

## Supplemental Information for

### Regioselective deacetylation based on teicoplanine-complexed Orf2\* crystal structures

*Hsiu-Chien Chan, Yu-Ting Huang, Syue-Yi Lyu, Chuen-Jiuan Huang, Yi-Shan Li,  
Yu-Chen Liu, Chia-Cheng Chou, Ming-Daw Tsai\* and Tsung-Lin Li\**

\*To whom correspondence should be sent. E-mail: [tlli@gate.sinica.edu.tw](mailto:tlli@gate.sinica.edu.tw);  
[mchtsai@gate.sinica.edu.tw](mailto:mchtsai@gate.sinica.edu.tw)

#### **This PDF file includes:**

##### Supplemental Data

Figures S1-21.....	pages S2-17
Figure S1.....	page S2
Figure S2.....	page S3
Figure S3.....	page S4
Figure S4.....	page S5-6
Figure S5.....	page S7
Figure S6.....	page S8
Figure S7.....	page S9-11
Figure S8.....	page S12-14
Figure S9.....	page S15-17
Figure S10.....	page S18
Table S1.....	page S19

##### Supplemental Experimental Procedures

CD analysis.....	page S20
NMR analysis.....	page S21-22

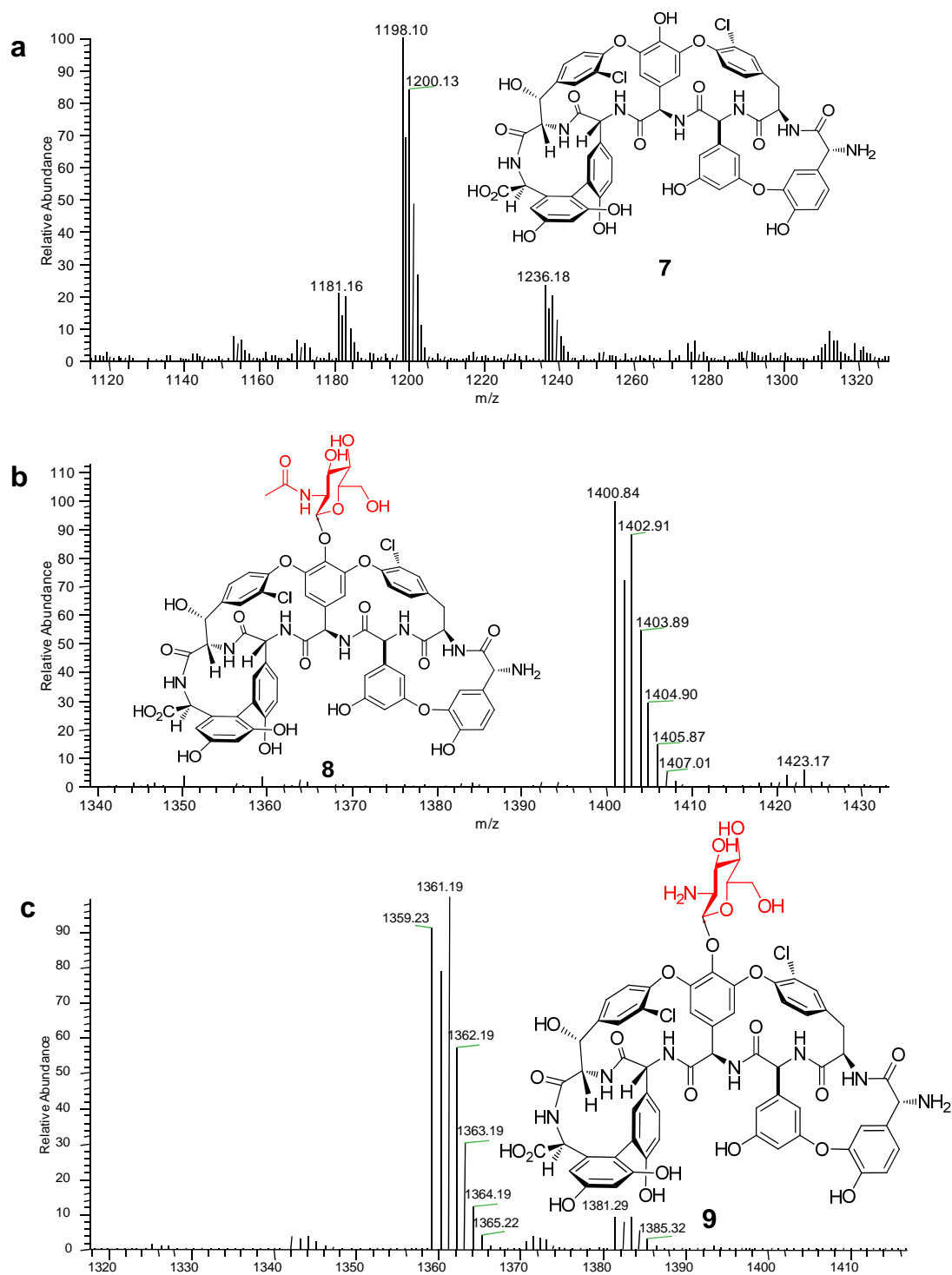


Figure S1. ESI/MS spectra for compounds shown in Fig. 2a. a) ESI/MS spectrum of aglycone (**7**). b) ESI/MS spectrum of compound (**8**). c) ESI/MS spectrum of compound (**9**).

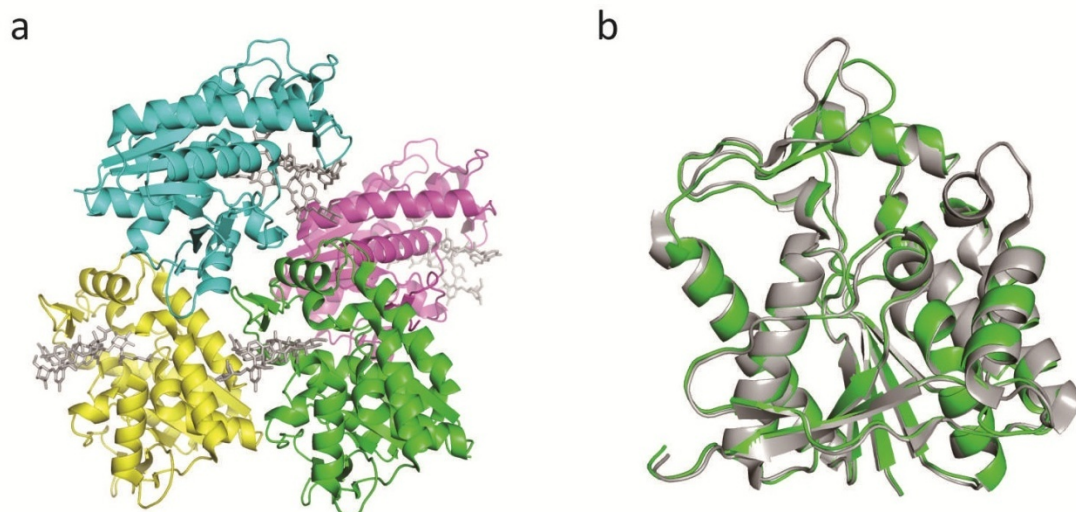


Figure S2. Tei-bound Orf2\* crystal structure and the structure superimposed with Tei-free structure. a) The crystal of the Tei-bound Orf2\* structure is packed in a  $P1$  space group (four polypeptide chains in an asymmetric unit). Tei is presented in grey sticks. b) Superimposed structures for native and Tei-bound Orf2\*. The native protein and the Tei-complexed structure are colored in in gray and green, respectively.

