Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/11/eaaz2094/DC1

Supplementary Materials for

Inverse control of Rab proteins by *Yersinia* ADP-ribosyltransferase and glycosyltransferase related to clostridial glucosylating toxins

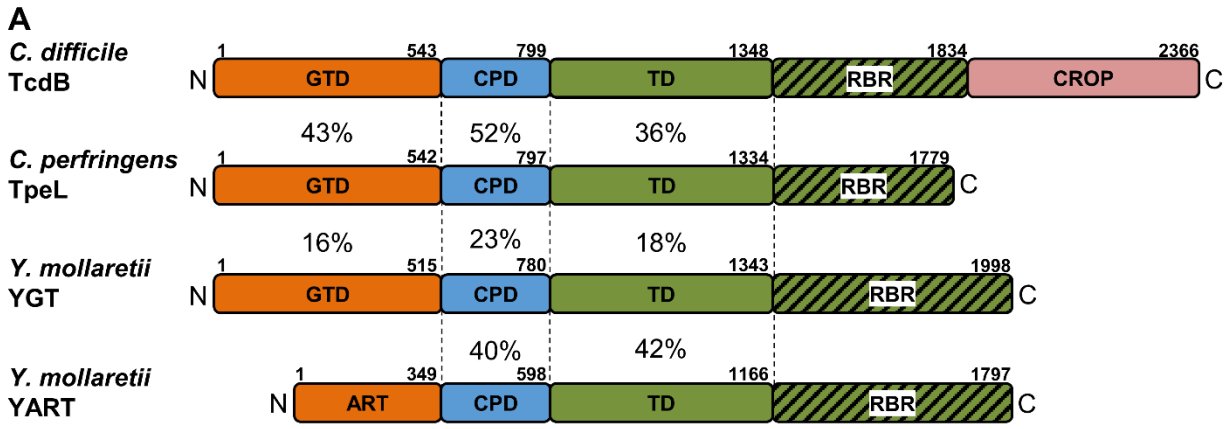
G. Stefan Ost, Christophe Wirth, Xenia Bogdanović, Wei-Chun Kao, Björn Schorch, Philipp J. K. Aktories, Panagiotis Papatheodorou, Carsten Schwan, Andreas Schlosser, Thomas Jank, Carola Hunte, Klaus Aktories*

*Corresponding author. Email: klaus.aktories@pharmakol.uni-freiburg.de

Published 11 March 2020, *Sci. Adv.* **6**, eaaz2094 (2020)
DOI: 10.1126/sciadv.aaz2094

This PDF file includes:

Supplementary Materials



B Sequence identity in %

	TpeL	YGT	YART
TcdB-GTD	43 %	16 %	9 %
TcdB-CPD	52 %	22 %	22 %
TcdB-TD	36 %	22 %	22 %
TcdB-RBR	26 %	9 %	7 %

C

```
YGT    1078VAVPLAGIAAGLPA1092LVEGFSTIA1100
YART    903VAVPLAGIAAGLPA 917LVGNYTTLA 925
TcdB   1092LLVPLAGISAGIPS1106LVNNELVLR1114
```

Fig. S1. Model of the domain architecture of YART and YGT, in comparison to *C. difficile* TcdB and *C. perfringens* TpeL. (**A**) Sequence alignments suggest an ADP-ribosyltransferase- (ART) and a glucosyltransferase domain (GTD) at the N-terminus of *Y. mollaretii* YART and YGT, respectively. It follows a cysteine protease domain (CPD), translocation domain (TD) and a receptor-binding region (RBR) within the translocation domain. The C-terminal located domain of combined repetitive oligopeptides (CROPs) is only present in TcdB but not in TpeL or YART and YGT. The amino acid sequences of the various toxin domains are compared between TcdB and TpeL, YGT and TpeL and YART and YGT. Sequence identities of domains/regions (indicated by amino acid numbers) are given in percent. (**B**) Sequence identities of protein domains/regions presented in (A) compared with the respective domains/regions of TcdB. (**C**) ClustalW alignment of TcdB, YART and YGT of a region essential for translocation. Mutation of the marked leucine residue (red box) inhibits membrane activity (see Fig. 1 B and C).

