**Supplementary Information for:**

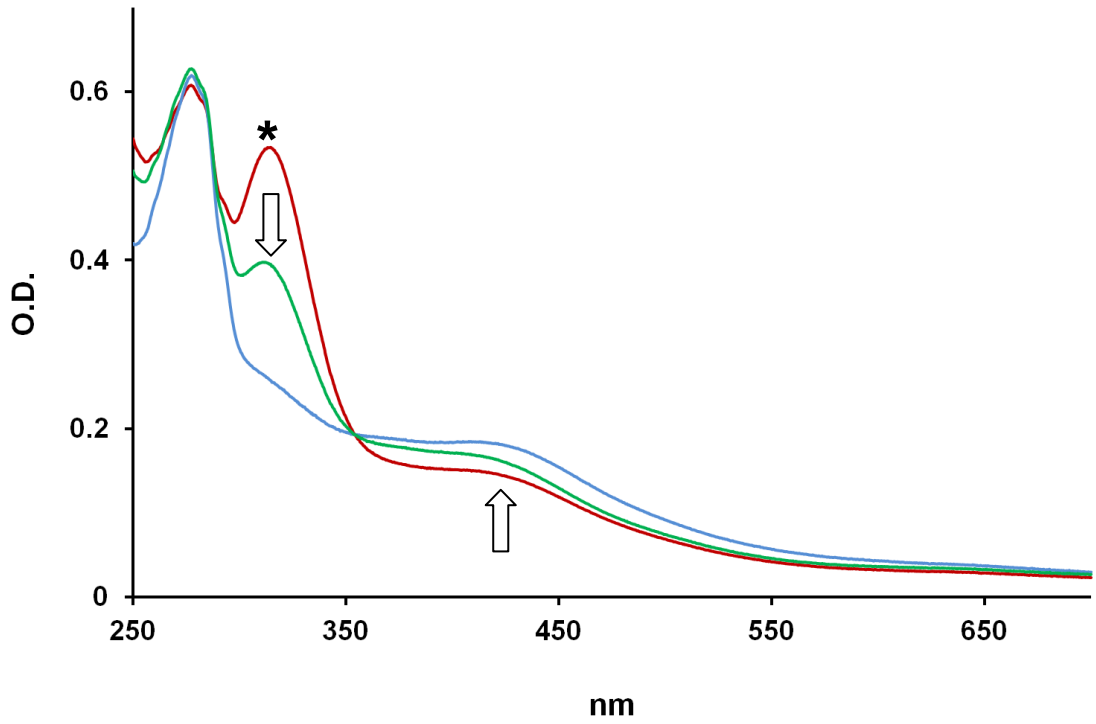
# Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in *Bacillus subtilis*

**Alhosna Benjdia[1], Alain Guillot[1], Pauline Ruffié[1], Jérôme Leprince[2] & Olivier Berteau[1]***

[1] Micalis Institute, ChemSyBio, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France

[2] Inserm U982, PRIMACEN, University of Rouen Normandy, F-76130 Mont-Saint-Aignan, France

**Olivier.Berteau@inra.fr**

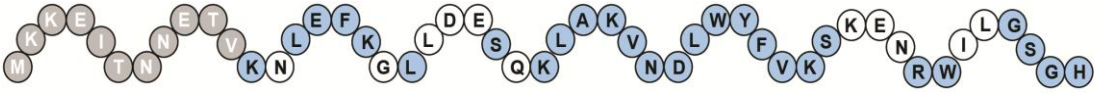



**Supplementary Figure 1 - UV-Visible absorption spectra of YydG after anaerobic reconstitution (blue trace) and reduction with sodium dithionite (2 mM) (red trace).** Following exposure to atmospheric oxygen (green trace), an increase in the region >350 nm was observed concomitantly with sodium dithionite oxidation (330 nm). The symbol "*" indicates absorbance due to reduced sodium dithionite. The concentration of the reconstituted protein was 84 µM. Protein sample was measured in a 2 mm pathlength quartz microcuvette.

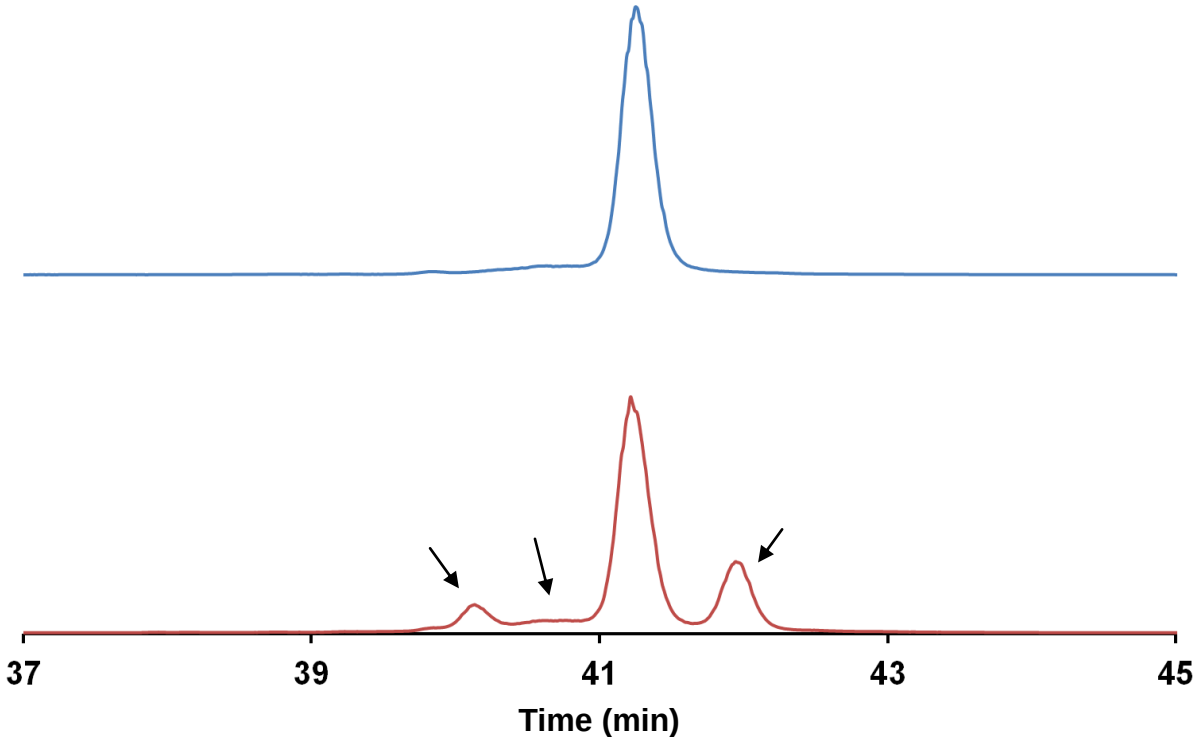
| Species | Accession Number | |
|---|---|---|
| *Bacillus subtilis* | WP_003226979.1 | M----------KKEITNNE-----T---VKNLEFKGLLDESQKLAKVNDLWYFVKSKENRWILGSGH |
| *Bacillus subtilis* | WP_014478520.1 | M----------KKENTNNE-----T---VKNLEFKGLLDESQKLAKVNDLWYFVKSKENRWILGSGH |
| *Bacillales* | WP_014115819.1 | M----------KKENTNNE-----P---VKNLEFKGLLDESQKLAKVNDLWYFVKSKENRWILGSGH |
| *Salinibacillus aidingensis* | WP_044160161.1 | M----------KKEN-----TNNET---VKNLEFKGLLDESQKLAKVNDLWYFVKSQDSRWILGSGH |
| *Bacillus coagulans* | WP_029142965.1 | MS----------KEN-----TQNSN---VKNLEFKSLVEESQKLAKVNDLWYFVKSKGNRWIVGSGH |
| *Staphylococcus pseudintermedius* | WP_014614717.1 | M----------KKEINSYKSTKENT---MKDLEFKKLVNDSEKLAKVNDLWYFVKSQSNRWIVGSGH |
| *Staphylococcus equorum* | WP_002511723.1 | M-----------KENLKVEK-QNKKEV-MKDLEFKTLINDSQKLAKVNDLWYFVKSKQNRWVVGSGH |
| *Staphylococcus condimenti* | WP_047132254.1 | MKGKGDI----KK-NKDVQIQKKDKKDAMKNLEFKNLVNDSEKLAKVNDLWYFVKSKSHRWIVGSGH |
| *Staphylococcus epidermidis* | WP_002486687.1 | MNKDLHNQKNN-K--Q----DV------MKDLEFKNLVNNSEKLAKVNDLWYFVKSKANRWVVGSGH |
| *Paenibacillus sp.* | WP_036610589.1 | M----------KK------------------LEIKELISKSEKLAKVNDLWYFVRSGEGAWIVGSG- |
| *Enterococcus caccae* | WP_010770382.1 | M----------KEL---E---------MKELV--E---KSEKLAKVNDLWYFVKSSSGAWIAGSGR |
| *Enterococcus faecalis* | WP_010710160.1 | M----------KEL---E---------MKELV--E---KSEKLAKINDLWYFVKSKGGAWIAGSG- |
| *Enterococcus faecalis* | WP_002383388.1 | MWNNYKGDIIM-KEL---E---------MKELV--E---KSEKLAKINDLWYFVKSKGGAWIAGSGK |
| *Streptococcus agalactiae* | AFV73078.1 | MTIEIKNIQREVKPILN-D---------MS---FAKVLTKKNKLDNVNDLWYFVRNSKNRWVAGSAH |
| *Streptococcus agalactiae* | WP_000810831.1 | MKQ--KNIQREVKPILN-D---------MS---FAKVLTKKNELDNVNDLWYFVRSSKNRWVAGSAH |
| *Streptococcus agalactiae* | WP_001867162.1 | M----------S---FAKVLTKKNKLDNVNDLWYFVRNSKNRWVAGSAH |


**Supplementary Figure 2 – (a) Representative YydF homologs identified in *Bacilli*, *Streptococci*, *Enterococci* and *Staphylococci* species using Blast search (E-value >50).** (**b**) **Sequence alignment of the yydF peptides using WebLogo** (weblogo.berkeley.edu).

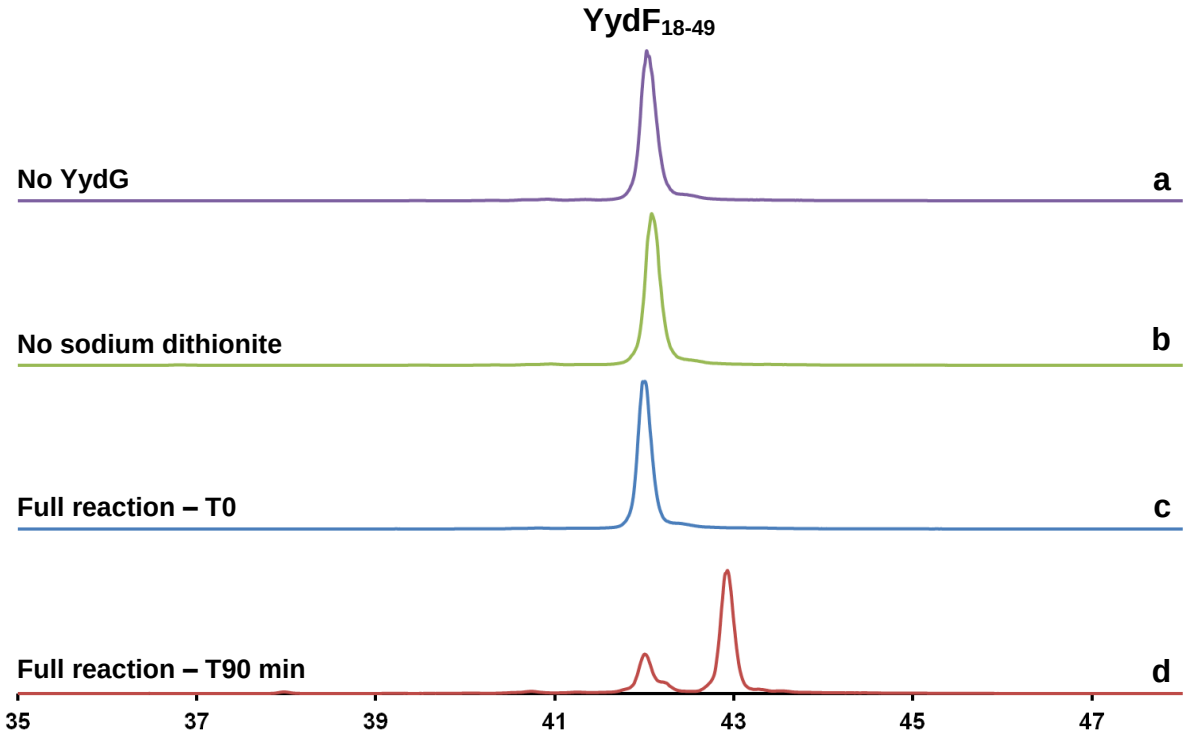
**a**



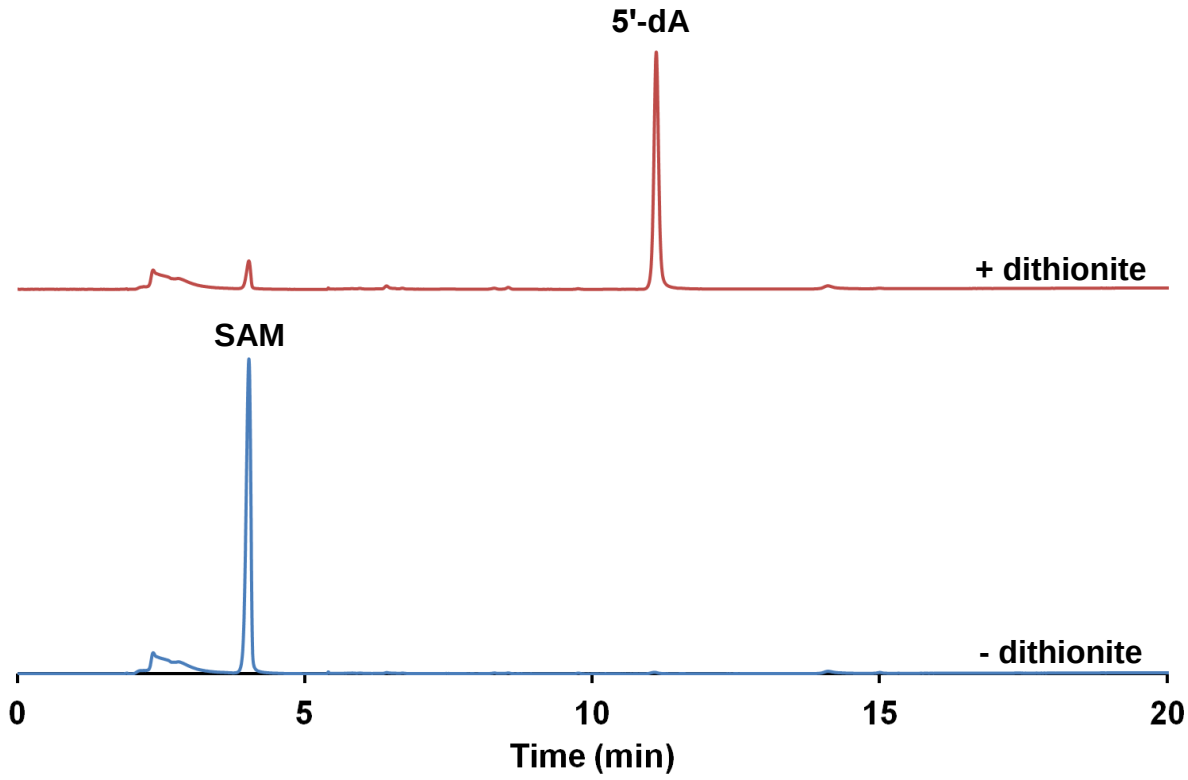

**b**



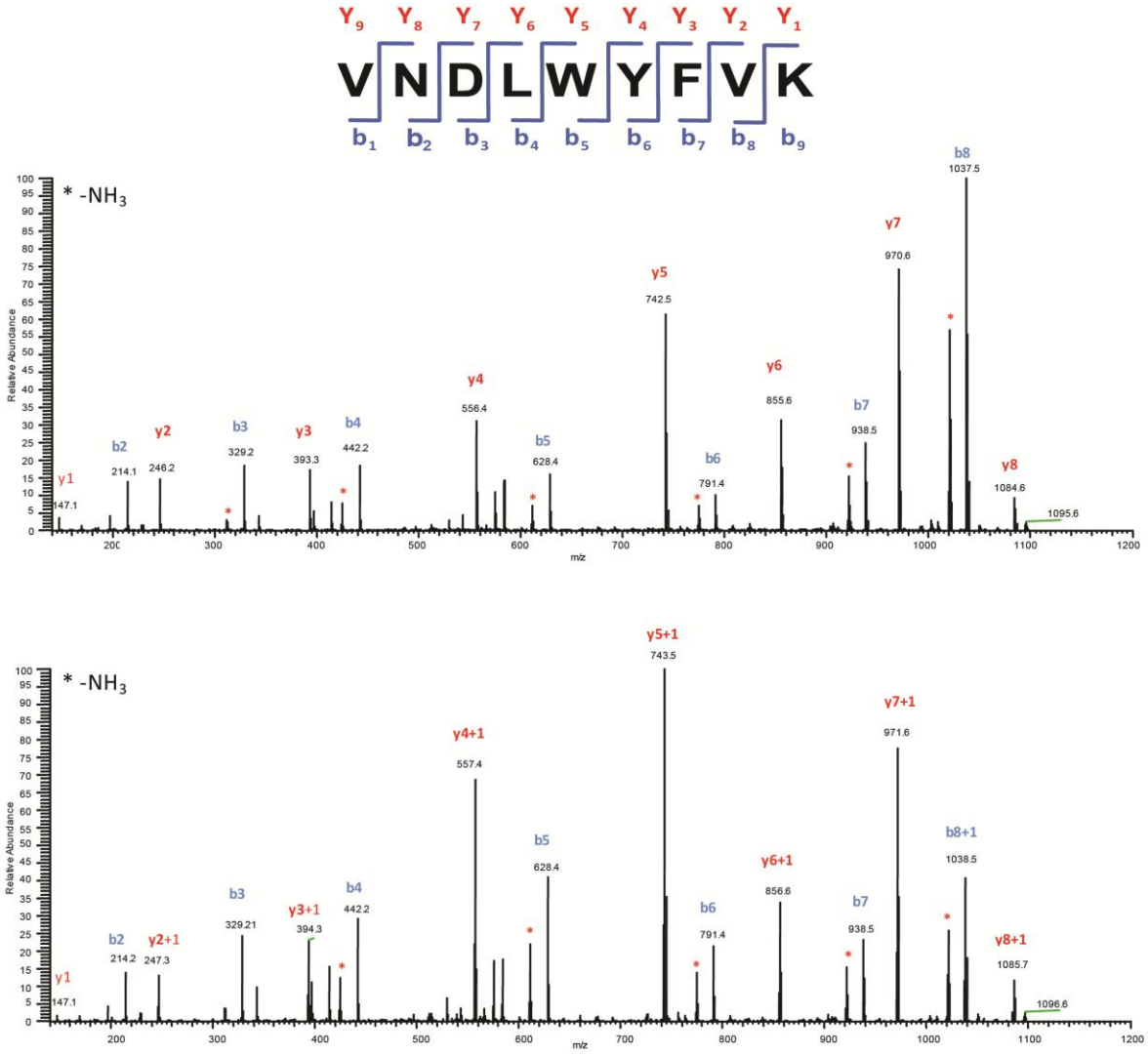

**Supplementary Figure 3 – Sequence (a) and HPLC analysis (b) of YydF at T0 (upper trace) and after 90 min (lower trace) of incubation in the presence of YydG**. Reconstituted YydG (100µM) was incubated under anaerobic conditions in the presence of DTT (6 mM), SAM (1.5 mM) and peptide substrate (1 mM) with sodium dithionite (2 mM). Peptides were analyzed using reverse phase C18 HPLC coupled with fluorescence detection (Ex/Em = 280/350 nm). Arrows indicated the formation of peptide products.



