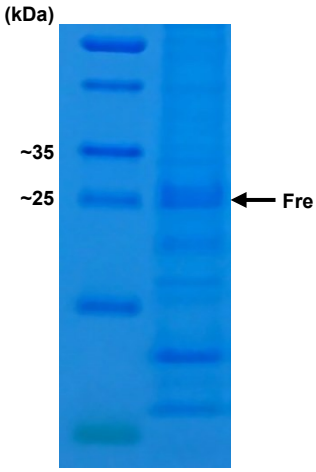


Figure S6. Purification of Fre as analyzed by SDS-Page.

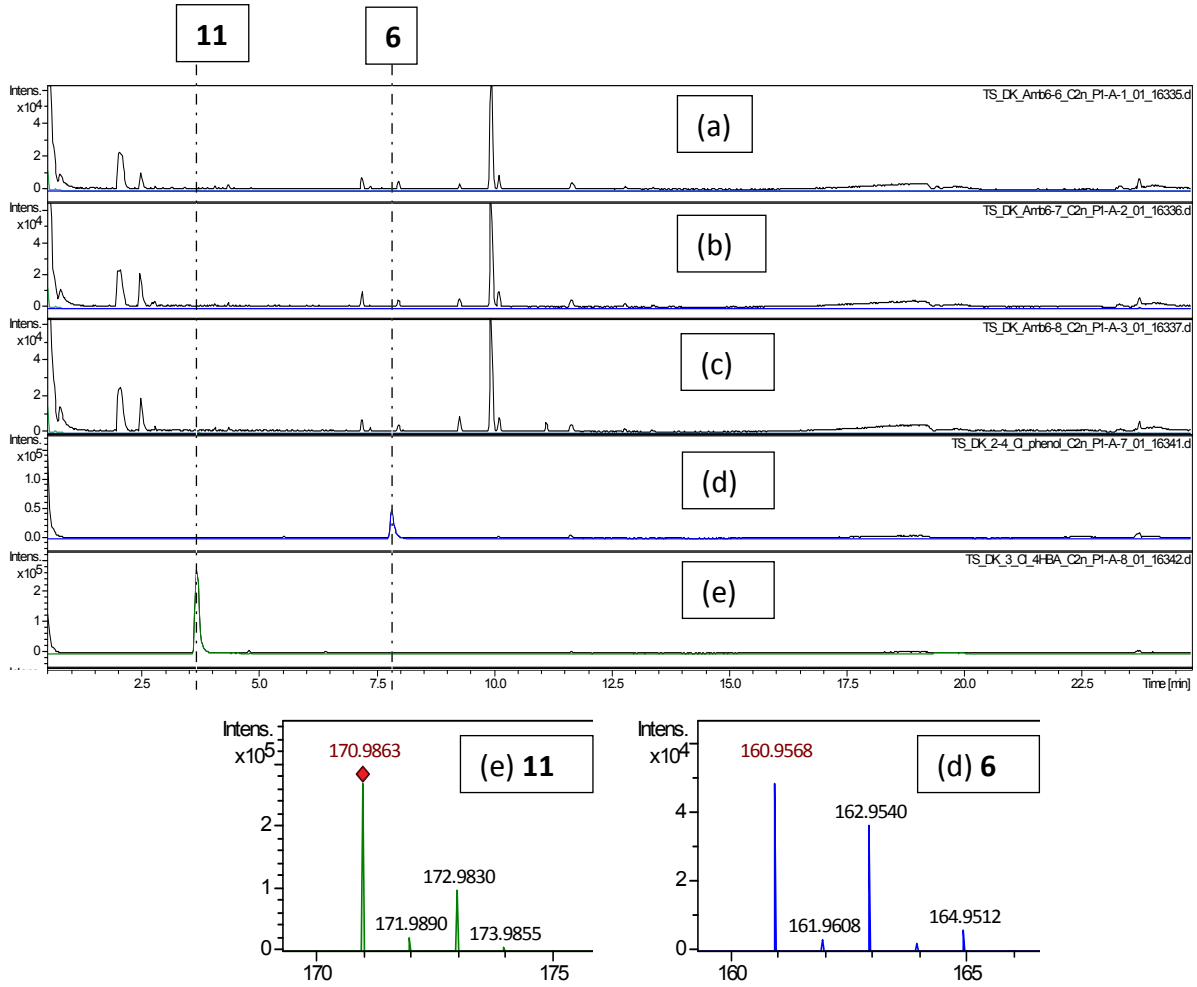


Figure S7. LCMS data for the in vitro assay using 4-HBA (**7**) as the substrate of Ab1, Ab10, or both. The assay was conducted (a) only with Ab10, (b) only with Ab1, (c) with Ab1 and Ab10. The (d) standard of 2,4-dichlorophenol (**6**), and the (e) standard of 3-Cl-4-HBA (**11**) served for comparison. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides for (d) and (e) close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Displayed in green are the extracted ion chromatograms (EIC) of the $[M-H]^-$ peak of **11** (theoretical m/z = 170.9843). Shown in blue are the extracted ion chromatograms of the $[M-H]^-$ peak of **6** (theoretical m/z = 160.9555). As apparent from the images, neither substance **6** nor **11** could be detected.

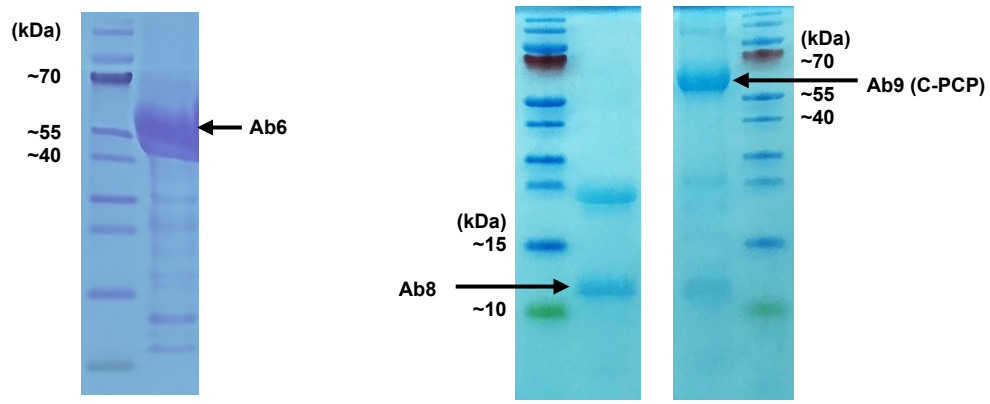


Figure S8. Purification of (a) Ab6, (b) Ab8, and (c) Ab9 (C-PCP) as analyzed by SDS-Page.

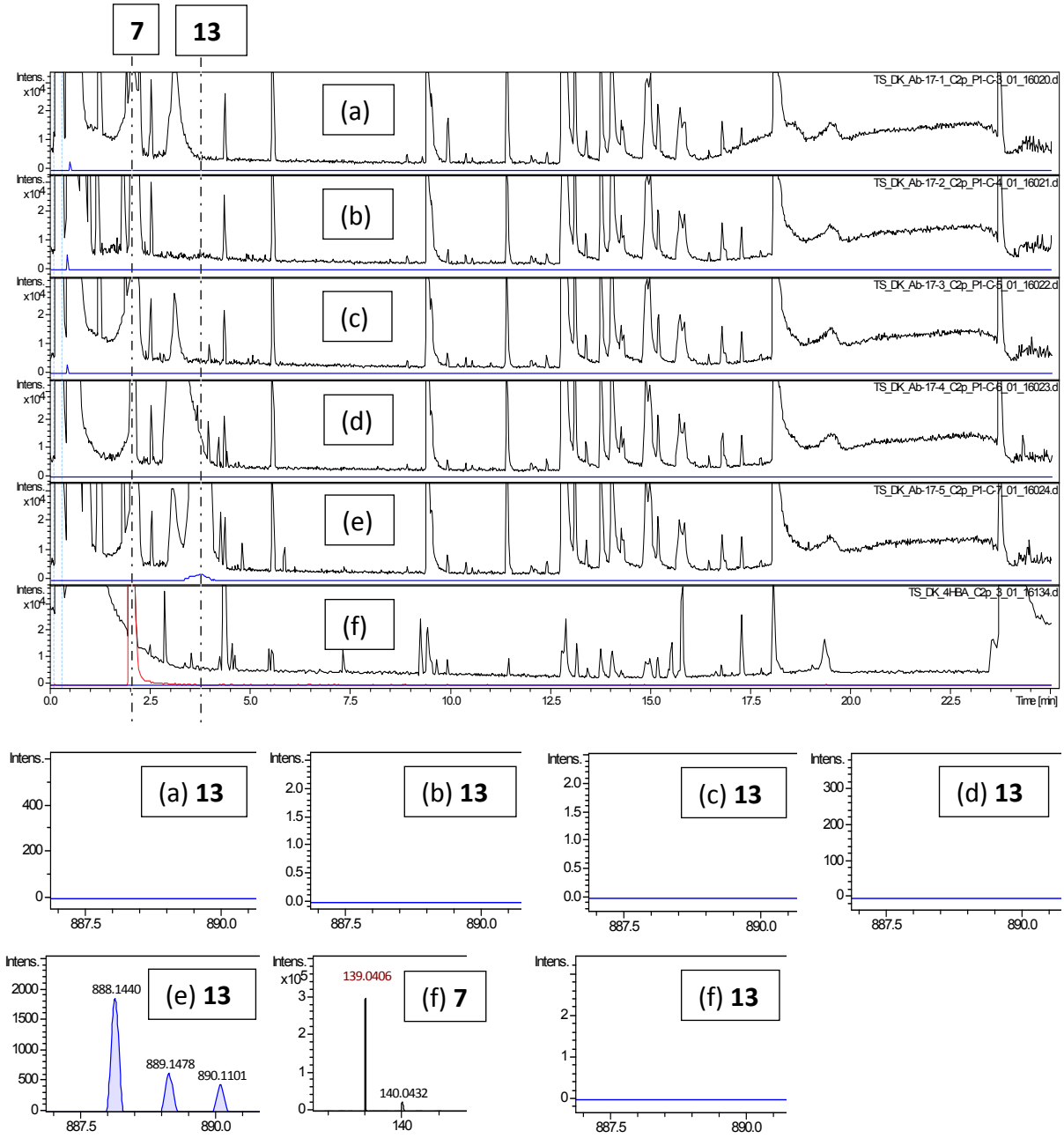


Figure S9. LCMS data of the Ab6 *in vitro* assay. The assay was performed (a) without ATP, (b) without CoA**,** (c) without the substrate 4-HBA (**7**), (d) without Ab6, and as (e) complete reaction. For comparison a (f) standard of **7** was used. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. The red chromatogram in (f) illustrates the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of the substrate standard **7** (m/z theoretical = 139.0395). The blue chromatogram shows the EIC of the $[M+H]^+$ peak of the product 4-HBA-CoA (**13**) (m/z theoretical = 888.1436).

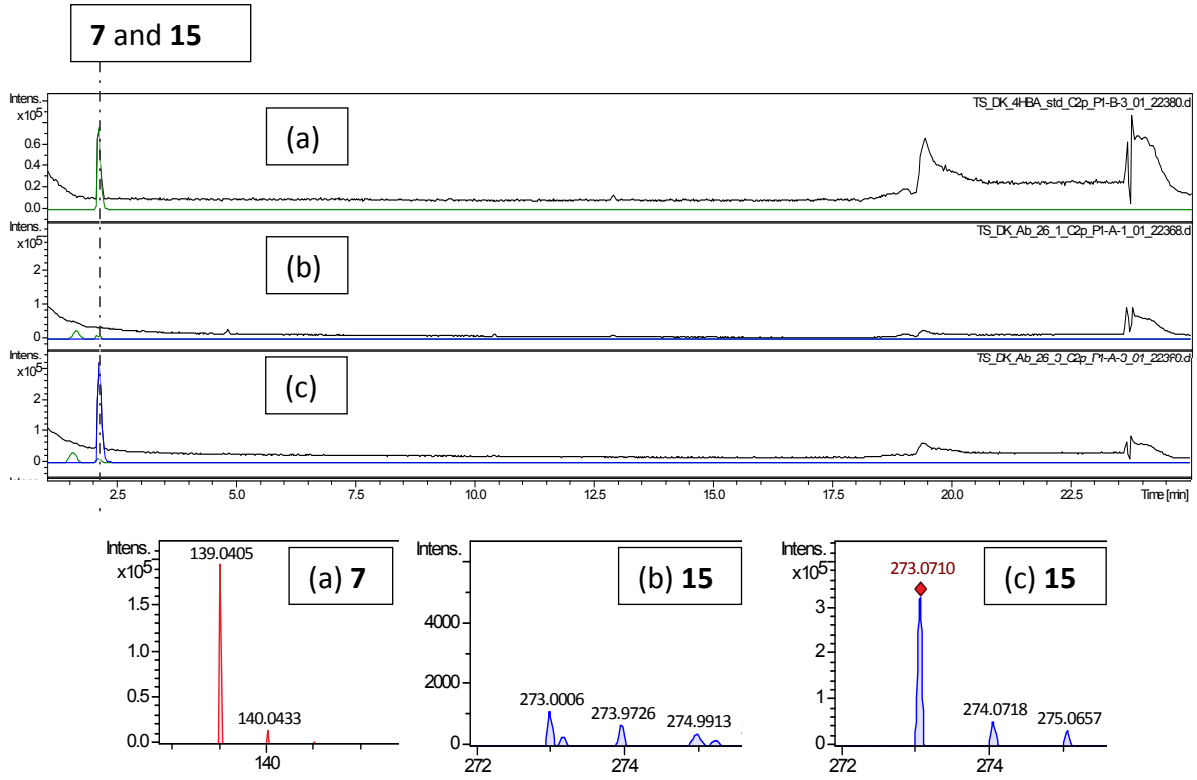


Figure S10. LCMS analysis of the control experiment for the *in vitro* carrier protein assay. For this control trial the substrate 4-HBA (**7**) was incubated with cysteamine in the presence or absence of Ab6. The (a) standard of substrate **7** served for comparison. The assay was performed (b) without Ab6, and (c) with Ab6. Mass spectra were measured in positive mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Displayed in green are the extracted ion chromatograms (EIC) of the $[M+H]^+$ peak of the substrate **7**. The blue chromatogram shows the EIC of the $[M+H]^+$ peak of the product "4-HBA-cystamine" (**15**) (theoretical m/z = 273.0726). As apparent from the images, product **15** could be detected in sample (c).