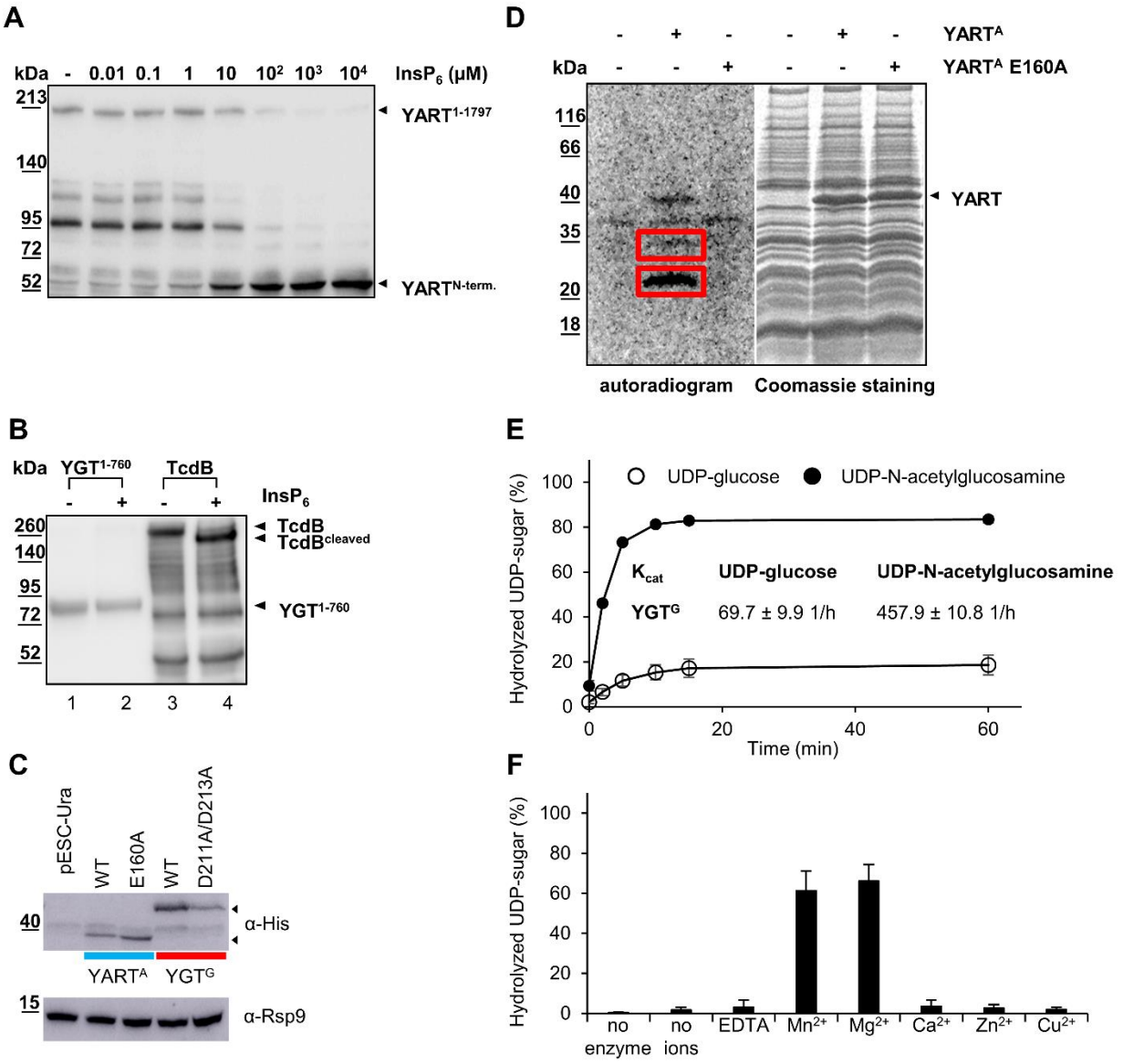


Fig. S2. Processing, expression, and enzyme activity of YART and YGT. (**A**, **B**) InsP6-dependent auto-catalytic cleavage of YART and YGT. (A) InsP6 induced processing of $YART^{1-1797}$ at increasing concentrations of InsP6 after incubation for 1 h at 37 °C. An ~52 kDa fragment was released ($YART^{N\text{-}term}$). The visualisation of the cleavage products was done by SDS-PAGE, Western-blot and subsequent staining with anti-His antibody. (B) YGT^{1-760} was incubated for 1 h at 37 °C with 10 mM InsP6 (lanes 1 and 2) without protein cleavage. As positive control, TcdB (1 µg, lanes 3 and 4) was treated with 10 mM InsP6 for 1 h at 37 °C, resulting in TcdB cleavage ($TcdB^{cleaved}$). The visualisation of the cleavage products was done as described in (A). (**C**) Expression of YART and YGT proteins in yeast. For examination of the expression of YART (wild type (WT) and inactive mutant E160A) and YGT (wild type (WT) and inactive mutant D211A/D213A), yeast cells were lysed after growing for 3 days under galactose conditions. As control, the empty vector (pESC-Ura) was transfected. The expression of the proteins was analyzed by SDS-PAGE, Western-blot and immunostaining with an anti-His