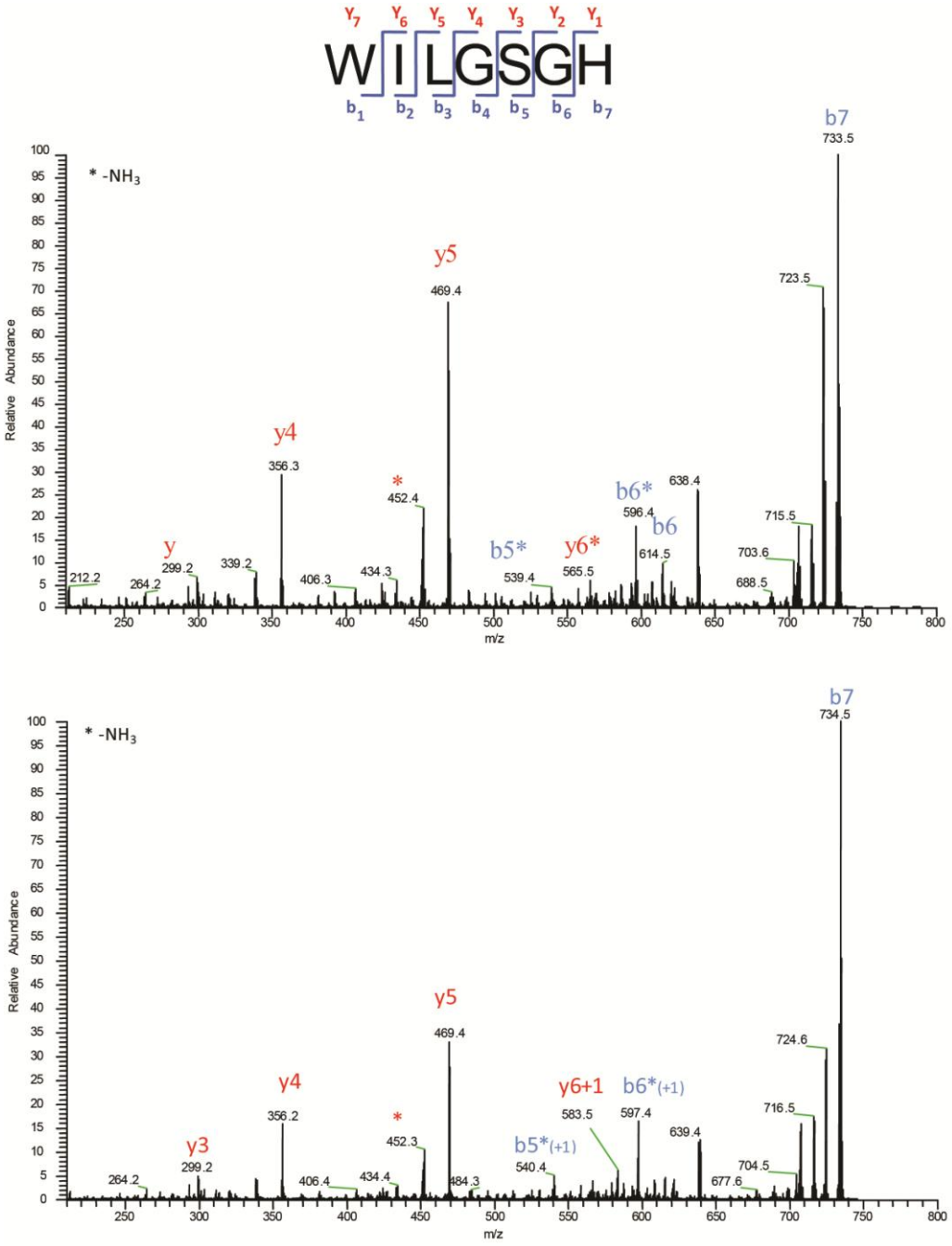

**Supplementary Figure 4 - HPLC analysis of the YydF$_{18-49}$ peptide (1 mM) incubated under various conditions with or without YydG**. (**a**) YydF$_{18-49}$ (1 mM) was incubated 90 minutes in the presence of SAM (1.5 mM), DTT (6 mM) and sodium dithionite (2 mM), without the YydG enzyme (purple trace). (**b**) YydF$_{18-49}$ was incubated 90 minutes in the presence of the YydG enzyme (100 µM), SAM (1.5 mM), DTT (6 mM) without sodium dithionite (green trace). (**c**) YydF$_{18-49}$ was incubated in the presence of the reconstituted YydG enzyme (100 µM), SAM (1.5 mM), DTT (6 mM) with sodium dithionite (2 mM) at T0 (blue trace) and after 90 minutes (**d**) (red trace). Peptide were analyzed using reverse phase C18 HPLC coupled with fluorescence detection (Ex/Em = 278/350 nm).



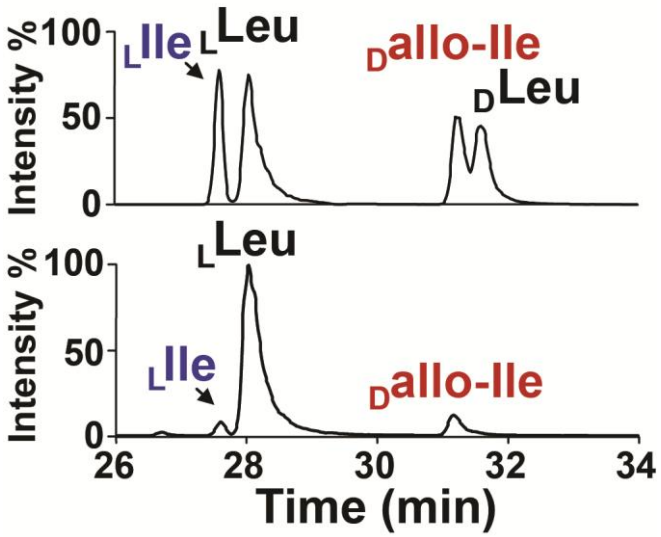

**Supplementary Figure 5 - HPLC analysis of the metabolite content of the YydG enzyme after anaerobic reconstitution before (blue trace) and after addition of sodium dithionite (2 mM) (red trace).** Metabolites were analyzed using a reverse phase C18 HPLC coupled with UV detection (257 nm).



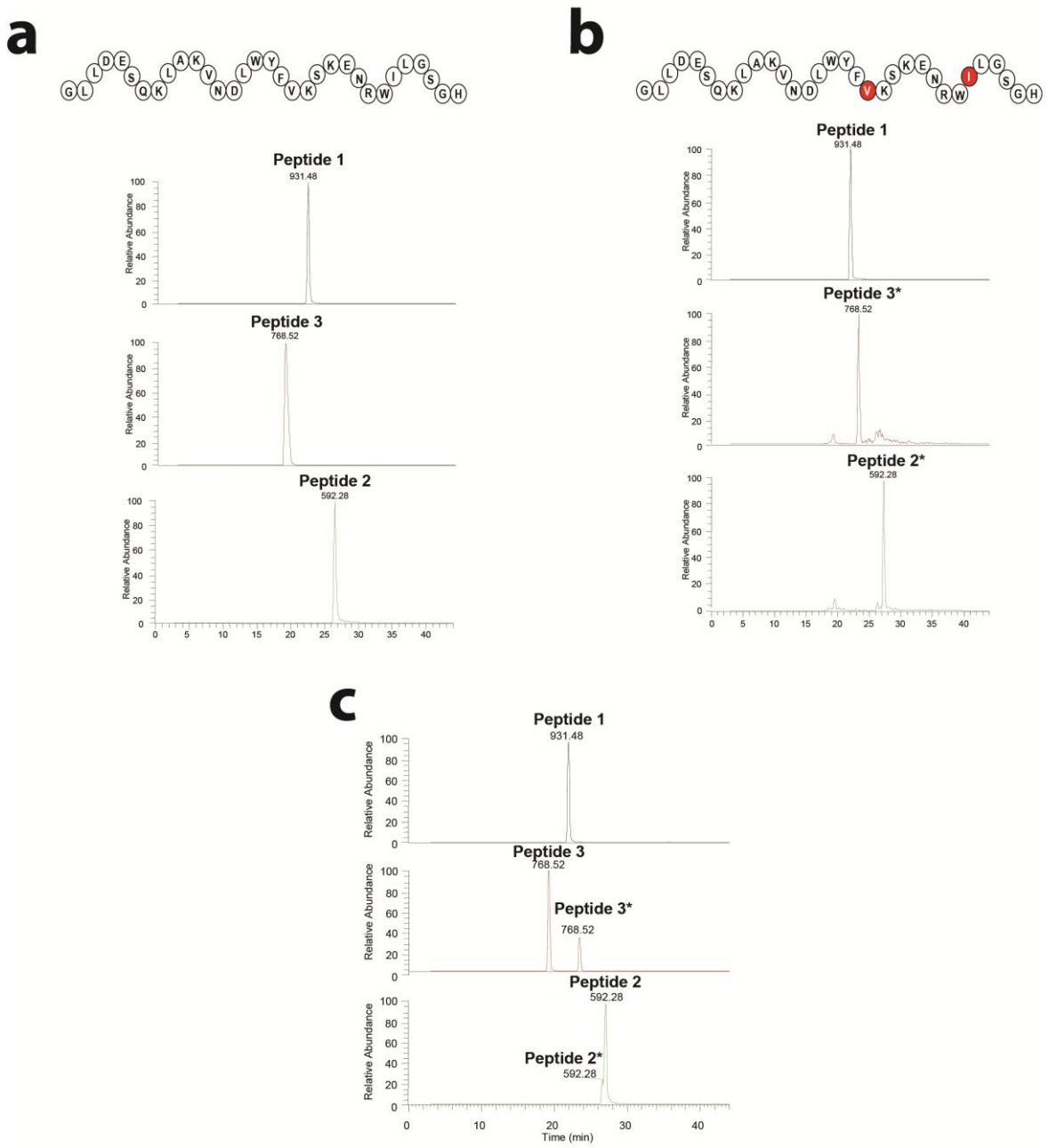

**Supplementary Figure 6 - Tryptic peptide mapping of Peptide 2 (upper panel) and Peptide 2* (lower panel) by LC-MS/MS analysis.** YydG (100µM) was reconstituted under anaerobic conditions and incubated with the $YydF_{18\text{-}49}$ peptide (1 mM) at 25°C for 90 min in the presence of SAM (1.5 mM), DTT (6 mM) and sodium dithionite (2 mM) in $D_2O$ buffer. After purification, the peptide was digested with trypsin and analyzed. (*For full assignment see Supplementary Table 1, 2 & 4*).



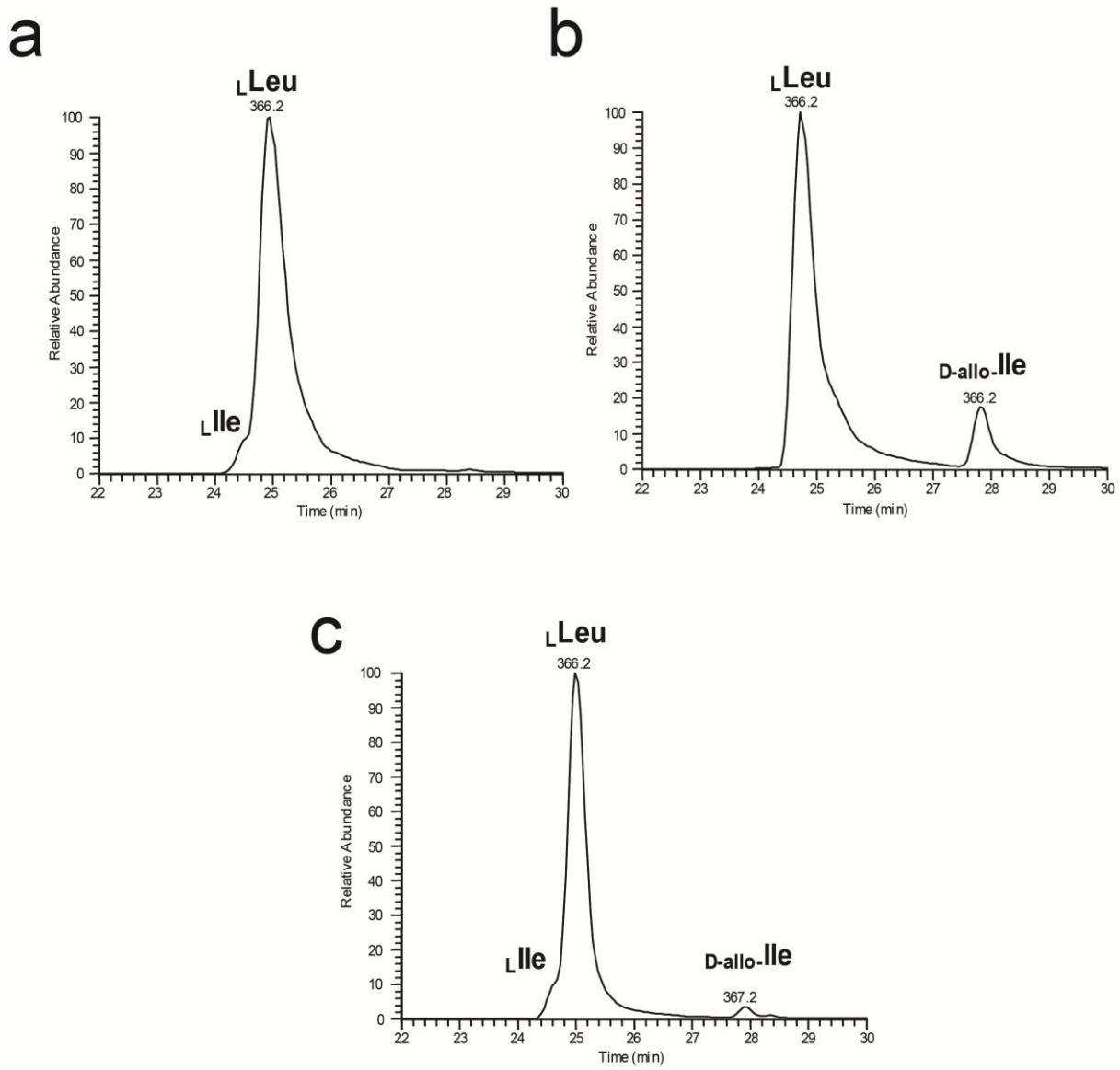

**Supplementary Figure 7 – Tryptic peptide mapping of Peptide 3 (upper panel) and Peptide 3* (lower panel) by LC-MS/MS**. YydG (100µM) was reconstituted under anaerobic conditions and incubated with the $YydF_{18\text{-}49}$ peptide (1 mM) at 25°C for 90 min, in the presence of SAM (1.5 mM), DTT (6 mM) and sodium dithionite (2 mM) in $D_2O$ buffer. After purification, the peptide was digested with trypsin and analyzed. (*For full assignment see Supplementary Table 1, 2 & 5*).