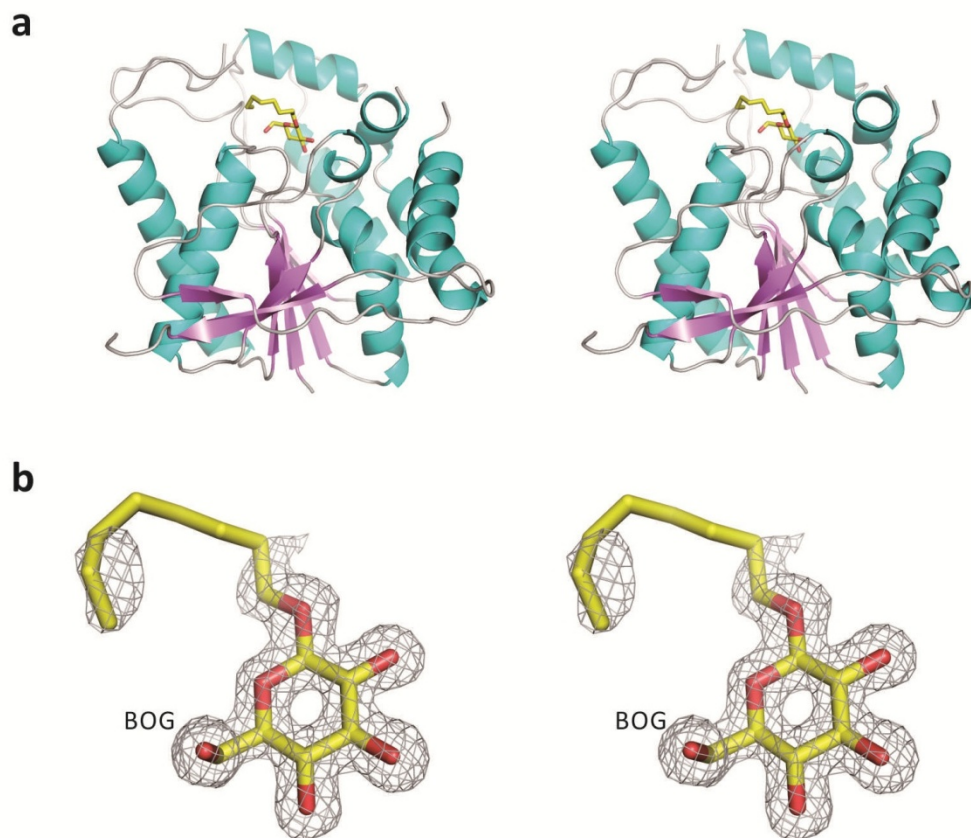
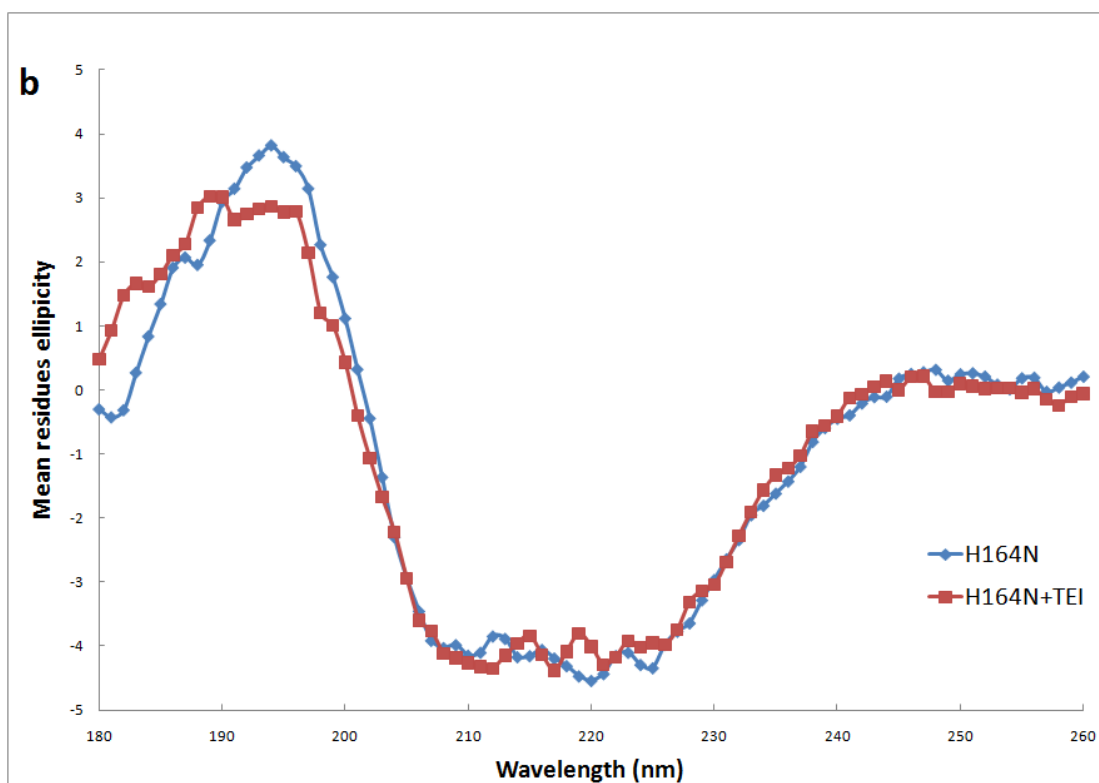
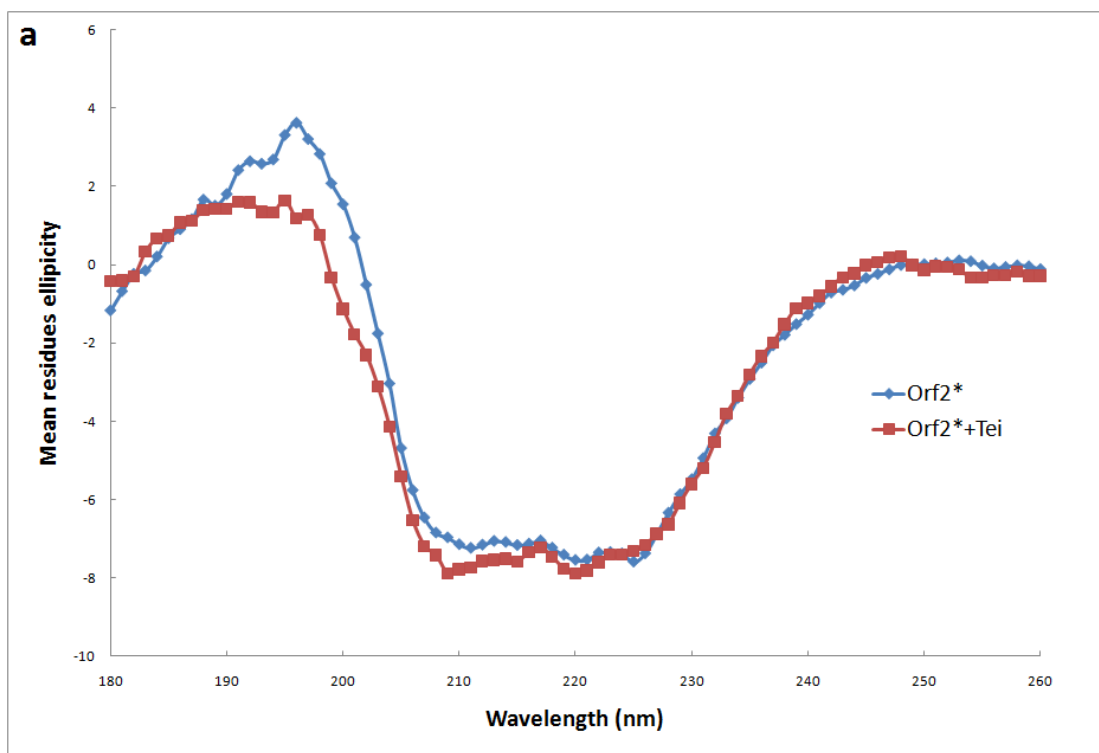


Figure S3. Stereo views of BOG-bound structure. a) Stereo view of overall BOG-bound Orf2\* structure. b) Stereo view of the  $2F_o-F_c$  electron density map for BOG contoured at  $1.0 \sigma$  (gray mesh).



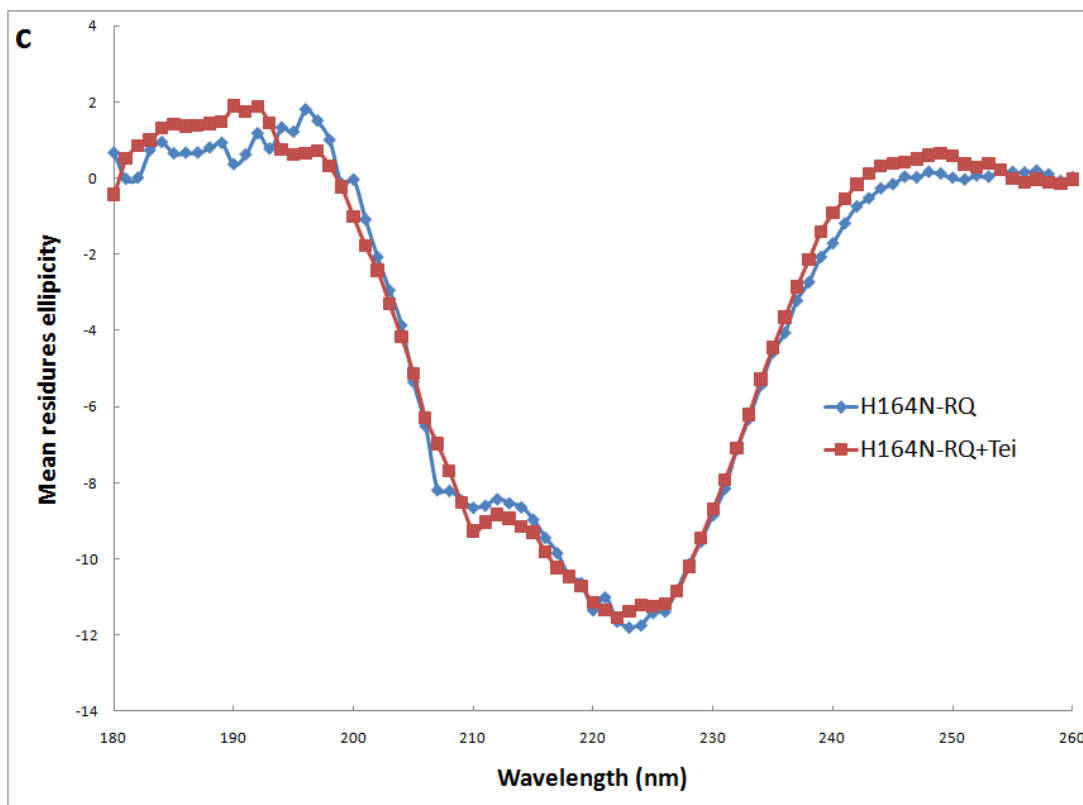


Figure S4. CD spectra of WT and mutants. a) Orf2\* (◆), Orf2\*+Tei (■). b) H164N (◆), H164N+Tei (■). c) H164N-RQ (RQ represents R116A/Q117A double mutation) (◆), and H164N-RQ+Tei (■). The CD curves of H164N-RQ in the presence or absence of Tei are almost identical, which is similar to those of Orf2\* and H164N mutant in the presence of Tei. Therefore, the capping loop may be fixed as in a closed mode in this double-mutation mutant.