antibody (proteins are indicated by arrow heads). Loading control was tested with antibody against the ribosomal subunit Rsp9. (**D**) $YART^{A}$ ADP-ribosylates several proteins in membrane fractions of yeast. The membrane fraction of yeast lysate was incubated for 1 h at 21 °C together with $YART^{A}$ and [^{32}P]NAD^{+}. As negative control, lysate was incubated with [^{32}P]NAD^{+} without $YART^{A}$ or with the inactive mutant $YART^{A}$ E160A. The two additional bands that occurred after incubation with $YART^{A}$ are boxed in red. Shown is the phosphorimager readout together with the respective Coomassie-stained gel. (**E**, **F**) Glycohydrolase activity of YGT. (E) Identification of the preferred substrate hydrolyzed by YGT. YGT^{G} (1.8 µM) was incubated with 110 µM UDP-[^{14}C]Glc or UDP-[^{14}C]GlcNAc for the indicated times. Then, samples were subjected to thin layer chromatography. Calculation of K_{cat} values was done by using time points 0, 2, and 5 min. (F) Cation-dependency of hydrolase activity of YGT. UDP-[^{14}C]GlcNAc (15 µM) was incubated with 0.5 µM YGT^{G} in presence of different divalent cations. The reaction was performed at 21 °C in buffer containing 50 mM Hepes, 100 mM KCl and 3 mM of the indicated cations or 1 mM of EDTA. Data (+/-SD) are given from 3 independent experiments. Hydrolyzed sugars were visualized by autoradiography and the signal intensities quantified with ImageQuant.

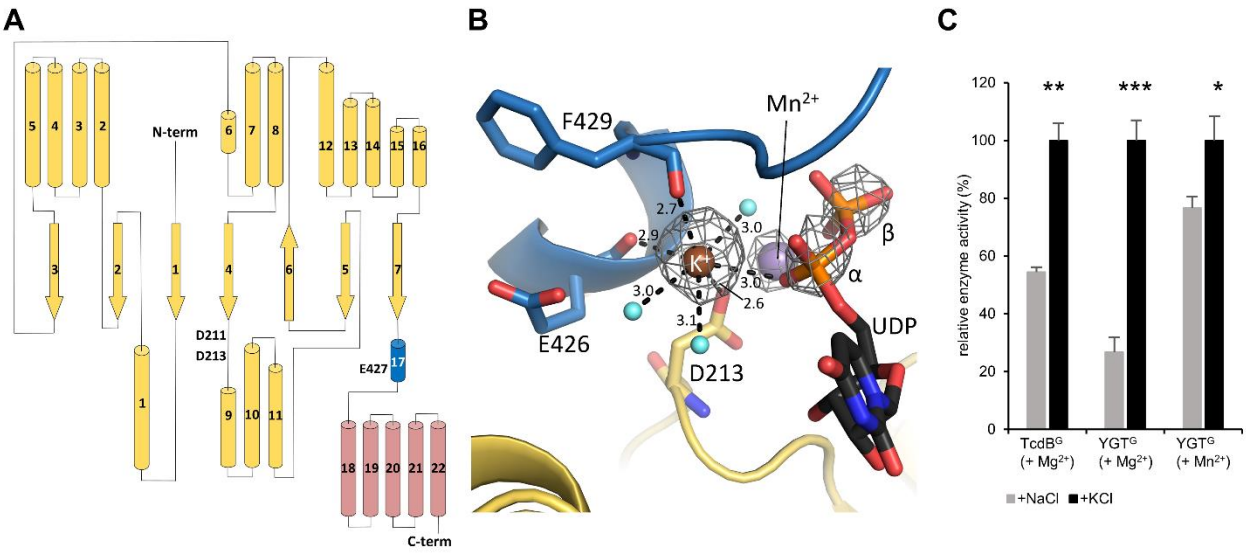


Fig. S3. Topological representation of YGT^G and potassium ion coordination in the ligand-bound YGT^G structure. (**A**) The core domain is represented in yellow, as in Fig. 4. The seven stranded mixed β-sheet is represented centrally, surrounded by the α-helical structural elements. The position of the DxD motif at the end of strand 4 is indicated. Helix 17, comprising Glu427, as well as the loop covering the UDP binding site are represented in blue. The 5 helix bundle at the C-terminus of YGT^G is represented in light red. (**B**) A strong peak in the anomalous Fourier electron density map obtained from native-SAD data (mesh presentation contoured at 5σ) supports presence of a potassium ion adjacent to the phosphate moieties of UDP and Mn^{2+}. The potassium ion (brown sphere) is coordinated by main chain carbonyl groups from Glu426 of helix 17 and from Phe429 of the subsequent loop that covers the di-phosphate moiety of the nucleotide. In addition, Asp213 coordinates both K^+ and Mn^{2+} via its carboxylic side chain. Further contributors to the K^+ coordination are water molecules and an oxygen atom from the α-phosphate from UDP. (**C**) Cation dependency of Rab5A glycosylation by YGT^G in present of Mg^{2+} or Mn^{2+} compared to RhoA glycosylation by TcdB^G with Mg^{2+}. Substrate proteins were separated by SDS-PAGE and analyzed by phosphorimaging. Shown are the means (+/-SD) of 3 independent experiments. Unpaired two sample t-test was used (*, $p< 0.05$;**, $p< 0.01$; ***, $p< 0.001$).