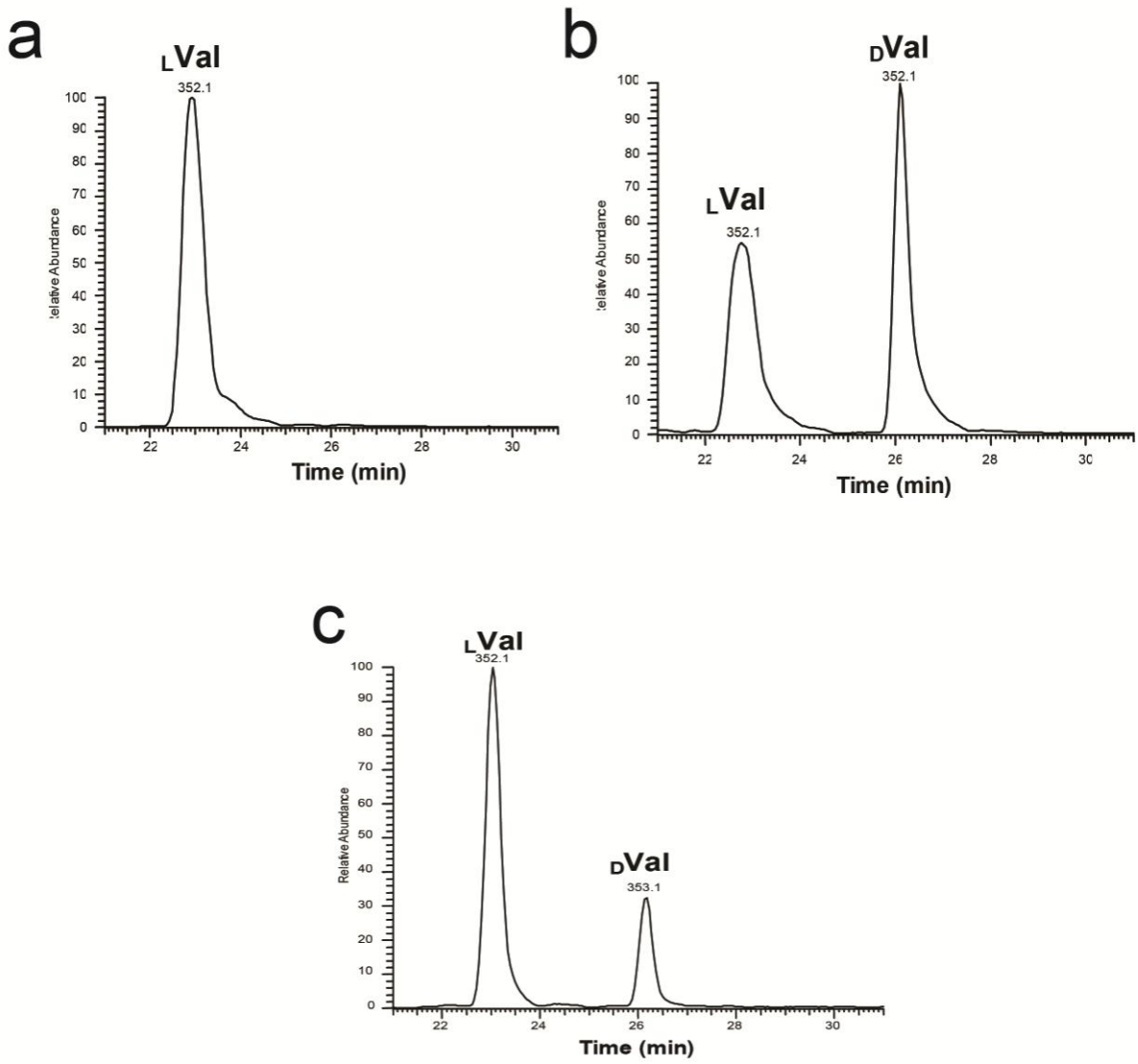

**Supplementary Figure 8 - LC-MS/MS analysis of Ile and Leu obtained after acid hydrolysis and L-FDVA derivatization of $YydF_{18-49}$ incubated with YydG in $H_2O$ buffer (lower traces) or standard amino acids (upper trace).** The ion extracted in MS/MS experiments corresponds to the major daughter ion observed for **Ile**-FDVA derivative ion (*m/z*= 412.18). For incubation conditions see legends in supplementary Fig. 6 & 7.



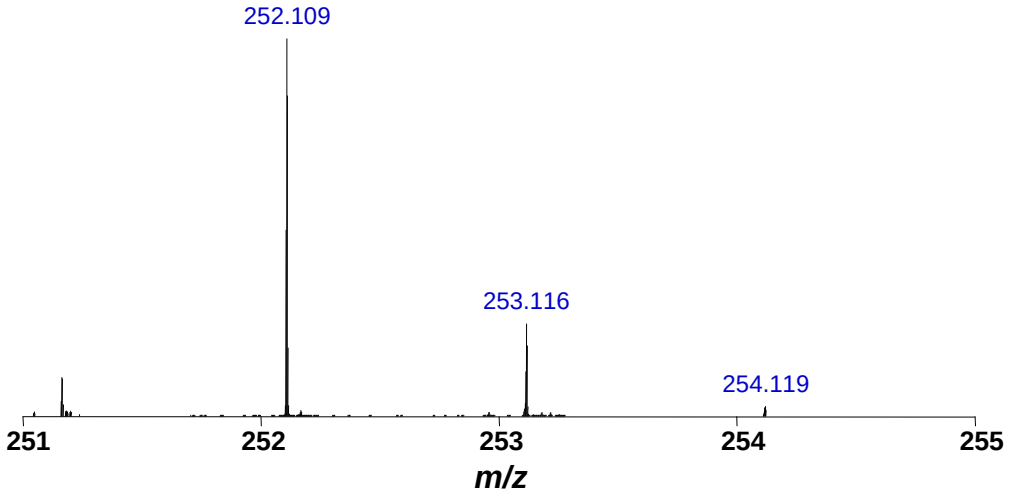

**Supplementary Figure 9 - Tryptic peptide mapping and LC-MS analysis of (a) YydF$_{18-49}$, (b) YydF$_{18-49}$ containing a D-Val and D-allo-Ile in positions 36 and 44, respectively and (c) YydF$_{18-49}$ after incubation with YydG.** In plain text the molecular weight of the corresponding tryptic peptides are indicated. For incubation conditions see legend in supplementary Fig. 6 & 7. In the peptide sequence, amino acids with a D-configuration are highlighted in red.



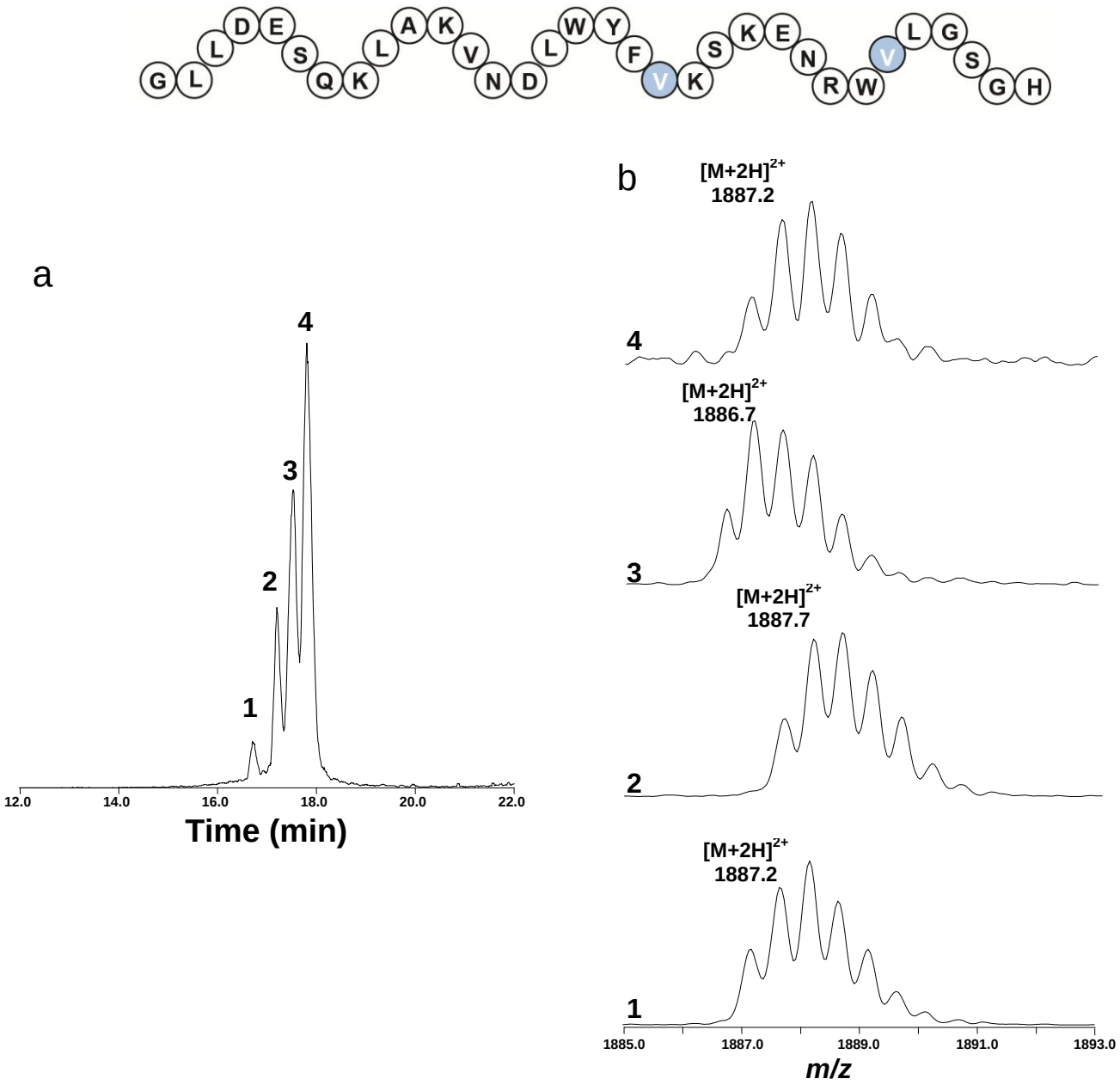

**Supplementary Figure 10 - LC-MS/MS analysis of the amino acids obtained after acid hydrolysis of: (a) YydF$_{18\text{-}49}$, (b) a synthetic peptide containing a D-Val and a D-allo-Ile in position 36 and 44, respectively and (c) YydF$_{18\text{-}49}$ after incubation with YydG in deuterated buffer.** The ion extracted in MS/MS experiments corresponds to the major daughter ion observed for Ile-FDVA derivative ion (*m/z*= 412.18). For incubation conditions see legends in supplementary Fig. 6 & 7.



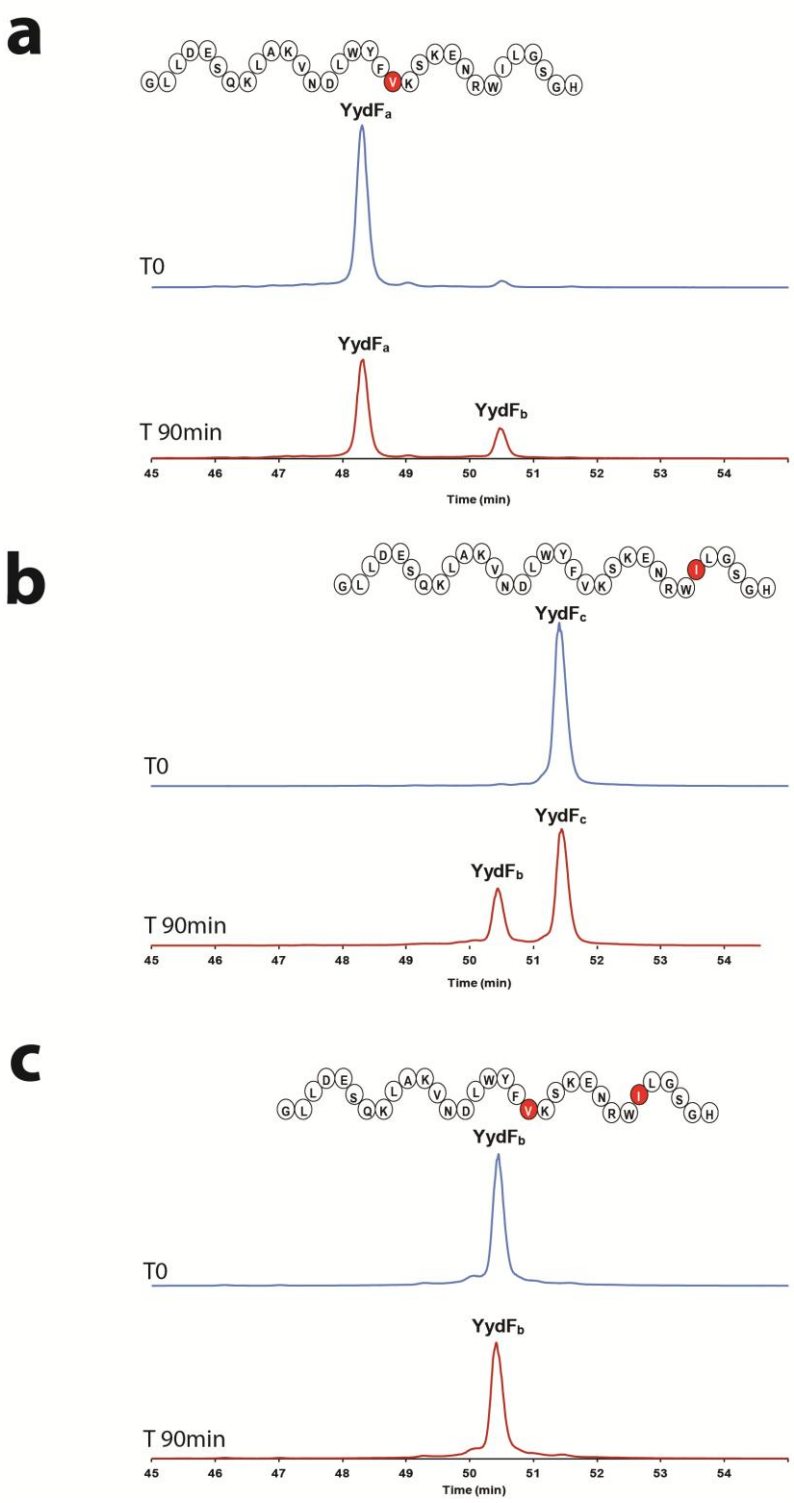

**Supplementary Figure 11 - LC-MS/MS analysis of the amino acids obtained after acid hydrolysis of: (a) $YydF_{18-49}$, (b) a synthetic $YydF_{18-49}$ containing a D-Val and a D-allo-Ile in position 36 and 44, respectively and (c) $YydF_{18-49}$ after incubation with YydG in deuterated buffer.** The ion extracted in MS/MS experiments corresponds to the major daughter ion observed for Val-FDVA derivative ion (*m/z*= 398.16). For incubation conditions see legends in supplementary Fig. 6 & 7.



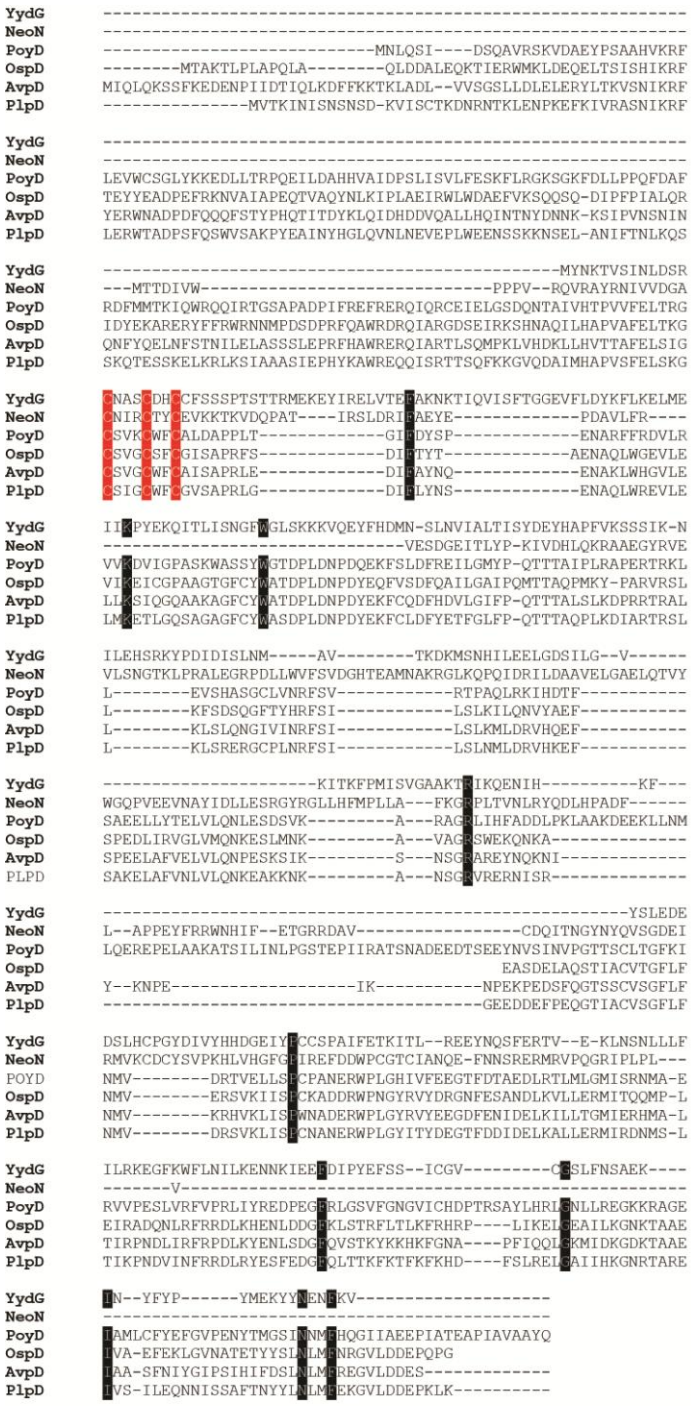

**Supplementary Figure 12 - High-resolution LC-MS analysis of 5'-dA produced by YydG during reaction in deuterated buffer in the presence of the YydF$_{18-49}$ peptide.** YydG was reconstituted under anaerobic conditions and incubated with the peptide substrate at 25°C for 90 min in the presence of SAM (1.5 mM), DTT (6 mM) and sodium dithionite (2 mM) in $D_2O$ buffer.