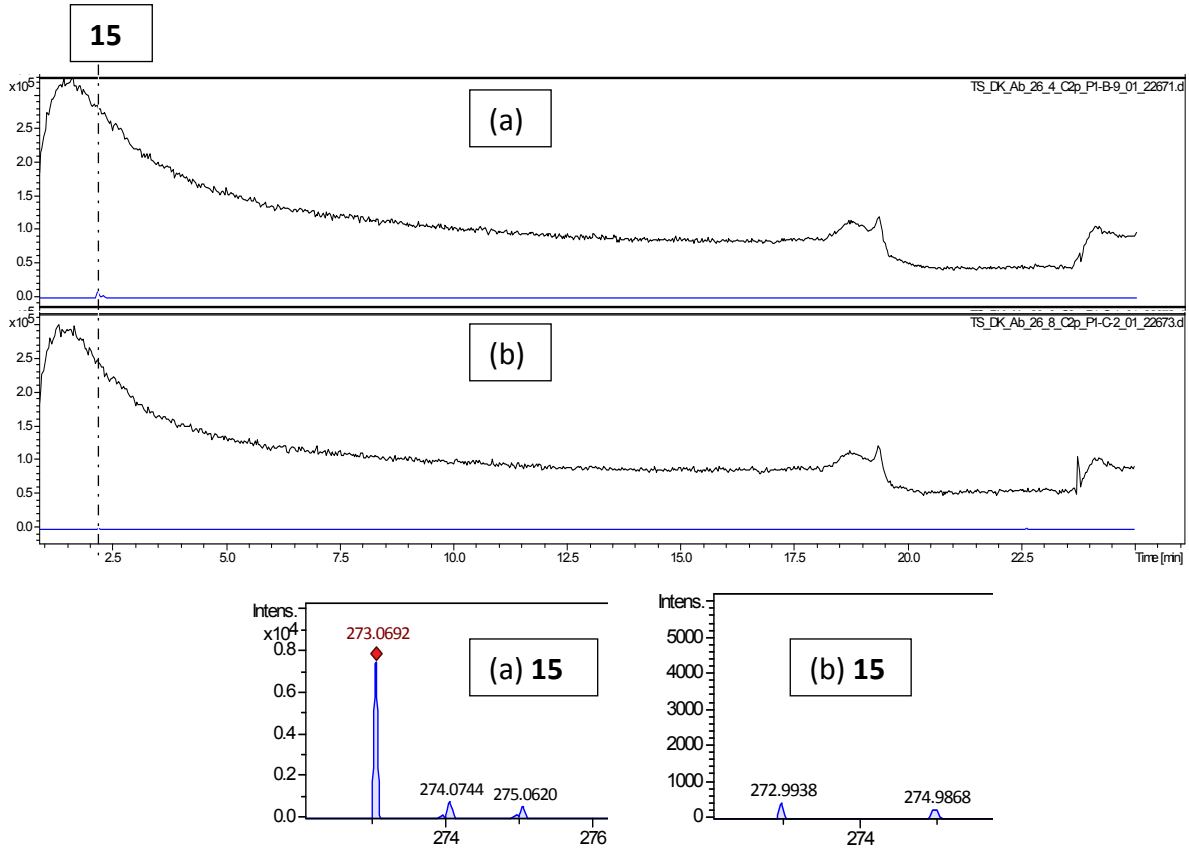


Figure S11. LCMS analysis of the Ab8 and Ab9 (C-PCP) carrier protein *in vitro* assay. The assay was performed in the presence of Ab6 using 4-HBA (**7**) as substrate. Prior to cysteamine treatment, Ab8 and Ab9 (C-PCP) were purified by HPLC. Mass spectra were measured in positive mode. In the assay either (a) Ab8 or, (b) Ab9 (C-PCP) served as a carrier protein. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Displayed in blue are the extracted ion chromatograms (EIC) of the $[M+H]^+$ peak of the product "4-HBA-cystamine" (**15**) (theoretical m/z = 273.0726). As apparent from the images, product **15** could be detected only when Ab8 was used as a carrier protein.

```
CLUSTAL O(1.2.4) multiple sequence alignment

Ab10 ------------------------------------------------------------
McnD MKTVEFLAYLNSLQINLWAENDKLRYRSPQGVMTPELLGKLKERKEQLIALLRQKAEDLG 60
ApdC MKTVNFLSHLNDLGINVWVENDKLRYRSPKGVIIPELLQELKERKEELIAFLRQQAEDLN 60
AerJ MKTVEFLSDLNHLGVTIWMEGDKLRYRSPQGVMTPDLLEQLKEHKEELIVLLREQADNFS 60


Ab10 MKNIYDVAIC**G**S**G**LA**G**LTLARQLKLKMPDISVVVLDRLARPLPEAGFKVGESSVEVGAFY 60
McnD EAEVYDVVIC**G**G**G**LA**G**LTLARQLKLQKLNISVIVLDKIARPLPEASFKVGESTVEVGAFY 120
ApdC QAETYDVVIC**G**G**G**LA**G**LTLARQLKLQKPNMAIAVLDKMSGLSPEASFKVGESTVEVGAFY 120
AerJ SETDYDVAIC**G**G**G**LA**G**LTLGRQLKLKQPNLSVVVLDKMARPLPEAGFKVGESTVEVGAFY 120
         ***.***.*******.*****:  :::: ***:::   ***.******:*******
                **GxGxxG motif**

Ab10 LAHIVQLEDYLEKQHLHKLGLRYFLGDTKGPFHKRPEIGLSKYHFPNSYQLDRGKLENDL 120
McnD LANTLQLTDYFEEQHLVKLGLRYFFNNSATNFQDRPELGLSEFHLPNSYQIDRGKLENDL 180
ApdC LANTLQLTDYFEEQHLVKLGLRYFFNNSATNFQERPELGLSEFHAPNSYQIDRGKLENDL 180
AerJ LANTLQLTDYFDHQHLPKLGLRYFFKPQETEFHKRPELGLSEFHAPKSYQIDRGKLENDL 180
     **: :** **::.*** *******:      *:.***:***::* *:***:*********

Ab10 RSINTEAGVELLEGCLVKDIELGDP-QQLHQIIYTQEN--NKATQAIQARWVVDSMGYRR 177
McnD RAFNVEAGVELRENCLVNEIELAVGLQQHHKIVYTQDKGDHKKTNVIQARWVVDAMGRRR 240
ApdC RQFNMEAGVELREGCLVNEIALAEGLQQHHKVVYTQGDGDNRKNKIIKARWVVDAMGRRR 240
AerJ RQFNIEAGIDLKENCSVKDIEFAEGLQQQHKIIYTQGSGANQKTHCIKSRWVVDAMGRRR 240
     * :* ***::* *.* *::* :.   ** *:::*** .  :: .: *::*****:** **

Ab10 FLQRKLGLAKPKNSQFSAVWFRVEGRFDVSDFVPSTEIEWHERVPHNNRYYSTNHLCGEG 237
McnD FLQKKLGLDKPNNAQFGAVWFRVEGRFDISDFVPSSEEKWHNRVPNKNRYYSTNHLCGEG 300
ApdC FLQRKLGLDKPNNDNFGAVWFRVNGRFDIGDFVPSSEEKWHNRVPNKNRYYSTNHLCGEG 300
AerJ FIQKKLGLAKPNHNNYSAVWFRVEGRFDVSNFVPASEEKWHRRVPNNNRYYSTNHLCGEG 300
     *:*:**** **:: ::.******:****:.:***::* :**.***::**************

Ab10 Y**W**V**W**L**IP**LSTGYTSIGIVTNEEIHPFGTYHTYEKAFQWLEKHEPVVAFHLKSNPPVDFMK 297
McnD Y**W**V**W**T**IP**LSTGHTSIGIVARQDIHPLKTYYNYELAYQWLQKNEPTLAFHLADKQPEDFRK 360
ApdC Y**W**V**W**T**IP**LSTGHTSIGIVARQDIHPLKNYYNYELAYQWLQKNEPVLACHLKDKEPEDFRK 360
AerJ Y**W**V**W**L**IP**LSTGYTSIGIVARQDIHPLKNYHNYELAFQWLRENEPVLAAHLEGKSPEDFRK 360
     **** ******:******:.::***: .*:.** *:***.::**.:* ** .: * ** *
      **WxWxIP motif**

Ab10 IPQYSYSSNQVFSINRWACVGVAGVFADPFYSPGTDLIGFGNSLITQMVELDRENKLTPE 357
McnD MPKYSYSSKQVFSYNRWACVGEAGTVPDPFYSPGTDNIGFGNSLTTQLITLDLEGKLTQE 420
ApdC MPKYSYSSKQVFSSNRWACVGEAGTFPDPFYSPGTDNIGFGNSLTTQLIALDLEDKLTKE 420
AerJ MPKYSYSSKQVFSFNRWACVGEAGLFPDPFYSPGSDSIGFGNSLTTQMIELDLKGQLTPE 420
     :*:*****:**** ******* ** . *******:* ******* **:: ** :.:** *

Ab10 IVNEANRFLITYNESVTSNIHNAYLCFGNETVMVMKYIWDVLSAWAFSAPMMFNSLFLDS 417
McnD KVKDANHFYLSYSDGVTLNIQNAYNCLGNGMVMATKFIWDTLSGWTFSGLMMFNSIYLDQ 480
ApdC KVEDANYFYISYSDGVTLNIQNAYNCLGNGIVMATKFIWDTLSGWTFSGLMMFNSIFLDQ 480
AerJ RVDDANHFYLTYHDGLTFNIQNSYNCMGNGIVMATKMIWDTLAGWTFGCLMMFNSIFLDP 480
      *.:** * ::* :.:* **:*:* *:**  **. * ***.*:.*:*.  *****::**

Ab10 DKRAKVRKGTGQFFLLAQRMNQLFRDWAVQSQRRTSFEFIDYLQIPFVRELRARNLKTNK 477
McnD EFRTKVQQISSKFFPLSYRMQQLFKDWANQSLHRVNFEFIDYLAIPFVEELRSRNLRYNQ 540
ApdC DFRTRVQEINSKFFPLSYRMQQLFRDWANKSLNRVNFEFIDYLAIPFVDELRTRNLKSNK 540
AerJ ELKMKVQQINAEFFPLSYRIQQLFRDWANQSLGRVSFEFIDYLAIPFVNELRTRNLQSNK 540
     : : :*:: ..:** *: *::***:*** :*  *..******* **** ***:***: *:

Ab10 TEQELIDDHLASIKLFEELAQVIFLLALEDTMPEKSADFPSPVWLNAWVVSLDDKRWEID 537
McnD TESEIIENYLSSLKLLEEVAQVMFHLALEDIMPEMLPKVSSNSWLNAWAISLNASKWESD 600
ApdC TKSEIIENYLSSLKLLEEVAQVMFHLALEDTMPDMLSKVNSNSWLNAWAISLDISKWEAD 600
```

```
AerJ TEAEIIEAYRASLKLLEEFAQVIFQIALEDTSPELLTKINEHPWLNAWAISLDASKWETD 600
     *: *:*: : :*:**:**.***:* :****  *:   .. .  *****.:**: .:** *

Ab10 GLFRPTSKPRDLRPMMEQLWQNIHFRAADRDSSLITA 574
McnD GLFAPKSEPRNLDLIKQQLWEAIGK------------ 625
ApdC GLFTPKSEPRNLNLIKQQLWEAIGK------------ 625
AerJ GLFSPKNEPRNLHSIKKQYLEAINR------------ 625
     *** *..:**:*  : :*  : *
```

Figure S12. Alignment of Ab10 with McnD, ApdC and AerJ using CLUSTALO (1.2.4) multiple sequence alignment (https://www.uniprot.org/align/). GxGxxG and WxWxIP motifs are highlighted in bold letters.

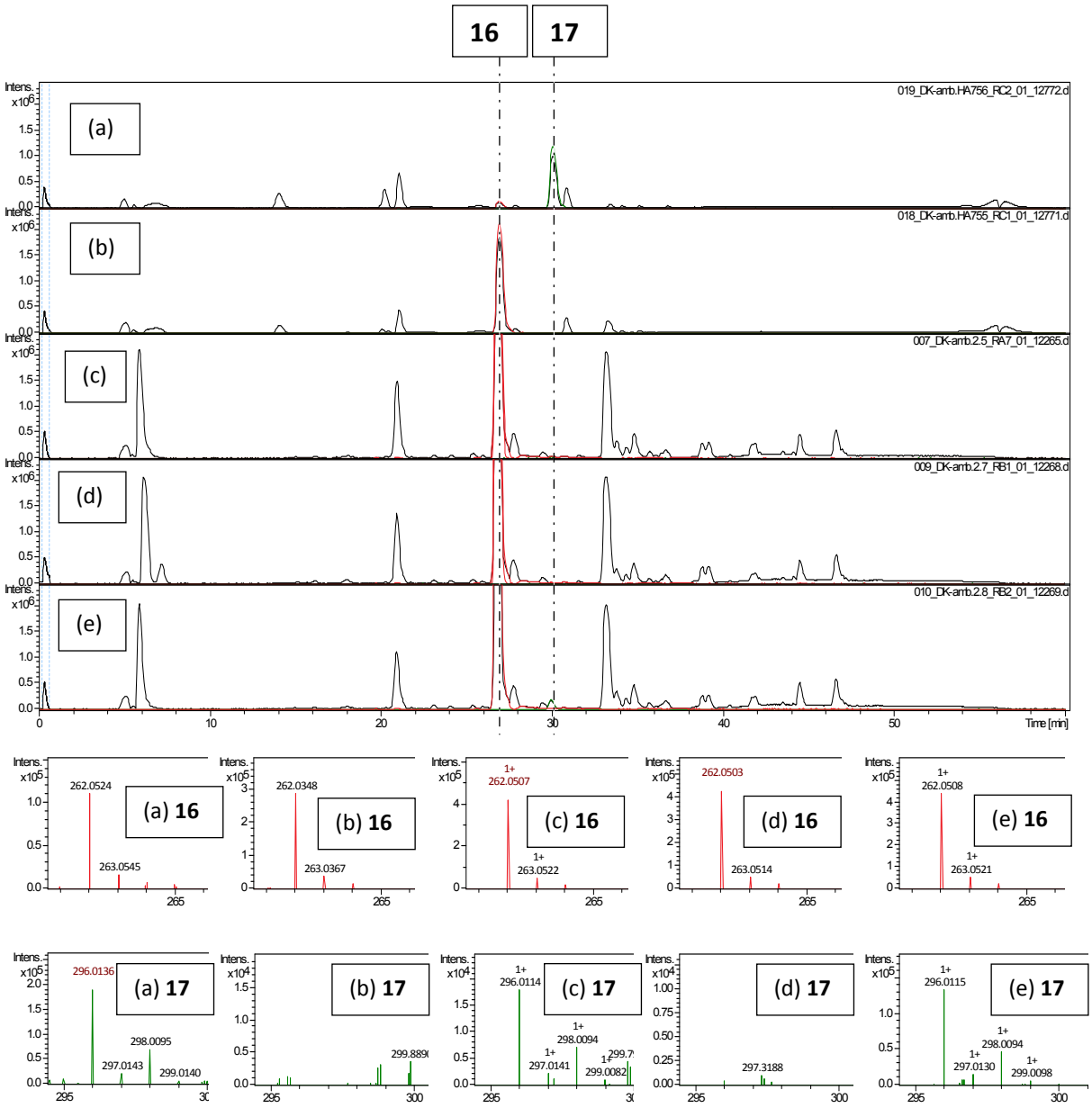


Figure S13. LCMS analysis of the Ab10 *in vitro* assay. For comparison a (a) standard of the product 3-Cl-4-HBA-SNAC (**17**), and a (b) standard of substrate 4-HBA-SNAC (**16**) were used. The assay itself was performed (c) without Fre, (d) without Ab10, and as (d) complete reaction. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Displayed in red are the extracted ion chromatograms (EIC) of the $[M+Na]^+$ of the substrate **16** (m/z theoretical = 262.0508). The green chromatogram shows the EIC to $[M+Na]^+$ of the product **17** (m/z theoretical = 296.0119).