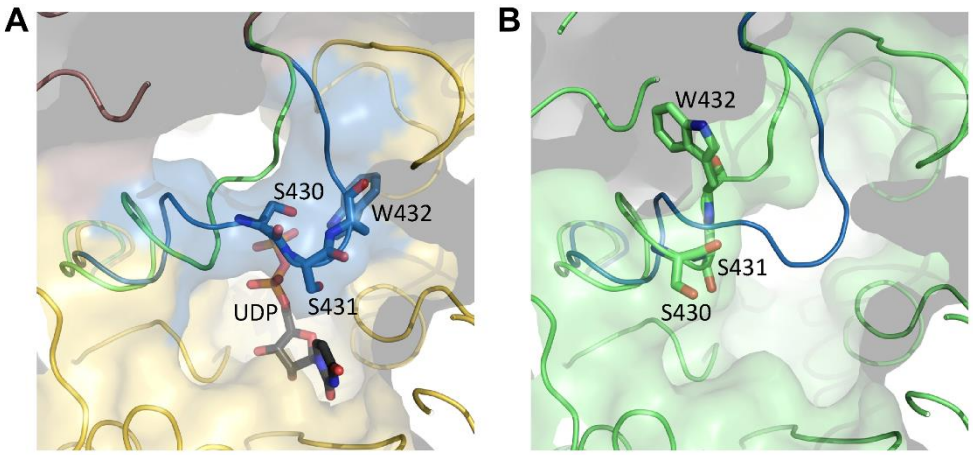


Fig. S4. C-terminal loop of the glycosyltransferase domain encloses the nucleotide-binding site in the UDP-bound state and leaves it open in the ligand-free state. (**A**) The structure of UDP-bound YGT^{G} is displayed as ribbon and with its surface represented with colors as in Fig. 4. The UDP molecule and the diphosphate-coordinating residues within the loop are represented as sticks. The loop of the ligand-free structure is superimposed and shown as green ribbon. It is shifted towards the 5-helix bundle with a distance of 8.5 Å for the Cα-position of Trp432. The side chain of the latter reorients away from the ligand binding site. (**B**) The ligand-free structure is represented as green ribbon and surface, with residues relevant for di-phosphate coordination displayed as sticks. The conformation of the loop leaves the binding pocket open. The loop is resolved in one of the YGT^{G} protomers, and disordered in the other. For visualization of the closed conformation, the loop of the UDP-bound YGT^{G} structure is superimposed and shown as blue ribbon.