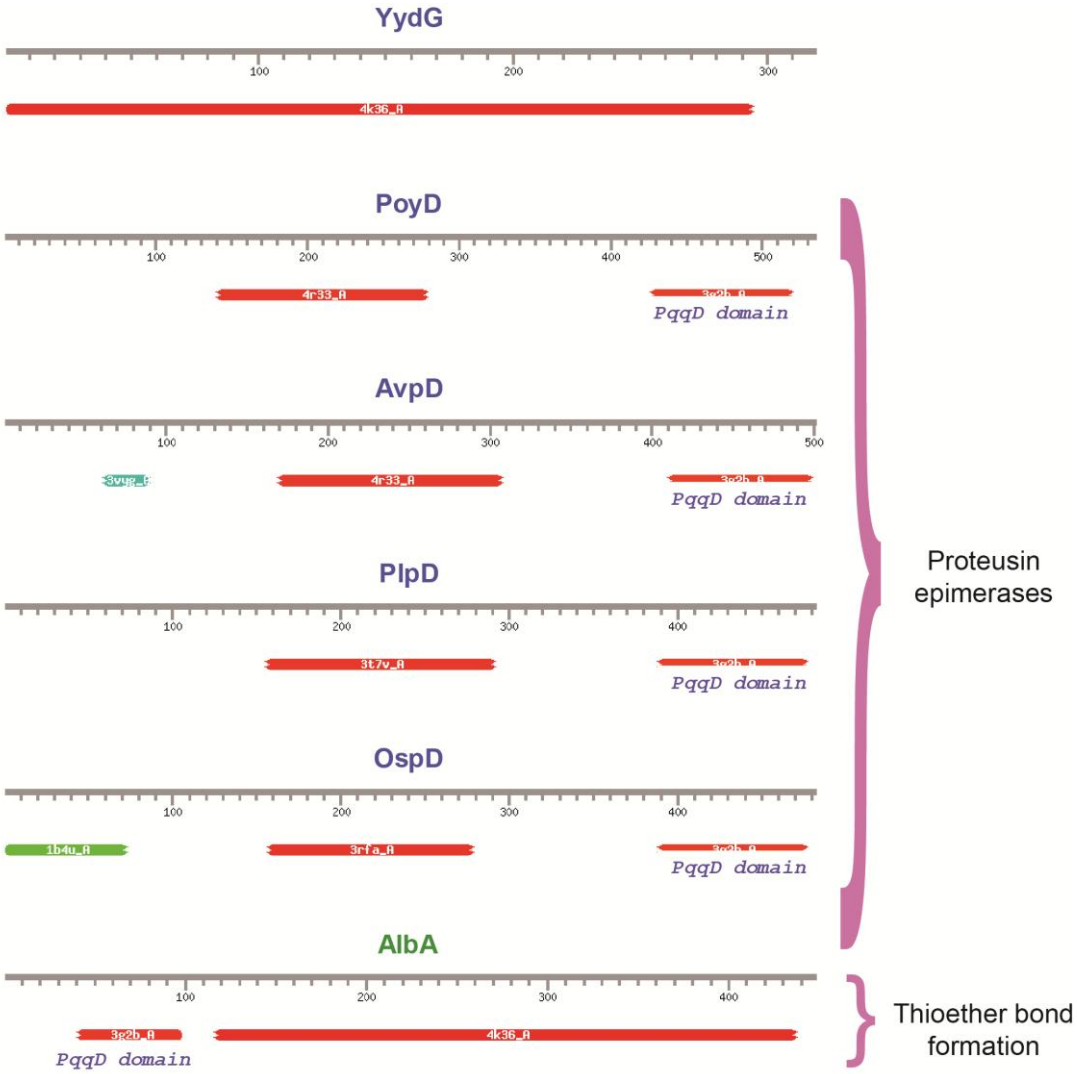

**Supplementary Figure 13 – (a) LC-MS analysis of the YydF$_{18-49}$ peptide possessing two per-deuterated valine residues in position 36 and 44, respectively ([M+2H]$^{2+}$ = 1887.7) (YydF$_{18-49}$-VD$_8$) after incubation with YydG.** (**b**) **MS spectra of the peaks 1, 2, 3 and 4 (left panel) corresponding to the substrate** (**peak 2) and the peptides with one (peaks 1&4) or two epimerized amino acids (peak 3).** YydG (100 µM) was reconstituted under anaerobic conditions and incubated in the presence of SAM (1.5 mM), sodium dithionite (2 mM) and DTT (6 mM) at 25°C for 90 min in the presence of the YydF$_{18-49}$-VD$_8$ peptide. In blue are indicated per-deuterated valine residues.



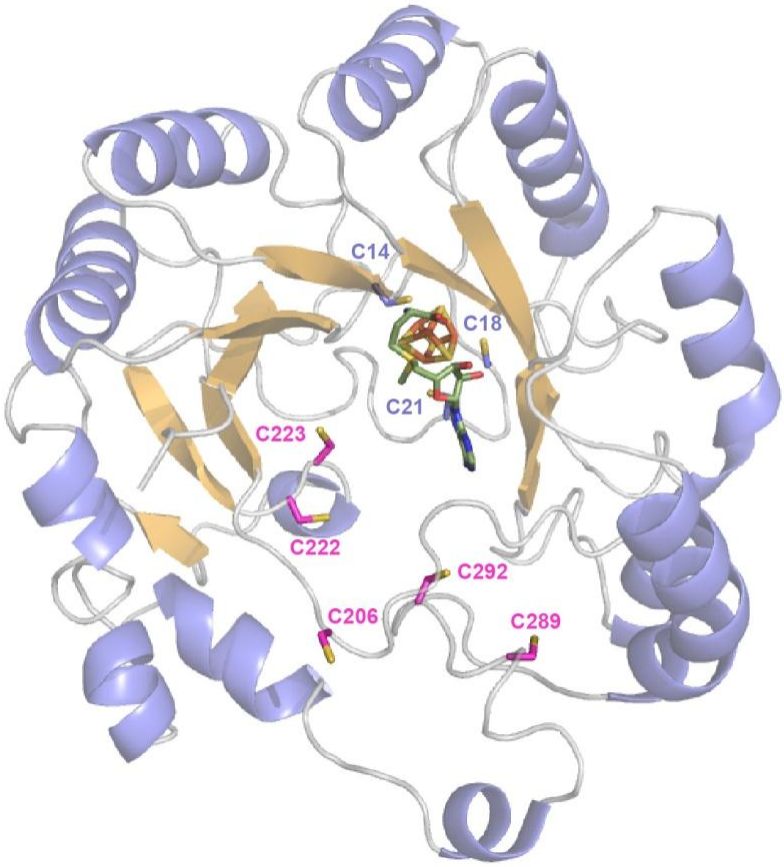

**Supplementary Figure 14 - HPLC analysis of the YydF$_a$ (a), YydF$_c$ (c) or YydF$_b$ (c) peptides (1 mM) incubated with YydF.** YydG (100 µM) was incubated 90 minutes in the presence of sodium dithionite (2 mM), SAM (1.5 mM) and DTT (6 mM) with each peptide (1 mM). Peptides were analyzed using reverse phase C18 HPLC coupled with fluorescence detection (Ex/Em = 280/350 nm). In the sequence, the D-amino acid residues are highlighted in red.




**Supplementary Figure 15 - Multiple sequence alignment of the radical SAM peptide epimerases: AvpD** (YP_321651.1)**, PlpD** (WP_019503883.1), **OspD** (CBN58198.1) **and PoyD** (AFS60640.1) **with YydG** (WP_003242609) **and NeoN** (Q53U14). Sequences were aligned using CLUSTAL Omega (1.2.1), accession numbers are indicated between brackets. In black are indicated strictly conserved residues between YydG and the proteusin epimerases.

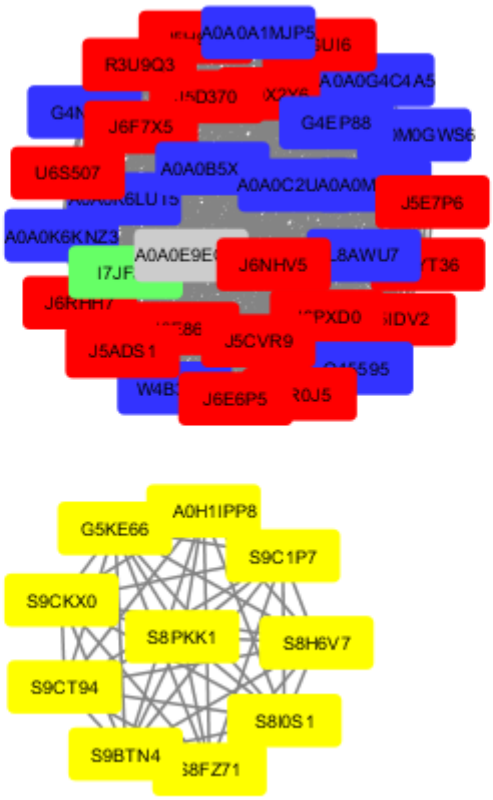



**Supplementary Figure 16 – Protein analysis using the HHpred server**[1]. Each protein was individually queried using HHPred and the pdb70_23Apr16 database. All other options were left on their default values. In red, blue and green are indicated the homologous domains found in the protein sequence with the corresponding PDB ID. The colors are related to HHpred probabilities for a homologous relationship: probability > 95% is indicated in red, probability > 70% is shown in green, probability > 55% is depicted in blue. Homology to PqqD, indicative of the presence of a RiPP precursor peptide recognition element (RREs or PqqD-like domain), is specified by PDB ID "3g2b_A" for Coenzyme PQQ synthesis protein D; helix-turn-helix, (also labeled as "PqqD domain"). The other homologous domains found in the protein are: 4k36_A for Anaerobic sulfatase-maturating enzyme; 4r33_A for NOSL; radical SAM enzyme/beta-alpha barrel, tryptophan lyase; 3vyg_C for Thiocyanate hydrolase subunit gamma; cysteine-SULF metalloenzyme; 3t7v_A for Methylornithine synthase PYLB; TIM-barrel fold,; 1b4u_A for LIGA, LIGB, protocatechuate 4,5-dioxygenase; non-heme iron protein.

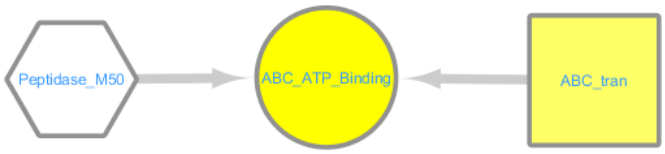



**Supplementary Figure 17 - Structural model prediction of YydG using i-Tasser[2]**. The best identified structural analog in PDB is anSME (PDB ID: 4k36A). SAM is depicted in green and colored by atom elements; iron-sulfur is shown in orange and yellow. Cysteine residues ($Cys_{14,}$ $Cys_{18}$, $Cys_{21}$) involved in the coordination of the radical SAM [4Fe-4S] cluster are colored in purple, while the 4 cysteine residues ($Cys_{206}$, $Cys_{222}$, $Cys_{289}$ and $Cys_{292}$), likely involved in the coordination of an additional [4Fe-4S] cluster, and the $Cys_{223}$ the enzyme H-atom donor, are labeled in pink.