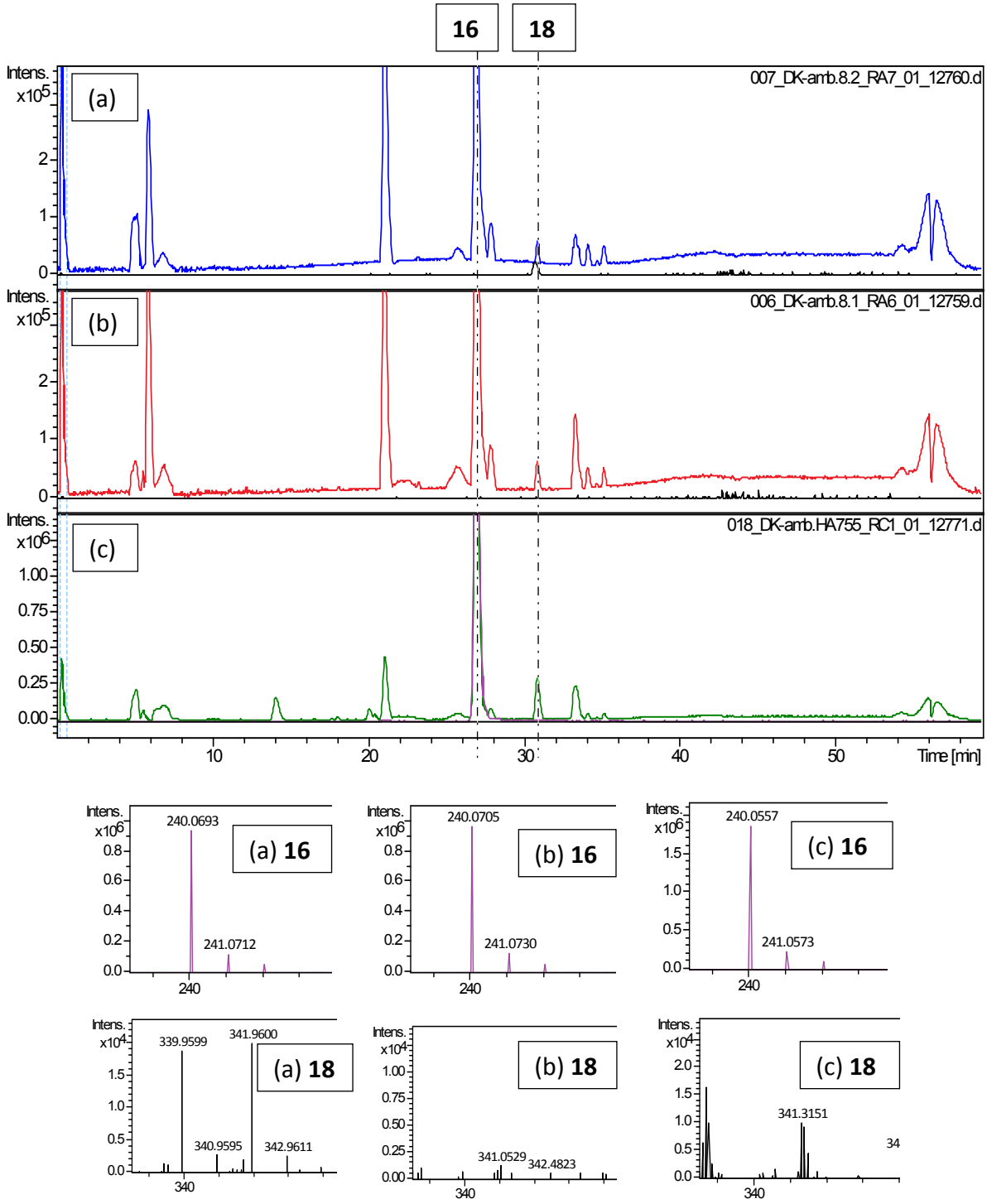


Figure S14. LCMS analysis of the *in vitro* assay of Ab10 using bromide as the halogen source for investigating Ab10 substrate specificity. The assay was performed as a (a) complete reaction, and (b) without Ab10. A (c) standard of the substrate 4-HBA-SNAC (**16**) was used for comparison. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) are shown in blue, red and green, respectively. The purple chromatogram in (c) illustrates the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of the substrate **16** (m/z theoretical = 240.0694). Black chromatograms in "a", "b" and "c" show the EIC of the $[M+H]^+$ peak of the product 3-Br-4-HBA-SNAC **18** (m/z theoretical = 339.9619).

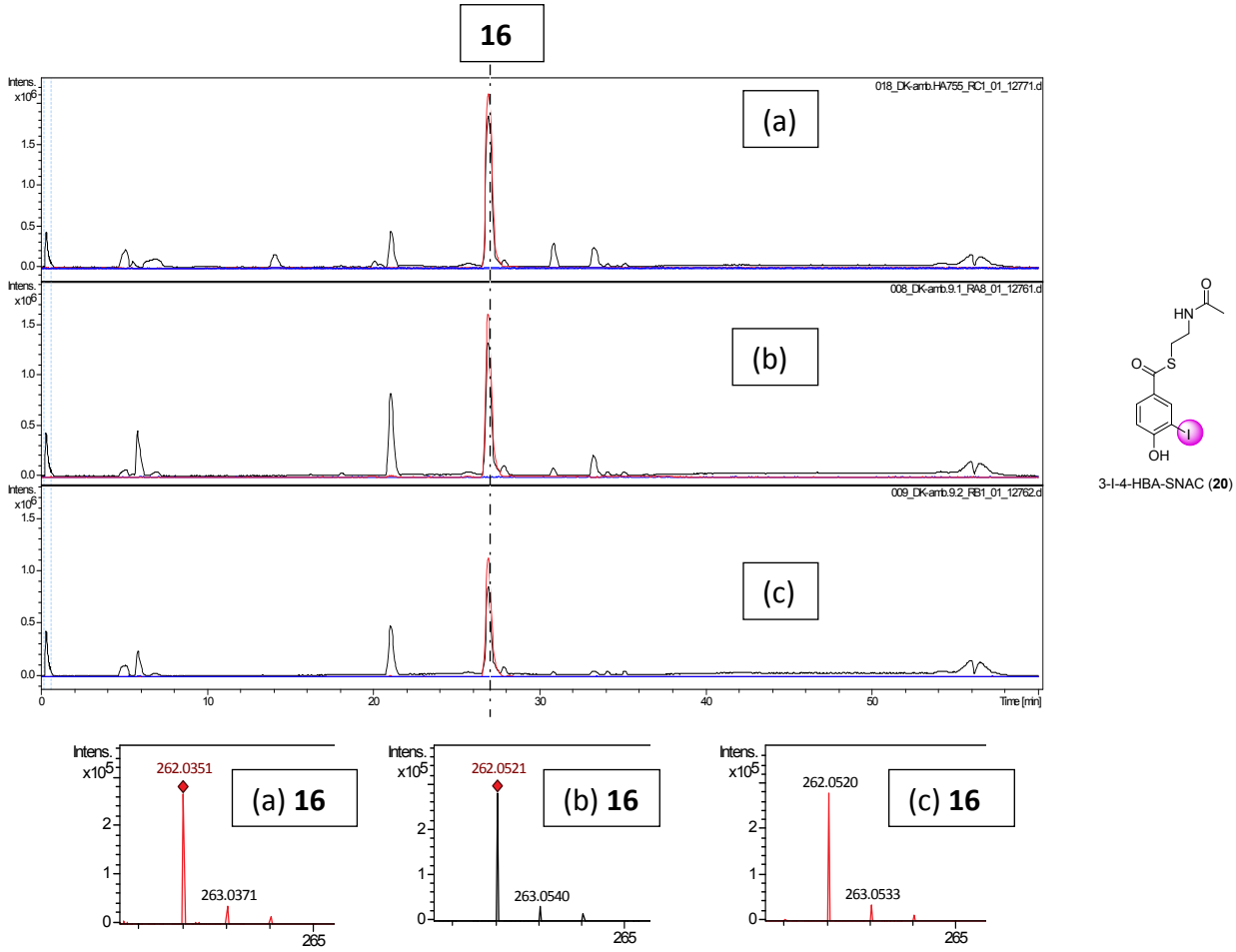


Figure S15. LCMS analysis of the *in vitro* assay of Ab10 using iodide as the halogen source. Mass spectra were measured in positive mode. For comparison the (a) standard of the substrate 4-HBA-SNAC (**16**) was used. The assay was performed (b) without Ab10, and as a (c) complete reaction. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Red chromatograms show the extracted ion chromatogram (EIC) of the $[M+Na]^+$ peak of **16** (theoretical m/z = 262.0508). Blue chromatograms show the EIC of the $[M+H]^+$, $[M+Na]^+$, and $[M+K]^+$ peaks (theoretical m/z = 365.9655, 387.9457, 403.9214 respectively) of the expected product 3-I-4-HBA-SNAC (**20**). As apparent from the images, no peak of the expected product **20** could be observed.

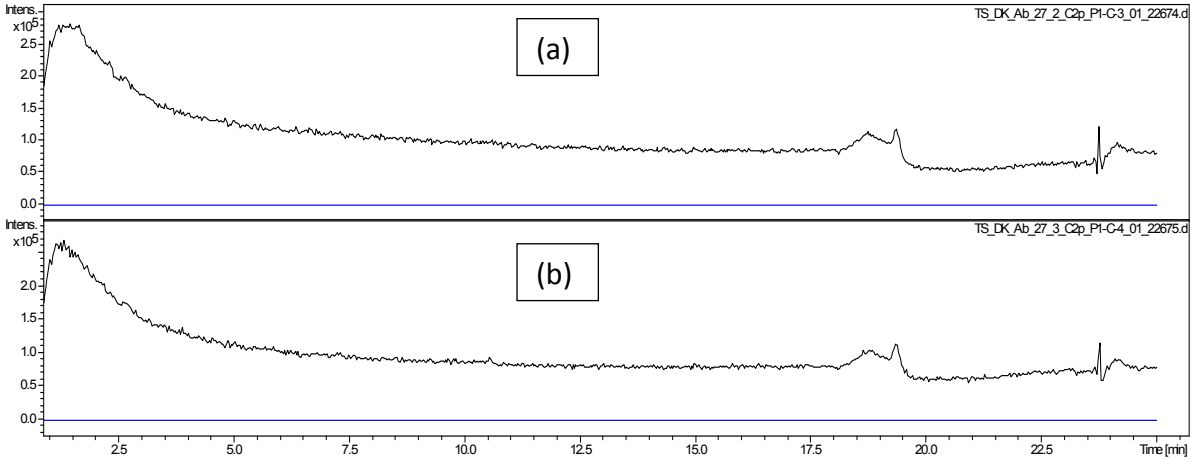


Figure S16. LCMS analysis of the *in vitro* assay of the C domain from Ab9 using 4-HBA-SNAC (**16**) as the substrate. The assay was performed using (a) only Ab9 (C-PCP) without substrate **16**, and as (b) complete reaction. Mass spectra were measured in positive mode. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of the product "4-HBA-cystamine" (**15**). As apparent from the images, product **15** could not be detected, neither in the control sample (a) nor in the complete reaction (b).

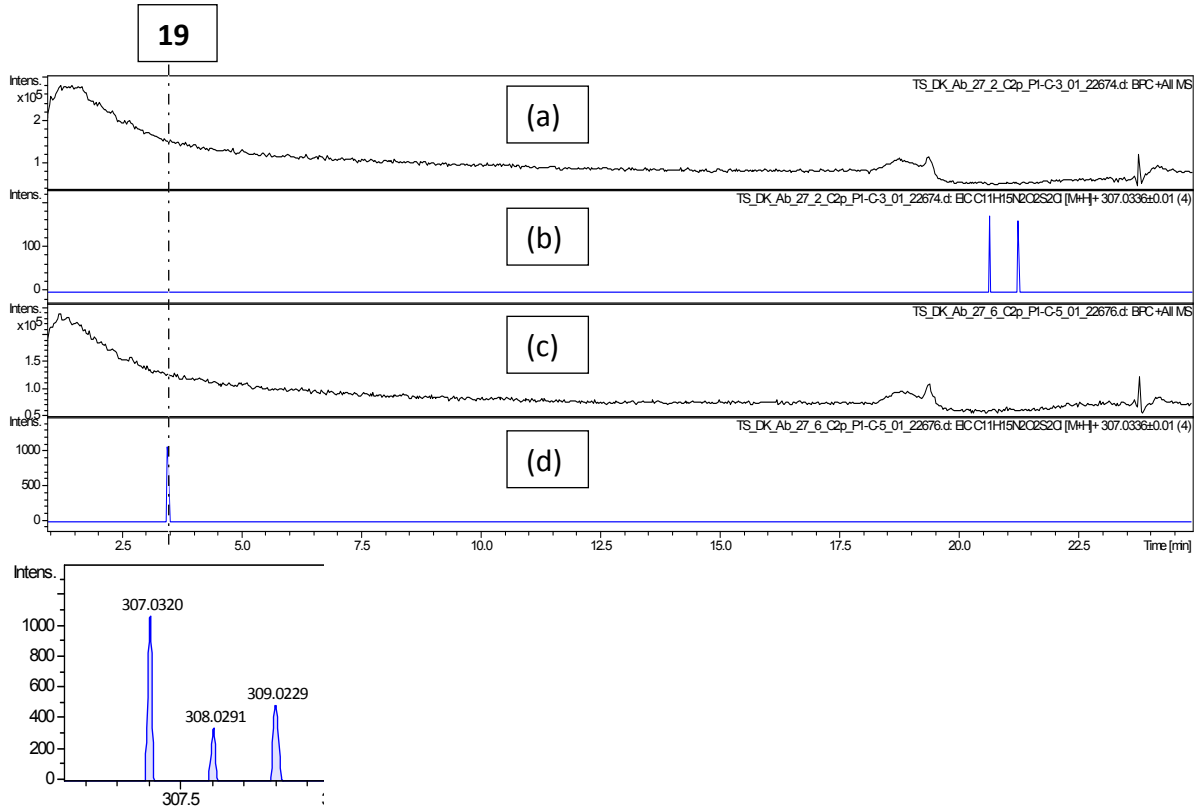


Figure S17. LCMS analysis of the *in vitro* assay of the C domain from Ab9 using 3-Cl-4-HBA-SNAC (**17**) as the substrate. The assay was performed using (a, b) only Ab9 (C-PCP) without substrate **17**, and as a (c, d) complete reaction. Mass spectra were measured in positive mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of the product "3-Cl-4-HBA-cystamine" **19** (theoretical m/z = 307.0336). As apparent from the images, product **19** could only be detected in the complete reaction sample.

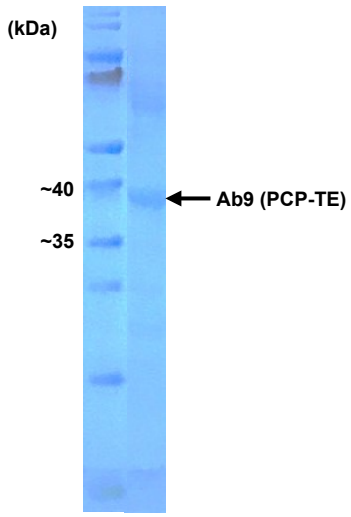


Figure S18. Purification of Ab9 (PCP-TE) as analyzed by SDS-Page.