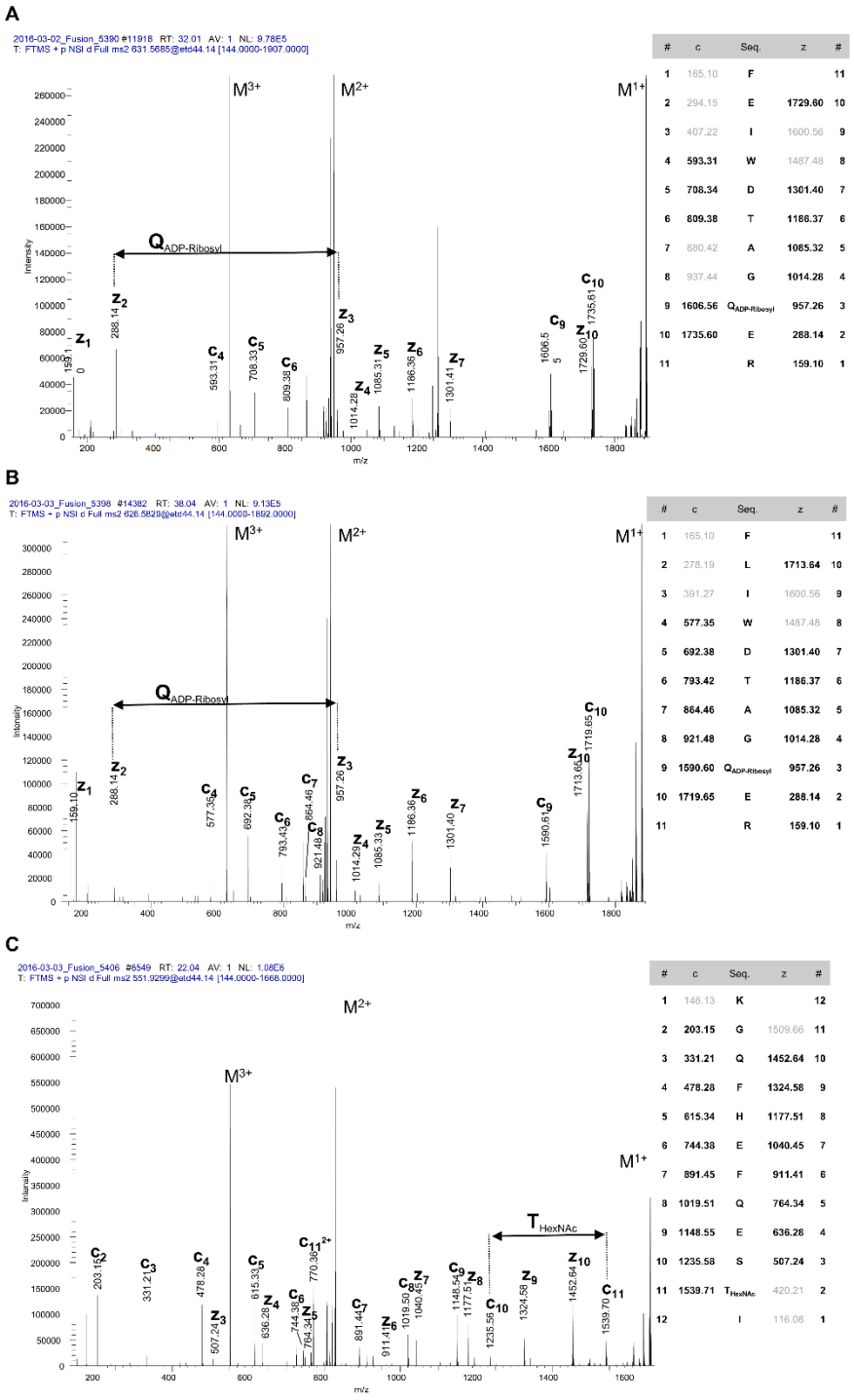


Fig. S5. ETD MS/MS spectra. (**A**) ETD MS/MS spectrum and assigned fragment ions (bold) of the peptide FEIWDTAGQER, ADP-ribosylated at Gln79 of Rab5A. This peptide was identified with a Mascot Score of 39. (**B**) ETD MS/MS spectrum and assigned fragment ions (bold) of the peptide FLIWDTAGQER, ADP-ribosylated at Gln64 of Rab31. This peptide was identified with a Mascot Score of 66. (**C**) ETD MS/MS spectrum and assigned fragment ions (bold) of the peptide KGQFHEFQESTI, HexNAc-modified at Thr52 of Rab5A. This peptide was identified with a Mascot Score of 90.