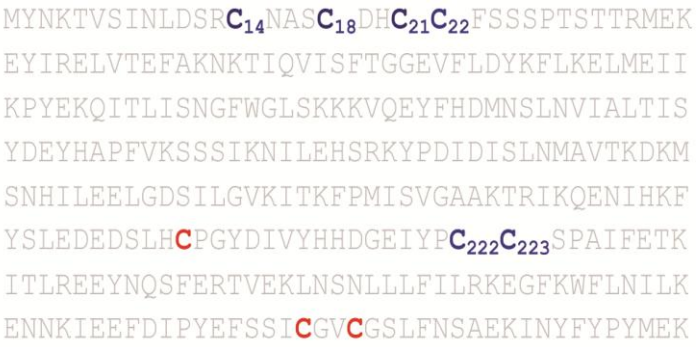
**a**



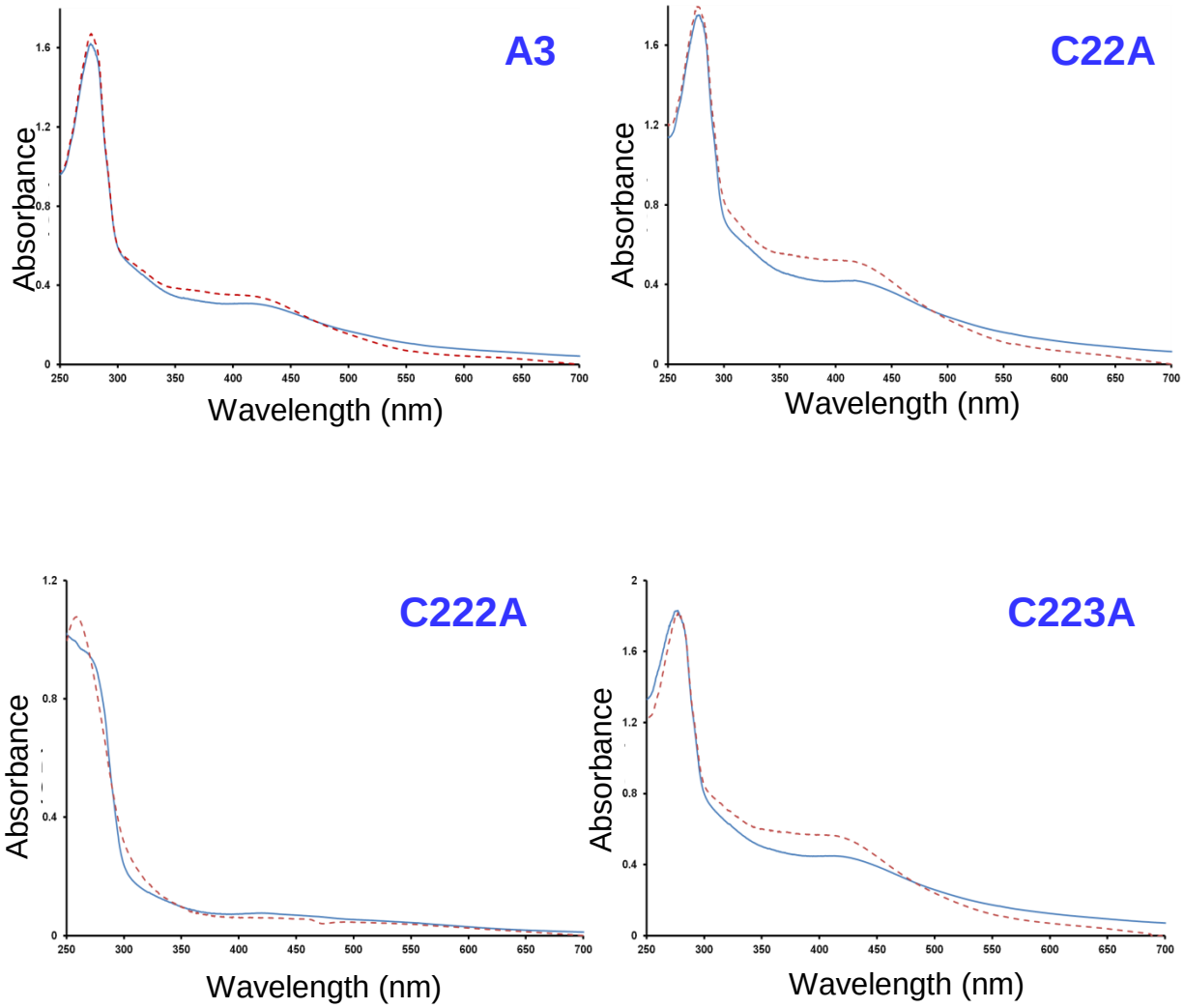

**b**



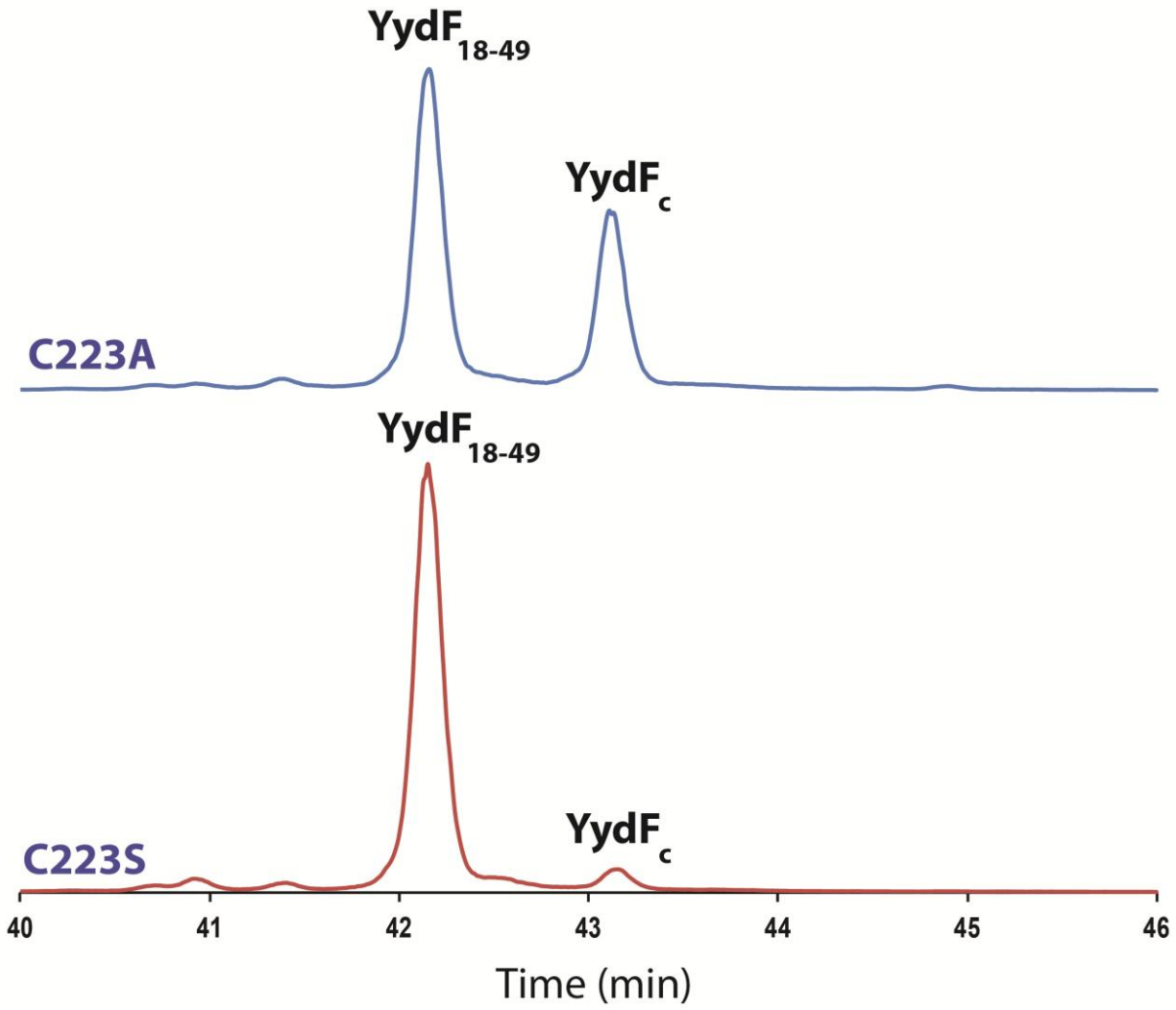

**Supplementary Figure 18 - (A) Sequence similarity network of YydG homologs generated using the EFI-EST server (http://enzymefunction.org/)[3].** Sequence homologs were retrieved using and E-value > 60 and a Score >90. *Bacillus* species (blue), *Enterococcus* species (red), *Streptococcus* species (yellow), *Staphylococcus* species (green) and other (gray) are indicated. **(B) Genome neighborhood network (EFI-GNT) built using the YydG sequence similarity network.**



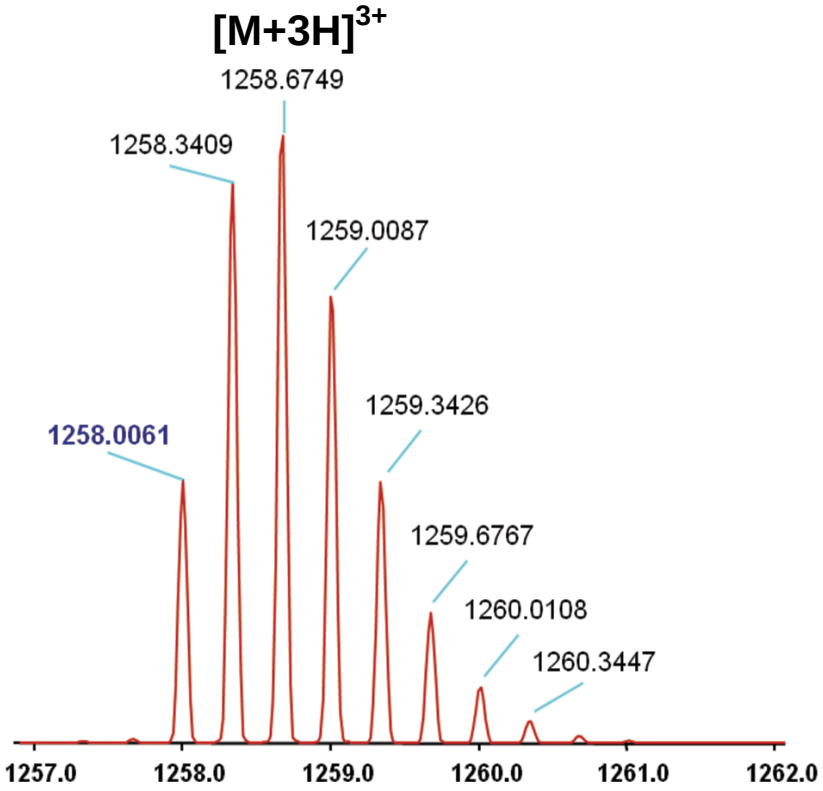

**Supplementary Figure 19 - Amino acid sequence of YydG with cysteine residues highlighted in blue or red.** Numbers indicate residues mutated in this study.




**Supplementary Figure 20 - UV-visible spectra of the A3 (*i.e.* AxxxAxxA), C22A, C222A and C223A mutants after aerobic purification (solid blue line) and anaerobic enzyme reconstitution (dashed red line).**

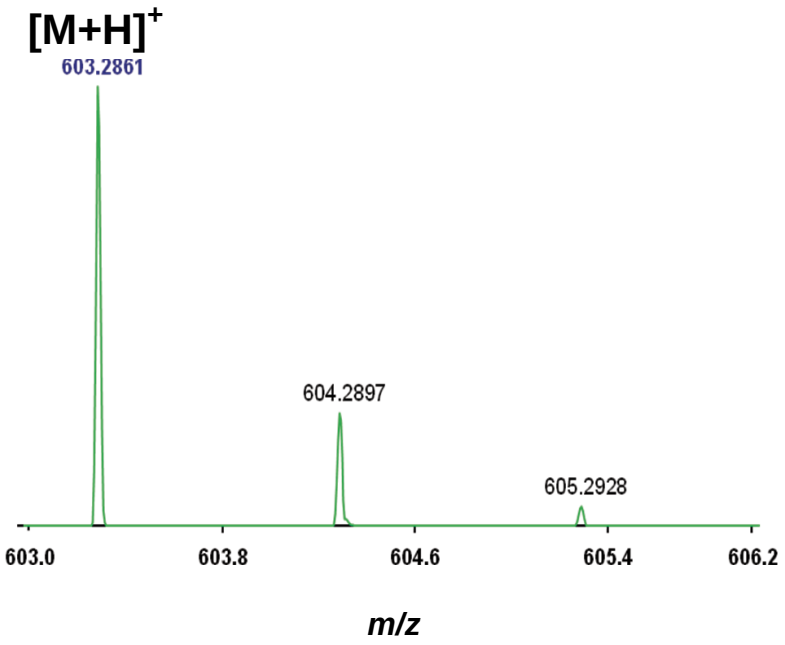



**Supplementary Figure 21 – HPLC analysis the $YydF_{18-49}$ peptide (1 mM) incubated with the C223A or C223S YydG mutants**. Each mutant (100µM) was incubated 90 minutes in the presence of SAM (1.5 mM), DTT (6 mM) and sodium dithionite (2 mM) with $YydF_{18-49}$ (1 mM). Peptides were analyzed using reverse phase C18 HPLC coupled with fluorescence detection (Ex/Em = 280/350 nm).

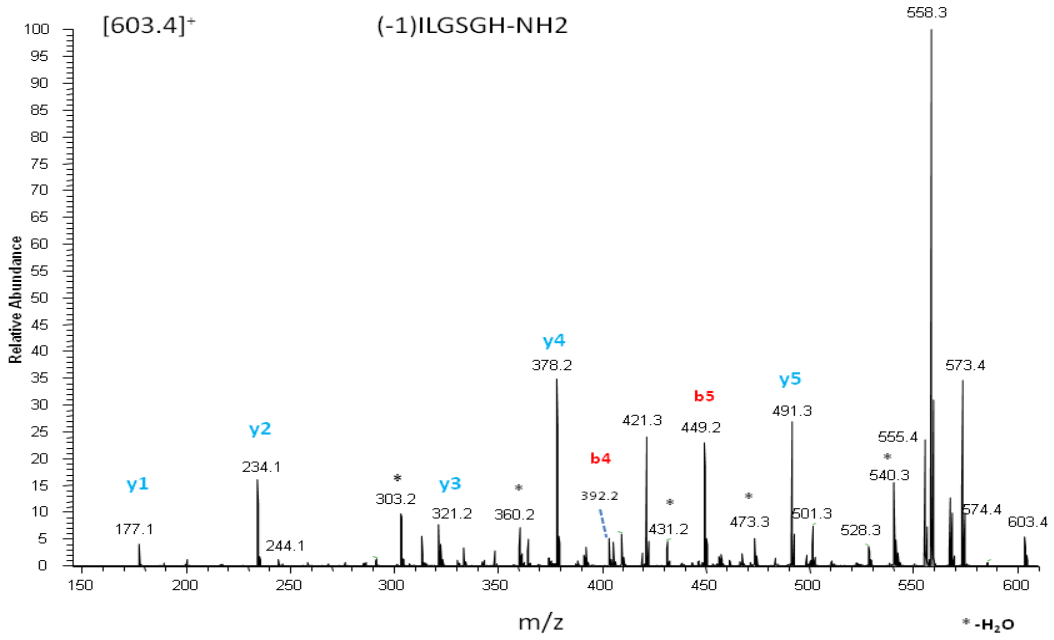



**Supplementary Figure 22 – High-resolution mass spectrometry analysis of the YydF$_{18-49}$ peptide**.

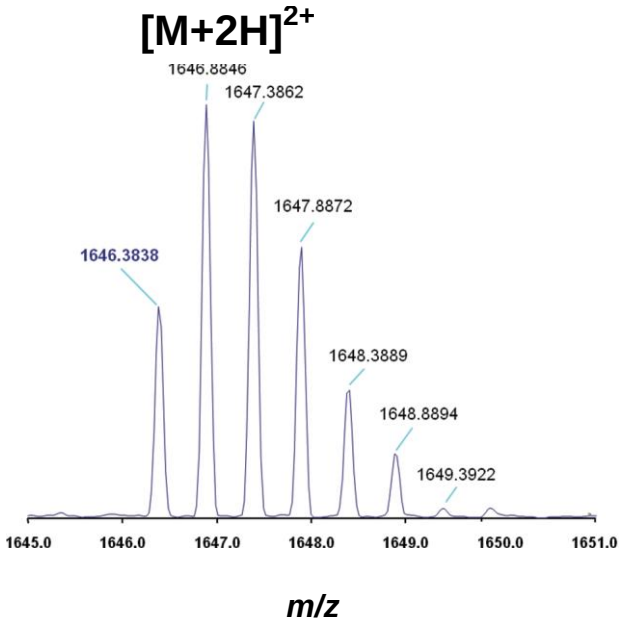
a



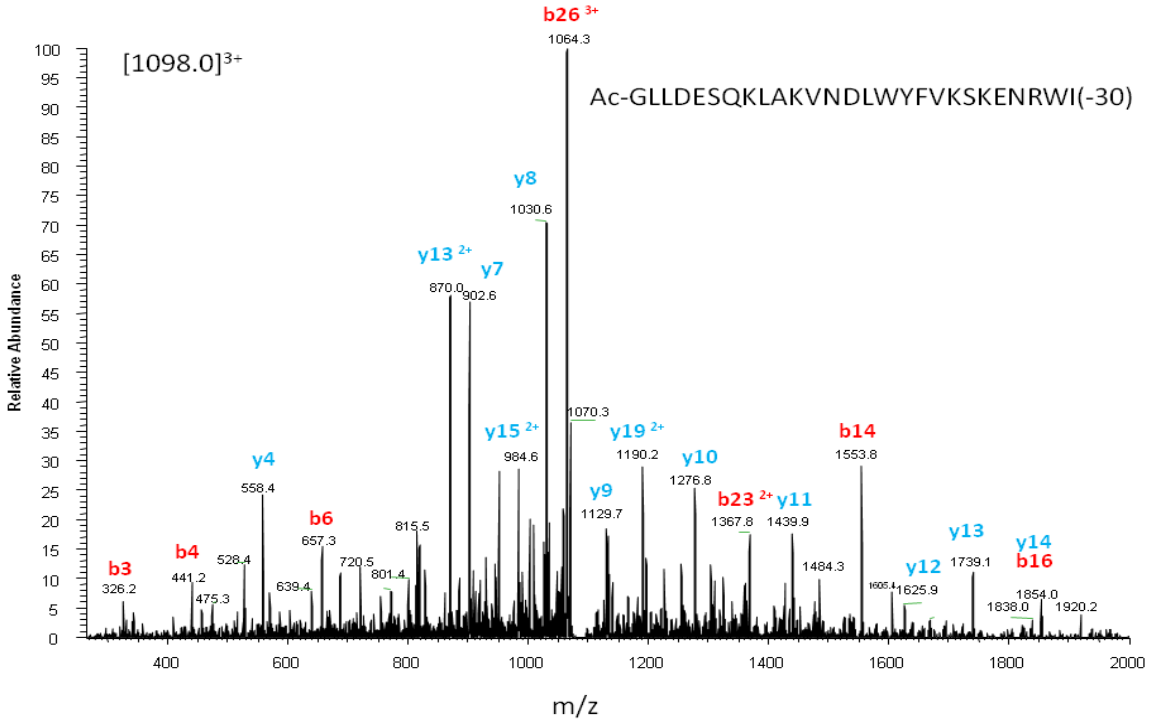

b



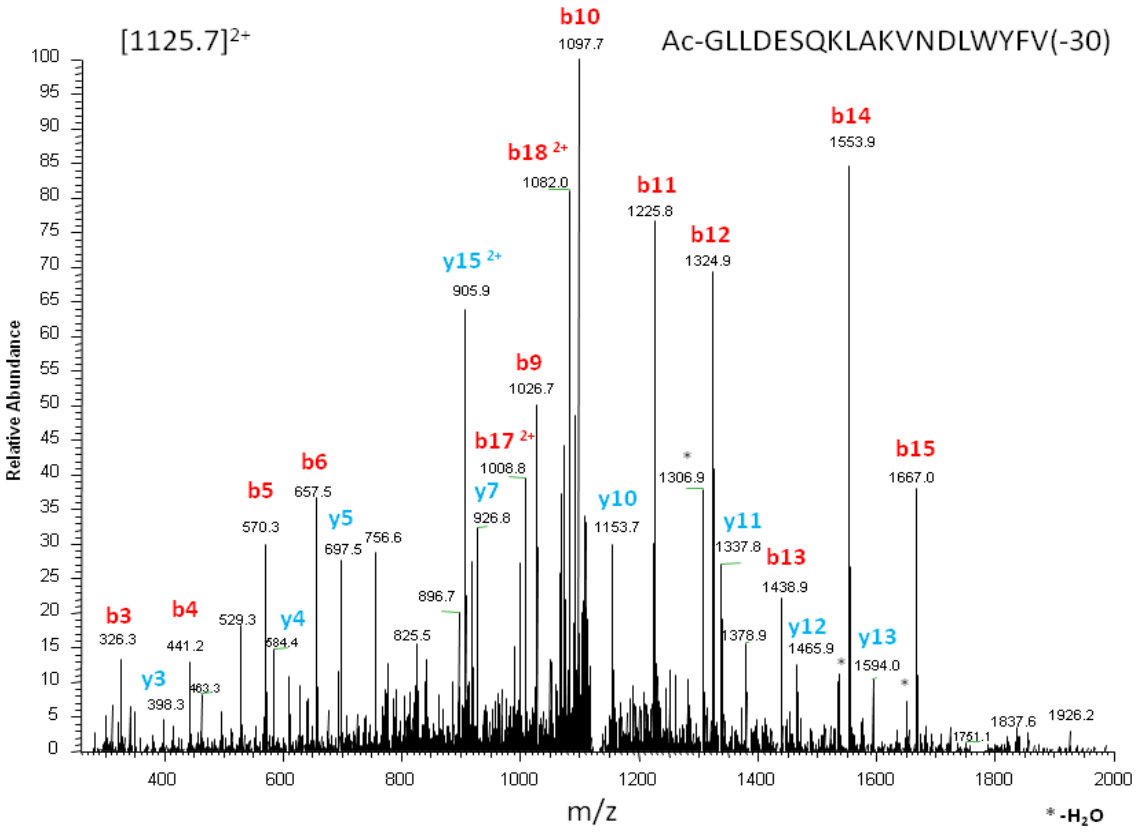

**Supplementary Figure 23 - LC-MS/MS analysis of the $I_{44}^{\Delta 1}$-$H_{49}$ peptide produced by the C223A mutant. (a) High-resolution mass spectrometry analysis of the $I_{44}^{\Delta 1}$-$H_{49}$ peptide and (b) mass fragmentation.** Relevant ions are indicated (*see Supplementary Tables 7-9 for full assignment*).