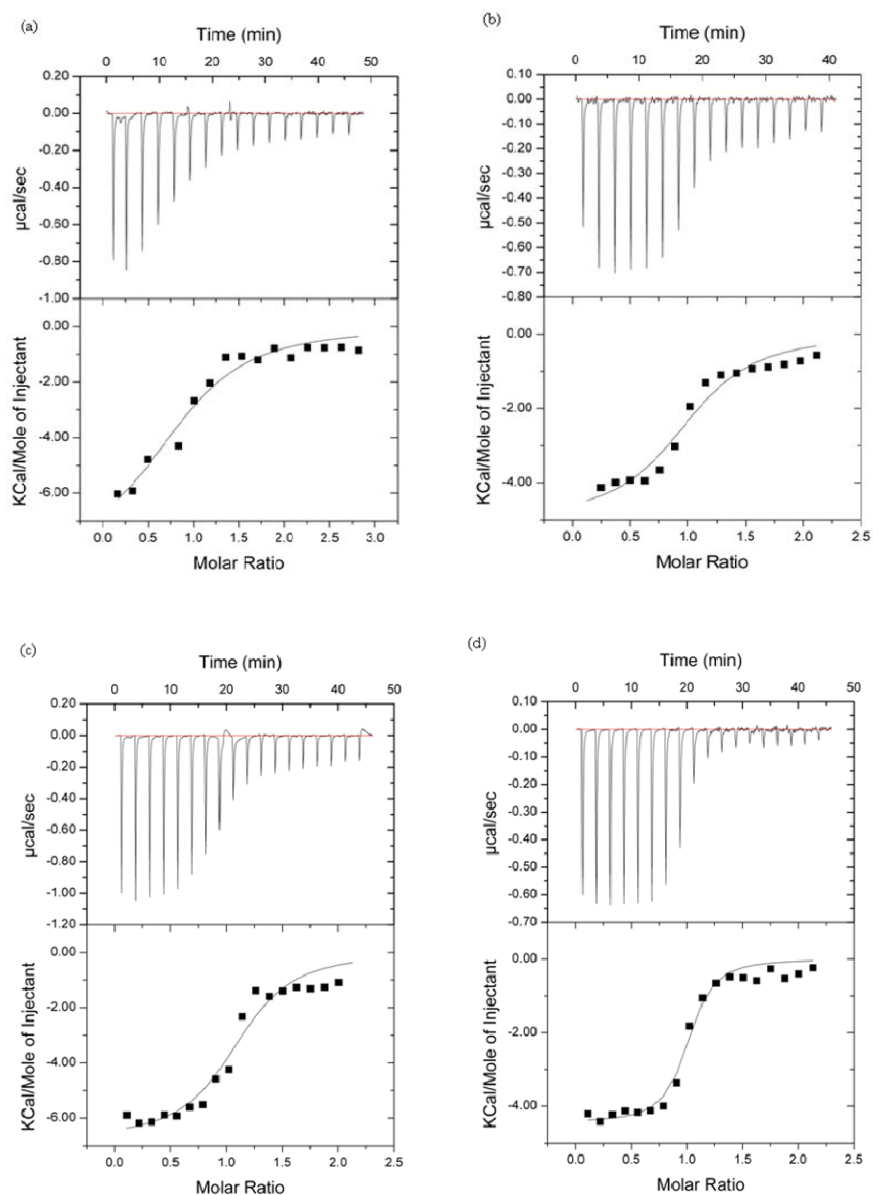


Figure S5. Binding affinity of Orf2\* mutant (H164N) vs. pseudoaglycone addended with a lipid side chain in various lengths. a) ITC thermographs for N-butyl (C4)-Glc Tei-pseudoaglycone, b) ITC thermographs for N-hexyl (C6)-Glc Tei-pseudoaglycone, c) ITC thermographs for N-octyl (C8)-Glc Tei-pseudoaglycone d) ITC thermographs for N-decanoyl (C10)-Glc Tei-pseudoaglycone.

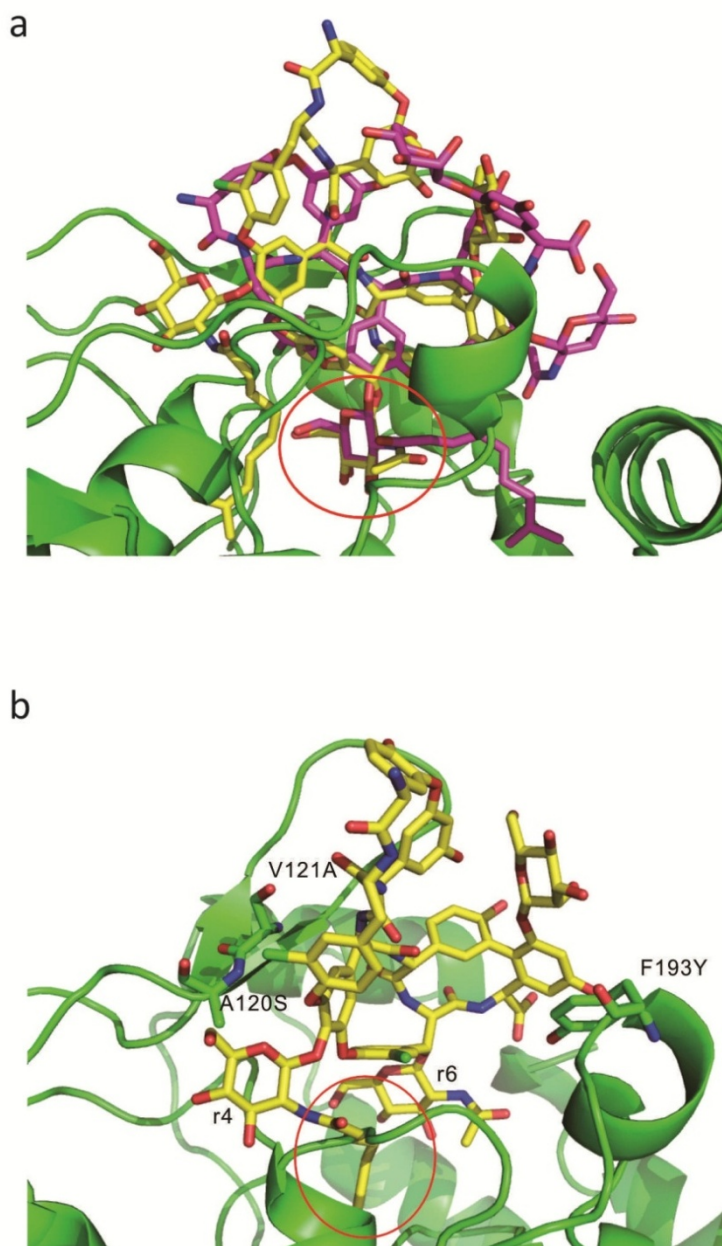


Figure S6. Docking models of r6 GlcNac Tei-bound Orf2\*. a) Superposition of r6 GlcNac (colored in yellow) with r4 N-acyl Gln (colored in magenta; highlighted by a red circle). Tei was rotated 90° in the Tei-bound Orf2\* model to superimpose r6 GlcNac with the r4 N-acyl Gln. First, the remaining part was manually optimized from colliding into surrounding residues. Then, the whole structure without the acyl side chain was subjected to the CHARMM energy minimization routine. b) A considerable collision resulted from the acyl side chain (added back here after the energy minimization) with surrounding residues as the r4-N-acyl-Gln moiety rotated 90° (highlighted by a red circle). The r4 N-acyl Gln was considered necessary to remove for the deacylation reaction to occur at r6 GlcNac. Based on this re-oriented model, residues A120, V121 and F193 were selected and mutated into A120S, V121A and F193Y to enhance the deacetylation reaction at r6 HlcNac.