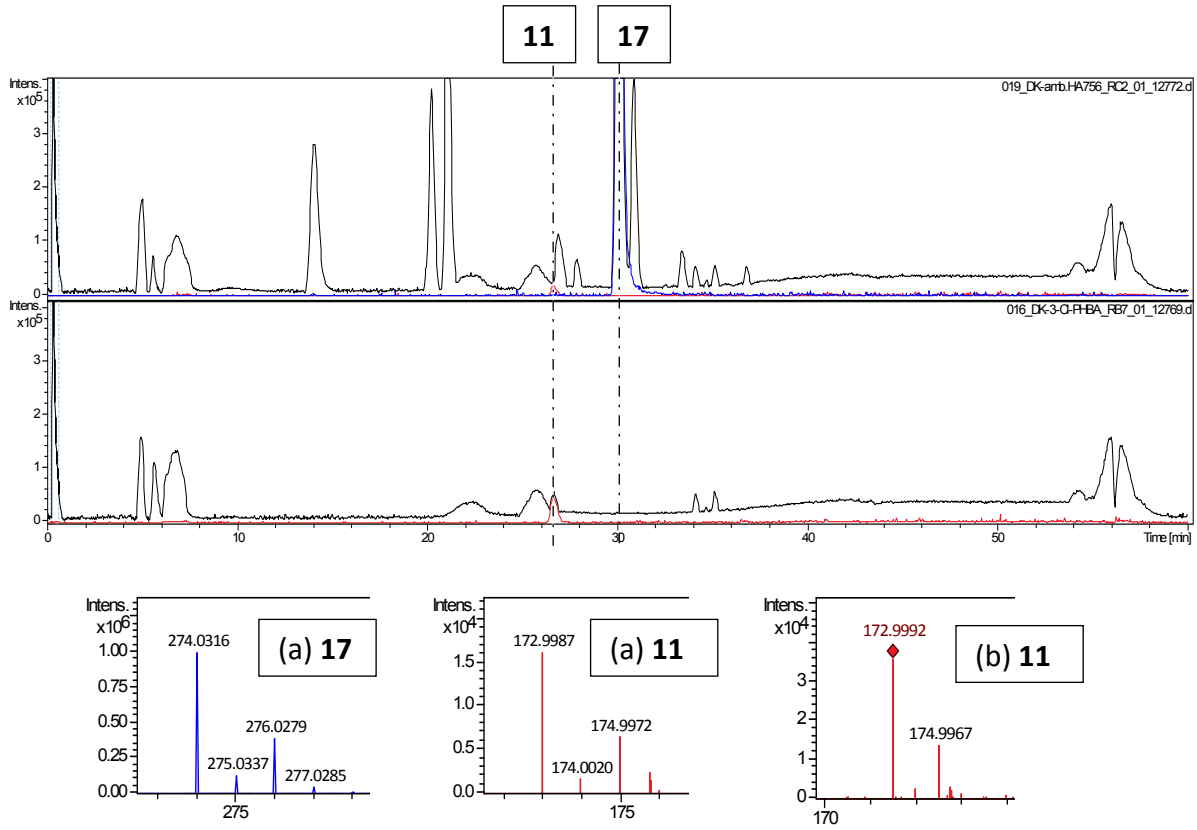
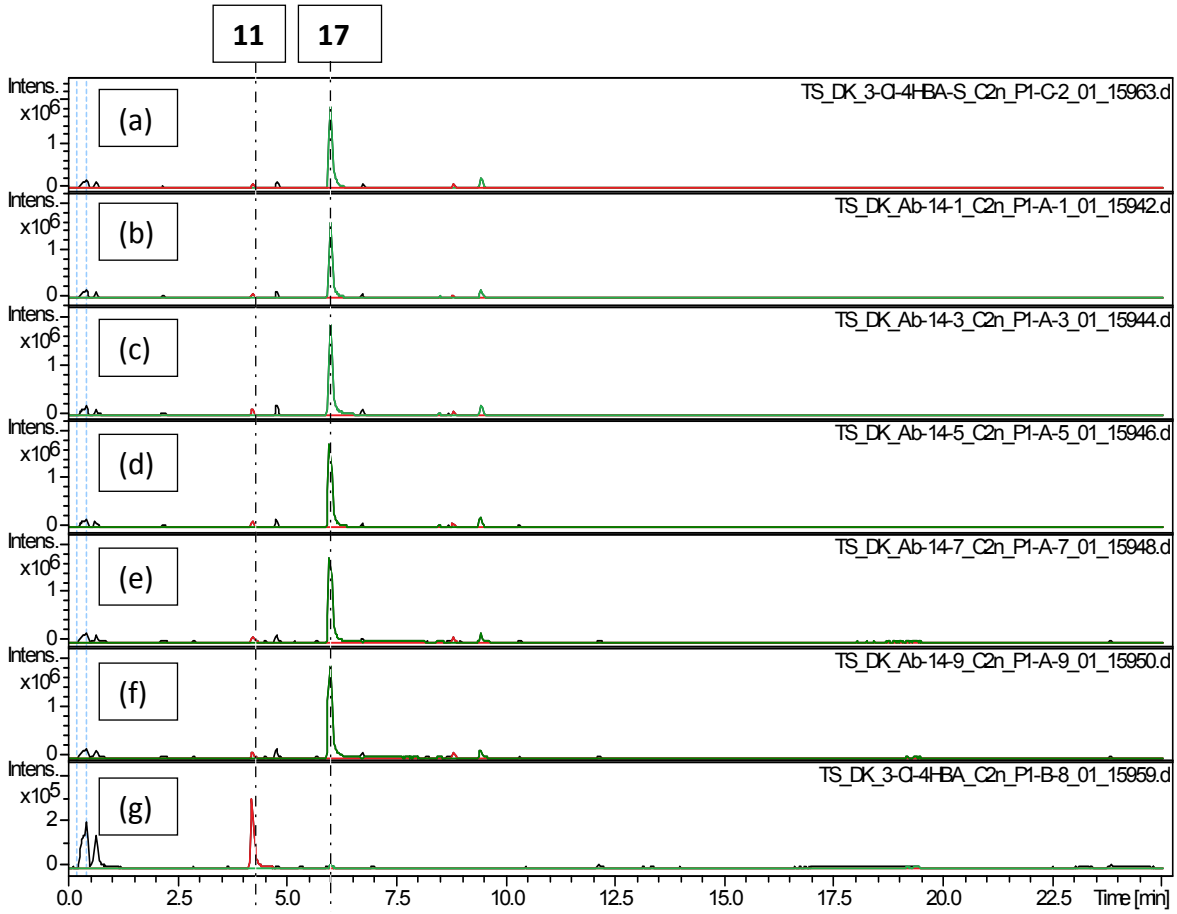
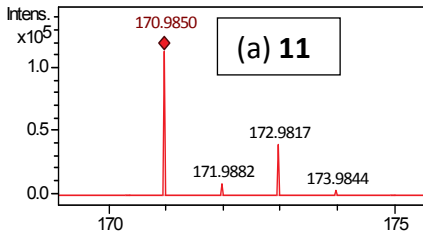
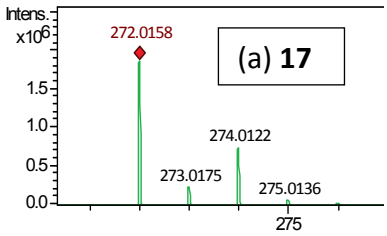
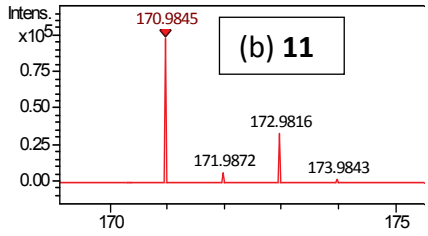


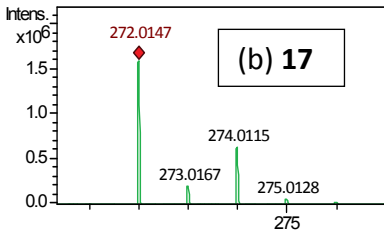
Figure S19. LCMS analysis of the (a) standard 3-Cl-4-HBA-SNAC (**17**) in comparison to the (b) standard of 3-Cl-4-HBA (**11**). Mass spectra were measured in positive mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Red chromatograms show the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of **11** (theoretical m/z = 173.0000). Blue chromatograms show the EIC of the $[M+H]^+$ peak of **17** (theoretical m/z = 274.0299). As apparent from the images, there is a contamination of **11** in the standard of **17**.

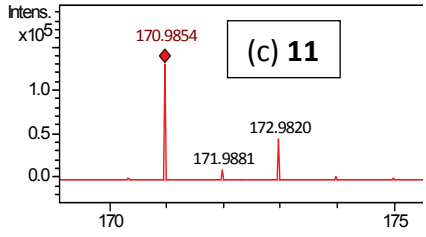

11
17
(a)
TS_DK_3-Cl-4HBA-S_C2n_P1-C-2_01_15963.d
(b)
TS_DK_Ab-14-1_C2n_P1-A-1_01_15942.d
(c)
TS_DK_Ab-14-3_C2n_P1-A-3_01_15944.d
(d)
TS_DK_Ab-14-5_C2n_P1-A-5_01_15946.d
(e)
TS_DK_Ab-14-7_C2n_P1-A-7_01_15948.d
(f)
TS_DK_Ab-14-9_C2n_P1-A-9_01_15950.d
(g)
TS_DK_3-Cl-4HBA_C2n_P1-B-8_01_15959.d
Intens.
Time [min]

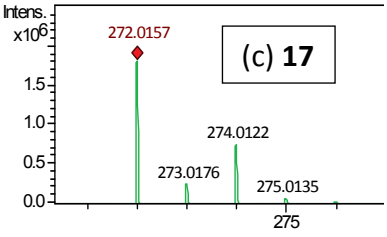

(a) 11
170.9850
172.9817
171.9882
173.9844


(a) 17
272.0158
274.0122
273.0175
275.0136


(b) 11
170.9845
172.9816
171.9872
173.9843


(b) 17
272.0147
274.0115
273.0167
275.0128


(c) 11
170.9854
172.9820
171.9881


(c) 17
272.0157
274.0122
273.0176
275.0135

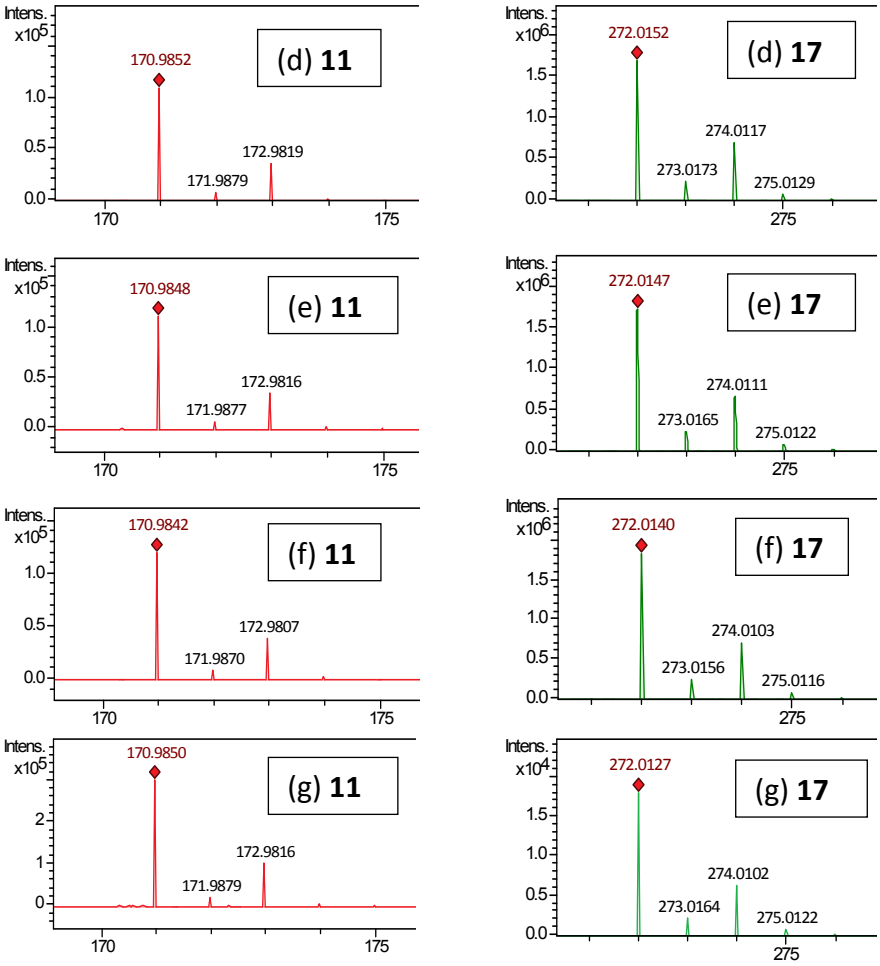
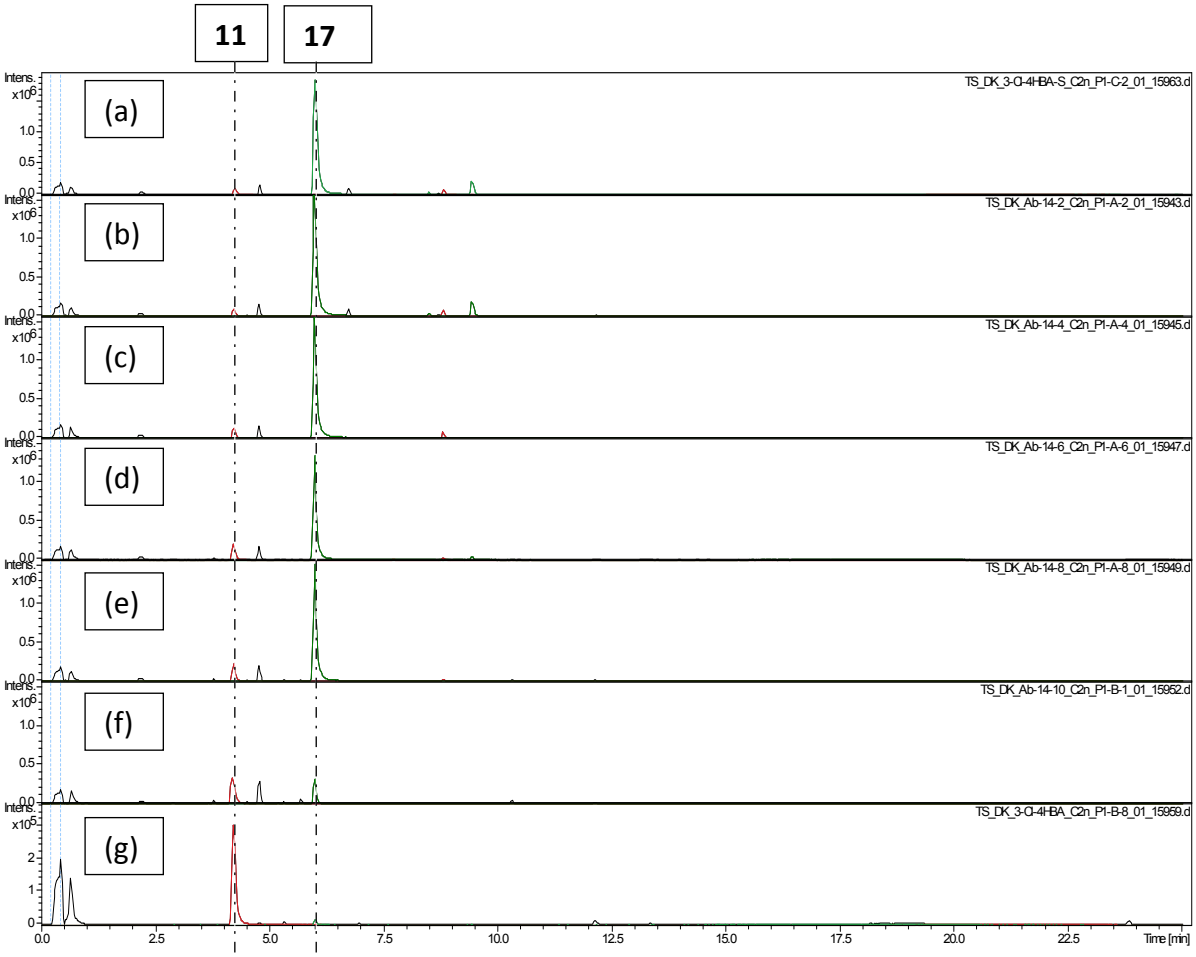
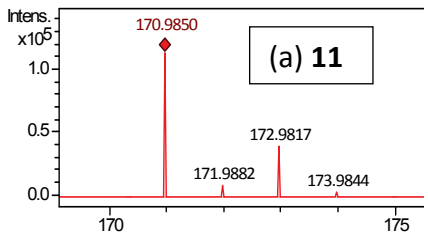
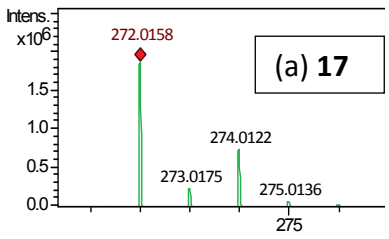
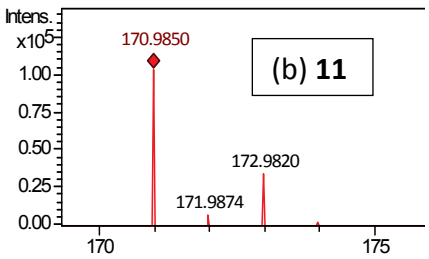