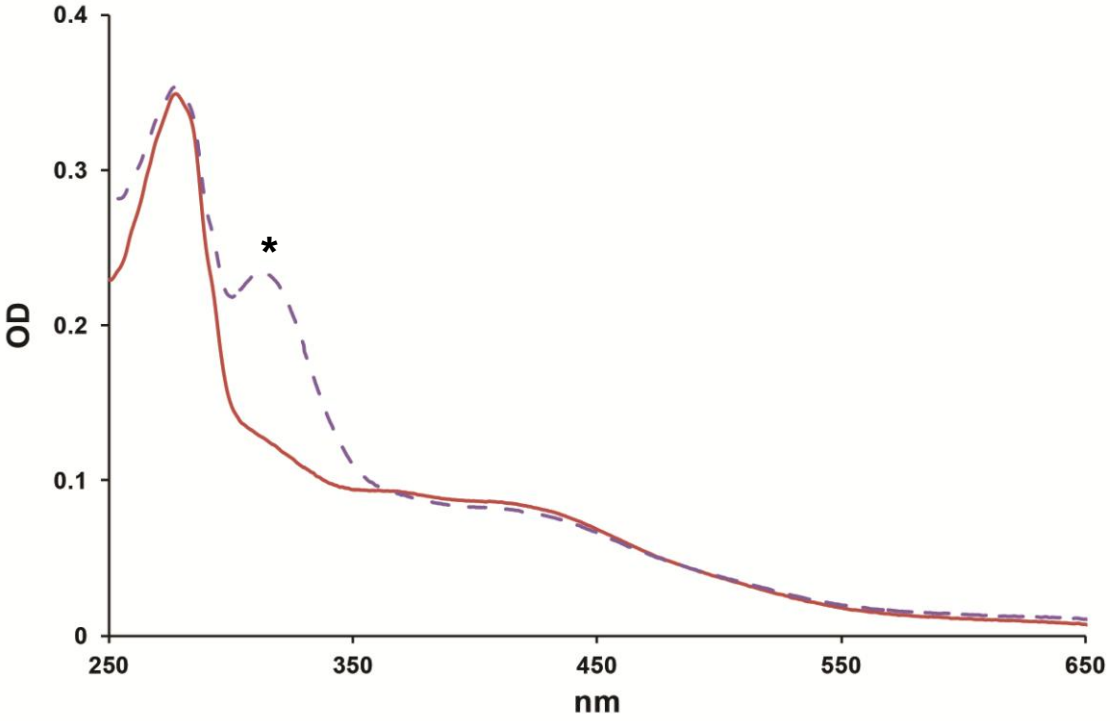

b



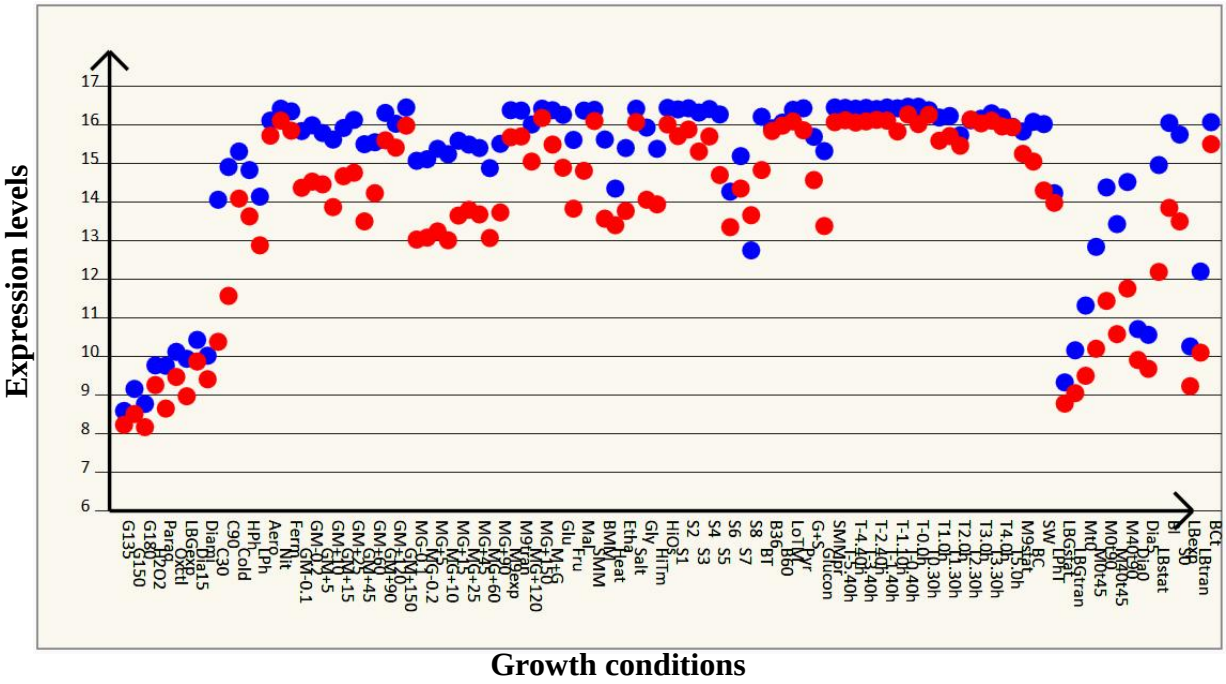

**Supplementary Figure 24 - LC-MS/MS analysis of the $G_{18}$-$I_{44}^{\Delta 30}$ peptide produced by the C223A mutant. (a) High resolution mass spectrometry analysis of the $G_{18}$-$I_{44}^{\Delta 30}$ peptide and (b) mass fragmentation.** Relevant ions are indicated (*see Supplementary Tables 7, 10 & 11 for full assignment*).



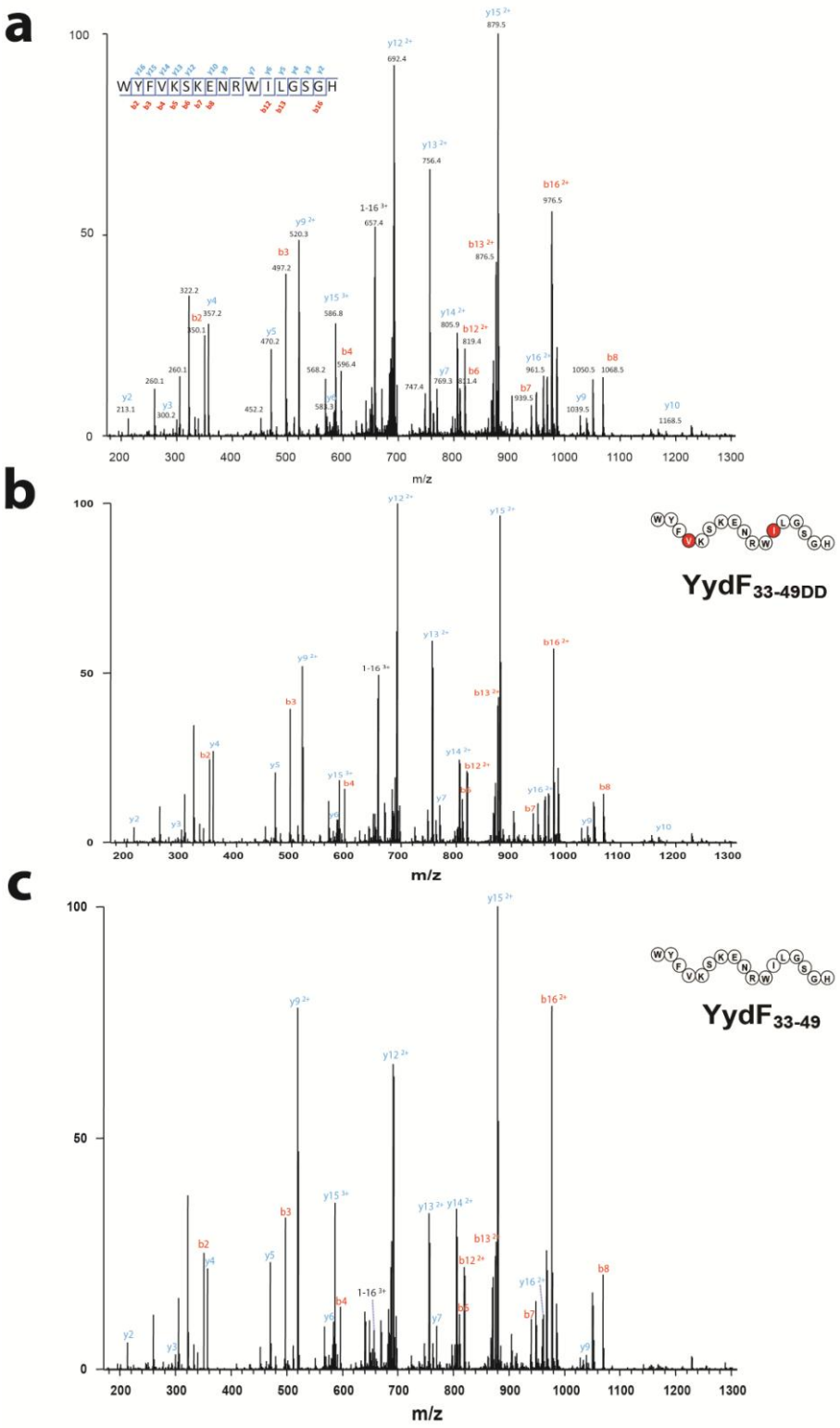

**Supplementary Figure 25 - LC-MS/MS analysis of the $G_{18}$-$V_{36}^{\Delta 30}$ peptide produced by the C223A mutant.** Relevant ions are indicated (*see Supplementary Tables 7 & 12 for full assignment*).



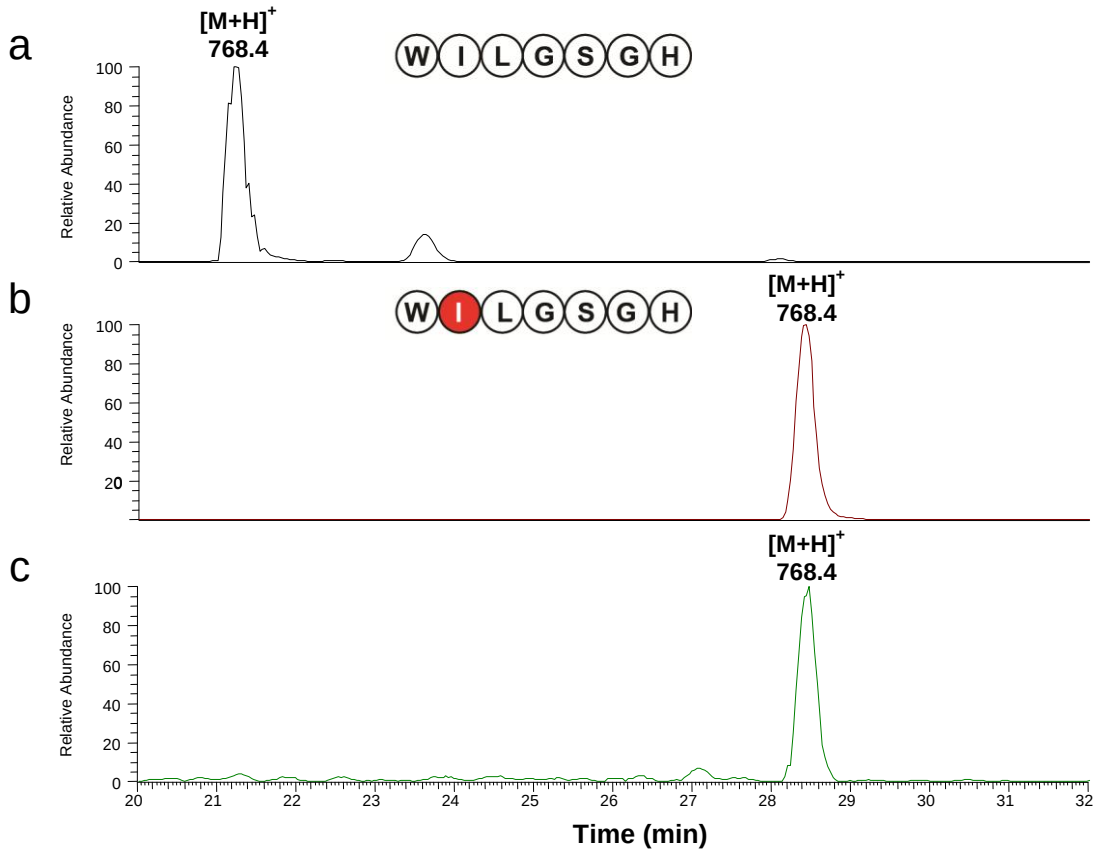

**Supplementary Figure 26 - UV-visible spectra of the A3 (*i.e.* AxxxAxxA) mutant (90 µM) after anaerobic reconstitution (solid line) and after incubation with sodium dithionite (2 mM) (dashed line) under anaerobic conditions.** The "*" symbol indicates reduced dithionite. Protein sample was measured in a 1 mm pathlength quartz microcuvette.




**Supplementary Figure 27 - Expression levels of YydF (red dots) and the killing factor SdpC (blue dots) by *B. subtilis* under 104 growth conditions[4].** The correlation coefficient between the two genes was determined to be 0.92, indicating the strongest matches between these two genes. Expression Data & Growth conditions: http://www.subtiwiki.uni-goettingen.de.




**Supplementary Figure 28 - LC-MS/MS analysis of (a) the peptide secreted by *B. subtilis*, (b) synthetic YydF$_{33\text{-}49DD}$ peptide containing two D-amino acids residues and (c) synthetic YydF$_{33\text{-}49}$ peptide.** Relevant ions are indicated. See *Supplementary Tables 13 for full assignment*. In the peptide sequence, amino acids with a D-configuration are indicated in red.




**Supplementary Figure 29 - LC-MS analysis of the tryptic peptides obtained from the synthetic peptide YydF$_{18-49}$ (a), the synthetic peptide YydF$_{18-49}$ containing a D-allo-Ile in position 44 (b) and the peptide isolated from *B. subtilis* (c).** The sequence of the tryptic peptide 3 is indicated and epimerized residues are labeled in red. *See Supplementary Fig. 9 for references.*

| Name | Sequence | z/adduct | Formula | m/z Theor. | *m/z* | m/z (HRMS) | Δm HRMS (ppm) |
|---|---|---|---|---|---|---|---|
| **YydF$_{18-49}$** | **Ac-GLLDESQKLAKVNDLWYFVKSKENRWILGSGH-NH$_2$** | 3/H | $C_{173}H_{270}N_{47}O_{48}$ | 1258.00438 | 1258.3 | 1258.0061 | **1.4** |
| **Peptide 1** | **Ac-GLLDESQK** | 1/H | $C_{39}H_{67}N_{10}O_{16}$ | 931.47365 | 931.7 | | |
| **Peptide 2** | **VNDLWYFVK** | 2/H | $C_{59}H_{84}N_{12}O_{14}$ | 592.31150 | 592.5 | | |
| **Peptide 3** | **WILGSGH-NH$_2$** | 1/H | $C_{36}H_{54}N_{11}O_8$ | 768.41568 | 768.6 | | |

**Supplementary Table 1: Sequences and molecular weights of the YydF$_{18-49}$ peptide and its major tryptic peptides.** Calculated molecular weights (*m/z* Theor) are associated to their experimentally measured masses by an LTQ-mass spectrometer (*m/z*) or a high resolution orbitrap mass analyzer (*m/z* HRMS). Z/adduct: Peptide charge state/adduct (H or Na) as observed. Δm are calculated in ppm and measured by high resolution spectroscopy. Fragments ions are described in the Supplementary Table 2-5. MS/MS spectra are reported in the Supplementary Figures 6 & 7.