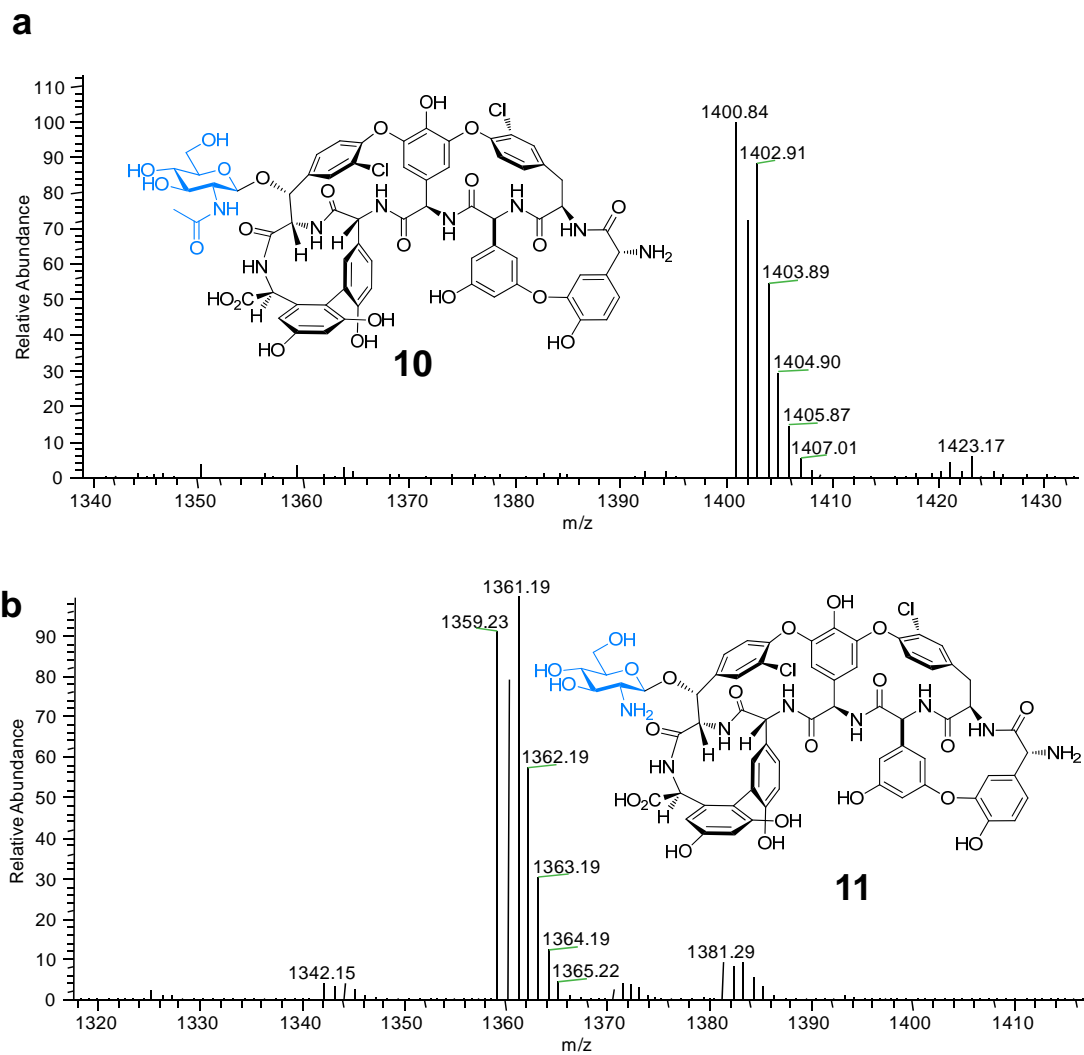
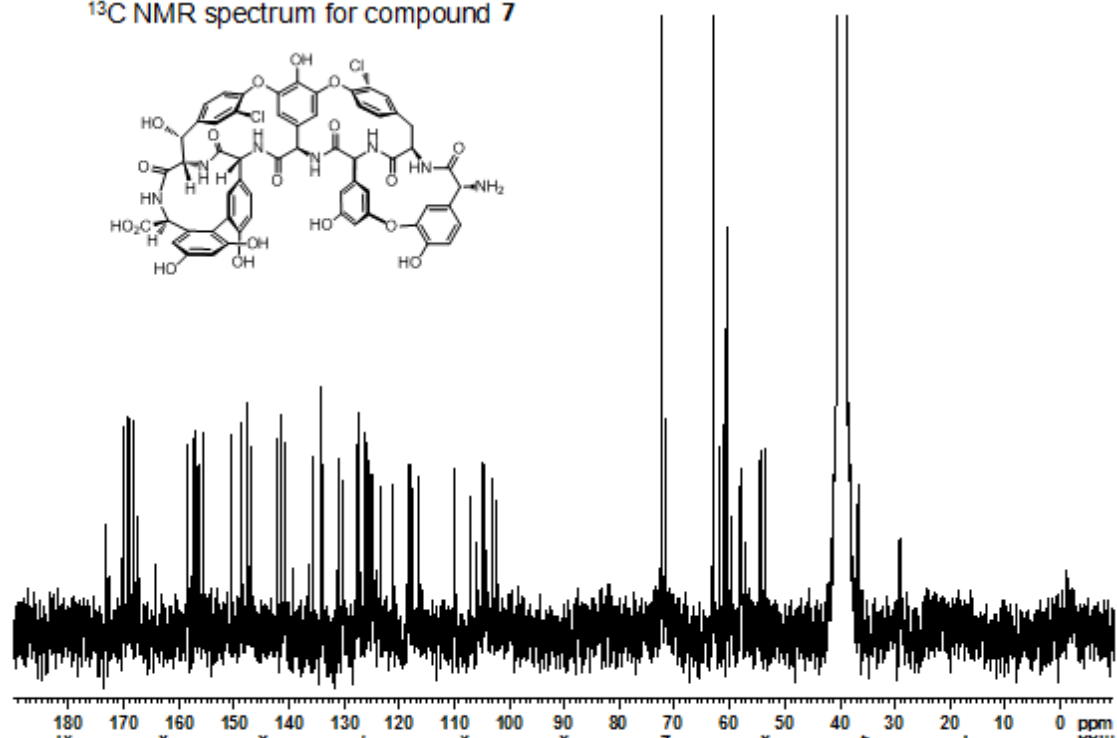
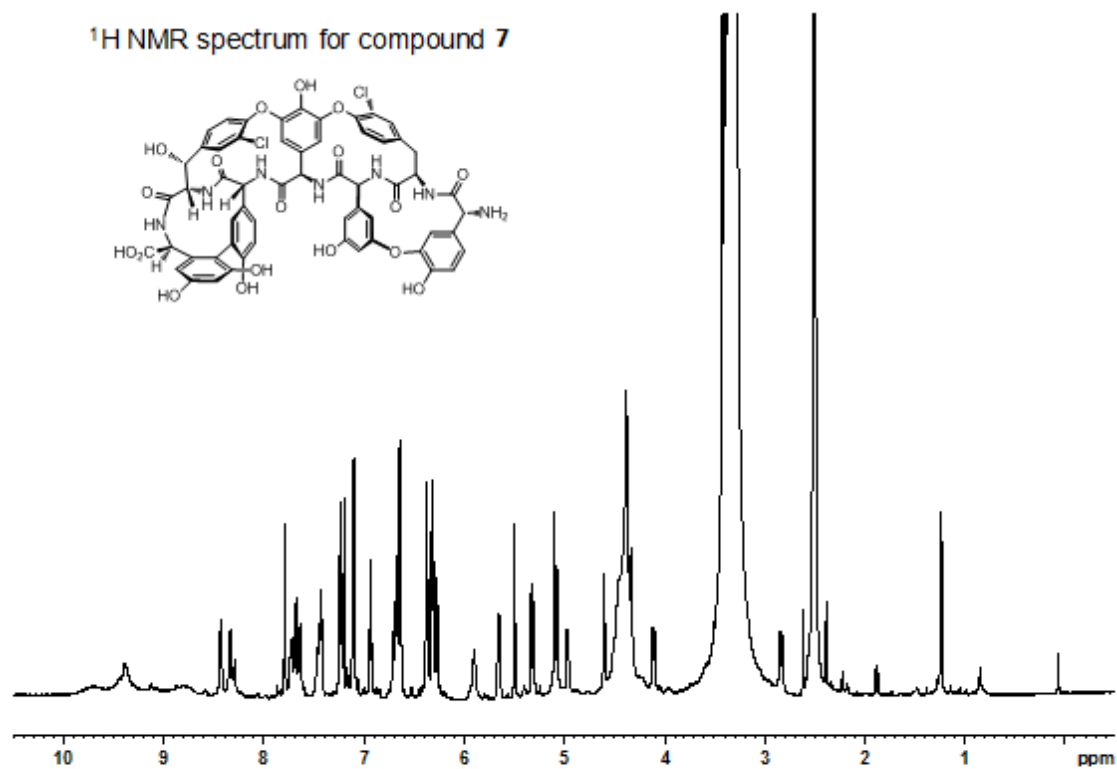


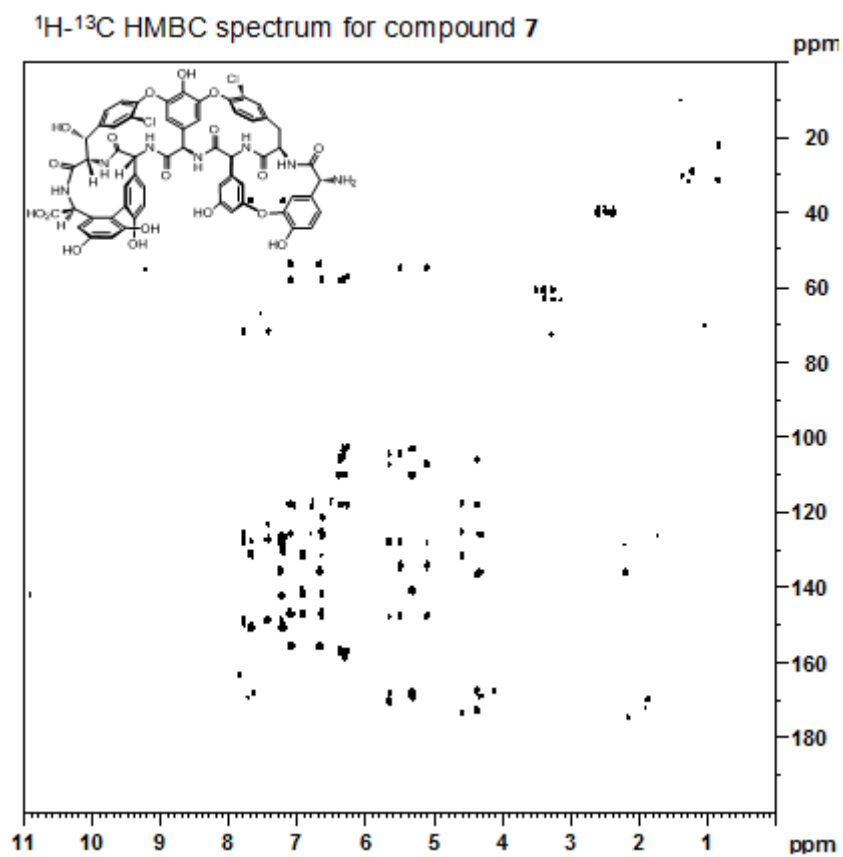
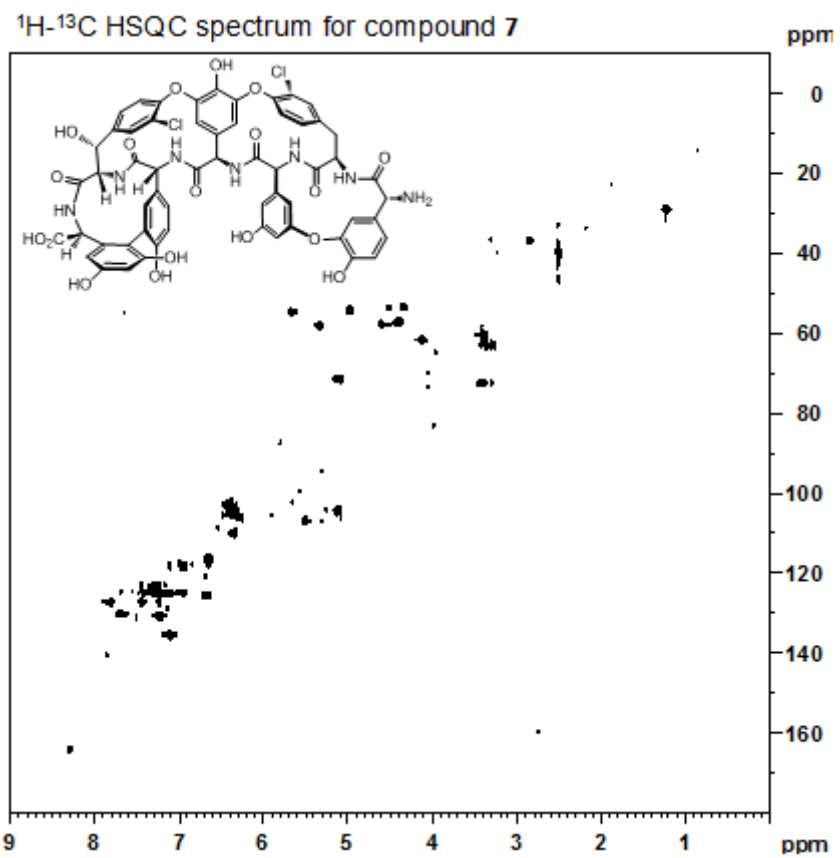
Figure S7. ESI/MS spectra for compounds shown in Fig. 2b. a) ESI/MS spectrum of aglycone (**10**). b) ESI/MS spectrum of compound (**11**).

