Supplementary Information:

Structural insight into receptor and substrate specificity of Cytotoxic Necrotizing Factor CNF1 from *Escherichia coli*

Shuangshuang Dong[1], Peer Lukat[1], Stefan Schmelz[1], Susanne zur Lage[1], Clement E. Blanchet[2], Ole Tiemann[1], Theresia E. B. Stradal[3,4] & Wulf Blankenfeldt[1,5,#]

[1]Structure and Function of Proteins, Helmholtz Centre for Infection Research, Inhoffenstr. 7, 38124 Braunschweig, Germany

[2]European Molecular Biology Laboratory Hamburg Outstation c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany

[3]Cell Biology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, 38124 Braunschweig, Germany

[4]Institute of Zoology, Technische Universität Braunschweig, Spielmannstr. 7, 38106 Braunschweig, Germany

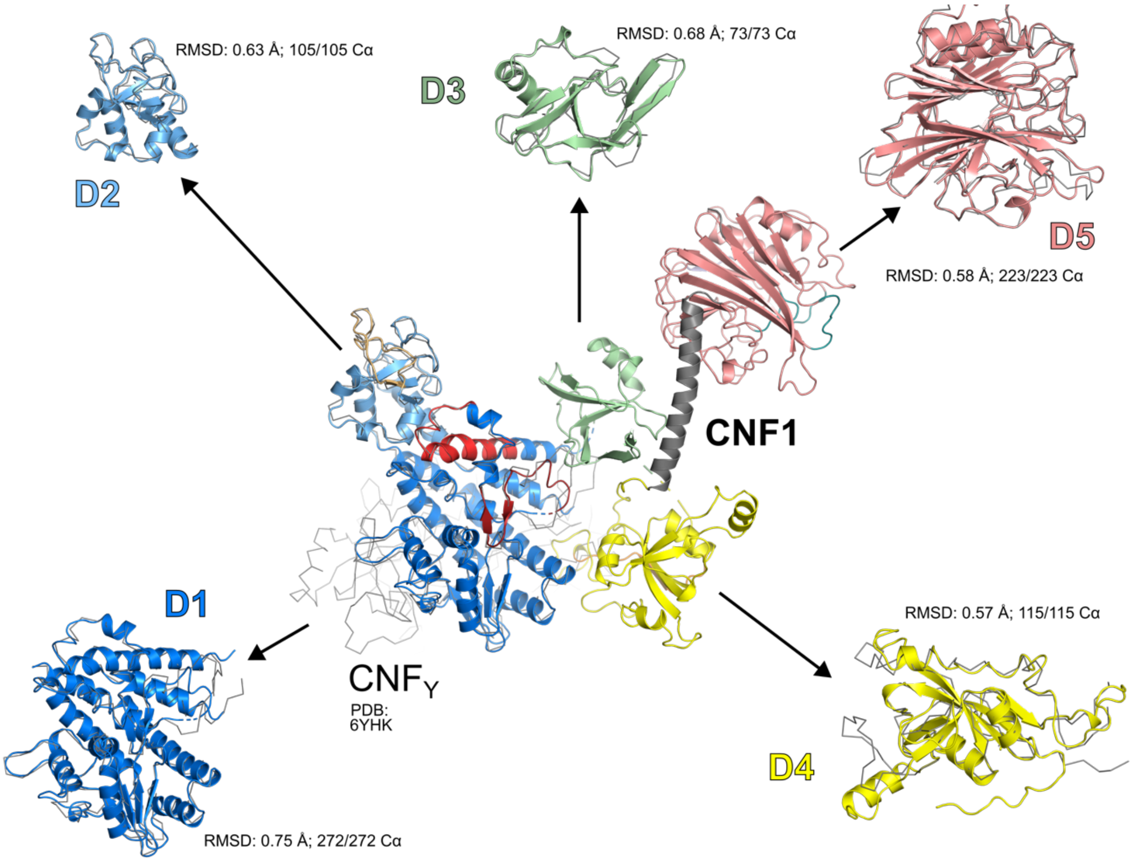
[5]Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig, Spielmannstr. 7, 38106 Braunschweig, Germany