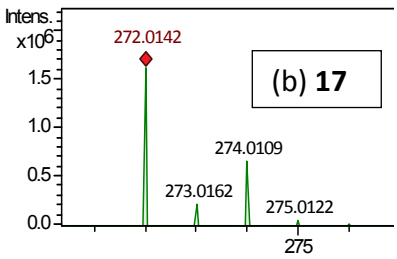
Figure S20. LCMS analysis of the control reaction (without Ab9 (PCP-TE)) of the time dependent Ab9 (PCP-TE) *in vitro* assay. For comparison a (a) standard of substrate 3-Cl-4-HBA-SNAC (**17**) and a (g) standard of the product 3-Cl-4-HBA (**11**) were used. The reactions (b), (c), (d), (e), (f) were incubated for 0, 5, 30, 60 min and 24 hours respectively. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Green chromatograms show the extracted ion chromatogram (EIC) of the $[M-H]^-$ peak of the substrate **17** (m/z theoretical = 272.0148). Red chromatograms show the EIC of the $[M-H]^-$ peak of the product **11** (m/z theoretical = 170.9848).

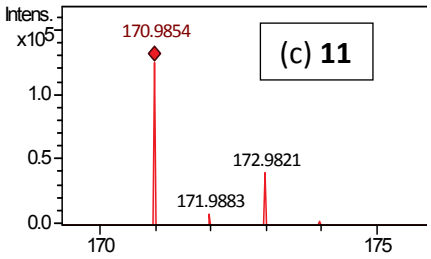

11
17
(a)
(b)
(c)
(d)
(e)
(f)
(g)
TS_DK_3-Cl-4HBA-S_C2n_P1-C-2_01_15963.d
TS_DK_Ab-14-2_C2n_P1-A-2_01_15943.d
TS_DK_Ab-14-4_C2n_P1-A-4_01_15945.d
TS_DK_Ab-14-6_C2n_P1-A-6_01_15947.d
TS_DK_Ab-14-8_C2n_P1-A-8_01_15949.d
TS_DK_Ab-14-10_C2n_P1-B-1_01_15952.d
TS_DK_3-Cl-4HBA_C2n_P1-B-8_01_15959.d
Intens.
Time [min]

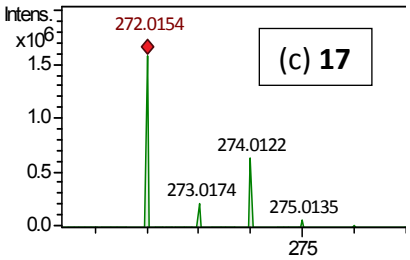

(a) 11
170.9850
172.9817
171.9882
173.9844


(a) 17
272.0158
274.0122
273.0175
275.0136


(b) 11
170.9850
172.9820
171.9874


(b) 17
272.0142
274.0109
273.0162
275.0122


(c) 11
170.9854
172.9821
171.9883


(c) 17
272.0154
274.0122
273.0174
275.0135

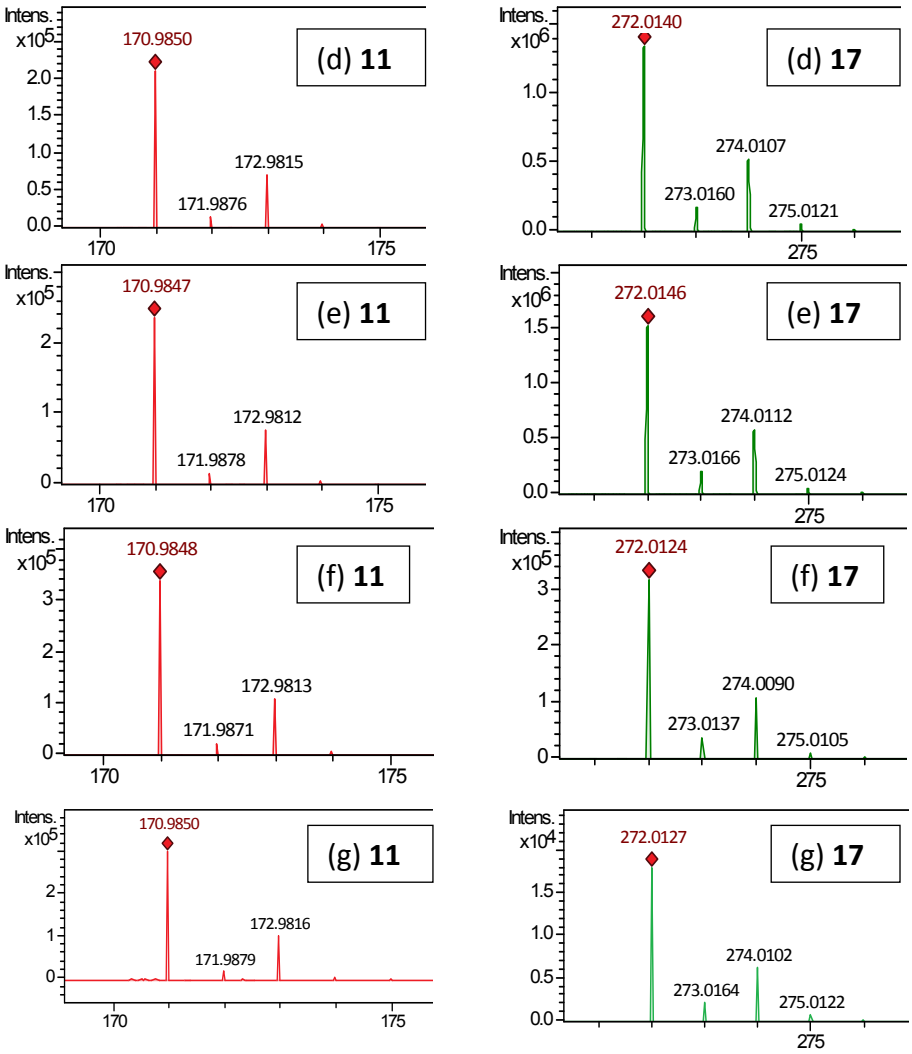


Figure S21. LCMS analysis of the complete Ab9 (PCP-TE) treated reaction of the time dependent Ab9 (PCP-TE) *in vitro* assay. For comparison a (a) standard of substrate 3-Cl-4-HBA-SNAC (**17**) and a (g) standard of the product 3-Cl-4-HBA (**11**) were used. The reactions (b), (c), (d), (e), (f) were incubated for 0, 5, 30, 60 min and 24 hours respectively. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Green chromatograms show the extracted ion chromatogram (EIC) of the $[M-H]^-$ peak of the substrate **17** (m/z theoretical = 272.0148). Red chromatograms show the EIC of the $[M-H]^-$ peak of the product **11** (m/z theoretical = 170.9848).

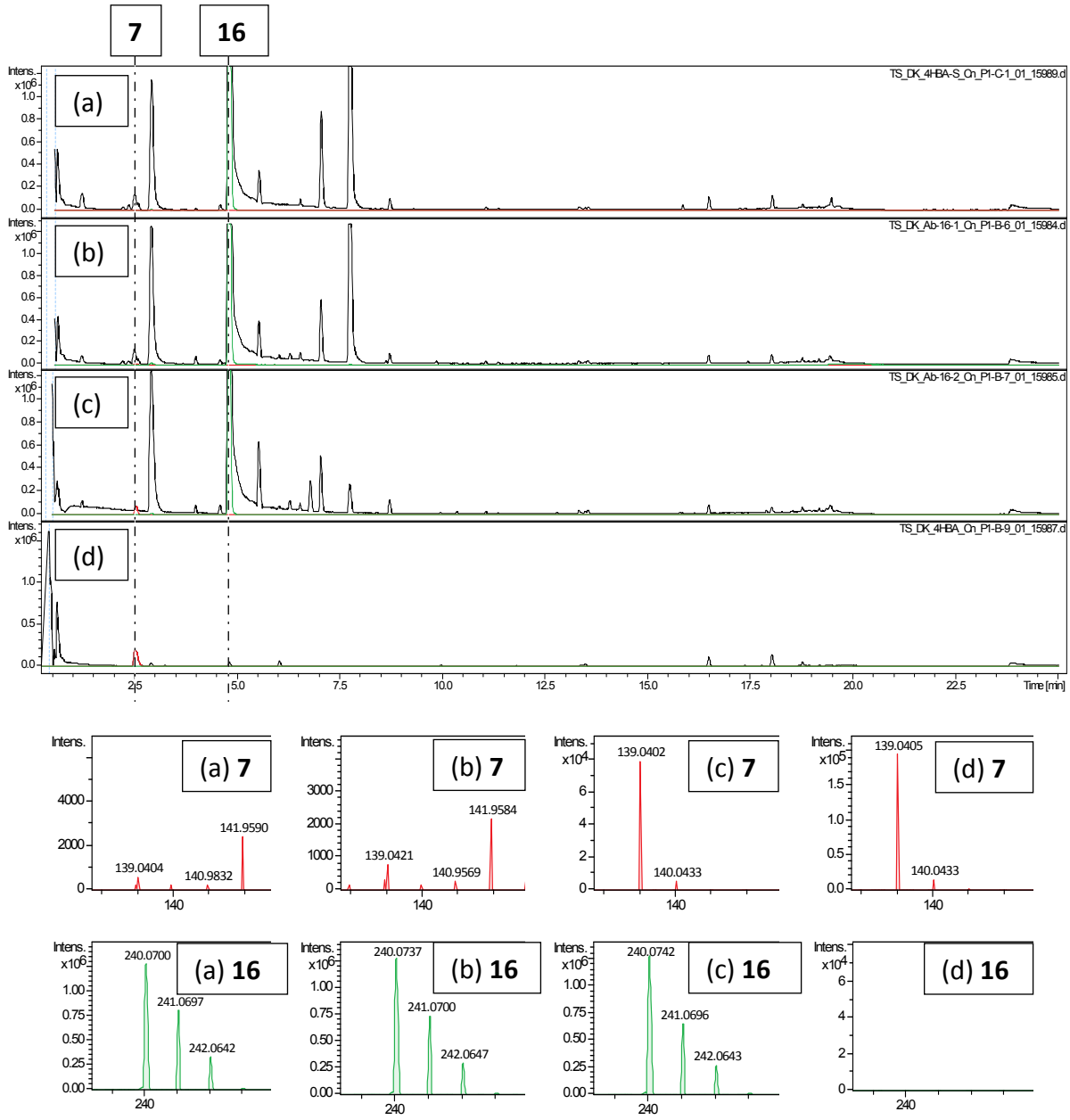


Figure S22. LCMS analysis of Ab9 (PCP-TE) substrate specificity experiment with respect to the substrate 4-HBA-SNAC (**16**). For comparison a (a) standard of the substrate **16**, and a (d) standard of the product 4-HBA (**7**) were used. The assay was carried out (b) without Ab9 (PCP-TE) (control), and as a (c) complete reaction. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Green chromatograms show the extracted ion chromatogram (EIC) of the $[M+H]^+$ peak of the substrate **16** (m/z theoretical = 240.0694). Red chromatograms show the EIC of the $[M+H]^+$ peak of the product **7** (m/z theoretical = 139.0395).