$^{13}\text{C}$  NMR spectrum for compound 7



$^1\text{H}$  NMR spectrum for compound 7





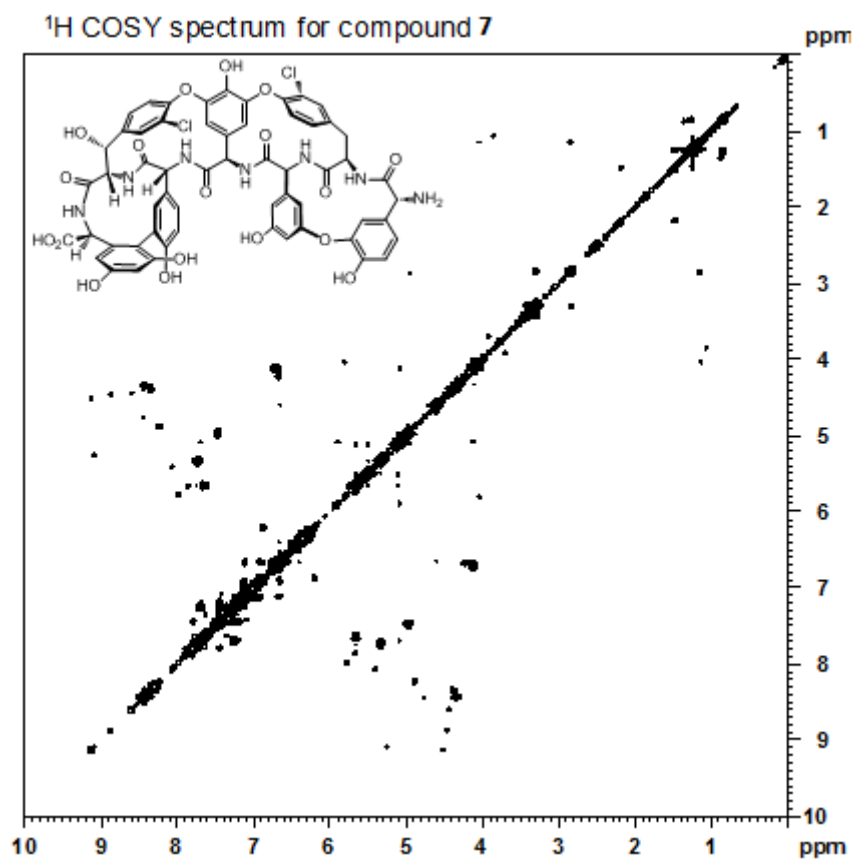
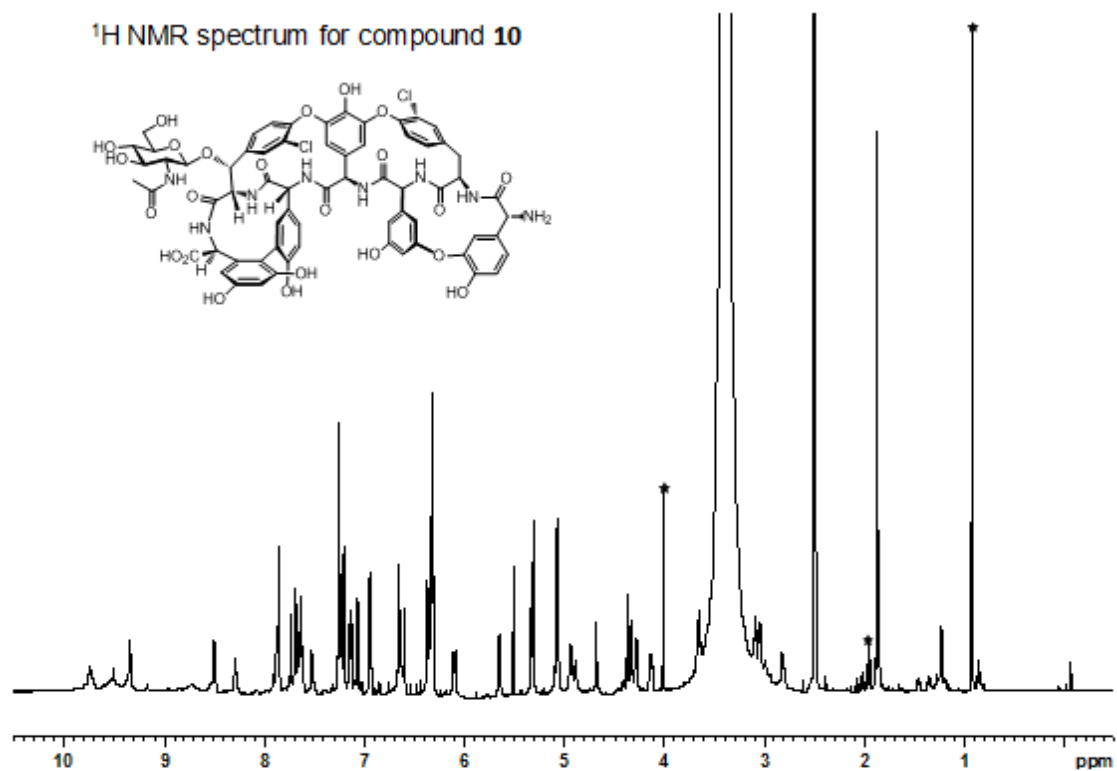
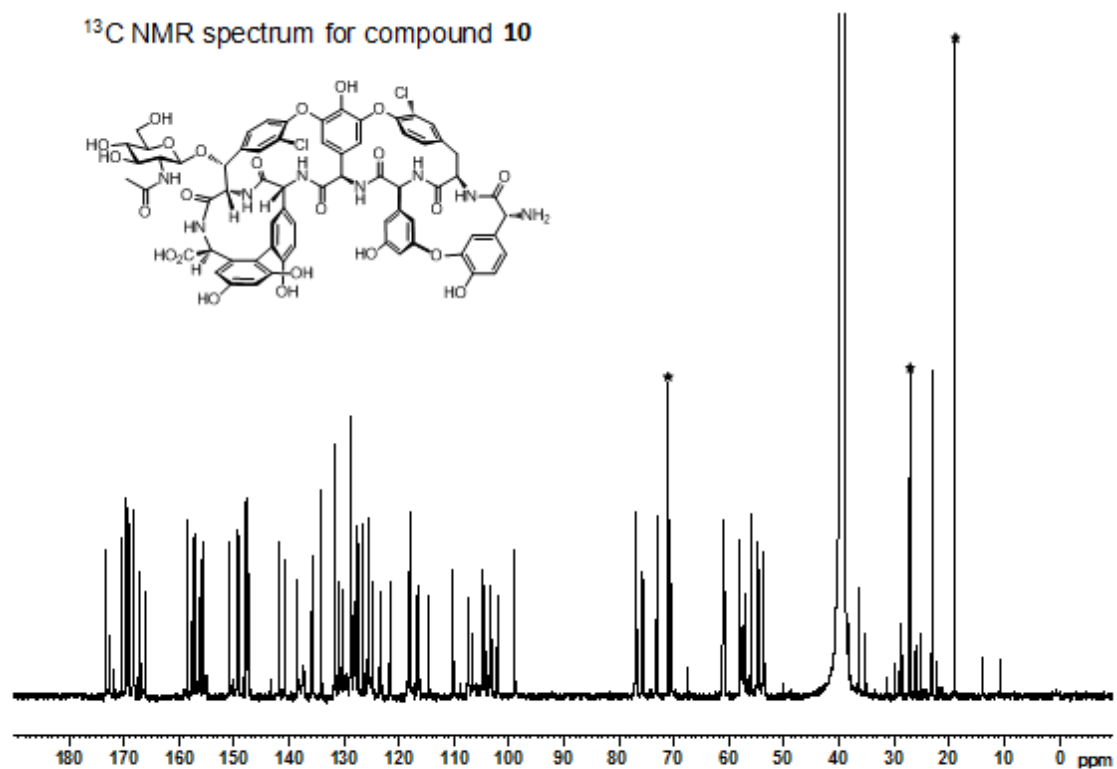