[#]corresponding author; wulf.blankenfeldt@helmholtz-hzi.de

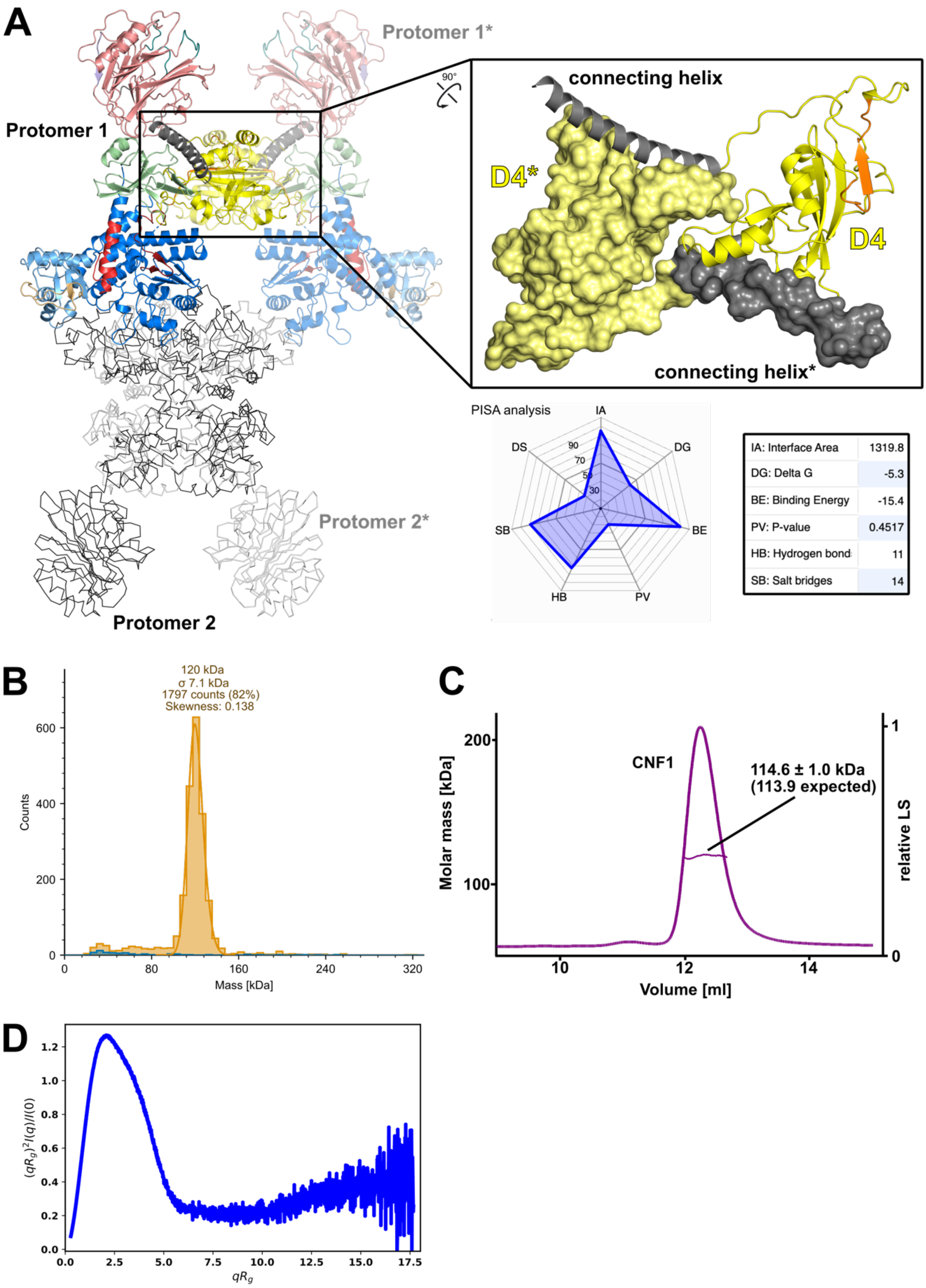
Supplementary Table 1: X-ray data collection and refinement statistics.