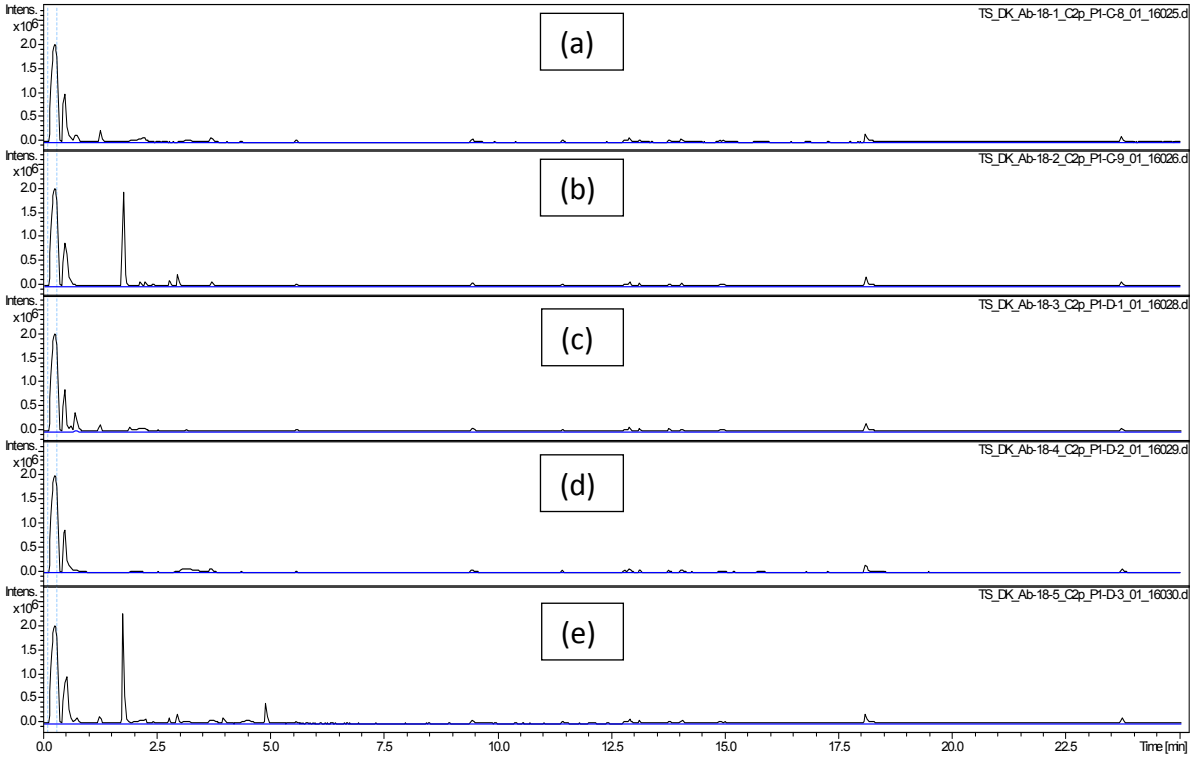


Figure S23. LCMS data of the *in vitro* investigation of Ab6 substrate specificity with respect to the substrate 3-Cl-4-HBA (**11**). The assay was carried out (a) without ATP, (b) without CoA, (c) without substrate **11**, (d) without Ab6, and as (e) complete reaction. Mass spectra were measured in positive mode. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M+H]^+$ and $[M+Na]^+$ peaks of 3-Cl-4-HBA-CoA (theoretical m/z = 922.1046 and 944.0866 respectively). As apparent from the EIC, the expected product could not be detected.

```
Ab1  MSNLPKSTKVLVVGGGPAGTTAATLLAREGFDITLLEREVFPRYHIGESLLPSSLK-VLD 59
CmlS ----MTRSKVAIIGGGPAGSVAGLTLHKLGHDVTIYERSAFPRYRVGESLLPGTMS-ILN 55
CndH MSTRPEVFDLIVIGGGPGGSTLASFVAMRGHRVLLLEREAFPRHQIGESLLPATVHGICA 60
             .: ::****.*:. .  :   *. : : **..***:::******.::  :

Ab1  LLGVRDKIDAHGFQYKPGGHYHWGDEHWDLNFSDLS-----GNITHSYQVRRDEFDKLLL 114
CmlS RLGLQEKIDAQNYVKKPSATFLWGQDQAPWTFSFAAPKVAPWVFDHAVQVKREEFDKLLL 115
CndH MLGLTDEMKRAGFPIKRGGTFRWGKEPEPWTFGFTRHPDDPY--GFAYQVERARFDDMLL 118
      **: :::.  .:  * .. : **.:    .*.            .: **.* .**.:**

Ab1  DHAKSQGVKVFDGIGVSSLSFENERPKSAIWSQTNDKNHTGEISFDFLIDATGRYGLMAN 174
CmlS DEARSRGITVHEETPVTDVDLSDPD--RVVLT-VRRGGESVTVESDFVIDAGGSGGPI-S 171
CndH RNSERKGVDVRERHEVIDVLFEGERAVGVRYR--NTEGVELMAHARFIVDASGNRTRV-S 175
      .:. :*: * :   * .: :..     .     .  .        *::** *    : .

Ab1  HHLKNREYHDVFQNVAIWGYWKNADRLDNGREGAIIIESLKDGWLWGIPLHDGTISVGLV 234
CmlS RKLGVRQYDEFYRNFAVWSYFKLKDPFEGDLKGTTYSITFEDGWVWMIPIKDDLYSVGLV 231
CndH QAVGERVYSRFFQNVALYGYFENGKRLPAPRQGNILSAAFQDGWFWYIPLSDTLTSVGAV 235
     : :  * *  .::*.*::.*::  . :    :*     :::***.* **: *   *** *

Ab1  VHKTIYKEKRSKSLKDIYLEGIAESLDLKRLLEPGEL-----ASEVRSEQDYSYAADSFA 289
CmlS VDRSKSAEVREQGADAFYSSTLAKCAKAMDILGGAEQ-----VDEVRIVQDWSYDTEVFS 286
CndH VSREAAEAI-KDGHEAALLRYIDRCPIIKEYLAPATRVTTGDYGEIRIRKDYSYCNTSFW 294
     * :       ... .      : ..      *  .        .*:*  :*:**    *

Ab1  GQGYFMIGDAACFLDPLLSTGVHLATFSGLLSAASLASVIRNHITEEQAISFFERTYKQA 349
CmlS ADRFFLCGDAACFTDPLFSQGVHLASQSAVSAAAAIDRITRHGDEKDAVHAWYNRTYREA 346
CndH KNGMALVGDAACFVDPVFSSGVHLATYSALLVARAINTCLAGEMSEQRCFEEFERRYRRE 354
      :   : ****** **::* *****: *.:  * ::         ::     ::* *:.

Ab1  YLRLMAMVSAFYEN---SKKESYFWQAQQLTKTRQSNEDKEKLHQMFLNVVSG--MEDMS 404
CmlS YEQYHQFLASFYTFASFTEPDSEFWRKRRITESDDDRLTRKKW----FESLAGNGPEDPS 402
CndH YGNFYQFLVAFYDMN--QDTDSYFWSARKIINTEERAN------EAFVRLIAGRSNLDEP 406
     * .   :: :**      . :* **  ::: :: :            .. ::*    *

Ab1  DAEENSEELFLEL------SERLRENWS--------LRHKQT----ANDL---------- 436
CmlS GTV-------ASF------RDRASTMIA--------IGRHQRP-ELSDDF---------- 430
CndH VFQSVAKDFFTEREGFGAWFGGLVTSMAKGDGGGLMVGEGATDATESTGFAPENFMQGFT 466
                .               :        : .       : .:

Ab1  -----------DQTE--EEKLRASNQFVSRLNGLFSLSK---ESAVEGLYIVTTPQLGLV 480
CmlS -----------SEAELNPARVRWISDLTKRLNSITRFKWTGGKAVLKQHYRVEPIGFRLE 479
CndH REITELQHLAMFGEDRGPETPLWSGGLVPSRDGLAWAVESGEDAAG-------------- 512
                   :         . :.   :.:        .:.

Ab1  QVN---------------------------------------------------------- 483
CmlS QREVLANGEGLDMAQYPMDDEARQIFQDLAEEEFGYKTLVKRLGAVGRQELSTQIVVRLM 539
CndH ------------------------------------------------------------


Ab1  ---------------------------------
CmlS EAGLLTGYDAQGEKVFVQGRLHFGGVGVEYEV 571
CndH ---------------------------------
```

Figure S24. Alignment of Ab1 against ClmS and CndH with GxGxxG and WxWxIP motifs being highlighted in bold red letters. BLASTp alignment of Ab1 against the database shows, that almost all of the 100 first hits are tryptophan 7-halogenases or NAD(P)/FAD-dependent oxidoreductases with the highest percentage of identity being 72.52%. However, none of those hits is so far experimentally characterized. A refined BLASTp search using the pdb database and including only experimentally characterized enzymes, revealed the first two hits to be ClmS and CndH. However, as can be concluded from the query cover, both of them do not share AA sequence in their C-terminal region. Differences in the C-terminal region of Flavin-dependent halogenases (FDH) are the origin of their regio- and substrate specificity, as described by the characterization of PyrH[1]. The N-terminus usually displays the region which forms the FAD binding site[2].

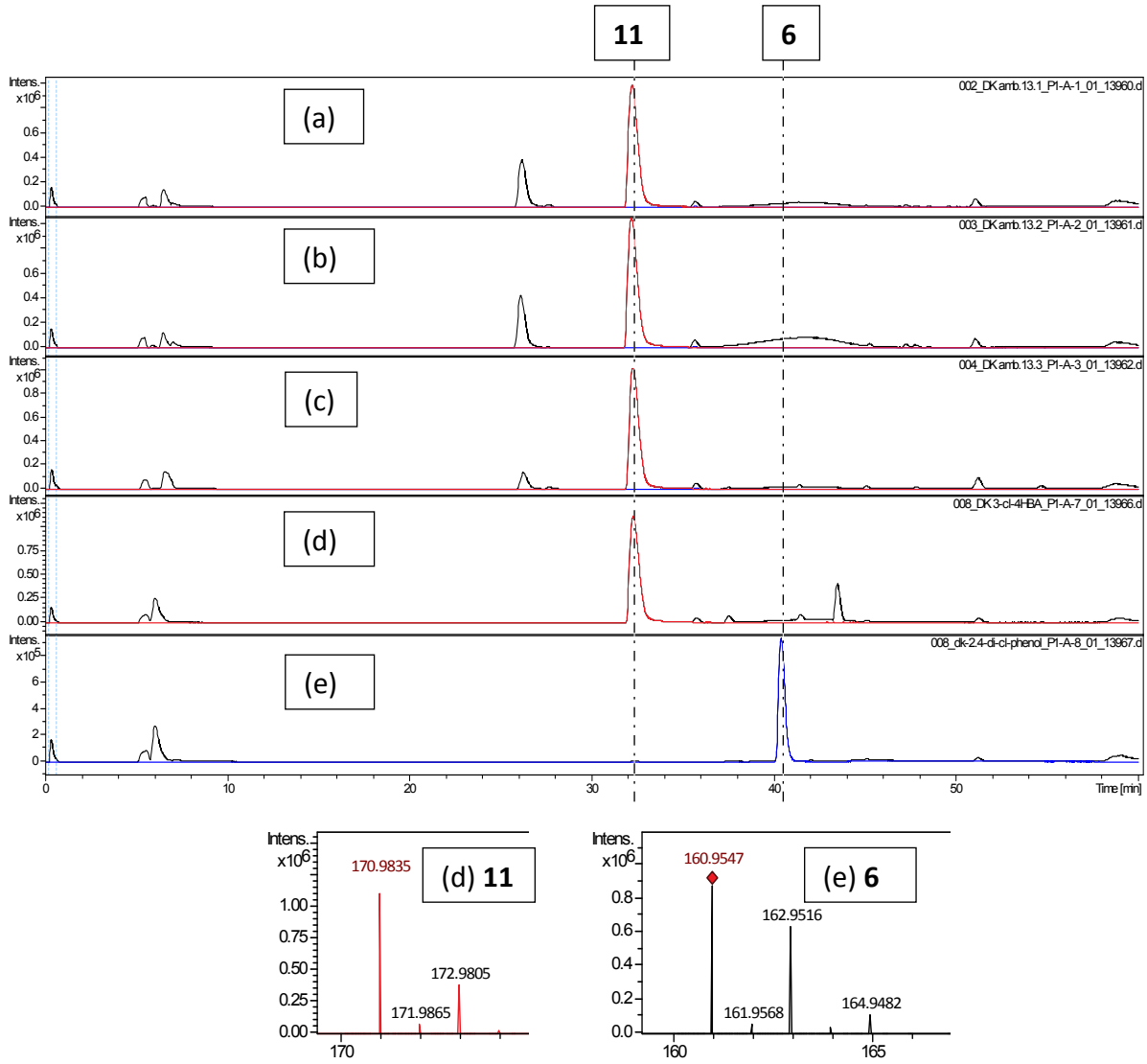


Figure S25. LCMS analysis of the *in vitro* assay using 3-Cl-4-HBA (**11**) as the substrate of Ab1. The assay was carried out (a) without Fre, as (b) complete reaction, (c) without Ab1. The (d) standard of substrate **11**, and the (e) standard of the product 2,4-dichlorophenol (**6**) served for comparison. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Red chromatograms show the extracted ion chromatogram (EIC) of the $[M-H]^-$ peak of **11** (theoretical m/z = 170.9843). Blue chromatograms show the EIC of the $[M-H]^-$ peak of **6** (theoretical m/z = 160.9555). As apparent from the images, the expected product **6** could not be observed.

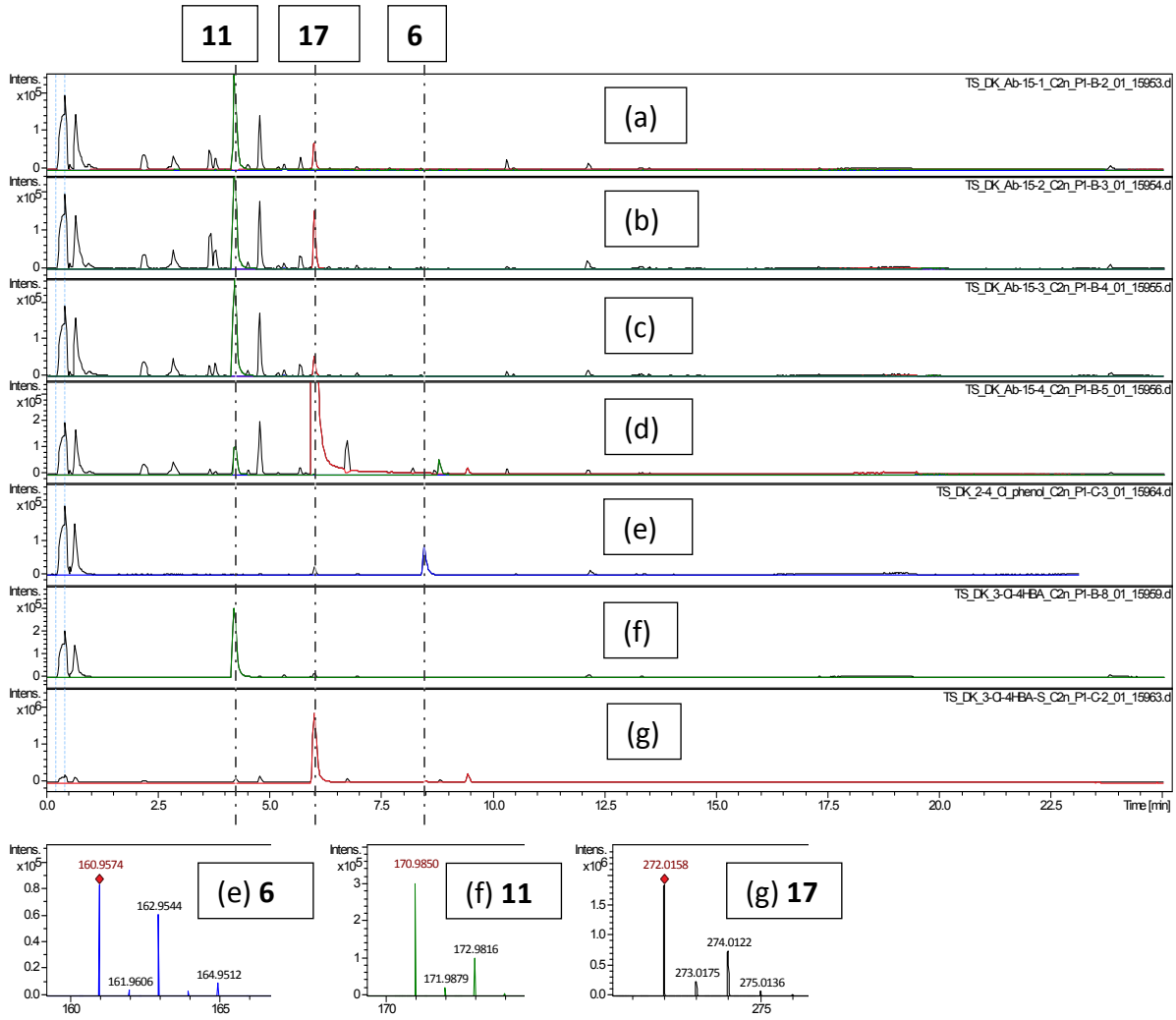


Figure S26. LCMS analysis of the *in vitro* assay using 3-Cl-4-HBA-SNAC (**17**) as the substrate of Ab1 in the presence of Ab9 (PCP-TE). The assay was carried out (a) without Ab1, as (b) complete reaction (with Ab1 and Ab9 (PCP-TE)), (c) without Fre, and (d) without Ab9 (PCP-TE). The (e) standard of 2,4-dichlorophenol (**6**), the (f) standard of 3-Cl-4-HBA (**11**), and the (g) standard of substrate **17** served for comparison. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M\text{-}H]^-$ peak of **6** (theoretical m/z = 160.9555). Green chromatograms show the EIC of the $[M\text{-}H]^-$ peak of **11** (theoretical m/z = 170.9843), while the red chromatograms display the EIC of the $[M\text{-}H]^-$ peak of **17** (theoretical m/z = 272.0143). The significant decrease of **17** and concomitant increase of **11** in the reactions (a)-(c) relative to (d) can be attributed to the activity of Ab9 (PCP-TE) which was not present in reaction (d). However, as apparent from the images, under various condition, the expected product **6** could not be observed.