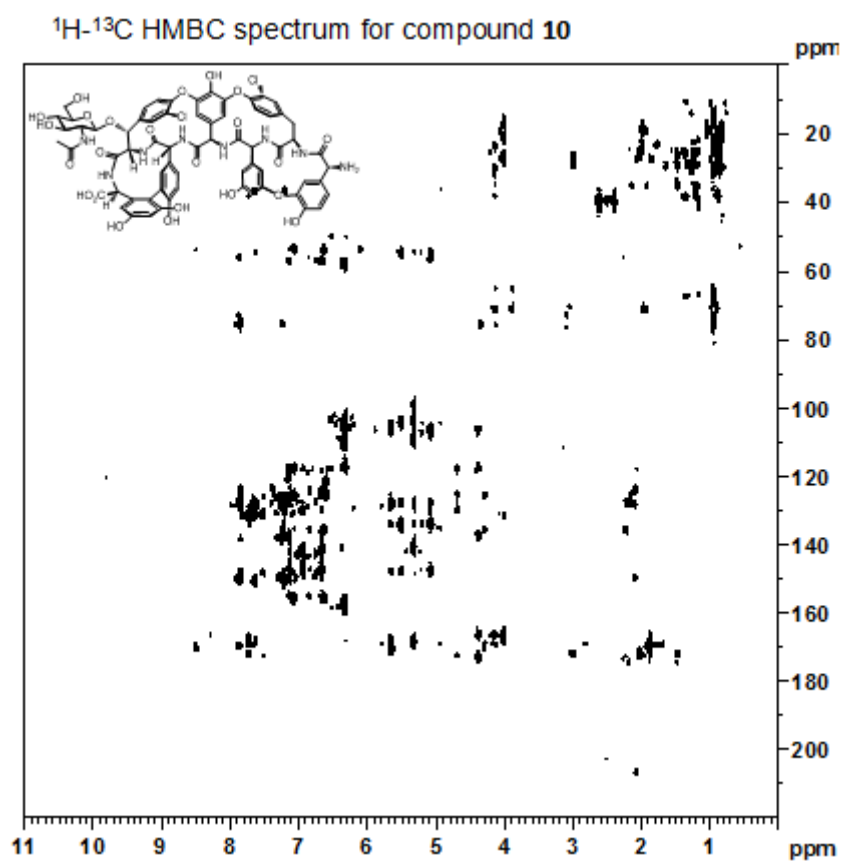
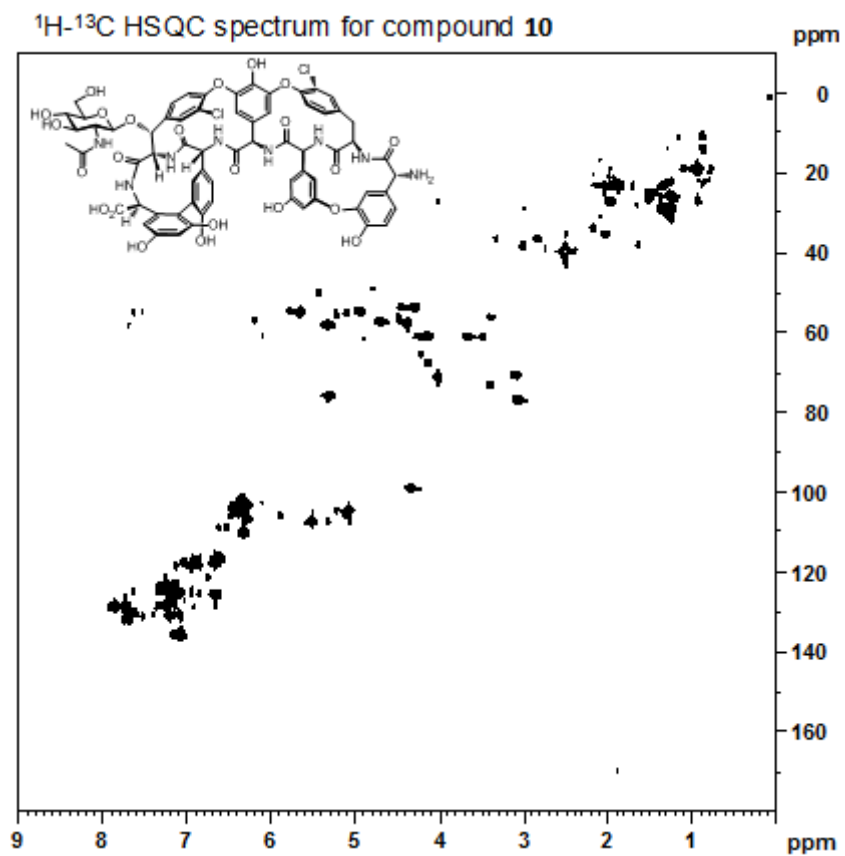
Figure S8. NMR information for compound 7. NMR spectra include <sup>1</sup>H, <sup>13</sup>C, HSQC, HMBC and COSY.

<sup>1</sup>H NMR spectrum for compound 10



<sup>13</sup>C NMR spectrum for compound 10





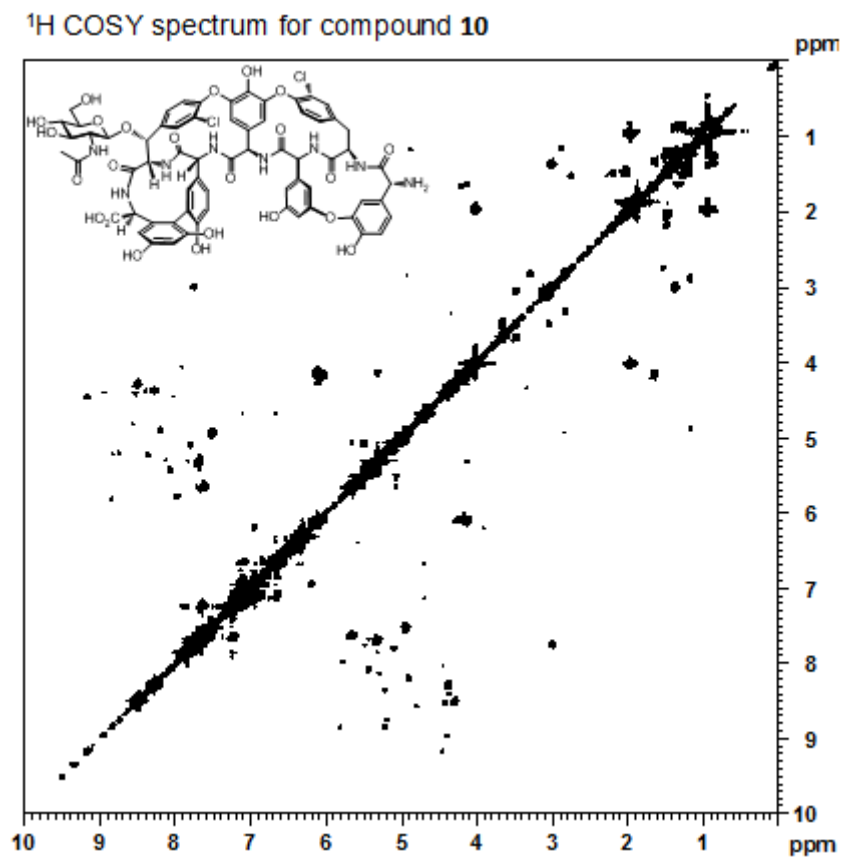
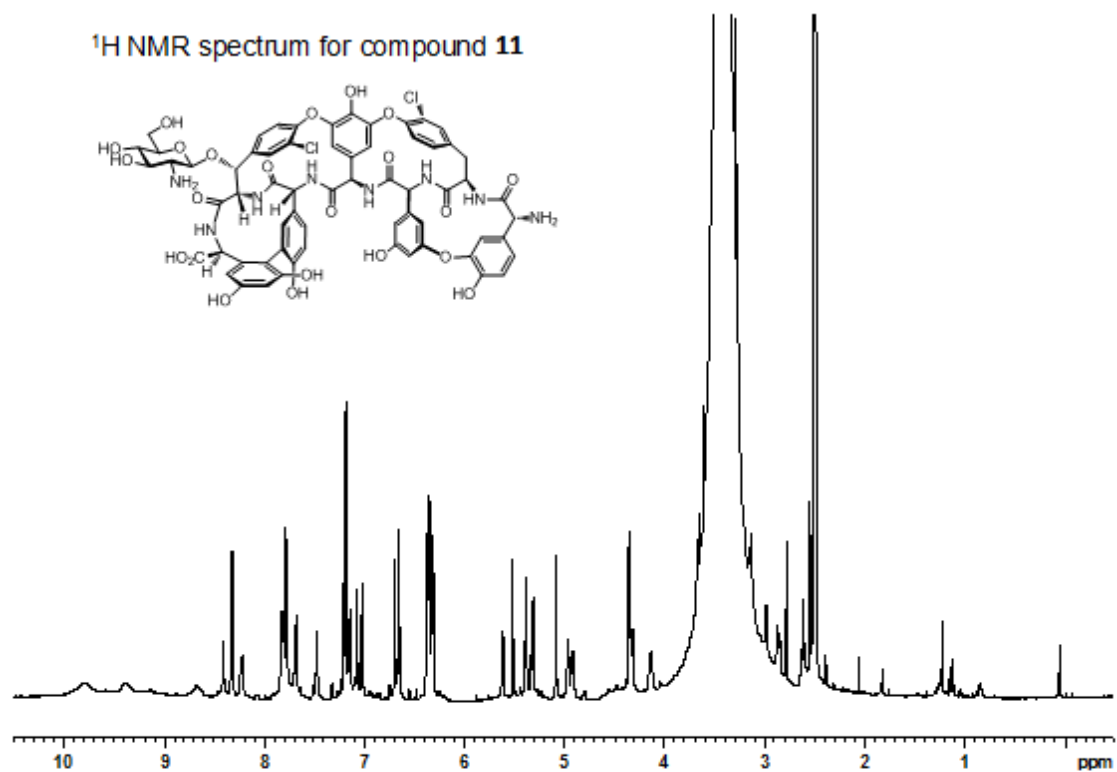
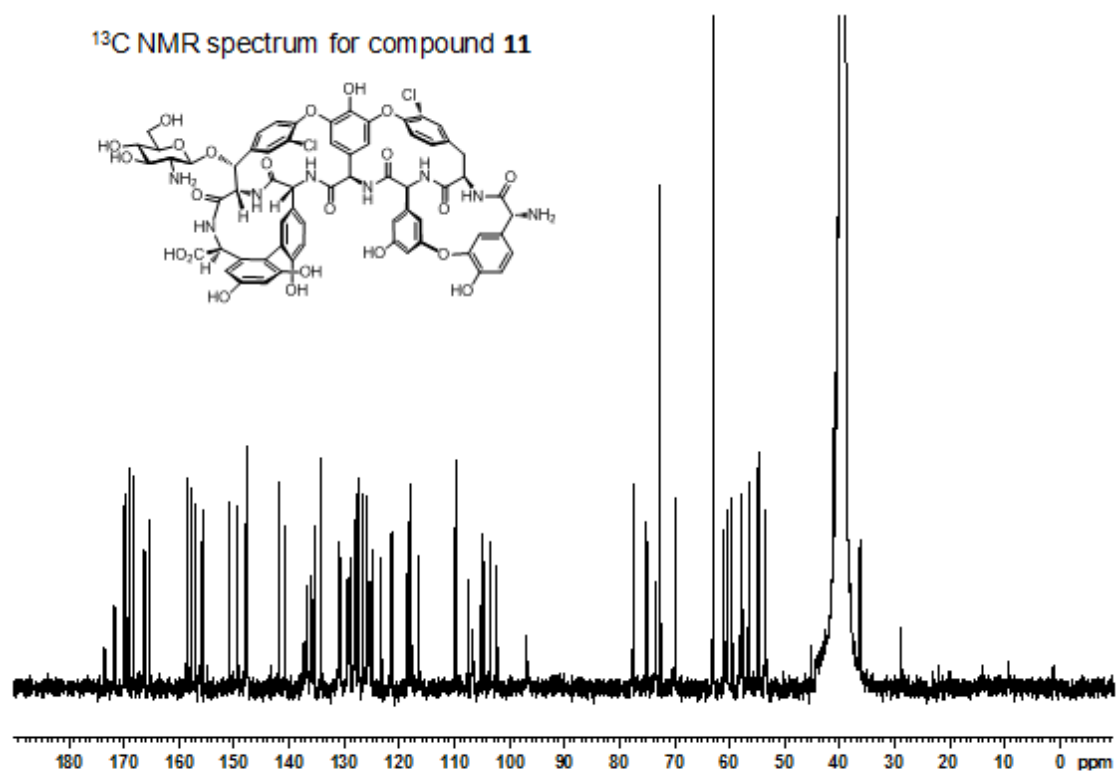