Structure	CNF1 (C866S)	CNF1-D5 + Lu/BCAM-D1-D2	RhoA/CNF1-D5 (C866S) + Lu/BCAM-D1-D2
PDB-ID:	29YG	29YH	29YI
Data collection			
Beamline	DESY P11	DESY P11	DESY P11
Wavelength (Å)	1.03	1.03	1.03
Space group	C2	$P2_12_12_1$	$P2_12_12$
Cell dimensions			
a, b, c (Å)	244.64, 79.86, 193.33	46.65, 113.54, 120.91	110.70, 131.36, 69.54
α, β, γ (°)	90, 115.03, 90	90, 90, 90	90, 90, 90
Resolution (Å)[a]	48.31 – 3.85 (4.06 – 3.85)	46.65 – 2.15 (2.22 – 2.15)	47.75 – 3.44 (3.50 – 3.44)
Unique reflections[a]	32315 (4699)	34105 (2945)	13785 (580)
R_{merge} (%)[a]	16.3 (99.1)	21.4 (335.1)	27.0 (72.2)
R_{pim} (%)[a]	6.7 (39.1)	5.9 (90.4)	11.4 (44.4)
$I/\sigma I$[a]	7.4 (2.2)	8.5 (2.2)	7.6 (2.1)
Completeness (%)[a]	99.4 (99.9)	95.5 (96.6)	98.6 (86.3)
Redundancy[a]	7.1 (7.3)	14.1 (14.1)	6.3 (3.5)
$CC_{1/2}$ (%)[a]	99.3 (75.0)	99.5 (75.1)	95.2 (22.6)
Refinement			
Resolution (Å)	3.85	2.15	3.44
No. reflections	32295	34007	13742
R_{work}/ R_{free} (%)	22.5/27.4	17.0/20.9	20.2/25.2
No. atoms	17646	4924	6094
Protein	17646	4666	6061
Ligand	0	48	32
Ions	0	0	1
Water	0	210	0
B-factors (Å^2)	158	51	93
Protein	158	51	93
Ligand	-	69	138
Ions	-	-	87
Water	-	50	-
R.m.s deviations			
Bond lengths (Å)	0.003	0.007	0.003
Bond angles (º)	0.67	0.63	0.61
Ramachandran statistics			
Favored	91.57	96.54	91.09
Allowed	8.28	3.46	8.91
Outliers	0.15	0.00	0.00
Clashscore (MolProbity)	6.32	1.55	2.59
MolProbity score	1.85	1.15	1.56

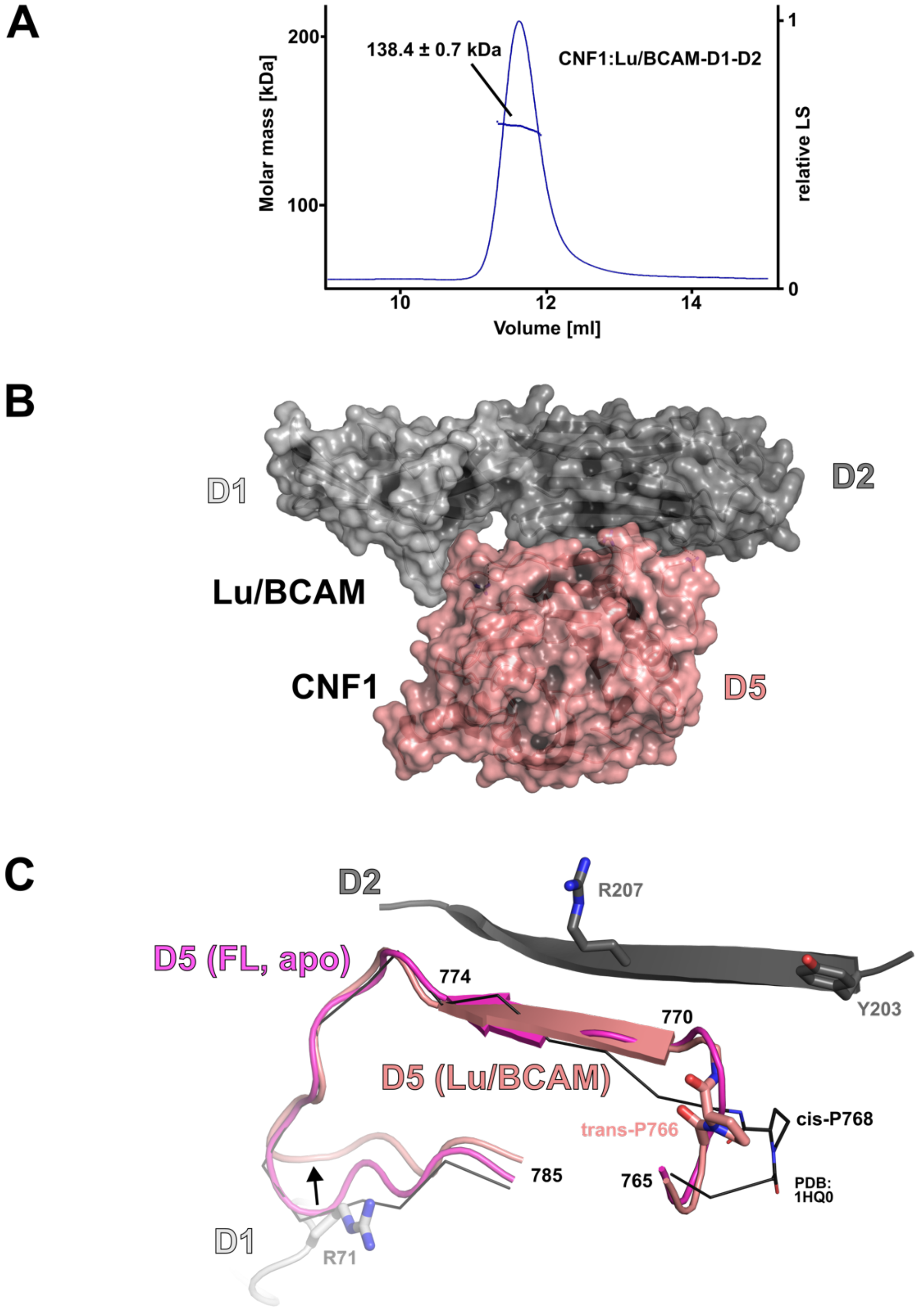
[a]Values for the highest resolution shell are shown in parentheses.