## Supplementary Table 2 - Mass fragments of YydF18-49 peptide

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **G** | 1 | 100.03989 | 3771.99751 | 32 |
| **L** | 2 | 213.12396 | 3672.96549 | 31 |
| **L** | 3 | 326.20802 | 3559.88143 | 30 |
| **D** | 4 | 441.23496 | 3446.79736 | 29 |
| **E** | 5 | 570.27756 | 3331.77042 | 28 |
| **S** | 6 | 657.30958 | 3202.72783 | 27 |
| **Q** | 7 | 785.36816 | 3115.6958 | 26 |
| **K** | 8 | 913.46312 | 2987.63722 | 25 |
| **L** | 9 | 1026.54719 | 2859.54226 | 24 |
| **A** | 10 | 1097.5843 | 2746.4582 | 23 |
| **K** | 11 | 1225.67926 | 2675.42108 | 22 |
| **V** | 12 | 1324.74768 | 2547.32612 | 21 |
| **N** | 13 | 1438.7906 | 2448.25771 | 20 |
| **D** | 14 | 1553.81755 | 2334.21478 | 19 |
| **L** | 15 | 1666.90161 | 2219.18784 | 18 |
| **W** | 16 | 1852.98092 | 2106.10377 | 17 |
| **Y** | 17 | 2016.04425 | 1920.02446 | 16 |
| **F** | 18 | 2163.11267 | 1756.96113 | 15 |
| **V** | 19 | 2262.18108 | 1609.89272 | 14 |
| **K** | 20 | 2390.27604 | 1510.82431 | 13 |
| **S** | 21 | 2477.30807 | 1382.72934 | 12 |
| **K** | 22 | 2605.40303 | 1295.69731 | 11 |
| **E** | 23 | 2734.44563 | 1167.60235 | 10 |
| **N** | 24 | 2848.48855 | 1038.55976 | 9 |
| **R** | 25 | 3004.58966 | 924.51683 | 8 |
| **W** | 26 | 3190.66898 | 768.41572 | 7 |
| **I** | 27 | 3303.75304 | 582.33641 | 6 |
| **L** | 28 | 3416.8371 | 469.25235 | 5 |
| **G** | 29 | 3473.85857 | 356.16828 | 4 |
| **S** | 30 | 3560.89059 | 299.14682 | 3 |
| **G** | 31 | 3617.91206 | 212.11479 | 2 |
| **H** | 32 | 3753.98695 | 155.09333 | 1 |

## Supplementary Table 3 - Mass fragments of Peptide 1

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **G** | 1 | 100.03989 | 931.47369 | 8 |
| **L** | 2 | 213.12396 | 832.44167 | 7 |
| **L** | 3 | 326.20802 | 719.35760 | 6 |
| **D** | 4 | 441.23496 | 606.27354 | 5 |
| **E** | 5 | 570.27756 | 491.24660 | 4 |
| **S** | 6 | 657.30958 | 362.20400 | 3 |
| **Q** | 7 | 785.36816 | 275.17197 | 2 |
| **K** | 8 | 913.46312 | 147.11340 | 1 |

## Supplementary Table 4 - Mass fragments Peptide 2

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **V** | 1 | 100.07628 | 1183.61521 | 9 |
| **N** | 2 | 214.11921 | 1084.5468 | 8 |
| **D** | 3 | 329.14615 | 970.50387 | 7 |
| **L** | 4 | 442.23022 | 855.47693 | 6 |
| **W** | 5 | 628.30953 | 742.39287 | 5 |
| **Y** | 6 | 791.37286 | 556.31355 | 4 |
| **F** | 7 | 938.44127 | 393.25022 | 3 |
| **V** | 8 | 1037.50969 | 246.18181 | 2 |
| **K** | 9 | 1165.60465 | 147.1134 | 1 |

## Supplementary Table 5 - Mass fragments of Peptide 3

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **W** | 1 | 187.08718 | 768.41572 | 7 |
| **I** | 2 | 300.17125 | 582.33641 | 6 |
| **L** | 3 | 413.25531 | 469.25235 | 5 |
| **G** | 4 | 470.27677 | 356.16828 | 4 |
| **S** | 5 | 557.3088 | 299.14682 | 3 |
| **G** | 6 | 614.33027 | 212.11479 | 2 |
| **H** | 7 | 750.40516 | 155.09333 | 1 |

| | Species | Number of Sequences |
|---|---|---|
| **Human microbiome** | *Enterococcus faecalis* | 54 |
| | *Streptococcus agalactiae* | 32 |
| | *Corynebacterium diphtheriae* | 10 |
| | *Enterococcus caccae* | 3 |
| | *Staphylococcus pseudintermedius* | 2 |
| | *Staphylococcus epidermidis* | 2 |
| | | |
| **Animal Microbiome** | *Staphylococcus equorum* | 2 |
| | | |
| **Pathogens** | *Bacillus cereus* | 3 |
| | | |
| **Environment** | *Bacillus subtilis* | 42 |
| | *Bacillus coagulans* | 4 |
| | *Clostridium sp. BNL1100* | 4 |
| | *Bacillus sp. YP1* | 2 |
| | *Bacillus* | 57 |

**Supplementary Table 6 – YydF homologs listed in the NCBI database (E-value >60).** The number of sequences indicates the number of individual sequence found in the respective genomes.

| Name | Sequences | z/Adduct | m/z Formula | m/zTheor. | m/z | m/z (HRMS) | Δm HRMS (ppm) |
|---|---|---|---|---|---|---|---|
| **YydF18-49** | **Ac-GLLDESQKLAKVNDLWYFVKSKENRWILGSGH-NH2** | 3/H | $C_{173}H_{270}N_{47}O_{48}$ | **1258.00438** | 1258.2 | **1258.0061** | **1.4** |
| $G_{18}$-$V_{36}$ | **Ac-GLLDESQKLAKVNDLWYFV** | 2/H | $C_{107}H_{164}N_{24}O_{31}$ | **1140.59972** | | | |
| $G_{18}$-$V_{36}^{\Delta 30}$ | **Ac-GLLDESQKLAKVNDLWYFV* (-COH2)** | 2/H | $C_{106}H_{162}N_{24}O_{30}$ | **1125.59443** | **1125.7** | | |
| $G_{18}$-$I_{44}$ | **Ac-GLLDESQKLAKVNDLWYFVKSKENRWI** | 2/H | $C_{154}H_{239}N_{39}O_{43}$ | **1661.38570** | | | |
| $G_{18}$-$I_{44}^{\Delta 30}$ | **Ac-GLLDESQKLAKVNDLWYFVKSKENRWI* (-COH2)** | **2/H** | $\mathbf{C_{153}H_{237}N_{39}O_{42}}$ | **1646.38042** | **1646.5** | **1646.3838** | **2.1** |
| | Ac-GLLDESQKLAKVNDLWYFVKSKENRWI* (-CNH4) | 2/H | $C_{153}H_{235}N_{38}O_{43}$ | 1646.36851 | | 1646.3838 | 9.3 |
| | Ac-GLLDESQKLAKVNDLWYFVKSKENRWI* (-C2H6) | 2/H | $C_{152}H_{233}N_{39}O_{43}$ | 1646.36222 | | 1646.3838 | 13.1 |
| $I_{44}$-$H_{49}$ | **ILGSGH-NH2** | 1/H | $C_{25}H_{44}N_9O_7$ | **582.33637** | | | |
| $I_{44}^{\Delta 1}$-$H_{49}$ | ILGSGH-NH2 | 1/Na | $C_{25}H_{43}N_9NaO_7$ | 604.31831 | | | |
| | ***ILGSGH-NH2 (-NH3+O)** | **1/Na** | $\mathbf{C_{25}H_{40}N_8NaO_8}$ | **603.28668** | **603.4** | **603.2861** | **-1.0** |
| | *ILGSGH-NH2 (-H) | 1/Na | $C_{25}H_{42}N_9NaO_7$ | 603.31049 | | 603.2861 | -40.4 |
| | *ILGSGH-NH2 (-OH+NH2) | 1/Na | $C_{25}H_{44}N_{10}NaO_6$ | 603.33430 | | 603.2861 | -79.9 |

**Supplementary Table 7 - Sequences and molecular weights of the $YydF_{18-49}$ peptide and the radically cleaved peptides produced by the C223A mutant.** Calculated molecular weights (*m/z* Theor) are associated to their experimentally measured masses by an LTQ-mass spectrometer (*m/z*) or a high resolution orbitrap mass analyzer (*m/z* HRMS). Z/adduct: Peptide charge state/ adduct (H or Na) as observed. Δm are calculated in ppm and measured by high resolution spectroscopy. Fragments ions are described in the Supplementary Table 8-12. MS/MS spectra are reported in the Supplementary Figures 22-25.

## Supplementary Table 8 - Mass fragments for peptide $I_{44}$-$H_{49}$

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **I** | 1 | 114.09193 | 582.33641 | 6 |
| **L** | 2 | 227.176 | 469.25235 | 5 |
| **G** | 3 | 284.19746 | 356.16828 | 4 |
| **S** | 4 | 371.22949 | 299.14682 | 3 |
| **G** | 5 | 428.25095 | 212.11479 | 2 |
| **H** | 6 | 564.32584 | 155.09333 | 1 |

## Supplementary Table 9 - Mass fragments for peptide $I_{44}^{\Delta 1}$-$H_{49}$

| Amino Acids | b ions | Mass of b ions $[M+Na]^+$ | Mass of y ions $[M+Na]^+$ | y ions |
|---|---|---|---|---|
| **I** | 1 | 135.042275 | 604.318355 | 6 |
| **L** | 2 | 248.126345 | 491.234295 | 5 |
| **G** | 3 | 305.147805 | 378.150225 | 4 |
| **S** | 4 | 392.179835 | 321.128765 | 3 |
| **G** | 5 | 449.201295 | 234.096735 | 2 |
| **H** | 6 | 585.276185 | 177.075275 | 1 |

## Supplementary Table 10 - Mass fragments for peptide $G_{18}$-$I_{44}$

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **G** | 1 | 100.03989 | 3321.7636 | 27 |
| **L** | 2 | 213.12396 | 3222.73158 | 26 |
| **L** | 3 | 326.20802 | 3109.64752 | 25 |
| **D** | 4 | 441.23496 | 2996.56345 | 24 |
| **E** | 5 | 570.27756 | 2881.53651 | 23 |
| **S** | 6 | 657.30958 | 2752.49392 | 22 |
| **Q** | 7 | 785.36816 | 2665.46189 | 21 |
| **K** | 8 | 913.46312 | 2537.40331 | 20 |
| **L** | 9 | 1026.54719 | 2409.30835 | 19 |
| **A** | 10 | 1097.5843 | 2296.22429 | 18 |
| **K** | 11 | 1225.67926 | 2225.18717 | 17 |
| **V** | 12 | 1324.74768 | 2097.09221 | 16 |
| **N** | 13 | 1438.7906 | 1998.0238 | 15 |
| **D** | 14 | 1553.81755 | 1883.98087 | 14 |
| **L** | 15 | 1666.90161 | 1768.95393 | 13 |
| **W** | 16 | 1852.98092 | 1655.86986 | 12 |
| **Y** | 17 | 2016.04425 | 1469.79055 | 11 |
| **F** | 18 | 2163.11267 | 1306.72722 | 10 |
| **V** | 19 | 2262.18108 | 1159.65881 | 9 |
| **K** | 20 | 2390.27604 | 1060.59039 | 8 |
| **S** | 21 | 2477.30807 | 932.49543 | 7 |
| **K** | 22 | 2605.40303 | 845.4634 | 6 |
| **E** | 23 | 2734.44563 | 717.36844 | 5 |
| **N** | 24 | 2848.48855 | 588.32585 | 4 |
| **R** | 25 | 3004.58966 | 474.28292 | 3 |
| **W** | 26 | 3190.66898 | 318.18181 | 2 |
| **I** | 27 | 3303.75304 | 132.1025 | 1 |

## Supplementary Table 11 - Mass fragments for peptide $G_{18}$-$I_{44}^{\Delta 30}$

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **G** | 1 | 100.03989 | 3291.7601 | 27 |
| **L** | 2 | 213.12396 | 3192.72808 | 26 |
| **L** | 3 | 326.20802 | 3079.64402 | 25 |
| **D** | 4 | 441.23496 | 2966.55995 | 24 |
| **E** | 5 | 570.27756 | 2851.53301 | 23 |
| **S** | 6 | 657.30958 | 2722.49042 | 22 |
| **Q** | 7 | 785.36816 | 2635.45839 | 21 |
| **K** | 8 | 913.46312 | 2507.39981 | 20 |
| **L** | 9 | 1026.54719 | 2379.30485 | 19 |
| **A** | 10 | 1097.5843 | 2266.22079 | 18 |
| **K** | 11 | 1225.67926 | 2195.18367 | 17 |
| **V** | 12 | 1324.74768 | 2067.08871 | 16 |
| **N** | 13 | 1438.7906 | 1968.0203 | 15 |
| **D** | 14 | 1553.81755 | 1853.97737 | 14 |
| **L** | 15 | 1666.90161 | 1738.95043 | 13 |
| **W** | 16 | 1852.98092 | 1625.86636 | 12 |
| **Y** | 17 | 2016.04425 | 1439.78705 | 11 |
| **F** | 18 | 2163.11267 | 1276.72372 | 10 |
| **V** | 19 | 2262.18108 | 1129.65531 | 9 |
| **K** | 20 | 2390.27604 | 1030.58689 | 8 |
| **S** | 21 | 2477.30807 | 902.49193 | 7 |
| **K** | 22 | 2605.40303 | 815.4599 | 6 |
| **E** | 23 | 2734.44563 | 687.36494 | 5 |
| **N** | 24 | 2848.48855 | 558.32235 | 4 |
| **R** | 25 | 3004.58966 | 444.27942 | 3 |
| **W** | 26 | 3190.66898 | 288.17831 | 2 |
| **I** | 27 | 3273.74954 | 102.099 | 1 |

## Supplementary Table 12 - Mass fragments for peptide $G_{18}$-$V_{36}^{\Delta 30}$

| Amino Acids | b ions | Mass of b ions $[M+H]^+$ | Mass of y ions $[M+H]^+$ | y ions |
|---|---|---|---|---|
| **G** | 1 | 100.03989 | 2250.18108 | 19 |
| **L** | 2 | 213.12396 | 2151.14906 | 18 |
| **L** | 3 | 326.20802 | 2038.065 | 17 |
| **D** | 4 | 441.23496 | 1924.98093 | 16 |
| **E** | 5 | 570.27756 | 1809.95399 | 15 |
| **S** | 6 | 657.30958 | 1680.9114 | 14 |
| **Q** | 7 | 785.36816 | 1593.87937 | 13 |
| **K** | 8 | 913.46312 | 1465.82079 | 12 |
| **L** | 9 | 1026.54719 | 1337.72583 | 11 |
| **A** | 10 | 1097.5843 | 1224.64177 | 10 |
| **K** | 11 | 1225.67926 | 1153.60465 | 9 |
| **V** | 12 | 1324.74768 | 1025.50969 | 8 |
| **N** | 13 | 1438.7906 | 926.44128 | 7 |
| **D** | 14 | 1553.81755 | 812.39835 | 6 |
| **L** | 15 | 1666.90161 | 697.37141 | 5 |
| **W** | 16 | 1852.98092 | 584.28734 | 4 |
| **Y** | 17 | 2016.04425 | 398.20803 | 3 |
| **F** | 18 | 2163.11267 | 235.1447 | 2 |
| **V** | 19 | 2232.17052 | 88.07629 | 1 |

**Supplementary Table 13 - Mass fragments for peptide YydF$_{33-49}$ isolated from *B. subtilis***

| Sequence | | b+ | b++ | y+ | y++ | |
|---|---|---|---|---|---|---|
| W | 1 | 187.08718 | 94.04753 | 2107.08779 | 1054.04783 | 17 |
| Y | 2 | 350.15051 | 175.57919 | 1921.00848 | 961.00818 | 16 |
| F | 3 | 497.21892 | 249.11340 | 1757.94515 | 879.47651 | 15 |
| V | 4 | 596.28734 | 298.64760 | 1610.87674 | 805.94230 | 14 |
| K | 5 | 724.38230 | 362.69509 | 1511.80833 | 756.40810 | 13 |
| S | 6 | 811.41433 | 406.21110 | 1383.71336 | 692.36062 | 12 |
| K | 7 | 939.50929 | 470.25858 | 1296.68133 | 648.84460 | 11 |
| E | 8 | 1068.55188 | 534.77988 | 1168.58637 | 584.79712 | 10 |
| N | 9 | 1182.59481 | 591.80134 | 1039.54378 | 520.27582 | 9 |
| R | 10 | 1338.69592 | 669.85190 | 925.50085 | 463.25436 | 8 |
| W | 11 | 1524.77523 | 762.89155 | 769.39974 | 385.20381 | 7 |
| I | 12 | 1637.85930 | 819.43358 | 583.32043 | 292.16415 | 6 |
| L | 13 | 1750.94336 | 875.97562 | 470.23637 | 235.62212 | 5 |
| G | 14 | 1807.96483 | 904.48635 | 357.15230 | 179.08009 | 4 |
| S | 15 | 1894.99685 | 948.00236 | 300.13084 | 150.56935 | 3 |
| G | 16 | 1952.01832 | 976.51309 | 213.09881 | 107.05334 | 2 |
| H | 17 | 2089.07723 | 1045.04255 | 156.07735 | 78.54261 | 1 |

| | |
|---|---|
| (M) | 2106.07997 |
| $(M+H)^{+}$ | 2107.08779 |
| $(M+2H)^{2+}$ | 1054.04783 |
| $(M+3H)^{3+}$ | **703.03451** |
| $(M+4H)^{4+}$ | 527.52785 |

**MS/MS spectra are reported in the Supplementary Figure 28.**

# Supplementary methods

## Cloning, expression and purification of the YydG protein

The *yydG* gene (Gene ID: 937720) from *Bacillus subtilis* str. 168 was optimized for *E. coli* expression and synthesized by Life Technologies. The synthetized *yydG* gene was then cloned into a pET28 plasmid (Novagen) as Strep-tag fusion. The expression of the YydG was performed in *E. coli* BL21 (DE3) star (Life Technologies). An overnight culture of *E. coli* BL21 (DE3)/pET28-Strep-tag-yydG was used to inoculate 9 liters of LB medium containing 50 $\mu g.mL^{-1}$ kanamycine (Sigma-aldrich). The culture was grown at 37°C until the OD at 600 nm reached ~0.6. The protein expression was then induced upon addition of 200 µg anhydrotetracycline (Sigma-Aldrich) per 1 liter followed by overnight incubation at 21°C. The cells were harvested and disrupted by ultra-sonication in buffer A (Tris 50 mM, KCl 300 mM, pH 7.5) supplemented with protease inhibitor cocktail (Complete, EDTA-free, Roche). The bacterial suspension was centrifuged at 45,000 x g for 1.5 hours and the protein supernatant was loaded onto a Streptactin high-capacity (IBA GmbH) gel column previously equilibrated with buffer A. The YydG protein was eluted with 6 mL of buffer A containing desthiobiotine (0.6 $mg.mL^{-1}$, Sigma-Aldrich) and further concentrated with Amicon concentrator (molecular cut-off of 10 kDa, Millipore). The protein purity was checked by SDS–polyacrylamide gel electrophoresis.