Figure S9. NMR information for compound **10**. NMR spectra include <sup>1</sup>H, <sup>13</sup>C, HSQC, HMBC and COSY. (a coeluted impurity identified is designated as \*)

<sup>1</sup>H NMR spectrum for compound **11**

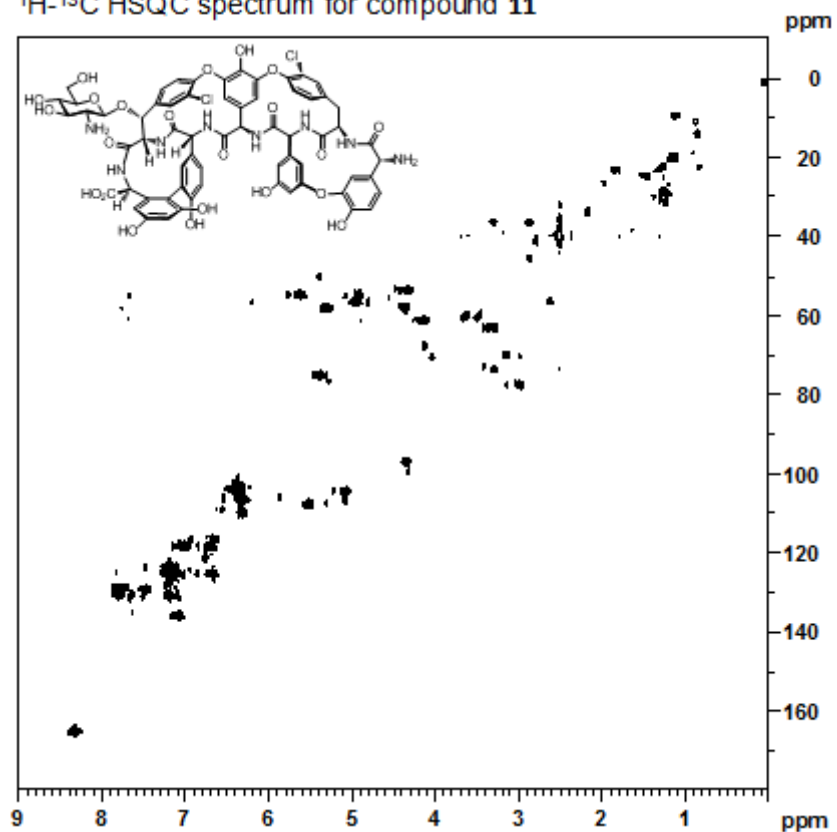


<sup>13</sup>C NMR spectrum for compound **11**

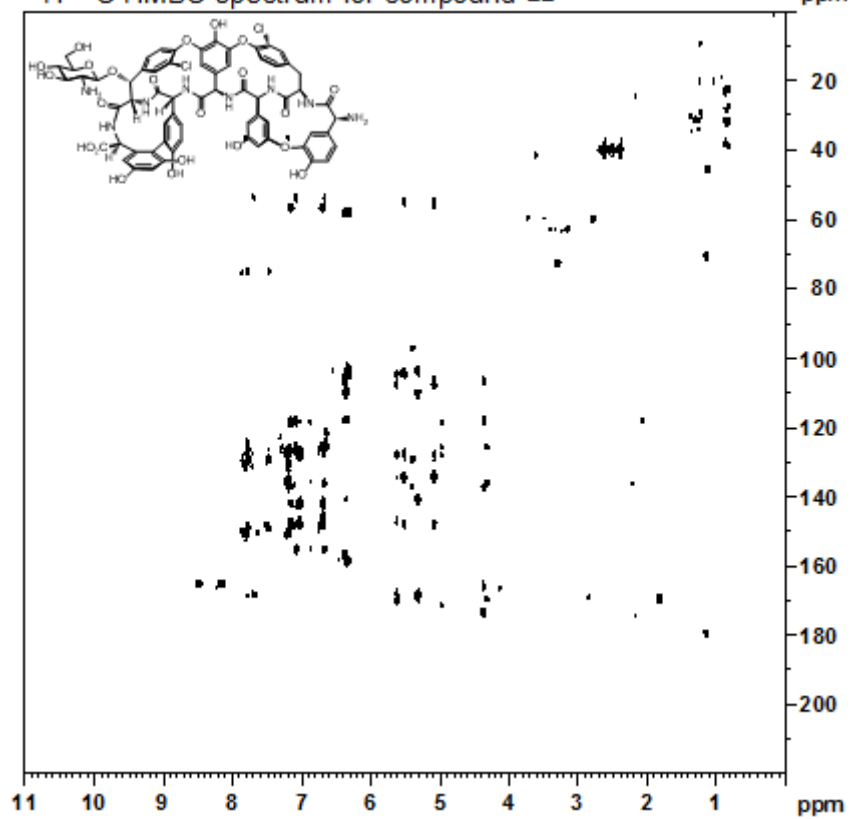




$^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum for compound **11**



$^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum for compound **11**



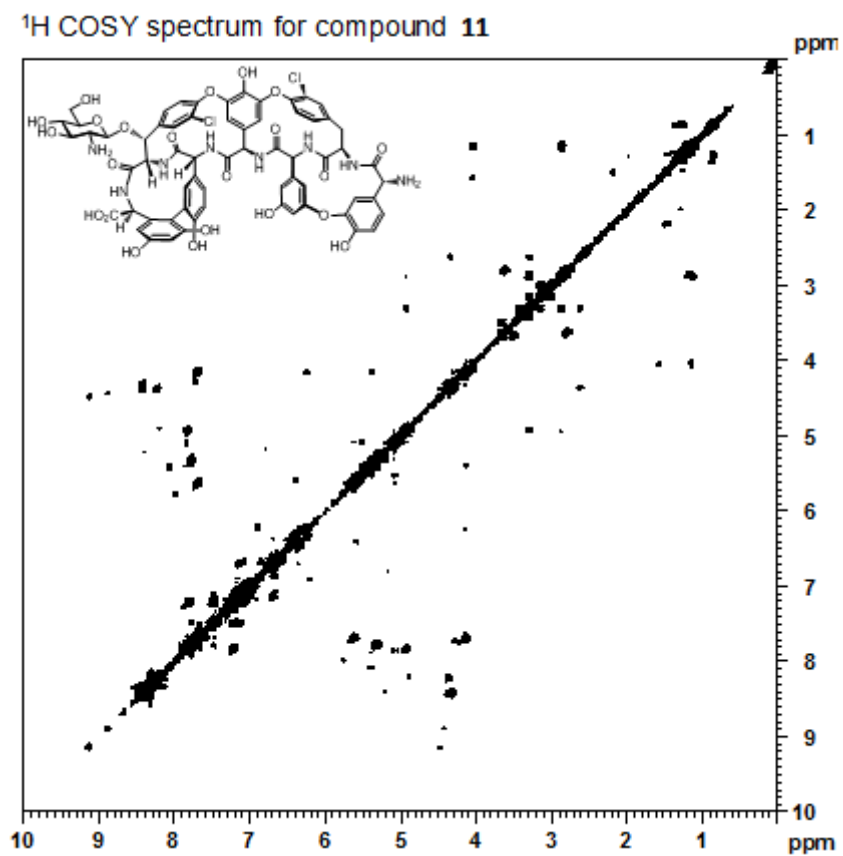


Figure S10. NMR information for compound **11**. NMR spectra include  $^1\text{H}$ ,  $^{13}\text{C}$ , HSQC, HMBC and COSY.