Supplementary Fig. 1: *E. coli* CNF1 and *Y. pseudotuberculosis* consist of similar structural building blocks. The crystal structures of *E. coli* CNF1 and of *Y. pseudotuberculosis* CNFY (PDB entry 6YHK) have been superimposed at their D1-D2 domains, revealing different relative positions of the domains D3 to D5 (center), whereas the individual domains themselves are highly similar in both toxins.



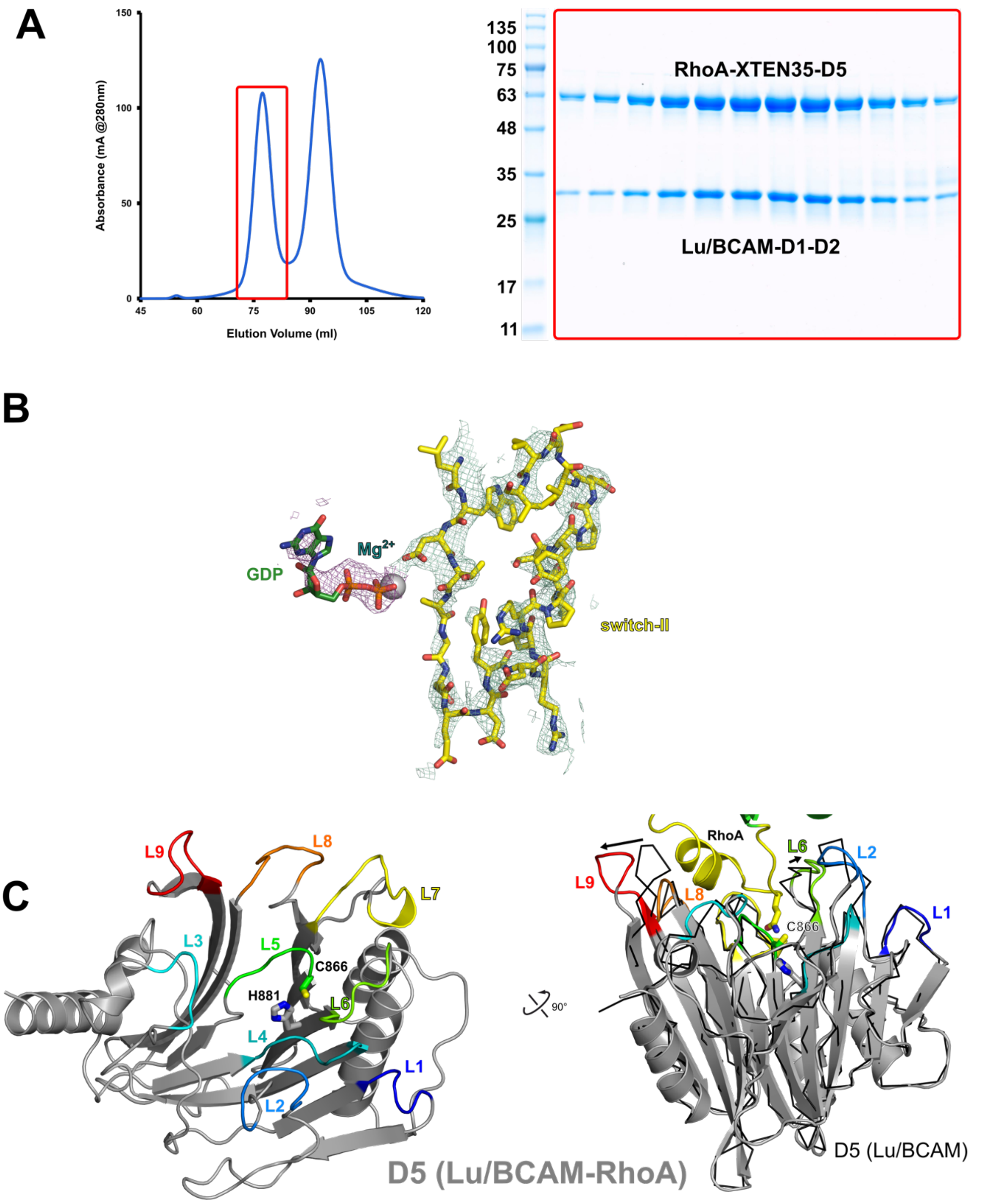
Supplementary Fig. 2: Oligomerization and flexibility analysis of CNF1. A. The asymmetric unit of the crystal form obtained for CNF1 contains two monomers in a similar conformation. Both chains reveal similar crystallographic dimers, mediated by interactions of the connecting helix (grey) with the D4-domain (yellow) of the crystallographic neighbor (magnified insert). PISA analysis suggests that this interaction could lead to a stable dimer. B. Mass photometry analysis of CNF1 suggests that the toxin is monomeric in solution. C. Similarly, SEC-MALS analysis of CNF1 also indicates that CNF1 is monomeric in solution. Note that the concentration for protein crystallization was approx. 90 µM, whereas 10 nM was used in B and 3 µM where injected in C. D. Kratky plot of the small-angle X-ray scattering data for CNF1 shown in Fig. 3. The pronounced peak indicates that the protein is predominantly folded, whereas deviations from a bell shape hint at internal flexibility.