## Production of the YydG mutant proteins

The yydG gene with the appropriate mutations (C14A/C18A/C21A; C22A; C222A or C223A) were optimized for expression in *E. coli* and synthesized by Life Technologies. The C223S mutant was obtained by site-directed mutagenesis with the plasmid pET28-Strep-tag-yydG used as DNA template. The primer sequences used to introduce a serine residue at

position 223 were: 5′-GAAATCTATCCGTGTTCTAGTCCGGCAATTTTCG-3′ and 5′-CGAAAATTGCCGGACTAGAACACGGATAGATTTCAC-3′. The PCR reaction containing 1µM of each primer, 250 µM of each dNTP, 1 µl of PfuUltra II Fusion HS DNA polymerase (Agilent), ~20 ng DNA template was performed using the following PCR cycling parameters: 1 cycle at 95°C for 2 min followed by 30 cycles of 94°C for 30 sec (denaturation), 55°C for 30 sec (hybridation), 72°C for 2 min (extension) and 1 cycle at 72°C for 10 min. Twenty units of DpnI was added to the PCR product to digest the DNA template. Afetr 1 hour incubation at 37°C, this digestion product was used to transform a chemically competent *E. coli* BL21 (DE3) star strain. Clones were selected on LB agar plate containing 50 $\mu g.mL^{-1}$kanamycine. One clone was selected after we checked the sequence of the targeted gene by DNA sequencing. The same procedures (described above) were used for the cloning in pET28 as Strep-tag fusion, expression and purification of all mutant proteins.

**Enzyme reconstitution**

YydG was reconstituted under anaerobic conditions in a Bactron IV anaerobic chamber. The reconstitution included an addition of 3 mM of DTT (Sigma-Aldrich) to the protein (incubation for 15 min at 12°C) followed by an incubation for 4h at 12°C of the protein with an 8-fold molar excess of both $Na_2S$ (Sigma-Aldrich) and $(NH_4)_2Fe(SO_4)_2$ (Sigma-Aldrich).

The protein was desalted using a Sephadex G25 column (GE Healthcare) equilibrated with buffer A containing 3 mM of DTT and the protein fraction was concentrated using Amicon Ultra-4 (Millipore). Protein concentration was determined by Nanodrop ($\varepsilon_{280\ nm}$ = 46300 $M^{-1}.cm^{-1}$). The iron–sulfur clusters of all mutant proteins were reconstituted using the same procedure.

**Enzyme assay**

YydG (100 µM) was incubated with 6 mM DTT, 1.5 mM SAM and 1 mM peptide substrate unless otherwise indicated. The reactions were triggered upon addition of 2 mM of sodium dithionite (Sigma-Aldrich). Incubations were performed at 25°C under strict anaerobic conditions and 10µL aliquots sampled overtime for analysis.

**Iron-sulfur content determination**

**Iron assay:** Samples of YydG protein at 84 µM (25, 75 or 100 µL with samples 25 and 75 µl made up to a volume of 100 µl in $H_2O$) are mixed with 100 µL of 1 % HCl. These samples are incubated at 80°C for 10 min and then cooled down in ice. We added sequentially to each sample: 500 µl of 7.5% (w/v) ammonium acetate, 100 µl of 4% ascorbic acid, 100 µl of 2.5% SDS, 100 µl of 1.5% ferene. Samples are centrifuged at high speed (14000 rpm) for 10 min. The supernant of each sample was transferred to a 1 mL cuvette and the absorbance was measured at 593 nm. A calibration curve was made using the same protocol and Mohr's salt solutions as standard. Based on this calibration curve, we determined the iron content in the YydG protein and mutants.

**Sulfur assay:** Samples of YydG protein (10, 20, 50 or 100 µL at 84 µM), are made up to a volume of 200 µL in $H_2O$. We added sequentially to each sample 600 µL of 1% zinc acetate, 50 µL of 7% NaOH. Samples were incubated for 15 min at room temperature. We then added to each sample 150 µL of 0.1% DMPD and 150 µL of 10 mM $FeCl_3$ solution. Samples have been vortexed vigorously for 30 sec and then centrifuged at high speed (14 000 rpm) for 10 min. The supernant of each sample was transferred to a 1 mL cuvette and after 20 min, the absorbance of methylene blue was measured at 670 nm. A calibration curve was made using the same protocol and $Na_2S$ as standard solutions (for more accuracy, concentrations of the

prepared solutions were calculated with the $\varepsilon_{670\,nm}$ = 26.7 $mM^{-1}cm^{-1}$). Based on this calibration curve, we determined the $S^{2-}$ content of the YydG protein and mutants.

**HPLC analysis**

HPLC analysis was performed on an Agilent 1200 series infinity equipped with a reversed phase column (LiChroCART RP-18e 5 µm) (Merck Millipore). A gradient from solvent A ($H_2O$, 0.1% TFA) to B (80 % $CH_3CN$, 19.9 % $H_2O$, 0.1% TFA) was applied as follow: 0-1 min: 100% A/0% B; 1-24 min: a linear gradient with 2% of solvent B per minute at a flow rate of 1 $mL.min^{-1}$. 24-64 min: a linear gradient with 1% of solvent B per minute at a flow rate of 1 $mL.min^{-1}$. Detection was performed at 257 & 278 nm with a diode array detector and by fluorescence (ex/em: 278/350 nm).

**UV-visible analysis**

UV-visible analysis were conducted anaerobically with a JASCO v700 connected with an optical fiber.

**Peptides purification on SPE C18 cartridge**

Peptides $YydF_{18\text{-}49(a,b,c)}$ were purified by SPE C18 cartridge (Sep-PaK C18, 100mg, Waters) and eluted by 1 mL of 80% acetonitrile, 0.1%TFA, and dried centrifugal vacuum concentrator for 2 hours.

**Enantiomeric analysis of amino-acids**

Purified peptides were subjected to vapor-phase acid hydrolysis at 110°C for 18h. Samples were then dried and derivatized by addition of 10 excess of *N*α-(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA, suspended in 1% acetone). Samples were heated at 42°C for 1 hour. The reaction was stopped by addition of of HCl 1N and the mixture diluted 1/10 in 20% acetonitrile, 0.1% formic acid before analysis.

**Peptide hydrolysis**

Trypsin hydrolysis was performed at 37°C for 3 hours by addition of 200 ng of trypsin (200 $ng.\mu L^{-1}$ (Promega) was suspended in 50mM acetic acid and diluted 1/10 in ammonium bicarbonate 50 mM just before use). The hydrolysis was stopped by addition of formic acid (1% final concentration). The sample was dilute 1/100 before analysis by nanoLC-MS/MS.

**Liquid chromatography–mass spectrometry/mass spectrometry analysis**

LC-MS/MS analysis was performed using an LTQ standard mass spectrometer (Thermo) with a nanoelectrospray ion source and an Ultimate 3000 LC system (Dionex) for routine m/z mass analysis. An LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) with a nano-electrospray ion source associated to an Ultimate nano-RSLC system (Dionex, Thermo Fisher Scientific) was used for high accuracy mass analysis. Molecule separation was performed on a nano-column Pepmap100 C18 (0.075 by 15 cm, 100 Å, 3 µm) by direct injection of 1µL of dilute sample (1/100 or 1/10). Gradient was as followed: 20 to 50% buffer B for $YydF_{18\text{-}49\ (a,\ b,\ c)}$ peptide, 0 to 50% buffer B for tryptic peptides and 30 to 100 % buffer B for L-FDVA derivatives (buffer A: acetonitrile 20%, formic Acid 0.1% ; buffer B: 80% acetonitrile, 0.1% formic acid.). Mass analysis was performed in positive mode. The *m/z* ions corresponding to: peptides $YydF_{18\text{-}49}$ (a, b and c), tryptic peptides (18-25, 29-37 and 43-49), peptides produced by the C223A YydG mutant ($I_{44}{}^{\Delta 1}$-$H_{49}$; $G_{18}$-$V_{36}{}^{\Delta 30}$; $G_{18}$-$I_{44}{}^{\Delta 30}$), peptide $YydF_{33\text{-}49}$, Peptide $YydF_{33\text{-}49DD}$ and the $_{L/D\text{-}allo}$-Leu/Ile, $_{L/D}$Val FDVA derivatives, were selected for fragmentation at 35% NCE and acquired in profile and enhanced resolution mode on LTQ mass spectrometer.

**High resolution mass spectrometry (HRMS):** 5'-deoxyadenosine (5'-dA), the $YydF_{18\text{-}49}$ and $I_{44}{}^{\Delta 1}$-$H_{49}$ and $G_{18}$-$I_{44}{}^{\Delta 30}$ peptides were analyzed with an LTQ-Orbitrap Discovery mass

spectrometer at high resolution (30000) on the high accuracy Orbitrap mass analyser (<2ppm). Auto-calibration was made using dimethylcyclosiloxane (*m/z*=445.12003).

**Bioinformatic analysis**

**Multiple protein alignment**. Multiple protein alignment was performed using CLUSTAL Omega (1.2.1) (http://www.ebi.ac.uk/Tools/msa/clustalo) and the following sequences: AvpD (YP_321651.1), PlpD (WP_019503883.1), OspD (CBN58198.1) and PoyD (AFS60640.1) with YydG (WP_003242609) and NeoN (Q53U14).

**RiPP precursor peptide recognition element prediction.** RiPP precursor peptide recognition element prediction was performed using the HHpred server (http://toolkit.tuebingen.mpg.de/hhpred) and the pdb70_23Apr16 database with default options.

**Phylogenetic tree analysis.** Phylogenetic tree was built using Mega7[5]. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model[6]. The bootstrap consensus tree[5] inferred from 1000 replicatesis taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Initial tree for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The analysis involved 26 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 83 positions in the final dataset.

**Structural model prediction for YydG.** The structural model for YydG was built using i-Tasser[2] (http://zhanglab.ccmb.med.umich.edu/I-TASSER).

**Sequence similarity and genome neighborhood networks.** Sequence similarity network (SSN) and genome neighborhood network were built using the EFI-EST server (http://enzymefunction.org/)[3]. The sequence YydG from *B. subtilis* was used with an E-value >60 and a score >90 to query the database. The SSN built was used to built a genome neighborhood network in order to identify the genes co-occurring with *yydG*.

### Inhibition assay on liquid medium

An overnight culture of *B. subtilis* was freshly inoculated to sterile LB liquid medium. After 4 hours of bacterial growth at 37°C, the medium was diluted to 1/100 and inoculated into sterile liquid LB medium. Peptide solution was added to a final concentration of 100µM unless otherwise indicated and OD 600 nm was recorded continuously using a multi-mode plate reader Synergy 2 (BioTek).

### Peptides synthesis

The following peptides were synthesized by solid-phase synthesis by ProteoGenix except the deuterated substrate YydF_18-49-VD8.

| Peptide used for in vitro study | Peptide sequence | Molecular mass |
|---|---|---|
| YydF | MKKEITNNETVKNLEFKGLLDESQKLAKVNDLWYFVKSKENRWILGSGH | 5780.7 |
| YydF_18-49 | Ac-GLLDESQKLAKVNDLWYFVKSKENRWILGSGH-NH2 | 3773.34 |
| YydF_a | Ac-GLLDESQKLAKVNDLWYF{dV}KSKENRWILGSGH-NH2 | 3773.34 |
| YydF_c | Ac-GLLDESQKLAKVNDLWYFVKSKENRW{d-allol}LGSGH-NH2 | 3773.34 |
| YydF_b | Ac-GLLDESQKLAKVNDLWYF{dV}KSKENRW{d-allol}LGSGH-NH2 | 3773.28 |
| YydF_18-49-VD8 | Ac-GLLDESQKLAKVNDLWYF**[D8-V]**KSKENRW**[D8-V]**LGSGH-NH2 | 3775.336 |
| **Peptide used for in vivo assay** | | |
| YydF_18-49 | Ac-GLLDESQKLAKVNDLWYFVKSKENRWILGSGH | 3774.33 |
| YydF_18-49_a | Ac-GLLDESQKLAKVNDLWYF{dV}KSKENRWILGSGH | 3774.33 |
| YydF_18-49_c | Ac-GLLDESQKLAKVNDLWYFVKSKENRW{d-allol}LGSGH | 3774.33 |
| YydF_18-49_b | Ac-GLLDESQKLAKVNDLWYF{dV}KSKENRW{d-allol}LGSGH | 3774.26 |
| YydF_33-49 | WYFVKSKENRWILGSGH | 2107.42 |
| YydF_33-49DD | WYF{D-V}KSKENRW{D-allo-I}LGSGH | 2107.42 |

**The peptide YydF_18-49-VD8 was synthesized as follow:**

**Chemicals and reagents**: L-amino acid residues were purchased from Senn Chemicals (Dielsdorf, Switzerland). Fmoc-Val-d8-OH was from Sigma-Aldrich (Saint-Quentin-Fallavier, France). Rink amide MBHA resin was obtained from VWR (Fontenay-sous-Bois, France). O-benzotriazol-1-yl-*N*,*N*,*N'*,*N'*-tetramethyluronium hexafluorophosphate (HBTU) was purchased from Life Technologies (Villebon-sur-Yvette, France). Acetonitrile, and *N*-methylpyrrolidone (NMP) were from Biosolve Chimie (Dieuze, France). Diisopropylethylamine (DIEA), piperidine, trifluoroacetic acid (TFA) and other reagents were from Sigma-Aldrich.

**Peptide synthesis:** YydF_18-49_VD8 (Ac-GLLDESQKLAKVNDLWYFV*KSKENRWV*LGSGH-$NH_2$), encompassing perdeuterated valine in positions 19 and 27 (V*), was synthesized (0.1 mmol scale) on a Rink amide MBHA resin using an Applied Biosystems model 433A automatic peptide synthesizer (Life Technologies) and the standard procedures, as previously described[7,8]. All Fmoc-amino acids (1 mmol, 10 eq.) were coupled by *in situ* activation with HBTU (1.25 mmol, 12.5 eq.) and DIEA (2.5 mmol, 25 eq.) in NMP. Reactive side-chains were protected as follows: Arg, pentamethyldihydrobenzofuran (Pbf) sulfonylamide; Asn and Gln, trityl (Trt) amide; His, Trt amine; Ser and Tyr, tert-butyl (*t*Bu) ether; Asp and Glu, O-tert-butyl (O*t*Bu) ester; Lys and Trp, tert-butyloxycarbonyl (Boc) carbamate. The peptide was deprotected and cleaved from the resin by TFA as previously described[9]. Crude peptide was purified by reversed-phase HPLC (RP-HPLC) on a Vydac 218TP1022 $C_{18}$ column (2.2 x 25 cm; Grace Discovery Sciences Alltech, Templemars, France) using a linear gradient (10-60% over 60 min) of acetonitrile/TFA (99.9:0.1, v/v) at a flow rate of 10 mL/min. The peptide was analyzed by RP-HPLC on a Vydac 218TP54 $C_{18}$ column (0.46 x 25 cm; Grace Discovery Sciences Alltech) using a linear gradient (10-60% over 25 min) of acetonitrile/TFA at a flow rate of 1 mL/min. The purity of the peptide was higher than 99.9%. The peptide was characterized by MALDI-TOF mass spectrometry on a Voyager DE-PRO (Applera France) in the reflector mode with $\alpha$-cyano-4-hydroxycinnamic acid as a matrix.

**Peptides purification from *Bacillus subtilis***

*B. subtilis* was grown in M9 minimum medium as follow: Buffer solution 5X (Na2HPO4: 17g ; KH2PO4: 7.5g, NaCl: 1.25g; NH4Cl : 2.5g in 500 mL) ; Trace element solution : $MnCl_2$: 20mg; $ZnCl_2$: 34mg; $CuCl_2$: 8.6mg; $CoCl_2$: 12mg; $Na_2MoO_4$ :12mg; in 200 mL

For 100 mL of medium: 20 mL of buffer solution was added with 1 mL of trace element solution, 600µL of glucose 100µL of $CaCl_2$ (0.1M); 100µL of MgSO4 (0.1M) and 100µL of $FeCl_3$.

Bacteria were grown until OD =1 then the supernatant was harvested and purified on SPE cartridge (2g, Waters). Peptides were eluted by washing steps of 20% $CH_3CN$ with 0.1% TFA. The $YydF_{33-49}$ peptide was eluted with 80% $CH_3CN$. LC-MS analysis was performed as described above.

# Additional references


1. Soding, J., Biegert, A. & Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-8 (2005).
2. Yang, J. et al. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7-8 (2015).
3. Gerlt, J.A. et al. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim Biophys Acta* **1854**, 1019-37 (2015).
4. Nicolas, P. et al. Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. *Science* **335**, 1103-6 (2012).
5. Zharkikh, A. & Li, W.H. Estimation of confidence in phylogeny: the complete-and-partial bootstrap technique. *Mol Phylogenet Evol* **4**, 44-63 (1995).
6. Jones, D.T., Taylor, W.R. & Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275-82 (1992).
7. Leprince, J. et al. Structure-activity relationships of a series of analogues of the octadecaneuropeptide ODN on calcium mobilization in rat astrocytes. *J Med Chem* **41**, 4433-8 (1998).
8. Leprince, J. et al. Synthesis, conformational analysis and biological activity of cyclic analogs of the octadecaneuropeptide ODN. Design of a potent endozepine antagonist. *Eur J Biochem* **268**, 6045-57 (2001).
9. Chatenet, D. et al. Structure-activity relationships of a novel series of urotensin II analogues: identification of a urotensin II antagonist. *J Med Chem* **49**, 7234-8 (2006).