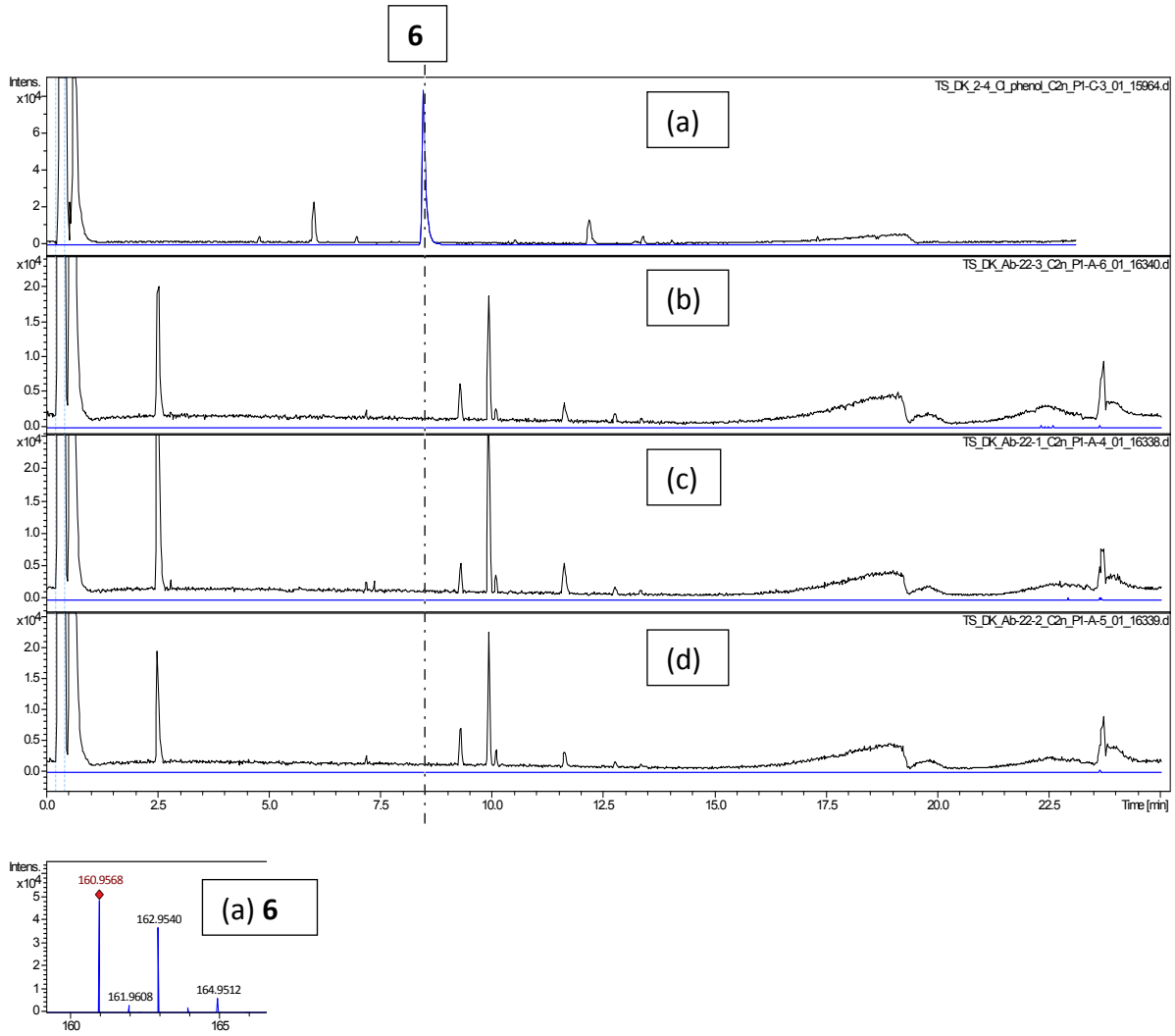
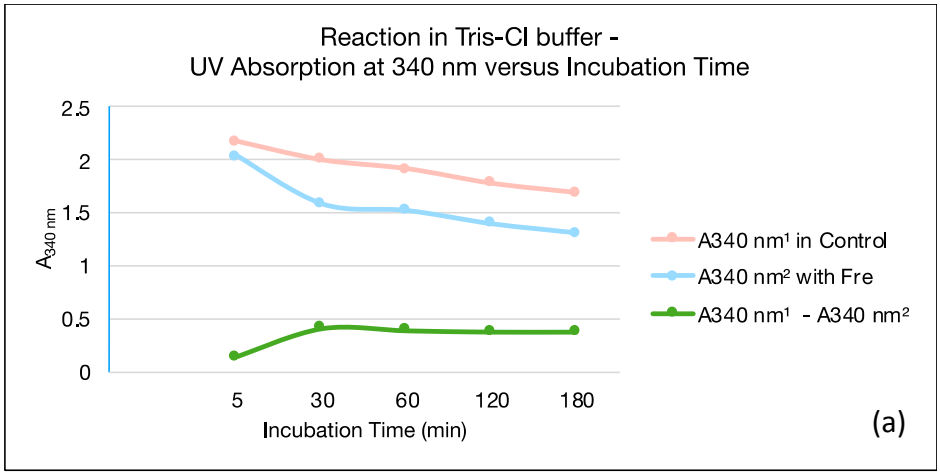
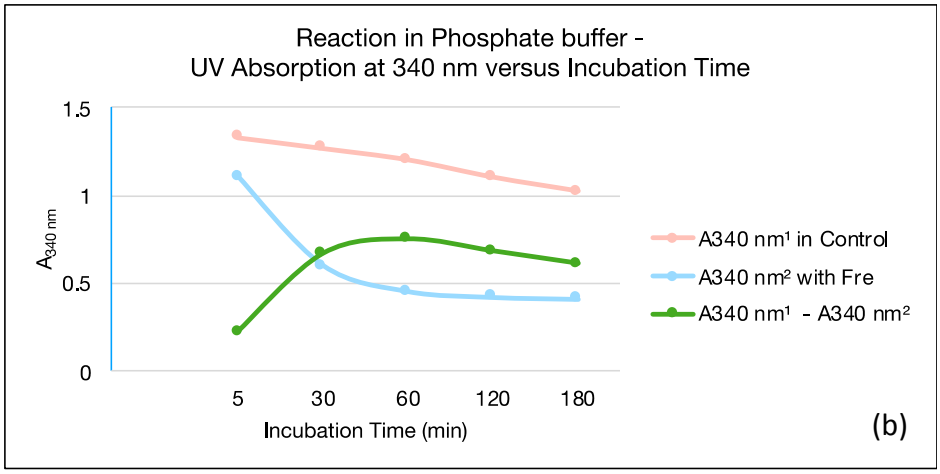


Figure S27. LCMS analysis of the *in vitro* assay using 2-chlorophenol as the substrate of Ab1. For comparison a (a) standard of the product 2,4-dichlorophenol (**6**) was employed. The assay was carried out as (b) complete reaction, (c) without Fre, and (d) without Ab1. Mass spectra were measured in negative mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M\text{-}H]^-$ peak of **6** (theoretical m/z = 160.9555). As apparent from the images, the expected product **6** could not be detected.





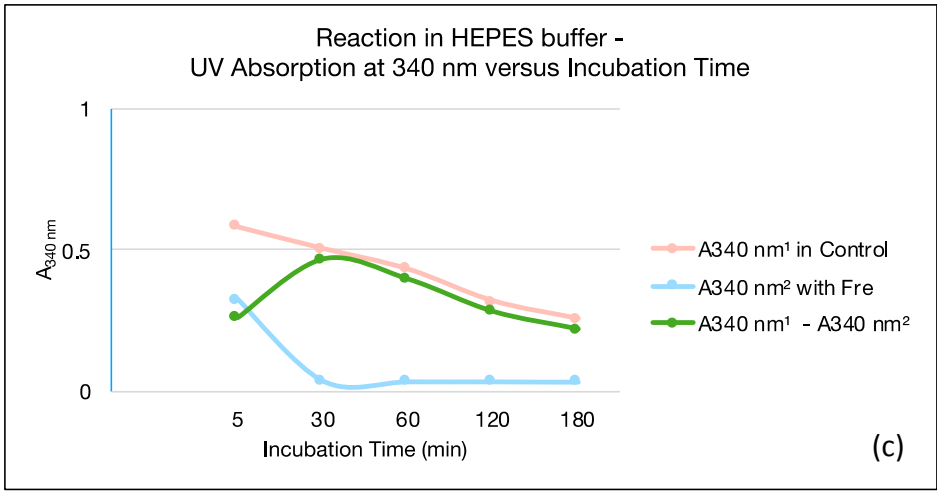


Figure S28. Graphic representation of Table S5 showing NADH consumption (green line) by Fre in (a) Tris-Cl, (b) Phosphate, and (c) HEPES buffer. Without Fre, the absorption ($A_{340\ nm}$) was decreasing over time, too. Therefore, it is likely that NADH naturally degraded over time. However, when Fre was present, the decrease of $A_{340\ nm}$ was accelerated, indicating that NADH as substrate of Fre was consumed. According to our results, Fre is active in all buffers tested.

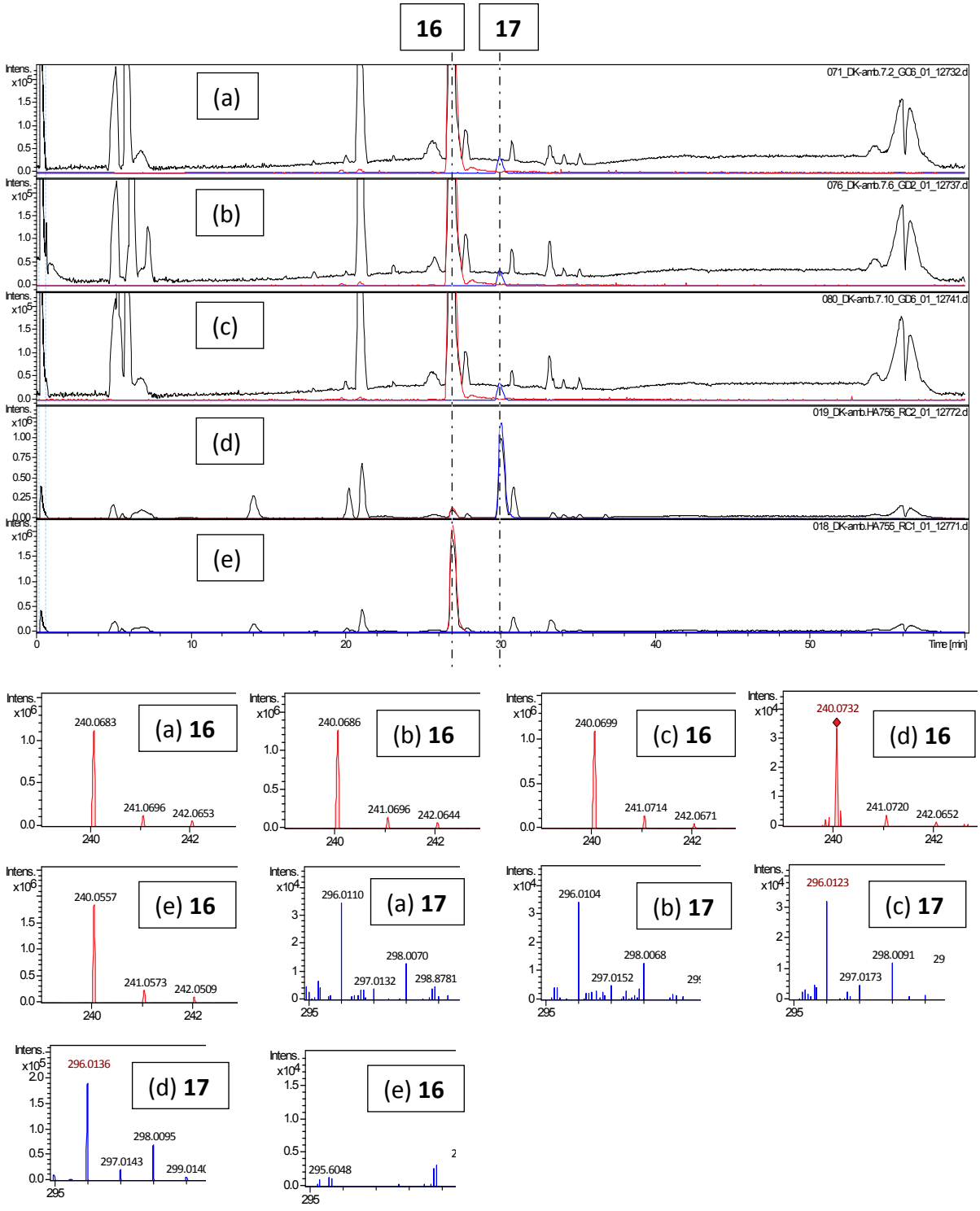


Figure S29. LCMS analysis of the *in vitro* assay of Ab10 in three different buffers. The utilized buffers were (a) Tris-Cl, (b) phosphate, and (c) HEPES, respectively. The (d) standard of the product 3-Cl-4-HBA-SNAC (**17**) and the (e) standard of the substrate 4-HBA-SNAC (**16**) served for comparison. Mass spectra were measured in positive mode. The upper part of the figure shows the respective chromatograms, while the lower part provides close-ups of the corresponding mass spectra. Base peak chromatograms (BPC) of each sample are shown in black. Blue chromatograms show the extracted ion chromatogram (EIC) of the $[M+Na]^+$ peak of **17** (theoretical m/z = 296.0119), while red chromatograms display the EIC of the $[M+H]^+$

peak of the substrate **16** (theoretical m/z = 240.0689). As shown in the images, the reaction product **17** was produced in comparable amounts in the three different buffers used (judging by the peak intensity), thus indicating that Ab10 works equally well in all three buffer systems.

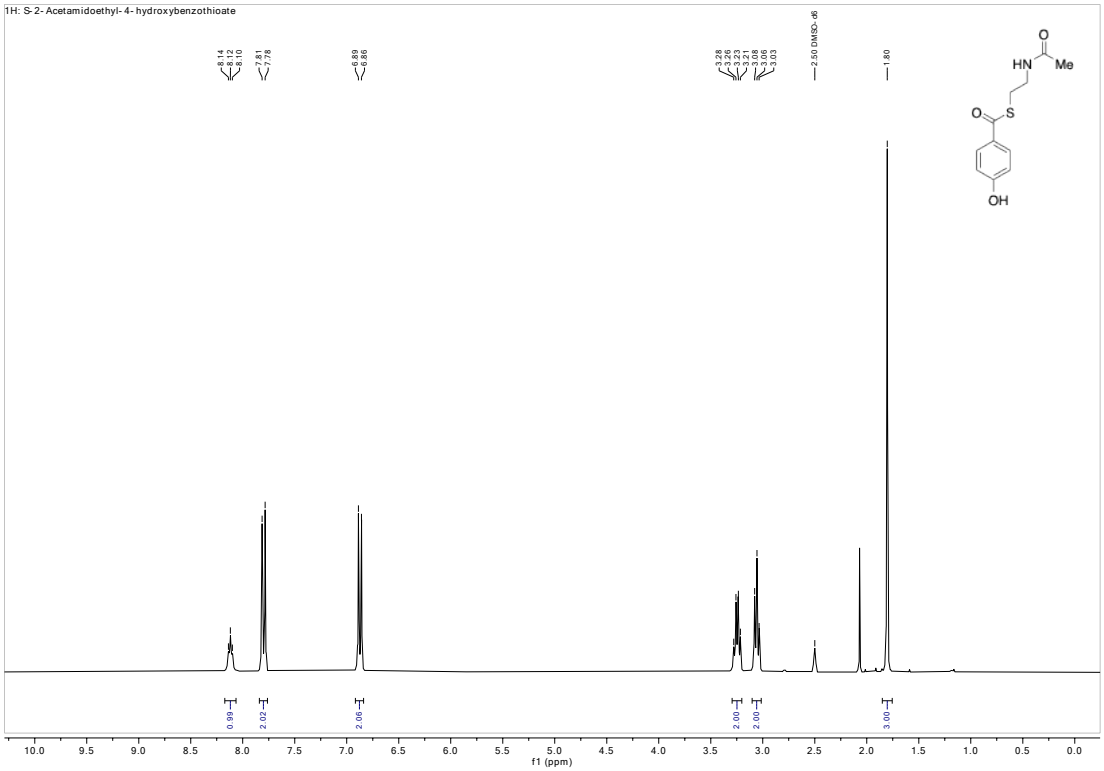


Figure S30. ^{1}H spectrum of *S*-2-acetamidoethyl 4-hydroxybenzothioate (**16**).