Supplementary Fig. 3: Complex formation between CNF1 and Lu/BCAM-D1-D2. A. SEC-MALS analysis of CNF1 with Lu/BCAM-D1-D2 shows formation of a 1:1 complex. B. Surface representation of the CNF1-D5:Lu/BCAM-D1-D2 complex. The orientation is the same as in Fig. 3B. C. Details of the main interactions between CNF1 and Lu/BCAM. Note the small movement (4 Å) around position 780 of D5 to accommodate the region around R71 in Lu/BCAM-D1 (arrow), as well as the extension of the interacting β-strand at residues 770-774. D5 in the full-length structure is shown in magenta, the respective region of the apo-D5 crystal structure (PDB entry 1HQ0) is shown as a black ribbon).

Supplementary Table 2: Loop nomenclature for CNF1-D5 after Paillares et al. (2021)

Loop	**Residues**
L1	764-768
L2	789-795
L3	812-816
L4	833-838
L5	862-866
L6	884-889
L7	940-948
L8	964-970
L9	996-1002



Supplementary Fig. 4: Substrate recognition by CNF1. A. SEC elution profile of RhoA-XTEN35-D5 with Lu/BCAM-D1-D2, revealing complex formation. The peak highlighted in red is shown in SDS-PAGE analysis and was used for crystallization experiments. B. Omit electron density maps of GDP/Mg^{2+} and the switch-II region of RhoA, displayed at 1 σ. C. Position of the loops L1 to L9 (see Supplementary Tab. 2), highlighted in rainbow colors. The view on the right has been overlayed with the structure of the D5 domain in the Lu/BCAM-D1-D2 complex to highlight movements occurring on RhoA binding (indicated by arrows).