Table S1 Primer pairs used in the mutagenesis studies

Name of primer	Sequence
<b>For Orf2* mutant</b>	
R75Q_f (forward)	5'cgacgcggtgctctacc <u>agc</u> ggaaagaagac3'
R75Q_r (reverse)	5' gtcttctttccgctggtagagcaccgcgtcg3'
S98A_f	5'ccggttcctcgacgccatctaccgtaag3'
S98A_r	5'cttacggtagatggcgtcgaggaaccgg3'
D97L_f	5'ggccggttcctc <u>ctc</u> tccatctaccgtaagtg3'
D97L_r	5'caacttacggtagatggagaggaggaaccggcc3'
D97N_f	5' ggccggttcctc <u>aact</u> cctctaccgtaagtg3'
D97N_r	5' caacttacggtagatggaggtgaggaaccggcc3'
D163L_f	5'cggcgaacaccccc <u>ctc</u> catgagggcagc3'
D163L_r	5'gcgtcgcctcatggttgggggttcgccg3'
D163N_f	5'cggcgaacacccca <u>acc</u> atgagggcagc3'
D163N_r	5'gcgtcgcctcatggttgggggttcgccg3'
Y190F_f	5'ggaggacctgcc <u>gtt</u> cgcgcttcaaate3'
Y190F_r	5'gatttgaagaccggaaccggcaggtcctcc3'
A55L_f	5'ccttattcccc <u>gct</u> ggcgcagcgaatg3'
A55L_r	5'cattcgtcgcgcagcggggaataagg3'
A120S_f	5'aggcagaagct <u>gtc</u> gtcaacgatcac3'
A120S_r	5' gtgatcgttgacggacagcttctgcct3'
F193Y_f	5'ccgatgcggtctacaatcaggtgcg3'
F193Y_r	5'cgcacctgattgtagaccgatacgg3'
V121A_f	5'ggcagaagctggccccaacgatcactcgc3'
V121A_r	5' gcgagtgatcgttggcggccagcttctgcc3'
H161A_f	5'ggccatcggcgaagccccgacctgagggc3'
H161A_r	5'gcctcatggtcgggggcttcgccgatggcc3'
H164N_f	5'ggcgaacaccccgacaatgagggcagcgcg3'
H164N_r	5'cgcgtcgcctcattgtcgggggttcgcc3'
R116A_f	5'ccatgtggagggc <u>gcg</u> cagaagctggc3'
R116A_r	5'gccagcttctgcgcgcctccacatgg3'
Q117A_f	5' ccatgtggagggcaggg <u>cga</u> agctggcc3'
Q117A_r	5' ggccagcttcgccctgccctccacatgg3'
R116AQ117A_f	5' ccatgtggagggc <u>gcg</u> cgaagctggcc3'
R116AQ117A_r	5' ggccagcttcgccgcgcctccacatgg3'

## Supplemental Experimental Procedures

### CD analysis

Far UV CD spectra of wild-type Orf2\* and mutants were recorded between 240 and 190 nm, on a Jasco CD spectropolarimeter at 25°C. Proteins were present at 1 mg ml<sup>-1</sup> in a standard buffer solution. Each spectrum represents an average of three scans, wherein the buffer background is subtracted.

### NMR analysis

NMR analyses were performed on a Bruker Avance 600 spectrometer equipped with CryoProbe™, with tetramethylsilane (TMS) as an internal standard. Compounds were dissolved in deuterated dimethyl sulfoxide (DMSO-d<sub>6</sub>) if not otherwise stated, and spectra were recorded at room temperature:

Compound **11**. HRMS ES(+): 680.1510 [M+2H]<sup>2+</sup>, calc. For C<sub>64</sub>H<sub>58</sub>Cl<sub>2</sub>N<sub>8</sub>O<sub>22</sub> 680.152 [M+2H]<sup>2+</sup> (SI Figure 20). <sup>1</sup>H NMR (600 MHz, DMSO-d<sub>6</sub>): 9.81 (m, 2H), 9.38 (m, 1H), 8.67 (m, 1H), 8.30 (m, 3H), 7.84 (m, 4H), 7.67 (m, 2H), 7.48 (m, 1H), 7.19 (m, 3H), 7.14 (m, 1H), 7.07 (m, 1H), 7.02 (d, J 8.2, 1H), 6.70 (m, 1H), 6.65 (m, 2H), 6.35 (m, 6H), 5.62 (d, J 8.2, 1H), 5.52 (m, 1H), 5.39 (s, 1H), 5.32 (d, J 10.4, 1H), 5.08 (m, 1H), 4.98 (m, 1H), 4.91 (m, 1H), 4.33 (m, 4H), 4.14 (m, 1H), 2.99 (m, 4H), 2.87 (m, 3H), 2.78 (m, 1H), 2.62 (m, 3H), 2.55 (s, 1H), 1.24 (m, 1H). <sup>13</sup>C NMR (600 MHz, DMSO-d<sub>6</sub>): 170.08, 169.61, 169.21, 168.29, 166.32, 165.28, 158.62, 157.71, 157.08, 156.11, 155.50, 150.84, 149.38, 147.93, 147.83, 147.67, 141.95, 140.78, 136.82, 136.08, 135.53, 134.33, 130.94, 130.61, 129.28, 129.02, 127.91, 127.39, 126.45, 125.80, 125.74, 125.46, 124.92, 123.28, 121.43, 118.51, 118.27, 117.87, 116.52, 110.04, 109.58, 107.66, 105.02, 104.49, 103.50, 102.19, 77.59, 75.01, 73.46, 72.56, 69.91, 63.05, 60.93, 60.44, 59.57, 58.14, 56.55, 56.40, 54.81, 54.77, 53.56, 41.10,

40.42, 36.39. An expanded version of HSQC, HMBC and  $^1\text{H}$  COSY spectra of the deacetylated compound **11** is shown in SI Figure 10.

Compound **10**. HRMS ES(+): 712.1442  $[\text{M}+\text{H}+\text{Na}]^{2+}$ , calc. for  $\text{C}_{66}\text{H}_{59}\text{Cl}_2\text{N}_8\text{O}_{23}\text{Na}$  712.1484  $[\text{M}+\text{H}+\text{Na}]^{2+}$  (SI Figure 20).  $^1\text{H}$  NMR (600 MHz,  $\text{DMSO-d}_6$ ): 9.75 (s, 1H), 9.52 (m, 1H), 9.34 (s, 1H), 8.52 (m, 1H), 8.29 (m, 1H), 7.88 (m, 2H), 7.75 (m, 1H), 7.69 (m, 1H), 7.62 (m, 2H), 7.52 (m, 1H), 7.24 (m, 4H), 7.14 (m, 1H), 7.05 (s, 1H), 6.95 (d, J 8.3, 1H), 6.68 (m, 1H), 6.63 (m, 1H), 6.59 (m, 1H), 6.34 (m, 5H), 6.10 (d, J 11.5, 1H), 5.65 (d, J 8.1, 1H), 5.51 (m, 1H), 5.34 (m, 1H), 5.29 (m, 1H), 5.06 (m, 2H), 4.95 (m, 1H), 4.87 (m, 1H), 4.68 (s, 1H), 4.37 (d, J 6.1, 1H), 4.33 (d, J 7.3, 1H), 4.27 (m, 1H), 4.13 (m, 1H), 3.66 (m, 2H), 3.10 (m, 2H), 3.04 (m, 2H), 2.98 (m, 1H), 2.82 (m, 1H), 1.85 (s, 3H), 1.22 (m, 1H), 0.82 (m, 1H).  $^{13}\text{C}$  NMR (600 MHz,  $\text{DMSO-d}_6$ ): 173.23, 170.45, 169.72, 169.27, 169.12, 168.24, 167.03, 166.22, 158.57, 157.55, 157.07, 156.23, 155.59, 150.76, 149.28, 147.76, 147.60, 147.41, 141.75, 140.87, 138.43, 135.80, 135.61, 134.28, 131.76, 131.67, 130.97, 130.37, 128.76, 128.68, 127.90, 127.53, 126.66, 125.65, 125.57, 125.50, 124.91, 123.43, 121.67, 118.47, 117.82, 116.54, 114.59, 110.09, 104.99, 104.48, 103.37, 102.15, 98.92, 76.90, 75.74, 73.23, 70.60, 61.06, 60.97, 58.09, 57.31, 56.13, 54.77, 54.62, 53.80, 36.58, 23.19. An expanded version of HSQC, HMBC and  $^1\text{H}$  COSY spectra of the compound **10** is shown in SI Figure 9.

Compound **7**. HRMS ES(+): 1220.2099  $[\text{M}+\text{Na}]^+$ , calc. for  $\text{C}_{58}\text{H}_{45}\text{Cl}_2\text{N}_7\text{O}_{18}\text{Na}$  1220.2096  $[\text{M}+\text{Na}]^+$  (SI Figure 20).  $^1\text{H}$  NMR (600 MHz,  $\text{DMSO-d}_6$ ): 9.38 (m, 2H), 8.45 (m, 1H), 8.29 (m, 1H), 7.79 (s, 1H), 7.67 (m, 2H), 7.45 (m, 1H), 7.24 (m, 2H), 7.10 (m, 2H), 6.94 (d, J 8.3, 1H), 6.69 (m, 3H), 6.34 (m, 4H), 5.65 (d, J 8.2, 1H), 5.49 (m, 1H), 5.32 (d, J 10.6, 1H), 5.10 (m, 2H), 4.96 (m, 1H), 4.62 (s, 1H), 4.37 (m, 11H), 4.11 (d, J 11.8, 1H), 2.84 (m, 1H), 1.24 (m, 1H).  $^{13}\text{C}$  NMR (600 MHz,  $\text{DMSO-d}_6$ ): 170.27, 169.21, 169.07, 168.12, 167.44, 158.52, 157.31, 157.13, 156.38, 155.51,

150.60, 148.58, 147.56, 147.43, 146.93, 142.25, 141.40, 140.83, 135.66, 135.54,  
134.15, 130.91, 130.23, 127.69, 127.57, 127.28, 126.15, 125.82, 125.65, 125.06,  
124.80, 123.23, 121.25, 118.35, 117.74, 117.50, 116.44, 110.03, 107.04, 104.87,  
104.27, 103.10, 102.34, 72.49, 71.56, 63.07, 61.70, 61.02, 60.47, 59.73, 58.01, 57.86,  
54.72, 54.28, 53.62, 36.67. An expanded version of HSQC, HMBC and  $^1\text{H}$  COSY  
spectra of the compound 5 is shown in SI Figure 8.