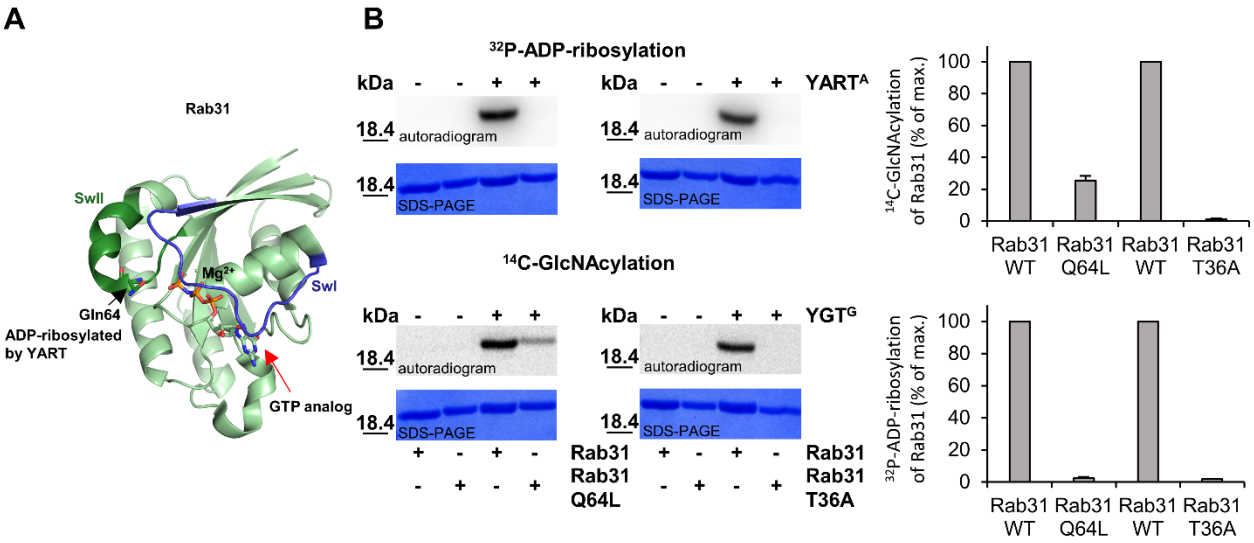


Fig. S6. YART- and YGT-induced modifications of Rab31. (**A**) Localization of the sites of Rab31 that are modified by YART-induced ADP-ribosylation. Modification sites are from MS analyses and introduced into the Rab31 structure (PDB 2FG5). Switch-1 (SwI) and switch-2 regions (SwII) are indicated in blue and dark green, respectively. Light green spheres represent the Mg^{2+} ion, while the GTP analog is indicated by red arrow. YART ADP-ribosylates Gln64 (black arrow). (**B**) ADP-ribosylation or GlcNAcylation of mutant Rab31 proteins. Wild-type and mutant Q79L or T52A Rab5 proteins (each 2 µg) were incubated with 1 µM of $YART^{A}$ or YGT^{G}, [^{32}P]NAD^{+} or UDP[^{14}C]GlcNAc for 1 h at 21 °C. The modifications of proteins were analyzed by SDS-PAGE and phosphorimaging. The amount of modified Rab proteins was normalized to wild-type Rab protein (100%). Shown are means (+/-SD) of 3 independent experiments and one representative autoradiogram and Coomassie gel.

Table S1. Data collection and refinement statistics.

	YGT^G (UDP + $MnCl_2$) 12.8 keV	YGT^G (Apo) 12.4 keV	YGT^G (UDP + $MnCl_2$) 6.2 keV
Data collection			
Wavelength (Å)	0.9686	1.000	2.000
Space group	$P4_32_12$	$P2_12_12_1$	$P4_32_12$
Cell dimensions			
a, *b*, *c* (Å)	164.3, 164.3, 47.7	47.2, 161.7, 163.0	164.5, 164.5, 47.7
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50-1.9 (2.0-1.9)*	50-3.4 (3.67-3.4)	50-2.3 (2.4-2.3)
R_{merge} (%)	18.4 (208.3)	47.3 (150.8)	31.5 (139.9)
CC(1/2) (%)	99.6 (52.5)	96.9 (88.0)	100.0 (93.7)
$I / \sigma I$	12.07 (1.3)	7.0 (2.2)	41.87 (6.23)
Completeness (%)	100 (100)	99.7 (99.4)	97.7 (92.2)
Redundancy	19.4 (19.2)	13.0 (12.2)	393.6 (329.4)
Refinement			
Resolution (Å)	45-1.9	30-3.4	
No. reflections	52064	17726	
R_{work} / R_{free}	18.6/22.3	22.7/24.9	
No. atoms			
Protein	4158	7978	
Ligands/ions	59	0	
Water	408	0	
B-factors			
Protein	52.4	46.1	
Ligands/ions	45.3	n.a.	
Water	50.1	n.a.	
R.m.s. deviations			
Bond lengths (Å)	0.005	0.0085	
Bond angles (°)	0.65	1.03	

Each dataset was collected from a single crystal
*Values in parentheses are for highest-resolution shell.

Table S2. Proteins structurally homologous to YGT^{G} as identified by DALI and showing a *Z* score higher than 10.

	PDB code	Z-score[a]	Rmsd (Å)	Lali	Nres	Identity (%)	Description	Organism
1	2vkd	33.9	3.3	396	542	21	Catalytic domain of lethal toxin	*Clostridium sordellii*
2	2bvl	33.7	3.2	397	543	21	Catalytic domain of toxin B (TcdB)	*Clostridium difficile*
3	4r04	32.1	4.0	419	1793	22	Catalytic domain of toxin A (TcdA)	*Clostridium difficile*
4	2vk9	26.9	3.2	391	540	20	Catalytic domain of the alpha toxin	*Clostridium novyi*
5	4mix	18.4	2.8	196	277	23	Glycosyltransferase domain of PaTox toxin	*Photorhabdus asymbiotica*
6	2wzf	15.5	3.0	213	515	14	LpGT (Lgt1) glycosyltransferase	*Legionella pneumophila*
7	3jsz	15.1	3.3	218	520	13	LpGT (Lgt1) glycosyltransferase	*Legionella pneumophila*
8	5h60	12.7	3.0	191	304	14	Salmonella effector SseK1	*Salmonella typhimurium*
9	6dus	12.4	3.1	190	310	14	Salmonella effector SseK3	*Salmonella typhimurium*
10	5h5y	12.0	3.0	180	286	13	Non-LEE encoded effector protein NleB	*Salmonella typhimurium*