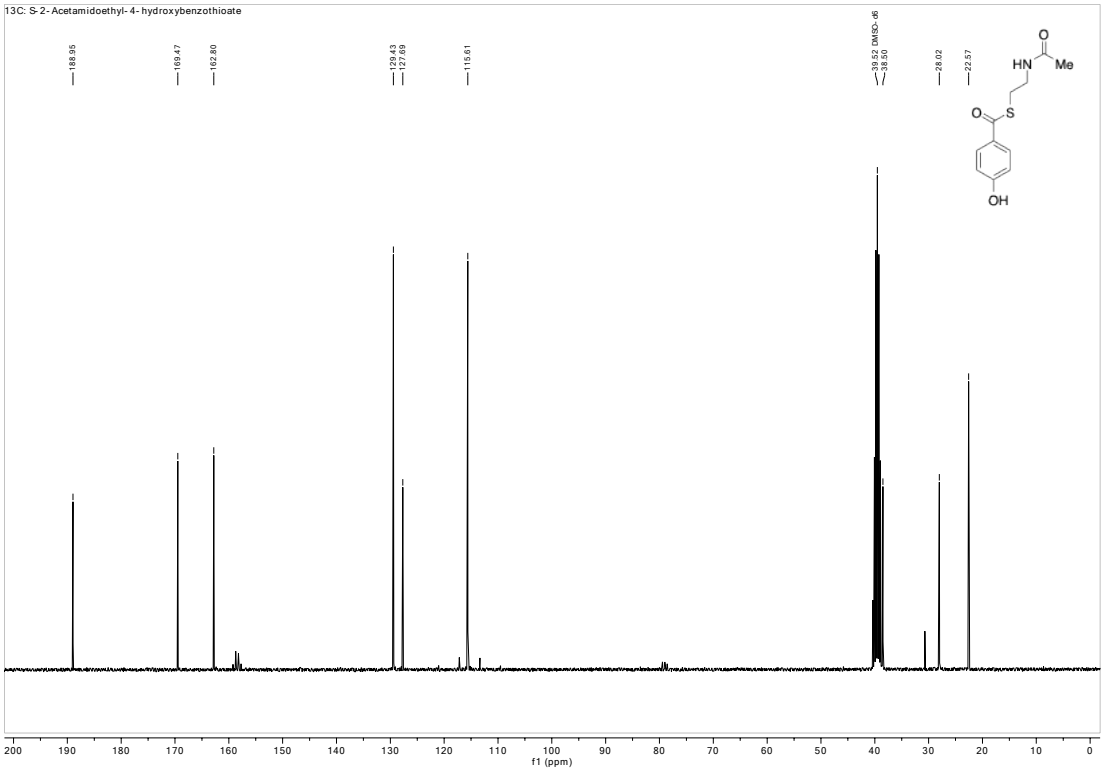


Figure S31. ^{13}C spectrum of *S*-2-acetamidoethyl 4-hydroxybenzothioate (**16**).

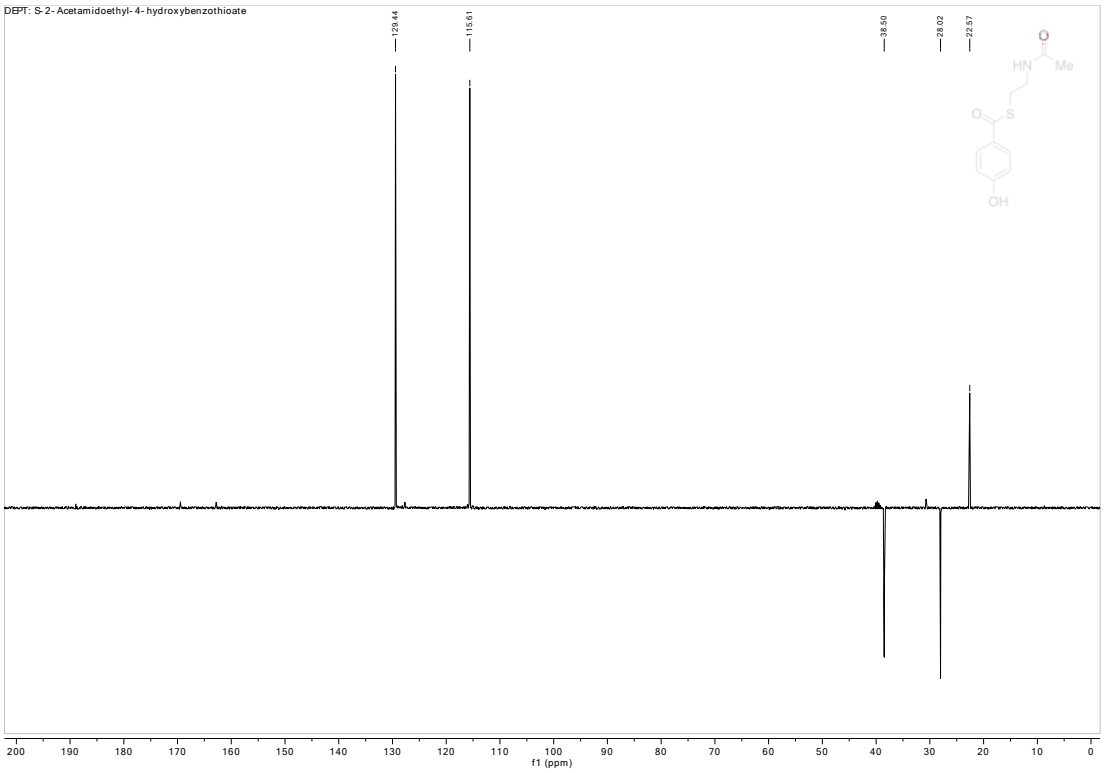


Figure S32. DEPT spectrum of *S*-2-acetamidoethyl 4-hydroxybenzothioate (**16**).

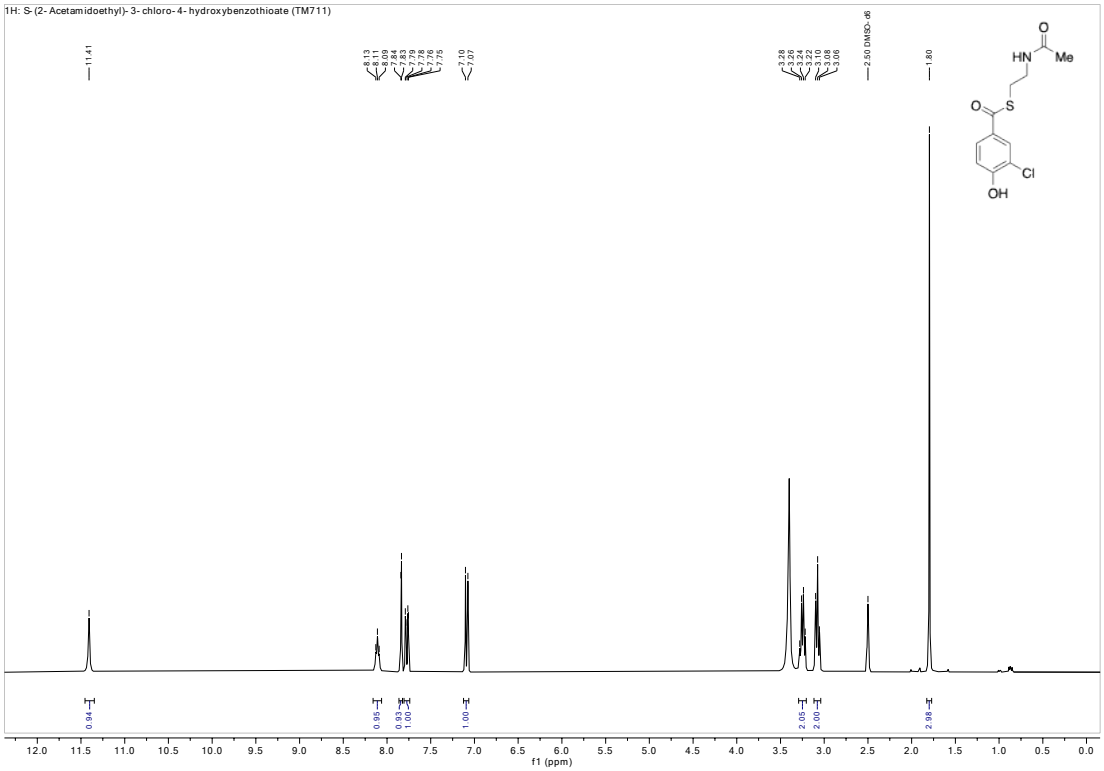


Figure S33. ^{1}H spectrum of *S*-2-acetamidoethyl 3-chloro-4-hydroxybenzothioate (**17**).

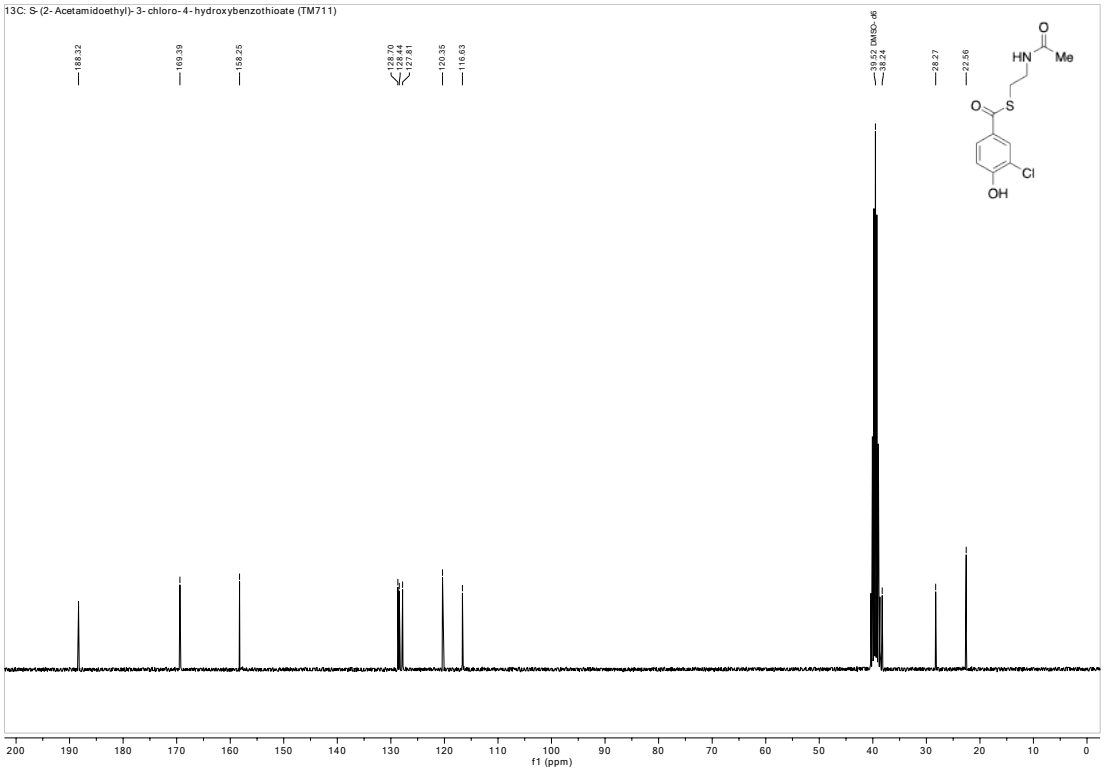


Figure S34. ^{13}C spectrum of *S*-2-acetamidoethyl 3-chloro-4-hydroxybenzothioate (**17**).

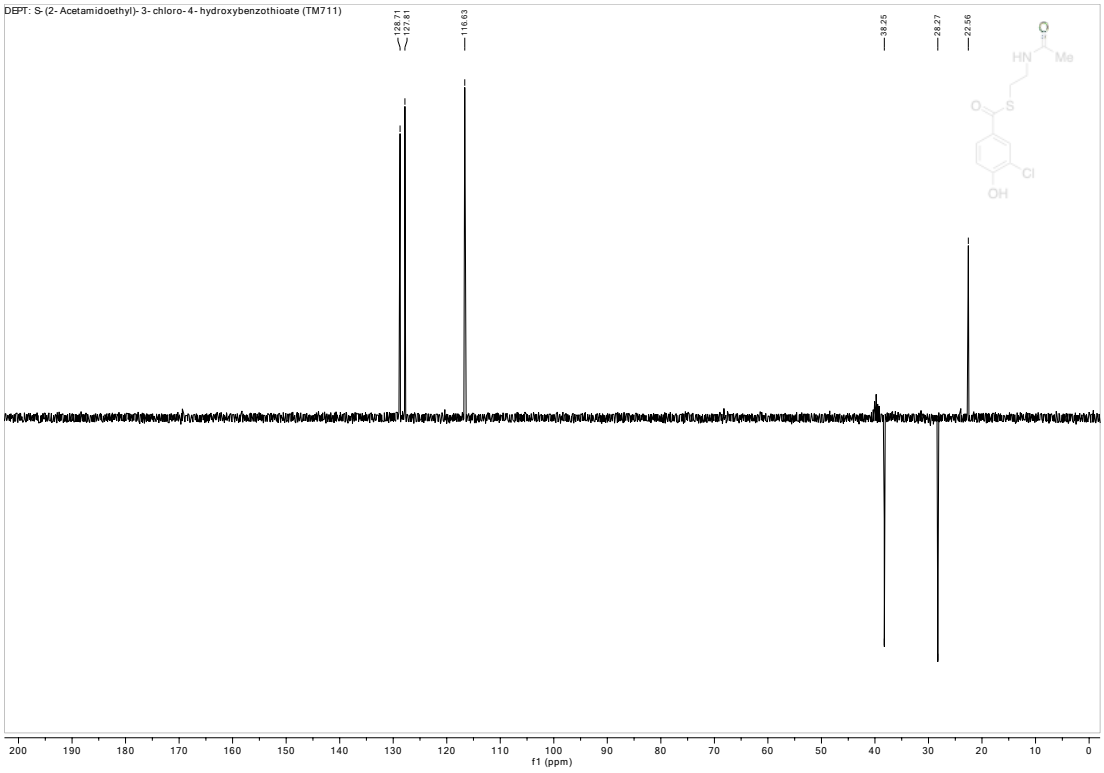


Figure S35. DEPT spectrum of *S*-2-acetamidoethyl 3-chloro-4-hydroxybenzothioate (**17**).

References

[1] X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pée, J. H. Naismith, Journal of Molecular Biology, 2009, 391, 74-85

[2] K. Podzelinska, R. Latimer, A. Bhattacharya, L. C. Vining, D. L. Zechel, Z. Jia, Journal of Molecular Biology, 2010, 397, 316-331.