Abbreviation: Z-score, statistical significance of the similarity between YGT^{G} and structurally related proteins; RMSD, root mean square deviation of C atoms given in Ångstroms; Lali, number of structurally equivalent residues; Nres, corresponds to the total number of amino acid residues in the hit protein; % identity, corresponds to the percentage of identical over structurally equivalent amino acid residues

The search was performed using the PDB90 database to avoid multiple occurrences of the same protein.

Table S3. Genomic DNA, primer, gblocks, plasmids, and vectors.

genomic DNA	
Yersinia mollaretii ATCC 43969	
primer	
Name	**Sequence (5→3)**
YART^{FL} fw	CCGAAGACGGAGGTATGAAAAAGAATAGAATTAATTCATTTAATGTTGATAACTACCC
YART^{FL} rv	CTACTCGAGTTAGTTGCTGACCCAGCCTCCTTGTTTAATC
YART^{A} for *E. coli* optimized fw	CTTGGATCCTTAAAAAAAAACCGCATTAACAGCTTC
YART^{A} for *E. coli* optimized rv	CTTAAGCTTTCACAGTACGCGCAATTTTTTCGTGTTTTTTTC
YART$^{1-632}$ fw	CCCAGAATTCATGAAAAAGAATAGAATTAATTCATTTAATGTTGATAACTACCC
YART$^{1-632}$ rv	CCTCTCGAGTTGACTCACCTGACGGTGCATTTCTTCC
YART^{A} for pESC-Ura fw	CTGGATCCATGTTAAAAAAAAACCGCATTAACAGCTTC
YART^{A} for pESC-Ura rv	GAGTCGACCGATTCCCACACACTAAAATCAATAG
YGT^{FL} fw	CAGAAGACGGAGGTATGAGCACTGATCTACCTAAATCGAACAAAAAC
YGT^{FL} rv	CCGTCGACTTAAACTACGCTTTTATGCGACAGCACCTCATAACTATCCC
YGT^{G} fw	CCATGG GCACTGATCTACCTAAATCG
YGT^{G} rv	GCGGCCGC ATCTGCCAGTGTTGCAAGC
YGT^{G} for pESC-Ura fw	ATGGATCCAGCACTGATCTACCTAAATCG
YGT^{G} for pESC-Ura rv	GACTCGAGTCACAACTGACCATAGGATTTCATAG
quick change (QC) YART E160A fw	CATAAAGGGCAAAAAGCGATATCTGTTATCAAGG
QC YART^{A} E160A (gBlock *3) fw	GAAACACAAGGGACAAAAAGCAATTAGCGTGATAAAGGAG
QC YART^{A} E160A (gBlock *1) fw	ACATAAGGGACAGAAAGCAATTAGCGTTATTAAAG
QC YGT^{G} D211/213A (gBlock *2) fw	GGTGTACATTGCCACGGCCCTTCTTCCCCACAT
QC YGT D211/213A fw	GGGTGTCTATATTGCTACAGCTTTGCTTCCACACATC
QC YART C500A fw	CGCTCAATCTCACCGGAGCTTCATTACTTAACCCG
QC YGT C668A fw	CTGAATATGGCGGGCGCTTCGTTATTCGATACC
QC YGT L781A fw	GAAATCCTATGGTCAGGCGGGCACACTATCACAAC
QC Rab5A T52A fw	CATGAATTTCAAGAGAGTGCCATAGGGGCTGCTTTTCTAA
QC Rab5A Q79L fw	TATGGGATACAGCTGGTCTAGAACGATACCATAGCTTAGC
QC Rab31 T36A fw	GACCACAACATCAGCCCTGCTATTGGGGCATCTTTTATG
QC Rab31 Q64L fw	CATCTGGGACACTGCTGGTCTAGAACGGTTTCATTCATTGGC

gblocks			
Number	**gblock**	**Comment**	**Amino acid seq.**
*1	YART^{A}	Codon optimized for *E. coli*	*Y. mollaretii*
*2	YGT^{G}	Codon optimized for *H. sapiens*; kozak sequence N-terminal, 3x FLAG-tag C-terminal	*Y. mollaretii*
*3	YART^{A}	Codon optimized for *H. sapiens*; kozak sequence N-terminal, 3x FLAG-tag C-terminal	*Y. mollaretii*
*4	RabGAP-5	Codon optimized for *E. coli*	*H. sapiens*
*5	R5BD-EEA1		*H. sapiens*
*6	R5BD-Rabenosyn-5		*H. sapiens*
*7	Rab5B	Codon optimized for *E. coli*	*H. sapiens*
*8	Rab5C	Codon optimized for *E. coli*	*H. sapiens*
*9	Rab17	Codon optimized for *E. coli*	*H. sapiens*
*10	Rab20	Codon optimized for *E. coli*	*H. sapiens*
*11	Rab21	Codon optimized for *E. coli*	*H. sapiens*

*12	Rab23	Codon optimized for *E. coli*	*H. sapiens*
*13	Rab24	Codon optimized for *E. coli*	*H. sapiens*
*14	Rab26	Codon optimized for *E. coli*	*H. sapiens*
*15	Rab6A	Codon optimized for *E. coli*	*H. sapiens*

plasmids and vectors

Number	Plasmid	Insert	Mutant	Comment/ gblock no.
#1	pE-SUMOstar	$YART^{1-1797}$ ($YART^{FL}$)	L917K	WP_004874725.1 (NCBI ref.-seq.)
#2		YGT^{1-1998} (YGT^{FL})	L781A; C668A; L1029K	WP_004876989.1 (NCBI ref.-seq.)
#3	pET28a	YGT^{1-520} (YGT^{G})	D211/213A	
#4		YGT^{1-780}		
#5		$YART^{1-630}$	C500A	
#6	pColdII	$YART^{1-349}$ ($YART^{A}$)	E160A	*1; 7; 8; 9; 10; 12; 14; 15
#7		Rab5B		
#8		Rab5C		
#9		Rab6A		
#10		Rab17		
#11		Rab20		
#12		Rab23		
#13		Rab26		
#14	pET28a	Rab5A	Q79L; T52A	AS-seq.: Canis lupus familaris
#15		Rab21		*11; 13;
#16		Rab24		
#17		Rab1A		AS-seq.: *H. sapiens*
#18		Rab31	Q64L; T36A	
#19	pGEX4-T3	Rab7		
#20		Rab10		
#21		H-Ras		
#22		RhoA		
#23		Rac1		
#24		Cdc42		
#25	pMAL-C2	RabGAP-5		*4; 5; 6
#26	pGEX4-T2	EEA1: $R5BD^{1-209}$		
#27		Rabenosyn-5: $R5BD^{1-40}$		
#28	pEGFP-C1	YGT^{G}	D211/213A	*2; 3
#29		$YART^{A}$	E160A	
#30	pmCherry-C1	Rab5A		AS-seq.: Canis lupus familaris
#31	pcDNA5/FRT/TO	FLAG-Rab5A		Addgene plasmid #28043
#32	pESC-Ura	YGT^{G}	D211/213A	
#33		$YART^{A}$	E160A	

Table S4. Buffer and conditions of protein purification.

PN	AR	IC (µM)	ET (°C)	ED	T	LB	WB1	WB2	EB	SB
#1	amp	50	10	ON	His, SUMO	B1			WB1 + 500 mM imidazole	B6
#2										
#3	kan	1000	23		His	B2			WB1 + 500 mM imidazole	
#6	amp	45	10			B1			WB1 + 500 mM imidazole	
#4	kan	300	23							
#5		300	20							
#7	amp	100	29			WB1 + 5 mM $MgCl_2$	B3	WB1; pH 8.0	WB2 + 500 mM imidazole	
#8										
#9										
#10										
#11										
#12										
#13										
#14	kan									
#15										
#16										
#17										
#18										
#19	amp				GST					
#20										
#21										
#22										
#23										
#24										
#25		300	37	2 h	MBP	B4		WB1; pH 7.4	WB2 + 10 mM maltose	
#26		500	37	4 h	GST	B5		PBS		PBS
#27										

Abbreviation: PN, plasmid number; AR, antibiotic resistance; IC, IPTG concentration; ET, expression temperature; ED expression duration; T, tag; LB, lysis buffer; WB, washing buffer; EB, elution buffer; SB, storage buffer; amp, ampicillin; kan, kanamycin; ON, overnight; MBP, maltose-binding protein; GST, glutathione S-transferase; B, buffer

B1: 300 mM NaCl; 20 mM Tris-HCl; 20 mM imidazole; 10 % glycerol; pH 8.0
B2: 150 mM KCl; 50 mM Tris-HCl; 20 mM imidazole; 3 mM ß-ME-EtOH; pH 7.4
B3: 300 mM NaCl; 20 mM Tris-HCl; pH 7.4
B4: 200 mM NaCl; 20 mM Tris-HCl; 1 mM EDTA; pH 8.0
B5: PBS + 0.1 % Triton x100
B6: PBS + 10 % glycerol